Turun yliopisto
University of Turku

NOVEL ^{19}F NMR SENSORS FOR THE CHARACTERIZATION OF HIGHER-ORDER SECONDARY STRUCTURES OF DNA AND RNA

Lotta Granqvist

TURUN YLIOPISTON JULKAISUJA - ANNALES UNIVERSITATIS TURKUENSIS
Sarja - ser. AI osa - tom. 580 | Astronomica - Chemica - Physica - Mathematica | Turku 2018

Turun yliopisto
University of Turku

NOVEL ^{19}F NMR SENSORS FOR THE CHARACTERIZATION OF HIGHER-ORDER SECONDARY STRUCTURES OF DNA AND RNA

Lotta Granqvist

TURUN YLIOPISTON JULKAISUJA – ANNALES UNIVERSITATIS TURKUENSIS
Sarja - ser. A I osa - tom. 580 | Astronomica - Chemica - Physica - Mathematica | Turku 2018

University of Turku

Faculty of Science and Engineering
Department of Chemistry
Laboratory of Organic Chemistry and Chemical Biology
Doctoral Programme in Physical and Chemical Sciences

Supervised by

Professor Pasi Virta
Department of Chemistry
University of Turku
Turku, Finland

Ass. Prof. Tuomas Lönnberg
Department of Chemistry
University of Turku
Turku, Finland

Custos

Professor Pasi Virta
Department of Chemistry
University of Turku
Turku, Finland

Reviewed by

Professor Ronald Micura
Institute of Organic Chemistry
Leopold Franzens University
Innsbruck, Austria

Professor Michal Hocek
Department of Organic Chemistry
Charles University
Prague, Czech Republic

Opponent

Professor Jacek Jemielity
Centre of New Technologies and
the Division of Biophysics
University of Warsaw
Warsaw, Poland

The originality of this thesis has been checked in accordance with the University of Turku quality assurance system using the Turnitin OriginalityCheck service.

ISBN 978-951-29-7252-4 (PRINT)
ISBN 978-951-29-7253-1 (PDF)
ISSN 0082-7002 (Print)
ISSN 2343-3175 (Online)
Painosalama Oy – Turku, Finland 2018

ABSTRACT

UNIVERSITY OF TURKU
Department of Chemistry/Faculty of Science and Engineering

GRANQVIST, LOTTA: Novel ^{19}F NMR Sensors for the Characterization of Higher-Order Secondary Structures of DNA and RNA

Doctoral thesis, 156 p.
Laboratory of Organic Chemistry and Chemical Biology
Bioorganic Chemistry
May 2018

Both DNA and RNA can fold into a variety of non-canonical structures. Non-canonical structures, such as triplexes and G-quadruplexes, are an active research area due to their biological significance and therapeutic potential. As structural complexity and conformational transitions are essential for the diverse biological roles of nucleic acids, characterization of the dynamic nature of nucleic acid is vital to understand their functional properties. This thesis focuses on ^{19}F NMR spectroscopy as a tool that can be used to investigate the conformational polymorphism of nucleic acids and the dynamic nature of nucleic acid-ligand interactions. The utility of ^{19}F NMR is based on covalently incorporated fluorine labels that act sensitive reporters upon conformational transition.

In this study, six novel fluorine-labelled building blocks were synthesized and incorporated into oligonucleotides applying standard solid-phase oligonucleotide synthesis. The building blocks were successfully used to investigate DNA and RNA triplexes, RNA invasion and bistable hairpin-G-quadruplex RNA structures. Melting of triplexes could be followed from well-distinguish ^{19}F signals of the triplex, duplex and single strand species, and melting temperatures of the structures were obtained. The temperature dependent ^{19}F NMR data of the bistable RNAs enabled to characterize melting processes, melting temperatures and thermodynamic parameters. In addition, ion induced changes at the hairpin-G-quadruplex equilibrium positions were successfully monitored. In general, the ^{19}F NMR experiments provided new information on investigated structures and demonstrated that five of the building blocks can be considered suitable for further ^{19}F NMR applications.

Key words: ^{19}F NMR, DNA, RNA, triplex, G-quadruplex, bistable

TIIVISTELMÄ

TURUN YLIOPISTO
Kemian laitos/Luonnontieteiden ja tekniikan tiedekunta

GRANQVIST, LOTTA: Uusia ^{19}F NMR sensoreita DNA:n ja RNA:n korkeamman asteen sekundaarirakenteiden karakterisointiin

Väitöskirja, 156 s.
Orgaanisen kemian ja kemiallisen biologian laboratorio
Bio-orgaaninen kemia
Toukokuu 2018

DNA ja RNA voivat laskostua erilaisiksi ei-kanonisiksi rakenteiksi. Ei-kanoniset rakenteet, kuten kolmoiskierteet ja G-kvadrupleksit, ovat aktiivinen tutkimusalue niiden biologisen merkittävyyden ja niihin liittyvien terapeuttisten mahdollisuuksien takia. Koska rakenteellinen monimutkaisuus ja konformaatiomuutokset ovat välttämättömiä nukleliinihappojen biologisille toiminnalle, rakenteiden ja niiden dynamiikan karakterisointi on tärkeää toiminnallisten ominaisuuksien ymmärtämiselle. Tässä väitöskirjassa keskityttiin ^{19}F NMR spektroskopiaan tutkimusmenetelmänä, jolla voidaan tutkia nukleiinihappojen konformationaalista monimuotoisuuutta ja nukleiinihappojen ja niiden ligandien välisiä vuorovaikutuksia. ^{19}F NMR perustuu nukleiinihappoihin kovalenttisesti liitettyihin fluorileimoihin, jotka toimivat herkkinä reporttereina konformaatiomuutosta detektoitaessa.

Tutkimuksessa syntetisoitiin kuusi fluorileimatttua rakennusyksikköä, jotka liitettiin oligonukleotideihin käyttämällä vakiintunutta kiintokantajalla tapahtuvaa oligonukleotidien synteesimenetelmää. Rakenneyksiköillä tutkittiin onnistuneesti DNA ja RNA kolmoiskierteitä, RNA:n invaasiota ja kahtaispysyviä hiusneula-G-kvadrupleksi RNA rakenteita. Kolmoiskierteiden denaturaatiota pystyttiin seuraamaan yksityiskohtaisesti hyvin erottuvista kolmoiskierrettä, kaksoiskierrettä ja yksinauhaista rakennetta vastaavista fluorisignaaleista. Kahtaispysyvien rakenteiden lämpötilariippuvainen käyttäytyminen pystyttiin karakterisoimaan, mikä mahdollisti sulamislämpötilojen ja termodynaamisten arvojen määrittämisen. Tämän lisäksi ionien vaikutusta hiusneula-G-kvadrupleksi tasapainotiloihin havannoitiin onnistuneesti. Yleisesti ^{19}F NMR tutkimuksilla saatiin uutta tietoa tutkituista rakenteista. Lisäksi tutkimukset osoittivat että viiden syntetisoidun rakennusyksikön ominaisuudet soveltunevat myös muihin ^{19}F NMR tutkimuksiin.

Asiasanat: ^{19}F NMR, DNA, RNA, kolmoiskierre, G-kvadrupleksi, kahtaispysyvä

PREFACE

This thesis is based on experimental work carried out in the Laboratory of Organic Chemistry and Chemical Biology at the Department of Chemistry, University of Turku (Finland) during the years 2013-2018. The financial support from the Academy of Finland, the Graduate School of University of Turku and the Finnish Cultural Foundation are gratefully acknowledged.

First and foremost, I would like to express my deepest gratitude to my supervisor Professor Pasi Virta for giving me the opportunity to work under his continuous guidance and encouragement. Working on this thesis has not been an easy journey and I truly appreciate his patience and support. I would also like to thank Professor Harri Lönnberg for his advices and for inspiring me to study chemistry during my undergraduate studies.

I wish to thank Professor Ronald Micura and Professor Michal Hocek for reviewing of this thesis and presenting their valuable comments and Professor Jacek Jemielity for accepting to act as my opponent.

I am grateful for Ville Tähtinen and Andrzej Kraszewski for their contribution on the paper IV.

I sincerely thank all the people in the Bioorganic group whom I have worked during my studies. I would also like to extend my thanks to Dr. Jari Sinkkonen, Kirsi Laaksonen, and Dr. Kaisa Ketomäki for their kind assistance. Also other colleagues that have helped me during this time are highly appreciated. Very special thanks are devoted to Mauri Nauma and Kari Loikas for fixing all the problems related to instruments, although I do think that one particular device would need an appointment with an exorcist.

I would like to express my warmest thanks to all my friends in and outside academia for their support. Finally, my deepest thanks are devoted to my parents and my brothers and sisters for always supporting me. Kiitos. Special thanks go to my little brother. He made me always laugh when I lost my sense of humour during this process. So dear siblings, just imagine my favourite emoji character here and after that a smiley face.

Turku, May, 2018

CONTENTS

LIST OF ORIGINAL PUBLICATIONS

This thesis is based on the following publications:

I Granqvist, L. and Virta, P. P. 4´-*C*-[(4-Trifluoromethyl-1*H*-1,2,3-triazol-1-yl)-methyl]thymidine as a sensitive ^{19}F NMR sensor for the detection of oligonucleotide secondary structures. *J. Org. Chem.* **2014**, *79*, 3529-3536.

II Granqvist, L. and Virta, P. 2´-*O*-[(4-CF_3-triazol-1-yl)methyl]uridine – A sensitive ^{19}F NMR sensor for the detection of RNA secondary structures. *J. Org. Chem.* **2015**, *80*, 7961-7970.

III Granqvist, L. and Virta, P. Characterization of G-quadruplex/hairpin transitions of RNAs by ^{19}F NMR spectroscopy. *Chem. Eur. J.* **2016**, *22*, 15360-15372.

IV Granqvist, L., Kraszewski,A., Tähtinen,V. and Virta,P. Synthesis of aminoglycoside-2´-*O*-methyl oligoribonucleotide fusions. *Molecules* **2017**, *22*, 760.

LIST OF RELATED PUBLICATIONS

Tähtinen, V.; Granqvist, L.; Virta, P. Synthesis of *C*-5, *C*-2′ and *C*-4′-neomycin-conjugated triplex forming oligonucleotides and their affinity to DNA-duplexes. *Bioorg. Med. Chem.* **2015**, *23*, 4472-4480. (Ref. 36)

Tähtinen, V.; Granqvist, L.; Murtola, M.; Strömberg, R.; Virta, P. ^{19}F NMR spectroscopic analysis of the binding modes in triple-helical peptide nucleic acid (PNA)/MicroRNA complexes. *Chem. Eur. J.* **2017**, *23*, 7113-7124. (Ref. 95)

ABBREVIATIONS

1D, 2D	one dimensional, two dimensional
A	adenosine
Ac	acetyl
Ac_2O	acetic anhydride
AcOH	acetic acid
Ade	adenine
AdoHcy	S-adenosyl-L-homocysteine
B	base
BnSTet	benzylthiotetrazole
Bz	benzoyl
C	cytidine
CD	circular dichroism
CPG	controlled pore glass
Cyt	cytosine
DCM	dichloromethane
DIEA	*N*,*N*-diisopropylethylamine
DMAP	4-dimethylaminopyrimidine
DMF	*N*,*N*-dimethylformamide
DMSO	dimethyl sulfoxide
DMTr	4,4´-dimethoxytrityl
DNA	deoxyribonucleic acid
ds	double-stranded
ENE	element for nuclear expression
ESI-TOF MS	electrospray ionization time-of-flight mass spectrometry
Et_3N	trimethylamine
Et_2O	diethyl ether
EXCY	exchange spectroscopy
FAAF	*N*-(2´-deoxyguanosin-8-yl)-7-fluoro-2-acetylaminofluorene
FABP	*N*-(2´-deoxyguanosin-8-yl)-4´-fluoro-4-aminobiphenyl
FAF	*N*-(2´-deoxyguanosin-8-yl)-7-fluoro-2-aminofluorene
FMN	flavine mononucleotide
G	guanosine
Gua	guanine
H-DNA	naturally occurring triple helical DNA structure
HIV-1, HIV-2	human immunodeficiency virus type 1 and type 2
HOESY	heteronuclear Overhauser effect spectroscopy
Hp	hairpin conformer
HP NMR	hyperpolarized NMR
ITC	isothermal titration calorimetry
K_d	dissociation constant
KSHV	Kaposi's sarcoma-associated herpesvirus
lncRNA	long noncoding RNA

MALAT1	metastasis-associated lung adenocarcinoma transcript 1
Me	methyl
MeCN	acetonitrile
MENβ	multiple endocrine neoplasia-β
MeOH	methanol
M.*Hha*I	DNA cytosine-5 methyltransferase *Hha*I
miRNA	microRNA
MRI	magnetic resonance imaging
mRNA	messenger RNA
MS	mass spectrometry
N	any nucleotide
NBSCl	2-nitrobenzenesulfenyl chloride
ncRNA	noncoding RNA
NMR	nuclear magnetic resonance
NOESY	nuclear Overhauser effect spectroscopy
nt	nucleotide
ON	oligonucleotide
O_R3	operator DNA sequence of bacteriophage lambda
PAN RNA	polyadenylated nuclear RNA
PDB	Protein Data Bank
PEG	polyethylene glycol
PNA	peptide nucleic acid
pre-mRNA	precursor messenger RNA
$preQ_1$	7-aminomethyl-7-deazaguanine
PrP^C	normal cellular prion protein
PrP^{Sc}	infectious misfolded form of prion protein
Py	pyridine
Qd	G-quadruplex conformer
REDOR	rotational-echo double-resonance
rF	2,4-difluorotoluyl ribonucleoside
RNA	ribonucleic acid
RP HPLC	reverse-phase high-performance liquid chromatography
rRNA	ribosomal RNA
SAM-II	S-adenosylmethionine type II riboswitch
SIIS	solvent-induced isotope shift
snRNA	small nuclear RNA
ss	single-stranded
T	thymidine
T_m	melting temperature
TAR	transactivation response element
TBA	thrombin binding aptamer
TBAF	tetrabutylammonium fluoride
TBDMS	*tert*-butyldimethylsilyl
TERRA	telomeric repeat-containing RNA
Tf	trifluoromethanesulfonyl
TFO	triplex-forming oligonucleotide

TFPNA	triplex-forming peptide nucleic acid
THF	tetrahydrofurane
Thy	thymine
TMSOTf	trimethylsilyl trifluoromethanesulfonate, trimethylsilyl triflate
TOM	triisopropylsilyloxymethyl
tRNA	transfer RNA
$tRNA^{Lys}_{3}$	lysine tRNA
$tRNA_f^{Met}$	*N*-formylmethionine tRNA
$tRNA^{Met}$	methionine tRNA
$tRNA^{Phe}$	phenylalanine tRNA
$tRNA^{Val}$	valine tRNA
TROSY	transverse relaxation optimized spectroscopy
U	uridine
Ura	uracil
UV	ultraviolet spectroscopy

1. INTRODUCTION

Fluorine-19 nuclear magnetic resonance (^{19}F NMR) spectroscopy has numerous applications in chemical biology.[1-12] The high detection sensitivity of fluorine combined with the intrinsic sensitivity of fluorine chemical shifts to the local environment has led to many successful ^{19}F NMR applications also for investigation of biological macromolecules.[1-6,8,9,11,12] Fluorine is an unnatural element in macromolecules that is essentially absent from biological systems. Therefore, an obvious prerequisite for ^{19}F NMR spectroscopy is the incorporation of fluorine into investigated molecules. However, the lack of interfering background signals provides a great advantage in NMR spectroscopy. ^{19}F NMR spectroscopy was recognized as a potential method to investigate biologically important macromolecules over 40 years ago.[13,14] Employing fluorine as a spin label to investigate macromolecules by introducing it either into the macromolecule[13] or into its ligand[14] was first established with proteins, and this concept was embraced a few years later in the nucleic acid field. In 1977, two independent studies reported the potential of the fluorine labelling approach for ^{19}F NMR studies of nucleic acids.[15,16] Marshall and Smith reported their ^{19}F NMR studies on fluorine-labelled 5S ribosomal RNA (rRNA)[15], and Horowitz and co-workers published their research on fluorine-labelled transfer RNA (tRNA)[16]. Since then, ^{19}F NMR has been used to investigate several artificial and biologically important nucleic acid structures, such as aptamers[17-19], viral regulatory hairpins[20-25], riboswitches[26-29], ribozymes[30-32], DNA adducts[33,34], triplexes[35,36] and telomeric G-quadruplexes[37-40].

Technical and synthetic advances have played and are still playing a significant role in the progress of ^{19}F NMR spectroscopy applications. Substantial advances have been made in NMR instrumentation that increase feasibility and sensitivity of ^{19}F detection.[4,5,8] Recent advances in the hyperpolarized NMR technique (HP NMR) have greatly enhanced the sensitivity of ^{19}F detection, and HP ^{19}F NMR is expected to contribute generally to future macromolecule research.[41,42] Significant progress in synthetic fluorine chemistry has provided straightforward access to numerous fluorine containing macromolecules and small molecules.[6,12,43-45] In addition, advances in oligonucleotide synthesis techniques have played a crucial role in nucleic acid-related research in general.[46-51]

Progress in the nucleic acid field has been slower than in the protein field, but interest in nucleic acid-related ^{19}F NMR research has dramatically increased in recent years. The structural and functional diversity of nucleic acids have increased tremendously and this progress has also impacted ^{19}F NMR research. In addition, the ^{19}F NMR spectroscopy-based techniques employed in drug development have started to shift from the protein field to the nucleic acid field.[17,40,52] Moreover, ^{19}F NMR spectroscopy based methods that utilize fluorinated oligonucleotides as hybridization probes[53-58] have recently been developed and have also received interest in magnetic resonance imaging (MRI)[53,54,57].

The growing knowledge of nucleic acids provides limitless research opportunities, and considering instrumental and synthetic developments as well as the variety of potential applications, ^{19}F NMR spectroscopy may become a routine detection method in nucleic acids research.

1.1 ^{19}F NMR and nucleic acids

Fluorine-19 has several advantageous NMR properties: it has a spin of ½, a high gyromagnetic ratio, and it has a natural abundance of 100%.[2,59,60] ^{19}F sensitivity is high, 83% of the sensitivity of ^{1}H, and much higher than that of ^{13}C, ^{15}N and ^{31}P (Table 1). ^{19}F has a wide chemical shift range (over 400 ppm), and ^{19}F shifts are highly sensitive to changes in the local environment. In addition, ^{19}F is absent from nucleic acids and the background signal interference is thus negligible with typical NMR measurement conditions. The disadvantage of ^{19}F NMR spectroscopy is that the ^{19}F reporter group (label, probe) must be incorporated into nucleic acids or into nucleic acid ligands.[1,2,5,8,12] Because an introduced modification may significantly alter the native properties of nucleic acids, the influence of the fluorine substitution should be carefully evaluated. The chemical and physical properties of fluorine, however, are highly beneficial for modulating pharmacological properties of small molecular ligands.[44]

Table 1.[2,60] NMR properties of selected nucleus

Nucleus	Spin (I)	Gyromagnetic ratio (10^{-7} rad s^{-1} T^{-1})	Natural abundance (%)	Relative sensitivity
^{1}H	½	26.7519	99.98	1.0
^{19}F	**½**	**25.181**	**100**	**0.83**
^{13}C	½	6.7283	1.108	0.0159
^{15}N	½	-2.712	0.37	0.00134
^{31}P	½	10.841	100	0.066

Much of the interest in ^{19}F NMR spectroscopy arises from the simplicity and informational richness of one-dimensional (1D) ^{19}F NMR spectra.[1,2,4-6,8,9,12] In particular, the strength of ^{19}F NMR is its ability to provide insight into the complex nature of the dynamic intra- and intermolecular structural equilibria of nucleic acids and dynamic behaviour of nucleic acid-ligand complexes.[1,2,5,8,12] In addition to solution-state NMR, solid-state ^{19}F NMR methods have been used to investigate nucleic acids.[22,61-65] A common problem with ^{1}H NMR studies on large molecules is the abundance of protons that need to be assigned and the severe signal overlap due to the narrow chemical shift range (10-15 ppm) of protons.[66-72] Introducing a spin label such as ^{19}F into the nucleic

acid may hence provide clear advantages.[1,2,5,8,72] Compared to spectrophotometric methods commonly used to investigate nucleic acids such as ultraviolet spectroscopy (UV)[73-75], circular dichroism (CD)[76-78], and fluorescence[79,80], NMR methods are inherently less sensitive. However, as with other NMR nuclei ^{19}F NMR spectroscopy has other advantageous properties, such as its ability to provide more detailed information at local levels and information on coexisting structures.[2,5,8,12] ^{19}F NMR is often used in conjunction with other methods, and it is considered an independent complementary method rather than a competitive alternative to NMR with traditional nuclei or other methods.

Although this thesis is focused on 1D ^{19}F NMR applications, multidimensional homo- and heteronuclear NMR experiments, such as ^{19}F/^{19}F NOESY (nuclear Overhauser effect spectroscopy) and ^{19}F/^{1}H HOESY (heteronuclear Overhauser effect spectroscopy), have been used to obtain further information on nucleic acids structures.[21,23,25,81-85]. In addition, multidimensional NMR methods are central techniques for ^{19}F signal assignment[21,23,25,29,83], and fluorine can be used as a tool to aid the assignment of ^{1}H signals[84-86].

1D ^{19}F NMR applications for nucleic acids are mainly derived from the intrinsic sensitivity of the fluorine chemical shift to the local environment.[2,60] In general, fluorine can be introduced either into a nucleic acid (nucleic acid observed, *internal probes*) or its ligand (ligand observed, *external probes*). Due to the high environmental sensitivity of fluorine, ^{19}F signals are expected to reflect distinct state of a process of interest, such as conformational states or ligand-bound and unbound states (Figure 1a). Depending on the exchange rate of the process, three basic exchange regimes can be identified: fast (with well-behaving coalescence signal), intermediate or slow exchange (Figure 1b).[11,87,88] When the exchange rate is slow on the NMR timescale, two distinct signals are observed corresponding to individual states (free and bound or distinct conformations A and B). The relative peak areas of the resolved signals directly reflect the distribution of the states (i.e., molar fractions), and thus the integration of the signals allows a quantitation of the states. In the fast exchange regime, a single signal is detected as a weighted average signal of both states. In the case of an intermediate exchange process, a broad signal is detected. Depending on the goals, ^{19}F NMR data can be used to obtain quantitative or qualitative information. Commonly, the interest has been in detecting structural equilibria or complex formation.[1,2,5,8,12] In addition, ^{19}F NMR data have been used to extract different parameters, such as dissociation constants (K_d)[17,20,24,25,38,52,89], pK_a values[90-92], melting temperatures (T_m)[19,24,36-39,55,89,93-95] and thermodynamic[19,31,37,90,96] and kinetic[30,32,96,97] parameters.

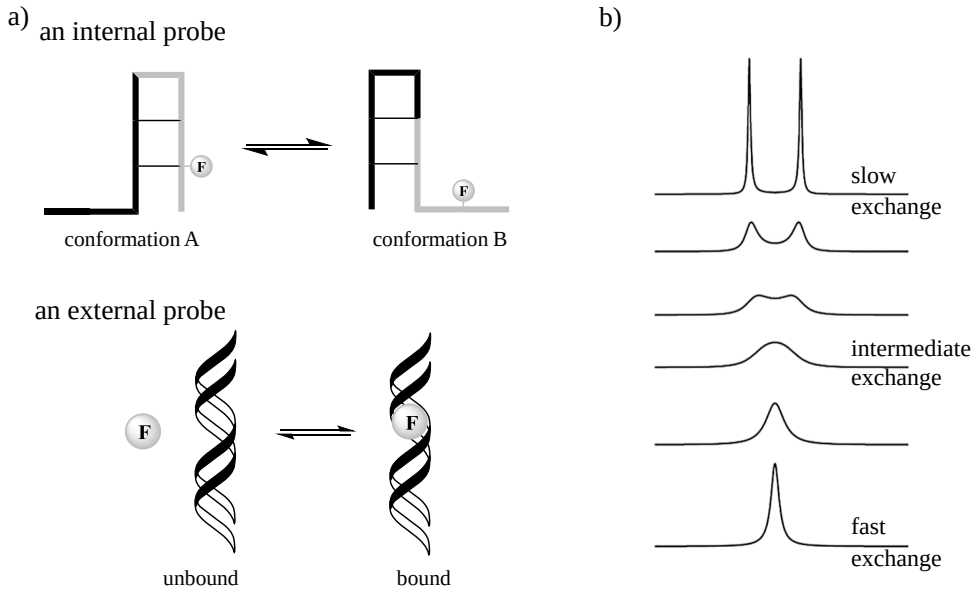


Figure 1. (a) A schematic presentation of internal and external fluorine probes and (b) the chemical exchange rates.

The structural and conformational analysis and ligand-nucleic acid interactions studies have mainly utilized internal fluorine probes. Consequently, a large portion of literature review is focused on internal fluorine probes (Section 1.2). Using external small molecule fluorine probes as structural probes has been demonstrated as an alternative approach to internal fluorine probes, and this concept is presented in Section 1.3. Section 1.4 is a brief summary of small molecule-nucleic acid interactions studies.

1.2 Internal fluorine probes

1.2.1 General properties of fluorine labels and synthesis of fluorine-labelled oligonucleotides

Fluorine labels can be incorporated into oligonucleotides either enzymatically or by chemical synthesis. Although standard oligonucleotide synthetic techniques[48,69,98] may be applied, challenges in the synthesis of appropriate fluorine-containing building blocks or their poor incorporation efficiency into oligonucleotides can significantly limit the use of the label. The most common synthetic method is the standard solid-phase oligonucleotide synthesis[98] using fluorine-modified phosphoramidite building blocks.

This method requires that a fluorine-containing nucleoside phosphoramidite is compatible with automated solid-phase synthesis, including the synthesis cycle and cleavage conditions. An enzymatic synthesis of fluorine-labelled oligonucleotides requires a triphosphate analogue of a fluorine-containing nucleoside.[21,23,29,91,99-101] The triphosphates are usually incorporated into RNA by *in vitro* transcription using T7 RNA polymerase[21,23,29,91,99] or into DNA using DNA polymerase[100,101]. In contrast to solid-phase synthesis, which allows the site-specific incorporation of a fluorine label, enzymatic synthesis typically produces a uniformly fluorine-labelled oligonucleotide.[21,23,48,98,99] In general, enzymatic synthesis allows synthesis of larger oligonucleotides compared with chemical synthesis.

In fluorinated nucleoside analogues, the fluorine moiety can be incorporated either in the base or the sugar.[12] In addition, a few non-nucleosidic fluorine modifications have been attached to the 5´-end of oligonucleotides.[18,37,53,102] Some of the analogues are well-known from other applications. For example, 5-fluorouridine is a nucleoside analogue of the anticancer drug 5-fluorouracil[103], 2´-deoxy-5-trifluoromethyluridine (trifluridine) is an antiviral drug[104], and 2´-deoxy-2´-fluoronucleosides are known modifications in the antisense applications[105]. However, there has been growing interest in the development of novel fluorinated derivatives that are specifically designed for ^{19}F NMR applications. The goal of development has been to improve the sensitivity and practicality of ^{19}F NMR measurements.[18,19,37,39,53,55,58,101,106,107] A direct approach to improve sensitivity is to increase the amount of magnetically equivalent fluorine atoms and to introduce fluorine atoms in an isolated spin system. The measurements may then be performed at micromolar RNA concentrations without plausibly interfering ^{19}F-^{1}H couplings (if the proton-fluorine decoupling technique is not available). In addition to generally improving ^{19}F NMR spectroscopic properties, derivatives have been developed for more specific purposes. For example, dual probes bearing both a ^{19}F- and a fluorescence-detecting moiety have been described.[39,100] An important aspect for the development of fluorinated derivatives, however, is their effect on the native structure of the nucleic acid. In general, fluorine modification should not cause secondary structural perturbations or affect the stability of the structure. Although a non-invasive modification is typically highly preferable, in some applications, for example in the case of hybridization probes[53,55,56,58], the role of perturbations is less important. The following chapters (1.2.1.1–1.2.1.3) describe fluorinated nucleoside and non-nucleoside derivatives covering some general issues, but the utility of the labelled derivative depends on the application.

1.2.1.1 Base modifications

The most common site for fluorine modification is the C5 position of pyrimidines (Figure 2). Substituents in the C5 position are oriented the major groove in the double

helix.[108] Large substituents at this position are generally well tolerated and do not significantly disturb the structure or stability of the duplex. Sterically demanding groups accommodated in the major groove may, as expected, significantly affect groove-mediated interactions.

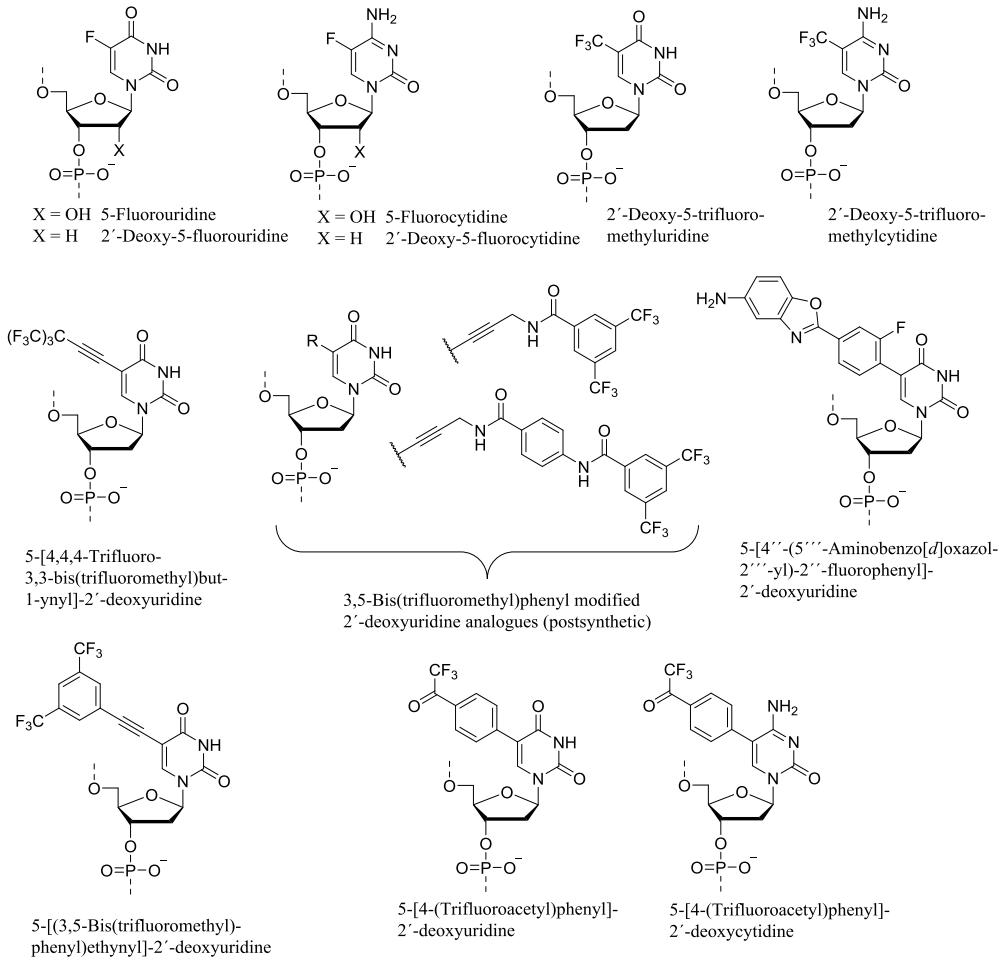


Figure 2. Structures of C5-substituted pyrimidines.

5-Fluoropyrimidines (i.e., 5-fluorouridine[15,16,20,23,25,26,28,30,64,65,83,90,97,109-111], 2´-deoxy-5-fluorouridine[35,109,111-117], 5-fluorocytidine[23,25,27,28,31,32,83,111] and 2´-deoxy-5-fluorocytidine[111,116,118]) have been extensively used in ^{19}F NMR studies. 5-Fluoropyrimidine-modified nucleic acids may be synthesized by solid-phase synthesis using 5-fluoropyrimidinenucleoside phosphoramidites.[111,119] In addition, 5-fluoropyrimidine (both U and C) -substituted RNAs can be produced efficiently by enzymatic methods.[23,99] The impact of 5-fluoro modification on nucleic acid structures has been thoroughly investigated by several methods showing that this modification has

a negligible effect on their structure or stability.[23,83,92,111,120-124] Fluorine is sterically comparable to hydrogen, but due to its high electronegativity[125], it decreases the apparent pK_a values of cytidine N3 from 4.2 to 2.3[126] and that of uridine from 9.2 to 7.6[23,127]. These decreased pK_a values have no observed effects on base pairing at a physiological pH.[23,83,92,111,120-124] Fluorine substitution, however, increases the opening rates of the canonical Watson-Crick 5-fluoropyrimidine-adenine base pairs within the DNA or RNA duplex. In contrast, the opening rates are unchanged with canonical 5-fluorocytosine-guanine base pairs. Overall, 5-fluoro pyrimidine modification has proven to be sensitive non-invasive label. From the ^{19}F NMR perspective, measurements have been conducted either with or without proton decoupling techniques.

2´-Deoxy-5-trifluoromethyluridine bears three magnetically equivalent fluorine atoms in an isolated spin system.[128] The trifluoromethyl group is, however, readily converted to a cyano group in concentrated ammonia, which prevents the incorporation of this nucleoside derivative into oligonucleotides by the standard solid-phase oligonucleotide synthesis.[129] Although this problem can be circumvented by applying milder cleavage conditions[56,129], the modest stability of the 5-trifluoromethyl group may be reason why few ^{19}F NMR experiments employ this nucleoside analogue[55,56,63,128,130]. (An enzymatic synthetic method has also been described.[131,132]) A 5-trifluoromethyluracil-adenine base pair within a DNA duplex is slightly destabilizing (compared to T-A), but the duplex forms a B-type duplex.[128,129,133] Recently, a 2´-deoxy-5-trifluoromethylcytidine was incorporated into DNA by solid-phase synthesis and was demonstrated to be a sensitive analogue for ^{19}F NMR studies.[130]

5-[4,4,4-Trifluoro-3,3-bis(trifluoromethyl)but-1-ynyl]-2´-deoxyuridine bears nine magnetically equivalent fluorine atoms in an isolated spin system.[106] This nucleoside has been introduced into oligonucleotides in place of thymidine[106] or uridine[24,89] by solid-phase oligonucleotide synthesis, and it has been reported to be slightly destabilizing. Other sterically demanding analogues contain a 3,5-bis(trifluoromethyl)phenyl group.[55] This is a sensitive reporter group containing six magnetically equivalent fluorine atoms that are not coupled to protons. In contrast to other fluorinated nucleoside analogues, the fluorine moieties of 3,5-bis(trifluoromethyl)phenyl-modified 2´-deoxyuridine analogues have been introduced into DNA postsynthetically (i.e., 3,5-bis(trifluoromethyl)benzoic acid and 4-[3,5-bis(trifluoromethyl)benzamido]benzoic acid were coupled to 5-(3-aminopropyn-1-yl)-2´-deoxyuridine residue after solid-phase synthesis). These 3,5-bis(trifluoromethyl)phenyl-modified analogues do not substantially effect the stability of the DNA duplexes; the duplexes maintain a B-type conformation, but deviations from the native secondary structures can be detected. 5-[4´´-(5´´´-Aminobenzo[*d*]oxazol-2´´´-yl)-2´´-fluorophenyl]-2´-deoxyuridine is an example of a dual probe, but its utility has mainly been in fluorescence detection.[100] Recently, the applicability of a 5-[(3,5-bis(trifluoromethyl)phenyl)ethynyl]-2´-deoxyuridine was demonstrated.[19] In contrast to the 3,5-bis(trifluoromethyl)phenyl group bearing 2´-deoxyuridine analogues discussed above, it has been incorporated into

DNA as a phosphoramidite. In addition, an enzymatically incorporated 4-(trifluoroacetyl)phenyl-modified 2´-deoxyuridine and 2´-deoxycytidine were recently presented as novel derivatives for ^{19}F NMR studies.[101]

In addition to fluorinated natural pyrimidine nucleobases, the well-known nonpolar pyrimidine isosteres 2,4-difluorotoluene and 2,4-difluorobenzene nucleoside analogues have been employed in a few ^{19}F NMR studies (Figure 3).[84,85,93,134] In these studies, the nucleoside analogues have been incorporated into oligonucleotides as corresponding nucleoside phosphoramidites. The fluorine atoms are magnetically non-equivalent and coupled to protons. Because these isosteres have been extensively used as base mimics, their properties have been thoroughly investigated.[135-137] The shape of a 2,4-difluorotoluene/2,4-difluorobenzene-adenine base pair closely resembles the natural thymine/uracil-adenine Watson-Crick base pair, but there are no hydrogen bonding interactions. Therefore, the 2,4-difluorotoluene/2,4-difluorobenzene substitution is typically structurally non-perturbing, but it decreases the stability of the double helix.

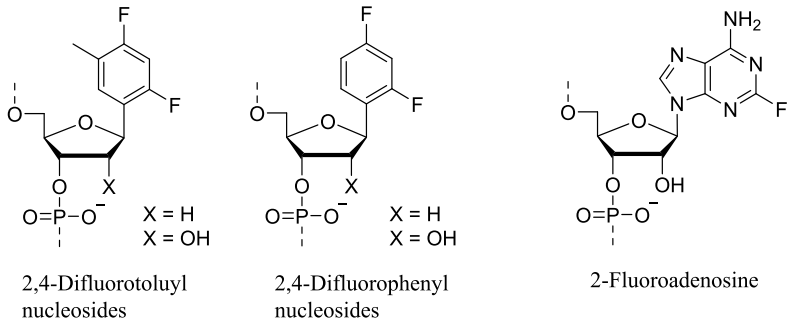


Figure 3. Structures of pyrimidine isosteres and 2-fluoroadenosine.

In contrast to fluorine-modified pyrimidine bases, the availability of fluorine-modified purine bases is limited, presumably due to the modest stability of these molecules.[138-140] Currently, only 2-fluoroadenosine has been employed more generally in ^{19}F NMR studies (Figure 3).[21,25,29,91] 2-Fluoroadenosine-5´-triphosphates can be incorporated efficiently into RNA by enzymatic methods.[21,25,29] The fluorine substituent is unstable in chemical RNA synthesis[139,140], and thus the utility of 2-fluoroadenosine has been limited to uniformly substituted enzymatically produced RNAs[21,25,29]. Studies on uniformly substituted RNAs have shown that 2-fluoroadenosine is a sensitive non-invasive analogue for ^{19}F NMR studies.[21,25,29] The fluorine is oriented into the minor groove. When introduced in a double-helical RNA, 2-fluoroadenine forms a stable base pair with uracil and mild destabilizing perturbations observed in local base pair stabilities are compensated by favourable base stacking interaction. In addition to 2-fluoroadenosine, some other purine analogues, including 7-fluoro-7-deaza-2´-deoxyadenosine[140], 7-fluoro-7-deaza-2´-deoxyguanosine[140], 2-fluoro-7-deazaadenosine[91], 7-fluoro-7-deazaadenosine[91], 6-trifluoromethylpurine ribonucleoside[141], and fluorinated biaryl[100]

and 4-(trifluoroacetyl)phenyl-modified[101] 7-deazapurine 2´-deoxynucleosides, have been proposed as potential derivatives for ^{19}F NMR studies.

Figure 4 presents more specialized fluorine-labelled nucleoside analogues, and these analogues are compatible with standard solid-phase synthesis. Fluorinated 2´-deoxy-*N*^{4}-(6-aminopyridin-2-yl)cytidine derivatives have been designed to interact with duplex DNA via Hoogsteen hydrogen bonding and have been employed to investigate triple-helical DNA structures.[58] 8-[(3,5-Bis(trifluoromethyl)phenyl)ethynyl]-2´-deoxyguanosine is a dual probe, and it has been developed for the detection of G-quadruplexes.[39] Due to the 3,5-bis(trifluoromethyl)phenyl group at the 8-position, guanine favours a *syn* orientation, which is favourable in the G-quadruplexes that have been investigated by this analogue. In addition, fluorinated arylamine guanosine DNA adducts[33,34] have been exploited in ^{19}F NMR studies, and these are presented in Section 1.2.2.10.

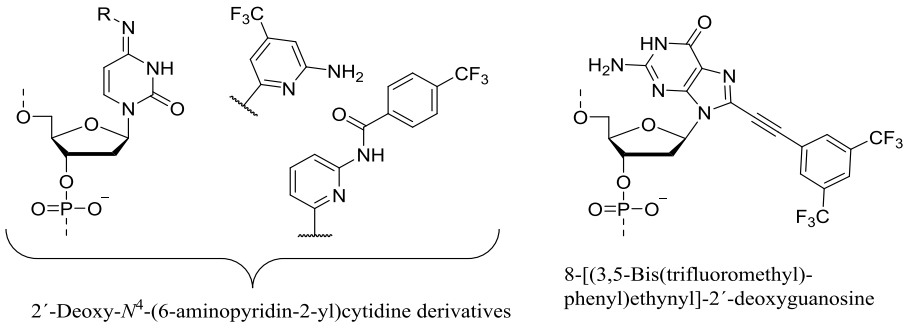


Figure 4. Structures of fluorinated nucleoside derivatives designed for specific nucleic acid structures.

1.2.1.2 Sugar modifications

Few studies have utilized sugar-modified nucleoside analogues and only two sugar modifications have been employed. One modification is a 2´-deoxy-2´-fluoro (2´-F)[17,22,61,82,142], and the other is a 2´deoxy-2´-trifluoromethylthio (2´-SCF_3) modification, which has been designed for ^{19}F NMR applications[94,107,143] (Figure 5a). Recently, 2´-fluoro-2´-deoxyarabinocytidine was proposed as a possible analogue for monitoring small molecule-DNA i-motif interactions.[144] One advantage of 2´-F and 2´-SCF_3 modifications is that the label can be incorporated into any position in the RNA.[94,107,142,143] Modifications in the sugar part, however, affect conformational properties and may cause significant alterations in the structure and stability of nucleic acids.[82,94,107,143,145] In addition, plausible 2´-hydroxyl group-mediated interactions are disrupted.

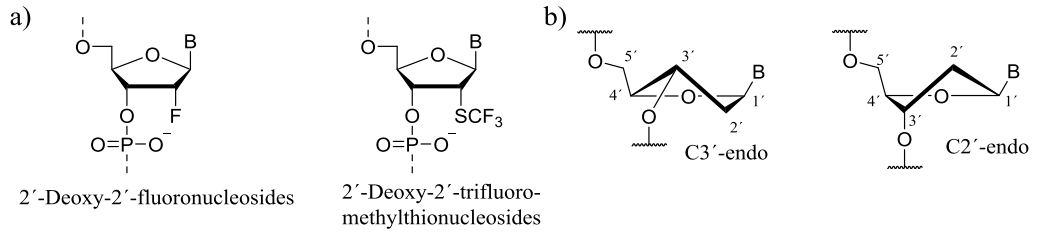


Figure 5. (a) Sugar-modified nucleoside analogues. Base (B) is adenine (Ade), uracil (Ura), cytosine (Cyt) or guanine (Gua). (b) C3′-endo (*N*-type, A-form) and C2′-endo (*S*-type, B-form) sugar conformations.[146]

2′-F and 2′-SCF_3 modifications can be incorporated into oligonucleotides using standard solid-phase oligonucleotide synthesis.[94,107,142,143] 2′-Deoxy-2′-fluoronucleosides predominantly adopt a C3′-endo sugar conformation that prefers RNA duplexes (A-form) (Figure 5b).[147-150] Thus, a 2′-F substitution in an RNA duplex is usually structurally non-perturbing and slightly increases the stability of the double helix. The strong preference to C3′-endo pucker, however, affects the C2′/C3′-endo equilibrium, and a 2′-F modification may cause structural perturbations, for instance when positioned into a loop[82]. Compared to 2′-F modification, the 2′-SCF_3 group is a more sensitive fluorine label with three magnetically equivalent and spin-isolated fluorine atoms.[94,107,143] In contrast to 2′-F, the 2′-SCF_3-substituted ribose favours a C2′-endo conformation and thus significantly decreases the stability of an RNA duplex.

1.2.1.3 Other modifications

A few ^{19}F NMR studies have utilized non-nucleosidic modifications at the 5′-terminus (Figure 6). As discussed, 3,5-bis(trifluoromethyl)phenyl is a sensitive reporter group for ^{19}F NMR studies. This group has been incorporated via an appropriate linker into the 5′-end of both DNA[18,53] and RNA[37] oligonucleotides using solid-phase synthesis, and the applicability of the group for the detection of G-quadruplexes has been demonstrated[18,37,38]. In addition, some preliminary ^{19}F NMR studies have been conducted with oligodeoxynucleotide 5′-fluorodiphosphates.[102]

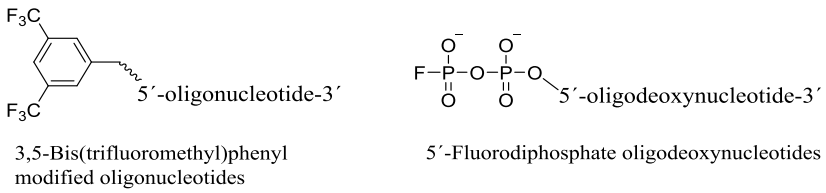


Figure 6. Structures of non-nucleosidic fluorine modifications at the 5′-end of oligonucleotides.

1.2.2 ^{19}F NMR spectroscopy studies of nucleic acids

1.2.2.1 Early studies and tRNA

Horowitz and co-workers published their research on 5-fluorouridine-modified tRNA in 1977[16], and in the same year, Marshall and Smith reported their ^{19}F NMR studies on 5-fluorouridine-modified 5S rRNA.[15] Fluorine-labelled tRNAs were then extensively studied over the following years[99,151-164], but little research was conducted with fluorine-labelled 5S rRNA[165]. tRNA studies have primarily focused on *Escherichia coli* tRNA^{Val}. 5-Fluorouridine-substituted tRNAs were isolated from *E. coli* cells grown in the presence of 5-fluorouracil.[16,166] This biosynthetic method provided tRNAs in which over 90 % of uridine residues were replaced by 5-fluorouridine[16], and this method was later replaced by enzymatic *in vitro* synthesis[99]. The fluorinated tRNA^{Val} contained fourteen 5-fluorouridine residues (Figure 7a).

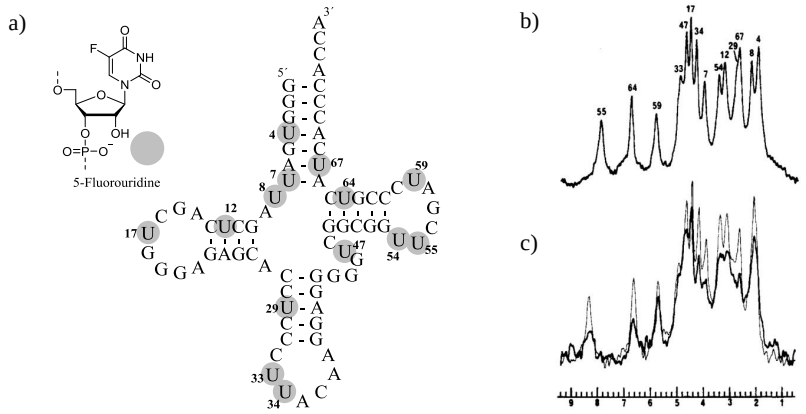


Figure 7. (a) 5-Fluorouracil-substituted *E. Coli* tRNA^{Val} cloverleaf structure (residues that are replaced by 5-fluorouridine are indicated as grey spheres) (b) and its ^{19}F NMR spectrum. (c) ^{19}F NMR spectra of 5-fluorouracil-substituted *E. Coli* tRNA^{Val} in the absence (thin line) and in the presence of valyl-tRNA-synthetase (solid line). Spectra reprinted with permission from Chu, W. C.; Horowitz, J. *Biochemistry* **1991**, *30*, 1655-1663. Copyright (1991) American Chemical Society.[159]

In addition to currently employed modern NMR strategies (HOESY and NOESY)[155,157,161], a variety of other approaches were used to assign the fluorine signals[16,99,152-161]. For instance chemical modification, pH dependency, solvent accessibility, mutation and interaction studies were performed[16,99,152-161], and this laborious signal assignment was not completed until 1991-1992[160,161]. In summary, a variety of investigations were conducted with fluorine-labelled tRNAs, including studies

of tRNA^{Val}-RNA[153,156], tRNA^{Val}-protein (cognate synthetase)[159,163,164] (Figure 7b and c) and tRNA^{Val}-small molecule[154,162] binding interactions, ion induced conformational changes[154,156,157] and temperature-dependent unfolding[158], thus demonstrating the sensitivity of the fluorine to local environment and the diversity of ^{19}F NMR-based approaches.

1.2.2.2 Short oligonucleotides and conformational heterogeneity

Many structural studies have utilized ^{19}F-labelled nucleosides that are able to provide well-separated signals in single-stranded (ss) and double-stranded (ds) environments. Employing this labelling strategy, both the intra- and intermolecular secondary structure equilibria of short oligonucleotides have been successfully analysed by ^{19}F NMR spectroscopy. Additionally, studies of more complex nucleic acid structures often rely on labels' single strand-double strand recognition ability.

Intramolecular hairpin-hairpin equilibria have been investigated to demonstrate the general ability of labels to distinguish coexisting nucleic acid structures quantitatively and to evaluate the influence of the fluorine modifications on nucleic acid structures. As a representative example, Figure 8 shows the ^{19}F NMR spectra of 2´-F-labelled RNAs.[142] As observed, the 2´-F labels were sensitive to secondary structures, and unambiguous signals were detected for alternative hairpin conformers. Consequently, the ratios of coexisting conformers were easily determined from the relative peak areas of the ^{19}F signals. However, the 2´-F pyrimidines shifted the equilibrium approximately 20-25 % toward the structures in which 2´-F pyrimidines exist in a double-helical region. In contrast, the equilibrium positions were unaffected with 2´-F purines.

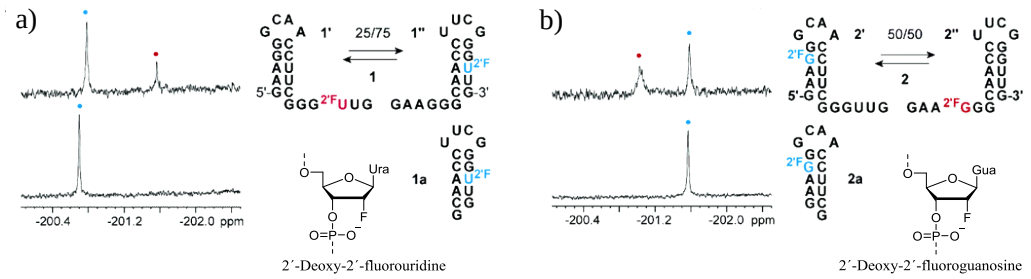


Figure 8. ^{19}F NMR spectra (^{1}H decoupled) of (a) 2´-F uridine ($^{2'F}$U)- and (b) 2´-F guanosine ($^{2'F}$G)-modified bistable RNAs and their reference hairpins. Reprinted and modified with permission from Kreutz, C.; Kählig, H.; Konrat, R.; Micura, R. *J. Am. Chem. Soc.* **2005**, *127*, 11558-11559. Copyright (2005) American Chemical Society.[142]

Furthermore, 5-fluoropyrimidine-modified DNA and RNA sequences adopting mutually exclusive hairpin structures have been investigated comprehensively.[111] In each case, the 5-fluoro pyrimidine label provided well separated ^{19}F signals, and the equilibrium position remained unaffected, thus demonstrating the non-invasive nature of the 5-F modification. Interestingly, it was noted that the chemical shift difference between DNA folds (i.e., between ss and ds environments) was clearly smaller than the chemical shift difference between RNA folds. It was suggested that the difference between chemical shift dispersion was caused by the different stacking patterns within A- and B-form duplexes, and therefore the fluorine experiences a much larger environment change between RNA folds compared with DNA folds. Likewise, with the 2´-F and 5-fluoro pyrimidine labels, 2´-SCF_3 labels have provided well-separated ^{19}F signals for alternative hairpin structures, but the effect of the 2´-SCF_3 group on the equilibrium position depends strongly on the labelling site.[107,143]

The effect of the surrounding conditions to induce the dimerization of short oligonucleotides has also been analysed with 5-fluoropyrimidines. These studies have taken advantage of the suitability of ^{19}F NMR spectroscopy to a variety of measurement conditions. For example, the dimerization of an RNA hairpin and its tethered model and corresponding DNA models (Figure 9a) were successfully investigated at different pH and in different Mg^{2+}, Na^+ and PEG400 (polyethylene glycol) concentrations.[115,116] In another study the effect of Na^+, K^+, Sr^{2+}, Mg^{2+}, Ca^{2+} or Ba^{2+} on the hairpin-dimer equilibrium of an RNA was examined (Figure 9b).[90] ^{19}F NMR data provided information on the contribution of electrostatic, entropic and enthalpic factors to the hairpin-dimer equilibrium and revealed that some of the cations induced formation of additional hairpin conformations.

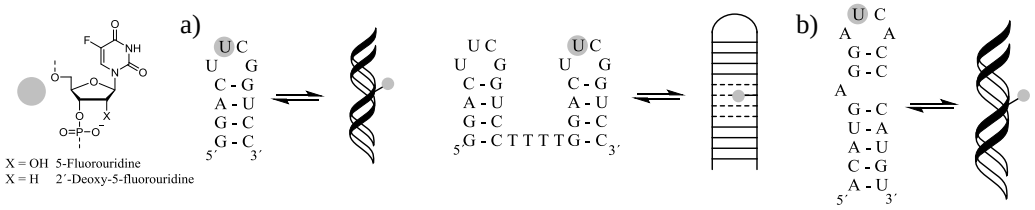


Figure 9. Schematic presentation of hairpin-dimer equilibria. (a) An RNA hairpin and its tethered model containing a 2´-deoxy-5-fluorouridine. In the case of DNA analogues U is T. (b) 5-Fluorouridine-modified RNA hairpin model of the translational operator of the MS2 bacteriophage. Also a hairpin model containing three 5-fluorouridine residues was used.

In a recent study, 2´-deoxy-5-trifluoromethylcytidine and 2´-deoxy-5-trifluoromethyl-uridine were used to probe a salt-induced formation of a left-handed Z-DNA.[130] The labels were able to distinguish B-form, Z-form and ssDNA structures. Consequently, the distribution of these structures was easily determined from well-separated ^{19}F signals in varying ionic conditions and different temperatures.

The ability of the labels to discriminate between structural species has enabled the use of ^{19}F NMR to study the thermal stability of nucleic acid structures. The melting of nucleic acid structures can be monitored in the fluorine signals upon stepwise heating[33,34,93,158], and thermodynamic parameters[19,37,90,96] and T_m values[93] can be extracted from the temperature-dependent ^{19}F NMR data. The T_m determination method[93] has gained popularity and has been employed to characterize the melting behaviour of a variety of structures.[19,24,36-39,55,89,94,95] This application was demonstrated with an intermolecular RNA duplex and an intramolecular RNA stem loop structure using a 2,4-difluorotoluyl ribonucleoside.[93] Melting of the intramolecular stem loop structure occurred in a fast exchange regime (Figure 10). When the chemical shift of the ^{19}F signal was plotted against temperature, a sigmoidal curve was obtained, and the T_m value of the stem loop structure was obtained from its inflection point. The melting of the intermolecular duplex was in the slow exchange regime. The sigmoidal melting curve and thus the T_m value was obtained by plotting a fraction of a single strand against temperature. Importantly, in both cases, the T_m value was comparable with the value obtained by temperature-dependent UV spectroscopy. Additionally, it was simultaneously possible to monitor the temperature-dependent behaviour of coexisting structures from their distinct ^{19}F signals. This is a clear advantage over traditional spectrophotometric methods (UV and CD)[73-78], which often cannot provide information on the melting of coexisting structures. Compared to UV and CD spectroscopy, ^{19}F reflects melting at a local level. Therefore, ^{19}F NMR can be used to probe the temperature-dependent behaviour of a specific part of a nucleic acid structure, which typically is difficult with UV and CD spectroscopy. ^{1}H NMR of imino protons can also be used for melting analysis, but spectral crowding and the T_m-independent exchange rate of the imino protons can interfere with this method.[167,168] It should be noted that the chemical shift of ^{19}F is inherently very sensitive to temperature[11,59], and thus the temperature-dependent behaviour of the reporter group needs to be considered when analyzing the data[24].

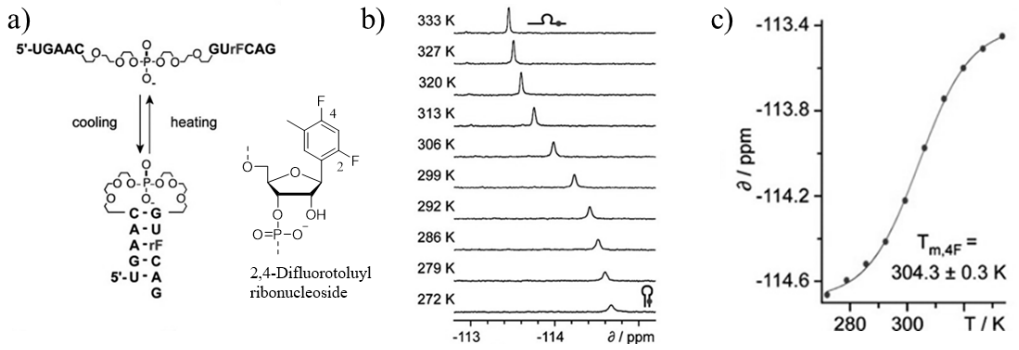


Figure 10. (a) 2,4-Difluorotoluyl ribonucleoside (rF)-modified intramolecular stem loop structure. (b) ^{19}F NMR spectra (^{1}H decoupled) at different temperatures. Only the signal of fluorine in position 4 is shown. (c) Melting profile obtained from the ^{19}F NMR data. Reprinted and modified with permission from Graber, D.; Moroder, H.; Micura, R. *J. Am. Chem. Soc.* **2008**, *130*, 17230-17231. Copyright (2008) American Chemical Society.[93]

1.2.2.3 Spliceosomal U2-U6 snRNA complex

The spliceosome is a ribonucleoprotein complex that removes introns from eukaryotic precursor messenger RNAs (pre-mRNAs) in a process called splicing.[169] The spliceosome is composed of numerous proteins and five small nuclear RNAs (snRNAs), including U2 and U6 snRNAs. The interaction between U2 and U6 snRNAs creates the splicing scaffold. 5-Fluorocytidine has been employed to investigate the conformational heterogeneity of the protein-free human spliceosomal U2-U6 snRNA complex, specifically to investigate the formation of a four-helix conformation and a tree-helix conformation and to study the potential equilibrium between these conformers (Figure 11).[31,32] The ^{19}F NMR spectrum of the U2-U6 snRNA complex exhibits a major ^{19}F signal, which was attributed to a double-stranded form and thus indicates the formation of the four-helix junction conformation. In addition, a very broad signal was detected in the single-stranded region corresponding to the three-way junction conformation. The complexity of the signal was thought to originate from multiple alternative conformations or intermediates. Several fluorine-labelled mutation models were investigated to confirm the signal assignment. Because an analogous conformational equilibrium of the yeast complex has been shown to depend strongly on Mg^{2+} concentration, the effect of Mg^{2+} on the human U2-U6 snRNA complex was also investigated. In the presence of Mg^{2+} a slight shift in the equilibrium was detected; the percentage of the heterogeneous minor conformer increased from ~14% to ~17%.

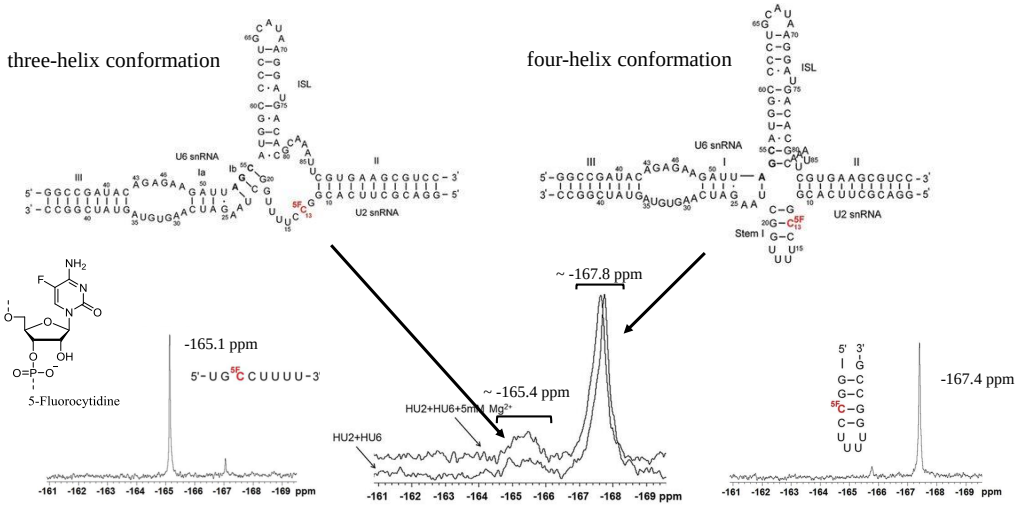


Figure 11. Proposed secondary structures of the human U2-U6 snRNA complex model. The ^{19}F NMR spectra of control RNAs and the U2-U6 snRNA complex in the absence and in the presence of magnesium. The 5-fluorocytidine (^{5F}C) is in red. Reprinted and modified with permission from Zhao, C.; Bachu, R.; Popović, M.; Devany, M.; Brenowitz, M.; Schlatterer, J. C.; Greenbaum, N. L. *RNA* **2013**, *19*, 561-573. © 2013 RNA Society.[31]

Two-dimensional ^{19}F-^{19}F EXSY (exchange spectroscopy) experiments were additionally conducted to obtain information about conformational exchange dynamics[32]; this method had previously been successfully applied to the hairpin-hairpin equilibria of a short 5-fluorouridine-modified RNA[97]. The ^{19}F-^{19}F EXSY data indicated that there was a dynamic exchange between different conformations, but the complexity and low quality of the spectra prevented more detailed kinetic analysis.[32] Overall, the ^{19}F NMR results were consistent with biochemical studies.[31,32] It was concluded that the predominant four-helix conformation is in dynamic equilibrium with the minor three-helix conformation, and facile interconversion between these conformations occurs.

1.2.2.4 Riboswitches

Riboswitches are ligand-sensing regulatory elements most commonly found in the noncoding regions of some bacterial mRNAs.[170,171] A typical riboswitch consists of two domains: an aptamer domain, which contains a selective high-affinity binding site for a ligand, and an adjoining expression platform, which controls gene expression. The formation of a stable aptamer-ligand complex induces conformational changes in the expression platform and leads to the regulation of gene expression. ^{19}F NMR spectroscopy has been used to investigate bacterial riboswitches, including preQ$_1$ class I riboswitches[26,28,94,111], an S-adenosylmethionine type II (SAM-II) riboswitch[27],

a guanine-sensing riboswitch[29], and an artificial tobramycin-sensing riboswitch[107]. Conformational equilibria of the riboswitches in the absence and presence of ligands can be detected and quantified by ^{19}F NMR analysis and complemented by other detection methods to offer insight into folding pathways. In addition, ^{19}F NMR has provided information about how different factors, such as Mg^{2+} ions[27,28,94], temperature[27,28], base mutations[28] or competitive artefact[28], affect structural equilibria. The $preQ_1$ class I riboswitches are discussed in more detail below; ^{19}F NMR spectroscopy has been applied to the detection of both aptamer-ligand complexes[28,94,111] and structural changes in the expression platform in response to ligand binding[26].

The minimal aptamer domain (34 nt) of a *Fusobacterium nucleatum* $preQ_1$ class I riboswitch forms a pseudoknot structure upon 7-aminomethyl-7-deazaguanine ($preQ_1$) binding.[172] ^{19}F NMR spectroscopy was used to analyse the structural behaviour of the wild-type aptamer and its mutant form (a single base pair mutation, A33G/U9C, Figure 12b) using 5-fluoro pyrimidine labels.[28] The ^{19}F NMR data indicated that the mutant aptamer adopted two conformational states in the absence of $preQ_1$ (Figure 12b), and this result was further confirmed with ^{1}H and ^{15}N NMR spectroscopy. In addition to the stem-loop structure, the aptamer adopted a pseudoknot conformation that structurally resembles the ligand-bound complex and exists in a slow dynamic equilibrium with the stem-loop structure. The pseudoknot formation stabilized at a low temperature and in the presence of Mg^{2+} ions.

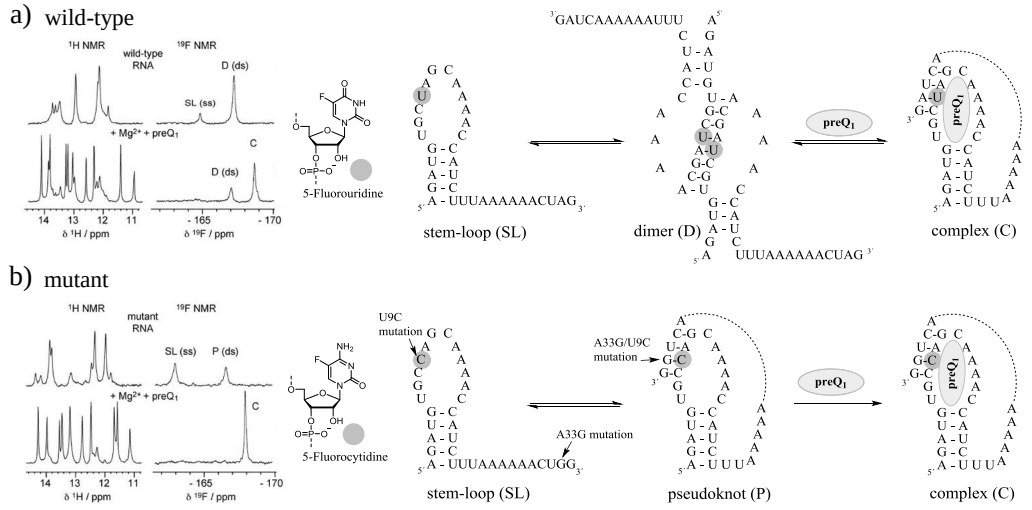


Figure 12. Comparison of the conformational behaviour of wild-type and mutant *F. nucleatum* $preQ_1$ riboswitch aptamers. ^{1}H NMR and ^{19}F NMR spectra of free (upper) and $preQ_1$-bound (lower) (a) 5-fluorouridine-modified wild-type RNA and (b) 5-fluorocytidine-modified mutant RNA. Proposed structures are shown. Spectra reprinted and modified with permission from Santner, T.; Rieder, U.; Kreutz, C.; Micura, R. *J. Am. Chem. Soc.* **2012**, *134*, 11928-11931. Copyright (2012) American Chemical Society.[28]

The ^{19}F NMR data showed that the wild-type aptamer was unable to adopt a preorganized pseudoknot structure.[28] Instead, a clear indication of a proposed dimeric structure was observed. As observed in Figure 12a, the dimer was a dominant structure in the absence of the ligand. In addition, a trace of the dimer was detected in the presence of $preQ_1$, whereas only the ^{19}F signal attributed to the $preQ_1$ complex was observed for the mutant aptamer (Figure 12b). This difference thought to be caused by the single base pair mutation because it not only stabilizes the pre-organized pseudoknot fold but also destabilizes the dimer formation. The dimer was assigned as an artefact formed in high RNA concentrations, and it was confirmed using a fluorescence experiment that the wild-type aptamer adopted a pre-organized pseudoknot structure in the presence of magnesium at a low RNA concentration (0.5 µM versus 600 µM used in ^{19}F NMR experiments). The direct evidence of a competitive structure provided an explanation for data previously obtained by small-angle X-ray scattering. Similarly, ligand-free state, a Mg^{2+}-induced pseudoknot structure and a ligand bound state ($preQ_1$) of another $preQ_1$ class I riboswitch aptamer domain have been successfully monitored with a 2´-SCF_3 guanosine (Figure 13a).[94]

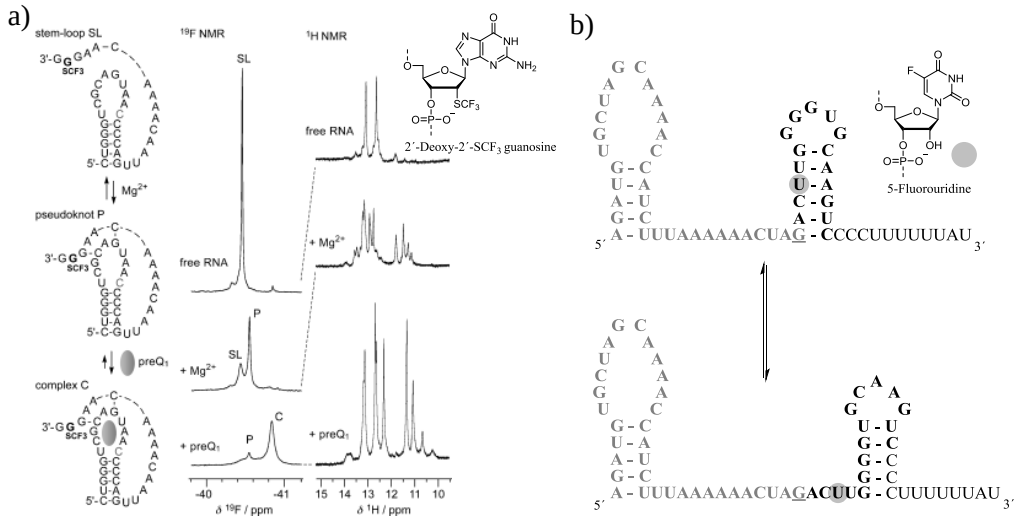


Figure 13. (a) ^{19}F NMR and ^{1}H NMR analysis of ligand-free, Mg^{2+}-induced pseudoknot fold and $preQ_1$-bound state of 2′-SCF_3 guanosine-modified *Thermoanaerobacter tengcongensis* $preQ_1$ class I riboswitch. Reprinted and modified with permission from Jud, L.; Košutić, M.; Schwarz, V.; Hartl, M.; Kreutz, C.; Bister, K.; Micura, R. *Chem. Eur. J.* **2015**, *21*, 10400-10407 © 2015 The Authors.[94] (b) 5-Fluorouridine-modified bistable secondary structure model of the *F. nucleatum* $preQ_1$ riboswitch. The minimal aptamer domain (grey bold letters, same as Figure 12a) and the bistable hairpin sequence within the expression platform (black bold letters) share a single nucleotide (G34 underlined). The binding of $preQ_1$ (aptamer domain adopts a $preQ_1$-bound state) shifts the equilibrium position towards the lower stem-loop structure (i.e., termination hairpin).

^{19}F NMR spectroscopy has also been used to detect the ligand-induced rearrangement on the expression platform of the *F. nucleatum* preQ$_1$ class I riboswitch (Figure 13b).[26] Chemical and enzymatic structural probing indicated that the expression platform contains a hairpin-hairpin equilibrium that is sensitive to ligand-induced rearrangement at the aptamer domain. Since a full-length riboswitch (61 nt) was also required for more detailed investigations, ^{1}H NMR spectroscopy was unsuitable method because of spectral crowding. Using 5-fluorouridine labelling, it was possible to confirm that the expression platform adopted a hairpin-hairpin equilibrium and to directly monitor the preQ$_1$-induced changes in the equilibrium position.

1.2.2.5 Hammerhead ribozyme

The metal ion-induced folding of hammerhead ribozyme has been monitored by ^{19}F NMR to obtain local information on the folding process.[30] Hammerhead ribozyme is a small catalytic RNA that undergoes autocatalytic self-cleavage in the presence of Mg^{2+} ions. The catalytic core of the minimal hammerhead ribozyme is flanked by three helical stems. To acquire a functional conformation the core undergoes structural rearrangements in the presence of Mg^{2+} ions. One of the 2´-*O*-methyl-5-fluorouridines (U4 in Figure 14) was introduced into domain 1 and the other was located at the interface between domains 1 and 2 (U7 in Figure 14) of the cloverleaf construct.

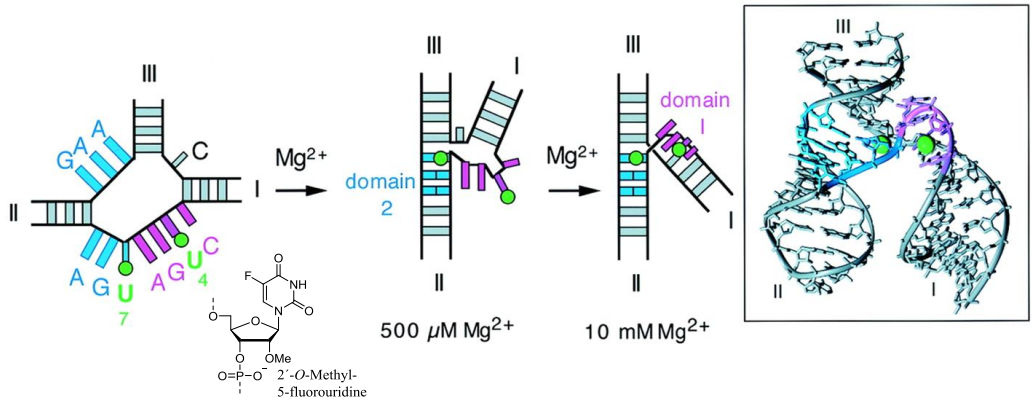


Figure 14. Schematic presentation of a two-stage folding process for a hammerhead ribozyme. The 2´-*O*-methyl-5-fluorouridines U4 and U7 are shown as green spheres. Crystal structure of hammerhead ribozyme. Reprinted and modified with permission from Hammann, C.; Norman, D. G.; Lilley, D. M. J. *PNAS* **2001**, *98*, 5503-5508. Copyright (2001) National Academy of Sciences.[30]

Both of the labels were sensitive to titration with Mg^{2+} ions; thus, it was possible to monitor the folding process from the chemical shifts and linewidths. With label U7, two

different transitions were detected upon the addition of Mg^{2+} ions. The first transition at a low Mg^{2+} concentration (500 µM) was attributed to an initial folding event i.e., the formation of domain 2. As the linewidth of the signal significantly broadened between 0 and 500 µM Mg^{2+}, and then subsequently narrowed between 500 µM and 1 mM Mg^{2+}, the folding process was interpreted as an intermediate exchange regime (a 500 MHz instrument was used for the analysis). It was estimated that the formation of domain 2 occurs on approximately the millisecond timescale. The second transition was detected at millimolar Mg^{2+} concentrations. The same transition was also observed with label 4, and thus, it was attributed to the formation of domain 1. The chemical shift data of the second transition showed non-cooperative binding of Mg^{2+} ions with apparent association constants of approximately 100-500 M^{-1}. In summary, the ^{19}F NMR data supported the two-stage ion-induced folding mechanism in which the sequential conformational transitions of domains 2 and 1 create the active conformation of the minimal hammerhead ribozyme.

1.2.2.6 Trans-activation response element

The trans-activation response element (TAR) is a conserved RNA stem loop structure found at the beginning of human immunodeficiency virus (HIV) transcripts.[173] The interaction between TAR and the viral protein Tat has an essential role in viral transcription, and TAR has been extensively studied as a potential therapeutic target.[174,175] ^{19}F NMR spectroscopy has been used to investigate the binding of metal ions[20], argininamide[20,25] and invader oligonucleotides[24,89] to fluorine-labelled TAR models. In addition, solid-state ^{19}F NMR spectroscopy has been applied to investigate Tat peptide-TAR interactions.[22,64,65]

Metal ion-TAR interactions have been explored using 5-fluorouridine-modified HIV type-1 (HIV-1) TAR models (Figure 15).[20] Several 5-fluorouridines were introduced within TAR models to determine the binding sites and K_d values for metal ions (Mg^{2+}, Ca^{2+} and $Co(NH_3)_6^{3+}$). The applicability of this approach was first evaluated by monitoring the chemical shift changes induced by the known TAR ligand argininamide as a function of its concentration. Identifying the locations of binding sites for small molecules has also been conducted previously with other RNA targets.[162,176]

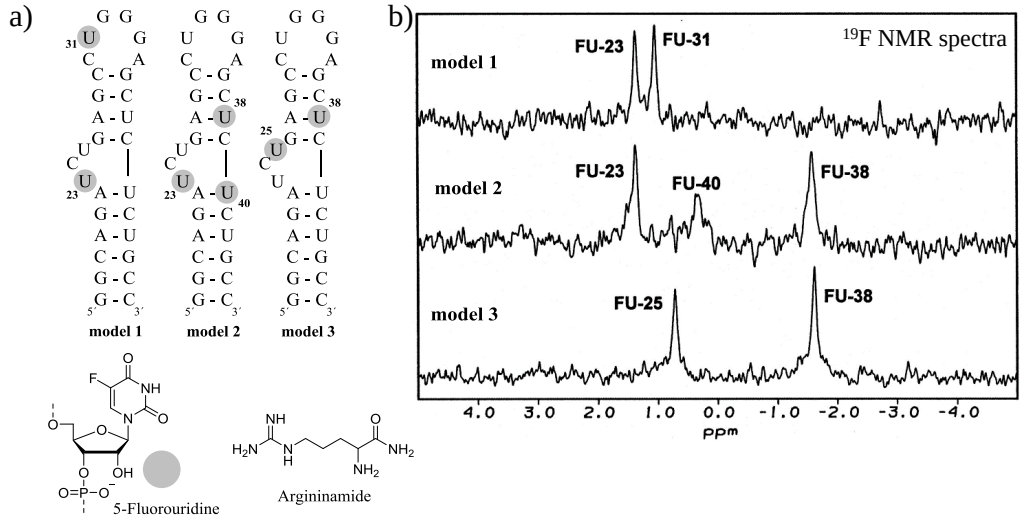


Figure 15. (a) 5-Fluorouridine-modified HIV-1 TAR models and (b) their ^{19}F NMR spectra. Spectra reprinted and modified from Olejniczak, M.; Gdaniec, Z.; Fischer, A.; Grabarkiewicz, T.; Bielecki, L.; Adamiak, R. W. *Nucleic Acids Res.* **2002**, *30*, 4241-4249, by permission of Oxford University Press Copyright 2002, Oxford University Press.[20]

Structural studies have shown that the bulge region undergoes major structural rearrangement upon the binding of argininamide (Figure 16b and c).[177-179] The residue U23 undergoes significant secondary structural rearrangement and may participate in base triplet (U23-A27-U38) formation within the argininamide-TAR complex. As expected, titration experiments showed that the labelled residues positioned within the bulge region (i.e., FU-23, FU-25, FU-38 and FU-40) were sensitive to the addition of argininamide, whereas negligible changes were observed for residue FU-31, which is positioned at the loop.[20] Notably, a large chemical shift change was detected for FU-23. Because the argininamide titration experiment was consistent with structural studies and was able to provide site-specific information, the binding of Mg^{2+}, Ca^{2+} and $Co(NH_3)_6^{3+}$ ions was investigated by the same technique. The titration data showed that the metal ions bind to the bulge region. In addition, the K_d values for argininamide (0.3 mM), Mg^{2+} (0.9 mM) and Ca^{2+} (2.7 mM) were obtained from the titration data. $Co(NH_3)_6^{3+}$ caused the irreversible aggregation of RNA, and thus its K_d value could not be determined.

A similar approach was successfully used to investigate the argininamide binding in uniformly labelled HIV-2 TAR models (Figure 16a).[25] The concept was the same as described above; ^{19}F chemical shift changes were followed upon the consecutive addition of argininamide, and the obtained data were used for the localization of the binding site and to determine the K_d values. It was proposed that this technique could be used to identify novel TAR-binding small molecules. In addition to the studies with uniformly labelled tRNAs, this is one of the few examples that has utilized a uniform labelling strategy. These TAR models have also been used for solvent accessibility

measurements (i.e., solvent-induced isotope shift (SIIS) studies) and detailed structural studies.[21,23,25,83]

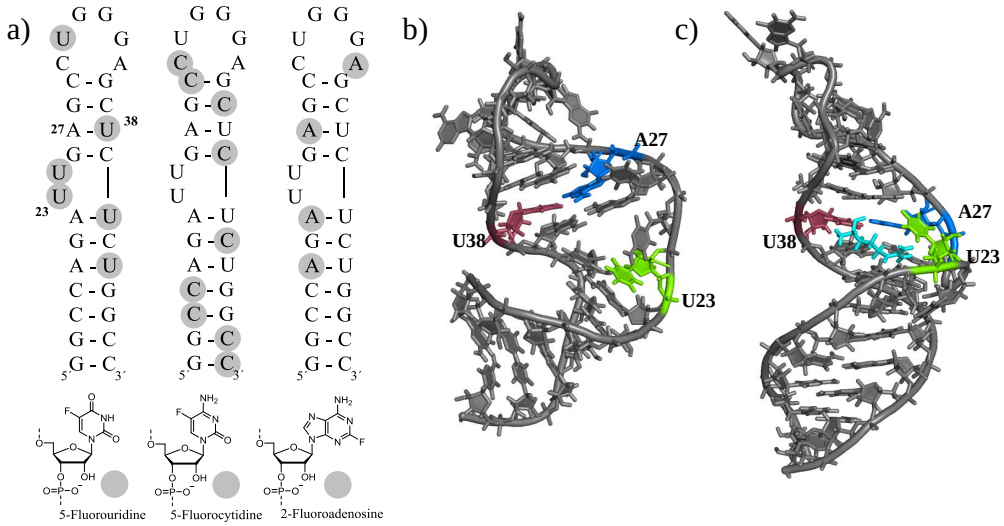


Figure 16. (a) Uniformly labelled HIV-2 TAR models. Structures of unlabelled (b) HIV-1 TAR (PDB ID 1ANR)[177] and (c) HIV-2 TAR argininamide (cyan) complex (PDB 1AJU)[179]. Residues U23, A27 and U38 are coloured to illustrate the secondary structure rearrangement of the bulge region. Consistent numbering is used for the HIV-1 (bulge U23, C24, U25) and HIV-2 (bulge U23, U25) TAR models.

As described above, residue U23 is part of a marked secondary structural rearrangement upon the binding of argininamide. Similarly, Tat peptide binding induces a significant conformational rearrangement at the bulge region and affects the position of residue U23.[180,181] The binding of Tat peptide has been monitored by rotational-echo double-resonance (REDOR) NMR using fluorine-labelled HIV-1 TAR models.[22,64,65] REDOR NMR is a solid-state NMR technique for measuring dipolar couplings and hence distances between a heteronuclear spin pair. Using a ^{31}P-^{19}F REDOR experiment, the distance between a 2´-F-modified U23 and a phosphorothioate modified A27 was shown to change from 10.3 to 6.6 Å upon Tat peptide binding.[22] In addition, the Tat-TAR interactions were investigated by combining several REDOR experiments using ^{15}N- and ^{13}C-labelled Tat peptide and a 5-fluorouridine-modified TAR model.[64,65] Both the intramolecular[22] and intermolecular[64,65] distances were comparable with solution NMR studies, thus demonstrating the utility of the REDOR method to study RNA-peptide interactions and RNA conformational changes.

The invasion of 2´-*O*-methyl oligoribonucleotides into a fluorine-labelled HIV-1 TAR model is a more complex example of an intermolecular nucleic acid-nucleic acid interaction that has been successfully monitored by ^{19}F NMR.[24,89] The invasion of oligonucleotides was clearly detected from distinct ^{19}F signals of the hairpin and of the

macro-loop invasion complex (Figure 17a and 17b), and K_d values could be extracted from the titration data. Thus, ^{19}F NMR provided a direct way to compare the ability of different oligonucleotides to invade the fluorine-labelled TAR model. The invasion was also monitored in the presence of the aminoglycoside neomycin. Neomycin is known to bind to the minor groove below the bulge. It was suggested that this interaction might affect the invasion, and titration data confirmed that neomycin clearly promoted the invasion. Moreover, it was later shown that the covalent conjugation of neomycin to 2´-*O*-methyl oligoribonucleotides significantly promotes invasion.[89] In addition to titration studies, the temperature-dependent behaviour of invasion complexes was monitored by ^{19}F NMR.[24,89] As illustrated in the Figure 17a, the macro-loop invasion complex/open chain invasion complex/single strand conversion is coupled to the hairpin/single strand conversion, and thus the overall melting process is a highly complex, in particular when the interaction between the invader and TAR is weak. The applicability of ^{19}F NMR spectroscopy for monitoring these dynamic intramolecular and intermolecular equilibria was demonstrated.

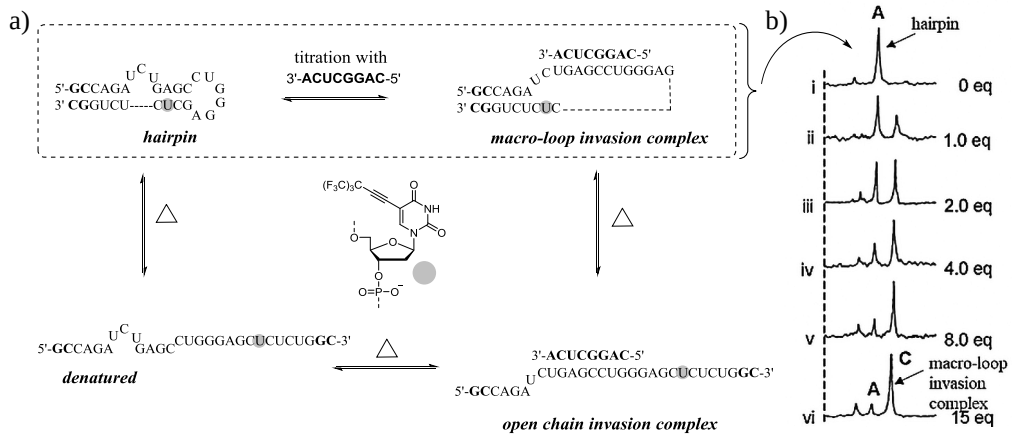


Figure 17. (a) Schematic presentation of the invasion of 8 mer 2´-*O*-methyl oligoribonucleotide into a 5-[4,4,4-trifluoro-3,3-bis(trifluoromethyl)but-1-ynyl]-2´-deoxyuridine-modified HIV-1 TAR model (boxed) and the temperature-dependent behaviour of the invasion complexes. (b) The ^{19}F NMR spectra upon the titration of 2´-*O*-methyl oligoribonucleotide. Reprinted and modified with permission from Kiviniemi, A.; Virta, P. *J. Am. Chem. Soc.* **2010**, *132*, 8560-8562. Copyright (2010) American Chemical Society.[24]

1.2.2.7 DNA triple helix

Triple-helical nucleic acid structures are formed when a third strand (i.e., triplex-forming strand) binds a sequence-specifically in the major groove of a double-helical nucleic acid structure.[182-184] (A brief description of triple helices is presented in Section 3.2.1). A pyrimidine motif DNA triplex has been studied by ^{19}F NMR spectroscopy using a 2´-deoxy-5-fluorouridine-modified third strand (Figure 18).[35] Because the formation of C-rich pyrimidine motif triplexes requires the protonation of the cytosine bases[185-189], the model is strongly pH-dependent and stabilized under acidic conditions.

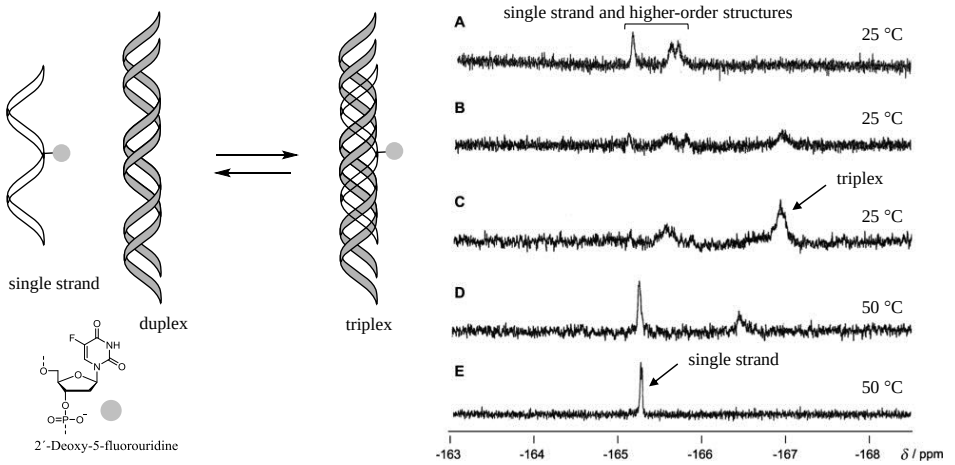


Figure 18. Schematic presentation of triplex formation and ^{19}F NMR spectra of 2-deoxy-5-fluorouridine-substitued single strand alone (A and E) and in the presence of duplex (B, C and D). Spectra reprinted and modified from Tanabe, K.; Sugiura, M.; Nishimoto, S. *Bioorg. Med. Chem.* **2010**, *18*, 6690-6694, Copyright (2010), with permission from Elsevier.[35]

The ^{19}F-NMR spectra of the third strand alone was complex and multiple signals indicated the formation of secondary structures (Figure 18).[35] These signals were attributed to non-specific structures, because C-rich strands can adopt unspecified structures under acidic conditions[190,191]. This suggestion was further confirmed by the fact that a single ^{19}F signal was observed at a neutral pH or at an elevated temperature. The association of the third strand to the DNA duplex at 25 °C and the dissociation of the third strand upon heating could be clearly detected from distinct ^{19}F chemical shifts. As a result, this labelling choice allowed the detection of triplex and single-stranded structures and to some extent, competitive unspecified structures. These results demonstrated that DNA triplexes can be monitored by ^{19}F NMR. In another study, neomycin-conjugated triplex-forming oligonucleotides were investigated by ^{19}F NMR spectroscopy.[36] This work has been conducted in our laboratory; some of the preliminary work is included in the Results and Discussion Section and thus it is not presented here.

1.2.2.8 G-quadruplexes

G-quadruplexes are four-stranded nucleic acid structures formed by G-rich DNA and RNA sequences.[192-194] These topologically diverse structures have received considerable attention due to their possible biological roles, for instance, in replication, translation, telomere biology and transcription.[193,195-201] A short description of G-quadruplexes is presented in Section 3.5.

A dimeric model of a human telomeric RNA G-quadruplex has been studied with a 3,5-bis(trifluoromethyl)phenyl label (Figure 19a).[37] A single ^{19}F signal assigned to a dimeric G-quadruplex was observed at a low RNA concentration, and an additional signal was detected at a higher RNA concentration (Figure 19b). This ^{19}F signal was attributed to a higher-order G-quadruplex in which two dimeric G-quadruplex subunits are stacked together. Consequently, the ^{19}F NMR allowed the behaviour of the dimeric and stacked G-quadruplex structures to be investigated simultaneously in various conditions.

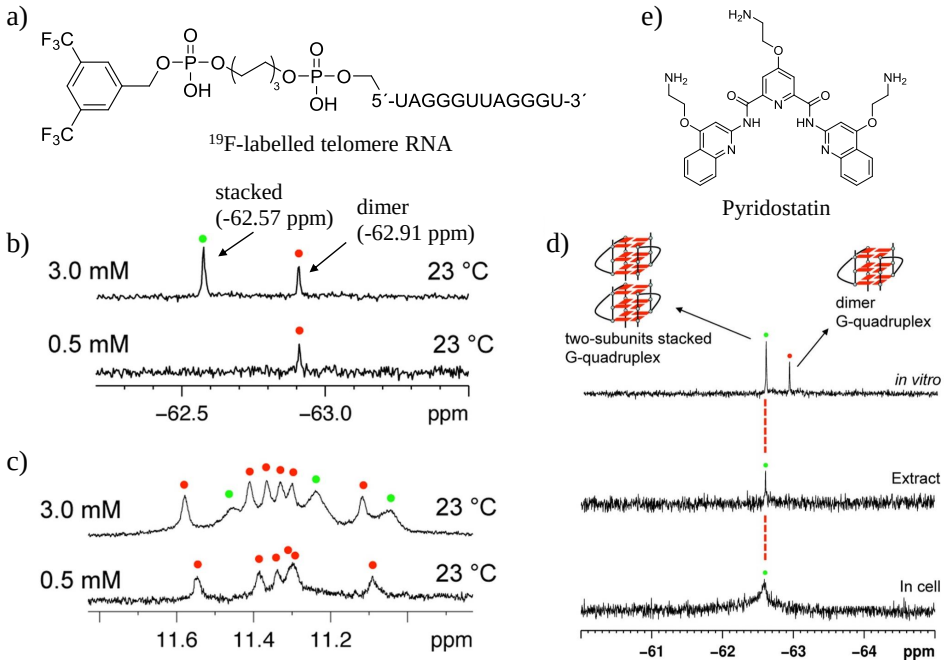


Figure 19. (a) Structure of ^{19}F-labelled telomere RNA. According to CD spectroscopy the labelling did not affect the G-quadruplex structure, but the T_m values were not reported. (b) ^{19}F NMR spectra and (c) ^{1}H NMR spectra of labelled RNA at different concentrations. (d) ^{19}F NMR spectra of labelled RNA *in vitro*, within *Xenopus laevis* oocyte extract and in *Xenopus laevis* cells. The peaks corresponding to dimeric G-quadruplex are marked with red dots and those of the stacked G-quadruplex with green dots. (e) Structure of pyridostatin. Reprinted and modified with permission from Bao, H.; Ishizuka, T.; Sakamoto, T.; Fujimoto, K.; Uechi, T.; Kenmochi, N.; Xu, Y. *Nucleic Acids Res.* **2017**, *45*, 5501-5511. Copyright (2017) The Authors.[37]

The temperature-dependent conversion of stacked G-quadruplex/dimer G-quadruplex/single strand was monitored by ^{19}F NMR.[37] T_m values of 52.4 °C and 67.8 °C were obtained for stacked and dimeric G-quadruplex, respectively, at a 3.0 mM concentration. Moreover, thermodynamic parameters were extracted from the temperature-dependent ^{19}F NMR data by applying van't Hoff analysis. The thermodynamic parameters indicated that the stabilizing stacking interactions are of an enthalpic origin. In particular, the ΔH value for the stacking interaction between two G-quadruplex subunits (ΔH -61.8 kJ/mol) was nearly comparable to the formation of the single G-tetrad (ΔH -86.9 kJ/mol).

Interestingly, the conformational behaviour of the G-quadruplex was investigated in cells using ^{19}F NMR spectroscopy (i.e., in-cell ^{19}F NMR experiment).[37] The ^{19}F-labelled RNA was injected into *Xenopus laevis* oocytes. The single observable ^{19}F signal associated with stacked G-quadruplex demonstrated that the higher-order G-quadruplex structure was formed in living cells (Figure 19d). Similarly, the ^{19}F signal of the stacked G-quadruplex was also detected in oocyte lysate. The stacked G-quadruplex-dimeric G-quadruplex equilibrium was additionally investigated under molecular crowding conditions. The ^{19}F NMR data showed that molecular crowding agents, including acetonitrile, dimethyl sulfoxide, ethanol and PEG200, induced and stabilized the formation of a stacked G-quadruplex structure. The binding of a well-known G-quadruplex binding molecule, pyridostatin (Figure 19e), to dimeric and stacked G-quadruplexes was also studied.[38] The binding stoichiometry (including K_d values) and temperature-dependent behaviour of the complexes could be monitored by ^{19}F NMR spectroscopy. The results indicated that pyridostatin preferably binds to the dimeric rather than the stacked G-quadruplex. Finally, it was demonstrated that interaction between a telomeric repeat binding protein and the dimeric G-quadruplex was detectable by ^{19}F NMR.

DNA G-quadruplexes have also been investigated by ^{19}F NMR spectroscopy. The intended purpose of the 5´-end 3,5-bis(trifluoromethyl)phenyl-modified thrombin binding aptamer (TBA) was to use it as a potassium ion sensor (Figure 20a).[18] A unique ^{19}F signal was detected that responded to K^+ addition in a concentration-dependent manner. Unfortunately, the modification decreased the thermal stability of TBA (ΔT_m -7.0 °C compared to native TBA) and thus reduced the detection range of potassium ions. In a recent study, TBA was investigated with a ^{19}F-modified guanine derivative (Figure 20b), and as expected, the derivative stabilized the G-quadruplex (ΔT_m +5.0 °C compared to native TBA).[39] Like the 5´-end modification, the guanine label was clearly able to distinguish single-stranded and G-quadruplex structures. In addition to TBA, the formation of intermolecular human telomeric DNA G-quadruplex structures was monitored with the guanine label. The derivative was also used to monitor the thermal melting of telomeric DNA G-quadruplex and for the detection of G-quadruplex structures in *Xenopus laevis* oocytes. In another recent study, the thermodynamic

properties of 5-[(3,5-bis(trifluoromethyl)phenyl)ethynyl]-2´-deoxyuridine-modified TBAs were analysed by ^{19}F NMR (Figure 20c).[19]

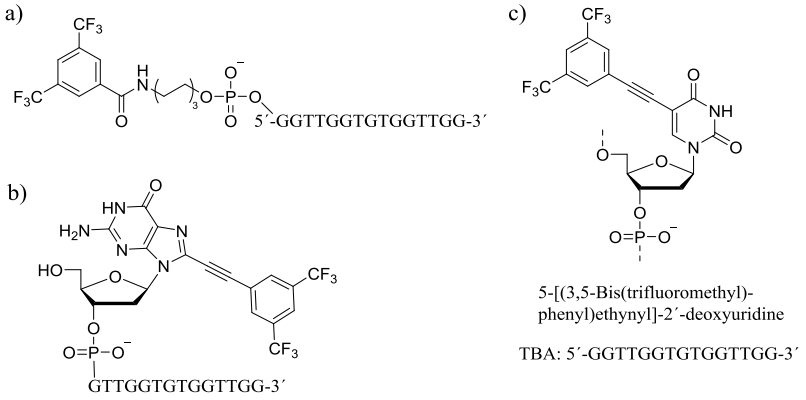


Figure 20. ^{19}F-labelled thrombin binding aptamers. (a) A non-nucleosidic fluorine modification at the 5´-end of TBA. (b) 8-[(3,5-Bis(trifluoromethyl)phenyl)ethynyl]-2´-deoxyguanosine-modified TBA. (c) 5-[(3,5-Bis(trifluoromethyl)phenyl)ethynyl]-2´-deoxyuridine, which has been incorporated in place of one of the thymidine residues within TBA.

1.2.2.9 Protein-nucleic acid interactions

Nucleic acid-protein interactions play fundamental roles in biology, and similar to nucleic acid-nucleic acid interactions, ^{19}F NMR spectroscopy can elucidate the structural and dynamic properties of these macromolecule complexes. ^{19}F NMR data have provided information about binding sites and interactions as well as conformational changes and structural distortions. These studies include interactions between *E. coli* $tRNA^{Val}$ and its cognate synthetase[159,163,164], between bacteriophage T7 RNA polymerase and its promoter[109], between Lambda phage cro repressor and its operator DNA O_R3[112,113] and between *E. coli* lac repressor and its lac operator[114].

In addition to probing binding interactions, ^{19}F NMR spectroscopy has been successfully used to investigate the dynamic conformational behaviour of enzyme-DNA complexes.[117,118,134] For example, ^{19}F NMR spectroscopy has been exploited to investigate the flipping of cytosine by DNA cytosine-5 methyltransferase *Hha*I (M.*Hha*I).[118] The dynamic conformational behaviour of the target cytosine (for methylation) in the binary (DNA-enzyme) and ternary (DNA-enzyme-cofactor) complexes was investigated by replacing it with 5-fluorocytosine (Figure 21a). ^{19}F NMR and gel shift mobility data revealed three conformational states for the 5-fluorocytosine: a stacked conformation in the DNA duplex, an ensemble of extrahelical flipped-out conformations, and an extrahelical flipped-out locked conformation in the enzyme active site. In particular, the binding of a cofactor analogue S-adenosyl-L-homocysteine

(AdoHcy) to the binary complex (i.e., the formation of a ternary complex) induced the formation of the extrahelical locked conformation (Figure 21b). This conformation was clearly detected by ^{19}F NMR spectroscopy because a new signal appeared approximately 11 ppm from the signal of the stacked 5-fluorocytosine. Because the data indicated that the base flipping was not solely dependent on the binding of the cytosine in the active site of the enzyme, it was proposed that the enzyme has an active role in the opening of the DNA duplex.

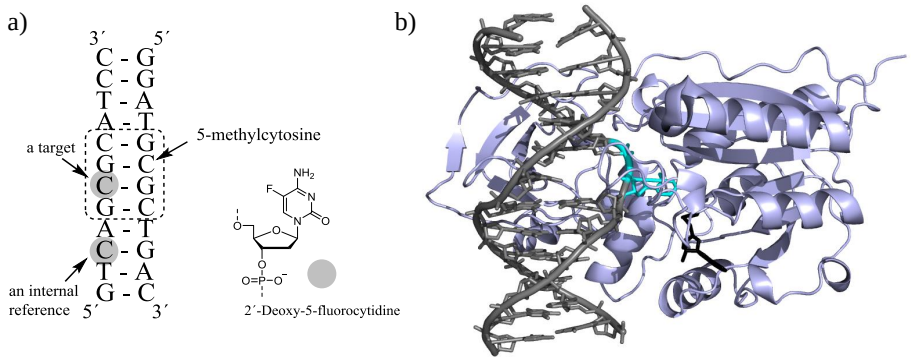


Figure 21. (a) ^{19}F-labelled DNA used to investigate base flipping. The recognition site of the methyltransferase is boxed. The target cytosine was replaced by 5-fluorocytosine, and the other 5-fluorocytosine served as an internal reference. A cytosine of the complementary strand was methylated to prevent interaction of the enzyme with that cytosine. (b) Structure of the ternary M.*Hha*I (light grey)-DNA (dark grey, flipped out cytidine residue is in cyan)-AdoHcy (black) complex with non-labelled DNA sequence (PDB ID 3MHT[202]).

In addition, ^{19}F NMR spectroscopy has been utilized to investigate the conformational dynamics of vaccinia type IB topoisomerase-DNA[117] and uracil DNA glycosylase-DNA[134] complexes. The catalytic mechanism of tRNA methylation has also been investigated.[110] All of these studies mentioned above were conducted between 1985 and 2004, and despite these early studies, this area has gained little interest during the last decade. Recent studies have demonstrated the capability of labels to detect the formation of protein-nucleic acid complexes.[19,38,101,107] For example, the interaction between DNA duplex and a core domain of p53 (a tumour suppressor protein[203]) was monitored with a 5-[4-(trifluoroacetyl)phenyl]-2′-deoxycytidine.[101] Protein-nucleic acid complexes have also been investigated using fluorine-labelled proteins.[204,205]

1.2.2.10 DNA adducts

Arylamines are mutagenic compounds that form covalent adducts with DNA *in vivo* and have been implicated in the development of various cancers.[206,207] The "adducted DNA"

can adopt multiple conformers that exist in dynamic equilibria, and the mutagenicity partly originates from this conformational heterogeneity. ^{19}F NMR spectroscopy has been used to investigate the complex conformational behaviour of FAF, FAAF and FABP adducts, which are fluorinated analogues of extensively studied C8-substitued deoxyguanosine arylamine DNA adducts of 2-aminofluorene and 4-aminobiphenyl (Figure 22a).[33,34,96,208-221] ^{19}F NMR spectroscopy has enabled the direct detection of various conformations and revealed how the distribution and stability of the conformers depend on surrounding features, such as DNA sequence composition or temperature (Figure 22b and c). ^{19}F NMR in conjunction with other detection methods has provided valuable insight into conformational heterogeneity and how a certain adduct affects the mutagenic and repair outcomes.

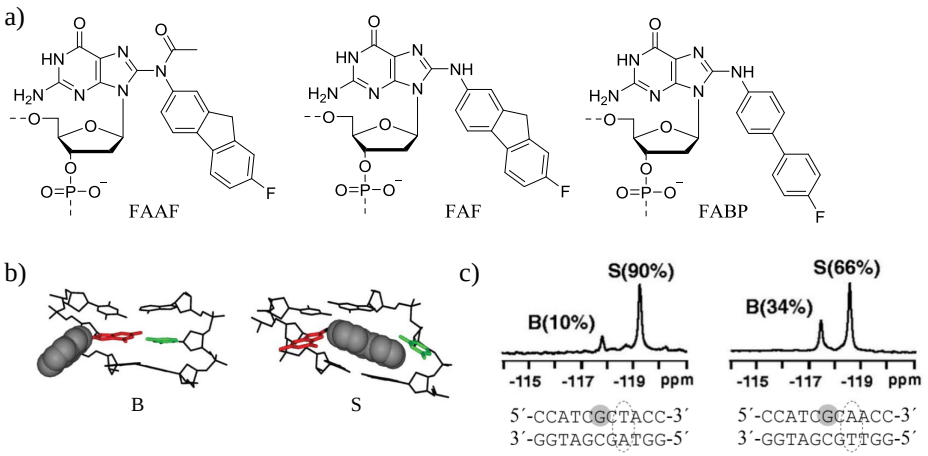


Figure 22. (a) Structures of FAF [*N*-(2´-deoxyguanosin-8-yl)-7-fluoro-2-aminofluorene], FAAF [*N*-(2´-deoxyguanosin-8-yl)-7-fluoro-2-acetylaminofluorene] and FABP [*N*-(2´-deoxyguanosin-8-yl)-4´-fluoro-4-aminobiphenyl] adducts. (b) Conformational states adopted by non-fluorinated FAF. A major groove B-type conformation (B) and a stacked conformation (S). (c) ^{19}F NMR spectra of FAF (grey G) modified DNA duplexes in different sequence contexts (the inverted base pair is circled). Reprinted and modified from Jain, V.; Hilton, B.; Lin, B.; Patnaik, S.; Liang, F.; Darian, E.; Zou, Y.; MacKerell, A. D.; Cho, B. P. *Nucleic Acids Res.* **2013**, *41*, 869-880 by permission of Oxford University Press © The Authors 2012.[217]

1.2.2.11 Peptide nucleic acid-nucleic acid interactions

Peptide nucleic acids (PNAs) are synthetic homologs of nucleic acids in which the phosphate-sugar backbone is replaced by a peptide-like backbone.[222] As illustrated in Figure 23a, PNAs can form a triplex structure via strand invasion in addition to triplex. ^{19}F NMR has been used to investigate interactions between triplex-forming peptide nucleic acids (TFPNAs) and ^{19}F-labelled miRNA hairpin models.[95] (This study has been

conducted in our laboratory and the used fluorine-modified nucleoside analogues are presented in the Results and Discussion Section). Using ^{19}F NMR, it was possible to distinguish these triplex structures and thus analyse the binding modes of differently modified TFPNAs. ^{19}F NMR provided information about binding stoichiometry, transitions between binding modes, and the temperature-dependent behaviour of coexisting complexes.

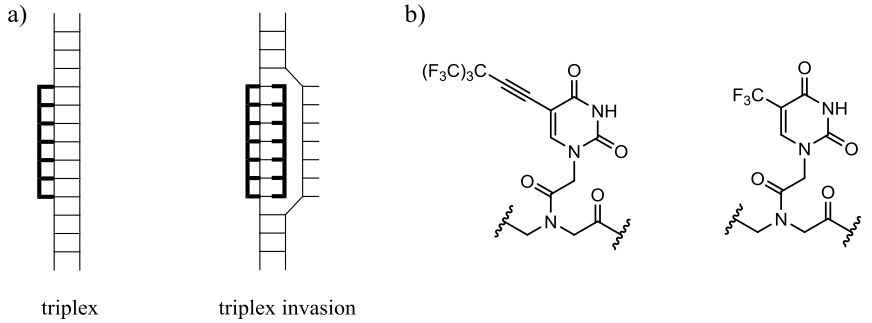


Figure 23. (a) Schematic presentation of PNA triplexes. PNA sequences are represented in bold. (b) Fluorinated PNA analogues.

In addition, PNA-nucleic acid interactions have been investigated by incorporating a fluorine label into the PNA (Figure 23b).[223] Using this strategy, ^{19}F NMR spectroscopy was successfully applied to monitor the hybridization of fluorine-labelled PNAs with complementary parallel or antiparallel RNA or DNA sequences and to monitor more complicated competitive hybridization processes.

1.2.2.12 Hybridization probes

Fluorine-labelled oligonucleotides have been used as ^{19}F NMR-based hybridization probes to detect target sequences and structural distortions. A variety of approaches have been described. In one strategy, a photo-cross-linking-induced ^{19}F chemical shift change was combined with a hybridization chain reaction.[56] Using this strategy it was possible to detect miRNAs at a nanomolar level using 3-cyanovinylcarbazole- and 2´-deoxy-5-trifluoromethyluridine-modified DNA probes (Figure 24a). In a different approach, a potential DNA hybridization probe for MRI was designed to utilize a paramagnetic relaxation enhancement effect.[53] As illustrated in Figure 24b, the probe adopts a hairpin structure in which close proximity of a 3´-end chelate to the 5´-end fluorine moiety was shown to quench the ^{19}F signal, whereas a detectable ^{19}F signal was observed upon hybridization with a target sequence. Additionally, fluorinated DNA-polymer conjugates[57] and fluorinated DNA gold nanoparticle complexes[54] have been proposed as potential MRI probes.

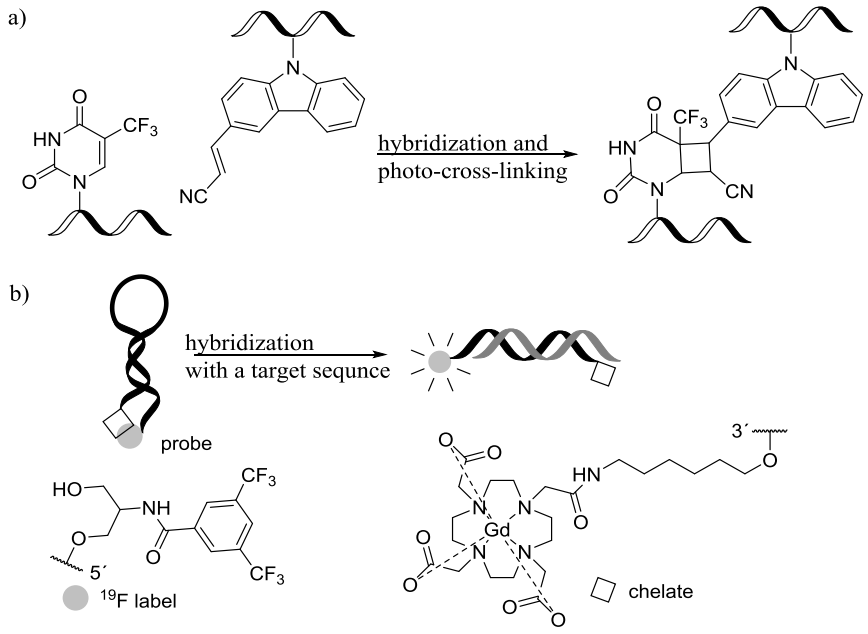


Figure 24. (a) Photo-cross-linking reaction between 3-cyanovinylcarbazole and 2´-deoxy-5-trifluoromethyluridine. (b) Schematic presentation of the mechanism of a DNA probe containing a fluorine and a chelate part.

3,5-Bis(trifluoromethyl)phenyl-modified oligodeoxynucleotides have been used as hybridization probes to detect DNA mismatches and single nucleotide bulges (Figure 25).[55] In this approach the hybridization of the probe with the target sequences is expected to provide a characteristic ^{19}F signal for each structure. As summarized in Figure 25, both of the probes were able to distinguish some of the distortions, but the discrimination ability of the fluorine moiety was better with a longer linker.

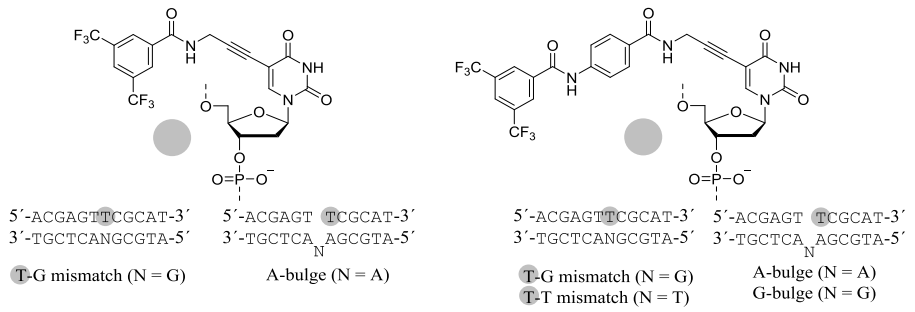


Figure 25. 3,5-Bis(trifluoromethyl)phenyl-modified hybridization probes and investigated mismatch and bulge structures. The non-complementary sequences that were recognized are presented below the labels.

In contrast to the hybridization probes discussed above, fluorinated 2´-deoxy-N^4-(6-aminopyridin-2-yl)cytidine-modified triplex-forming oligodeoxynucleotides have been used to recognize mismatches in double-stranded DNA structures.[58] One of the cytidine derivatives and its interaction with a duplex is shown in Figure 26. This probe was able to distinguish fully complementary sequence and all three mismatch structures simultaneously.

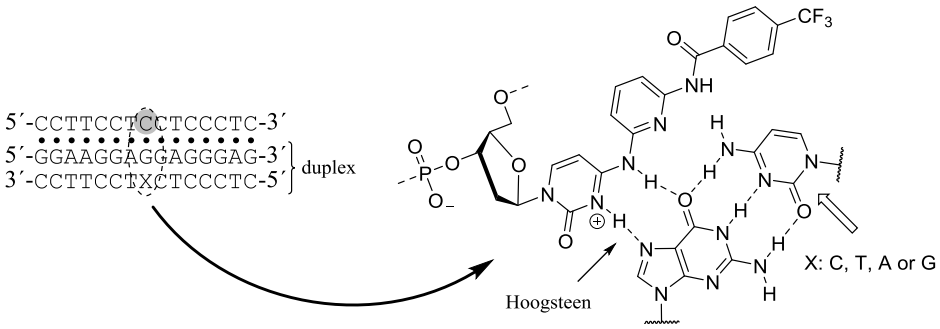


Figure 26. The proposed interaction of fluorinated 2´-deoxy-N^4-(6-aminopyridin-2-yl)cytidine derivative with the DNA duplex. (• Hoogsteen hydrogen bonding)

1.3 External fluorine probes

Fluorinated small molecules have been proposed as an alternative approach to internal fluorine probes.[224] One advantage of small molecule external probes is that the labelling of the target nucleic acid is avoided. However, the obvious challenge of this approach is to find a suitable fluorinated small molecule that readily reflects the process of interest. The concept of this approach has been demonstrated with fluorinated diaminocyclopentanes[224], but the use of external fluorine probes as structural probes has received little interest thus far.

The interaction of racemic fluorinated diaminocyclopentanes (Figure 27a) with $tRNA^{Lys}_3$ was investigated by ^{19}F NMR to evaluate their properties as external ^{19}F NMR probes.[224] Among these compounds, a (±)-4-fluorocyclopentane-1,3-diamine was the most sensitive probe. Its interaction with $tRNA^{Lys}_3$ led to a chiral resolution and distinct ^{19}F signals were observed for each enantiomer (Figure 27b). ^1H NMR and TROSY (transverse relaxation optimized spectroscopy) measurements revealed, that the enantiomers bind at the same binding site located in the T-arm with K_d values in the millimolar range.

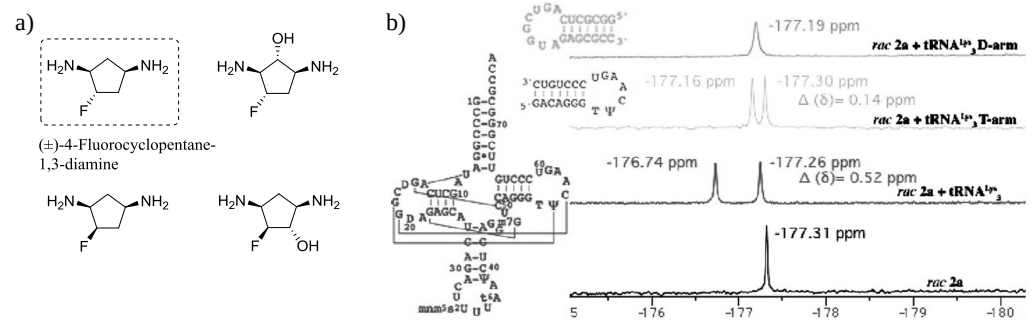


Figure 27. (a) Structures of fluorinated diaminocyclopentanes. (b) ^{19}F NMR spectra of (±)-4-fluorocyclopentane-1,3-diamine (*rac* 2a refers to (±)-4-fluorocyclopentane-1,3-diamine) with tRNA$^{Lys}_3$ D-arm, tRNA$^{Lys}_3$ T-arm and tRNA$^{Lys}_3$. Reprinted and modified with permission from Moumné, R.; Pasco, M.; Prost, E.; Lecourt, T.; Micouin, L.; Tisné, C. *J. Am. Chem. Soc.* **2010**, *132*, 13111-13113. Copyright (2010) American Chemical Society.[224]

The applicability of (±)-4-fluorocyclopentane-1,3-diamine as a structural probe was further demonstrated by monitoring its interaction with other tRNAs including *E.coli* tRNA$_f^{Met}$ and tRNA^{Met} and yeast tRNA^{Phe}.[224] A different signal splitting was observed in each case, demonstrating the sensitivity of the probe to its local environment. In addition, (±)-4-fluorocyclopentane-1,3-diamine was successfully used to monitor the thermal melting of the T-arm hairpin. The non-invasive properties of (±)-4-fluorocyclopentane-1,3-diamine were also demonstrated by analysing its interaction with neomycin aptamer.[52] Neomycin aptamer adopts a less structured conformation in the absence of a ligand and upon neomycin binding adopts a well-structured conformation. ^{1}H NMR experiments showed that the interaction of (±)-4-fluorocyclopentane-1,3-diamine with the neomycin aptamer did not affect conformational equilibrium whereas a higher affinity probe clearly shifted the equilibrium towards a bound state. Because the (±)-4-fluorocyclopentane-1,3-diamine was displaced upon neomycin binding, it was possible to monitor the conformational capture of the aptamer via its ^{19}F signals. It was proposed that this concept might be suitable for investigating riboswitches. The (±)-4-fluorocyclopentane-1,3-diamine has also been employed in ligand binding studies (Section 1.4).

Another example of an external fluorine probe is 3,5-bis(trifluoromethyl)phenyl-modified bisbenzimide H 33258 (Figure 28a).[225,226] It is a dual probe designed to discriminate between dsDNA structures. In contrast to (±)-4-fluorocyclopentane-1,3-diamine, its interaction with the target is not based on a weak non-invasive interaction. Bisbenzimide H 33258 (Hoechst 33258) is a well-known fluorescent agent that binds to the minor groove of an AATT region of dsDNA. The ^{19}F NMR data showed that the ^{19}F shift of the probe was sensitive to the adjacent base-pairs of the AATT binding site (Figure 28b) and it was demonstrated that all four neighbouring base pairs were recognized in a mixture of sequences.[225] Moreover, it was shown that the probe

could distinguish single nucleotide bulge structures near the AATT binding site (Figure 28c) and thus provide characteristic ^{19}F signals for each structure.[226]

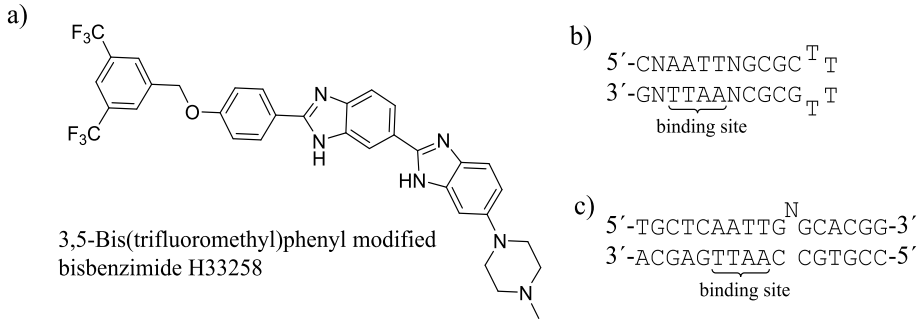


Figure 28. (a) Structure of 3,5-bis(trifluoromethyl)phenyl-modified bisbenzimide H 33258. (b and c) Investigated structures. Binding site of the probe is shown.

1.4 Small molecule binding studies

Nucleic acid structures are attractive targets for small molecules, and ^{19}F NMR-based methods have also been used to investigate small molecule-nucleic acid interactions. The studies employing internal fluorine probes have focused on investigating known small molecule-nucleic acid interactions.[17,20,25,38,162,176] Previously mentioned TAR-argininamide[20,25] and G-quadruplex-pyridostatin[38] interactions are examples. In addition, several examples have been reported in which ^{19}F NMR spectroscopy has been utilized to provide detailed information on the binding interactions of fluorinated small molecules[62,227-238], such as fluorinated daunomycin[234-237] or tripeptide[231,232] or acridine[229,233] derivatives, usually with polynucleotides or short oligonucleotides. Although ^{19}F NMR can be used to investigate small molecule-nucleic acid complexes in a detailed manner, it also has great potential to identify novel small molecules in biologically relevant nucleic acid targets. ^{19}F NMR spectroscopy has been recognized as a valuable tool in drug development in the protein field, and ^{19}F NMR based methods for K_d determination and small molecule screening are well established.[4,11,239-241] Fluorine is extensively employed to modulate the pharmacological properties of drugs, and thus, fluorine is not introduced into small molecules only for detection purposes.[44] Despite ^{19}F NMR-based methods being extensively used in protein-based drug discovery, few corresponding studies with nucleic acids have been conducted. The basic principles of small molecular ligand-RNA screening are presented below.

As discussed earlier, internal fluorine probes can be used to investigate small molecule-nucleic acid interactions. A general concept of this method has been demonstrated by investigating the binding of both high-affinity and low-affinity ligands to 2′-F-labelled

tobramycin and flavine mononucleotide (FMN) aptamers.[17] The changes to the ^{19}F signals upon the consecutive addition of ligands were characteristic of the strength of the ligand-RNA interactions. In other words, the binding of high-affinity ligands occurred in a slow exchange regime, whereas the binding of low-affinity ligands occurred in a fast exchange regime. It was additionally demonstrated that the method is suitable for larger RNAs. The 53 nt RNA construct contained the ligand binding sites for tobramycin and FMN, and binding of these ligands could be detected simultaneously (Figure 29). Apparently, this concept has not been used to identify novel small molecules binding to biologically relevant targets.

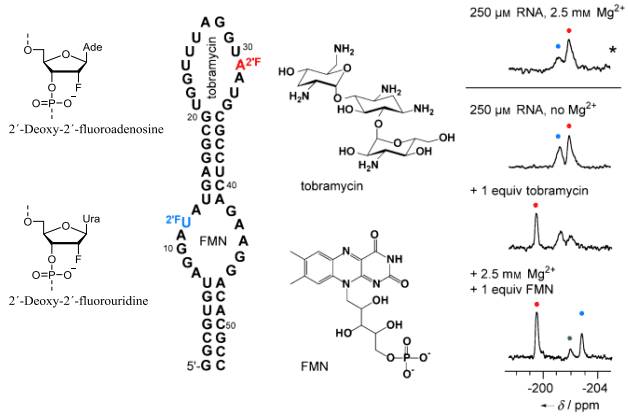


Figure 29. RNA construct (53 nt) bearing ligand binding sites for tobramycin (i.e., hairpin loop containing 2´-F adenosine (red)) and FMN (i.e., internal loop containing 2´-F uridine (blue)). ^{19}F NMR spectra upon the addition of tobramycin and FMN. Reprinted and modified with permission from Kreutz, C.; Kählig, H.; Konrat, R.; Micura, R. *Angew. Chem. Int. Ed.* **2006**, *45*, 3450-3453. © 2006 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.[17]

In an alternative approach, the small molecule contains a fluorine.[4,11,239] This approach can be further divided into a ligand-based binding (direct ligand observation) and a competition ligand-based binding (a reporter displacement method) (Figure 30). In the ligand-based binding experiment, the ligand of interest contains fluorine and its binding to a nucleic acid target is monitored directly from its fluorine signal. The competition ligand-based binding experiment requires a fluorine-labelled molecule, called a spy or reporter molecule, which has a weak to medium affinity to a nucleic acid target. A non-fluorinated ligand of interest displaces the spy molecule from the target, and the binding of the ligand can be observed via the fluorine signal of the spy molecule.

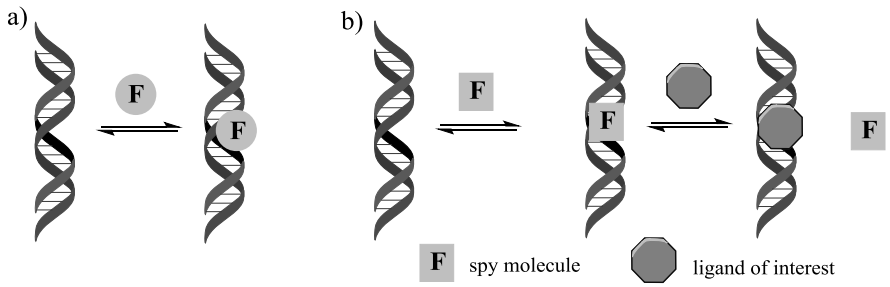


Figure 30. Schematic presentation of (a) a ligand-based binding and (b) a competition ligand-based binding.

In a recent study, ^{19}F NMR-based fragment screening was applied for the first time to a nucleic acid target.[40] The target is a telomeric repeat RNA (TERRA) that folds into G-quadruplexes and is considered to be a potential anticancer target. A library of 355 chemically and structurally diverse fluorinated fragments was screened against $TERRA_{16}$ (composed of 16 r(UUAGGG) repeats). Fragments that bind to $TERRA_{16}$ were identified by comparing ^{19}F spectra in the absence and in the presence of $TERRA_{16}$. An increase in linewidth and a decrease in the intensity of the ^{19}F signal was detected upon binding, whereas the ^{19}F signal of nonbinding compound remained unaffected (Figure 31a). Using this technique, 20 primary hits were identified, and the binding properties of seven of these were further evaluated. The same ^{19}F NMR technique was employed to investigate whether these seven hits bind to $tRNA^{Phe}$; ^{1}H NMR spectroscopy was used to investigate whether they bind to a shorter $TERRA_2$ construct (Figure 31b) or a DNA analogue of $TERRA_2$ or dsDNA. This selectivity validation showed that some of the compounds were selective for G-quadruplexes and, interestingly, favoured the parallel propeller-like conformation. Apparently, this is the only research that has utilized ^{19}F NMR-based fragment screening of a nucleic acid target.

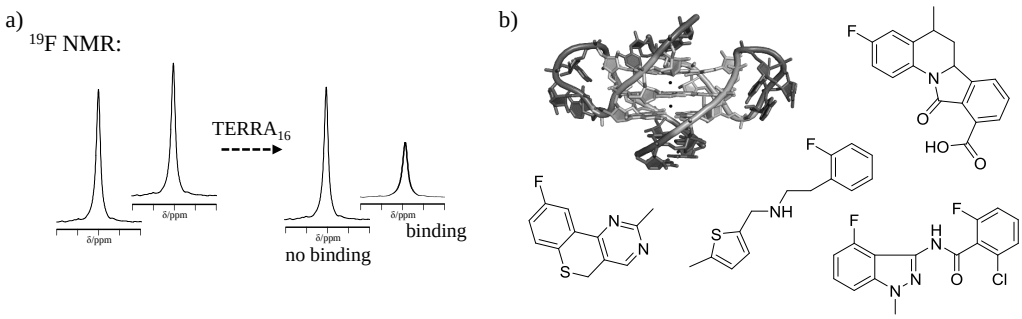


Figure 31. (a) Schematic presentation of the ^{19}F NMR based screening method. (b) Structures of some of the compounds that bind to $TERRA_2$ (PDB ID 3IBK[242]).

The spy-molecule approach has been demonstrated by investigating aminoglycoside binding into a 16S rRNA A-site model (Figure 32a).[52] (Aminoglycosides are antibiotics that function by binding in the bulge region of the 16S rRNA A site.[243]) (±)-4-Fluorocyclopentane-1,3-diamine, which has also been employed as an external structural probe (Section 1.3), and a fluorinated paromamine derivative were used as spy molecules (Figure 32b).[52] Although the binding of aminoglycosides (Figure 32c) could be monitored with (±)-4-fluorocyclopentane-1,3-diamine (K_d 2 mM), the obtained K_d values for high- and low-affinity aminoglycosides were comparable. Therefore, the fluorinated paromamine derivative was employed to rank high-affinity aminoglycosides, but due to its higher binding affinity, it was unsuitable as a spy molecule using the same measurement conditions as (±)-4-fluorocyclopentane-1,3-diamine. However, by increasing the salt concentration, the binding affinity of the fluorinated paromamine derivative decreased (K_d 300 μM vs. 1.1 mM), and its interaction with the target was shifted from a slow exchange intermediate to a fast exchange regime. Consequently, in these conditions, it was possible to use it as a spy molecule, and it reflected the binding strengths of aminoglycosides better than (±)-4-fluorocyclopentane-1,3-diamine. This research clearly demonstrated that the spy molecule approach can be used to identify both high- and low-affinity molecules in a nucleic acid target, but apparently this method has not yet been applied to the discovery of novel RNA or DNA binders.

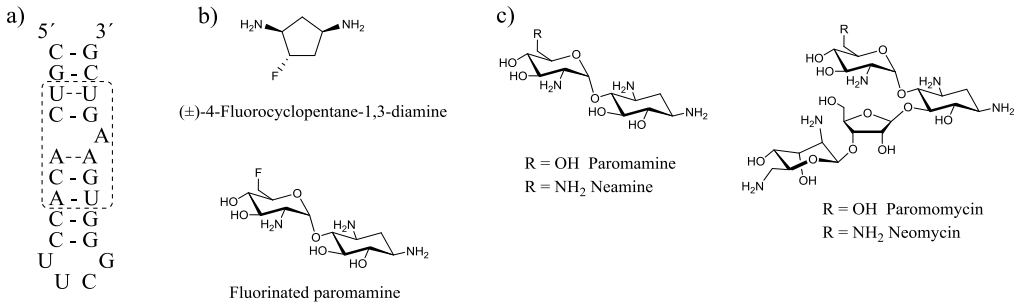


Figure 32. (a) Structure of 16S rRNA A-site model. Binding site of aminoglycosides is circled. (b) Structures of spy molecules and (c) investigated aminoglycosides.

2. AIMS OF THE THESIS

Nucleic acids are complex dynamic molecules that may adopt a variety of secondary structures. The structural complexity, conformational flexibility, and dynamic and adaptive nature of nucleic acid ligand interactions make them challenging biopolymers to investigate. ^{19}F NMR spectroscopy has shown its potential to investigate diverse nucleic acid structures. Like other spin label-based techniques, ^{19}F NMR spectroscopy may be used to provide quantitative information and detailed real-time information about structural equilibria and ligand-nucleic acid interactions. As the interest of ^{19}F NMR spectroscopy for the detection of nucleic acids has increased, there is increasing demand for novel ^{19}F-labelled nucleoside analogues with improved ^{19}F spectroscopic properties. Other properties would also be desirable for these analogues such as facile incorporation into oligonucleotides, stability of the analogue in a variety of conditions and a non-invasive nature. The main aim of this thesis was to develop novel ^{19}F-labelled building blocks that can be used to characterize DNA and RNA secondary structures. In the present study, fluorine-containing reporter groups were introduced into the sugar part. Compared to base modifications, the sugar part provides a uniform modification site. Thus, the modification can readily be expanded to all four conventional nucleosides, allowing labelling at any site of the target oligonucleotide. On the downside, sugar modifications can affect the sugar pucker, causing deviations to the overall conformation and decreasing the stability of the nucleic acid structures. Modification at the 2´-site may also disturb interactions that are related to the 2´-hydroxyl group of the RNA. As the purpose was to apply conventional solid-phase oligonucleotide synthesis for the incorporation of fluorine labels site-specifically into oligonucleotides, phosphoramidite building blocks of fluorinated nucleosides were prepared. The influence of the analogues on the stability and conformation of the native oligonucleotide structures were studied by distinct methods (UV- and CD-melting profiles and ^{1}H NMR of imino protons), and then the applicability of the sensors for the ^{19}F NMR-based detection of the secondary structural transitions of the oligonucleotides was evaluated. DNA- and RNA-triplex/duplex/single strand conversions and RNA G-quadruplex/hairpin-transitions were studied.

The aims of this thesis may be summarized as follows:

- To synthesize novel non-invasive fluorinated building blocks that can be efficiently incorporated into oligonucleotides by an automated solid-phase synthesis (I-III)
- To evaluate the applicability of these building blocks to characterize DNA and RNA secondary structures (including DNA (I) and RNA triplexes (II), RNA invasion (I) and RNA hairpin-G-quadruplex equilibria (III)) by ^{19}F NMR spectroscopy
- To provide new information about higher-order secondary structures of oligonucleotides, more specifically, triple-helical DNAs and RNAs and bistable G-quadruplex/hairpin RNAs

3. RESULTS AND DISCUSSION

3.1 Fluorine-labelled building blocks

The ^{19}F-labelled nucleoside and non-nucleoside derivatives that have been introduced into nucleic acids to characterize RNA and DNA secondary structures were presented in Section 1.2.1. Although they have been successfully applied to investigate nucleic acid structures, several shortcomings have been described, such as synthetic challenges with building blocks, compatibility with oligonucleotide synthesis methods, the general stability of the label, changes in the stability and/or structure of the nucleic acid, poor shift discrimination, possible sensitivity limitations with a single fluorine atom and the need for proton decoupling techniques. The ^{19}F-modified thymidine analogues are all C5 position modified pyrimidines (Figure 2, Section 1.2.1.1)[19,55,100,101,106,111,128,130], and several of them bear sterically demanding reporter groups that will likely affect major groove mediated interactions. The two sugar modifications, 2´-F[17,82,142] and 2´-SCF_3[94,107,143] presented in Section 1.2.1.2 have been used to investigate RNA structures and are not generally compatible with B-form DNA. In fact, sugar-modified fluorine probes have not yet been generally applied to characterize DNA secondary structures. Although 2´-F is sterically non-perturbing[125] and prefers the C3´-endo conformation[147-150], it bears only one fluorine atom coupled to protons[142]. 2´-SCF_3-modified ribonucleosides are more sensitive, but the strong preference for C2´-endo pucker has restricted the utility of these analogues in single-stranded RNA regions.[94,107,143] Fluorine-modified nucleobases for RNA bear only one fluorine atom[21,111], although 5-[4,4,4-trifluoro-3,3-bis(trifluoromethyl)but-1-ynyl]-2´-deoxyuridine has been employed as an RNA environment as well[24,89]. The incompatibility of 2-fluoroadenosine for solid-phase oligonucleotide synthesis[139,140] also narrows the repertoire of available purine analogues for site-specific labelling. The utility of 5´-end fluorine modifications[18,37,53,102] (Section 1.2.1.3) is generally narrow for diverse structural studies. Overall, the choices of ^{19}F-modified derivatives in place of ribonucleosides are still rather limited, and studies have mainly relied on 5-fluoropyrimidines. For these reasons, there is demand for novel and improved labelling strategies. The fluorine-labelled building blocks presented in this thesis were developed from these aspects and are intended to improve or overcome weaknesses described above and thus provide viable alternatives to the currently available building blocks.

The fluorine-labelled building blocks (**1**–**6**) developed in this thesis are shown in Figure 33. 4´-*C*-[(4-Trifluoromethyl-1*H*-1,2,3-triazol-1-yl)methyl]thymidine (**1**)^{I} is designed for the detection of DNA secondary structures, whereas 4´-*C*-[4-(trifluoromethyl)phenyl]uridine (**2**)^{II} and 2´-*O*-(4-trifluoromethyl-1*H*-1,2,3-

triazol-1-yl)methyl ribonucleosides (**3**–**6**)[II,III] are designed for the detection of RNA secondary structures. Each of the building blocks was incorporated into oligonucleotides as phosphoramidite. To obtain the fluorine signals directly as singlets without the need for fluorine-proton decoupling and to improve the sensitivity, the reporter groups, i.e., (4-trifluoromethyl-1*H*-1,2,3-triazol-1-yl)methyl and 4-(trifluoromethyl)phenyl, contain a trifluoromethyl (CF_3) group in a quasi-isolated spin system. The triazoyl moiety was synthetically easily accessible. It is a weak base[244] and is uncharged in the pH range used in ^{19}F NMR measurements (i.e., pH 5.5–7.0).

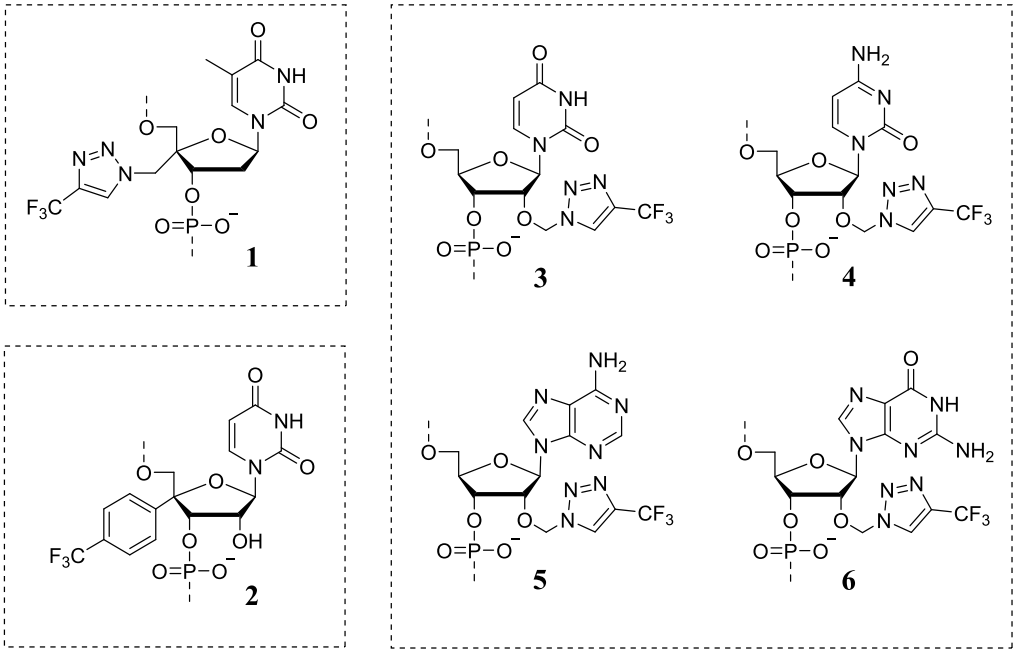


Figure 33. Synthesized fluorine-labelled building blocks.

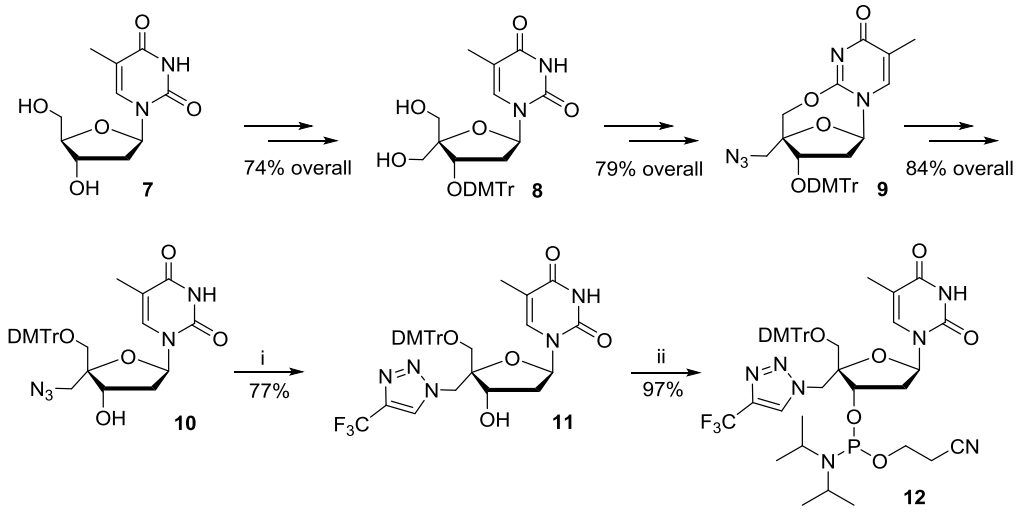
It has been shown that the 4´-*C*-modifications face the minor groove in double-helical DNA and usually induce only minor changes to the stability of the double helix.[245-254] Therefore, it was expected that introducing **1** into the DNA may be a non-disturbing substitution but one that is sensitive to secondary structural rearrangements. In addition, the nucleoside of **1** predominantly has an *S*-type conformation typically adopted in B-type duplexes (Section 3.1.3). The applicability of **1** as a non-invasive derivative for ^{19}F NMR studies was demonstrated with DNA triplexes (Section 3.2.2).[I] On an RNA duplex 4´-*C*-modifications orient outward from the helix. Therefore, a modest sensitivity to structural rearrangements was expected for 4´-*C*-reporter groups in an RNA environment upon hybridization. Despite this possible shortcoming, a 4´-*C*-modified building block **2** was synthesized.[II] As an advantage over the other labelled building blocks, it bears a free 2´-OH group. However, its ^{19}F NMR properties were found to be disadvantageous.

Its properties compared to **1** and **3** are presented in Section 3.3.2. Whilst **1** is designed for a DNA environment (due to the orientation of the reporter group, the lack of a 2´-OH group and the preference for an *S*-type conformation, as discussed above), its applicability in an RNA environment was demonstrated as well (Section 3.3.1).[I] Similar to other 2´-modifications, the possible shortcomings of 2´-*O*-(4-CF_3-1*H*-1,2,3-triazol-1-yl)methyl modifications may arise from sugar puckering and the disturbance of 2´-OH related interactions. The sugar conformation analysis of the nucleoside of **3** showed the preferred *N* conformation that is desired in RNA (Section 3.1.3). It was also shown to be a non-invasive substitution in RNA triplexes (Section 3.4).[II] Due to these promising results, 2´-*O*-(4-CF_3-1*H*-1,2,3-triazol-1-yl)methyl modification was incorporated into cytidine and purine nucleosides as well. These ^{19}F-labelled building blocks (**4**–**6**) were used to characterize bistable hairpin-G-quadruplex structures (Section 3.5).[III]

3.1.1 Synthesis of fluorine-containing phosphoramidite building blocks

3.1.1.1 4´-C-[(4-Trifluoromethyl-1H-1,2,3-triazol-1-yl)methyl]thymidine phosphoramidite

The synthesis of 4´-*C*-[(4-CF_3-1*H*-1,2,3-triazol-1-yl)methyl]thymidine phosphoramidite (**12**) is outlined in Scheme 1.[I] The multistep synthesis of 4´-*C*-azidomethyl-5´-*O*-(4,4´-dimethoxytrityl)thymidine (**10**) starting from thymidine (**7**) has previously been described.[253,255,256] Accordingly, 3´-*O*-(4,4´-dimethoxytrityl) (DMTr) -protected 4´-*C*-hydroxymethylthymidine (**8**) was first synthesized from thymidine (**7**), and the primary hydroxyl groups of **8** were then converted to trifluoromethanesulfonyl (Tf) groups. The obtained triflate nucleoside was cyclized into an $O^{2,5´}$-anhydronucleoside[257], and the remaining Tf-group was replaced with an azide ion to obtain **9**. The hydrolysis of the anhydronucleoside linkage, exposure of the 3´-OH group, and DMTr-protection of the 5´-OH group provided **10** in 49 % overall yield from **7**. The copper(I)-catalyzed azide-alkyne cycloaddition[258,259] between the 4´-*C*-azidomethyl group and gaseous 3,3,3-trifluoropropyne in the presence of sodium ascorbate afforded **11** bearing a (4-CF_3-1*H*-1,2,3-triazoyl)methyl group. Compound **11** was then quantitatively phosphitylated to the phosphoramidite **12**.

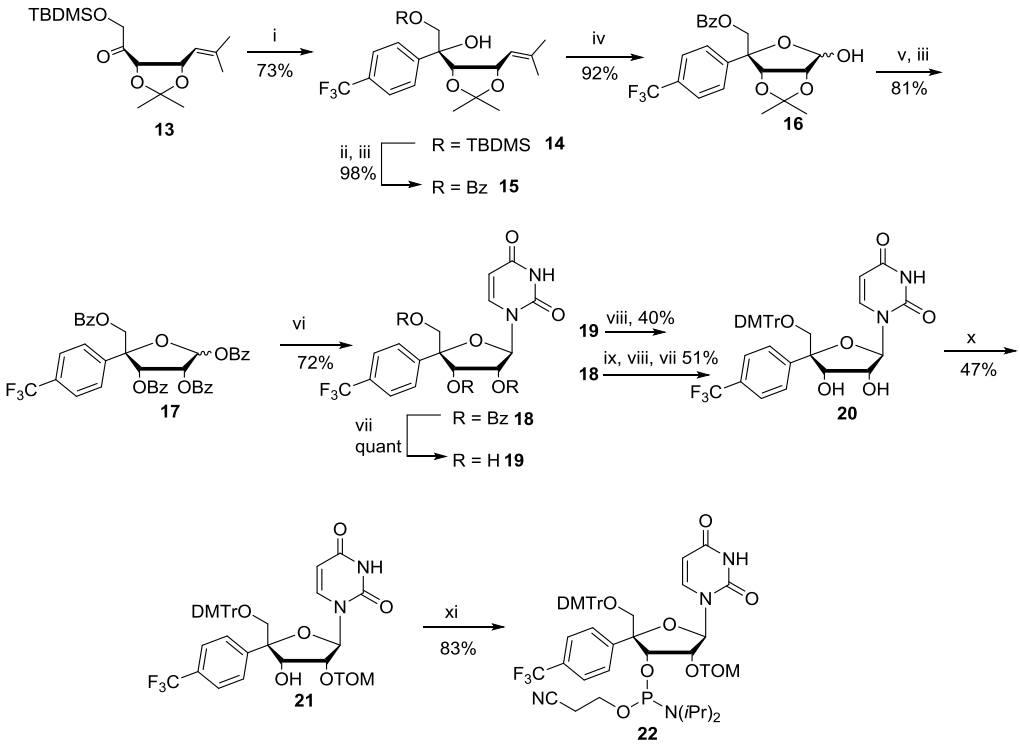


Scheme 1. Synthesis of phosphoramidite building block **12**. Conditions: i) 3,3,3-trifluoroprop-1-yne, $CuSO_4$, sodium ascorbate, H_2O, DMSO ii) 2-cyanoethyl *N*,*N*-diisopropylchlorophosphoramidite, Et_3N, DCM.

3.1.1.2 4´-C-[4-(Trifluoromethyl)phenyl]uridine phosphoramidite

The synthesis of the 4´-*C*-[4-(trifluoromethyl)phenyl]uridine phosphoramidite building block (**22**) is depicted in Scheme 2.[II] The starting compound **13** could be synthesized from ribose on the multigram scale following a previously published method.[260] A Grignard reaction between 4-(trifluoromethyl)phenylmagnesiumbromide and ketone **13** stereoselectively gave **14** (73% yield, 2*R*:2*S*, 7:1, *n/n*). The *tert*-butyldimethylsilyl group (TBDMS) was then replaced by a benzoyl group (Bz) (**15**) to avoid premature pyranose formation during the subsequent steps. The oxidation of the alkene with OsO_4 followed by periodic acid treatment provided **16**. The isopropylidene group was then removed, and the deprotected hydroxyl groups were protected with Bz-groups (**17**). The *N*-glycosidation of **17** gave **18** in 72% yield. The *β*-anomer was the major product due to the neighbouring group participation via the 2´-*O*-Bz group. The 2D NOESY correlation between the $H^{1´}$ and aromatic protons of 4-(trifluoromethyl)phenyl (CF_3Ph) verified the 4´*R*-configuration. The major problem of the synthesis was the 4,4´-dimethoxytritylation of the 5´-OH group of **19** (**20** in 40% yield). The reactivity of the 5´-OH group was comparable to that of the 2´-OH group due to the steric hindrance of the 4´-CF_3Ph group, and therefore a substantial amount of the 2´,5´-bistritylated derivative was formed together with the desired product (**20**). Several synthetic approaches were attempted to overcome this problem (Experimental Section). For example, the 5´-*O*-Bz group of **18** was removed selectively by [*t*BuSnOHCl]$_2$ in methanol[261], the exposed 5´-OH group was tritylated and the 2´-*O*- and 3´-*O*-Bz

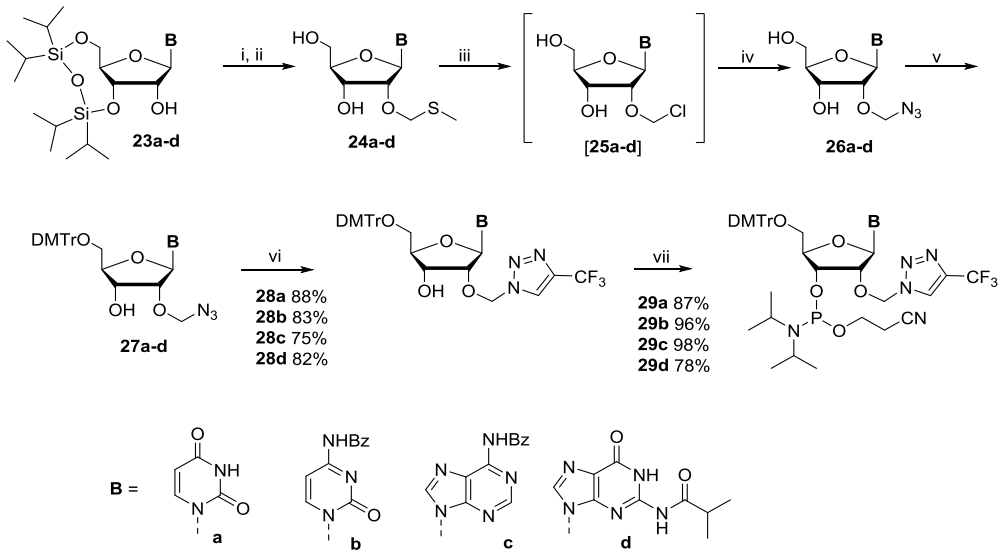
protections were removed by NaOMe treatment. By this strategy, the yield was slightly improved (51%). However, the direct tritylation of **19** proved to be a less complex method. Moreover, **19** could be readily recovered from the bistritylated sideproduct. Finally, the 2´-OH was protected with a triisopropylsilyloxymethyl group (TOM) (**21**) and the 3´-OH group was phosphitylated to give phosphoramidite **22**. The overall yield from **13** to **22** was 6 %. Instead of the TOM group, the TBDMS group was initially used for protecting the 2´-OH group. However, both the synthesis of the corresponding 2´-*O*-TBDMS protected phosphoramidite and its introduction into oligoribonucleotides were difficult (Experimental Section).



Scheme 2. (i) 4-Bromobenzotrifluoride, Mg, Et_2O; (ii) TBAF, THF; (iii) BzCl, DMAP, Py; (iv) 1: OsO_4, 4-methylmorpholine *N*-oxide, acetone; 2: H_5IO_6, THF; (v) HCl, H_2O, dioxane; (vi) TMSOTf, 2,4-bis(trimethylsilyloxy)uridine, MeCN; (vii) NaOMe, MeOH; (viii) DMTrCl, Py; (ix) [*t*BuSnOHCl]$_2$, MeOH; (x) DIEA, Bu_2SnCl_2, (triisopropylsiloxy)methyl chloride, 1.2-dichloroethane; (xi) 2-cyanoethyl *N*,*N*-diisopropylchlorophosphoramidite, Et_3N, DCM.

3.1.1.3 2´-O-(4-Trifluoromethyl-1H-1,2,3-triazol-1-yl)methyl modified phosphoramidites

The syntheses of 2´-*O*-(4-CF_3-1*H*-1,2,3-triazol-1-yl)methyl modified phosphoramidite building blocks are outlined in Scheme 3.[II,III] A straightforward synthesis of 2-´*O*-(azidomethyl)-5´-*O*-(4,4´-dimethoxytrityl)ribonucleosides (**27a-d**) from commercially available 3´,5´-*O*-(1,1,3,3-tetraisopropyldisiloxane-1,3-diyl)-protected ribonucleosides (**23a-d**) bearing appropriate base protections has previously been reported.[262-264] Accordingly, **23a-d** were treated with a mixture of dimethyl sulfoxide (DMSO), acetic anhydride (Ac_2O) and acetic acid (AcOH).[262] The 3´,5´-cyclic silyl protecting groups were then removed with tetrabutylammonium fluoride (TBAF) to obtain **24a-d**. The 2´-*O*-methylthiomethyl groups were converted to 2´-*O*-chloromethyl intermediates **25a-d** by 2-nitrobenzenesulfenyl chloride (NBSCl), and subsequent treatment with NaN_3 gave 2´-*O*-azidomethyl derivatives **26a-d**.[263,264] In the case of purine derivatives, NBSCl was added in the presence of trifluoromethanesulfonic acid to prevent a side reaction between the 2´-*O*-chloromethyl intermediate and the N3 of a purine nucleobase to give a cyclonucleoside.[263]The 5´-OH groups were protected with DMTr group (**27a-d**). The 2´-*O*-azidomethyl groups were then exposed to copper(I)-catalyzed azide-alkyne cycloaddition with gaseous 3,3,3-trifluoropropyne to give **28a-d**. Finally, compounds **28a-d** were converted to the corresponding phosphoramidites **29a-d** using 2-cyanoethyl *N*,*N*-diisopropylchlorophosphoramidite. The overall yields were 64-80 % (calculated from azides **27a-d**).



Scheme 3. (i) DMSO, Ac_2O, AcOH; (ii) TBAF, THF; (iii) NBSCl, DMF (CF_3SO_3H); (iv) NaN_3, DMF; (v) DMTrCl, Py; (vi) 3,3,3-trifluoroprop-1-yne, $CuSO_4$, sodium ascorbate, H_2O, DMSO; (vii) 2-cyanoethyl *N*,*N*-diisopropylchlorophosphoramidite, Et_3N, DCM.

3.1.2 Synthesis of fluorine-labelled oligonucleotides

Fluorine-labelled oligonucleotides were synthesized on the 1.0 µM scale using an automatic DNA/RNA synthesizer applying standard phosphoramidite coupling chemistry. Benzylthiotetrazole (BnSTet) was used as an activator. Coupling times of 20 s and 300 s were used for the standard commercially available DNA and RNA phosphoramidite building blocks, respectively. ^{19}F-labelled phosphoramidites **29a-c** were incorporated into oligonucleotides as an integral part of the automatic synthesis by increasing the coupling time to 600 s (no other adjustments were made to the standard coupling cycles). The coupling yields of **29a-c** were comparable to those of commercially available phosphoramidite building blocks (> 98 %) according to a trityl assay. **29d** was manually coupled due to its low solubility in acetonitrile. Accordingly, a solution of **29d** (0.20 mol L^{-1} in dry acetonitrile, 200 µL, 40 µmol) was mixed with a BnSTet solution (0.25 mol L^{-1}, in dry acetonitrile, 160 µL, 40 µmol). The opaque solution of **29d** cleared up after the addition of BnSTet. This mixture was suspended with a CPG (controlled pore glass) support bearing 1 µmol of the sequence before **29d** and mixed for 10 min under nitrogen. The mixture was then loaded back onto the synthesis column and filtered. Capping (with a mixture of Ac_2O, lutidine and *N*-methylimidazole in tetrahydrofurane (THF), 5:5:8:82, *v/v/v/v*) and oxidation (with a mixture of 0.02 mol L^{-1} I_2 in pyridine (Py) and H_2O in THF 1:21:213, *v/v/v*) steps were performed manually followed by washings with acetonitrile. The column was then coupled to the synthesizer, and the automatic chain elongation was continued. According to the trityl assay, **29d** coupled nearly quantitatively (> 98 %). Also, for **12** and **22**, better coupling efficiencies (**12**: 98% and **22**: 90%) were achieved by applying manual coupling, but smaller excesses of **12** and **22** could be used (10 equiv). It is worth noting that the 2´-*O*-TBDMS-protected 4´-*C*-[4-(trifluoromethyl)phenyl]uridine phosphoramidite (**36**, Experimental Section) coupled inefficiently (coupling yield ~20%) due to the steric hindrance of the TBDMS group. The oligonucleotides were released from the support and deprotected by standard ammonolysis (2´-deoxyoligonucleotides in concentrated ammonia at 55 ºC overnight, oligoribonucleotides in a mixture of concentrated ammonia and ethanol 3:1, *v/v*, 3.5 h at 55 ºC and overnight at rt). The 2´-*O*-TBDMS and 2´-*O*-TOM -protections were removed by a standard protocol with triethylamine trihydrofluoride treatment followed by cartridge filtration to remove traces of reagents and protecting groups. The crude oligonucleotides were purified with reverse-phase high-performance liquid chromatography (RP HPLC) using a semipreparative column at either ambient temperature or elevated temperature (60 °C) (Figure 34). The homogeneity of the oligonucleotides was confirmed by RP HPLC, and the authenticity was verified by electrospray ionization time-of-flight mass spectrometry (ESI-TOF MS).

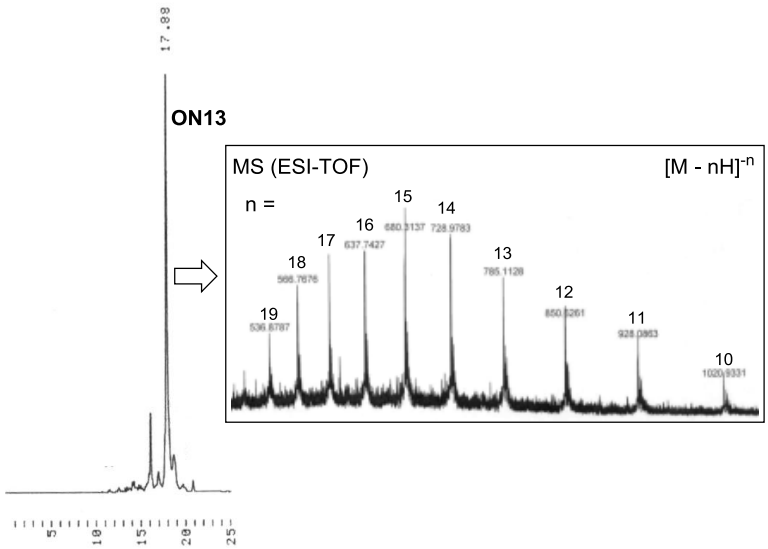


Figure 34. A representative example of the RP HPLC trace of a crude oligoribonucleotide containing a 2´-*O*-[(4-CF_3-1*H*-1,2,3-triazol-1-yl)methyl]uridine (**3**) (**ON13**, Figure 43). Conditions: a semipreparative RP HPLC column (C-18, 250 × 10 mm, 5 μm), a gradient elution of 0–50% acetonitrile in 0.1 mol L^{-1} triethylammonium acetate in 25 min at 25 °C, flow rate 3.0 ml min^{-1}, detection at 260 nm. MS (ESI-TOF) spectra of the purified oligoribonucleotide (**ON13**).

3.1.3 Sugar conformation of the fluorinated nucleosides

Substituents can affect the sugar ring conformation and hence significantly decrease the stability of the nucleic acids. The sugars in B-type dsDNA prefer the C2´-endo conformation (*S*-type), whereas those in A-type dsRNA prefer the C3´-endo conformation (*N*-type).[146] The sugar pucker preference can, however, vary from those observed in canonical dsDNA and dsRNA, particularly in single-stranded and non-canonical regions.[107,143,146,193] The sugar puckers of the corresponding nucleosides of **1**, **2** and **3** (i.e., **30**^{I}, **31**^{II} and **32**^{II}, Figure 35a) were determined using the optimized Karplus relation for ^{1}H NMR H1´-H2´ coupling constants ($J_{H1´-H2´}$)[265]. 4´-*C*-[(4-CF_3-1*H*-1,2,3-triazol-1-yl)methyl]thymidine (**30**, Figure 35b) favoured the DNA-type C2´-endo conformation (*S* 72%, $J_{H1´-H2´}$ = 7.3 Hz), and 2´-*O*-[(4-CF_3-1*H*-1,2,3-triazol-1-yl)methyl]uridine (**31,** Figure 35c) existed predominantly as an RNA-type C3´-endo conformation (*N* 67%, $J_{H1´-H2´}$ = 3.3 Hz). The population of the C3´-endo conformation for 4´-*C*-[4-(trifluoromethyl)phenyl]uridine (**32**) was *N* 48% ($J_{H1´-H2´}$ = 5.3 Hz) and the corresponding value for uridine was *N* 53% ($J_{H1´-H2´}$ = 4.8 Hz). The sugar conformation of **3** was additionally determined in a short single-stranded RNA sequence, 5´-AU**3**A-3´. The obtained C3´-endo population of **3** in this sequence was *N* 32%

($J_{H1'–H2'}$ = 6.9 Hz), and the corresponding values (*N* %) for the other residues were 48% ($J_{H1'–H2'}$ = 5.3 Hz, A), 50% ($J_{H1'–H2'}$ = 5.1 Hz, A) and 46% ($J_{H1'–H2'}$ = 5.5 Hz, U). In the non-labelled sequence, 5′-AUUA-3′, the corresponding values (*N* %) were 50% ($J_{H1'–H2'}$ = 5.1 Hz, A), 46% ($J_{H1'–H2'}$ = 5.5 Hz, A), 44% ($J_{H1'–H2'}$ = 5.7 Hz, U) and 46% ($J_{H1'–H2'}$ = 5.5 Hz, U). It may be worth noting that 2′-*O*-[(4-CF_3-1*H*-1,2,3-triazol-1-yl)methyl]uridine (**3**) hence existed as relatively low C3′-endo populations (*N* 32%) in this short RNA single strand, but so did the unmodified ribonucleotides (*N* 44–50%).

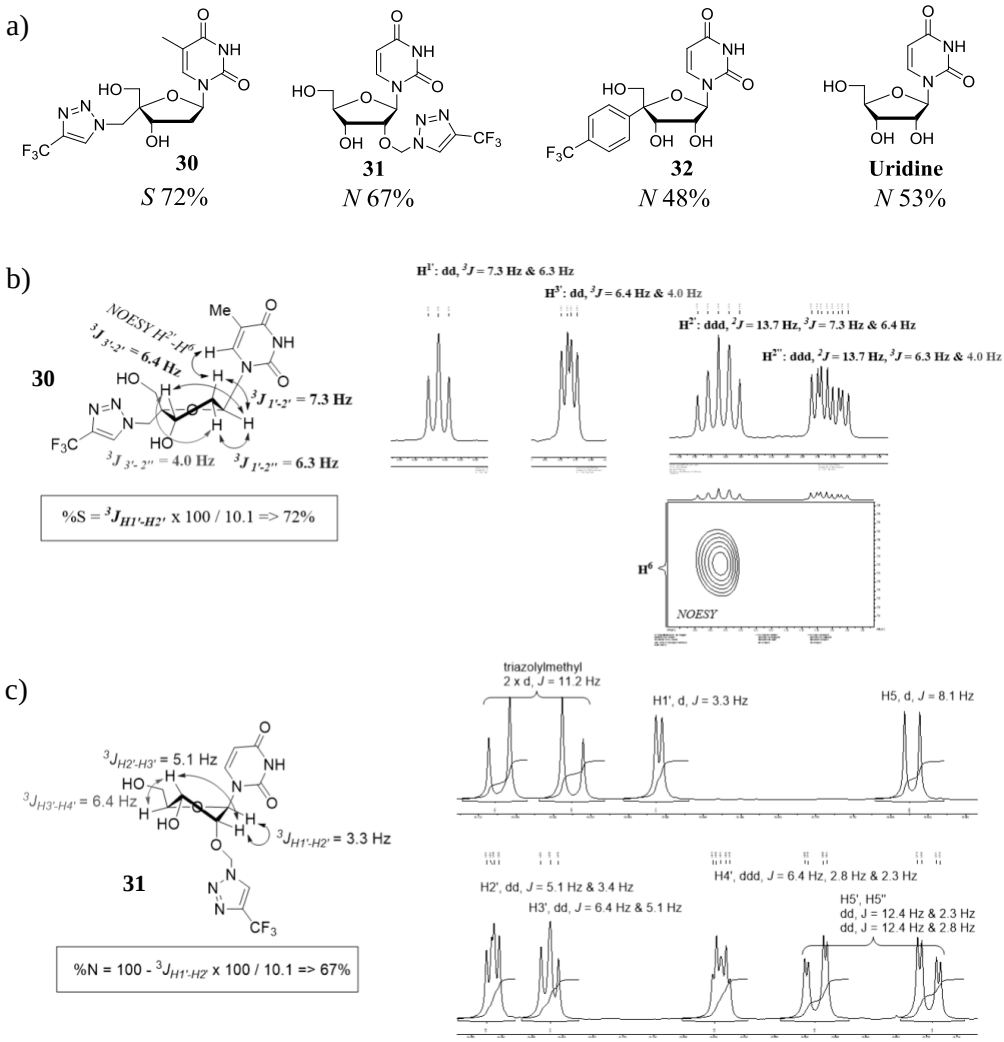


Figure 35. Sugar puckering. (a) Structures and sugar pucker of fluorine-labelled nucleosides and uridine. (b) Sugar puckering of 4′-*C*-[(4-CF_3-1*H*-1,2,3-triazol-1-yl)methyl]thymidine (**30**) and ^{1}H NMR and NOESY spectra used for analysis. (c) Sugar puckering of 2′-*O*-[(4-CF_3-1*H*-1,2,3-triazol-1-yl)methyl]uridine (**31**) and ^{1}H NMR used for analysis. ^{1}H NMR (500 MHz) measurements were carried out in MeOD.

3.2 DNA triplexes

3.2.1 Introduction to triple helices

Both artificial triplex-forming oligonucleotides (TFOs) and naturally occurring triple helices have been investigated intensively for decades. TFOs have gained enormous attention as potential therapeutic agents, because they can bind sequence-specifically to genomic DNA and affect gene expression (i.e., antigene strategy).[266,267] In general, TFOs can be used as tools to site-specifically manipulate DNA, and a variety of biological applications have been described.[266-270] For instance, TFOs have been used to inhibit transcription and replication and to induce mutations and recombination. Naturally occurring DNA triplexes (H-DNAs) may be involved in gene expression regulation.[269,271-273] H-DNAs induce genetic instability, and have been associated with several diseases[269,274,275] such as Friedreich's ataxia[276], autosomal dominant polycystic kidney disease and lymphomas[269]. In contrast to DNA triplexes, RNA triplexes have attracted little attention, but their biological importance has recently been recognized (Section 3.4).[273,277,278] As TFO-based targeting has heavily focused on DNA, it may also be worth noting that the targeting of the double-helical regions of RNA by TFPNAs has shown promising results.[95,279,280] In addition, accumulating evidence suggests that some noncoding RNA (ncRNA, non-protein coding RNA) regulatory mechanisms are mediated through triplex formation with genomic DNA.[273,278,281-285] Overall, triple-helical nucleic acids are a diverse, constantly emerging multidisciplinary research area.

Various methods are used to investigate triplexes: for example, UV[75] and CD[76,77] spectroscopy, fluorescence based techniques[286,287], isothermal titration calorimetry (ITC)[187], gel electrophoresis[184,288] and NMR[183,289,290]. Among the used methods, the UV melting study[75] is probably the most used technique. However, this technique has several limitations that can complicate the interpretation of the melting profile such as overlapping or wide melting transitions and small absorbance changes. In addition, intermediate structures or coexisting species are difficult or impossible to detect and analyse. As UV spectroscopy does not provide structural information, it is often used in combination with CD spectroscopy. Although CD spectroscopy can be used to characterize triplexes[76,77], it has rather similar caveats to those of UV spectroscopy for analysing intermediate structures, coexisting species or large nucleic acids. Although several methods are available, the interest in ^{19}F NMR to investigate triplexes arises from its ability to probe structural species in a quantitative manner in different surrounding conditions. A deeper understanding of the conformational transitions of triple helical oligonucleotides may increase the medicinal potential of these complex secondary structures. In this thesis, the ^{19}F NMR spectroscopic detection of both DNA[I] and RNA[II] triplexes has been investigated, and new information could be provided.

Nucleic acids can form both intra- (i.e., the triple helical structure is formed within a nucleic acid molecule) and intermolecular (i.e., the triple helical structure is formed between distinct nucleic acid molecules) triple helical structures.[182-184] In a triple helical structure, a single-stranded third strand is bound sequence-specifically to the major groove of a polypurine-polypyrimidine duplex. The third strand can be oriented either parallel or antiparallel relative to the polypurine strand of the duplex. The investigated DNA (Section 3.2.2) and RNA triplexes (Section 3.4) in this thesis are parallel pyrimidine motif triplexes, and antiparallel triplexes are not discussed. In the parallel pyrimidine motif triplexes, T/U•A-T/U and C^+•G-C base triplets (• Hoogsteen base pairing, - Watson-Crick base pairing, C^+ a N3-protonated cytosine) are formed *via* Hoogsteen hydrogen bonding (Figure 36).[182-184] Several factors influence the formation kinetics and stability of the triplex, such as the sequence length, base and backbone composition, pH, salt concentration and temperature.[184,186,187,291-293] The stability of C-rich pyrimidine motif triplexes is strongly dependent on pH, because the C^+•G-C base triplet requires a N3-protonated cytosine (pK_a value for N3 of cytidine is 4.2[126]).[185-189] Therefore, the formation of C-rich triplexes is typically favoured in acidic conditions, which limits their occurrence in physiological conditions. Under acidic conditions, C-rich sequences can also form competitive structures such as CC^+-duplexes and i-motifs that disturb the triplex formation.[190,191] In addition, ligands can affect the formation and stability of triplexes.[294,295] For example, neomycin (an aminoglycoside antibiotic, Figure 36) has been shown to significantly stabilize triplexes.[296-305]

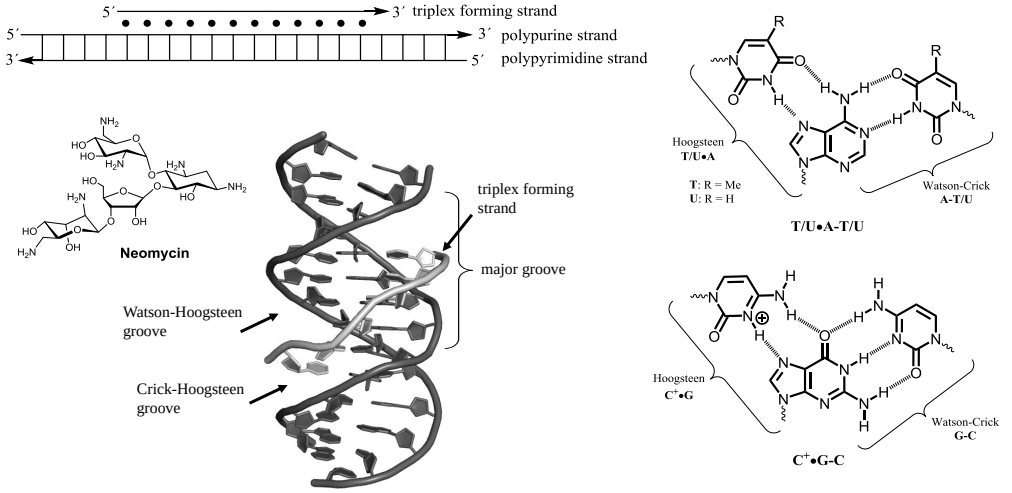


Figure 36. Schematic presentation of parallel pyrimidine motif triplex (• Hoogsteen hydrogen bonding, - Watson-Crick base pairing). T/U•A-T/U and C^+•G-C base triplets. Structure of a parallel pyrimidine motif DNA triplex (PDB ID 1BWG[290]). Neomycin binds in the triplex major groove (i.e., the Watson-Hoogsteen groove).[300]

Several factors limit the efficient triplex targeting and formation in physiological conditions, for example, pH- and cation-dependency, charge and steric repulsion, kinetic properties, self-associated structures, thermal stability, base triplet code, modest affinity, cellular delivery and stability to intracellular nucleases.[266,267,269,294,306,307] Therefore, a considerable number of chemical modifications and conjugate groups have been introduced into TFOs to improve their properties.[266,267,294,306-308] Publication IV describes the syntheses and hybridization properties of triplex-forming aminoglycoside 2´-*O*-methyl oligoribonucleotide conjugates in which aminoglycoside modifications have been incorporated into desired positions of oligonucleotide chains as phosphoramidite building blocks. However, as it is not related to ^{19}F NMR spectroscopy, it is not discussed in this thesis, even though it is included in the PhD studies of the author.

3.2.2 Characterization of DNA triplexes

DNA triplexes were investigated using 4´-*C*-[(4-CF_3-1*H*-1,2,3-triazol-1-yl)methyl]thymidine (**1**).[I] As discussed, due to the orientation of the reporter group and the *S*-type sugar conformation, it was expected that **1** is a sensitive non-invasive derivative in DNA environments. The triplex structure chosen for the studies is the same pH-dependent parallel pyrimidine motif triplex **ON1**/**ON4**/**ON5** (Figure 37) that was previously investigated by Tanabe *et al.*[35] using 2´-deoxy-5-fluorouridine (Section 1.2.2.7). They had incorporated the label in the third strand and showed that it was able to detect triplex formation. In addition, they monitored the pH- and temperature-dependent behaviour of the triplex-forming strand. Similarly to this previous study, **1** was positioned at the middle of a pyrimidine strand (**ON1**). **ON1** is the third strand in the **ON1**/**ON4**/**ON5** triplex (i.e., paired via the Hoogsteen face) (Figure 37). To also monitor the behaviour of the underlying duplex, an **ON1**/**ON2**/**ON3** triplex in which the **ON1** is the pyrimidine strand of the duplex (i.e., paired via the Watson-Crick face) was investigated as well (Figure 37). Tanabe *et al.*[35] did not report whether 2´-deoxy-5-fluorouridine can be used to distinguish triplex and duplex structures when positioned in a Watson-Crick strand.

The influence of **1** on the thermal stability of the structures was investigated by UV spectroscopy. As seen in Table 2, **1** had a minor effect on the stability. The CD-profiles additionally confirmed that the labelled sequences formed triplex structures similarly to their native counterparts.

Table 2. UV-melting temperatures (T_m/ºC) of the **1**-modified secondary structures

entry	oligonucleotides	T_m/ºC (ΔT_m/°C)
1	**ON1/ON2**-duplex	54.9 (-0.6)
2	**ON1/ON2/ON3**-triplex	32.9 (-0.6)
3	**ON1/ON4/ON5**-triplex	31.0 (-1.7)

ΔT_m in comparison to those obtained with unmodified oligonucleotides (X = thymidine). Conditions: 10 mmol L^{-1} NaH_2PO_4 (pH = 6.0), 0.1 mol L^{-1} NaCl, 2.0 mmol L^{-1} MgCl and 2.0 μmol L^{-1} of each oligonucleotide. UV-detection at 260 nm.

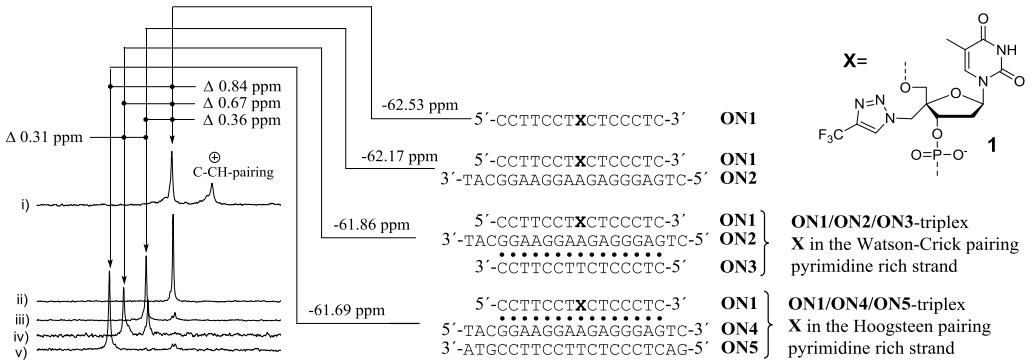


Figure 37. ^{19}F NMR signals of labelled probe in different DNA-hybridization modes. Conditions: i) 100 μmol L^{-1} of **ON1**, 10 mmol L^{-1} sodium phosphate (pH = 5.5) 2 mmol L^{-1} Mg_2Cl and 0.1 mol L^{-1} NaCl in D_2O-H_2O (1:9, v/v) at 25 ºC, ii) the same as i), but pH = 6.0 at 37 ºC, iii) 50 μmol L^{-1} of **ON1**, 45 μmol L^{-1} **ON2**, 10 mmol L^{-1} sodium phosphate (pH = 6.0) 2 mmol L^{-1} Mg_2Cl and 0.1 mol L^{-1} NaCl in D_2O-H_2O (1:9, *v/v*) at 37 ºC, iv) 50 μmol L^{-1} of **ON1**, 50 μmol L^{-1} **ON2**, 50 μmol L^{-1} **ON3**, 10 mmol L^{-1} sodium phosphate (pH = 6.0) 2 mmol L^{-1} Mg_2Cl and 0.1 mol L^{-1} NaCl in D_2O-H_2O (1:9, *v/v*) at 50 ºC, the passive temperature-dependent shift has been eliminated and the shifts have been interpolated to those at 37 ºC, v) 100 μmol L^{-1} of **ON1**, 100 μmol L^{-1} **ON3**, 100 μmol L^{-1} **ON4**, 10 mmol L^{-1} sodium phosphate (pH = 6.0) 2 mmol L^{-1} Mg_2Cl, 0.1 mol L^{-1} NaCl in D_2O-H_2O (1:9, *v/v*) at 37 ºC. (Note: in spite of the virtual homogeneity of the oligonucleotides, the shoulders of the ^{19}F NMR signals most likely originate from impurities).

^{19}F NMR studies were then carried out (detailed measurement conditions are presented in the caption of Figure 37). As seen in Figure 37ii, **ON1** at pH 6.0 gave a sharp signal at -62.53 ppm that corresponds to the single strand, whereas at pH 5.5 (Figure 37i), a broad signal at ~ 0.6 ppm upfield was additionally detected, attributed to unspecific C-CH^+ interactions. This pH-dependent behaviour of **ON1** is consistent with that described by Tanabe *et al.*[35] Next, the **ON1/ON4/ON5** triplex formation was investigated by the titration of **ON1** with the duplex **ON4/ON5**. As seen in Figure 38a, upon the consecutive addition of **ON4/ON5** duplex, a new unique signal appeared at -61.69 ppm indicating the formation of **ON1/ON4/ON5** triplex. As well-separated signals ($\Delta\delta$ = 0.84 ppm) were detected, the temperature-dependent behaviour of the **ON1/ON4/ON5** triplex was then studied to evaluate the applicability of **1** for melting temperature determination. The temperature-dependent ^{19}F NMR spectra are shown in

Figure 38b. When the molar fraction of **ON1** was plotted against the temperature (Figure 38c), a sigmoidal curve resulted. The T_m value could be extracted from the inflection point (66 °C). The obtained T_m value is not comparable to that determined by UV spectroscopy (31 °C) due to the concentration dependency of the triplex formation (100 µmol L^{-1} vs. 2 µmol L^{-1}).

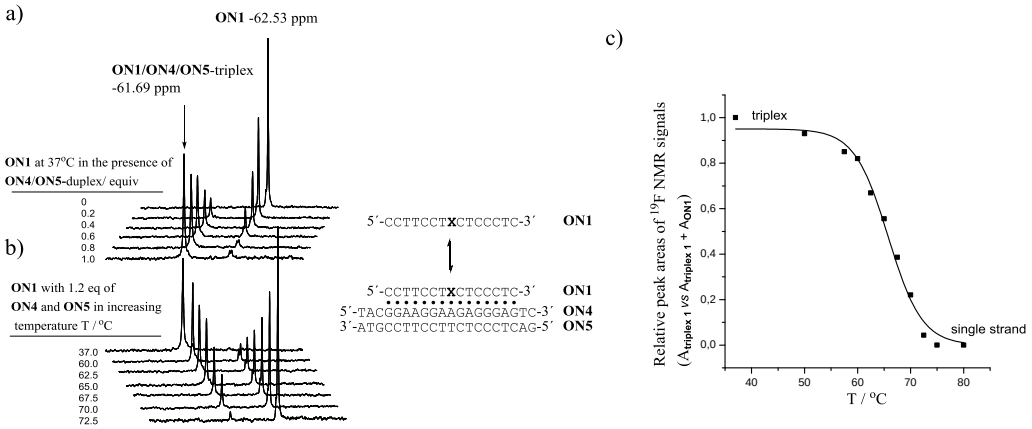


Figure 38. (a) DNA-triplex formation and (b and c) its thermal melting followed by ^{19}F NMR spectroscopy. Conditions: 100 µmol L^{-1} **ON1** (X = **1**), 0-120 µmol L^{-1} **ON4** and **ON5** (1:1, *n/n*), 10 mmol L^{-1} sodium phosphate (pH = 6.0), 2 mmol L^{-1} MgCl and 0.1 mol L^{-1} NaCl in D_2O-H_2O (1:9, *v/v*).

The formation and melting of **ON1/ON2/ON3** triplex in which **ON1** is paired via the Watson-Crick face was investigated in a similar manner. A lower ON concentration (50 µmol L^{-1}) was, however, used for the measurements (a seemingly direct conversion of **ON1/ON2/ON3** triplex to **ON1** was observed at a 100 µmol L^{-1} ON concentration due to the concentration dependency of the triplex formation). As seen in Figures 35 iii and iv, distinct ^{19}F signals were detected for the duplex (-62.17 ppm) and the triplex (-61.86 ppm). Consequently, by this labelling strategy, it was possible to monitor the melting of the triplex and the duplex. The molar fractions of the structural species upon heating were obtained from the relative peak areas, and the T_m values could be extracted from the ^{19}F NMR data (Figure 39). In addition, the denaturation of **ON1/ON2/ON3** triplex was monitored in the presence of neomycin. As seen in Figure 39, a stoichiometric amount of neomycin stabilized the triplex (ΔT_m = 4 °C) whereas the duplex stability remained almost unaffected as expected.

In conclusion, **1** was proven to be suitable for the characterization of DNA triplexes providing well-separated signals for structural species. Compared to the results reported by Tanabe *et al.*[35] **1** provided clearly sharper signals than 2′-deoxy-5-fluorouridine (Figure 18). In addition, the temperature-dependent data obtained using **1** could be used to extract T_m values.

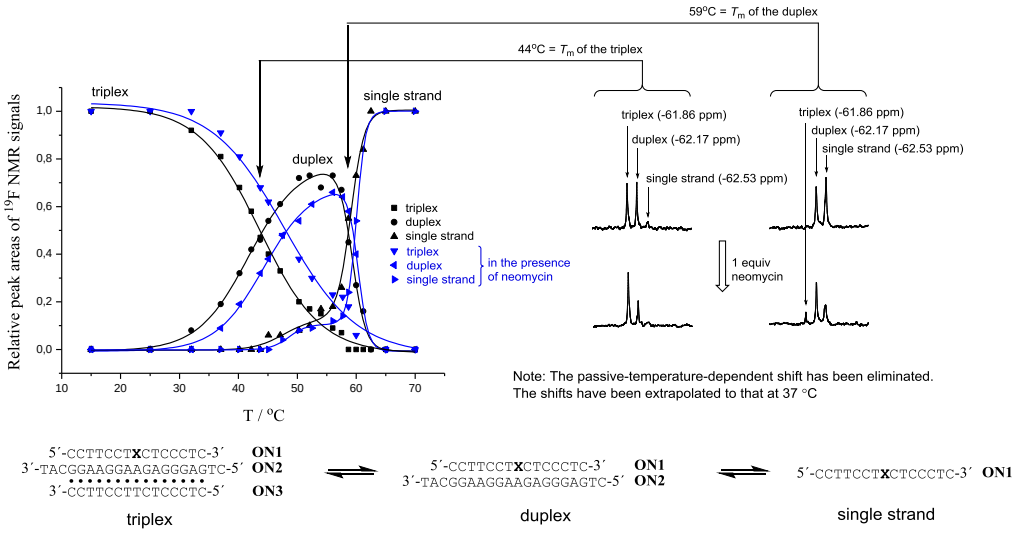


Figure 39. DNA-triplex/duplex/single strand conversions followed by ^{19}F NMR spectroscopy. Conditions: 50 μmol L^{-1} **ON1** (X = **1**), **ON2** and **ON3**, 10 mmol L^{-1} sodium phosphate (pH = 6.0), 2 mmol L^{-1} MgCl and 0.1 mol L^{-1} NaCl in D_2O-H_2O (1:9, v/v).

3.3 Comparison of ^{19}F NMR properties

3.3.1 1 vs 5-[4,4,4-trifluoro-3,3-bis(trifluoromethyl)but-1-ynyl]-2´-deoxyuridine on a HIV-1 TAR-model

Our laboratory has previously applied ^{19}F NMR spectroscopy to characterize the invasion of 2´-*O*-methyl oligoribonucleotides into a ^{19}F-labelled HIV-1 TAR model (Section 1.2.2.6).[24,89] In these previous studies, the 5-[4,4,4-trifluoro-3,3-bis(trifluoromethyl)but-1-ynyl]-2´-deoxyuridine was positioned in the upper stem of the HIV-1 TAR hairpin, and it was able to reflect the denaturation of the hairpin (i.e., **A**/**B** equilibrium) as well as the melting of the macro-loop complex to the open chain invasion complex (i.e., **C**/**D** equilibrium) (Figure 40a). Both of these equilibria were able to be monitored from separate albeit partly coalescent signals, indicating that the equilibria between these intramolecular secondary structures were slow on the NMR time scale. As intramolecular hybridizations are usually fast on the NMR time scale, it was attributed that the 5-[4,4,4-trifluoro-3,3-bis(trifluoromethyl)but-1-ynyl]-2´-deoxyuridine itself deviated the equilibrium rate. Therefore, the same TAR model was modified with **1**, and the invasion studies were repeated.[I]

Three **1**-modified HIV-1 TAR models were synthesized in which **1** was positioned at the loop (**ON6**), at the upper stem (**ON7**) or at the lower stem (**ON8**) (Figure 40a). The labelling site of **ON7** was the same as that previously used for 5-[4,4,4-trifluoro-3,3-bis(trifluoromethyl)but-1-ynyl]-2´-deoxyuridine. As seen in Table 3, the effect of **1** was either marginal or slightly destabilizing, depending on the substitution site.

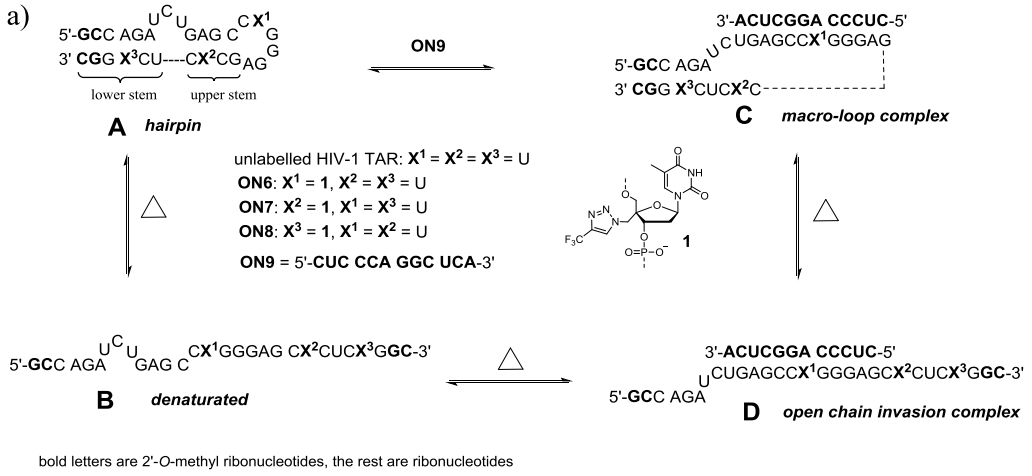
Table 3. UV-melting temperatures (T_m/ºC) of the **1**-modified HIV-1 TAR models

entry	oligonucleotides	T_m/ºC (ΔT_m/°C)
1	**ON6**	72.4 (+0.5)
2	**ON7**	71.0 (-0.9)
3	**ON8**	69.2 (-2.7)

Conditions: 10 mmol L^{-1} sodium cacodylate (pH = 7.0), 0.1 mol L^{-1} NaCl, and 2.0 µmol L^{-1} of each oligonucleotide. ΔT_m (in parentheses) in comparison to that obtained with unlabelled HIV-1 TAR model. UV-detection at 260 nm.

The temperature-dependent behaviour of the ^{19}F-labelled HIV-1 TAR models were monitored in the presence and absence of 1.0 equivalent of a 12 mer 2´-*O*-methyl oligoribonucleotide (**ON9**) (Figure 40a). ^{19}F NMR measurements were carried out using a 50 µmol L^{-1} oligonucleotide concentration in a 10 mmol L^{-1} sodium cacodylate buffer containing 0.1 mol L^{-1} NaCl, pH 7.0, D_2O-H_2O (1:9, *v/v*). (Previous conditions: 20 µmol L^{-1} ^{19}F-labelled HIV-1 TAR with **ON9** in 0.1 mol L^{-1} NaCl with 25 mmol L^{-1} sodium phosphate, pH = 6.5, D_2O-H_2O (1:9, *v/v*)).[24] As seen in Figure 40b, the labels positioned into the double-helical regions (**ON7** and **ON8**) allowed the monitoring of the hairpin melting (**A**/**B** equilibrium). In contrast, the label positioned in the loop region (**ON6**) was insensitive to the melting of the hairpin and shifted continuously (over 0.5 ppm) over the measured temperature range. The T_m values obtained for **ON7** (71 °C) and **ON8** (68 °C) from ^{19}F NMR data were close to those determined by UV spectroscopy (**ON7** and **ON8** 71.0 °C and 69.2 °C, respectively). Compared to the previous study, in which the melting occurred in a slow exchange regime[24], well-behaving coalescence signals were now obtained (Figure 41a). By **ON8**, it was possible to monitor the equilibrium between the macro-loop complex and the open chain invasion complex (**C**/**D** equilibrium). Also here, an averaged ^{19}F signal between the structures (**C**/**D**) was detected (Figure 41b). The signal was broad indicating that the process was slower compared to the melting of the hairpin (**A**/**B** equilibrium). A T_m value (39 °C) was obtained from the inflection point of the melting profile (Figure 40b). The label located in the upper stem of the TAR (**ON7**) was also able to sense the **C**/**D** equilibrium, but the obtained shift versus temperature profile was unclear. Although the signals of the hairpin (**A**) and the macro loop complex (**C**) overlapped (Figure 40b), the titration with **ON9** could be followed at 55ºC, *i.e.*, under conditions where the hairpin (**A**) was directly converted to the open chain invasion complex (**D**) (Figure 41c). Finally, the thermal denaturation of the open chain invasion complex (**B**/**D** equilibrium) could be monitored from the relative peak areas using **ON6** (Figure 41d). As a result, these studies showed that **1** is also an informative relative non-invasive modification in RNA environments,

despite the expected modest discrimination due to the orientation of the 4´-*C* modification and the preferred C2´-endo conformation (cf. Section 3.1.3).



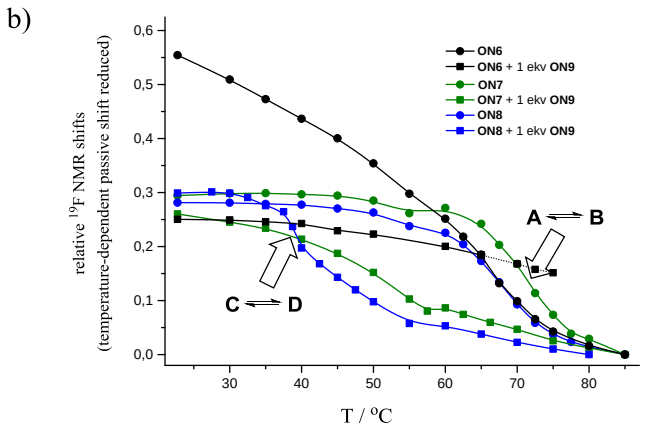


Figure 40. (a) Fluorine-labelled TAR models and schematic presentation of temperature-dependent behaviour of invasion complex. (b) ^{19}F NMR shift difference versus temperature profiles of ^{19}F-labelled HIV-TAR models (the passive temperature-dependent shift reduced). Conditions: 50 μmol L^{-1} ^{19}F-labelled HIV-1 TAR (**ON6–8**) with and without 50 μmol L^{-1} of **ON9**, 10 mmol L^{-1} sodium cacodylate (pH = 7.0) and 0.1 mol L^{-1} NaCl in D_2O-H_2O (1:9, *v/v*).

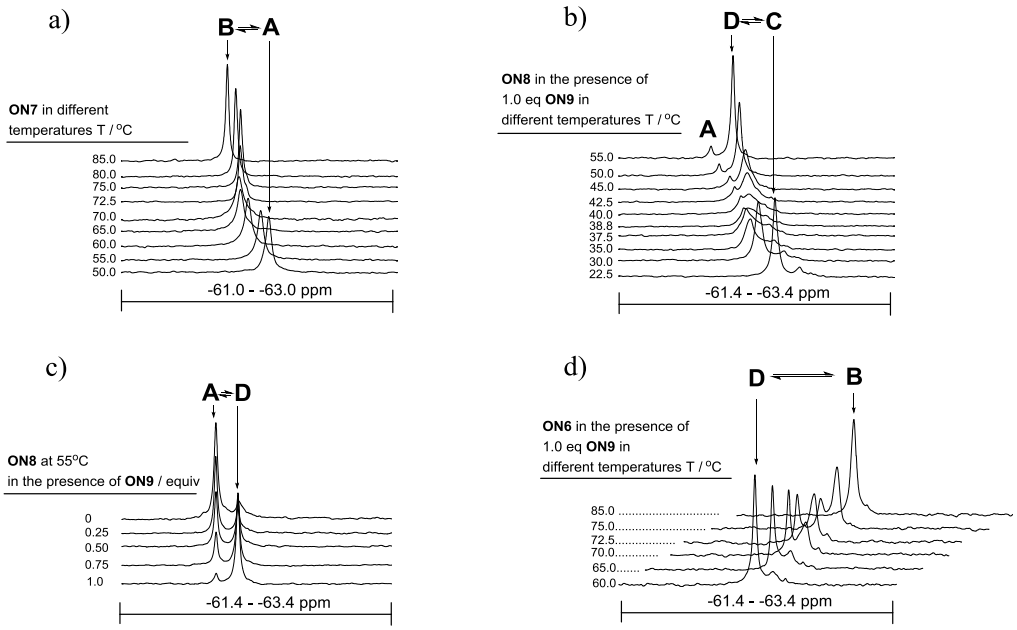


Figure 41. ^{19}F NMR spectra of ^{19}F-labelled HIV-TAR models and their invasion complexes under increasing temperature. Conditions: 50 μmol L^{-1} labelled HIV-1 TAR model (**ON6–ON8**) with and without 50 μmol L^{-1} **ON9** (titration carried out in c), 10 mmol L^{-1} sodium cacodylate (pH = 7.0) and 0.1 mol L^{-1} NaCl in D_2O-H_2O (1:9, *v/v*).

In summary, the invasion mechanism could be characterized with **1**, and the results were consistent with those of previous ^{19}F NMR studies[24,89]. However, compared to the temperature-dependent data obtained with 5-[4,4,4-trifluoro-3,3-bis(trifluoromethyl)-but-1-ynyl]-2´-deoxyuridine[24], the intramolecular melting transitions (**A**/**B** and **C**/**D** equilibria) now occurred under a fast exchange regime, as expected for intramolecular hybridizations. Thus, these results seem to suggest that a highly modified label such as 4,4,4-trifluoro-3,3-bis(trifluoromethyl)but-1-ynyl may misrepresent the denaturation rate, although the modification does not significantly destabilize the secondary structure (ΔT_m -2.7 °C compared to native TAR[24]). As a limitation, **1** was unable to reflect the equilibrium between the hairpin and macro-loop invasion complex (**A**/**C** equilibrium), while 5-[4,4,4-trifluoro-3,3-bis(trifluoromethyl)but-1-ynyl]-2´-deoxyuridine provided well-separated signals for these structures. Therefore, it may be concluded that 5-[4,4,4-trifluoro-3,3-bis(trifluoromethyl)but-1-ynyl]-2´-deoxyuridine was more suitable for titration studies, whereas **1** reflected the temperature-dependent behaviour of complexes more reliably.

3.3.2 1, 2 and 3 on a miR-215 hairpin

Several studies presented in the Introduction Section (1) are based on the labels' ability to distinguish single- and double-stranded environments. This shift dispersion ability is typically used to evaluate novel fluorinated nucleoside analogues for further studies, as this property can be readily applied to characterize more complex structures. To evaluate and compare the properties of **1**–**3**, they were introduced into a pri-miRNA-215 hairpin model[309] (**ON10**–**ON12**, Figure 42), and their ability to reflect an intramolecular melting process was monitored.[II] UV melting studies did not show differences between the derivatives on the thermal stability of the hairpin, and the modifications had only a minor effect on the stability of the hairpin (in each case, ΔT_m < -1.0 °C compared to the unlabelled pri-miRNA-215 hairpin model). The ^{19}F NMR measurements were carried out using an RNA concentration of 50 μmol L^{-1} in a mixture of 0.1 mol L^{-1} NaCl and 10 mmol L^{-1} sodium cacodylate at pH 6.0, D_2O–H_2O (1:9, *v/v*). Each of the probes gave a coalescence signal that discriminated between the hairpin (**E**) and denaturated form (**F**), as expected for intramolecular hybridization. As seen in Figure 42, in each case, a sigmoidal melting profile was obtained, and comparable T_m values (~ 55 °C) were obtained from the inflection points. However, clear differences could be observed in the shift dispersion between the single- (**F**) and double-stranded (**E**) environment. **3** provided the largest shift dispersion ($\Delta\delta$ = 0.82 ppm, **ON12**), which is likely related to the shielding of the H-5´,5´´ protons of the preceding nucleotide in the helix. The modest shift dispersions of **1** ($\Delta\delta$ = 0.23 ppm, **ON10**) and **2** ($\Delta\delta$ = 0.12 ppm, **ON11**) are likely related to the orientation of the reporter groups. As discussed, the substituent at the 4´-*C* site orients outward from the RNA helix, and thus the reporter groups probably encounter only modest changes in the local environment upon hybridization. The applicability of **1** an RNA environment was already demonstrated with HIV-1 TAR studies. However, **1** is designed for DNA, and its shift dispersion between the single- and double-stranded DNA environments is better than that in RNA. In addition to the poor shift response, the synthetic properties of **2** are also disadvantageous: the synthesis of the corresponding phosphoramidite of **2** (i.e., **22**) was complex, and it coupled with lower efficiency. Therefore, no further studies were conducted with **2**. In conclusion, among these derivatives, **3** seems to be the most promising for RNA studies: it provided a large shift dispersion, the synthesis of the corresponding phosphoramidite (i.e., **29a**) was facile, and it coupled efficiently in the DNA/RNA synthesizer.

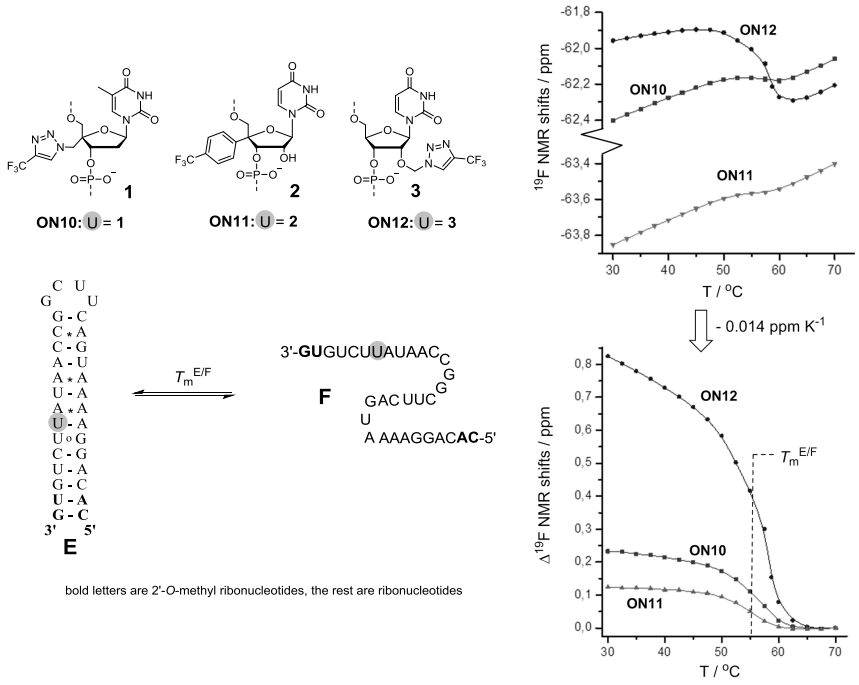


Figure 42. ^{19}F NMR shift versus temperature profiles and profiles of the shift differences versus temperature after the subtraction of a passive temperature-dependent shift (0.014 ppm K^{-1}) obtained by **1**, **2** and **3** upon the thermal denaturation of a PRI-miR-215 hairpin model (**E**/**F**). Conditions: 50 µmol L^{-1} of ON in 0.1 mol L^{-1} NaCl and 10 mmol L^{-1} sodium cacodylate at pH 6.0, D_2O–H_2O (1:9, *v/v*).

3.4 RNA triplexes

Similar to DNA, RNA can form parallel pyrimidine motif triplexes *via* Hoogsteen hydrogen bonding (Section 3.2.1, U•A-U and C$^+$•G-C base triplets Figure 36). In fact, the first reported triplex was a poly(rU)poly(rA)poly(rU) RNA triplex.[310] Although triplex interactions are ubiquitous structural motifs in RNA tertiary structures[311], the biological importance of RNA triplexes has been recognized more recently[273,277,278]. Triplexes have been found to play essential functional roles, for instance, in SAM-II riboswitch[312], telomerase[313,314], spliceosome[315] and ribosomal frameshift pseudoknots[316]. Moreover, a triple helix forming RNA structure, called an element for nuclear expression (ENE, previously named an expression and nuclear retention element), has been recognized to act as a general stability and regulatory element.[317-323]

In the present thesis, two triple-helical structures were chosen for ^{19}F NMR experiments: one is an artificial model[288] and the other is a viral stability and regulatory element called

PAN ENE[317]. Both of these structures were labelled with 2´-*O*-[(4-CF_3-1*H*-1,2,3-triazol-1-yl)methyl]uridine (**3**), and their temperature-dependent behaviour was monitored by ^{19}F NMR spectroscopy.[II] In contrast to the above investigated DNA triplexes, these RNA triplexes are stable at neutral pH.

3.4.1 An artificial RNA triplex

The structure of the artificial RNA triplex is shown in Figure 43. It is a parallel pyrimidine motif poly U•A-U triplex construct interrupted by two C^+•G-C triplets, and the strands are tethered to each other by hexaethylene glycol spacers. It has been preciously studied by UV spectroscopy and gel shift mobility assays.[288] To monitor both triplex and duplex structures, **3** was positioned into the polypyrimidine Watson-Crick strand.

3.4.1.1 UV and CD spectroscopy studies

The UV melting curve and CD profile of **ON13** in comparison of those of the unlabelled RNA are shown in Figure 43. The sensor had a marginal effect on the duplex stability, and it also only slightly destabilized the triplex. As seen in the CD profiles (Figure 43b), the labelling did not induce structural distortions.

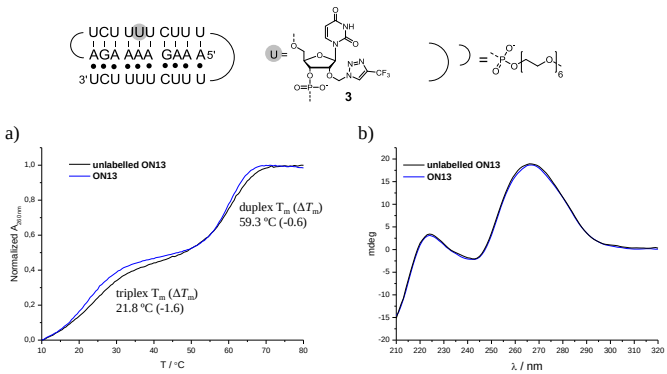


Figure 43. (a) UV-melting profile and T_m values of **ON13** (blue) in comparison to those of unlabelled RNA (black). ΔT_m values (in parentheses) in comparison to those obtained with unlabelled RNA. (b) CD-profile of **ON13** (blue) in comparison to that of unmodified RNA (black) at 10 °C. Conditions: 10 mmol L^{-1} sodium cacodylate (pH = 7.0), 0.1 mol L^{-1} NaCl, 2.0 µmol L^{-1} of each oligonucleotide.

3.4.1.2 ^{19}F NMR studies

The ^{19}F NMR measurements were performed using a mixture of 50 µmol L^{-1} **ON13** in 0.1 mol L^{-1} NaCl and 10 mmol L^{-1} sodium cacodylate at pH 7.0, D_2O–H_2O (1:9, *v/v*). As can be seen in Figure 44, the thermal denaturation of the intramolecular duplex (**H** → **I**) could be monitored by the chemical shift changes of the obtained ^{19}F NMR coalescence signal (Figure 44a). The melting temperature of the duplex ($T_m^{H/I}$ = 59 °C) could be extracted from the inflection point of the sigmoidal curve (Figure 44b) after the subtraction of the temperature-dependent ^{19}F NMR shift (Figure 44c). The obtained melting temperature is in good agreement with the T_m value determined by UV spectroscopy (59.3 °C). At a lower temperature (40 °C), a new distinct ^{19}F signal appeared (Figure 44a). The relative peak area of this signal increased when the temperature was decreased. At 32.5 °C, the signal was of equal size to the duplex-related signal. The new signal was assigned to a triplex that denatures at 32.5 °C. However, the distinct signals for the duplex and triplex indicated that the equilibrium between these structures may be an intermolecular rather than an intramolecular process. The ^{19}F NMR measurements were then conducted at a lower **ON13** concentration. As seen in Figure 44a, the triplex formation was indeed concentration-dependent: at a higher **ON13** concentration (50 µmol L^{-1}), the duplex and triplex signals were of equal size at 32.5 °C (= $T_m^{G/H}$), whereas at a lower **ON13** concentration (5 µmol L^{-1}), only a trace of the triplex signal was detected at the same temperature. Thus, the detected triplex was concentration-dependent and may refer to an intermolecular triplex dimer of **ON13** (**G**) instead of the expected intramolecular triplex (**J**). Finally, the molar fractions of the triplex (**G**), duplex (**H**) and single strand (**I**) (Figure 44c) were extracted from the relative peak areas (**G** vs **H**, Figure 44a), and the shift versus temperature profile after the subtraction of the passive temperature-dependent shift (**H** vs **I**, Figure 44b).

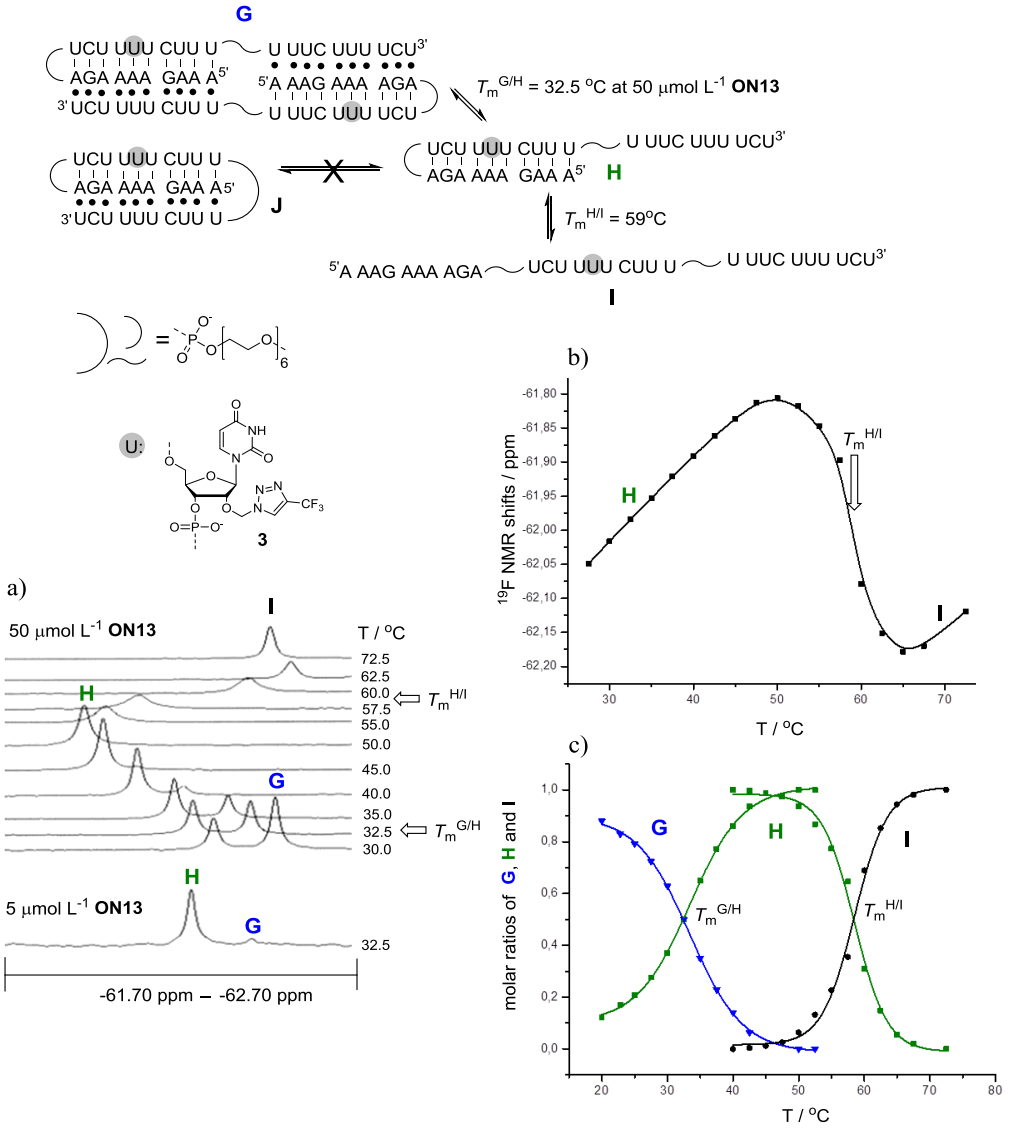


Figure 44. (a) ^{19}F NMR spectra of **ON13** at different temperatures (5 and 50 µmol L^{-1} of **ON13** in 0.1 mol L^{-1} NaCl and 10 mmol L^{-1} sodium cacodylate at pH 7.0, D_2O–H_2O (1:9, *v/v*)). (b) ^{19}F NMR shift versus temperature profile of **ON13** (**H** vs. **I**). (c) Molar fractions of the proposed dimeric triple helix **G**, duplex **H** and single strand **I** of **ON13**. The molar fractions of the triplex, duplex and single strand were extracted either from the relative peak areas or the shift versus temperature profile after the subtraction of the passive temperature-dependent shift.

3.4.2 PAN ENE polyA complex

The PAN ENE structure was identified in a viral transcript produced by an oncogenic human gammaherpesvirus, Kaposi's sarcoma-associated herpesvirus (KSHV).[317] During the lytic phase (production of infectious virions), KSHV produces a long noncoding RNA (lncRNA) polyadenylated nuclear (PAN) RNA.[324] The PAN RNA is a 1.1 kb long intronless transcript containing a 5´ cap and a 3´ poly(A) tail, and it is the most abundant viral transcript during the lytic phase. A 79-nt stem-loop structure, called ENE exists near the 3´-end. The ENE contains an asymmetric U-rich internal loop that interacts in *cis* with the poly(A) tail, forming a triple-helical structure (Figure 45).[317,324,325] This stabilizing interaction protects PAN RNA from a nuclear deadenylation-dependent decay pathway and hence enables its accumulation in the nucleus at very high levels. Since the characterization of the PAN-ENE polyA complex, ENE-like structures have been discovered, for example in other viruses[319], plants[322], fungi[322] and, interestingly, in human cancer associated lncRNAs metastasis-associated lung adenocarcinoma transcript 1 (MALAT-1)[318,320] and multiple endocrine neoplasia-β (MENβ)[318].

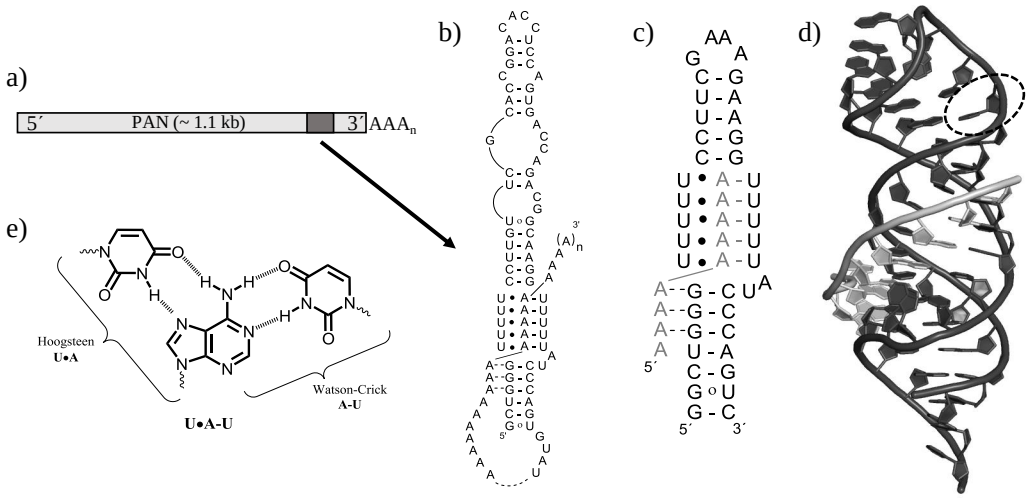


Figure 45. (a) Schematic presentation of PAN RNA. (b) The predicted secondary structure of the PAN ENE complexed with a polyA tail. (c) The schematic presentation of the core ENE model complexed with A_9 (9 mer adenine oligoribonucleotide) and (d) the crystal structure of this model (A_9 grey) (PDB ID 3P22[317]). The uridine residue that was replaced by **3** is circled. (e) U•A-U base triplet.

A 40 mer PAN ENE model bearing the core ENE structure used in previous studies[317,318] was chosen for the ^{19}F NMR investigations (Figure 45). **3** was incorporated into an upper stem of the model and in addition, four base pairs at the lower stem were replaced with corresponding 2´-*O*-methyl ribonucleotides (**ON14**, Figure 46).

3.4.2.1 UV and CD spectroscopy studies

UV-melting studies showed (Table 4) that the **3** had a marginal effect on the stability of the PAN ENE hairpin. The T_m values for the PAN ENE A_9 complexes, however, could not be determined from the melting curves due to the strong background absorbance of the hairpin. This is consistent with a previously reported UV study[318] in which an unclear melting profile was obtained for the complex. According to CD spectroscopy the substitution did not induce conformational changes.

Table 4. UV-melting experiments (T_m/ºC) of the ^{19}F-labelled **ON14**

entry	oligoribonucleotides	hairpin T_m/ºC	complex T_m/ºC
1	**ON14**	68.8 (-0.2)	-
2	**ON14** + A_9	68.3 (-0.8)	n.d.[a]

n.d: absorbance change referring to denaturation of **ON14** + A_9-complex severely overlapped with background absorbance of **ON14**. Inflection point cannot be determined. ΔT_m (in parentheses) in comparison to those obtained with unlabelled RNAs. Conditions: 10 mmol L^{-1} sodium cacodylate (pH = 7.0), 0.1 mol L^{-1} NaCl, 2.0 µmol L^{-1} of each oligonucleotide, UV-detection at 260 nm.

3.4.2.2 ^{19}F NMR studies

^{19}F NMR measurements were performed using a mixture of 50 µmol L^{-1} **ON14** and 50 µmol L^{-1} A_9 in 10 mmol L^{-1} sodium cacodylate buffer (pH 7.0) containing 0.1 mol L^{-1} NaCl, D_2O–H_2O (1:9, *v/v*). Although the **3** was clearly positioned outside the expected binding region of A_9, it was able to distinguish the triple helical complex (**K**) and hairpin (**L**) structures. As seen in Figure 46a, the complex behaved nearly as **ON13**: a well-behaving coalescence signal was obtained for the hairpin melting, whereas the intermolecular triplex complex gave a distinct signal. The molar fractions of the structures and the melting temperatures were extracted from the ^{19}F NMR data (Figure 46b), as in the case of **ON13**. The melting temperatures obtained for the triplex (**K**) and the duplex (**L**) were 20 ºC ($T_m^{K/L}$) and 68 ºC ($T_m^{L/M}$), respectively. As neomycin has been shown to stabilize RNA triplexes[297], the ^{19}F NMR measurements were additionally conducted in the presence of five equivalents of neomycin. The molar ratios of the complex (**K´**), duplex (**L´**) and single strand (**M´**) in the presence of neomycin as a function of temperature are shown in Figure 46b. As expected, neomycin significantly stabilized the triple helical complex ($T_m^{K´/L´}$ = 48 ºC, ΔT_m = 28 ºC) but also did so for the hairpin ($T_m^{L´/M´}$ = 83 ºC, ΔT_m = 15 ºC).

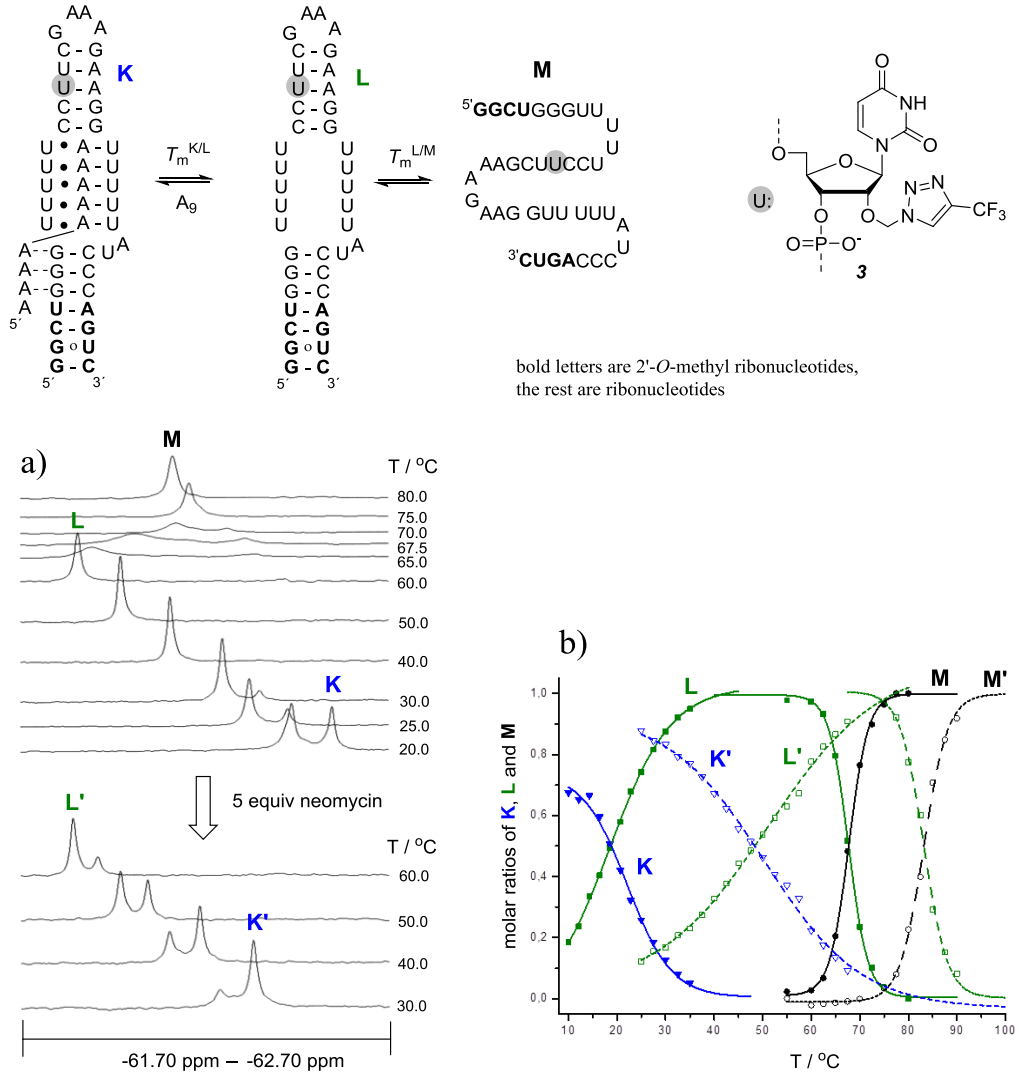


Figure 46. (a) ^{19}F NMR spectra of **ON14** + A_9 in different temperatures in the presence and absence of neomycin (5 equiv) (50 µmol L^{-1} **ON14** and A_9 in 0.1 mol L^{-1} NaCl and 10 mmol L^{-1} sodium cacodylate at pH 7.0, D_2O–H_2O (1:9, *v/v*)). (b) Molar fractions of different secondary structures at different temperatures (symbols **K**, **L**, **M** and solid lines refer to a mixture without neomycin, and symbols **K´**, **L´**, **M´** and dotted lines to a mixture in the presence of neomycin). The molar fractions of the complex, duplex and single strand were extracted either from the relative peak areas or the shift versus temperature profile after the subtraction of the passive temperature-dependent shift.

3.5 G-quadruplex-hairpin equilibria

RNA G-quadruplexes formed by guanine-rich RNA sequences have gained considerable attention in recent years. RNA G-quadruplexes may be involved in several biological processes, such as translation, telomere homeostasis and the processing of ncRNAs.[193,195,196,198,199,326,327] G-quadruplexes (both RNA and DNA G-quadruplexes) are four-stranded nucleic acid structures formed by stacked G-quartets connected by loops (Figure 47).[192-194] The core unit is a planar G-quartet (tetrad) in which four guanine residues are linked together by Hoogsteen hydrogen bonding (Figure 47a). The G-quadruplex structures can be intramolecular or intermolecular. The formation and stability of G-quadruplexes is dependent on the monovalent cation that is located in the central channel of the structure (Figure 47c). In particular, potassium ions strongly stabilize G-quadruplexes ($K^+ > Na^+ > Li^+$). RNA G-quadruplexes prefer a parallel folding topology (i.e., all four strands are oriented in the same direction, Figure 47b), which simplifies their topological diversity compared to DNA G-quadruplexes. The preference for a parallel topology is due to the 2´-hydroxyl groups, which restrict the guanine bases to the *anti* conformation. The antiparallel topology requires a *syn* conformation, and it is hence prevented in RNA. Several other features also contribute to the structural diversity and stability of the G-quadruplexes.[192-194,328,329]

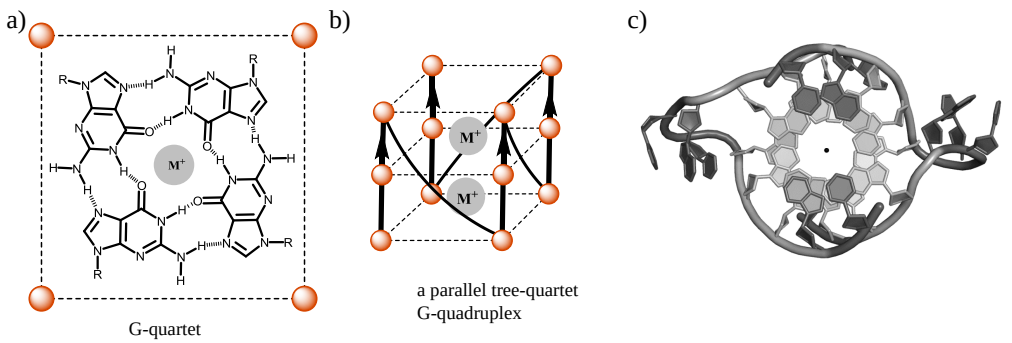


Figure 47. (a) Schematic structure of G-quartet and (b) a parallel tree-quartet G-quadruplex. (c) Crystal structure of bimolecular human telomeric RNA G-quadruplex (PDB ID 3IBK[242]). Potassium ions in central channel are shown as black dots, and guanine residues are in light grey colour.

Structural transitions between RNA secondary structures are common in many RNA-mediated regulatory processes. In some cases, an RNA G-quadruplex structure may be in equilibrium with an alternative secondary structure, usually a hairpin.[330-341] Such intramolecular secondary structural equilibria have recently attracted attention as the meanings of RNA G-quadruplexes in biological processes have started to emerge. Although few bistable hairpin-G-quadruplex (Hp/Qd) structures have been identified so far, the recently published human transcriptome-wide analysis suggested that bistable

Hp/Qd structures may be highly relevant structural motifs that play an important role in gene regulation.[342] Therefore, it might be expected that several Hp/Qd equilibria will be identified, as G-quadruplex involving structural switches are a little explored research area so far. The Hp/Qd equilibrium is sensitive to the surrounding conditions such as the cation content and temperature.[330-337] In addition, the equilibrium position may be modulated by small molecular ligands[330-333,336] and external nucleic acid interactions[341]. Mutations may also affect Hp/Qd transitions.[335] Bistable Hp/Qd structures have been identified both in non-coding and coding RNAs, and they may affect biologically important processes such as pre-miRNA maturation[336,337,339], mRNA splicing[335] and translation[332,340]. Consequently, it has been proposed that Hp/Qd equilibria may be involved in diseases[199], and hence they may also be potential novel therapeutic targets[199,333,334] in addition to RNA G-quadruplexes[193,197,199,343].

The characterization that a G-rich sequence can fold into a G-quadruplex is relatively straightforward[201] and can be assessed by combining several methods, typically by UV melting studies[74], CD spectroscopy[344] and ^{1}H NMR[71,345]. Despite the structural differences, the quantification of molar fractions in a Hp/Qd equilibrium is complex even within short bistable sequences.[330-337,339-341,346,347] Among the employed methods, ^{1}H NMR spectroscopy has been used to provide direct information of coexisting hairpin and G-quadruplex structures.[331-333,335,346] The ^{1}H NMR resonances of the imino protons of these structures exist in distinct spectral regions. The resonances at 10–12 ppm are characteristic of G-quadruplexes (i.e., Hoogsteen base pairing)[71,345], whereas the imino-proton resonances involved in Watson-Crick base pairing may be found at 12–14 ppm[66]. However, the spectral overlap and broadness of the resonance signals interfere with the distribution analysis of the conformers.[71,331-333,335,345,346] As one of the main advantage of ^{19}F NMR is that it provides direct quantitative information of coexisting structural species and it has been successfully used to investigate several intramolecular hairpin-hairpin equilibria[107,111,142,143], it might be a valuable method to investigate Hp/Qd equilibria as well. In the present thesis, a well-characterized artificial Hp/Qd RNA[331] (Figure 48) was chosen for a model system.[III] In addition, two sequences within prion mRNA (Figure 54) that have been previously shown to be Hp/Qd sequences[332] were chosen to probe the general utility of ^{19}F NMR spectroscopy to characterize such structural equilibria.

3.5.1 An artificial G-quadruplex-hairpin model

An artificial Hp/Qd RNA sequence has been constructed by combining a hairpin-forming sequence derived from a pseudoknot structure of a mouse mammary tumour virus and a G-quadruplex-forming sequence from human *N-RAS* mRNA.[331] This sequence (with additional 2´-*O*-methyl modifications) was labelled with purine derivatives **5** (**ON16**) and **6** (**ON17**).[III] The synthesized sequences and their schematic structures are presented in Figure 48. In addition, ^{19}F-labelled hairpin derivatives with a polyU tail (**ON19** and **ON20**) were synthesized as reference structures as well as the non-labelled analogues (**ON15** and **ON18**). The influence of the 2´-*O*-(4-CF_3-1*H*-1,2,3-triazol-1-yl)methyl reporter group was analysed by UV and CD spectroscopy and the ^{1}H NMR analysis of imino protons.

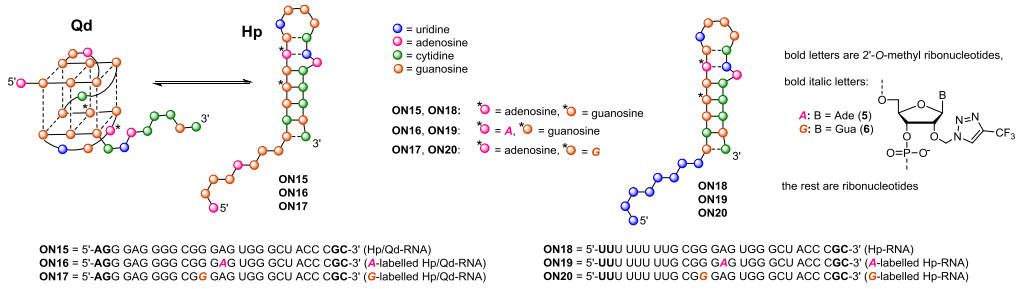


Figure 48. Synthesized sequences and their schematic structures.

3.5.1.1 UV melting studies

UV melting studies were performed using a 2 µmol L^{-1} ON concentration in 0.1 mol L^{-1} NaCl, 10 mmol L^{-1} sodium cacodylate (pH 7.0). The T_m values are listed in Table 5, and the corresponding UV melting profiles are shown in Figure 49. At 260 nm, Hp/Qd ONs (**ON15**–**17**) gave biphasic melting curves. The melting temperatures of the hairpins were obtained at the latter inflection points (Figure 49a). The T_m values of reference hairpins **ON18**–**20** were comparable with those of the bistable sequences (**ON15**–**17**), indicating that the nucleobase content of the overhang did not affect the hairpin stability. Interestingly, the melting curves of the reference hairpins (**ON18**–**20**) were also biphasic (Figure 49c). The melting of G-quadruplexes can be followed at 295 nm, and an inverted melting profile at this wavelength is a characteristic feature of G-quadruplexes.[74] As seen in Figure 49b, the bistable sequences **ON15**–**17** gave inverted melting profiles, indicating that the sequences were able to adopt G-quadruplex structures. Although the ^{19}F-labelled RNAs behaved like their non-labelled analogues, the 2´-*O*-(4-CF_3-1*H*-1,2,3-triazol-1-yl)methyl reporter group affected the thermal stability of the structures. The

reporter group destabilized the hairpin structures (ΔT_m^{Hp}: -3.2 – -6.3 °C), whereas it stabilized the G-quadruplex structures (ΔT_m^{Qd}: +2.2 and +4.3 °C).

Table 5. UV-melting temperatures (T_m/°C) of the oligoribonucleotides

entry	oligoribonucleotides	T_m^{Hp}/°C (260 nm)	T_m^{Qd}/°C (295 nm)
Hp/Qd RNAs			
1	**ON15**	74.5 ± 0.0	34.8 ± 0.5
2	**ON16 (*A*)**	70.0 ± 0.2 (-4.5)	37.0 ± 0.2 (+2.2)
3	**ON17 (*G*)**	68.2 ± 0.3 (-6.3)	39.1 ± 0.2 (+4.3)
Hp RNAs			
4	**ON18**	73.9 ± 0.1	n.a.
5	**ON19 (*A*)**	70.7 ± 0.1 (-3.2)	n.a.
6	**ON20 (*G*)**	68.7 ± 0.2 (-5.2)	n.a.

Conditions: 10 mmol L^{-1} sodium cacodylate (pH = 7.0), 0.1 mol L^{-1} NaCl, and 2.0 µmol L^{-1} of each oligonucleotide. ΔT_m (in parentheses) in comparison to that obtained with an unlabelled model (i.e., **ON15** or **ON18**)

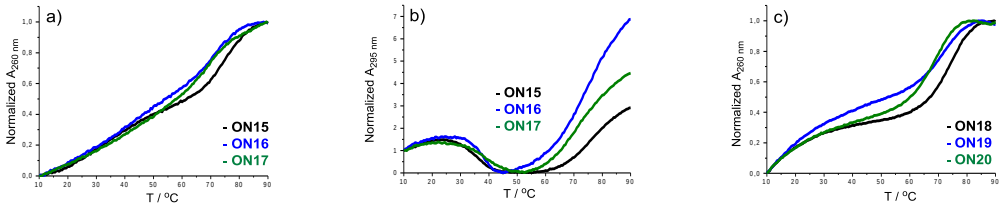


Figure 49. UV-Melting profiles of **ON15–20**. Conditions same as in Table 5.

3.5.1.2 CD spectroscopy studies

The CD spectra of ONs were recorded at 20 °C in a mixture of 10 mmol L^{-1} sodium cacodylate (pH 7.0) and 0.1 mol L^{-1} mM NaCl. The CD profiles of the Hp RNAs (**ON18–20**) displayed characteristic features for an A-form RNA helix[76,78], with a strong positive peak around 260–270 nm and negative bands at 235 nm and 210 nm (Figure 50b). To evaluate the effect of the labelling on G-quadruplex structures, the CD spectra of bistable structures were recorded in the presence of KCl, because the K^+ ions strongly shift the conformational equilibrium toward the G-quadruplex structure. At a 0.1 mol L^{-1} KCl concentration, the CD spectra of **ONs 15–17** showed typical features for a parallel G-quadruplex[344], with a negative peak at 240 nm and a positive peak around 260 nm, but **ON17** (***G***-label) additionally showed a negative turn between 210 and 220 nm, which is an Hp characteristic feature (Figure 50a). Overall, the labelling has minor effects on the secondary structures, except in the case of **ON17**. In addition, the temperature-dependent CD spectra of **ON17** (the same mixture as was used for NMR measurements) were

recorded and compared with those of the non-labelled analogue (**ON15**). As seen by comparing Figures 50 c and d, the CD profiles of **ON17** (50 µmol L^{-1}) were comparable to those of **ON15** (4 µmol L^{-1}) in 10 mmol L^{-1} sodium cacodylate (pH 7.0) with 0.1 mol L^{-1} NaCl (without KCl). Moreover, the Qd-Hp transition (indicated by arrows in the spectra) was observed in both spectra.

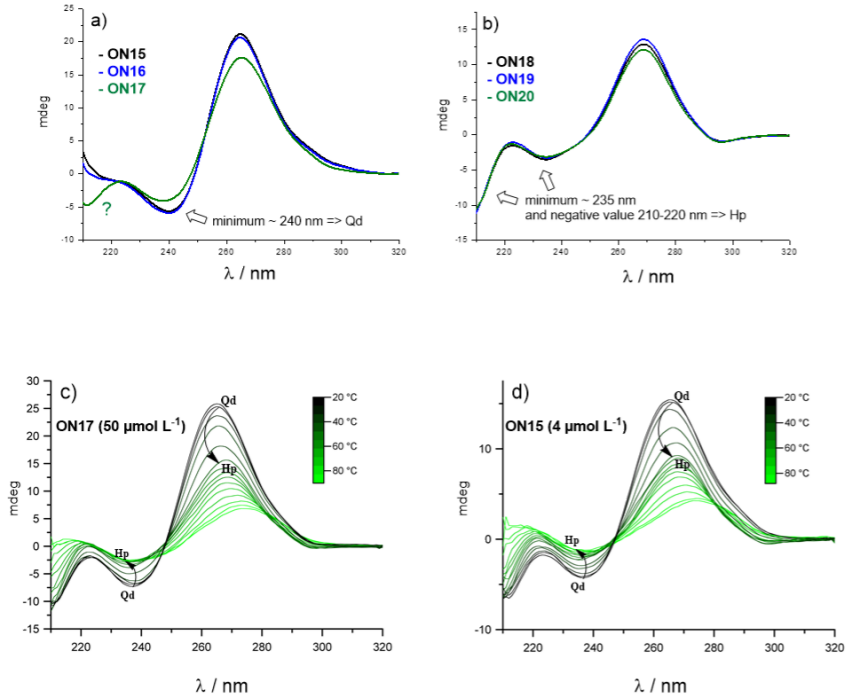


Figure 50. CD spectra. (a) and (b) CD spectra of **ON15–20** at 20 °C. Conditions: 2.0-4.0 µmol L^{-1} ON in (a) 10 mmol L^{-1} sodium cacodylate, 0.1 mol L^{-1} NaCl, 0.1 mol L^{-1} KCl and in (b) 10 mmol L^{-1} sodium cacodylate, 0.1 mol L^{-1} NaCl, pH 7.0, (c) CD spectra of the NMR sample of **ON17** (50 µmol L^{-1}) and (d) 4.0 µmol L^{-1} solution of **ON15** at different temperatures in 10 mmol L^{-1} sodium cacodylate, 0.1 mol L^{-1} NaCl, pH 7.0.

3.5.1.3 ^{1}H NMR studies

The recorded imino-proton ^{1}H NMR spectra of **ON15–20** are shown in Figure 51 (i-viii). The ^{1}H NMR measurements were first performed at a 50 µmol L^{-1} ON concentration in a mixture of 0.1 mol L^{-1} NaCl and 10 mmol L^{-1} sodium cacodylate at pH 7.0 in D_2O-H_2O (1:9, *v/v*). As seen (i-iii), the Hp RNAs (**ON18–20**) gave signals in the range 12–13 ppm, corresponding to regular Watson-Crick base pairs[66]. The 2´-*O*-(4-CF_3-1*H*-1,2,3-triazol-1-yl)methyl reporter group affected the resonance shifts (marked with

asterisks): a downfield shift was observed with the ***A***-label (**ON19**, ii) and an upfield shift with the ***G***-label (**ON20**, iii), but the intensity of the signals remained unchanged. The imino proton resonances of the Hp/Qd RNAs (**ONs 15-17**, iv-vi) were broad, which is consistent with a previously reported study[331]. In addition to a Watson-Crick fingerprint that resembled those of the Hp RNAs (**ON18–20**), broad resonances were detected in the Hoogsteen region (i.e., 10–12 ppm)[71,345]. Although some changes were observed, the reporter group neither caused a significant displacement of the Hp/Qd ratio nor induced local decreases in the signal intensity. Additionally, the Hp/Qd equilibrium was investigated in the presence of Mg^{2+} ions (3.0 mmol L^{-1}) that stabilize the Hp-conformer[331]. When $MgCl_2$ was added into the samples, the signals broadened, but the expected increase in the intensity of the resonances in the Watson-Crick region was barely detectable (iv/vii and vi/viii).

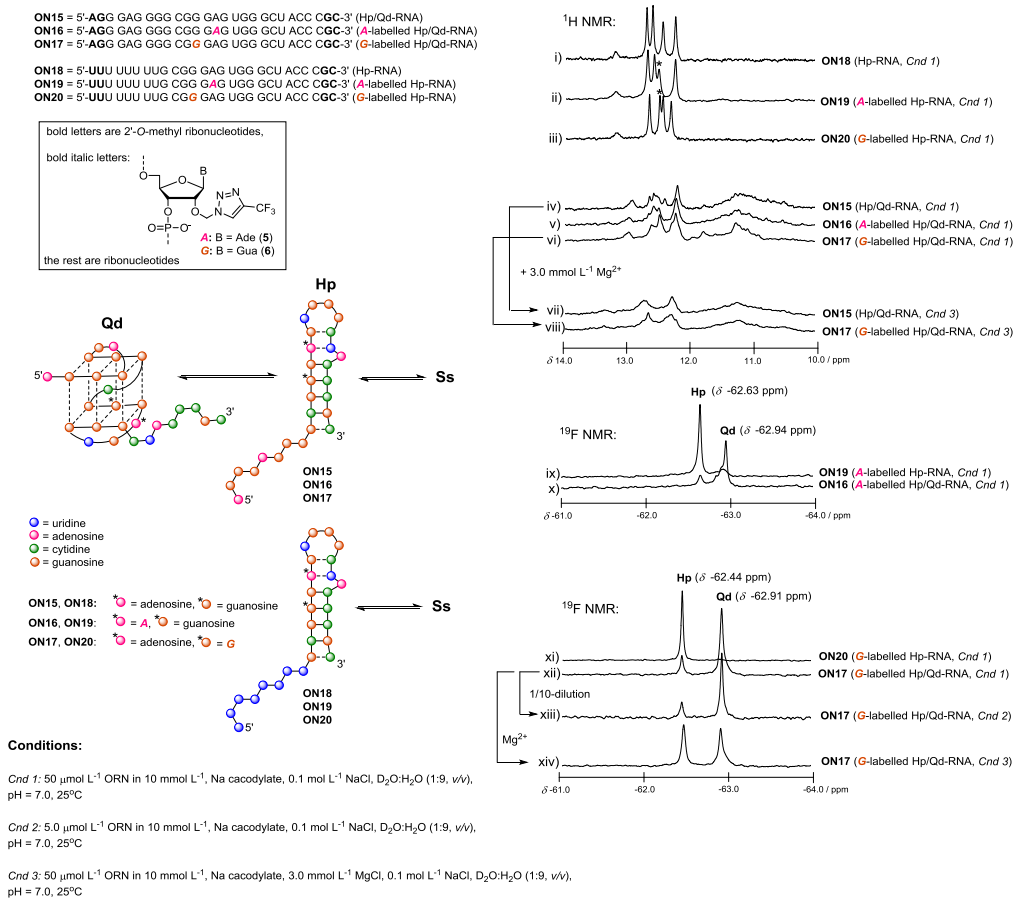


Figure 51. ^{1}H NMR- and ^{19}F NMR-spectra of **ONs 15–20** at 25°C.

3.5.1.4 ^{19}F NMR studies

The ^{19}F NMR spectra (Figure 51 ix-xiv) of **ON**s **16**, **17**, **19** and **20** were then recorded. Hp RNAs **ON19** and **ON20** showed sharp signals at -62.63 ppm (***A***-label, ix) and -62.44 ppm (***G***-label, xi), respectively, whereas two peaks were observed with Hp/Qd RNAs **ON16** and **ON17**. **ON16** (***A***-label, x) gave signals at -62.63 ppm and -62.94 ppm (broad), and **ON17** (***G***-label, xii) at -62.44 ppm and -62.91 ppm. As in both cases the shift of the minor signal corresponded to that of Hp RNAs **ON19** and **ON20,** those signals were attributed to Hp-conformers, and hence the major signals at -62.94 ppm (***A***-label) and -62.91 ppm (***G***-label) could be attributed to Qd-conformers. In both cases, the ratio of the peak areas was 0.25:0.75 (Hp/Qd). The effect of the ON concentration (5.0 μmol L^{-1}) and Mg^{2+} ions (3 mmol L^{-1} $MgCl_2$) on the Hp/Qd equilibrium of **ON17** was investigated. The distribution of the conformers remained unaffected at the lower **ON17** concentration (xiii), indicating that the Hp/Qd equilibrium is not dependent on the concentration (i.e., an intramolecular process). Compared to ^{1}H NMR data (vi/vii), a Mg^{2+} ion-induced shift to the Hp-conformer was clearly detected: the Hp/Qd ratio changed from 0.25:0.75 to 0.54:0.46 (xii/xiv).

Temperature-dependent measurements

The temperature-dependent measurements were carried out using 50 μmol L^{-1} of each ON in 10 mmol L^{-1} sodium cacodylate, 0.1 mol L^{-1} NaCl, D_2O–H_2O (1:9, *v/v*), pH 7.0. The temperature-dependent ^{19}F NMR spectra of Hp/Qd RNAs **ON16** and **ON17** (together with those of the Hp RNAs **ON19** and **ON20**) are shown in Figure 52b and 52d. As seen, the Qd-related ^{19}F-resonances of **ON16** and **ON17** decreased upon heating, and the Hp-related ones increased, existing as a predominant signal upon passing 50 ºC. The ratio of the signals (Hp/Qd peak areas) was reproducible, whether the sample was heated or cooled (2.5 ºC / 90 min) to the target temperature. As seen in Figure 52c, the shifts of the Hp-related resonances of the Hp/Qd RNAs (**ON16** and **ON17**) were in good agreement with those of the Hp RNAs (**ON19** and **ON20**) in almost the whole temperature range except for temporally incomplete Qd/Hp/Ss coalescence signals at 52.5–62.5 ºC (**ON16**) and 57.5–65.0 ºC (**ON17**). At temperatures over 65 ºC, the ^{19}F resonances of **ON16** and **ON17** followed regular upfield-shifted Hp/Ss-coalescence curves. Due to the modest slope of the Hp/Ss-coalescence curves of **ON16** and **ON19**, the denaturation range of the Hp-conformer was not easily extracted from the data. In contrast, the ***G***-label reflected the melting of the Hp-conformer better, and rather steep upfield-shifted Hp/Ss-coalescence curves were observed with **ON17** and **ON20.** A minor signal referring to an incomplete Hp/Ss coalescence signal and/or partially degraded RNA was observed after the complete denaturation of each RNA (marked with asterisks*, fused with the major signal at 80°C). The contribution of this signal (Ss*) was taken into account in determining the molar fractions of the Qd/Hp/Ss-transitions. The molar fractions of the **ON16** and **ON17** conformers as functions of temperature are shown in Figure 52a. In practice, the ratio of the Qd- and Hp-conformers was extracted from the relative peak areas (at 25–70 °C), and that of the Hp- and Ss-conformers (at

55–90 °C) from the shift versus temperature profiles. The contribution of the Ss* was taken into account by subtracting it from the molar fractions of the conformers, and an internal standard was used to confirm the changes in the peak areas. The molar fractions of the Qd- and Hp-conformers at temperatures below 50 °C could be determined accurately, but due to the modest shift discrimination between the Qd- and Ss-related signals (*) the molar fractions of the Qd-conformers at temperatures over 50 °C were approximations (the extent of the Ss-related signal * in the overlapped signals was based on that observed for Hp RNAs **ON19** and **ON20**).

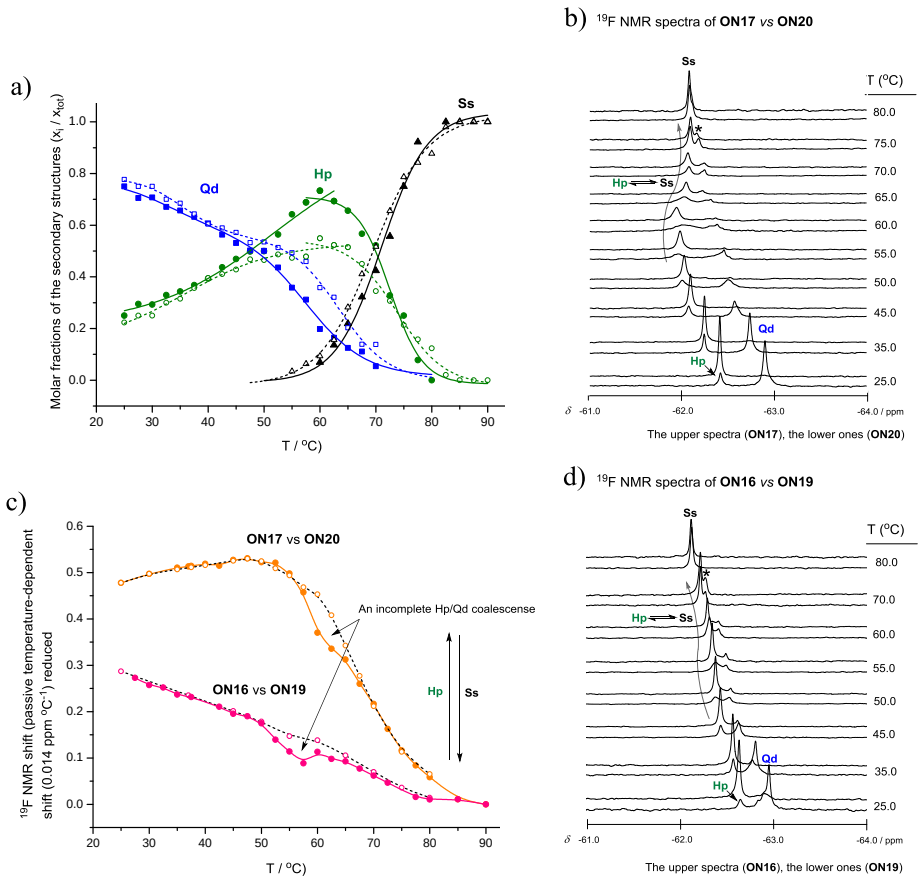


Figure 52. Monitoring of the temperature-dependent Qd/Hp-transitions of **ON16** and **ON17** by ^{19}F NMR spectroscopy. (a) Molar fractions (x) of the secondary structures: Qd (squares, blue lines), Hp (circles, green lines) and Ss (triangles, black lines), **ON16** (filled symbols, solid lines) and **ON17** (open symbols, dotted lines); (b) ^{19}F NMR spectra of **ON17** and **ON20** (50 µmol L^{-1} ON in 10 mmol L^{-1} sodium cacodylate, 0.1 mol L^{-1} NaCl, D_2O–H_2O (1:9, *v/v*), pH 7.0 at 25–80 °C); (c) ^{19}F NMR resonance shift versus temperature profiles of **ON17** (orange filled circles, solid line), **ON20** (orange open circles, dotted line), **ON16** (red filled circles, solid line) and **ON19** (red open circles, dotted line); (d) ^{19}F NMR spectra of **ON16** and **ON19** (50 µmol L^{-1} ON in 10 mmol L^{-1} sodium cacodylate, 0.1 mol L^{-1} NaCl, D_2O–H_2O (1:9, *v/v*), pH 7.0 at 25–80°C).

To characterize the thermodynamics of the Qd→Hp-transition, the stability constant (K = [Qd]/[Hp]) was determined at different temperatures. As seen in Figure 53, lnK was linearly dependent on the inverse temperature from 25 to 50 °C whereas severe dispersion was observed above 50 °C. Consequently, the thermodynamic parameters for the Qd→Hp-transition could be extracted from the linear region. The obtained parameters were as follows: **ON16**: ΔS^{θ} = 116 ± 2 J K^{-1} mol^{-1}, ΔH^{θ} = 37.3 ± 1.5 kJ K^{-1} mol^{-1}, ΔG^{θ} (298.2 K) = 2.7 ± 2.1 kJ K^{-1} mol^{-1} and **ON17**: ΔS^{θ} = 112 ± 5 J K^{-1} mol^{-1}, ΔH^{θ} = 37.1 ± 1.7 kJ K^{-1} mol^{-1}, ΔG^{θ} (298.2 K) = 3.8 ± 3.3 kJ K^{-1} mol^{-1}. Importantly, the parameters obtained for two different sites of the 2´-*O*-(4-CF_3-1*H*-1,2,3-triazol-1-yl)methyl reporter group (**ON16** *vs* **ON17**) were comparable to each other. The turning point of the Qd-Hp transition ([Qd] = [Hp], ΔG^{θ} = 0) was determined from the average values and was assigned to be 53 °C. The obtained midpoint transition differs markedly from the T_m^{Qd} values obtained by UV spectroscopy (**ON16**: 37.0 ± 0.2 ºC and **ON17**: 39.1 ± 0.2 ºC). The UV-based Qd-melting profiles at 295 nm were interfered with the denaturation of the Hp-conformers (the increased absorbance > 50°C, Figure 49b). The obtained UV-values suggested false T_m^{Qd} values detected from the first inflection points (**ON16**: 32.8 ± 6.7 ºC, **ON17**: 35.5 ± 1.5 ºC) of the ^{19}F NMR-based biphasic melting profiles (Figure 52a).

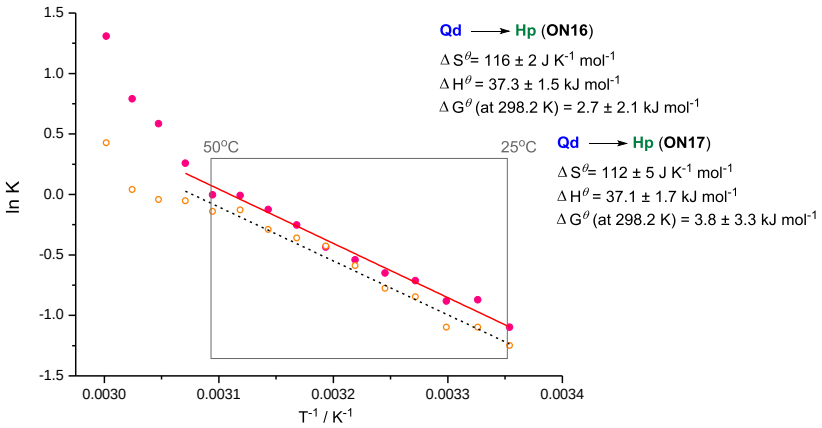


Figure 53. ln K/T^{-1} curves of **ON16** (filled red circles, solid line) and **ON17** (orange open circles, dotted line).

3.5.2 Prion sequences

Prion diseases, such as Jakob-Creutzfeldt, scrapie, bovine spongiform encephalopathy (mad cow disease), are neurodegenerative diseases caused by the conversion of a normal cellular prion protein (PrP^{C}) into an infectious misfolded form of the prion protein (PrP^{Sc}).[348,349] The mechanism that initiates this conversion is unknown. Currently, nucleic acids are considered one of the potential candidates to induce this misfolding.[350-352] Interestingly, it has been proposed that the G-quadruplexes might be involved in the conversion process[352], because G-quadruplex forming aptamers have been shown to bind with high affinity to PrP^{c} and to prevent the conversion of PrP^{c} to PrP^{Sc} [353-357] but also to induce formation of PrP^{Sc} by lowering the free energy barrier between PrP^{c} and PrP^{Sc} [358]. Therefore, it has been speculated that PrP's own mRNA might play a role in the conversion process because it contains several consecutive guanine stretches.[332,352]

Two previously investigated bistable Hp/Qd sequences within prion mRNA[332] were labelled with 2´-*O*-[(4-CF_3-1*H*-1,2,3-triazol-1-yl)methyl]cytidine (**4**). In contrast to the artificial sequence discussed above that can adopt a tree-quartet G-quadruplex, the bistable prion sequences can fold into two-quartet G-quadruplexes that are usually less stable than the tree-quartet ones[329,359]. In addition to ^{19}F-labelled Hp/Qd sequences (**ON22** and **ON24**), a ^{19}F-labelled hairpin without an overhang was synthesized as a reference structure (**ON26**) as well as the corresponding non-labelled analogues (**ON21**, **ON23** and **ON25**) (Figure 54). The influence of the **4** was evaluated as above by UV and CD spectroscopy and the ^{1}H NMR of imino protons.

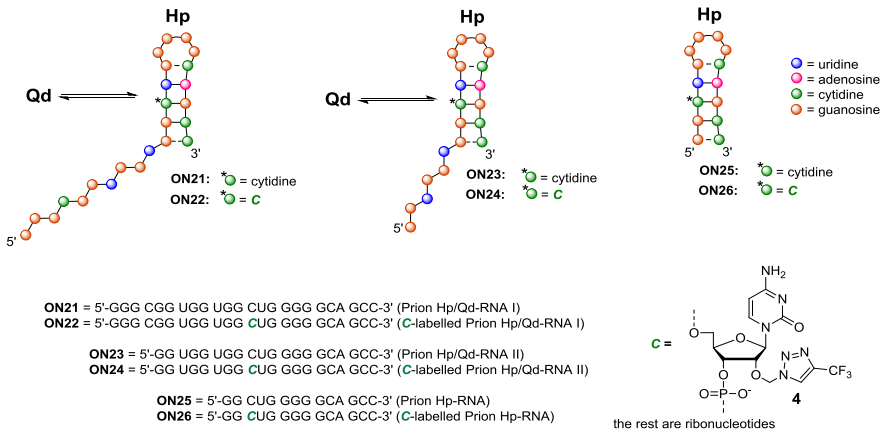


Figure 54. The synthesized prion sequences and their schematic structures.

3.5.2.1 UV melting studies

The T_m values are listed in Table 6, and the corresponding UV melting profiles are shown in Figure 55. UV melting studies were performed using a 2 µmol L^{-1} ON concentration in 0.1 mol L^{-1} NaCl, 10 mmol L^{-1} sodium cacodylate (pH 7.0). Hp/Qd RNAs **ON21** and **ON22** gave biphasic melting curves at 260 nm (Figure 55a), and the melting temperatures of the hairpins were obtained at the latter inflection points. The deletion of the overhang did not significantly affect the hairpin stability (Hp-RNAs **ON25** and **26** vs Hp/Qd RNAs **ON21**–**24**) (Figure 55a, c and d). The melting behaviours of the G-quadruplexes were monitored in the presence of K^+ ions. Hp/Qd-RNAs I (**ON21** and **ON22**) gave inverted melting profiles characteristic of G-quadruplexes[74] in the presence of 20 mmol L^{-1} KCl, whereas such profiles were not observed with Hp/Qd-RNAs II (**ON23** and **ON24**) even at a higher K^+ concentration (0.1 mol L^{-1}) (Figure 55b). However, the NMR data revealed G-quadruplex formation with **ON23** and **ON24**, which was attributed to the concentration dependency of these G-quadruplexes. The labelled RNAs (**ON22**, **ON24**, **ON26**) behaved like their corresponding non-labelled analogues (**ON21**, **ON23**, **ON25**), but the label destabilized the hairpin structures (ΔT_m^{Hp} = -2.7 – -3.9) and stabilized the G-quadruplex structure (ΔT_m^{Qd} = +3.2 °C). Similar effects were observed above with **ON**s **15**–**20**.

Table 6. UV-melting temperatures (T_m/°C) of the oligoribonucleotides

entry	oligoribonucleotides	T_m^{Hp}/°C (260 nm)	T_m^{Qd}/°C (295 nm)
Hp/Qd-RNAs I			
1	**ON21**	71.5 ± 0.3	55.2 ± 0.7[a]
2	**ON22 (*C*)**	68.8 ± 0.2 (-2.7)	58.4 ± 0.4 (+3.2)[a]
Hp/Qd-RNAs II			
3	**ON23**	73.0 ± 0.1	n.d.
4	**ON24 (*C*)**	69.1 ± 0.2 (-3.9)	n.d.
Hp-RNAs			
5	**ON25**	71.7 ± 0.3	n.a.
6	**ON26 (*C*)**	67.9 ± 0.3 (-3.8)	n.a.

Conditions: [a]2 µmol L^{-1} ON in 10 mmol L^{-1} sodium cacodylate (pH = 7.0), 0.1 mol L^{-1} NaCl, 20 mmol L^{-1} KCl, other experiments 10 mmol L^{-1} sodium cacodylate (pH = 7.0), 0.1 mol L^{-1} NaCl, and 2.0 µmol L^{-1} of each oligonucleotide. ΔT_m (in parentheses) in comparison to that obtained with unlabelled model.

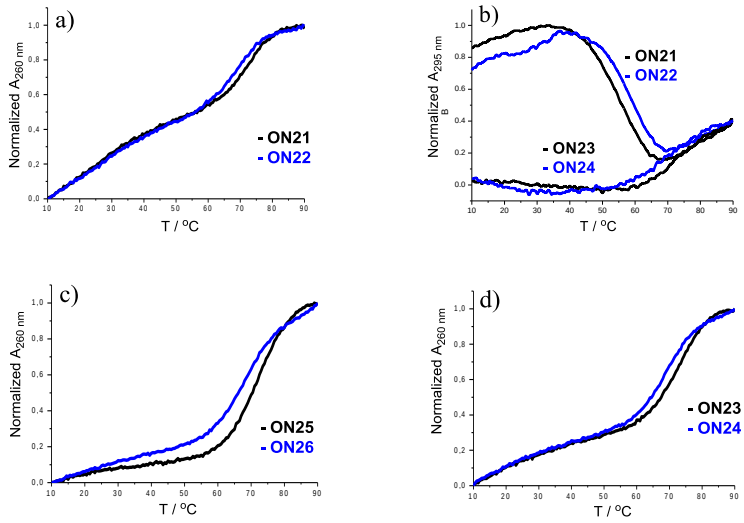


Figure 55. UV-Melting profiles of **ON21**–**26**. Conditions: (a,c,d): 2 µmol L^{-1} **ON21–26** in 10 mmol L^{-1} sodium cacodylate, 0.1 mol L^{-1} NaCl, pH 7.0, (b) 2 µmol L^{-1} **ON21** and **ON22** in 10 mmol L^{-1} sodium cacodylate, 0.1 mol L^{-1} NaCl, 20 mmol L^{-1} KCl, pH 7.0 and 2 µmol L^{-1}, **ON23** and **ON24** in 10 mmol L^{-1} sodium cacodylate, 0.1 mol L^{-1} NaCl, 0.1 mol L^{-1} KCl, pH 7.0.

3.5.2.2 CD spectroscopy studies

The CD spectra of prion ONs were recorded at 20 °C in a mixture of 10 mmol L^{-1} sodium cacodylate (pH 7.0) and 0.1 mol L^{-1} NaCl with or without added KCl. The CD profiles of the Hp RNAs (**ON25** and **26**) displayed characteristic features for an A-form RNA helix[76,78] (Figure 56c). At 20 mmol L^{-1} KCl the CD spectra of Hp/Qd-RNAs I (**ON21** and **ON22**) showed Qd-characteristic CD profiles[344] (Figure 56a). At 0.1 mol L^{-1} KCl the Hp/Qd RNAs II (**ON23** and **ON24**) displayed Qd-characteristic negative peaks at 240 nm, but also Hp-characteristic negative values between 210 and 220 nm (Figure 56b). Overall, the **4** did not induce significant structural perturbations. The temperature-dependent CD spectra of **ON22** used for NMR measurements were recorded and compared with those of the non-labelled analogue **ON21**. As seen, the CD profiles of **ON22** (50 µmol L^{-1}, Figure 56d) were comparable to those of **ON21** (4 µmol L^{-1}, Figure 56e) in 10 mmol L^{-1} sodium cacodylate (pH 7.0) with 0.1 mol L^{-1} NaCl and 20 mmol L^{-1} KCl. In contrast to **ONs 15** and **17** (Section 3.5.1.2), no clear transitions were observed, and Hp/Qd-RNAs I showed Qd-characteristic profiles throughout the temperature ramp.

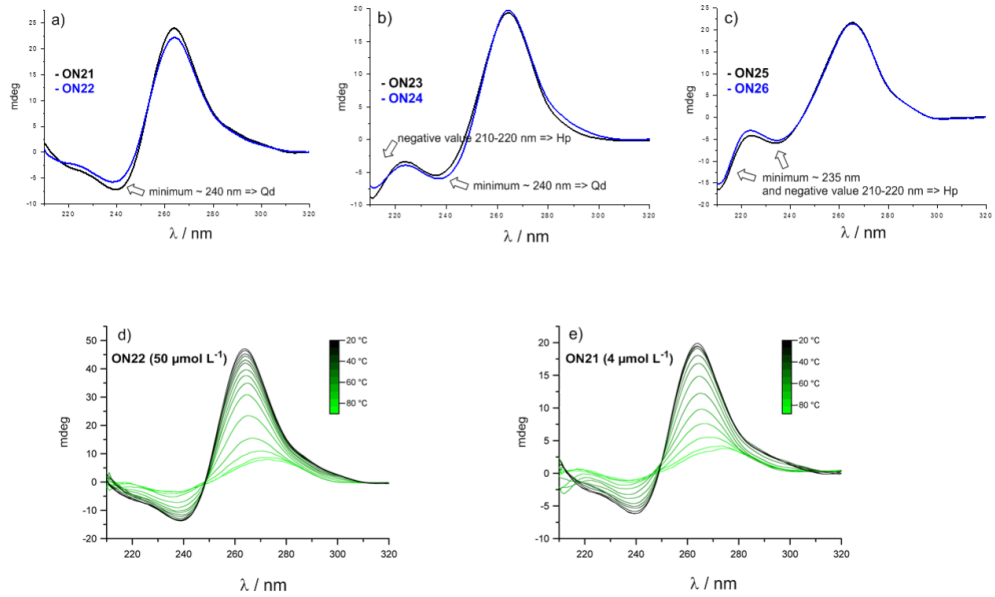


Figure 56. (a-c) CD-spectra of **ON21–26** at 20ºC. Conditions: 2.0–4.0 µmol L^{-1} ON in (a) 10 mmol L^{-1} sodium cacodylate, 0.1 mol L^{-1} NaCl, 20 mmol L^{-1} KCl, pH 7.0; (b) 10 mmol L^{-1} sodium cacodylate, 0.1 mol L^{-1} NaCl, 0.1 mol L^{-1} KCl, pH 7.0 and (c) 10 mmol L^{-1} sodium cacodylate, 0.1 mol L^{-1} NaCl, pH 7.0. (d) Temperature-dependent CD spectra of the NMR sample of **ON22** (50 µmol L^{-1}) and (e) 4.0 µmol L^{-1} solution of **ON21** in 10 mmol L^{-1} sodium cacodylate, 0.1 mol L^{-1} NaCl, 20 mmol L^{-1} KCl, pH 7.0.

3.5.2.3 ^{1}H NMR studies

As above with **ONs 15**–**20**, the effect of the labelling on prion mRNAs models was analysed by imino proton ^{1}H NMR spectroscopy. The imino proton ^{1}H NMR spectra are shown in Figure 57 (i-x). The ^{1}H NMR measurements were first performed at a 50 µmol L^{-1} ON concentration in a mixture of 0.1 mol L^{-1} NaCl and 10 mmol L^{-1} sodium cacodylate at pH 7.0, D_2O–H_2O (1:9, *v/v*) without added KCl (Figure 57 i-vi). As seen, a downfield-shifted Watson-Crick signal (marked with asterisks) was observed for the labelled sequences, but otherwise the ^{1}H NMR spectra of the labelled RNAs were comparable to those of the non-labelled analogues (i/ii, iii/iv, v/vi). Besides the Watson-Crick resonances, low-intensity resonances were observed in the Hoogsteen region of Hp/Qd-RNAs I (**ON21** iii and **ON22** iv), indicating the presence of Qd-conformers. In contrast, Hoogsteen-related resonances were not observed with the Hp/Qd-RNAs II (**ON23** v and **ON24** vi). The imino-proton ^{1}H NMR spectra were then recorded in the presence of KCl (Hp/Qd-RNAs I 20 mmol L^{-1} and Hp/Qd-RNAs II 0.1 mol L^{-1}). As expected, the intensity of the Hoogsteen-related resonances increased (**ON21** vii and

ON22 viii). Hoogsteen-related resonances were also observable for Hp/Qd-RNAs II (**ON23** ix and **ON24** x), showing that these sequences were able to adopt Qd-conformers. The obtained ^{1}H NMR data of the Hp/Qd-RNAs II was hence inconsistent with the UV melting results, which did not show Qd formation. As the difference between the measurements was the concentration (2 µmol L^{-1} vs 50 µmol L^{-1}), the observed Hoogsteen resonances of **ON23** and **ON24** may be attributed to an intermolecular process rather than to an intramolecular Hp/Qd transition.

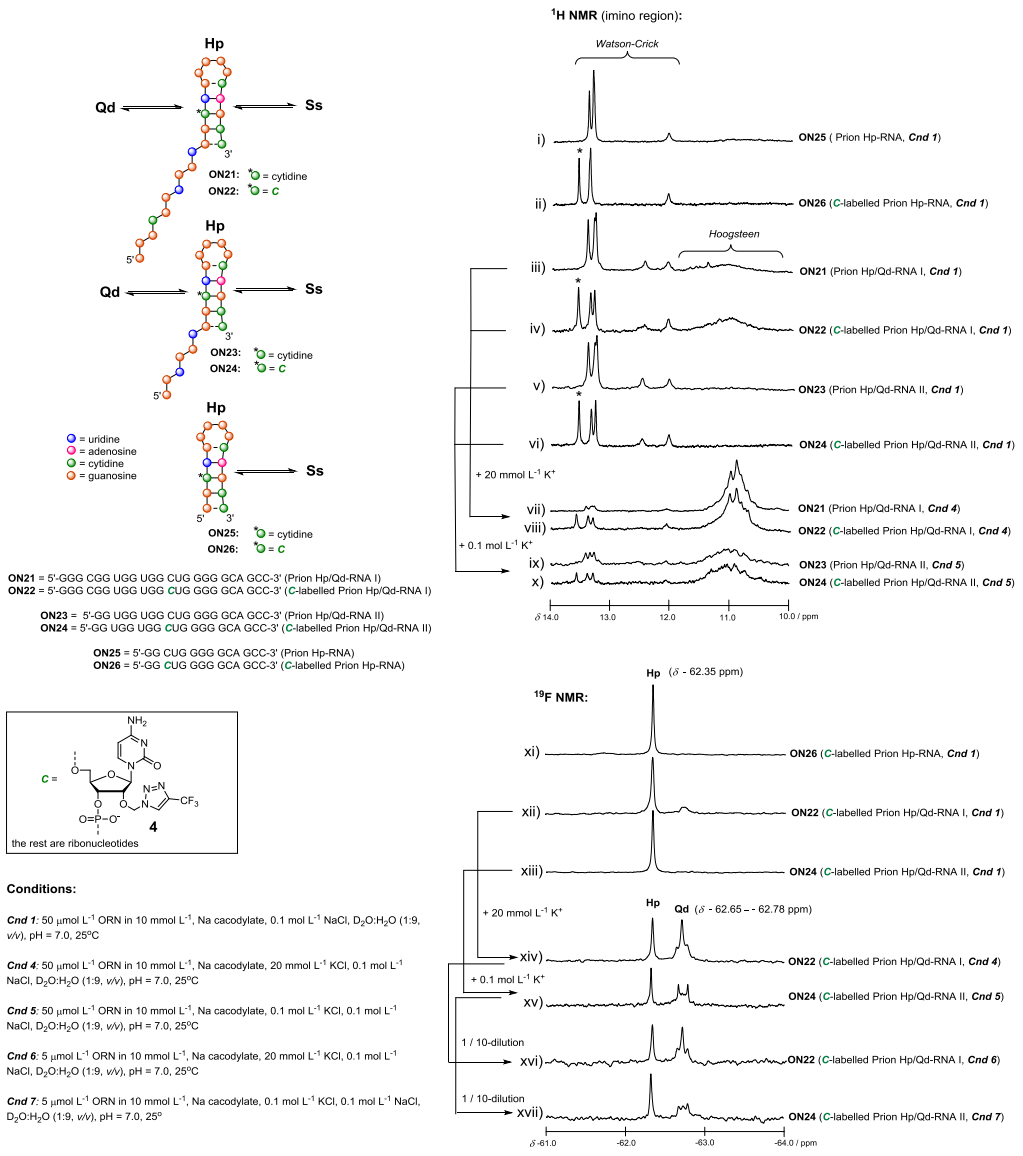


Figure 57. ^{1}H NMR- and ^{19}F NMR-spectra of **ONs 21–26** at 25°C.

3.5.2.4 ^{19}F NMR studies

The same samples used in the ^{1}H NMR experiments were used for the ^{19}F NMR experiments, and the recorded ^{19}F NMR spectra are shown in Figure 57 (xi-xvii). The hairpin (**ON26**) gave a sharp signal at -62.35 ppm (xi), and this signal was the predominant one with both of the Hp/Qd sequences (**ON22** xii and **ON24** xiii). Compared to the ^{1}H NMR spectrum of the **ON22**, only a trace of the signal related to the Qd-conformer was detected in the ^{19}F NMR spectrum (iv/xii). This difference may be related to the slow proton exchange of the Hoogsteen hydrogen bonds[71]. In the presence of KCl (**ON22** 20 mmol L^{-1} xiv and **ON24** 0.1 mol L^{-1} xv), the Qd-conformer was clearly detected with both sequences. However, in both cases, a set of Qd-signals were observed instead of a single signal, attributed to a mixture of Qd-conformers. To evaluate the concentration dependency of the structures, the ^{19}F NMR spectra of the ONs were recorded at lower RNA concentrations (5 µmol L^{-1} versus 50 µmol L^{-1}). As seen, the intensity of the Qd-related signals of **ON24** decreased (xv/xvii), indicating that intermolecular interactions affect the equilibrium. In contrast, the ratio of the conformers of **ON22** remained unaffected (xiv/xvi).

*The temperature-dependent behaviour of **ON22***

The measurements were carried out using 50 µmol L^{-1} **ON22** in 10 mmol L^{-1} sodium cacodylate, 0.1 mol L^{-1} NaCl, 20 mmol L^{-1} KCl, D_2O–H_2O (1:9, *v/v*), pH 7.0. The molar fractions of the structural species (Qd, Hp, Ss) as a function of temperature are shown in Figure 58a. Because a set of signals referring to a set of Qds interfered with the evaluation of the Hp/Qd ratio, an internal standard (i.e., 5-[4,4,4-trifluoro-3,3-bis(trifluoromethyl)but-1-ynyl]-2´-deoxyuridine) was used to evaluate the changes in the peak areas [molar fractions: x(Qd) = 1 – x(Hp + Ss), x(Hp + Ss) = A(Hp + Ss)/Atot compared to an standard]. The molar fractions of Hp and Ss were extracted from the shift versus temperature profile. The Hp-conformer of **ON22** did not accumulate as in the case of artificial **ON**s **16** and **17**. The Qd/Hp ratio remained almost unchanged throughout the temperature ramp, and the melting temperatures (T_m(Qd) = 70.3 ± 1.1 °C and T_m(Hp) = 72.7 ± 0.5 °C) of the conformers were close to each other. The obtained T_m value for the Qd-conformer was much higher than that determined by UV spectroscopy (58.4 °C). As seen in Figure 58b, changes in the peak set referring to Qd-conformers were observed upon heating. In addition, the equilibration rate of the Qd/Hp–ratio of **ON22** was questionably slow (despite the Qd/Hp-ratio being reproducible whether the sample was heated or cooled to the target temperature). Taken together, it may be concluded that intermolecular Hoogsteen interactions disturbed the Qd/Hp ratio in the concentration used for the ^{19}F NMR analysis (50 µmol L^{-1} of **ON22**), although the intramolecular Qd/Hp transition may take place at a lower concentration[332].

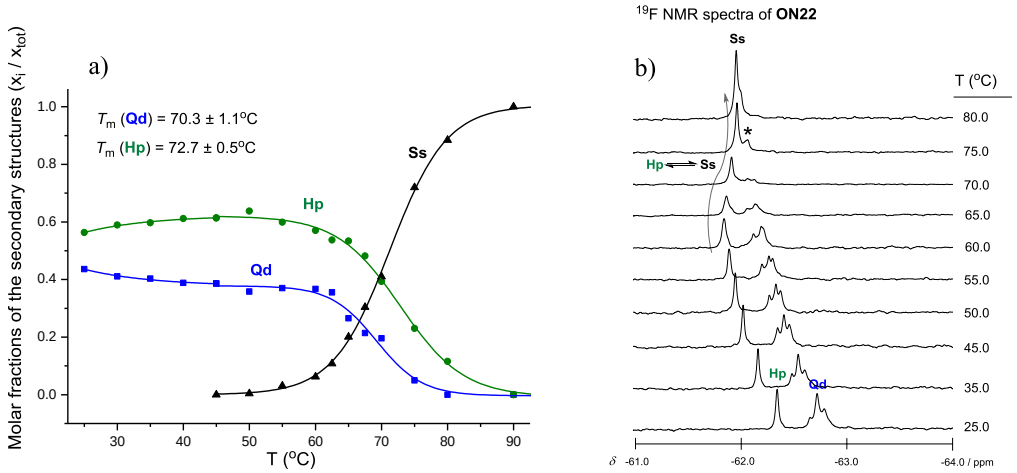


Figure 58. Monitoring of the temperature-dependent Qd/Hp-transition of **ON22** by ^{19}F NMR spectroscopy. (a) Molar fractions (x) of the secondary structures: Qd (squares, blue line), Hp (circles, green line) and Ss (triangles, black line); (b) ^{19}F NMR spectra of **ON22** (50 µmol L^{-1} **ON22** in 10 mmol L^{-1} sodium cacodylate, 0.1 mol L^{-1} NaCl, 20 mmol L^{-1} KCl, D_2O–H_2O (1:9, *v/v*), pH 7.0 at 25–90°C.

4. CONCLUSIONS

Six novel fluorine-labelled building blocks were synthesized, and their applicability as fluorine sensors was evaluated. Phosphoramidite building blocks of 4´-*C*-[(4-CF_3-1*H*-1,2,3-triazol-1-yl)methyl]thymidine (**12**) and 2´-*O*-[(4-CF_3-1*H*-1,2,3-triazol-1-yl)methyl]uridine (**29a**), -cytidine (**29b**), -adenosine (**29c**), guanosine (**29d**) were obtained in relatively high yields. The (4-CF_3-1*H*-1,2,3-triazol-1-yl)methyl reporter group was easily accessible in each case. Importantly, the 2´-*O*-(4-CF_3-1*H*-1,2,3-triazol-1-yl)methyl group could be introduced to all four common ribonucleosides, enabling incorporation at any position in RNAs. In contrast, the phosphoramidite building block of 4´-*C*-[4-(trifluoromethyl)phenyl]uridine (**22**) was obtained with a low yield after a multistep synthesis. All the phosphoramidite building blocks were stable under solid-phase synthesis and cleavage conditions, and no special adjustments were required for conventional synthesis procedures. Phosphoramidite building blocks **12** and **29a-d** were coupled efficiently, either manually or as an integral part of oligonucleotide synthesis, whereas a lower coupling efficiency was obtained with **22**.

The ^{19}F NMR spectroscopic experiments showed that the reporter groups provided sharp ^{19}F signals as singlets without proton decoupling techniques. However, in addition to the synthetic shortcomings, a modest discrimination between single- and double-stranded environments was observed with 4´-*C*-[4-(trifluoromethyl)phenyl]uridine (**2**), and it was not used to characterize secondary structures.

The applicability of 4´-*C*-[(4-CF_3-1*H*-1,2,3-triazol-1-yl)methyl]thymidine (**1**) was demonstrated in both DNA and RNA environments. It was proven to be a relatively non-perturbing substitution in both environments. It was able to reflect DNA triplex structures when bound to either a Watson-Crick or a Hoogsteen face. The molar fractions of the triplex/duplex/single strand were obtained from well-separated ^{19}F signals, and T_m values were obtained from temperature-dependent ^{19}F NMR data. In addition, a previously conducted RNA invasion study was repeated, and the temperature-dependent behaviour of the invasion complex was successfully monitored.

2´-*O*-[(4-CF_3-1*H*-1,2,3-triazol-1-yl)methyl]uridine (**3**) was applied to investigate the temperature-dependent behaviour of RNA triplexes, and the applicability of ^{19}F NMR spectroscopy to investigate RNA triplexes was demonstrated for the first time. **3** was proven to be a non-invasive analogue readily reflecting triplex, duplex and single strand environments, and the molar fractions of the structural specifies were easily determined. In addition to the T_m values, the concentration dependency of the artificial triplex model was revealed. Significantly, the temperature-dependent behaviour of a PAN ENE polyA complex was characterized, and T_m values for the triplex complex and hairpin were obtained in the absence and presence of neomycin.

2´-*O*-[(4-CF_3-1*H*-1,2,3-triazol-1-yl)methyl]cytidine (**4**), -adenosine (**5**) and -guanosine (**6**) were used to characterize RNA G-quadruplex/hairpin transitions. The molar fractions of the Hp- and Qd-conformers were determined under different temperatures and ionic conditions. **4**–**6** were also able to reflect the concentration dependency of the Qd-conformation, i.e., the formation of intra- and intermolecular Qd species. In addition, T_m values and thermodynamic parameters were extracted from the temperature-dependent ^{19}F NMR data. Although no significant structural perturbations were observed, except in the case of **6**-modified Qd-conformer, the analogues affected the stability of both conformers and hence the molar ratios of the conformers, and these effects need to be taken in account when directly compared to the native Hp/Qd equilibria. Whether the observed stability effects are general properties of the derivatives or may be overcome by labelling position changes needs to be assessed with further studies. Nevertheless, compared to UV and CD spectroscopy as well as ^{1}H NMR, the simplicity of determining the ratio of the conformers directly from ^{19}F signals in a variety of conditions was demonstrated, and bistable Hp-Qd RNA models were analysed by ^{19}F NMR spectroscopy for the first time.

In summary, due to the efficient synthesis of **12** and **29a-d**, their facile incorporation into oligonucleotides and the beneficial ^{19}F NMR spectroscopic properties, **1** and **3-6** may be considered potential analogues for further ^{19}F NMR studies.

The potential of ^{19}F NMR spectroscopy to investigate nucleic acids has been demonstrated with several studies, but also the common limitations of this method and the importance of the development of novel ^{19}F-modified analogues for further applications have been recognized. In general, the choice of the appropriate ^{19}F-modification for the given application depends on the nucleic acid of interest and the primary experimental goal. In addition to ^{19}F NMR spectroscopy, a variety of other methods are also available to investigate nucleic acids, such as NMR spectroscopy with other nuclei (^{1}H, ^{15}N, ^{13}C) and spectrophotometric methods, each with its strengths and limitations. NMR spectroscopy is in general invaluable tool to investigate nucleic acids. Commonly, spin-labels and multidimensional NMR methods are needed to overcome the reduced chemical shift dispersion of protons.[66-72] One of the advantage of ^{19}F NMR spectroscopy is the simplicity and informational richness of the 1D ^{19}F NMR spectra, and therefore the utility of ^{19}F NMR spectroscopy do not require NMR specialization in the same extent that is needed for studying complex nucleic acid structures or structural equilibria by ^{1}H NMR spectroscopy. Spectrophotometric methods, such as UV[73-75] and CD spectroscopy[76-78] as well as fluorescence based methods[79,80], are inherently more sensitive methods than NMR spectroscopy. Fluorescence based methods are extensively used to investigate nucleic acids and a vast amount of applications have been described.[79,80] Because nucleic acids are inherently non-fluorescent, it is necessary to introduce a fluorescent modification into a nucleic acid, and currently, a number of fluorescent analogues are available. In principle, fluorescence based methods and ^{19}F NMR spectroscopy share some common general aspects, because prerequisite for both

of these techniques is the incorporation of an unnatural modification into a target nucleic acid (or its ligand). For instance, challenges in development and synthesis of appropriate building blocks and introducing them efficiently into oligonucleotides can limit the utility of these techniques. In addition, the introduced modification can affect significantly on nucleic acid structure and/or stability. In general, the utility of a labelled building block depends not only its general properties, but also its behaviour in particular structure and its ability to reflect the process of interest. UV[73-75] and CD[76-78] studies are also commonly used to investigate nucleic acids. These methods do not require labelling of the nucleic acid of interest, but analysing intermediate structures, coexisting structural species or large nucleic acids is often difficult with these methods. In particular, the strength of ^{19}F NMR is that it provides direct quantitative information of coexisting structural species and it can be readily used to analyse changes in structural equilibria induced by environmental changes, such as ions, temperature and pH, or interaction partners, such as small molecules, nucleic acids and proteins. This valuable characteristic feature of ^{19}F NMR was also demonstrated in this thesis by investigating DNA and RNA triplexes, RNA invasion process and RNA G-quadruplex-hairpin equilibria with novel fluorine-labelled nucleoside analogues **1** and **3**-**6**.

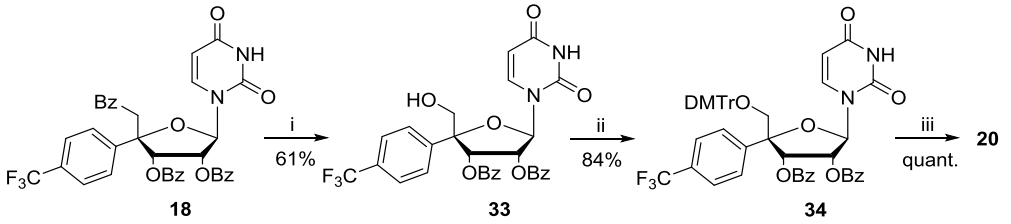
5. EXPERIMENTAL

5.1 General

The synthesis and characterization of the compounds are described in the original publications. The synthesized novel compounds were characterized by ^{1}H NMR, ^{13}C NMR, ^{31}P NMR ^{19}F NMR and MS when applicable.

*Synthesis of **20** by 5´-debenzoylation strategy*

The bottleneck of the synthesis of **22** was the tritylation step and the alternative synthetic routes are shortly summarized here. However, as stated in Section 3.1.1.1 the direct tritylation of **19** was practically straightforward approach, and importantly, the bis- and monotritylated sideproducts were readily converted back to the starting material by dicloroacetic acid treatment. The selective removal of the 5´-*O*-Bz protection was most successful with [*t*BuSnOHCl]$_2$[261,360] (Scheme 4). However, it has been reported that the selective 5´-debenzoylation with this tin catalyst is sensitive to the amount of catalyst, reaction time and temperature, and reaction can provide a complex mixture of products.[261] This was also the case with **18** and several test reactions were conducted varying the amount of catalyst, reaction time and temperature to obtain **33** in a moderate yield (61 %). The best result was obtained in dry MeOH with 30 mol% of catalyst at 66 °C for 68 h, but the obtained mixture of compounds complicated the purification. Tritylation of **33** with excess of DMTrCl for 120 h gave **34** in an 84% yield (yield from **18** 51%).

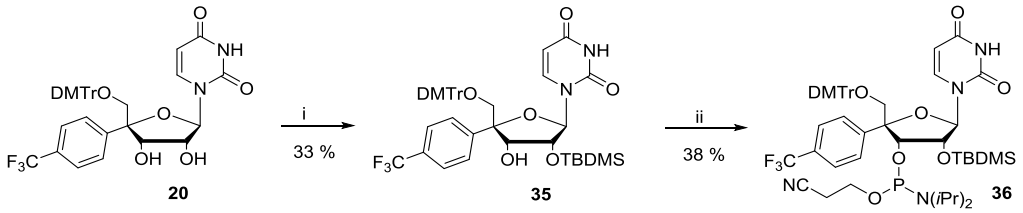


Scheme 4. (i) [*t*BuSnOHCl]$_2$, MeOH; (ii) DMTrCl, Py; (iii) NaOMe, MeOH.

Although NaOMe is unspesific reagent for selective debenzoylation, it was also tested for cleaving 5´-*O*-Bz group selectively. **33** was noticed to accumulate in some extent during debenzoylation reaction. Treatment of **18** with 0.1 ekv of NaOMe solution for 10-30 min provided a mixture of products, as expected. The obtained crude was then tritylated and **34** was obtained in 48% yield. Although the yield is comparable to that of the tin method, the NaOMe step was hard to control. 5´-Debenzoylation was also tested using 1 % iodine-methanol solution[361], but this attempt was unsuccessful.

Synthesis of 2´-O-TBDMS-protected 4´-C-[4-(trifluoromethyl)phenyl]uridine phosphoramidite

Synthesis of 2´-*O*-TBDMS-protected 4´-*C*-[4-(trifluoromethyl)phenyl]uridine phosphoramidite (**36**) from **20** is depicted in Scheme 5. First 2´-OH group was silylated with TBDMSCl in the presence of $AgNO_3$[362] (**35** in 33% yield). The reaction proceeded slowly and a mixture of products was obtained. **35** was then phosphitylated, and also this reaction was difficult due to the steric hindrance of 4´-*C*-[4-(trifluoromethyl)phenyl] and 2´-*O*-TBDMS groups. High excess of 2-cyanoethyl *N*,*N*-diisopropylchlorophosphoramidite (6 ekv) and prolonged reaction time at elevated temperature (48 h at 55°C) were required to obtain **36** in 38% yield. This building block coupled inefficiently due to the bulky TBDMS group. The coupling efficiency improved significantly, from 20 to 90%, when the 2´-OH group was protected with sterically less demanding TOM group (**22**, Section 3.1.2). In addition, the phosphitylation reaction was easier with TOM-protected compound (Scheme 2). Smaller excess of 2-cyanoethyl *N*,*N*-diisopropylchlorophosphoramidite (3.5 ekv) could be used and the reaction proceeded at ambient temperature (24 h) in a good yield (**22**, 86%).



Scheme 5. (i) TBDMSCl, Py, $AgNO_3$, THF; (ii) 2-cyanoethyl *N*,*N*-diisopropylchlorophosphoramidite, Et_3N, DCM.

5.2 Oligonucleotide synthesis

The detailed synthetic methods of the oligonucleotides are described in the original publications. Oligonucleotides were synthesized on an Applied Biosystems 3400 DNA/RNA synthesizer using standard protocols. The manual coupling of phosphoramidites **12**, **22** and **29d** is described in Section 3.2.1. The oligonucleotides were released from the support by standard ammonolysis (oligonucleotides in concentrated ammonia at 55 ºC overnight, oligoribonucleotides in a mixture of concentrated ammonia and ethanol 3:1, *v/v*, 3.5 h at 55 ºC and overnight at rt). The silyl protections were removed by a mixture of triethylamine trihydrofluoride (75 µL), triethylamine (60 µL) and DMSO (115 µL) (for 2.5 hours at 65 ºC). Then, 0.1 mol L^{-1} NaOAc (10 mL) was added to the mixture, and it was loaded into a cartridge (Poly Pak II, Glen Research). 0.1 mol L^{-1} aqueous triethylammonium acetate (6.0 mL, pH = 7.0)

was eluted though the cartridges to remove traces of reagents and protecting groups. The crude oligoribonucleotides were released by elution with 60 % aqueous acetonitrile. Oligonucleotides were purified with RP-HPLC using a semipreparative column, a Phenomenex Oligo-RP C18 (250×10mm, 5 µm), with a gradient elution, typically from 0 to 40-60 % acetonitrile in 0.1 mol L^{-1} aqueous triethylammonium acetate in 25 min at ambient or elevated temperature, flow rate 3.0 ml min^{-1}, and UV detection at 260 nm. The homogeneity of the purified oligonucleotides was analysed by an analytical RP HPLC column, a Thermo ODS Hypercil C18 (250 × 5 mm, 5µm), flow rate 1.0 ml min^{-1}, detection at 260 nm, and elution conditions the same as those of the semipreparative column. The authenticity of the oligonucleotides was verified by ESI-TOF MS.

5.3 UV melting studies and CD spectroscopy experiments

The detailed measurement conditions for the UV melting temperature and CD spectroscopy studies are described in the Results and Discussion section and in the original publications. The melting curves (absorbance versus temperature) were measured at 260 nm or 295 nm on a Perkin-Elmer Lambda 35 UV-vis spectrometer equipped with a multiple cell holder and a Peltier temperature controller. T_m values were determined as the maximum of the first derivative of the melting curve. The CD spectra were measured using an Applied Photophysics Chirascan spectropolarimeter. For both equipment, an internal thermometer was additionally used to verify the validity of the target temperature.

5.4 ^{19}F NMR and ^{1}H NMR experiments of the oligonucleotides

The detailed measurement conditions for ^{1}H NMR and ^{19}F NMR measurements are described in the Results and Discussion section and in the original publications. ^{19}F NMR spectra were recorded at a frequency of 470.6 MHz on a Bruker Avance 500 MHz spectrophotometer. Typical experimental parameters were as follows: ^{19}F excitation pulse 4.0 µs, acquisition time 1.17 s, prescan delay 6.0 µs, relaxation delay 0.8 s and usual numbers of scans (ns) 1024, 2048 and 20480. The parameters were optimized to obtain the signals with the longest relaxation rate. A macro command was used for the automatic temperature ramps using a 20 min equilibration time for each temperature. The sample temperatures were calibrated using known shifts of ethylene glycol at different temperatures. The ^{1}H NMR spectra of the imino regions at 25°C were recorded at a frequency of 500.1 MHz on a Bruker Avance 500 MHz spectrophotometer. Water was suppressed by an excitation sculpting technique.

6. REFERENCES

(1) Gerig, J. T. *Prog. Nucl. Magn. Reson. Spectrosc.* **1994**, *26*, 293-370.
(2) Rastinejad, F.; Evilia, C.; Lu, P. *Methods Enzymol.***1995**, *261*, 560-575.
(3) Danielson, M. A.; Falke, J. J. *Annu. Rev. Biophys. Biomol. Struct.* **1996**, *25*, 163-195.
(4) Dalvit, C. *Prog. Nucl. Magn. Reson. Spectrosc.* **2007**, *51*, 243-271.
(5) Cobb, S. L.; Murphy, C. D. *J. Fluorine Chem.* **2009**, *130*, 132-143.
(6) Kitevski-LeBlanc, J. L.; Prosser, R. S. *Prog. Nucl. Magn. Reson. Spectrosc.* **2012**, *62*, 1-33.
(7) Yu, J.; Hallac, R. R.; Chiguru, S.; Mason, R. P Prog. Nucl. Magn. Reson. Spectrosc. **2013**, *70*, 25-49.
(8) Chen, H.; Viel, S.; Ziarelli, F.; Peng, L. *Chem. Soc. Rev.* **2013**, *42*, 7971-7982.
(9) Marsh, E. N. G.; Suzuki, Y. *ACS Chem. Biol.* **2014**, *9*, 1242-1250.
(10) Tirotta, I.; Dichiarante, V.; Pigliacelli, C.; Cavallo, G.; Terraneo, G.; Bombelli, F. B.; Metrangolo, P.; Resnati, G. *Chem. Rev.* **2015**, *115*, 1106-1129.
(11) Gossert, A. D.; Jahnke, W. *Prog. Nucl. Magn. Reson. Spectrosc.* **2016**, *97*, 82-125.
(12) Guo, F.; Li, Q.; Zhou, C. *Org. Biomol. Chem.* **2017**, *15*, 9552-9565.
(13) Hull, W.; Sykes, B. *Biochemistry* **1974**, *13*, 3431-3437.
(14) Spotswood, T.; Evans, J.; Richards, J. *J. Am. Chem. Soc.* **1967**, *89*, 5052-5054.
(15) Marshall, A. G.; Smith, J. L. *J. Am. Chem. Soc.* **1977**, *99*, 635-636.
(16) Horowitz, J.; Ofengand, J.; Daniel, W. E.; Cohn, M. *J. Biol. Chem.* **1977**, *252*, 4418-4420.
(17) Kreutz, C.; Kählig, H.; Konrat, R.; Micura, R. *Angew. Chem. Int. Ed.* **2006**, *45*, 3450-3453.
(18) Sakamoto, T.; Hayakawa, H.; Fujimoto, K. *Chem. Lett.* **2011**, *40*, 720-721.
(19) Ishizuka, T.; Yamashita, A.; Asada, Y.; Xu, Y. *Org. Biomol. Chem.* **2017**, *2*, 8843-8848.
(20) Olejniczak, M.; Gdaniec, Z.; Fischer, A.; Grabarkiewicz, T.; Bielecki, Ł; Adamiak, R. W. *Nucleic Acids Res.* **2002**, *30*, 4241-4249.
(21) Scott, L. G.; Geierstanger, B. H.; Williamson, J. R.; Hennig, M. *J. Am. Chem. Soc.* **2004**, *126*, 11776-11777.
(22) Olsen, G. L.; Edwards, T. E.; Deka, P.; Varani, G.; Sigurdsson, S. T.; Drobny, G. P. *Nucleic Acids Res.* **2005**, *33*, 3447-3454.
(23) Hennig, M.; Scott, L. G.; Sperling, E.; Bermel, W.; Williamson, J. R. *J. Am. Chem. Soc.* **2007**, *129*, 14911-14921.
(24) Kiviniemi, A.; Virta, P. *J. Am. Chem. Soc.* **2010**, *132*, 8560-8562.
(25) Scott, L. G.; Hennig, M. *Methods Enzymol.* **2016**, *566*, 59-87.
(26) Rieder, U.; Kreutz, C.; Micura, R. *PNAS* **2010**, *107*, 10804-10809.
(27) Haller, A.; Rieder, U.; Aigner, M.; Blanchard, S. C.; Micura, R. *Nat. Chem. Biol.* **2011**, *7*, 393-400.
(28) Santner, T.; Rieder, U.; Kreutz, C.; Micura, R. *J. Am. Chem. Soc.* **2012**, *134*, 11928-11931.
(29) Sochor, F.; Silvers, R.; Müller, D.; Richter, C.; Fürtig, B.; Schwalbe, H. *J. Biomol. NMR* **2015**, *64*, 63-74.
(30) Hammann, C.; Norman, D. G.; Lilley, D. M. J. *PNAS* **2001**, *98*, 5503-5508.
(31) Zhao, C.; Bachu, R.; Popović, M.; Devany, M.; Brenowitz, M.; Schlatterer, J. C.; Greenbaum, N. L. *RNA* **2013**, *19*, 561-573.
(32) Zhao, C.; Anklin, C.; Greenbaum, N. L. *Methods Enzymol.* **2014**, *549*, 267-285.
(33) Zhou, L.; Rajabzadeh, M.; Traficante, D. D.; Cho, B. S. P. *J. Am. Chem. Soc.* **1997**, *119*, 5384-5389.
(34) Cho, B. P.; Zhou, L. *Biochemistry* **1999**, *38*, 7572-7583.
(35) Tanabe, K.; Sugiura, M.; Nishimoto, S. *Bioorg. Med. Chem.* **2010**, *18*, 6690-6694.
(36) Tähtinen, V.; Granqvist, L.; Virta, P. *Bioorg. Med. Chem.* **2015**, *23*, 4472-4480.

(37) Bao, H.; Ishizuka, T.; Sakamoto, T.; Fujimoto, K.; Uechi, T.; Kenmochi, N.; Xu, Y. *Nucleic Acids Res.* **2017**, *45*, 5501-5511.
(38) Bao, H.; Ishizuka, T.; Iwanami, A.; Oyoshi, T.; Xu, Y. *ChemistrySelect* **2017**, *2*, 4170-4175.
(39) Ishizuka, T.; Zhao, P.; Bao, H.; Xu, Y. *Analyst* **2017**, *142*, 4083-4088.
(40) Garavís, M.; López-Méndez, B.; Somoza, A.; Oyarzabal, J.; Dalvit, C.; Villasante, A.; Campos-Olivas, R.; González, C. *ACS Chem. Biol.* **2014**, *9*, 1559-1566.
(41) Lee, Y.; Zeng, H.; Ruedisser, S.; Gossert, A. D.; Hilty, C. *J. Am. Chem. Soc.* **2012**, *134*, 17448-17451.
(42) Meier, S.; Jensen, P. R.; Karlsson, M.; Lerche, M. H. *Sensors* **2014**, *14*, 1576-1597.
(43) Liang, T.; Neumann, C. N.; Ritter, T. *Angew. Chem. Int. Ed.* **2013**, *52*, 8214-8264.
(44) Gillis, E. P.; Eastman, K. J.; Hill, M. D.; Donnelly, D. J.; Meanwell, N. A. *J. Med. Chem.* **2015**, *58*, 8315-8359.
(45) Qiu, X.; Xu, X.; Qing, F. *Tetrahedron* **2010**, *66*, 789-843.
(46) Milligan, J.; Groebe, D.; Witherell, G.; Uhlenbeck, O. *Nucleic Acids Res.* **1987**, *15*, 8783-8798.
(47) Reese, C. B. *Org. Biomol. Chem.* **2005**, *3*, 3851-3868.
(48) Paredes, E.; Evans, M.; Das, S. R. *Methods* **2011**, *54*, 251-259.
(49) Caruthers, M. H. *Biochem. Soc. Trans.* **2011**, *39*, 575-580.
(50) Kosuri, S.; Church, G. M. *Nat. Methods* **2014**, *11*, 499-507.
(51) Hughes, R. A.; Ellington, A. D. *Cold Spring Harbor Perspect. Biol.* **2017**, *9*, a023812.
(52) Lombès, T.; Moumné, R.; Larue, V.; Prost, E.; Catala, M.; Lecourt, T.; Dardel, F.; Micouin, L.; Tisné, C. *Angew. Chem. Int. Ed.* **2012**, *51*, 9530-9534.
(53) Sakamoto, T.; Shimizu, Y.; Sasaki, J.; Hayakawa, H.; Fujimoto, K. *Bioorg. Med. Chem. Lett.* **2011**, *21*, 303-306.
(54) Kieger, A.; Wiester, M. J.; Procissi, D.; Parrish, T. B.; Mirkin, C. A.; Thaxton, C. S. *Small* **2011**, *7*, 1977-1981.
(55) Tanabe, K.; Tsuda, T.; Ito, T.; Nishimoto, S. *Chemistry* **2013**, *19*, 15133-15140.
(56) Nakamura, S.; Fujimoto, K. *Chem. Commun.* **2015**, *51*, 11765-11768.
(57) Sicilia, G.; Davis, A. L.; Spain, S. G.; Magnusson, J. P.; Boase, N. R. B.; Thurecht, K. J.; Alexander, C. *Polym. Chem.* **2016**, *7*, 2180-2191.
(58) Bhuma, N.; Tähtinen, V.; Virta, P. *Eur. J. Org. Chem.* **2017**.
(59) Pettinari, C.; Rafaiani, G. In *19 F NMR, Applications, Solution State;* Encyclopedia of Spectroscopy and Spectrometry; Elsevier Ltd: **1999**; pp 489–498.
(60) Gerig, J. T. *Fluorine NMR.* http://www.biophysics.org/img/jtg2001-2.pdf
(61) Olsen, G. L.; Louie, E. A.; Drobny, G. P.; Sigurdsson, S. T. *Nucleic Acids Res.* **2003**, *31*, 5084-5089.
(62) Mehta, A. K.; Shayo, Y.; Vankayalapati, H.; Hurley, L. H.; Schaefer, J. *Biochemistry* **2004**, *43*, 11953-11958.
(63) Louie, E. A.; Chirakul, P.; Raghunathan, V.; Sigurdsson, S. T.; Drobny, G. P. *J. Magn. Reson.* **2006**, *178*, 11-24.
(64) Huang, W.; Varani, G.; Drobny, G. P. *J. Am. Chem. Soc.* **2010**, *132*, 17643-17645.
(65) Huang, W.; Varani, G.; Drobny, G. P. *J. Biomol. NMR* **2011**, *51*, 347-356.
(66) Fürtig, B.; Richter, C.; Wöhnert, J.; Schwalbe, H. *ChemBioChem* **2003**, *4*, 936-962.
(67) Latham, M. P.; Brown, D. J.; McCallum, S. A.; Pardi, A. *ChemBioChem* **2005**, *6*, 1492-1505.
(68) Wachowius, F.; Höbartner, C. *ChemBioChem* **2010**, *11*, 469-480.
(69) Lu, K.; Miyazaki, Y.; Summers, M. F. *J. Biomol. NMR* **2010**, *46*, 113-125.
(70) Bothe, J. R.; Nikolova, E. N.; Eichhorn, C. D.; Chugh, J.; Hansen, A. L.; Al-Hashimi, H. M. *Nat. Methods* **2011**, *8*, 919-931.
(71) Adrian, M.; Heddi, B.; Phan, A. T. *Methods* **2012**, *57*, 11-24.
(72) Barnwal, R. P.; Yang, F.; Varani, G. *Arch. Biochem. Biophys.* **2017**, *628*, 42-56.

(73) Puglisi, J. D.; Tinoco, I. *Methods Enzymol.* **1989**, *180*, 304-325.
(74) Mergny, J.; Phan, A.; Lacroix, L. *FEBS Lett.* **1998**, *435*, 74-78.
(75) Mergny, J. L.; Lacroix, L. *Oligonucleotides* **2003**, *13*, 515-537.
(76) Gray, D. M.; Hung, S. H.; Johnson, K. H. *Methods Enzymol.* **1995**, *246*, 19-34.
(77) Kypr, J.; Kejnovská, I.; Renčiuk, D.; Vorlíčková, M. *Nucleic Acids Res.* **2009**, *37*, 1713-1725.
(78) Vorlíčková, M.; Kejnovská, I.; Bednářová, K.; Renčiuk, D.; Kypr, J. *Chirality* **2012**, *24*, 691-698.
(79) Xu, W.; Chan, K. M.; Kool, E. T. *Nat. Chem.* **2017**, *9*, 1043-1055.
(80) Rist, M. J.; Marino, J. P. *Curr. Org. Chem.* **2002**, *6*, 775-793.
(81) Metzler, W.; Leighton, P.; Lu, P. *J. Magn. Reson.* **1988**, *76*, 534-539.
(82) Luy, B.; Marino, J. P. *J Biomol. NMR* **2001**, *20*, 39-47.
(83) Hennig, M.; Munzarová, M. L.; Bermel, W.; Scott, L. G.; Sklenář, V.; Williamson, J. R. *J. Am. Chem. Soc.* **2006**, *128*, 5851-5858.
(84) Yi-Brunozzi, H. Y.; Brinson, R. G.; Brabazon, D. M.; Lener, D.; Le Grice, Stuart F J; Marino, J. P. *Chem. Biol.* **2008**, *15*, 254-262.
(85) Brinson, R. G.; Miller, J. T.; Kahn, J. D.; Le Grice, Stuart F J; Marino, J. P. *Methods Enzymol.* **2016**, *566*, 89-110.
(86) Chen, J. L.; VanEtten, D. M.; Fountain, M. A.; Yildirim, I.; Disney, M. D. *Biochemistry* **2017**, *56*, 3463-3474.
(87) Wemmer, D. E.; Williams, P. G. *Methods Enzymol.* **1994**, *239*, 739-767.
(88) Kleckner, I. R.; Foster, M. P. *Biochim. Biophys. Acta, Proteins Proteomics* **2011**, *1814*, 942-968.
(89) Kiviniemi, A.; Virta, P. *Bioconjugate Chem.* **2011**, *22*, 1559-1566.
(90) Arnold, J. R. P.; Fisher, J. *J. Biomol. Struct. Dyn.* **2000**, *17*, 843-856.
(91) Suydam, I. T.; Strobel, S. A. *J. Am. Chem. Soc.* **2008**, *130*, 13639-13648.
(92) Parker, J. B.; Stivers, J. T. *Biochemistry* **2011**, *50*, 612-617.
(93) Graber, D.; Moroder, H.; Micura, R. *J. Am. Chem. Soc.* **2008**, *130*, 17230-17231.
(94) Jud, L.; Košutić, M.; Schwarz, V.; Hartl, M.; Kreutz, C.; Bister, K.; Micura, R. *Chem. Eur. J.* **2015**, *21*, 10400-10407.
(95) Tähtinen, V.; Granqvist, L.; Murtola, M.; Strömberg, R.; Virta, P. *Chem. Eur. J.* **2017**, *23*, 7113-7124.
(96) Meneni, S. R.; Shell, S. M.; Gao, L.; Jurecka, P.; Lee, W.; Sponer, J.; Zou, Y.; Chiarelli, M. P.; Cho, B. P. *Biochemistry* **2007**, *46*, 11263-11278.
(97) Zhao, C.; Devany, M.; Greenbaum, N. L. *Biochem. Biophys. Res. Commun.* **2014**, *453*, 692-695.
(98) Roy, S.; Caruthers, M. *Molecules* **2013**, *18*, 14268-14284.
(99) Chu, W.; Horowitz, J. *Nucleic Acids Res.* **1989**, *17*, 7241-7252.
(100) Riedl, J.; Pohl, R.; Rulíšek, L.; Hocek, M. *J. Org. Chem.* **2012**, *77*, 1026-1044.
(101) Olszewska, A.; Pohl, R.; Hocek, M. *J. Org. Chem.* **2017**, *82*, 11431-11439.
(102) Baranowski, M. R.; Nowicka, A.; Rydzik, A. M.; Warminski, M.; Kasprzyk, R.; Wojtczak, B. A.; Wojcik, J.; Claridge, T. D. W.; Kowalska, J.; Jemielity, J. *J. Org. Chem.* **2015**, *80*, 3982-3997.
(103) Longley, D. B.; Harkin, D. P.; Johnston, P. G. *Nat. Rev. Cancer* **2003**, *3*, 330-338.
(104) De Clercq, E. *Biochem. Pharmacol.* **2013**, *85*, 727-744.
(105) Deleavey, G. F.; Damha, M. J. *Chem. Biol.* **2012**, *19*, 937-954.
(106) Barhate, N. B.; Barhate, R. N.; Cekan, P.; Drobny, G.; Sigurdsson, S. T. *Org. Lett.* **2008**, *10*, 2745-2747.
(107) Fauster, K.; Kreutz, C.; Micura, R. *Angew. Chem. Int. Ed.* **2012**, *51*, 13080-13084.
(108) Luyten, I.; Herdewijn, P. *Eur. J. Med. Chem.* **1998**, *33*, 515-576.
(109) Rastinejad, F.; Lu, P. *J. Mol. Biol.* **1993**, *232*, 105-122.

(110) Kealey, J. T.; Santi, D. V. *Biochemistry* **1995**, *34*, 2441-2446.
(111) Puffer, B.; Kreutz, C.; Rieder, U.; Ebert, M.; Konrat, R.; Micura, R. *Nucleic Acids Res.* **2009**, *37*, 7728-7740.
(112) Metzler, W. J.; Arndt, K.; Tecza, E.; Wasilewski, J.; Lu, P. *Biochemistry* **1985**, *24*, 1418-1424.
(113) Metzler, W.; Lu, P. *J. Mol. Biol.* **1989**, *205*, 149-164.
(114) Rastinejad, F.; Artz, P.; Lu, P. *J. Mol. Biol.* **1993**, *233*, 389-399.
(115) Kanyo, J. E.; Duhamel, J.; Lu, P. Z. *Nucleic Acids Res.* **1996**, *24*, 4015-4022.
(116) Evilia, C.; Zhang, X.; Kanyo, J.; Lu, P. *Nucleosides Nucleotides* **1997**, *16*, 1809-1820.
(117) Kwon, K.; Jiang, Y. L.; Song, F.; Stivers, J. T. *J. Biol. Chem.* **2002**, *277*, 353-358.
(118) Klimašauskas, S.; Szyperski, T.; Serva, S.; Wüthrich, K. *EMBO J.* **1998**, *17*, 317-324.
(119) Fischer, A.; Gdaniec, Z.; Biala, E.; Łożyński, M.; Milecki, J.; Adamiak, R. W. *Nucleosides Nucleotides* **1996**, *15*, 477-488.
(120) Sowers, L. C.; Eritja, R.; Kaplan, B. E.; Goodman, M. F.; Fazakerley, G. V. *J. Biol. Chem.* **1987**, *262*, 15436-15442.
(121) Stolarski, R.; Egan, W.; James, T. L. *Biochemistry* **1992**, *31*, 7027-7042.
(122) Sahasrabudhe, P. V.; Pon, R. T.; Gmeiner, W. H. *Nucleic Acids Res.* **1995**, *23*, 3916-3921.
(123) Sahasrabudhe, P. V.; Pon, R. T.; Gmeiner, W. H. *Biochemistry* **1996**, *35*, 13597-13608.
(124) Sahasrabudhe, P. V.; Gmeiner, W. H. *Biochemistry* **1997**, *36*, 5981-5991.
(125) O'Hagan, D. *Chem. Soc. Rev.* **2008**, *37*, 308-319.
(126) Oyelere, A. K.; Strobel, S. A. *J. Am. Chem. Soc.* **2000**, *122*, 10259-10267.
(127) Guéron, M.; Leroy, J. *Methods Enzymol.* 1995, 261, 383-413.
(128) Gmeiner, W. H.; Pon, R. T.; Lown, J. W. *J. Org. Chem.* **1991**, *56*, 3602-3608.
(129) Markley, J. C.; Chirakul, P.; Sologub, D.; Sigurdsson, S. T. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 2453-2455.
(130) Nakamura, S.; Yang, H.; Hirata, C.; Kersaudy, F.; Fujimoto, K. *Org. Biomol. Chem.* **2017**, *15*, 5109-5111.
(131) Holzberger, B.; Marx, A. *Bioorg. Med. Chem.* **2009**, *17*, 3653-3658.
(132) Suzuki, N.; Fukushima, M. *Nucleosides Nucleotides Nucleic Acids* **2010**, *29*, 896-904.
(133) Lahoud, G.; Timoshchuk, V.; Lebedev, A.; de Vega, M.; Salas, M.; Arar, K.; Hou, Y.; Gamper, H. *Nucleic Acids Res.* **2008**, *36*, 3409-3419.
(134) Jiang, Y. L.; McDowell, L. M.; Poliks, B.; Studelska, D. R.; Cao, C.; Potter, G. S.; Schaefer, J.; Song, F.; Stivers, J. T. *Biochemistry* **2004**, *43*, 15429-15438.
(135) Kool, E. T.; Sintim, H. O. *Chem. Commun. (Camb.)* **2006**, 3665-3675.
(136) Khakshoor, O.; Wheeler, S. E.; Houk, K. N.; Kool, E. T. *J. Am. Chem. Soc.* **2012**, *134*, 3154-3163.
(137) Egli, M. *Acc. Chem. Res.* **2012**, *45*, 1237-1246.
(138) Musumeci, D.; Irace, C.; Santamaria, R.; Montesarchio, D. *Med. Chem. Commun.* **2013**, *4*, 1405-1410.
(139) Peng, X.; Li, H.; Seela, F. *Nucleic Acids Res.* **2006**, *34*, 5987-6000.
(140) Seela, F.; Xu, K. *Org. Biomol. Chem.* **2008**, *6*, 3552-3560.
(141) Véliz, E. A.; Stephens, O. M.; Beal, P. A. *Org. Lett.* **2001**, *3*, 2969-2972.
(142) Kreutz, C.; Kählig, H.; Konrat, R.; Micura, R. *J. Am. Chem. Soc.* **2005**, *127*, 11558-11559.
(143) Košutić, M.; Jud, L.; Da Veiga, C.; Frener, M.; Fauster, K.; Kreutz, C.; Ennifar, E.; Micura, R. *J. Am. Chem. Soc.* **2014**, *136*, 6656-6663.
(144) Assi, H. A.; Harkness, R. W.; Martin-Pintado, N.; Wilds, C. J.; Campos-Olivas, R.; Mittermaier, A. K.; González, C.; Damha, M. J. *Nucleic Acids Res.* **2016**, *44*, 4998-5009.
(145) Evich, M.; Spring-Connell, A. M.; Germann, M. W. *Heterocycl. Commun.* **2017**, *23*, 155-165.
(146) Neidle, S. *Principles of nucleic acid structure*; Academic Press: New York, **2008**.

(147) Kawasaki, A. M.; Casper, M. D.; Freier, S. M.; Lesnik, E. A.; Zounes, M. C.; Cummins, L. L.; Gonzalez, C.; Cook, P. D. *J. Med. Chem.* **1993**, *36*, 831-841.
(148) Reif, B.; Wittmann, V.; Schwalbe, H.; Griesinger, C.; Worner, K.; JahnHofmann, K.; Engels, J. W.; Bermel, W. *Helv. Chim. Acta* **1997**, *80*, 1952-1971.
(149) Wilds, C. J.; Damha, M. J. *Bioconjugate chem.* **1999**, *10*, 299-305.
(150) Pallan, P. S.; Greene, E. M.; Jicman, P. A.; Pandey, R. K.; Manoharan, M.; Rozners, E.; Egli, M. *Nucleic Acids Res.* **2011**, *39*, 3482-3495.
(151) Hills, D. C.; Cotten, M. L.; Horowitz, J. *Biochemistry* **1983**, *22*, 1113-1122.
(152) Horowitz, J.; Cotten, M. L.; Hardin, C. C.; Gollnick, P. *Biochim. Biophys. Acta* **1983**, *741*, 70-76.
(153) Gollnick, P.; Hardin, C. C.; Horowitz, J. *Nucleic Acids Res.* **1986**, *14*, 4659-4672.
(154) Hardin, C. C.; Gollnick, P.; Kallenbach, N. R.; Cohn, M.; Horowitz, J. *Biochemistry* **1986**, *25*, 5699-5709.
(155) Hardin, C. C.; Horowitz, J. *J. Mol. Biol.* **1987**, *197*, 555-569.
(156) Gollnick, P.; Hardin, C. C.; Horowitz, J. *J. Mol. Biol.* **1987**, *197*, 571-584.
(157) Hardin, C. C.; Gollnick, P.; Horowitz, J. *Biochemistry* **1988**, *27*, 487-495.
(158) Chu, W. C.; Horowitz, J. *FEBS Lett.* **1991**, *295*, 159-162.
(159) Chu, W. C.; Horowitz, J. *Biochemistry* **1991**, *30*, 1655-1663.
(160) Chu, W. C.; Feiz, V.; Derrick, W. B.; Horowitz, J. *J. Mol. Biol.* **1992**, *227*, 1164-1172.
(161) Chu, W. C.; Kintanar, A.; Horowitz, J. *J. Mol. Biol.* **1992**, *227*, 1173-1181.
(162) Chu, W. C.; Liu, J. C. H,; Horowitz, J. *Nucleic Acids Res.* **1997**, *25*, 3944-3949.
(163) Liu, M. S.; Chu, W. C.; Liu, J. C. H.; Horowitz, J. *Nucleic Acids Res.* **1997**, *25*, 4883-4890.
(164) Horowitz, J.; Chu, W. C.; Derrick, W. B.; Liu, J. C. H.; Liu, M. S.; Yue, D. X. *Biochemistry* **1999**, *38*, 7737-7746.
(165) Marshall, A.; Smith, J. *Biochemistry* **1980**, *19*, 5955-5959.
(166) Horowitz, J.; Ching-Nan, O.; Ishaq, M.; Ofengand, J.; Bierbaum, J. *J. Mol. Biol.* **1974**, *88*, 301-312.
(167) Patel, D.; Pardi, A.; Itakura, K. *Science* **1982**, *216*, 581-590.
(168) Nikolova, E. N.; Al-Hashimi, H. M. *RNA* **2010**, *16*, 1687-1691.
(169) Papasaikas, P.; Valcárcel, J. *Trends Biochem. Sci.* **2016**, *41*, 33-45.
(170) Garst, A. D.; Edwards, A. L.; Batey, R. T. *Cold Spring Harbor Perspect. Biol.* **2011**, *3*, a003533.
(171) Serganov, A.; Nudler, E. *Cell* **2013**, *152*, 17-24.
(172) Eichhorn, C. D.; Kang, M.; Feigon, J. *Biochim. Biophys. Acta, Gene Regulatory Mechanisms* **2014**, *1839*, 939-950.
(173) Karn, J.; Stoltzfus, C. M. *Cold Spring Harbor Perspect. Med.* **2012**, *2*, a006916.
(174) Richter, S. N.; Palù, G. *Curr. Med. Chem.* **2006**, *13*, 1305-1315.
(175) Mousseau, G.; Mediouni, S.; Valente, S. T. *Curr. Top. Microbiol. Immunol.* **2015**, *389*, 121-145.
(176) Chen, Q.; Shafer, R. H.; Kuntz, I. D. *Biochemistry* **1997**, *36*, 11402-11407.
(177) Aboul-ela, F.; Karn, J.; Varani, G. *Nucleic Acids Res.* **1996**, *24*, 3974-3981.
(178) Puglisi, J.; Tan, R.; Calnan, B.; Frankel, A.; Williamson, J. *Science* **1992**, *257*, 76-80.
(179) Brodsky, A. S.; Williamson, J. R. *J. Mol. Biol.* **1997**, *267*, 624-639.
(180) Aboul-ela, F.; Karn, J.; Varani, G. *J. Mol. Biol.* **1995**, *253*, 313-332.
(181) Puglisi, J.; Chen, L.; Blanchard, S.; Frankel, A. *Science* **1995**, *270*, 1200-1203.
(182) Sun, J.; Helene, C. *Curr. Opin. Struct. Biol.* **1993**, *3*, 345-356.
(183) Radhakrishnan, I.; Patel, D. *Biochemistry* **1994**, *33*, 11405-11416.
(184) Sojfer, V. N.; Potaman, V. N. *Triple helical nucleic acids*; Springer-Verlag: New York, **1996**.
(185) Singleton, S. F.; Dervan, P. B. *Biochemistry* **1992**, *31*, 10995-11003.

(186) Plum, G.; Pilch, D.; Singleton, S.; Breslauer, K. *Annu. Rev. Biophys. Biomolec. Struct.* **1995**, *24*, 319-350.
(187) Plum, G. E. *Biopolymers* **1997**, *44*, 241-256.
(188) Asensio, J. L.; Lane, A. N.; Dhesi, J.; Bergqvist, S.; Brown, T. *J. Mol. Biol.* **1998**, *275*, 811-822.
(189) Sugimoto, N.; Wu, P.; Hara, H.; Kawamoto, Y. *Biochemistry* **2001**, *40*, 9396-9405.
(190) Mergny, J.; Lacroix, L.; Han, X.; Leroy, J.; Helene, C. *J. Am. Chem. Soc.* **1995**, *117*, 8887-8898.
(191) Lacroix, L.; Mergny, J.; Leroy, J.; Hélène, C. *Biochemistry* **1996**, *35*, 8715-8722.
(192) Burge, S.; Parkinson, G. N.; Hazel, P.; Todd, A. K.; Neidle, S. *Nucleic Acids Res.* **2006**, *34*, 5402-5415.
(193) Collie, G. W.; Parkinson, G. N. *Chem. Soc. Rev.* **2011**, *40*, 5867-5892.
(194) Malgowska, M.; Czajczynska, K.; Gudanis, D.; Tworak, A.; Gdaniec, Z. *Acta Biochim. Pol.* **2016**, *63*, 609-621.
(195) Simone, R.; Fratta, P.; Neidle, S.; Parkinson, G. N.; Isaacs, A. M. *FEBS Lett.* **2015**, *589*, 1653-1668.
(196) Agarwala, P.; Pandey, S.; Maiti, S. *Org. Biomol. Chem.* **2015**, *13*, 5570-5585.
(197) Neidle, S. *J. Med. Chem.* **2016**, *59*, 5987-6011.
(198) Fay, M. M.; Lyons, S. M.; Ivanov, P. *J. Mol. Biol.* **2017**, *429*, 2127-2147.
(199) Cammas, A.; Millevoi, S. *Nucleic Acids Res.* **2017**, *45*, 1584-1595.
(200) Hansel-Hertsch, R.; Di Antonio, M.; Balasubramanian, S. *Nat. Rev. Mol. Cell Biol.* **2017**, *18*, 279-284.
(201) Kwok, C. K.; Merrick, C. J. *Trends Biotechnol.* **2017**, *35*, 997-1013.
(202) O'Gara, M.; Klimasauskas, S.; Roberts, R. J.; Cheng, X. *J. Mol. Biol.* **1996**, *261*, 634-645.
(203) Kastenhuber, E. R.; Lowe, S. W. *Cell* **2017**, *170*, 1062-1078.
(204) Song, L.; Teng, Q.; Phillips, R. S.; Brewer, J. M.; Summers, A. O. *J. Mol. Biol.* **2007**, *371*, 79-92.
(205) Holzberger, B.; Rubini, M.; Möller, H.; Marx, A. *Angew. Chem. Int. Ed.* **2010**, *49*, 1324-1327.
(206) Luch, A. *Nat. Rev. Cancer* **2005**, *5*, 113-125.
(207) Kovacic, P.; Somanathan, R. *J. Appl. Toxicol.* **2014**, *34*, 810-824.
(208) Meneni, S.; Liang, F.; Cho, B. P. *J. Mol. Biol.* **2007**, *366*, 1387-1400.
(209) Meneni, S.; Shell, S. M.; Zou, Y.; Cho, B. P. *Chem. Res. Toxicol.* **2007**, *20*, 6-10.
(210) Jain, N.; Li, Y.; Zhang, L.; Meneni, S. R.; Cho, B. P. *Biochemistry* **2007**, *46*, 13310-13321.
(211) Jain, N.; Meneni, S.; Jain, V.; Cho, B. P. *Nucleic Acids Res.* **2009**, *37*, 1628-1637.
(212) Patnaik, S.; Cho, B. P. *Chem. Res. Toxicol.* **2010**, *23*, 1650-1652.
(213) Liang, F.; Cho, B. P. *Chem. Res. Toxicol.* **2011**, *24*, 597-605.
(214) Vaidyanathan, V. G.; Cho, B. P. *Biochemistry* **2012**, *51*, 1983-1995.
(215) Jain, V.; Hilton, B.; Patnaik, S.; Zou, Y.; Chiarelli, M. P.; Cho, B. P. *Nucleic Acids Res.* **2012**, *40*, 3939-3951.
(216) Vaidyanathan, V. G.; Liang, F.; Beard, W. A.; Shock, D. D.; Wilson, S. H.; Cho, B. P. *J. Biol. Chem.* **2013**, *288*, 23573-23585.
(217) Jain, V.; Hilton, B.; Lin, B.; Patnaik, S.; Liang, F.; Darian, E.; Zou, Y.; MacKerell, A. D.; Cho, B. P. *Nucleic Acids Res.* **2013**, *41*, 869-880.
(218) Sandineni, A.; Lin, B.; MacKerell, A. D.; Cho, B. P. *Chem. Res. Toxicol.* **2013**, *26*, 937-951.
(219) Jain, V.; Hilton, B.; Lin, B.; Jain, A.; MacKerell, A. D.; Zou, Y.; Cho, B. P. *Chem. Res. Toxicol.* **2013**, *26*, 1251-1262.
(220) Jain, V.; Vaidyanathan, V. G.; Patnaik, S.; Gopal, S.; Cho, B. P. *Biochemistry* **2014**, *53*, 4059-4071.
(221) Xu, L.; Cho, B. P. *Chem. Res. Toxicol.* **2016**, *29*, 213-226.
(222) Nielsen, P. E. *Chem. Biodivers.* **2010**, *7*, 786-804.

(223) Kiviniemi, A.; Murtola, M.; Ingman, P.; Virta, P. *J. Org. Chem.* **2013**, *78*, 5153-5159.
(224) Moumné, R.; Pasco, M.; Prost, E.; Lecourt, T.; Micouin, L.; Tisné, C. *J. Am. Chem. Soc.* **2010**, *132*, 13111-13113.
(225) Sakamoto, T.; Hasegawa, D.; Fujimoto, K. *Chem. Commun.* **2015**, *51*, 8749-8752.
(226) Sakamoto, T.; Hasegawa, D.; Fujimoto, K. *Analyst* **2016**, *141*, 1214-1217.
(227) Bolton, P.; Mirau, P.; Shafer, R.; James, T. *Biopolymers* **1981**, *20*, 435-449.
(228) Patel, D. J.; Gabbay, E. J. *PNAS* **1981**, *78*, 1351-1355.
(229) Mirau, P. A.; Shafer, R. H.; James, T. L.; Bolton, P. H. *Biopolymers* **1982**, *21*, 909-921.
(230) Mirau, P. A.; Shafer, R. H.; James, T. L. *Biochemistry* **1982**, *21*, 615-620.
(231) Shine, N. R.; James, T. L. *Biochemistry* **1985**, *24*, 4333-4341.
(232) Shine, N.; Bubienko, E.; James, T. *Biochemistry* **1985**, *24*, 4341-4345.
(233) Delbarre, A.; Brown, S. C.; James, T. L.; Shafer, R. H. *Biopolymers* **1988**, *27*, 1953-1975.
(234) Hammer, B.; Russell, R.; Warrener, R.; Collins, J. *Eur. J. Biochem.* **1989**, *178*, 683-688.
(235) Hammer, B.; Russell, R.; Warrener, R.; Collins, J. *FEBS Lett.* **1989**, *244*, 227-230.
(236) Hammer, B.; Russell, R.; Warrener, R.; Collins, J. *Eur. J. Biochem.* **1990**, *191*, 307-313.
(237) Barr, K.; Russell, R.; Warrener, R.; Collins, J. *FEBS Lett.* **1993**, *322*, 173-176.
(238) Wang, A.; Cottens, S.; Dervan, P.; Yesinowski, J.; Vandermarel, G.; Vanboom, J. *J. Biomol. Struct. Dyn.* **1989**, *7*, 101-117.
(239) Fielding, L. *Prog. Nucl. Magn. Reson. Spectrosc.* **2007**, *51*, 219-242.
(240) Dalvit, C. *Drug Discovery Today* **2009**, *14*, 1051-1057.
(241) Arntson, K. E.; Pomerantz, W. C. K. *J. Med. Chem.* **2016**, *59*, 5158.
(242) Collie, G. W.; Haider, S. M.; Neidle, S.; Parkinson, G. N. *Nucleic Acids Res.* **2010**, *38*, 5569-5580.
(243) Fourmy, D.; Recht, M. I.; Puglisi, J. D. *J. Mol. Biol.* **1998**, *277*, 347-362.
(244) Abboud, J. L. M.; Foces-Foces, C.; Notario, R.; Trifonov, R. E.; Volovodenko, A. P.; Ostrovskii, V. A.; Alkorta, I.; Elguero, J. *Eur. J. Org. Chem.* **2001**, 3013-3024.
(245) Maag, H.; Schmidt, B.; Rose, S. *Tetrahedron Lett.* **1994**, *35*, 6449-6452.
(246) Fensholdt, J.; Thrane, H.; Wengel, J. *Tetrahedron Lett.* **1995**, *36*, 2535-2538.
(247) Thrane, H.; Fensholdt, J.; Regner, M.; Wengel, J. *Tetrahedron* **1995**, *51*, 10389-10402.
(248) Wang, G. Y.; Seifert, W. E. *Tetrahedron Lett.* **1996**, *37*, 6515-6518.
(249) Ueno, Y.; Nagasawa, Y.; Sugimoto, I.; Kojima, N.; Kanazaki, M.; Shuto, S.; Matsuda, A. *J. Org. Chem.* **1998**, *63*, 1660-1667.
(250) Wengel, J. *Accounts Chem. Res.* **1999**, *32*, 301-310.
(251) Pfundheller, H. M.; Bryld, T.; Olsen, C. E.; Wengel, J. *Helv. Chim. Acta* **2000**, *83*, 128-151.
(252) Raunkjær, M.; Bryld, T.; Wengel, J. *Chem. Commun.* **2003**, 164-165.
(253) Kiviniemi, A.; Virta, P.; Lönnberg, H. *Bioconjugate Chem.* **2008**, *19*, 1726-1734.
(254) Kiviniemi, A.; Virta, P.; Lönnberg, H. *Bioconjugate Chem.* **2010**, *21*, 1890-1901.
(255) Jones, G.; Taniguchi, M.; Tegg, D.; Moffatt, J. *J. Org. Chem.* **1979**, *44*, 1309-1317.
(256) O-Yang, C.; Kurz, W.; Eugui, E.; McRoberts, M.; Verheyden, J.; Kurz, L.; Walker, K. *Tetrahedron Lett.* **1992**, *33*, 41-44.
(257) Wu, T.; Nauwelaerts, K.; Van Aerschot, A.; Froeyen, M.; Lescrinier, E.; Herdewijn, P. *J. Org. Chem.* **2006**, *71*, 5423-5431.
(258) Stepanova, N.; Galishev, V.; Turbanova, E.; Maleev, A.; Potekhin, K.; Kurkutova, E.; Struchkov, Y.; Petrov, A. *Zhurnal Org. Khimii* **1989**, *25*, 1613-1618.
(259) Rostovtsev, V. V.; Green, L. G.; Fokin, V. V.; Sharpless, K. B. *Angew. Chem. Int. Edit.* **2002**, *41*, 2596-2599.
(260) Chen, B.; Jamieson, E. R.; Tullius, T. D. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 3093-3096.
(261) Taylor, S. D.; Mirzaei, F.; Bearne, S. L. *J. Org. Chem.* **2008**, *73*, 1403-1412.
(262) Semenyuk, A.; Földesi, A.; Johansson, T.; Estmer-Nilsson, C.; Blomgren, P.; Brännvall, M.; Kirsebom, L. A.; Kwiatkowski, M. *J. Am. Chem. Soc.* **2006**, *128*, 12356-12357.

(263) Efimov, V. A.; Aralov, A. V.; Fedunin, S. V.; Klykov, V. N.; Chakhmakhcheva, O. G. *Russ. J. Bioorg. Chem.* **2009**, *35*, 250-253.
(264) Efimov, V. A.; Aralov, A. V.; Klykov, V. N.; Chakhmakhcheva, O. G. *Nucleosides Nucleotides Nucleic Acids* **2009**, *28*, 846-865.
(265) Davies, D. B. *Prog. Nucl. Magn. Reson. Spectrosc.* **1978**, *12*, 135-225.
(266) Praseuth, D.; Guieysse, A. L.; Helene, C. *Biochim. Biophys. Acta, Gene Struct. Expression* **1999**, *1489*, 181-206.
(267) Duca, M.; Vekhoff, P.; Oussedik, K.; Halby, L.; Arimondo, P. B. *Nucleic Acids Res.* **2008**, *36*, 5123-5138.
(268) Seidman, M. M.; Glazer, P. M. *J. Clin. Invest.* **2003**, *112*, 487-494.
(269) Jain, A.; Wang, G.; Vasquez, K. M. *Biochimie* **2008**, *90*, 1117-1130.
(270) Mukherjee, A.; Vasquez, K. M. *Biochimie* **2011**, *93*, 1197-1208.
(271) Frankkamenetskii, M.; Mirkin, S. *Annu. Rev. Biochem.* **1995**, *64*, 65-95.
(272) Zain, R.; Sun, J. S. *Cell. Mol. Life Sci.* **2003**, *60*, 862-870.
(273) Buske, F. A.; Mattick, J. S.; Bailey, T. L. *RNA Biol* **2011**, *8*, 427-439.
(274) Wang, G. L.; Vasquez, K. M. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 13448-13453.
(275) Wang, G.; Vasquez, K. M. *DNA Repair* **2014**, *19*, 143-151.
(276) Wells, R. D. *Faseb J.* **2008**, *22*, 1625-1634.
(277) Conrad, N. K. *WIREs RNA* **2014**, *5*, 15-29.
(278) Bacolla, A.; Wang, G.; Vasquez, K. M. *PLOS Genet.* **2015**, *11*, e1005696.
(279) Devi, G.; Zhou, Y.; Zhong, Z.; Toh, D. K.; Chen, G. *WIREs RNA* **2015**, *6*, 111-128.
(280) Endoh, T.; Hnedzko, D.; Rozners, E.; Sugimoto, N. *Angew. Chem. Int. Edit.* **2016**, *55*, 899-903.
(281) Bagasra, O.; Stir, A. E.; Pirisi-Creek, L.; Creek, K. E.; Bagasra, A. U.; Glenn, N.; Lee, J. S. *Appl. Immunohistochem.* **2006**, *14*, 276-290.
(282) Martianov, I.; Ramadass, A.; Barros, A. S.; Chow, N.; Akoulitchev, A. *Nature* **2007**, *445*, 666-670.
(283) Kanak, M.; Alseiari, M.; Balasubramanian, P.; Addanki, K.; Aggarwal, M.; Noorali, S.; Kalsum, A.; Mahalingam, K.; Pace, G.; Panasik, N.; Bagasra, O. *Appl. Immunohistochem.* **2010**, *18*, 532-545.
(284) Toscano-Garibay, J. D.; Aquino-Jarquin, G. *Biochim. Biophys. Acta, Gene Regul. Mech.* **2014**, *1839*, 1079-1083.
(285) Li, Y.; Syed, J.; Sugiyama, H. *Cell Chem. Biol.* **2016**, *23*, 1325-1333.
(286) Yang, M.; Ghosh, S.; Millar, D. *Biochemistry* **1994**, *33*, 15329-15337.
(287) Darby, R. a. J.; Sollogoub, M.; McKeen, C.; Brown, L.; Risitano, A.; Brown, N.; Barton, C.; Brown, T.; Fox, K. R. *Nucleic Acids Res.* **2002**, *30*, e39.
(288) Hoyne, P. R.; Gacy, A. M.; McMurray, C. T.; Maher, L. J. *Nucleic Acids Res.* **2000**, *28*, 770-775.
(289) Rajagopal, P.; Feigon, J. *Nature* **1989**, *339*, 637-640.
(290) Asensio, J. L.; Brown, T.; Lane, A. N. *Structure* **1999**, *7*, 1-11.
(291) Rougee, M.; Faucon, B.; Mergny, J. L.; Barcelo, F.; Giovannangeli, C.; Garestier, T.; Helene, C. *Biochemistry* **1992**, *31*, 9269-9278.
(292) Roberts, R.; Crothers, D. *Science* **1992**, *258*, 1463-1466.
(293) Han, H.; Dervan, P. B. *PNAS* **1993**, *90*, 3806-3810.
(294) Fox, K. R. *Curr. Med. Chem.* **2000**, *7*, 17-37.
(295) Biver, T. *Coord. Chem. Rev.* **2013**, *257*, 2765-2783.
(296) Arya, D. P.; Coffee, R. L. *Bioorg. Med. Chem. Lett.* **2000**, *10*, 1897-1899.
(297) Arya, D. P.; Coffee, R. L.; Willis, B.; Abramovitch, A. I. *J. Am. Chem. Soc.* **2001**, *123*, 5385-5395.
(298) Arya, D. P.; Coffee, R. L.; Charles, I. *J. Am. Chem. Soc.* **2001**, *123*, 11093-11094.
(299) Arya, D. P.; Xue, L.; Willis, B. *J. Am. Chem. Soc.* **2003**, *125*, 10148-10149.

(300) Arya, D. P.; Micovic, L.; Charles, I.; Coffee, R. L.; Willis, B.; Xue, L. *J. Am. Chem. Soc.* **2003**, *125*, 3733-3744.
(301) Arya, D. P. *Top. Curr. Chem.* **2005**, *253*, 149-178.
(302) Willis, B.; Arya, D. P. *Adv. Carbohydr. Chem. Biochem.* **2006**, *60*, 251-302.
(303) Xi, H.; Kumar, S.; Dosen-Micovic, L.; Arya, D. P. *Biochimie* **2010**, *92*, 514-529.
(304) Arya, D. P. *Acc. Chem. Res.* **2011**, *44*, 134-146.
(305) Xi, H.; Davis, E.; Ranjan, N.; Xue, L.; Hyde-Volpe, D.; Arya, D. P. *Biochemistry* **2011**, *50*, 9088-9113.
(306) Thuong, N.; Helene, C. *Angew. Chem. Int. Edit.* **1993**, *32*, 666-690.
(307) Robles, J.; Grandas, A.; Pedroso, E.; Luque, F. J.; Eritja, R.; Orozco, M. *Curr. Org. Chem.* **2002**, *6*, 1333-1368.
(308) Helene, C.; Giovannangeli, C.; Guieysse-Peugeot, A. L.; Praseuth, D. *Ciba Found. Symp.* **1997**, *209*, 94-102.
(309) Zengeya, T.; Gupta, P.; Rozners, E. *Angew. Chem. Int. Ed.* **2012**, *51*, 12593-12596.
(310) Felsenfeld, G.; Davies, D. R.; Rich, A. *J. Am. Chem. Soc.* **1957**, *79*, 2023-2024.
(311) Butcher, S. E.; Pyle, A. M. *Acc. Chem. Res.* **2011**, *44*, 1302-1311.
(312) Gilbert, S. D.; Rambo, R. P.; Van Tyne, D.; Batey, R. T. *Nat. Struct. Mol. Biol.* **2008**, *15*, 177-182.
(313) Theimer, C. A.; Blois, C. A.; Feigon, J. *Mol. Cell* **2005**, *17*, 671-682.
(314) Kim, N.; Zhang, Q.; Zhou, J.; Theimer, C. A.; Peterson, R. D.; Feigon, J. *J. Mol. Biol.* **2008**, *384*, 1249-1261.
(315) Fica, S. M.; Mefford, M. A.; Piccirilli, J. A.; Staley, J. P. *Nat. Struct. Mol. Biol.* **2014**, *21*, 464-471.
(316) Chen, G.; Chang, K.; Chou, M.; Bustamante, C.; Tinoco, I. *PNAS* **2009**, *106*, 12706-12711.
(317) Mitton-Fry, R. M.; DeGregorio, S. J.; Wang, J.; Steitz, T. A.; Steitz, J. A. *Science* **2010**, *330*, 1244-1247.
(318) Brown, J. A.; Valenstein, M. L.; Yario, T. A.; Tycowski, K. T.; Steitz, J. A. *PNAS* **2012**, *109*, 19202-19207.
(319) Tycowski, K.; Shu, M.; Borah, S.; Shi, M.; Steitz, J. *Cell Reports* **2012**, *2*, 26-32.
(320) Wilusz, J. E.; JnBaptiste, C. K.; Lu, L. Y.; Kuhn, C.; Joshua-Tor, L.; Sharp, P. A. *Genes Dev.* **2012**, *26*, 2392-2407.
(321) Brown, J. A.; Bulkley, D.; Wang, J.; Valenstein, M. L.; Yario, T. A.; Steitz, T. A.; Steitz, J. A. *Nat. Struct. Mol. Biol.* **2014**, *21*, 633-640.
(322) Tycowski, K.; Shu, M.; Steitz, J. *Cell Reports* **2016**, *15*, 1266-1276.
(323) Zhang, B.; Mao, Y. S.; Diermeier, S. D.; Novikova, I. V.; Nawrocki, E. P.; Jones, T. A.; Lazar, Z.; Tung, C.; Luo, W.; Eddy, S. R.; Sanbonmatsu, K. Y.; Spector, D. L. *Cell Reports* **2017**, *19*, 1723-1738.
(324) Conrad, N. K. *Virus Res.* **2016**, *212*, 53-63.
(325) Conrad, N. K.; Mili, S.; Marshall, E. L.; Shu, M.; Steitz, J. A. *Mol. Cell* **2006**, *24*, 943-953.
(326) Millevoi, S.; Moine, H.; Vagner, S. *WIREs RNA* **2012**, *3*, 495-507.
(327) Song, J.; Perreault, J.; Topisirovic, I.; Richard, S. *Translation* **2016**, *4*, e1244031.
(328) Lane, A. N.; Chaires, J. B.; Gray, R. D.; Trent, J. O. *Nucleic Acids Res.* **2008**, *36*, 5482-5515.
(329) Pandey, S.; Agarwala, P.; Maiti, S. *J. Phys. Chem. B* **2013**, *117*, 6896-6905.
(330) Gros, J.; Guédin, A.; Mergny, J.; Lacroix, L. *ChemBioChem* **2008**, *9*, 2075-2079.
(331) Bugaut, A.; Murat, P.; Balasubramanian, S. *J. Am. Chem. Soc.* **2012**, *134*, 19953-19956.
(332) Olsthoorn, R. C. L. *Nucleic Acids Res.* **2014**, *42*, 9327-9333.
(333) Su, Z.; Zhang, Y.; Gendron, T. F.; Bauer, P. O.; Chew, J.; Yang, W.; Fostvedt, E.; Jansen-West, K.; Belzil, V. V.; Desaro, P.; Johnston, A.; Overstreet, K.; Oh, S.; Todd, P. K.; Berry, J. D.; Cudkowicz, M. E.; Boeve, B. F.; Dickson, D.; Floeter, M. K.; Traynor, B. J.; Morelli,

C.; Ratti, A.; Silani, V.; Rademakers, R.; Brown, R. H.; Rothstein, J. D.; Boylan, K. B.; Petrucelli, L.; Disney, M. D. *Neuron* **2014**, *83*, 1043-1050.
(334) Haeusler, A. R.; Donnelly, C. J.; Periz, G.; Simko, E. A. J.; Shaw, P. G.; Kim, M.; Maragakis, N. J.; Troncoso, J. C.; Pandey, A.; Sattler, R.; Rothstein, J. D.; Wang, J. *Nature* **2014**, *507*, 195-200.
(335) Kralovicova, J.; Lages, A.; Patel, A.; Dhir, A.; Buratti, E.; Searle, M.; Vorechovsky, I. *Nucleic Acids Res.* **2014**, *42*, 8161-8173.
(336) Pandey, S.; Agarwala, P.; Jayaraj, G. G.; Gargallo, R.; Maiti, S. *Biochemistry* **2015**, *54*, 7067-7078.
(337) Mirihana Arachchilage, G.; Dassanayake, A.; Basu, S. *Chem. Biol.* **2015**, *22*, 262-272.
(338) Stefanovic, S.; Bassell, G. J.; Mihailescu, M. R. *RNA* **2015**, *21*, 48-60.
(339) Kwok, C. K.; Sahakyan, A. B.; Balasubramanian, S. *Angew. Chem. Int. Ed.* **2016**, *55*, 8958-8961.
(340) Endoh, T.; Rode, A. B.; Takahashi, S.; Kataoka, Y.; Kuwahara, M.; Sugimoto, N. *Anal. Chem.* **2016**, *88*, 1984-1989.
(341) Rode, A. B.; Endoh, T.; Sugimoto, N. *Angew. Chem. Int. Ed.* **2016**, *128*, 14527-14531.
(342) Kwok, C. K.; Marsico, G.; Sahakyan, A. B.; Chambers, V. S.; Balasubramanian, S. *Nat. Methods* **2016**, *13*, 841-844.
(343) Bugaut, A.; Balasubramanian, S. *Nucleic Acids Res.* **2012**, *40*, 4727-4741.
(344) Vorlickova, M.; Kejnovska, I.; Sagi, J.; Renciuk, D.; Bednarova, K.; Motlova, J.; Kypr, J. *Methods* **2012**, *57*, 64-75.
(345) da Silva, M. W. *Methods* **2007**, *43*, 264-277.
(346) Kuo, M. H.; Wang, Z.; Tseng, T.; Li, M.; Hsu, S. D.; Lin, J.; Chang, T. *J. Am. Chem. Soc.* **2015**, *137*, 210-218.
(347) Qiu, J.; Liu, J.; Chen, S.; Ou, T.; Tan, J.; Gu, L.; Huang, Z.; Li, D. *Org. Lett.* **2015**, *17*, 4584-4587.
(348) Prusiner, S. B. *PNAS.* **1998**, *95*, 13363-13383.
(349) Prusiner, S. B. *Annu. Rev. Genet.* **2013**, *47*, 601-623.
(350) Gomes, M. P. B.; Vieira, Tuane C R G; Cordeiro, Y.; Silva, J. L. *WIREs RNA* **2012**, *3*, 415-428.
(351) Silva, J. L.; Cordeiro, Y. *J. Biol. Chem.* **2016**, *291*, 15482-15490.
(352) Macedo, B.; Cordeiro, Y. *Int. J. Mol. Sci.* **2017**, *18*, 1023.
(353) Proske, D.; Gilch, S.; Wopfner, F.; Schatzl, H. M.; Winnacker, E. L.; Famulok, M. *ChemBioChem* **2002**, *3*, 717-725.
(354) Murakami, K.; Nishikawa, F.; Noda, K.; Yokoyama, T.; Nishikawa, S. *Prion* **2008**, *2*, 73-80.
(355) Mashima, T.; Matsugami, A.; Nishikawa, F.; Nishikawa, S.; Katahira, M. *Nucleic Acids Res.* **2009**, *37*, 6249-6258.
(356) Mashima, T.; Nishikawa, F.; Kamatari, Y. O.; Fujiwara, H.; Saimura, M.; Nagata, T.; Kodaki, T.; Nishikawa, S.; Kuwata, K.; Katahira, M. *Nucleic Acids Res.* **2013**, *41*, 1355-1362.
(357) Hayashi, T.; Oshima, H.; Mashima, T.; Nagata, T.; Katahira, M.; Kinoshita, M. *Nucleic Acids Res.* **2014**, *42*, 6861-6875.
(358) Cavaliere, P.; Pagano, B.; Granata, V.; Prigent, S.; Rezaei, H.; Giancola, C.; Zagari, A. *Nucleic Acids Res.* **2013**, *41*, 327-339.
(359) Zhang, A. Y. Q.; Balasubramanian, S. *J. Am. Chem. Soc.* **2012**, *134*, 19297-19308.
(360) Orita, A.; Hamada, Y.; Nakano, T.; Toyoshima, S.; Otera, J. *Chem. Eur. J.* **2001**, *7*, 3321-3327.
(361) Ren, B.; Cai, L.; Zhang, L.; Yang, Z.; Zhang, L. *Tetrahedron Lett.* **2005**, *46*, 8083-8086.
(362) Hakimelahi, G. H.; Proba, Z. A.; Ogilvie, K. K. *Tetrahedron Lett.* **1981**, *22*, 4775-4778.

Annales Universitatis Turkuensis

Turun yliopisto
University of Turku

Painosalama Oy, Turku, Finland 2018

ISBN 978-951-29-7252-4 (PRINT)
ISBN 978-951-29-7253-1 (PDF)
ISSN 0082-7002 (PRINT) | ISSN 2343-3175 (PDF)