UNIVERSITY OF TURKU

RECOGNITION OF DOUBLE-HELICAL DNA AND RNA BY MODIFIED TRIPLEX-FORMING OLIGONUCLEOTIDES AND PEPTIDE NUCLEIC ACIDS

Ville Tähtinen

TURUN YLIOPISTON JULKAISUJA – ANNALES UNIVERSITATIS TURKUENSIS
SARJA – SER. AI OSA – TOM. 617 | ASTRONOMICA – CHEMICA – PHYSICA – MATHEMATICA | TURKU 2019

UNIVERSITY OF TURKU

RECOGNITION OF DOUBLE-HELICAL DNA AND RNA BY MODIFIED TRIPLEX-FORMING OLIGONUCLEOTIDES AND PEPTIDE NUCLEIC ACIDS

Ville Tähtinen

TURUN YLIOPISTON JULKAISUJA – ANNALES UNIVERSITATIS TURKUENSIS
SARJA – SER. AI OSA – TOM. 617 | ASTRONOMICA – CHEMICA – PHYSICA – MATHEMATICA | TURKU 2019

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

Asst. Professor 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 Eriks Rozners
Department of Chemistry
Binghamton University
Binghamton, USA

Professor Mikko Oivanen
Department of Chemistry
University of Helsinki
Helsinki, Finland

Opponent

Professor Sabine Müller
Institute of Biochemistry
University of Greifswald
Greifswald, Germany

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

ISBN 978-951-29-7785-7 (PRINT)
ISBN 978-951-29-7786-4 (PDF)
ISSN 0082-7002 (Print)
ISSN 2343-3175 (Online)
Painosalama Oy - Turku, Finland 2019

UNIVERSITY OF TURKU
Department of Chemistry
Faculty of Science and Engineering
Laboratory of Organic Chemistry and Chemical Biology
Bioorganic Chemistry
TÄHTINEN, VILLE: Recognition of Double-Helical DNA and RNA by Modified Triplex-Forming Oligonucleotides and Peptide Nucleic Acids
Doctoral thesis, 153 p.
September 2019

ABSTRACT

Triplex-forming oligonucleotides (TFOs) have received considerable interest because of their ability to recognize double-helical DNA and RNA sequence-selectively. Sequence-selective binding to double-helical DNA would enable controlling the function of a certain gene through triplex formation, which makes TFOs potential therapeutic agents. In recent years, the recognition of double-helical RNA through triplex formation has attracted increasing interest because of the central role of noncoding, double-helical RNA in cellular processes and gene expression. With TFO-based recognition, it would be possible to detect these double-helical RNA molecules and interfere with their function. However, unmodified TFOs have certain limitations such as poor triplex stability at physiological pH and poor cellular delivery. To overcome these limitations, various chemical modifications have been introduced into TFOs.

In the present thesis, TFOs and triplex-forming peptide nucleic acids (TFPNAs) were modified to increase their affinity for DNA and RNA duplexes. The binding affinity for DNA duplexes could be increased by conjugating TFOs with neomycin, a known groove binder. Moreover, the importance of the conjugation site was demonstrated. The binding affinity and specificity of TFPNAs for RNA duplexes could be increased by introducing chiral backbone modifications. TFPNAs with an appropriate pattern of γ-(*R*)-hydroxymethyl and γ-(*S*)-guanidinylmethyl modifications preferred stoichiometric Hoogsteen-face binding to a microRNA-215 model, and the γ-(*S*)-guanidinylmethyl modifications also enhanced the cellular uptake of TFPNAs. Additionally, ^{19}F-labeled RNA target hairpins and ^{19}F NMR spectroscopy were used to reveal detailed information on the stoichiometry and transition between alternative binding modes, stoichiometric PNA/RNA triplex and ternary $(PNA)_2$/RNA triplex invasion complex. Furthermore, this thesis describes ^{19}F-labeled TFO probes able to recognize variable nucleobases in the pyrimidine strand of double-helical DNA.

KEY WORDS: triplex-forming oligonucleotides, peptide nucleic acids, ^{19}F NMR spectroscopy, DNA, RNA

TURUN YLIOPISTO
Luonnontieteiden ja tekniikan tiedekunta
Kemian laitos
Orgaanisen kemian ja kemiallisen biologian laboratorio
Bio-orgaaninen kemia
TÄHTINEN, VILLE: Kaksoiskierteisen DNA:n ja RNA:n tunnistaminen muokattujen kolmoisjuosteen muodostavien oligonukleotidien ja peptidinukleiinihappojen avulla
Väitöskirja, 153 s.
Syyskuu 2019

TIIVISTELMÄ

Kolmoisjuosteen muodostavat oligonukleotidit (TFOt) ovat herättäneet suurta mielenkiintoa, sillä ne pystyvät sitoutumaan kaksijuosteiseen DNA:han ja RNA:han sekvenssiselektiivisesti. Tämän ominaisuutensa ansiosta TFO-juosteilla on sovelluskohteita esimerkiksi uudenlaisina diagnostisina koettimina ja ne ovat myös potentiaalisia lääkeaineita. Sekvenssiselektiivinen sitoutuminen kaksijuosteiseen DNA:han voisi mahdollistaa haluttujen geenien toiminnan säätelemisen. Myös kaksijuosteisella, ei-koodaavalla RNA:lla on merkittävä rooli geenien ilmenemisessä. TFO-juosteiden avulla näitä kaksijuosteisia RNA-molekyylejä voitaisiin havaita ja niiden toimintaan voitaisiin vaikuttaa. TFO-juosteiden käyttökelpoisuutta rajoittavat kuitenkin esimerkiksi kolmoisjuosteen heikko pysyvyys fysiologisessa pH:ssa ja heikko kulkeutuminen soluun.

Tässä väitöskirjatyössä TFO-juosteita ja kolmoisjuosteen muodostavia peptidinukleiinihappoja (TFPNAt) muokattiin kemiallisesti, jotta niiden sitoutumisvoimakkuus DNA- ja RNA-kaksoisjuosteisiin kasvaisi. Sitoutumisvoimakkuutta DNA-kaksoisjuosteisiin saatiin kasvatettua liittämällä TFO-juosteisiin neomysiiniä, joka pystyy sitoutumaan kolmoisjuosteen isoon uurteeseen. TFPNA-juosteiden sitoutumisvoimakkuutta ja -spesifisyyttä taas saatiin kasvatettua kiraalisilla TFPNA-rakenteilla. TFPNAt, joiden sopiviin kohtiin oli liitetty γ-(*R*)-hydroksimetyyli- tai γ-(*S*)-guanidinyylimetyyliryhmiä, suosivat stoikiometristä, Hoogsteen-vetysidoksiin perustuvaa sitoutumista mikroRNA-215-malliin. γ-(*S*)-guanidinyylimetyyliryhmät myös paransivat TFPNA-juosteiden kulkeutumista soluun. Lisäksi ^{19}F-leimattujen RNA-kaksoisjuosteiden ja ^{19}F-NMR-spektroskopian avulla paljastui yksityiskohtaista tietoa TFPNA-juosteiden vaihtoehtoisista sitoutumismekanismeista. Lisäksi tässä väitöskirjatyössä kehitettiin ^{19}F-leimattuja TFO-koettimia, joiden avulla pystyttiin tunnistamaan yksittäisiä nukleoemäksiä DNA-kaksoisjuosteen pyrimidiinijuosteessa.

AVAINSANAT: kolmoisjuosteen muodostava oligonukleotidi, peptidinukleiinihappo, ^{19}F-NMR-spektroskopia, DNA, RNA

Contents

List of Original Publications

The thesis is based on the following publications:

I Tähtinen, V., Granqvist, L. and 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.

II Tähtinen, V., Granqvist, L., Murtola, M., Strömberg, R. and 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.

III Bhuma, N., Tähtinen, V. and Virta, P.: Synthesis and applicability of base-discriminating DNA-triplex-forming ^{19}F NMR probes. *Eur. J. Org. Chem.* **2018**, *2018*, 605–613.

IV Tähtinen, V., Verhassel, A., Tuomela, J. and Virta, P.: γ-(*S*)-Guanidinylmethyl Modified Triplex-Forming Peptide Nucleic Acids Increase Hoogsteen-Face Affinity for a MicroRNA and Enhance Cellular Uptake. *(accepted manuscript) ChemBioChem* 10.1002/cbic.201900393

List of Related Publications

Granqvist, L., Kraszewski, A., Tähtinen, V. and Virta, P.: Synthesis of aminoglycoside-2′-O-methyl oligoribonucleotide fusions. *Molecules* **2017**, *22*, 760–772.

Granqvist, L., Tähtinen, V. and Virta, P.: Synthesis of Glycosidic (β-1’’→6, 3’ and 4’) Site Isomers of Neomycin B and their Effect on RNA and DNA Triplex Stability. *Molecules* **2019**, *24*, 580–591.

Abbreviations and Symbols

A	adenine
Ac	acetyl
AE	aminoethylribose
AEEA	2-(2-aminoethoxy)ethoxy acetic acid
All	allyl
Arg	arginine
2',4'-BNA^{NC}	2'-*O*,4'-*C*-aminomethylene bridged nucleic acid
Boc	*tert*-butyloxycarbonyl
BOP	(benzotriazol-1-yloxy)tris(dimethylamino)phosphonium hexafluorophosphate
BQQ	benzo[*f*]quino[3,4-*b*]quinoxaline
C	cytosine
CD	circular dichroism
dA	2'-deoxyadenosine
DCM	dichloromethane
DIEA	*N*,*N*-diisopropylethylamine
DMF	dimethylformamide
DMSO	dimethyl sulfoxide
DMTr	4,4′-dimethoxytrityl
DNA	deoxyribonucleic acid
ds	double-stranded
dT	2'-deoxythymidine
dU	2'-deoxyuridine
ENA	ethylene-bridged nucleic acid
Et_3N	triethylamine
ESI-TOF MS	electrospray ionization time-of-flight mass spectrometry
Fmoc	fluorenylmethyloxycarbonyl
G	guanine
GAL4	galectin-4 protein
GPNA	guanidine-modified PNA
H-DNA	naturally occurring triple-helical DNA
heg	hexaethylene glycol

HF488	HiLyte Fluor 488
HIV	human immunodeficiency virus
HIV-TAR	human immunodeficiency virus trans-activation response element
HPLC	high performance liquid chromatography
HRMS	high resolution mass spectrometry
iC	5-methylisocytosine
ITC	isothermal titration calorimetry
LCAA-CPG	long chain alkylamine controlled pore glass
LNA	locked nucleic acid
lncRNA	long noncoding RNA
Lys	lysine
Me	methyl
^{Me}C	5-methylcytosine
MEG3	maternally expressed gene 3
Me_3P	trimethylphosphine
Me_3SiCl	trimethylsilyl chloride
^{Me}P	3-methyl-2-aminopyridine
miRNA	microRNA
miR-215	microRNA-215
mRNA	messenger RNA
NER	nucleotide excision repair
NIS	N-iodosuccinimide
NMP	N-methyl-2-pyrrolidone
NMR	nuclear macnetic resonance
ODN	2'-deoxyoligonucleotide
ON	oligonucleotide
ORN	oligoribonucleotide
Pbf	2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl
PMB	*p*-methoxybenzyl
PNA	peptide nucleic acid
PNHDMAP	dimethylaminopropyl phosphoramidate
–1 PRF	–1 programmed ribosomal frameshifting
PS	phosphorothioate
PvuII	a restriction enzyme from *Proteus vulgaris*
Py	pyridine
PyBOP	benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate
rA	adenosine
RNA	ribonucleic acid
RP-HPLC	reverse-phase high performance liquid chromatography
rt	room temperature

rU	uridine
SNP	single-nucleotide polymorphism
ss	single strand
T	thymine
TBTA	tris[(1-benzyl-1*H*-1,2,3-triazol-4-yl)methyl]amine
*t*Bu	*tert*-butyl
TFA	trifluoroacetic acid
Tfa	trifluoroacetyl
TFO	triplex-forming oligonucleotide
TFPNA	triplex-forming peptide nucleic acid
TGF-β	transforming growth factor beta
THF	tetrahydrofurane
T_m	thermal melting temperature
$T_m{}^3$	triplex thermal melting temperature
TMS	trimethylsilyl ether
Trt	trityl
TTS	triple-helix target sequence
U	uracil
UV	ultraviolet
VP16	Herpes simplex virus protein 16

1 Introduction

Triplex-forming oligonucleotides (TFOs) have received considerable interest because of their potential as therapeutic agents that sequence-selectively recognize double-helical DNA.[1,2] The sequence of the DNA base pairs in double-helical DNA holds the genetic information and gives the cell the instructions to perform its functions. The double-helical structure of DNA was discovered in 1953.[3–5] Four years later, Felsenfeld et al. reported the formation of a three-stranded structure between one poly(rA) and two poly(rU) chains in the presence of Mg^{2+}.[6] However, it took 30 years until researchers started to realize that genes and their function could be modulated by TFOs that bind sequence-specifically to the major groove of DNA duplexes. In 1987, short oligonucleotides were used to direct DNA double helix cleavage sequence-specifically through triplex formation.[7,8] At the time, the potential of TFOs in the control of gene expression was also suggested.[9,10] To date, the ability of TFOs to control gene expression site-specifically has been demonstrated in numerous studies.[1,2,11–13] For instance, TFOs have been shown to inhibit transcription and replication and to induce site-directed mutations, recombination and DNA repair. Furthermore, TFOs have been used to target cleaving and cross-linking agents, transcription factors and nucleases to their target site in DNA duplex. Despite their great potential, TFOs have not been applied in therapy so far. Regarding therapeutic use, unmodified TFOs have certain limitations. Some of the limitations are general for oligonucleotide therapeutics, such as poor cellular delivery and stability towards nucleases, but TFOs have also certain specific challenges: poor triplex stability at physiological pH, limited target sequences (polypurine target sequences are usually required), and delivery to nucleus. To overcome these limitations, various chemical modifications have been developed.

In recent years, the recognition of double-helical RNA through triplex formation has also received increasing interest because of the biological importance of double-helical RNA structures.[14–17] For instance, a large number of noncoding RNAs, RNA transcripts that are not transcribed into proteins, are either completely double-helical or contain short double-helical regions. MicroRNA (miRNA) precursors are an example of such noncoding RNAs. With TFO-based recognition, it would be possible to detect these double-helical RNA molecules and interfere with their function. However, TFOs usually have higher affinity for DNA duplexes than for RNA

duplexes. Recently, the recognition of double-helical RNA by TFPNAs has shown promising results.[18–31] TFPNAs have been shown to have significantly higher affinity for RNA duplexes compared to DNA duplexes.[18,20,25] For instance, TFPNAs that bind to double-helical regions of messenger RNA have been demonstrated to suppress translation *in vitro* and in cells and to stimulate programmed ribosomal frameshifting *in vitro*.[23,28]

1.1 DNA triplexes

Double-helical DNA is formed by two complementary strands that bind to each other through hydrogen bonding between adenine and thymine (A–T) and guanine and cytosine (G–C) nucleobases, the Watson–Crick base pairs (Figure 1a). The formation of the Watson–Crick base pairs leaves hydrogen bond acceptor and donor groups in the major groove of DNA duplex that do not participate in the Watson–Crick hydrogen bonding. These acceptor and donor groups may enable the binding of a third strand through Hoogsteen or reverse Hoogsteen hydrogen bonding (Figure 1d).[32–35] There are three classes of triple helices that all require an oligopurine-oligopyrimidine target duplex (TC triplex, GT triplex and GA triplex, Figure 1c).[36,37] In the TC triplex, the third strand is oriented parallel to the polypurine strand of the target duplex and forms T*A–T and C^+*G–C base triplets through Hoogsteen hydrogen bonding. The C^+*G–C base triplet requires protonation at the N3 site of cytosine ($pK_a \sim 4.5$), and the TC triplex formation is thus limited to acidic conditions. In the GT triplex, the third strand can be oriented either parallel or antiparallel to the polypurine strand of the target duplex. The parallel triplex is formed through Hoogsteen G*G–C and T*A–T triplets and the antiparallel triplex through reverse Hoogsteen G*G–C and T*A–T triplets. In the GA triplex, the third strand is antiparallel to the polypurine strand of the target duplex and forms reverse Hoogsteen G*G–C and A*A–T base triplets.

Because TFOs bind to DNA duplexes sequence-specifically, they enable targeting certain genes by triplex formation. Targeting oligonucleotides to genes is called the antigene strategy.[1] Compared to the antisense strategy,[39,40] which uses antisense oligonucleotides that bind to single-stranded messenger RNA (mRNA) by duplex formation, one advantage of the antigene strategy is that there are only two copies (two alleles) of the target gene in each cell, whereas there may be thousands of copies of the corresponding mRNA. Blocking mRNA translation even by sequence-specific mRNA cleavage does not prevent the transcription of the corresponding gene, and therefore, the mRNAs are resynthesized. The antigene strategy, on the other hand, may prevent the mRNA synthesis as long as the antigene oligonucleotide is bound to its target gene and not cleaved by nucleases. However, the antigene strategy has certain difficulties compared to the antisense strategy, such as delivering the therapeutic oligonucleotide to cell nucleus. One of the main limitations for the use of

TFOs is the requirement for an oligopurine-oligopyrimidine target duplex. However, triple-helix target sequences (TTS) are highly over-represented in the human genome, especially at promoter regions.[41,42] Orozco et al. suggested that even if TTS are not directly targeted by transcription factors, they would act as spacing fragments that help the correct positioning of transcription factors. Consequently, TTS would be important for gene functionality.

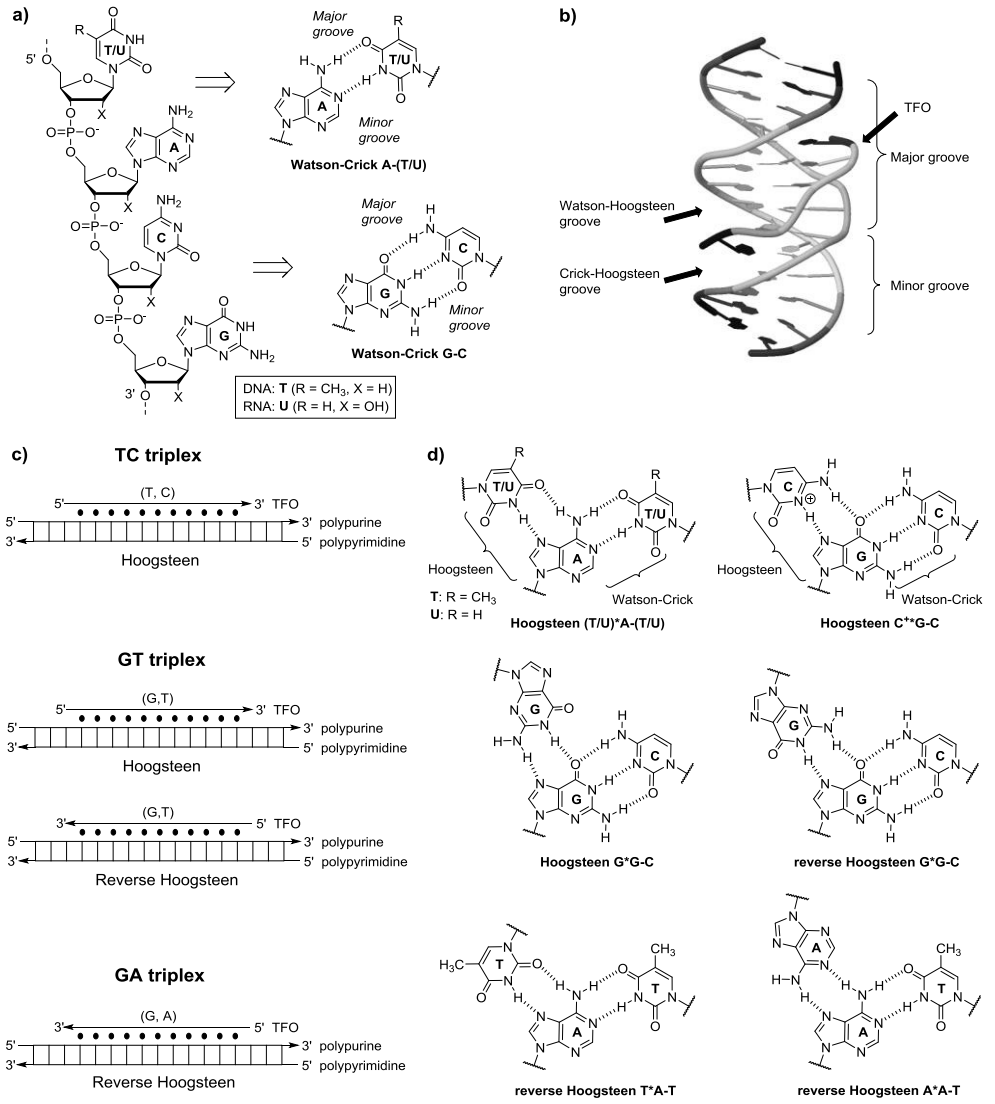


Figure 1. Schematic presentation of DNA triplexes. a) The structure of DNA and RNA and the Watson-Crick base pairs. b) The structure of a parallel TC triplex (PDB 1D 1BWG[38]). c) The three classes of DNA triplexes and the orientation of the triplex-forming strand. d) The Hoogsteen and reverse Hoogsteen base triplets.

DNA triplexes occur also in nature. H-DNAs are naturally occurring intramolecular DNA triplexes that have been suggested to be involved in gene expression regulation.[12,43–45] For example, an intramolecular triplex in the human c-MYC promoter was reported to modulate transcription.[46] H-DNAs also induce genome instability and they are related to several diseases,[12,47,48] such as autosomal dominant polycystic kidney disease, lymphomas and Friedreich's ataxia[49]. DNA triplexes have also been suggested to be present at the catalytic site of DNA polymerases.[50] Recently, MEG3 long noncoding RNA was reported to regulate the TGF-β pathway genes through RNA-DNA triplex formation.[51] The results suggest that RNA-DNA triplex formation may be a general mechanism by which long noncoding RNAs recognize their target gene.

1.1.1 Improving the binding properties and cellular delivery of TFOs

When TFO binds to its target duplex, the resulting triplex is usually thermodynamically less stable than the target duplex. In other words, the triplex melting temperature is usually lower than that of the duplex ($T_m^3 < T_m^2$). One reason for the lower thermal stability is the fact there are less hydrogen bonds in the Hoogsteen-face than in the Watson–Crick-face (cf. Figure 1d). Additionally, the low triplex stability at physiological conditions originates from the charge repulsion between the three negatively charged strands. Consequently, the triplex stability can be increased with high concentrations of multivalent cations, such as Mg^{2+},[52,53] or polyamines.[54,55] In the case of purine-rich TFOs, the triplex formation may be disturbed by competing structures, such as G-quadruplexes or GA homoduplexes. G-quadruplexes are stabilized by monovalent cations, such as K^+,[56] whereas GA homoduplexes are stabilized by divalent Mg^{2+}.[57] Additionally, under acidic conditions, C-rich TFOs can form competing structures that interfere with the triplex formation, such as CC^+ duplexes and i-motifs.[58,59]

The nuclear environment of living cells raises additional challenges for the potential *in vivo* applications of TFOs. The TFO must form a triplex at physiological pH and salt concentrations, be nuclease resistant and not be blocked in a stable secondary structure. Furthermore, the triplex must be stable enough to have time to perform its biological function on the target duplex, for example inhibit transcription or target DNA modifying agents. Moreover, the delivery of TFOs to the cell nucleus is a challenging task (section 1.1.1.5). To overcome these challenges and to improve the binding properties of TFOs, a number of chemical modifications have been introduced to TFOs.[2] The modifications may be introduced either in the nucleobases (section 1.1.1.1), in the phosphodiester backbone (1.1.1.3) or in the sugar moiety (1.1.1.2). Moreover, attempting to stabilize triplexes, triplex-binding ligands have been studied (1.1.1.4).

1.1.1.1 Nucleobase modifications

Aiming to increase the binding affinity of TFOs, various nucleobase modifications have been studied (Figure 2). The majority of the modifications have been introduced in pyrimidine bases. To improve the stability of TC triplexes at physiological pH, cytosine modifications have been developed. 5-methylcytosine (^{Me}C) was shown to stabilize triplexes at higher pH already in 1989,[60] and more recently, 2-aminopyridine derivatives (e.g. ^{Me}P) have been used.[61–64] Thymidine has also been replaced by several modifications, such as 2'-deoxyuridine (dU)[65] and 5-(prop-1-ynyl)-2'-deoxyuridine[66]. Further triplex stabilization has been obtained by other *C*-5-(alkyn-1-yl)-modified TFOs,[67–70] such as 5-(1-propargylamino)-2'-deoxyuridine[67]. The triplex-stabilizing effect of the *C*-5-(alkyn-1-yl)-modifications most likely arises principally from favorable base stacking interactions. Recently, a thiazole orange intercalator was conjugated to the *C*-5 site of a thymine nucleobase in TFOs to increase the triplex stability.[71] With three thiazole orange conjugates, the triplex was stabilized by 45 °C at pH 7. Moreover, the triplex formation could be detected as an increase in the fluorescence signal of thiazole orange.

Regarding the potential applications of TFOs, the requirement for polypurine-polypyrimidine target duplexes is one of the limitations. Because the binding of TFOs is sequence-specific, any single T-A or C-G base pair interruption reduces strongly the triplex stability.[72–74] To overcome this sequence limitation, two main approaches have been proposed.[75] The first approach consists of conjugating intercalators to the TFO, aiming to stabilize the triplex containing base pair interruptions in the purine strand.[76] The drawback of this strategy is the reduced sequence-specificity. The second approach is the specific base pair strategy. This strategy consists of synthesizing new modified nucleobases capable of forming hydrogen bonds with one or both partners of the T-A or C-G base pair inversions.[75,77–81] For instance, N^4-(2-guanidoethyl)-5-methylcytosine has been developed to recognize CG inversions and *N*-acetyl-2,7-diamino-1,8-naphtyridine to recognize TA inversions (Figure 2).[82] Consequently, nucleobase-modified TFOs capable of sequence-specific recognition of DNA duplexes containing all the four base pairs have been reported.[63,82] Section 1.2.1.1 describes more nucleobase modifications used for the recognition of double-helical RNA.

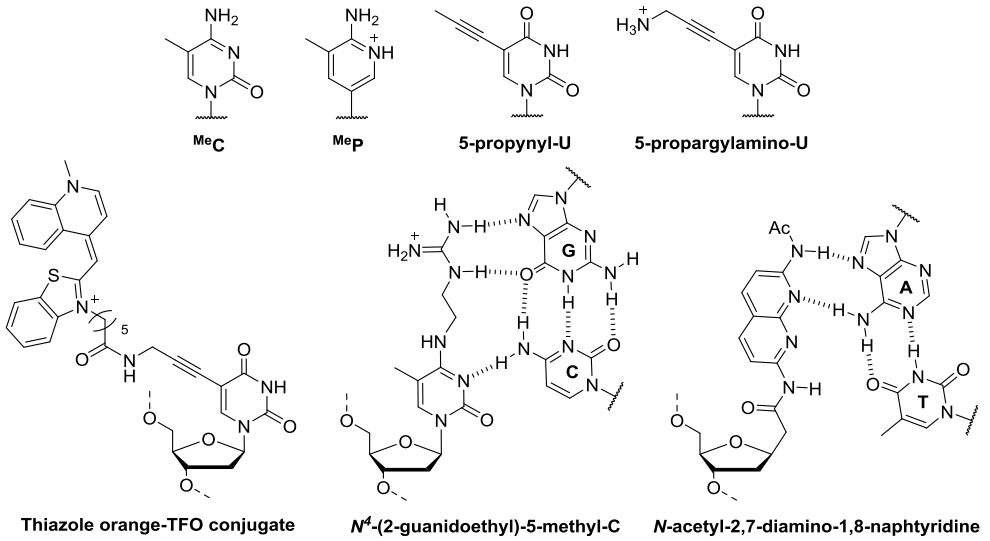


Figure 2. Examples of the nucleobase modifications introduced in TFOs.

1.1.1.2 Sugar modifications

Several modifications in the sugar moiety of TFOs have also been developed (Figure 3a). In 1992, Roberts and Crothers reported that RNA third strands form more stable triplexes with double-helical DNA compared to DNA third strands.[83] This observation attracted increasing interest in the development of various ribose analogues. 2'-methoxylation (2'-OMe) stabilizes the C3'-endo conformation of the ribose sugar (Figure 3b), thus causing less distortion to the target duplex and increasing the triplex stability. Furthermore, the 2'-OMe modification prevents degradation by RNA nucleases. The 2'-aminoethylribose (2'-AE) modification also stabilizes the C3'-endo conformation, and the positive charge of the amino group provides additional triplex stabilization.[84]

Another successful approach has been to lock the ribose sugar to the C3'-endo conformation using locked nucleic acid (LNA) monomers. LNA monomers have an *O*-2', *C*-4'-methylene link and they have been shown to increase the binding affinity for both single-stranded and double-stranded DNA and RNA.[85–88] LNA-modified oligonucleotides have also been reported to strand-invade double-stranded DNA through Watson–Crick base pairing.[87] In addition to the more favorable C3'-endo conformation, the reason for the increased binding affinity of LNA may be the conformational restriction itself: it is well established that preorganization may reduce the entropic penalty of complex formation.[89] However, even if LNA monomers significantly stabilize triplexes,[90,91] fully LNA-substituted TFOs do not form triplexes.[92] A number of LNA modifications have been developed during the years. Ethylene-bridged nucleic acid (ENA) is an LNA homologue that also stabilizes

triplexes, although to a smaller extent compared to LNA.[93,94] On the other hand, in contrast to LNA, fully ENA-substituted TFOs form stable triplexes. Furthermore, the stability of LNA-modified TFOs has been increased by introducing other modifications within the bridge of the LNA sugar. 2',4'-BNA^{NC} contains a 2'-*O*, 4'-*C*-aminomethylene bridge that allows the substitution of the amino nitrogen.[95] Moreover, a fully 2',4'-BNA^{NC}(NH)-substituted TFO still formed a stable triplex and had an improved nuclease resistance.

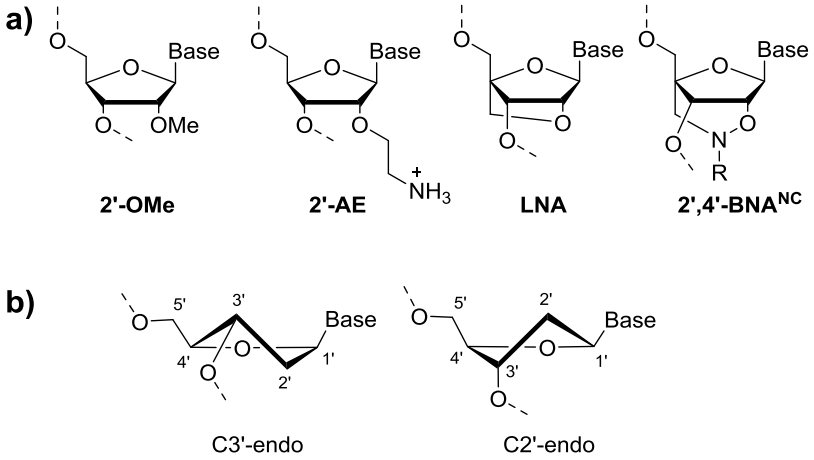


Figure 3. a) Examples of the sugar modifications introduced in TFOs. b) C3'-endo (*N*-type, A-form helices) and C2'-endo (*S*-type, B-form helices) sugar conformations.

1.1.1.3 Backbone modifications

The backbone modifications aim to reduce the electrostatic repulsion upon hybridization caused by the negatively charged phosphodiester backbone and to improve the nuclease resistance of TFOs. Examples of such backbone modifications are phosphorothioates (PS)[96], dimethylaminopropyl phosphoramidates (PNHDMAPs)[97] and peptide nucleic acids (PNAs)[98] (Figure 4a). Phosphorothioates, in which one of the non-bridging phosphate oxygens is replaced by sulfur, improve the nuclease resistance of oligonucleotides. Additionally, phosphorothioates have better pharmacokinetic properties than unmodified oligonucleotides.[39] This improvement results from their enhanced binding to plasma proteins, which prevents their rapid excretion in the urine and thus facilitates uptake by tissues. However, if the phosphorothioate content in oligonucleotide is high, nonspecific interactions with cellular proteins may cause toxic side effects.[99] Another drawback of phosphorothioates may be their tendency to accumulate in the liver and kidney,

although the liver is rich in gene targets for therapeutic oligonucleotides.[100,101] Phosphorothioate modifications are widely used in antisense oligonucleotides.

PNAs, on the other hand, are synthetic oligonucleotide analogues in which the naturally occurring sugar-phosphate backbone is replaced by a peptide backbone consisting of *N*-(2-aminoethyl)glycine units.[102] PNAs can hybridize with complementary DNA and RNA through Watson–Crick and Hoogsteen base pairing just as their natural counterparts, but with higher affinity and sequence-selectivity.[103–105] The improved affinity results from the neutral peptide backbone that causes no electrostatic repulsion upon hybridization, and the relative rigidity of the PNA backbone is suggested to be the reason for the improved sequence-selectivity. In addition, PNAs are resistant to nucleases and proteases[106] and have a low affinity for proteins[107]. Unlike the backbone of nucleic acids, the PNA backbone is non-chiral. Consequently, the binding of PNAs to complementary DNA or RNA is not very selective between parallel vs. antiparallel strands. The therapeutic potential of PNAs is limited by their poor cellular uptake. Cationic peptides conjugated to PNAs have been shown to enhance the delivery of PNAs in cultured cells,[27,31,108–112] but entrapment of PNAs within endosomes is another challenge yet to overcome (see section 1.1.1.5).

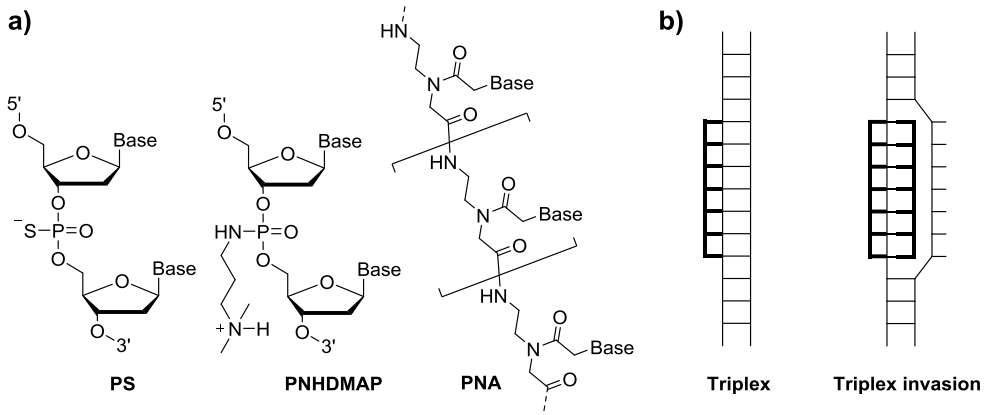


Figure 4. a) Examples of the backbone modifications introduced in TFOs. b) Schematic presentation of the dsDNA–PNA triplexes.

PNAs were originally designed for recognizing double-helical DNA through triplex formation.[102] However, it was soon realized that homopyrimidine PNA oligomers preferred the triplex invasion binding mode when binding to their oligopurine-oligopyrimidine target duplex (Figure 4b). In the triplex invasion complex, one PNA oligomer binds to the purine strand of the target duplex through Watson–Crick base pairing and the other PNA oligomer through Hoogsteen hydrogen bonding. This invasion complex is very stable, which has led to the development of bisPNA clamps, where two PNA strands, connected by a flexible linker, target the same purine

strand.[98,113] Interestingly, triplex-forming PNAs (TFPNAs) have significantly higher affinity for RNA duplexes compared to DNA duplexes.[18,20,25] Therefore, TFPNAs are discussed in more detail in section 1.2.1. Although not directly related to Hoogsteen-face recognition, an interesting recent approach for recognizing double-helical DNA and RNA is using backbone-modified bifacial PNAs.[114] These bifacial PNAs are capable of invading between the double helix and, owing to modified bifacial nucleobases, capable of forming Watson–Crick base pairs with both strands of the target duplex.

1.1.1.4 Triplex-stabilizing ligands

Aiming to increase the stability of triplexes, several triplex-stabilizing small molecular ligands have been studied.[115] The majority of the ligands have been heterocycles that intercalate between the triplex bases.[116–119] Another approach for stabilizing the triplex is using major or minor groove binders. Minor groove binders are usually not triplex-specific, and they often even destabilize triplexes because they prefer the DNA duplex.[120,121] Among the major groove binders, the aminoglycoside neomycin has shown the most potential.[115,121–129]

Neomycin B (Figure 5a) is one of the aminoglycoside antibiotics discovered to be active against tuberculosis in the 1940s.[130,131] Neomycin B is very specific for stabilizing the DNA triplex without affecting the duplex stability: for example, when 1.67 neomycin units/base triplet were added to a poly(dA)•2poly(dT) triplex at pH 6.8, the triplex melting temperature $T_m{}^3$ was increased by 25 °C (from 34 to 59 °C), whereas the duplex melting temperature $T_m{}^2$ was virtually unaffected.[122] Additionally, neomycin has been shown to improve the cationic lipid-mediated delivery of oligonucleotides in cells.[132] The triplex-stabilizing effect of neomycin originates from its binding to the Watson-Hoogsteen groove (Figure 5a). Most aminoglycosides with five or more amino groups are able to stabilize DNA triplexes without affecting the duplex stability. However, paramomycin stabilized DNA triplex 16 °C less compared to neomycin, although the only difference between neomycin and paramomycin is one amino group in ring I replaced by a hydroxyl group (cf. structures in Figure 5).[121] Ribostamycin, on the other hand, lacking the ring IV compared to neomycin, stabilized the triplex only to a small extent. These results indicate that the positively charged amino groups in rings I and IV are key factors for the binding of neomycin. Neomycin's affinity for RNA triplex is even higher, but it lacks selectivity between RNA triplex and duplex. The binding affinity of neomycin follows the trend: DNA duplex < DNA triplex < DNA/RNA hybrids < RNA duplex ~ RNA triplex.[115] In a recent study conducted in our laboratory, the effect of neomycin site isomers on DNA and RNA triplex stability was evaluated.[133] Three site isomers, separated by the binding site of rings I and II to the ribose sugar, were synthesized (Figure 5c). Interestingly, all the site isomers increased the triplex stability almost equally compared to neomycin. Consequently, the spatial orientation of the amino groups was favorable for triplex binding regardless of the site isomer.

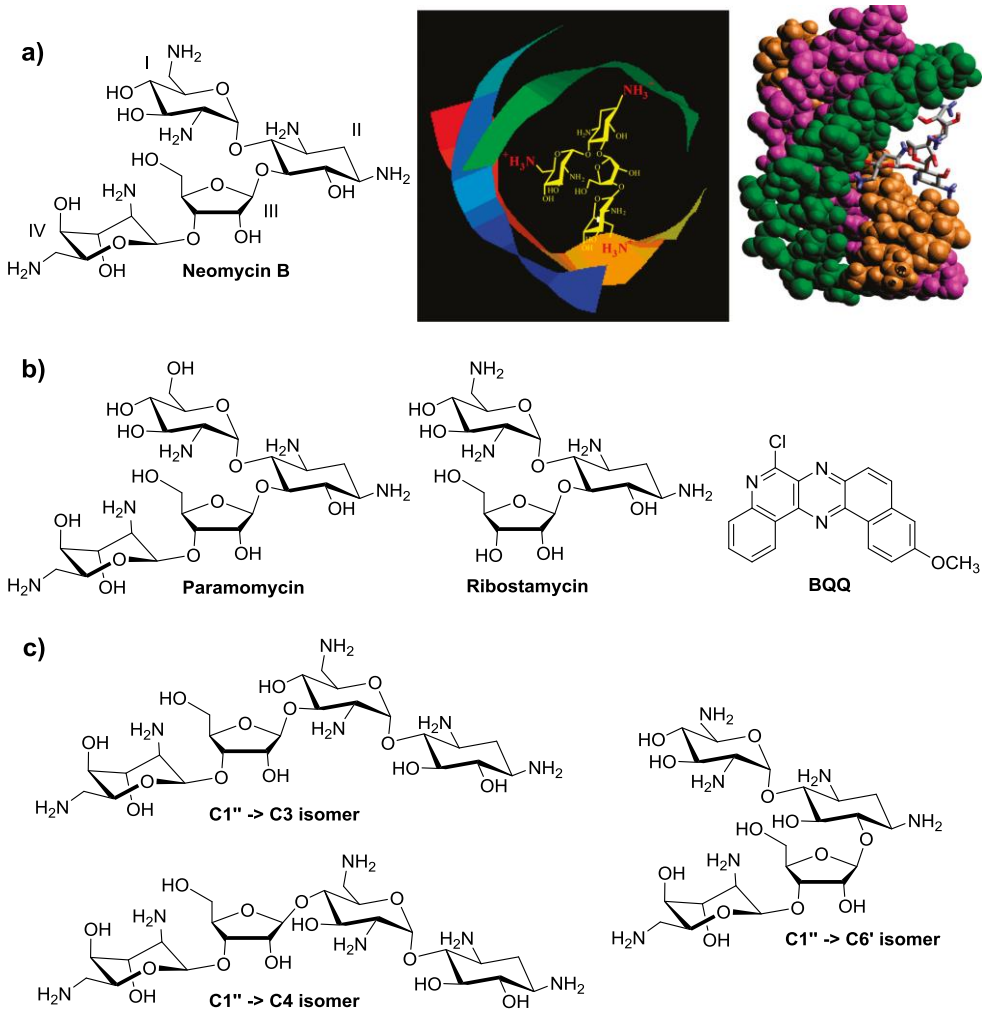


Figure 5. a) The structure of neomycin B with ring numbering and the proposed binding to the Watson–Hoogsteen groove of DNA triplex. Reprinted and modified with permission from Arya, D. P. *Acc. Chem. Res.* **2011**, *44*, 134–146. Copyright (2011) American Chemical Society.[115] b) The structures of the aminoglycosides paramomycin and ribostamycin and the intercalator BQQ. c) The structures of neomycin B site isomers.

Compared to neomycin, intercalator ligands often show equal or more triplex stabilization at low concentrations. At higher concentrations, on the other hand, intercalators often tend to stabilize the duplex as well, although more triplex-selective intercalators have been developed.[116,118,134,135] The most effective triplex-stabilizing intercalator to date is benzo[*f*]quino[3,4-*b*]quinoxaline (BQQ).[136] With 1.5 μmol L^{-1} oligonucleotide concentration and 7 equivalents of BQQ relative to the oligonucleotide, a TC triplex was stabilized by 51 °C (from 18 to 69 °C) at pH 6.2, whereas the duplex was stabilized by 4 °C (from 58 to 62°C). The conjugation of

BQQ to neomycin resulted in almost 1000-fold binding constant compared to neomycin alone.[137] Furthermore, several intercalators and other ligands have been covalently conjugated to TFOs to result in remarkable triplex stabilization, for example spermine,[138] acridine,[139] Hoechst 33258,[140] cyclopapyrroloindole,[141] benzopyridoindole[142] psoralen[143] and thiazole orange[71] (see section 1.1.1.1). However, these ligands have no such selectivity for triplex vs. duplex stabilization as compared to neomycin. Surprisingly, covalent conjugation of neomycin to TFOs and its effect on the triplex stability was not reported before our studies described in section 3.1. Several studies describe the conjugation of oligonucleotides and PNAs with neomycin. The neomycin–oligonucleotide conjugates have been reported to stabilize DNA and RNA duplexes to some extent. For example, neomycin, covalently attached to the *C*-5 position of a 2'-deoxyuridine residue in DNA-oligonucleotide, enhanced duplex formation with an α-sarcin loop RNA target.[144] Additionally, neomycin conjugated to 2'-OMe oligoribonucleotides enhanced invasion to a HIV-TAR RNA model.[145]

1.1.1.5 Improving the cellular delivery of TFOs

As with all therapeutic oligonucleotides, the delivery of TFOs to their intracellular target is a challenging task. The cellular uptake of unmodified oligonucleotides is inefficient.[146] The large size, hydrophilicity and the anionic backbone of oligonucleotides prevent their direct permeation through the cell membrane surface. Oligonucleotides enter the cells via either adsorptive or receptor-mediated endocytosis, thus resulting in vesicular uptake.[146–149] The ionic repulsion between the negatively charged oligonucleotide and cell membrane disturbs the adsorption to the cell surface. However, the cellular uptake cannot be improved by eliminating the negative charge alone, as demonstrated by the poor cellular uptake of neutral oligonucleotide analogs, such as PNA.[147] After entering the cytoplasm, the oligonucleotide vesicles are delivered to the endosome-lysosome compartment. Endosomes are membranous organelles that act as sorting compartments in cells. The oligonucleotide must be able to escape the endosome to reach its site of action in the cytoplasm or nucleus. If the oligonucleotide gets trapped in the endosome, as it often does, it is likely to be degraded by lysosomes.

To facilitate the cellular uptake of oligonucleotides, three main strategies have been developed.[150] One strategy is to use chemically modified oligonucleotides, such as phosphorothioates[39]. Unlike unmodified oligonucleotides, phosphorothioate oligonucleotides can slowly cross the endosomal lipid bilayer to escape endosomes into the cytoplasm or nucleus by an unknown mechanism called gymnosis.[151–155] The second strategy relies on molecular, macromolecular or supramolecular carriers that form non-covalent complexes with oligonucleotides. These carriers include cationic lipids, dendrimers, polymers, cyclodextrins and carbon nanotubes. For instance,

dendrimers, highly branched 3D molecules, were demonstrated to enhance the cellular uptake of TFOs by forming TFO–dendrimer nanoparticles.[156]

The third strategy for improving the cellular delivery of oligonucleotides is using covalently bound oligonucleotide conjugates, for instance cell-penetrating peptides, ligands for cell surface receptors and lipophilic substituents, such as cholesterol (Figure 6). Cholesterol conjugation to TFOs has been shown to enhance the cellular uptake by liver cells.[157] Furthermore, the cholesterol–TFO conjugate inhibited transcription *in vitro*.[157] Cationic groups have also been conjugated to TFOs, aiming to reduce the ionic repulsion upon hybridization and enhance cellular permeation. One potential cationic conjugate is the guanidinium group, the functional group of arginine. The guanidinium group is highly basic (pK_a = 12.5) and thus, it remains protonated under a wide pH range. Guanidinium modifications have been shown to improve the cellular uptake when introduced either in the nucleobase[158] or in the backbone[159] of oligonucleotides. Guanidinium modifications also improve the cellular uptake of PNA.[160,161] As discussed previously, the cellular uptake of PNAs can also be improved by conjugation with cationic peptides.[27,31,108–112] Conjugation of TFOs with ligands that bind to cell surface receptors, on the other hand, could allow cell-specific targeting.[162] Carbohydrates, in particular *N*-acetylcalactosamines (GalNAc),[163,164] are a widely studied example of such ligands.

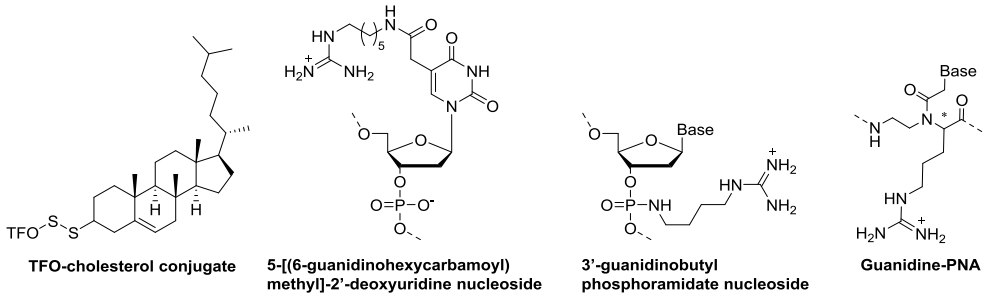


Figure 6. Examples of the conjugate groups reported to improve the cellular uptake of TFOs and other oligonucleotides.

Compared to antisense approaches for single-stranded RNA, targeting dsDNA in the cell nucleus raises an additional challenge. DNA is typically bound to histones and tightly packed into chromatin, which may prevent TFO from accessing its target sequence. However, various demonstrations of triplex formation in chromosomal environment have been reported.[165] In general, the nuclear entry itself may not be the rate-limiting step for therapeutic oligonucleotides, as phosphorothioate oligonucleotides are demonstrated to continuously shuttle between the nucleus and cytoplasm.[166]

1.1.2 Potential applications of TFOs

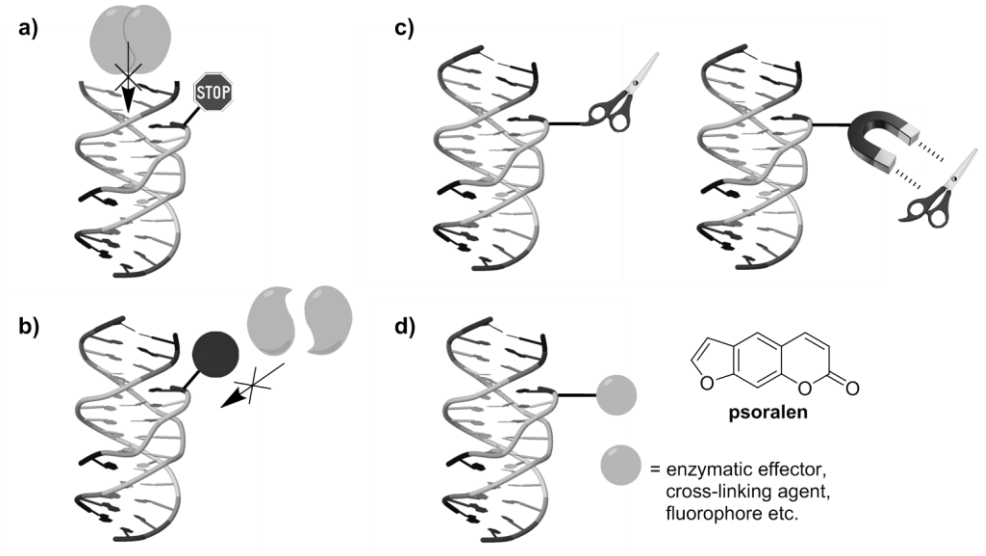


Figure 7. Examples of the potential applications of TFOs. a) Blockage of transcription or replication elongation. b) Blockage of transcription or elongation initiation complex. c) Site-specific cleavage of the target duplex. d) Site-specific targeting of drugs, enzymatic effectors (transcription factors, restriction enzymes), cross-linking agents, fluorophores etc.

1.1.2.1 Modulating transcription and replication via triplex formation

By binding to the major groove, TFOs can prevent specific proteins from recognizing the DNA duplex.[1,167–169] Various *in vitro* studies have shown that TFOs can modulate gene expression by inhibiting the transcription, by interrupting either the binding of transcription factors[170,171] or the formation of the initiation complex[172] (Figure 7b). TFOs can also stop transcription elongation, either alone or conjugated to psoralen or other DNA damaging agents (Figure 7 a & d).[173–175]

On the other hand, TFOs can be used to increase gene expression. Human β-globin disorders, such as sickle cell anemia and β-thalassemia, can be treated by increasing the transcription levels of certain genes. For example, TFO-psoralen conjugate was used to direct psoralen-induced mutation site-specifically to the γ-globin gene transcription factor binding site.[176] This mutation resulted in 4-fold activation of gene expression. Additionally, a TFO conjugated to transcriptional activation domains derived from Herpes simplex virus protein 16 (VP16) was demonstrated to activate the expression of the target gene.[177] The activation of transcription has also been demonstrated *in vivo* in *Saccharomyces cerevisiae* cells.[177] The authors used a TFO with a hairpin residue that has a binding site for GAL4 protein, a transcription activator. Furthermore, transcription can be activated by

targeting TFO to the gene repressor region, thus stimulating the expression of the gene in interest.[178,179] TFOs have also been reported to inhibit replication.[50,180] *In vitro* studies show that DNA polymerase elongation can be inhibited by targeting TFOs upstream[181] or downstream[182,183] the initiation site.

1.1.2.2 Introducing or correcting mutations site-specifically

TFOs have been used to induce DNA damage site-specifically to their target duplex, thus causing mutation and recombination to the target site both *in vitro* and *in vivo* (Figure 7 c & d).[184,185] UV light, alkylating agents and photoreactive molecules such as psoralen have been shown to cause recombination, but not site-specifically. However, conjugating DNA damaging agents to TFOs enables site-specific targeting. Various damaging agents have been conjugated to TFOs, such as photoactivable agents,[186] metal complexes,[187] orthophenantroline,[188] nucleases[189] and restriction enzymes. For instance, the restriction enzyme PvuII was used to cleave DNA site-specifically.[190] Psoralen–TFO conjugates, on the other hand, have been employed in targeting site-specific mutations through the formation of base pair -specific psoralen adducts.[191] Psoralen–TFO conjugates were also demonstrated to form cross-links in the target DNA in cultured cells.[192] TFOs itself, without conjugation, have also been reported to induce recombination through a nucleotide excision repair (NER) -dependent pathway.[193] Recently, TFPNAs have been demonstrated to stimulate the correction of gene mutations: TFPNAs stimulated the correction of a splice-site mutation in the β-globin gene,[194] and γ-modified TFPNAs induced the correction of anaemia in β-thalassemic mice *in vivo*.[195]

1.1.2.3 Tools for biotechnology and nanotechnology

In addition to being used in the modulation of gene expression, TFOs can also be employed as biotechnological tools.[2] TFOs can be employed in targeting drugs to specific sites of DNA, site-specific DNA labeling and in the recognition and purification of DNA. For instance, a TFO conjugated with camptothecin, a topoisomerase I inhibiting drug, stimulated cleavage at the DNA target site.[196] TFOs bearing a fluorescence probe, on the other hand, were used to label chromosomes in cell nucleus.[197] TFOs have also been used to purify their target duplex from a bacterial lysate by immobilizing the TFOs on an affinity chromatography support.[198] At low pH, the TFO bound its target duplex from the lysate, and the target duplex could eventually be released from the support by eluting with alkaline buffer. Additionally, TFOs can be applied for the recognition of protein binding sites on DNA, since the third strand can act as a competitor for the DNA-binding protein.[199] Furthermore, DNA triplexes have great potential in DNA nanotechnology, a rapidly developing research area.[200,201]

1.1.2.4 Recognition of the nucleobase content in the target DNA duplex by TFO probes

One potential application for TFOs is using them as sensors for recognizing nucleobase content in the target duplex. In cytogenetic research and clinical diagnosis, there is a strong need for the sequence-selective detection of nucleic acid sequences. For instance, single-nucleotide polymorphisms (SNPs) and other irregular DNA structures are considered to play a major role in the production of nonfunctional proteins.[202] Therefore, detecting SNPs would be important for monitoring various biological processes and identifying genetic mutations of inherited diseases. For clinical diagnosis, labeled TFOs could provide a convenient method to detect nucleic acid sequences directly from cell, with no need for amplification steps such as PCR (polymerase chain reaction). Several methods for detecting SNPs have been developed,[203–211] many of which rely on changes in hybridization energies. However, the differences in the hybridization energies tend to be small between complementary oligonucleotide strands and oligonucleotides containing a single nucleotide mismatch.

Fluorescence-labeled oligonucleotide probes have the potential to increase the sensitivity of the recognition, if the fluorescence label shows marked change in the fluorescence intensity between the hybridized and non-hybridized probe. Fluorescent-labeled oligonucleotide probes have been developed for recognizing SNPs in single-stranded DNA.[212,213] The recognition was based on either duplex formation[212] or triplex formation with a clamp-type TFO.[213] Additionally, fluorescent-labeled TFOs have been used for specific labeling of genomic DNA duplexes.[197] A fluorescent-labeled TFO probe that recognizes Watson–Crick base pair inversions in double-stranded DNA was demonstrated by Pedersen et al. (Figure 8).[214] A pyrene moiety at the 3'-position of a uridine residue was used as the fluorescent label. The TFO probe showed a drastic increase in fluorescence upon binding to complementary dsDNA, but only marginal change in the fluorescence was observed upon binding to dsDNAs with single base pair inversions. The single base pair inversions resulted in a Hoogsteen mismatch between the TFO and the dsDNA. Because all three Hoogsteen mismatches resulted in a very weak fluorescence signal, the TFO probe was unable to discriminate the Hoogsteen mismatches from each other.

Recently, ^{19}F NMR based sensors, because of their sensitivity to local environments, have also been developed for the recognition of SNPs and other irregular DNA structures. Tanabe et al. introduced a ^{19}F-labeled nucleobase modification into oligonucleotides and were able to detect mismatched and bulged structures in ssDNA.[215] Hocek et al. used biaryl-substituted nucleobases as dual fluorescent and ^{19}F NMR probes capable of detecting mismatches, deletions and hairpins in ssDNA.[216] These examples both required duplex formation between the target ssDNA and the labeled oligonucleotide strand. Sakamoto et al. developed an external ^{19}F NMR probe capable of detecting nucleobase content in dsDNA.[217] This

3,5-bis(trifluoromethyl)-phenyl-modified bisbenzimide H33258 probe binds to the minor groove of an AATT sequence of dsDNA. The ^{19}F NMR shift of the probe was sensitive to the base pairs adjacent to the AATT binding site, and the probe was able to recognize bulge structures and all the four different base pairs neighboring the binding site.[218] Section 3.2 describes ^{19}F-labeled TFO probes capable of discriminating variable nucleobases in the pyrimidine strand of dsDNA.

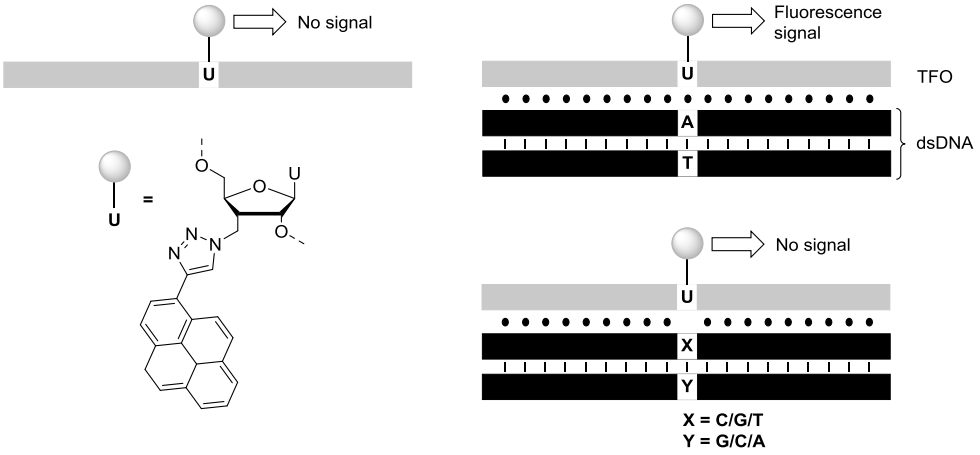


Figure 8. The principle of the pyrene-labeled fluorescent TFO probe for the recognition of Watson–Crick base pair inversions in dsDNA.[214]

1.2 RNA triplexes

Compared to DNA triplexes, RNA triplexes have received relatively little attention. However, the biological importance of RNA triple helices has started to reveal in recent years.[45,219–222] Naturally occurring RNA triplexes are ubiquitous and important for folding RNAs into complex tertiary structures that are crucial for their biological functions. Some examples of biologically important RNAs with triple-helical regions include long noncoding RNAs (lncRNAs),[223–226] telomerase RNAs,[227–232] metabolite-sensing riboswitches,[233–241] ribosomal RNAs,[242–244] mRNA pseudoknots that induce –1 ribosomal frameshifting,[245–250] group I introns[251–253] and group II introns[254]. Moreover, the molecular recognition of double-helical RNA through triplex formation has attracted increasing interest in recent years because of the central role of noncoding RNAs in cellular processes and gene expression.[14–17] A large number of noncoding RNAs, RNA transcripts that are not transcribed into proteins, are either completely double-stranded or contain short double-helical regions. For instance, microRNA (miRNA) precursors adopt a double-stranded form. Since the involvement of miRNAs in cancer development was demonstrated in 2002,[255] miRNAs have been

the mostly studied noncoding RNAs. Therefore, the TFO-based recognition of double-helical RNA is an appealing research field.

The TFOs that target double-helical RNA are usually pyrimidine strands that bind to the major groove of oligopurine-oligopyrimidine target duplex in parallel orientation (parallel pyrimidine motif RNA triplex, cf. TC triplex in Figure 1c).[222] The TFOs bind to the target duplex through Hoogsteen hydrogen bonding to form (T/U)*A–U and C^+*G–C base triplets. Also other types of major groove RNA triplexes may form.[256] One of the major limitations in targeting RNA duplexes by TFOs is that TFOs usually have higher affinity for DNA duplexes than for RNA duplexes. Because RNA duplex adopts an A-form helical structure, the major groove is relatively deep and narrow compared to B-form DNA duplex. For example DNA and 2'-OMe-RNA based TFOs only bind to DNA but not to RNA duplexes.[83,88,257–260] DNA as the third strand probably cannot find optimal geometry for binding to A-form RNA duplex, and the 2'-OMe may cause steric hindrance with the RNA duplex. In addition to Hoogsteen hydrogen bonding, naturally occurring RNA triplex is also stabilized by backbone–backbone hydrogen bonding between the 2'-OH from the third strand and a non-bridging phosphate oxygen from the purine strand of the target duplex. The additional 2'-OH–phosphate hydrogen bond is missing with DNA or 2'-OMe-RNA as the third strand. It is worth noting that 2'-OMe modification enhances the binding of TFOs to DNA duplexes, probably due to the geometrical compatibility with the major groove of DNA duplex. LNA modifications, on the other hand, have been shown to stabilize TFO/$(RNA)_2$ triplexes as well as TFO/$(DNA)_2$ triplexes.[88,261] This stabilization is probably caused by the TFO backbone preorganization and steric compatibility, despite the lack of the 2'-OH–phosphate hydrogen bond.

1.2.1 Recognition of double-helical RNA by triplex-forming peptide nucleic acids (TFPNAs)

One promising approach for the recognition of double-stranded RNA is using triplex-forming peptide nucleic acids (TFPNAs) instead of TFOs.[18–31,262,263] As described in section 1.1.1.3, PNAs were originally designed for the recognition of double-helical DNA. However, TFPNAs have been shown to have significantly higher affinity for RNA duplexes compared to DNA duplexes.[18,20,25] The reason for this difference in binding affinity is not yet completely understood and requires X-ray crystallography and NMR studies, but the neutral and flexible PNA backbone probably provides favorable geometry and backbone-backbone interactions to the deep and narrow major groove of RNA duplex.[222] A recent NMR study suggests that hydrogen bonding between N-H of TFPNA backbone and non-bridging phosphate oxygen of RNA may be behind the higher affinity of TFPNAs for RNA.[264] Aiming to improve the binding properties of TFPNAs, various modifications have been introduced either in the nucleobases or in the backbone of TFPNAs.

1.2.1.1 Nucleobase modifications

The standard major-groove U*A–U base triplet is very common in natural RNAs. In TFPNAs, the uracil residue is often replaced by thymine for the recognition of A–U base pairs to form T*A–U triplets. As discussed previously, the C^+*G–C triplet requires low, non-physiological pH to fully protonate C (pK_a ~ 4.5). To overcome this pH dependence, a number of modified nucleobases have been developed for the recognition of G–C base pairs (Figure 9). The pseudoisocytosine modification (**J**, a neutral analogue of protonated cytosine, Figure 9) significantly alleviates the pH dependency of PNA/$(DNA)_2$[113,265] and PNA/$(RNA)_2$[20,25] triplex formation. The binding affinity for RNA duplexes has been further increased by a 2-aminopyridine nucleobase modification (**M**, pK_a ~ 6.7, Figure 9) that forms M^+*G–C triplets.[25] **M**-modified PNAs form stable and sequence-selective triplexes at physiological conditions, and they have been reported to be highly selective for double-stranded RNA over DNA.[25] The positive charge of **M** is probably a key factor in the increased binding affinity. Additionally, **M** has been shown to enhance the cellular uptake of PNAs.[31] **L**, a thiolated derivative of the **J**, has also been reported to enhance the recognition of G–C base pairs at physiological conditions compared to **J**.[20] Importantly, the **L**-modified PNA bound selectively to double-stranded RNA, showing relatively weak binding to double-stranded DNA, single-stranded DNA and single-stranded RNA.

The binding of a pyrimidine strand through triplex formation requires a polypurine strand in the target RNA duplex. However, in naturally occurring RNA duplexes, the purine strand is often interrupted by inverted Watson–Crick U-A and C-G base pairs. For example in microRNAs and other noncoding RNAs, it is common to find eight or more consecutive purine bases interrupted by a couple of pyrimidines. For the recognition of these pyrimidine inversions, several modified nucleobases have been developed (Figure 9). For the recognition of the C-G inversion, PNAs modified with 5-methylisocytosine (**iC**)[19], 2-pyrimidinone (**P**)[266,267], $\mathbf{P_{ex}}$[24] and **Q**[262] have been reported. The **iC** modification slightly increases the affinity for the target RNA duplex, but the base-selectivity is low.[19] The **P** modification shows relatively good base-selectivity, but the binding affinity is low.[24] The binding affinity can be improved using the $\mathbf{P_{ex}}$ modification with an additional linker between the PNA backbone and **P** nucleobase.[24] The $\mathbf{P_{ex}}$ modification has modest base-selectivity at pH 5.5, but interestingly, the base-selectivity improves significantly at pH 6.25. The best C-G inversion recognition has been reported with the guanidine-modified **Q**.[262] **Q** shows high binding affinity, base-selectivity and selective binding to dsRNA over ssRNA. Furthermore, the incorporation of multiple **Q** residues to PNAs improves the cellular uptake. For the recognition of the U-A inversion, on the other hand, the 3-oxo-2,3-dihydropyridazine nucleobase (**E**) has shown promising results.[24] In the future, extended nucleobases, capable of recognizing both nucleobases of the base

pair, may be a feasible option for the recognition of pyrimidine inversions (cf. **Q**). Furthermore, as discussed previously, the RNA triplex stability can be increased with small molecular ligands, such as neomycin.[115]

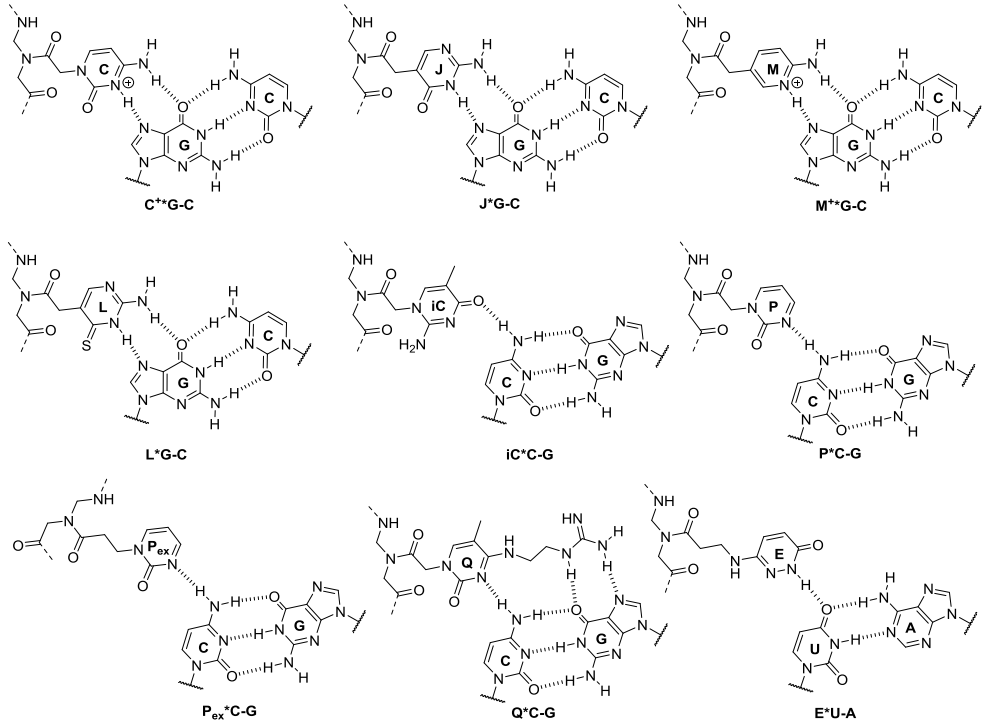


Figure 9. Modified PNA nucleobases developed for the recognition of G-C base pairs and pyrimidine inversions in double-stranded RNA.

1.2.1.2 PNA backbone modifications

The binding affinity of TFPNAs for dsRNA can also be improved by conjugating cationic peptides, such as lysine and arginine, to the backbone of TFPNAs.[27] Moreover, the cationic peptides maintained the sequence-selectivity of the TFPNAs and improved their cellular uptake. Another approach for enhancing the binding affinity and selectivity of PNAs is modifying their backbone with chiral units. The chiral units may provide suitable preorganization that improves binding to DNA and RNA targets. The binding of chiral PNAs to single-stranded RNAs has been thoroughly studied.[268–273] For example, α-guanidine modified PNA (α-GPNA, Figure 10a), where D-arginine was incorporated to every other position of the backbone instead of glycine, sustained strong and sequence-selective binding to complementary single-stranded RNA.[270] Furthermore, antisense PNAs bearing chiral arginine units at the α and γ positions have been shown to inhibit miR-210 activity in cells.[271] In

addition to beneficial preorganization, the guanidinium groups improved the cellular uptake of PNA and increased water solubility.[160,161] However, before our studies described in section 3.3, the ability of chiral TFPNAs to recognize RNA duplexes was demonstrated in only one report by Rozners et al.[26] In this report, the binding of triplex-forming α-GPNAs to double-helical RNA hairpins was studied. According to isothermal titration calorimetry (ITC), the α-GPNAs had reduced affinity and sequence-selectivity for the RNA hairpins. The data also suggested that α-GPNA preferred a 2:1 $(PNA)_2$/RNA triplex invasion complex instead of stoichiometric PNA/RNA triple helix (cf. Figure 10b).

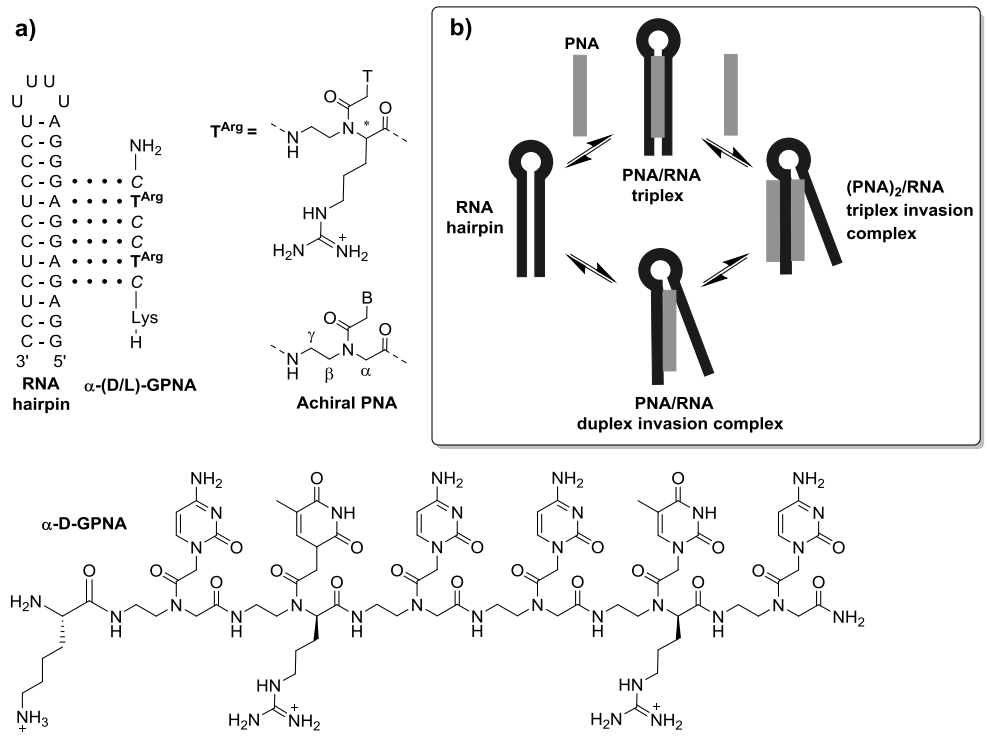


Figure 10. a) The recognition of a double-helical RNA hairpin by α-GPNAs (Rozners et al.)[26] and the potential sites for chiral substituents in the PNA backbone (α–γ). b) Examples of the potential binding modes of PNAs to RNA hairpins.

1.2.2 Potential applications

A large number of biologically important noncoding RNAs are either completely double-helical or contain short double-helical regions. Accordingly, TFO-based RNA recognition has various potential RNA targets and applications.[222] For instance, miRNAs are involved in many gene regulation processes associated with diseases such as cancer.[16,255] Cellular miRNAs are maturated from double-helical hairpin

precursors (pre-miRNAs) that are transported from the nucleus to the cytoplasm and eventually cleaved by the RNase enzyme Dicer into double-stranded miRNAs. MiRNA activity can be inhibited by antisense oligonucleotides, referred to as anti-miRs. However, the antisense-based recognition of dsRNA requires a strand invasion, where the antisense oligonucleotide must be able to replace one strand in dsRNA. The TFO-based recognition, on the other hand, does not disrupt the secondary structure of the target RNA, which presumably fastens the binding. In 2012, Zengeya et al. developed a nucleobase-modified TFPNA that binds to a pre-miRNA-215 hairpin model with high affinity at physiological conditions (see **PNA4** and **ORN2** in Figure 23).[25]

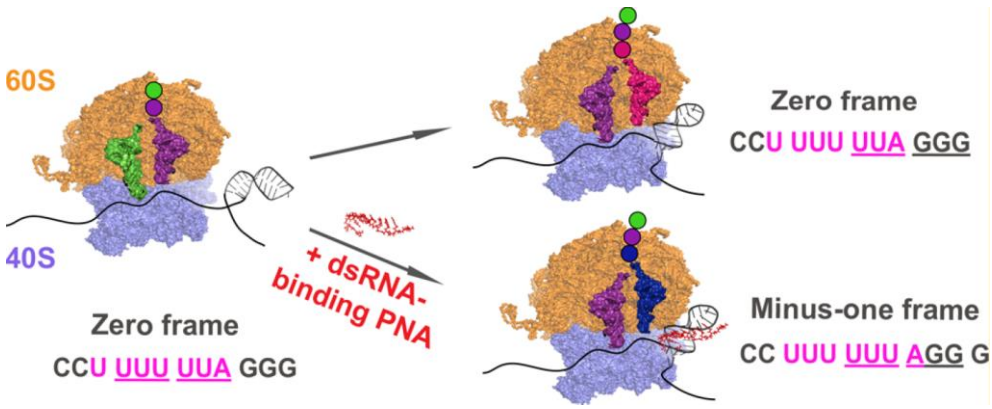


Figure 11. –1 programmed ribosomal frameshifting stimulation by TFPNAs. Reprinted with permission from Puah, R. Y.; Jia, H.; Maraswami, M.; Kaixin Toh, D.-F.; Ero, R.; Yang, L.; Patil, K. M.; Lerk Ong, A. A.; Krishna, M. S.; Sun, R.; Tong, C.; Huang, M.; Chen, X.; Loh, T. P.; Gao, Y.-G.; Liu, D. X.; Chen, G. *Biochemistry* **2018**, *57*, 149–159. Copyright (2018) American Chemical Society.[23]

To date, nucleobase-modified TFPNAs that bind to hairpin structures in mRNA have been demonstrated to suppress translation *in vitro* and in cells and to stimulate –1 programmed ribosomal frameshifting *in vitro* (Figure 11).[23,28] Many viruses and bacteria utilize –1 programmed ribosomal frameshifting (–1 PRF) to increase the information content in their genomes.[274] –1 PRF can be activated by several stimulatory elements such as a heptameric slippery sequence (X XXY YYZ) and an mRNA secondary structure (hairpin or pseudoknot) that is positioned 2–8 nucleotides downstream from the slippery site. For example, HIV-1 mutants with enhanced −1 PRF efficiencies are significantly less infectious.[275] Consequently, TFPNAs that stabilize –1 PRF stimulating RNA structures are potential antiviral agents. Recently, a nucleobase-modified TFPNA was also demonstrated to inhibit influenza viral replication by targeting a double-stranded region in viral RNA.[276] In addition,

TFPNAs may find applications in inhibiting the binding of proteins to dsRNA and in recognizing certain RNA duplexes.[222]

1.3 Methods for investigating triplexes

1.3.1 Conventional methods

Several methods have been used for studying triplexes, for example ultraviolet (UV),[277,278] circular dichroism (CD)[279,280] and nuclear magnetic resonance (NMR)[34,38,281] spectroscopy, fluorescence-based techniques,[282,283] gel electrophoresis[35,284] and isothermal titration calorimetry (ITC)[285,286]. Among these methods, the UV thermal denaturation study is probably the most used experiment. With the thermal denaturation study, the stability of the nucleic acid secondary structures, such as triplex or duplex, can be determined. The conformational changes of nucleic acids upon heating often result in changes in their UV absorbance. In most cases, the thermal denaturation experiment can be performed by recording the absorbance at 260 nm as a function of temperature. For example, duplex and triplex denaturation lead to an increase in the absorbance (hyperchromism) at this wavelength. To be able to determine duplex and triplex melting temperatures by UV spectroscopy, the difference between the duplex and triplex melting temperatures must be large enough. The denaturation of cytosine-containing triplexes can be detected more specifically by using 295 nm as the detection wavelength.[277] At 295 nm, the denaturation can be observed as a decrease in the absorbance (hypochromism), caused by the partial deprotonation of the cytosine nucleobases in the triplex-forming strand upon unfolding. CD spectroscopy can also be used to monitor triplex denaturation, even if triplexes do not have a characteristic CD spectrum like A-form helices, for instance.[280] ^{1}H NMR, on the other hand, allows the detection of Hoogsteen hydrogen bonding.[281]

1.3.2 ^{19}F NMR spectroscopy

^{19}F NMR spectroscopy has been shown to be a valuable tool for investigating biological events and biomolecules, including nucleic acids.[287] The ^{19}F nucleus has several favorable properties for NMR spectroscopy.[288–290] It has a spin of ½, 100% natural abundance and a high gyromagnetic ratio. ^{19}F has a high sensitivity, 83% of that of ^{1}H, and much higher compared to ^{13}C, ^{15}N and ^{31}P (approximately by a factor of 27, 250 and 8, respectively). Importantly, the chemical shift of ^{19}F is very sensitive to changes in the local environment of the nucleus, and the chemical shift range is wide, over 400 ppm for organofluorine compounds. Additionally, because of the absence of fluorine in nucleic acids and biological systems, there are no interfering background signals. ^{1}H NMR studies on large nucleic acid molecules are often

disturbed by the abundance of protons that need to be assigned and the severely overlapping signals within a narrow chemical shift range. The strength of ^{19}F NMR spectroscopy arises from the relatively simple one-dimensional (1D) spectra that can provide insight into complex nucleic acid structures. The drawback of ^{19}F NMR is that a fluorine probe or label has to be incorporated into the nucleic acids (or nucleic acid binding ligands) under study.[291] To make sure that the labeling causes no significant changes in the native properties of the nucleic acids, the influence of the fluorine modification should be thoroughly evaluated. Compared to spectrophotometric methods (UV, CD and fluorescence), NMR spectroscopy is relatively insensitive. However, ^{19}F NMR spectroscopy has been demonstrated to provide more detailed information on coexisting structures and local environments and local denaturation of nucleic acids.[287,288,290,291] ^{19}F NMR is often used in combination with other methods rather than being used as a competitive alternative for other techniques.

^{19}F NMR has been used to study nucleic acid structures such as aptamers,[292–294] viral regulatory hairpins,[295–300] riboswitches,[301–304] ribozymes,[305–307] DNA adducts[308,309] and G-quadruplexes[310–314]. Both DNA[315,316] and RNA[317] triplexes have also been studied by ^{19}F NMR. The monitoring of triplex formation by ^{19}F NMR was reported for the first time by Tanabe et al.[315] The authors introduced a 2'-deoxy-5-fluorouridine label into a TFO and monitored the binding of the TFO to its target duplex (Figure 12a). The TFO was a C-rich pyrimidine oligonucleotide, and thus required slightly acidic conditions for the triplex formation (pH 5.5). The TFO alone gave multiple signals at 25 °C, indicating the formation of unspecific secondary structures that are often formed by C-rich nucleic acids under acidic conditions.[58,59] This suggestion was confirmed by the fact that at neutral pH or at elevated temperature, only one ^{19}F resonance signal was observed. Triplex formation with the target duplex resulted in a marked change in the chemical shift of the ^{19}F resonance signal, albeit the signal was relatively broad. Additionally, our group developed a sensitive 4'-*C*-[(4-CF_3-1*H*-1,2,3-triazol-1-yl)methyl]thymidine sensor that allowed convenient monitoring of DNA single strand–duplex–triplex conversions by ^{19}F NMR (Figure 12b).[316] Furthermore, for the detection of RNA secondary structures, including triplexes, a 2'-O-[(4-CF_3-triazol-1-yl)methyl]uridine sensor was developed.[317] The sensors were incorporated into oligonucleotides as phosphoramidites and they caused only marginal effects on the stability of DNA and RNA models.

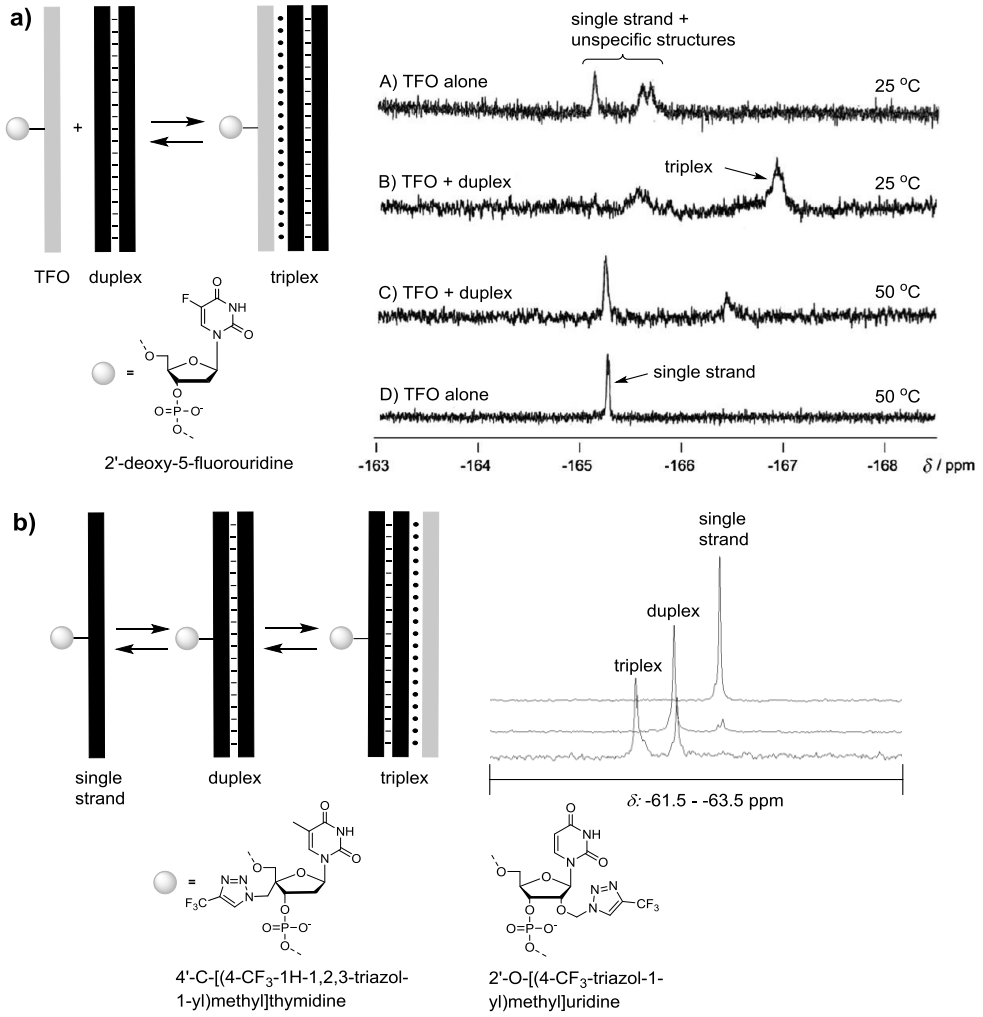


Figure 12. Triplex formation monitored by ^{19}F NMR spectroscopy using a) 2'-deoxy-5-fluorouridine sensor[315] or b) 4'-*C*-[(4-CF_3-1*H*-1,2,3-triazol-1-yl)methyl]thymidine sensor[316]. a) Spectra reprinted and modified from Tanabe, K.; Sugiura, M.; Nishimoto, S. I. *Bioorganic Med. Chem.* **2010**, *18*, 6690–6694, Copyright (2010), with permission from Elsevier.[315] b) Spectra reprinted and modified with permission from Granqvist, L.; Virta, P. *J. Org. Chem.* **2014**, *79*, 3529–3536. Copyright (2014) American Chemical Society.[316]

2 Aims of the Thesis

Triplex-forming oligonucleotides (TFOs) have received considerable interest because of their potential as therapeutic agents that sequence-selectively recognize double-helical DNA and RNA. The recognition of double-helical DNA is the basis of the antigene therapy, the aim of which is to control the function of a certain gene by triplex formation. In recent years, the recognition of double-helical RNA has become highly attractive because of the central role of noncoding RNAs in gene expression. A large number of noncoding RNAs are either completely double-stranded or contain short double-stranded regions. For instance, microRNA (miRNA) precursors adopt a double-helical form. The first aim of this thesis was to develop modified TFOs and triplex-forming peptide nucleic acids (TFPNAs) that have high affinity for DNA and RNA duplexes especially under physiological conditions. To increase the binding affinity for DNA duplexes, neomycin was conjugated to various positions of triplex-forming DNA oligonucleotides. To increase the binding affinity for RNA duplexes, chiral γ-(*R*)-hydroxymethyl and γ-(*S*)-guanidinylmethyl backbone modifications were introduced to TFPNAs.

The second aim of this thesis was to provide further information on the competing mechanisms by which TFPNAs bind to double-helical miRNA. These binding mechanisms include 1:1 PNA/RNA triplex and 2:1 $(PNA)_2$/RNA triplex invasion complex. This information was obtained especially by using ^{19}F-labeled miRNA targets and ^{19}F NMR spectroscopy, which is a simple option to discriminate different ways of binding. For the potential therapeutic use of TFPNAs in the future, it is crucial to discriminate the different binding modes and to ensure the selectivity of the binding. Furthermore, the objective of the chiral backbone modifications was to provide suitable preorganization for 1:1 PNA/RNA triplex formation, in addition to increasing the binding affinity. Moreover, the chiral modifications would have the potential to increase the cellular uptake of PNAs. The third aim of this thesis was to develop ^{19}F-labeled TFOs that are able to recognize variable nucleobase content in the pyrimidine strand of double-helical DNA. These ^{19}F-labeled TFOs may find applications in studies that aim to simultaneously detect changes of local environments in structurally resembling double-helical regions e. g. upon binding of nucleobase-specific ligands.

The aims of the thesis may be summarized as follows:

i. To develop modified TFOs (I; section 3.1) and TFPNAs (II and IV; section 3.3) that have high affinity for DNA and RNA duplexes especially under physiological conditions.
ii. To provide detailed information on the competing mechanisms by which TFPNAs bind to double-helical miRNA and to improve the selectivity for stoichiometric PNA/RNA triplex formation by chiral modifications (II and IV; section 3.3).
iii. To develop ^{19}F-labelled TFOs that are able to recognize nucleobase content in the pyrimidine strand of double-helical DNA (III; section 3.2).

3 Results and Discussion

3.1 DNA triplex stabilization by neomycin-conjugated TFOs

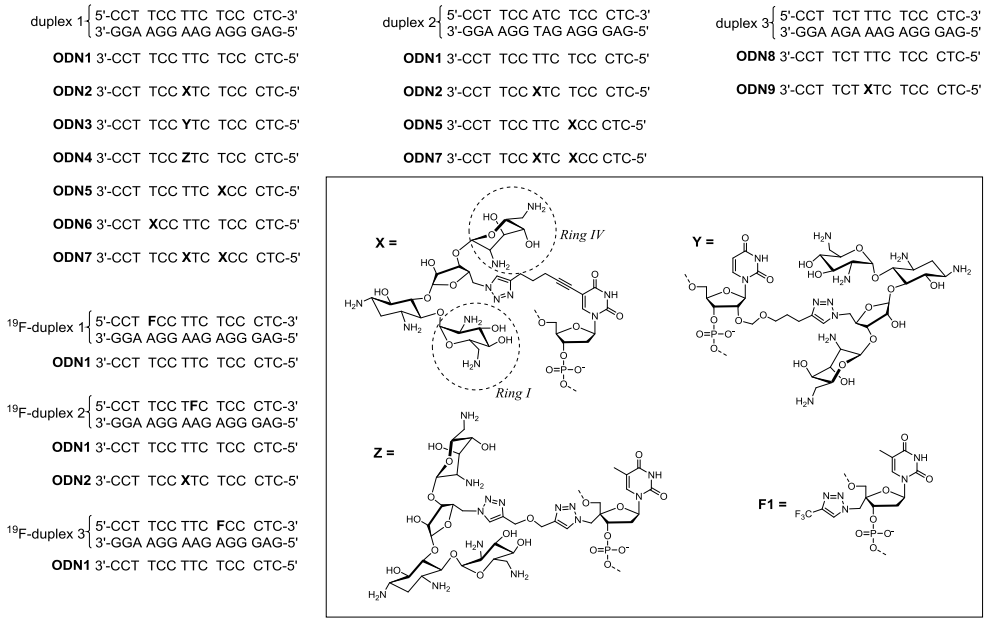


Figure 13. The DNA target duplexes and the neomycin-conjugated TFOs (**ODN2**–**ODN7** and **ODN9**) studied herein.

In the first section of this thesis, neomycin units were conjugated to variable sites of TFOs and their effect on DNA triplex stability was studied (**ODN2**–**7** and **ODN9**, Figure 13). The neomycin units were conjugated either to the *C*-5 site (**ODN2**), *C*-2' site (**ODN3**) or *C*-4' site (**ODN4**) of a thymidine residue, and the conjugation site within the TFO sequence was also varied (**ODN2** vs. **ODN5** and **ODN6**). For the synthesis of the conjugates, three different alkyne-modified phosphoramidite building blocks (**1**–**3**, Figure 14) were synthesized and incorporated into the TFOs by automated DNA synthesis. An azide-derived

neomycin (**4**) was then coupled to the terminal alkynes on solid support using the "click" chemistry conditions (1,3-dipolar Huisgen cycloaddition). Finally, the neomycin-conjugated TFOs were released from the solid support and deprotected using concentrated aqueous ammonia. The influence of the neomycin units on the triplex stability was evaluated at slightly acidic conditions (pH 5.5 and 6.0) by comparing the triplex melting temperatures (T_m[3]) obtained from the UV thermal melting profiles at 260 and 295 nm. Additionally, more detailed information on the triplex/duplex/single strand conversion was provided by ^{19}F NMR spectroscopic analysis using ^{19}F-labeled target duplexes (^{19}F-duplex 1–3, Figure 13).

3.1.1 Synthesis of C-5, C-2' and C-4'-neomycin-conjugated TFOs

3.1.1.1 Building blocks **1–4**

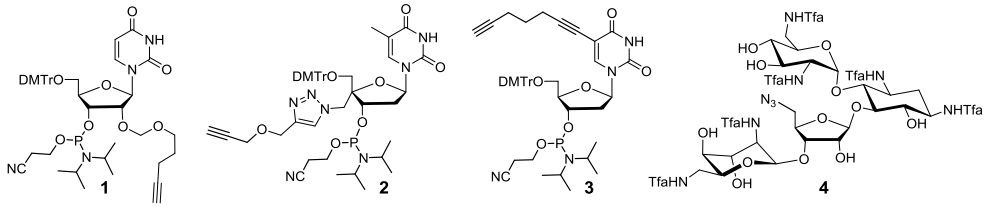
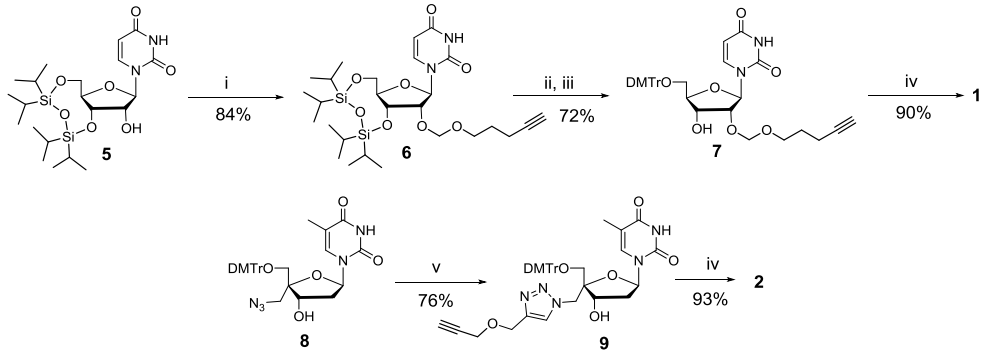


Figure 14. Alkyne-modified phosphoramidite building blocks **1–3** and azide-derived neomycin **4** used for the synthesis of neomycin-conjugated TFOs.

For the synthesis of the neomycin-conjugated TFOs, three nonstandard phosphoramidite building blocks (**1**–**3**) and an azide-derived neomycin (**4**) were synthesized (Figure 14). The synthesis of compounds **1** and **2** is described in Scheme 1. For the synthesis of compound **1**, commercially available 3',5'-*O*-(1,1,3,3-tetraisopropyl-1,3-disiloxanediyl)uridine (**5**) was first converted into acetal **6** via SN1 type alkylation with methyl[(pent-4-yn-1-yloxy)methyl]sulfide. The disiloxanediyl protection of compound **6** was then removed with tetrabutylammonium fluoride and the exposed 5'-OH group was protected with a 4,4'-dimethoxytrityl group to yield compound **7**. Finally, phosphitylation of the 3'-OH group gave compound **1** in 56% overall yield. For the synthesis of compound **2**, 5'-(4'4'-dimethoxytrityl)-4'-*C*-azidomethylthymidine (**8**) was first synthesized from thymidine in 49% yield following a protocol previously developed in our laboratory.[318] Compound **8** was then converted into compound **9** via Cu(I)-catalyzed 1,3-dipolar cycloaddition with dipropargyl ether. Finally, phosphitylation

of the 3'-OH group gave compound **2** in 35% overall yield (calculated from thymidine). In compound **2**, the triazolyl-bridge that links the alkynyl moiety to the *C*-4' of thymidine was introduced mainly for synthetic reasons. Previously, our group used a *C*-4'-alkynylaminomethyl thymidine derivative for the synthesis of neomycin-conjugated oligonucleotides, but the synthesis of this derivative was relatively complex.[319] Correspondingly, a *C*-4'-alkynyloxymethyl thymidine derivative turned out to be difficult to synthesize because of the facile formation of oxetanes.[319] Compound **3**[320,321] and compound **4**[319] were synthesized according to previously reported protocols. The amino groups of compound **4** were trifluoroacetyl protected to avoid undesired Cu(II) complexation of the neomycin moiety during the click conjugation.

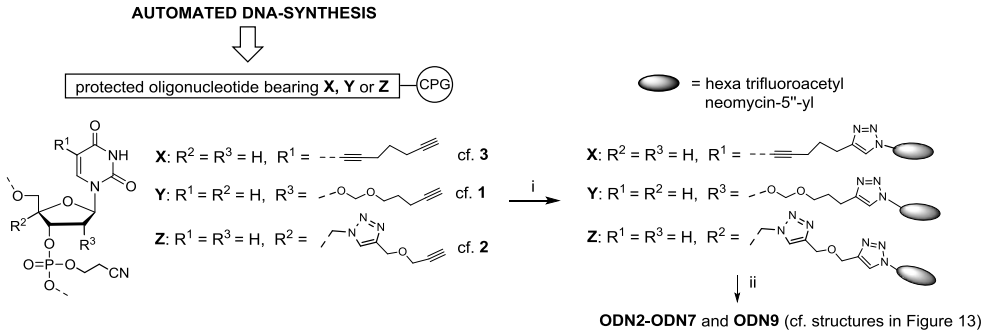


Scheme 1. (i) methyl[(pent-4-yn-1-yloxy)methyl]sulfide, NIS, CF_3SO_3H, THF, -40ºC, 2h; (ii) $Bu_4N^+F^-$, THF, 2 h, rt; (iii) DMTrCl, pyridine, overnight, rt; (iv) 2-Cyanoethyl N,N-diisopropylchlorophosphoramidite, Et_3N, DCM, 2 h, rt; (v) $CuSO_4$, sodium ascorbate, H_2O, dioxane, dipropargyl ether, overnight, rt.

3.1.1.2 Oligonucleotide synthesis

For the synthesis of the neomycin-conjugated oligonucleotides **ODN2**–**ODN7** and **ODN9** (Figure 13), alkyne modified oligonucleotides were first synthesized on solid support on a 1.0 μmol scale using an automatic DNA/RNA synthesizer (Scheme 2). The phosphoramidite building blocks **1**–**3** were introduced into the oligonucleotides using the standard phosphoramidite coupling cycle. The coupling yields of the building blocks **1**–**3** were >97%, comparable to commercial phosphoramidites. After the oligonucleotide synthesis, the azide-derived neomycin (**4**) was coupled to the solid-supported oligonucleotides using the click chemistry conditions: 10 equivalents of **4**, $CuSO_4$/TBTA, sodium ascorbate, overnight at room temperature. The neomycin-conjugated oligonucleotides were then released from the solid support by concentrated ammonia (overnight at 55 °C), purified by RP HPLC and lyophilized to white powders of **ODN2**–**ODN7** and **ODN9** in 20–

30% isolated yields. The authenticity of the oligonucleotides was verified by electrospray ionization time-of-flight mass spectrometry (ESI-TOF MS). The fluorine-labeled oligonucleotides were synthesized by introducing the phosphoramidite building block of 4'-*C*-[(4-CF_3-1*H*-1,2,3- triazol-1-yl)methyl]thymidine (**F1**, Figure 13) into the oligonucleotides as previously described by our group.[316]



Scheme 2. (i) **4**, TBTA, $CuSO_4$, sodium ascorbate, H_2O, DMSO, 2-butanol, overnight at rt, (ii) conc. aq ammonia, overnight at 55 ºC.

3.1.2 UV thermal melting studies

The UV thermal melting temperatures (Figure 15, Table 1) were determined at pH 5.5 and 6.0 by using 2.0 μmol L^{-1} or 1.0 μmol L^{-1} (entries 18–20) oligonucleotide concentration in 10 mmol L^{-1} sodium cacodylate and 0.1 or 0.3 mol L^{-1} NaCl in H_2O. The triplex stability was higher at pH 5.5 compared to pH 6.0 as expected because of the protonation of the cytosine bases in the triplex-forming strand. At pH 5.5, the melting curves of the duplexes and triplexes overlapped in most cases when λ = 260 nm was used as the detection wavelength. By using λ = 295 nm as the detection wavelength, on the other hand, the unwinding of the C^+*G-C triplet could be specifically monitored, and the triplex T_m values could be determined.[277] The influence of the conjugation site of the neomycin moiety on the triplex stability was first studied by comparing the T_m values of duplex 1 + **ODN1**–**4** (entries 1–4). The neomycin moiety was conjugated either to the *C*-5 site (**ODN2**), *C*-2' site (**ODN3**) or *C*-4' site (**ODN4**) of the thymidine residue. The *C*-5 conjugate **ODN2** increased the triplex melting temperature by = +7.3 °C (pH 5.5) and +7.7 °C (pH 6.0), while the C-2' and C-4' conjugates (**ODN3** and **ODN4**) showed no effect. The importance of the conjugation site on the triplex stability is understandable, since the *C*-5 site is oriented toward the Watson–Hoogsteen groove, where the binding of neomycin takes place.[115] The C-2' and C-4' sites, on

the other hand, are oriented outward from the triple helix, thus not providing favorable binding position for neomycin.

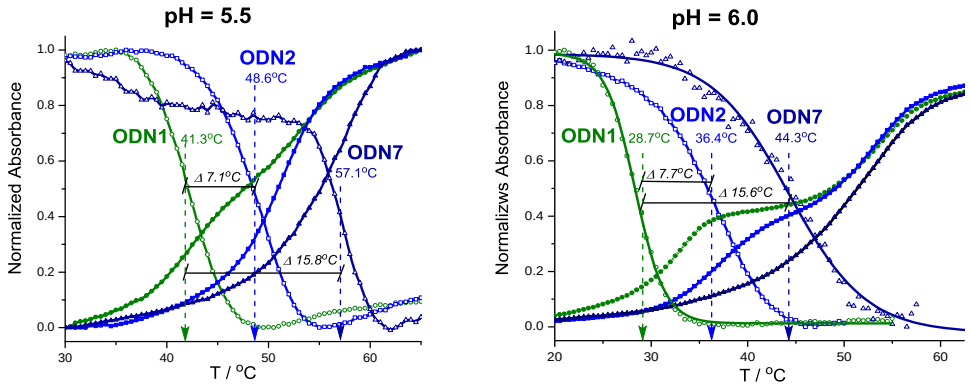


Figure 15. Normalized UV melting profiles of duplex 1 + **ODN1**/**ODN2**/**ODN7** at pH = 5.5 and 6.0. Green curves: duplex 1 + **ODN1**, blue curves: duplex 1 + **ODN2**, dark blue curves: duplex 1 + **ODN7**. Detection wavelength: 260 nm (positive curves) or 295 nm (negative curves). Conditions: 2.0 μmol L^{-1} of each oligonucleotide, 10 mmol L^{-1} sodium cacodylate, 0.1 mol L^{-1} NaCl in H_2O.

The influence of the position of neomycin within the TFO sequence was evaluated by comparing the binding affinity of **ODN2**, **ODN5** and **ODN6**, where the *C*-5 neomycin conjugate was placed at different sites within the TFO. **ODN2** and **ODN5** provided almost equal stabilization (pH 5.5: ΔT_m^3 = +7.3 °C vs. +8.1 °C; pH 6.0: ΔT_m^3 = +7.7 °C vs. +7.5 °C, entries 2 and 5). Surprisingly, **ODN6** showed no effect on the triplex stability (entry 6). In **ODN6**, the neomycin moiety was conjugated closer to the TFO terminus compared to **ODN2** and **ODN5**, but this hardly explains the reduced stability. **ODN6** and **ODN2** have the same adjacent triplets (TAT-TAX-CGC), so they cannot play a role, either. The reason for the reduced binding affinity of **ODN6** compared to **ODN2** and **ODN5** remained unresolved.

When duplex 1 was targeted with **ODN2**, the neomycin-conjugated X*A-T triplet was next to T*A-T and C^+*G-C triplets. To study if the neighboring triplets affect the binding affinity of the neomycin-conjugated TFOs, the binding of **ODN9** to duplex 3 was next evaluated. In the case of **ODN9 +** duplex 3, the X*A-T triplet is neighbored by two T*A-T triplets. This replacement of the C^+*G-C triplet with T*A-T triplet decreased the stabilizing effect by 4.7 °C at pH 6.0 (ΔT_m^3 = +3.0 °C vs. +7.7 °C, entry 13 vs. 2, column C), but at pH 5.5, the decrease was marginal (ΔT_m^3 = +6.4 °C vs. +7.1 °C, entry 13 vs. 2, column F).

To evaluate whether the neomycin-conjugated TFOs can stabilize the triplex also over a Hoogsteen mismatch, the binding of TFOs to duplex 2 was studied.

Duplex 2 contains one thymidine interruption (T vs. A, cf. duplex 1). At pH 6.0, **ODN2** showed no stabilization, whereas at pH 5.5, **ODN2** stabilized the triplex by +4.8 °C (entry 9). **ODN5**, on the other hand, stabilized the triplex almost equally with duplex 2 and duplex 1 at pH 5.5 (+8.2 °C vs. +8.1 °C, entry 10 vs. entry 5). This equal stabilization was expected, because the neomycin conjugation site in **ODN5** is far from the mismatch region.

Table 1. UV thermal melting temperatures of the duplexes and triplexes.

Entry	Triplex	A	B	C	D	E	F
		pH 5.5, 0.1 mol L^{-1} NaCl			pH 6.0, 0.1 mol L^{-1} NaCl		
		T_m^2 (260nm)/ºC	T_m^3 (295nm)/ºC	ΔT_m^3 (295nm)/ºC	T_m^2 (260nm)/ºC	T_m^3 (295nm)/ºC	ΔT_m^3 (295nm)/ºC
1	duplex 1 + **ODN1**	51.7 (± 0.4)	41.3 (± 0.7)	–	53.0 (± 0.2)	28.7 (± 0.3)	–
2	duplex 1 + **ODN2**	51.2 (± 0.0)	48.6 (± 0.0)	+7.3[a]	52.2 (± 0.1)	36.4 (± 0.1)	+7.7[a]
3	duplex 1 + **ODN3**	50.6 (± 0.1)	41.7 (± 0.0)	+0.4[a]	53.5 (± 0.1)	29.2 (± 0.3)	+0.5[a]
4	duplex 1 + **ODN4**	50.1 (± 0.3)	41.4 (± 0.4)	+0.1[a]	53.5 (± 0.2)	27.4 (± 0.3)	-1.3[a]
5	duplex 1 + **ODN5**	51.3 (± 0.0)	49.4 (± 0.6)	+8.1[a]	53.4 (± 0.3)	36.2 (± 0.1)	+7.5[a]
6	duplex 1 + **ODN6**	51.3 (± 0.4)	41.4 (± 0.1)	+0.1[a]	53.8 (± 0.1)	27.8 (± 0.1)	-0.9[a]
7	duplex 1 + **ODN7**	n.d.	57.1 (± 0.2)	+15.8[a]	51.3 (± 0.2)	44.3 (± 0.6)	+15.6[a]
8	duplex 2 + **ODN1**	52.6 (± 0.3)	25.1 (± 0.1)	–	54.1 (± 0.1)	14.9 (± 0.0)	–
9	duplex 2 + **ODN2**	53.1 (± 0.3)	29.9 (± 0.3)	+4.8[b]	54.2 (± 0.2)	15.1 (± 0.0)	+0.2[b]
10	duplex 2 + **ODN5**	53.1 (± 0.1)	33.2 (± 0.3)	+8.2[b]	54.7 (± 0.1)	17.4 (± 0.1)	+2.3[b]
11	duplex 2 + **ODN7**	52.2 (± 0.3)	40.8 (± 0.2)	+15.7[b]	54.3 (± 0.1)	24.1 (± 0.4)	+9.2[b]
12	duplex 3 + **ODN8**	45.6 (± 1.1)	40.3 (± 0.2)	–	50.6 (± 0.2)	30.6 (± 0.1)	–
13	duplex 3 + **ODN9**	48.3 (± 0.3)	46.7 (± 0.3)	+6.4[c]	50.5 (± 0.1)	33.6 (± 0.1)	+3.0[c]
14	^{19}F-duplex1 + **ODN1**	n.d.	41.8 (± 0.1)	+0.5[a]	53.5 (± 0.4)	29.3 (± 0.0)	+0.6[a]
15	^{19}F-duplex2 + **ODN1**	50.5 (± 0.0)	41.4 (± 0.1)	+0.1[a]	53.2 (± 0.2)	28.3 (± 0.0)	-0.5[a]
16	^{19}F-duplex2 + **ODN2**	49.6 (± 0.3)	48.4 (± 0.2)	-0.2[d]	52.6 (± 0.1)	34.3 (± 0.2)	-2.1[d]
17	^{19}F-duplex3 + **ODN1**	n.d.	40.6 (± 0.1)	-0.7[a]	54.0 (± 0.1)	28.0 (± 0.0)	-0.7[a]
Entry	Triplex	pH 5.5, 0.3 mol L^{-1} NaCl			pH 6.0, 0.3 mol L^{-1} NaCl		
18	duplex 1 + **ODN1**	57.7 (± 0.2)	40.7 (± 0.2)	–	58.3 (± 0.5)	28.7 (± 0.1)	–
19	duplex 1 + **ODN2**	57.2 (± 0.3)	45.6 (± 0.0)	+4.9[e]	57.5 (± 0.1)	33.1 (± 0.5)	+4.4[e]
20	duplex 1 + **ODN7**	56.4 (± 0.3)	50.9 (± 1.3)	+10.2[e]	57.0 (± 0.2)	41.9 (± 0.5)	+13.2[e]

Conditions: 2 µmol L^{-1} solution of each oligonucleotide (entries 18–20: 1 µmol L^{-1}), 10 mmol L^{-1} sodium cacodylate. Notes: ΔT_m3 values compared to [a]duplex 1 + **ODN1** in 0.1 mol L^{-1} NaCl, [b]duplex 2 + **ODN1**, [c]duplex 3 + **ODN8**, [d]duplex 1 + **ODN2** and [e]duplex 1 + **ODN1** in 0.3 mol L^{-1} NaCl; n.d: T_m^2 and T_m^3 values could not be distinguished because of the overlapped duplex/triplex melting curves.

Finally, the binding of **ODN7**, containing two neomycin units, to duplex 1 and duplex 2 was studied. With duplex 1, the melting temperature was increased by +15.8 °C at pH 5.5 and +15.6 °C at pH 6.0 (entry 7). With duplex 2, the melting temperature was increased by +15.7 °C at pH 5.5 and +9.2 °C at pH 6.0 (entry 11). It may be worth noting that at pH 5.5, the melting temperature of **ODN7** + duplex 1 was even higher than the melting temperature of duplex 1 (57.1 °C vs. 51.7 °C, entry 7 vs. entry 1). In this case, the neomycin ligands hold all three sequences together, and the triplex is denaturated directly into single strands upon heating. Interestingly, while one neomycin unit at a Hoogsteen mismatch site (**ODN2** + duplex 2) could not stabilize the triplex at pH 6.0 (entry 9, column F), the same conjugation site in **ODN7** gave stabilization (cf. ΔT_m^3 values of **ODN7**: +9.2 °C and **ODN5**: +2.3 °C, entries 11 and 10, column F). In conclusion, the increased overall triplex stability seemed to facilitate the neomycin-promoted triplex stabilization in the mismatch region. The same phenomenon can be observed by comparing the ΔT_m^3 values of **ODN2** + duplex 2 at pH 5.5 and 6.0: +4.8 vs. +0.2 °C (entry 9).

The triplex formation in the present study requires slightly acidic conditions, because the triplex model used in this study contains nine C$^+$*G–C triplets. The C$^+$*G–C triplet requires protonation at the N3 site of cytosine (pK_a=4.5), and hence, the triplex formation remains marginal at neutral pH. Although the triplex melting temperatures are lower at pH 6.0 than at pH 5.5, **ODN7** increased the melting temperature equally at pH 5.5 and 6.0 (ΔT_m^3 = +15.8 °C and +15.6 °C, entry 7). Previous studies with discrete neomycin have shown that the triplex-stabilizing effect of neomycin decreases when the pH is increased from 7.2 to 8.1, as the pH approaches the pK_a values of the neomycin amines, but neomycin is still effective at pH 8.1.[115,125,126,134]

The triplex melting temperatures discussed above were obtained in a relatively low 0.1 mol L^{-1} salt concentration. The T_m^3 values with different TFOs are comparable to each other, but the stabilizing effect of the polycationic neomycin unit may be overemphasized in low salt concentrations. Previous studies with a similar triplex model at pH 5.5 and 0.3 mol L^{-1} NaCl have shown that simple *C*-5-(1-propargylamino) (see Figure 2), *C*-5-hydroxyprop-1-ynyl and *C*-5-prop-1-ynyl modifications already affect the triplex stability by ΔT_m^3 = +4.6, +2.6 and –1.4 °C, and two modifications increased the triplex melting temperature by ΔT_m^3 = +7.6, +5.3 and +3.2 °C, respectively.[69] To evaluate how the salt concentration affects the binding affinity of the neomycin-conjugated TFOs, the T_m^3 values of duplex 1 + **ODN1**, **ODN2** and **ODN7** were determined also in 0.3 mol L^{-1} NaCl concentration. The ΔT_m^3 values were lower compared to those in 0.1 mol L^{-1} NaCl: +5.0 °C (pH 5.5) and +4.4 °C (pH 6.0) with **ODN2** (entry 19) and +10.4 °C (pH 5.5) and +13.2 °C (pH 6.0) with **ODN7** (entry 20). Although part of the triplex

stabilization obtained by the neomycin units may result from the *C*-5 spacer alone, the ΔT_m[3] values with two neomycin units (**ODN7**) were remarkable.

3.1.3 ^{19}F NMR studies

To monitor the thermal unwinding of the triplexes in more detail, ^{19}F NMR spectroscopy was used. 4'-*C*-[(4-CF_3-1*H*-1,2,3-triazol-1-yl)methyl]thymidine[316] (**F1**, Figure 13) was inserted into three different sites of duplex 1 (^{19}F-duplex 1–3, Figure 13) and the triplex/duplex/single strand conversion was monitored by ^{19}F NMR spectroscopy. The **F1** sensor was previously developed in our laboratory.[316] The ^{19}F resonance shift of the sensor was highly sensitive to environmental changes in the secondary structures of oligonucleotides, and the sensor caused no effect on the DNA triplex stabilities.

In the present study, the ^{19}F NMR spectroscopic measurements were performed in solutions of 50 μmol L^{-1} of each oligonucleotide, 10 mmol L^{-1} sodium cacodylate (pH 5.5 or 6.0) and 0.1 mol L^{-1} NaCl in D_2O/H_2O (1:9, *v*/*v*). According to the UV and CD spectroscopic thermal melting profiles, ^{19}F-duplex 1–3 + **ODN1** all had equal thermal stability (Table 1, entries 14, 15 and 17). However, the behavior of the ^{19}F NMR resonance signals showed notable differences depending on the ^{19}F labeling site. ^{19}F-duplex 2 + **ODN1** and ^{19}F-duplex 3 + **ODN1** behaved in a very similar manner (Figure 16 b and c). In both cases, when 1.0 equivalent of **ODN1** was added to the ^{19}F-labeled duplex, the ^{19}F resonance signal of the duplex was fully replaced by a signal resulting from the triplex (at 30 °C, signals at 61.93 ppm and 62.17 ppm). When the temperature was increased, the signal of the triplex was gradually converted into the signal of the duplex and ultimately to the signal of the single strand. According to the relative peak areas of the ^{19}F resonance signals, the molar fractions of triplex, duplex and single strand could be determined at each temperature.

^{19}F-duplex 1 + **ODN1**, on the other hand, behaved in drastically different manner compared to ^{19}F-duplexes 2–3. The addition of 1.0 equivalent of **ODN1** to ^{19}F-duplex 1 gave two relatively broad signals in 1:3 ratio (*n*/*n*) at 30 °C (Figure 16 a, 61.80 ppm and 62.07 ppm). Moreover, when the temperature was increased, the duplex signal appeared very close to the larger triplex signal. In the case of ^{19}F duplexes 2–3, by contrast, the triplex and duplex signals were clearly distinguishable. Complete explanation for the different behavior of ^{19}F-duplex 1 remained unknown. Previous studies that used the same triplex sequence suggested that an *i*-motif (or unspecific C^+*C base pairing) competes with the triplex formation.[315] Interestingly, the ^{19}F label in ^{19}F-duplex 1 is in the same region as the *C*-5 neomycin conjugate in **ODN6**, the only *C*-5 neomycin modified TFO that was unable to stabilize the triplex (duplex 1 + **ODN6**, Table 1, entry 6). It is noteworthy

that all three ^{19}F-duplexes with **ODN1** gave nearly identical UV and CD thermal melting profiles and triplex melting temperatures. Therefore, the different ^{19}F NMR spectra of ^{19}F-duplexes 1–3 + **ODN1** illustrated different local environments in the triplex model. The modest binding affinity of **ODN6** may result from a twisted groove environment at the neomycin conjugation site.

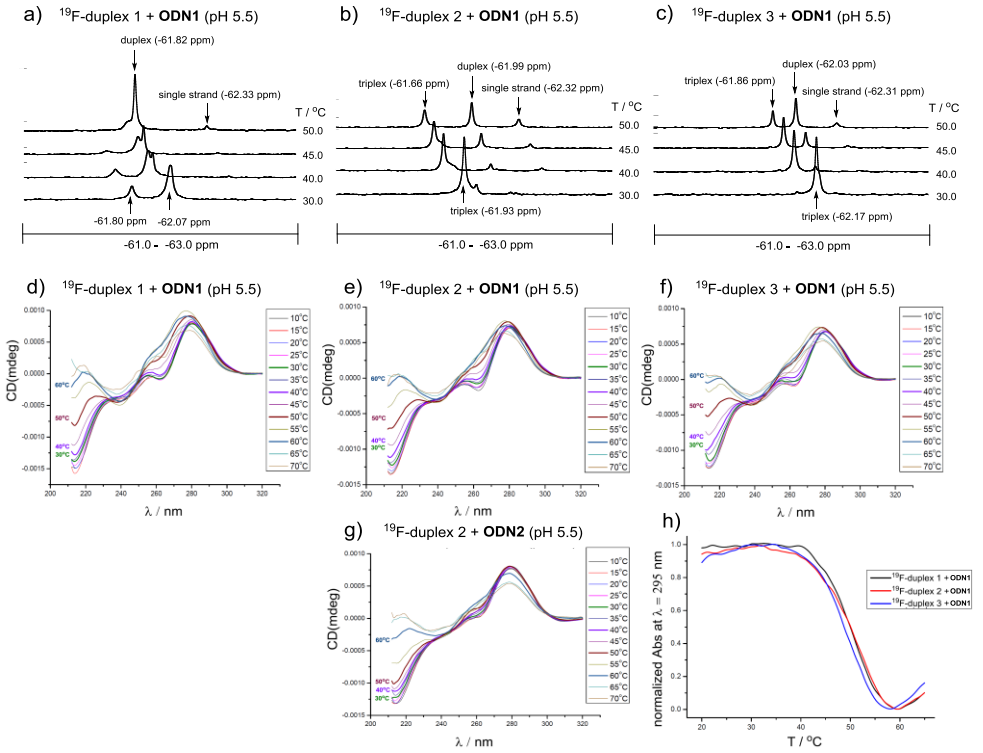


Figure 16. ^{19}F NMR spectra of ^{19}F-duplex 1–3 + **ODN1** at 30–50 °C (a–c), CD spectra at different temperatures (d–f) and UV thermal melting profiles at λ = 295 nm (h) from the same mixtures. CD spectra of ^{19}F-duplex 2 + **ODN2** (g) are also included in comparison to e.

Next, the binding of neomycin-conjugated **ODN2** to ^{19}F-duplex 2 was monitored by ^{19}F NMR spectroscopy at pH 5.5 and 6.0. The ^{19}F NMR spectra of ^{19}F-duplex 2 + **ODN1** and **ODN2** at different temperatures are shown in Figure 17 and the corresponding relative peak areas of the ^{19}F resonance signals as a function of temperature in Figure 18 a and b. In addition to the signals of triplex, duplex and single strand, small additional signals were also observed, probably related to minor secondary structures or aggregation. However, the relative peak areas could be adequately integrated. At pH 5.5, both **ODN1** and **ODN2** were fully Hoogsteen-hybridized, when the temperature was lower than 30 °C (Figure 18a). The ^{19}F

NMR based triplex denaturation curve for ^{19}F-duplex 2 + **ODN2** showed two inflection points (Figure 18a), whereas the UV thermal melting profile from the same mixture showed only one inflection point (Figure 18c). The latter inflection point of the ^{19}F NMR based triplex denaturation curve was close to the UV based $T_m{}^3$ value (ca. 58 °C, note the 25-fold oligonucleotide concentrations compared to the $T_m{}^3$ values in Table 1). According to the ^{19}F NMR measurements, this $T_m{}^3$ value resulted from the simultaneous denaturation of the Hoogsteen and Watson–Crick hydrogen bonds. Triplex was the predominant secondary structure throughout the whole denaturation range, 40–60 °C, although notable amount of duplex was also present at temperatures 45–57.5 °C.

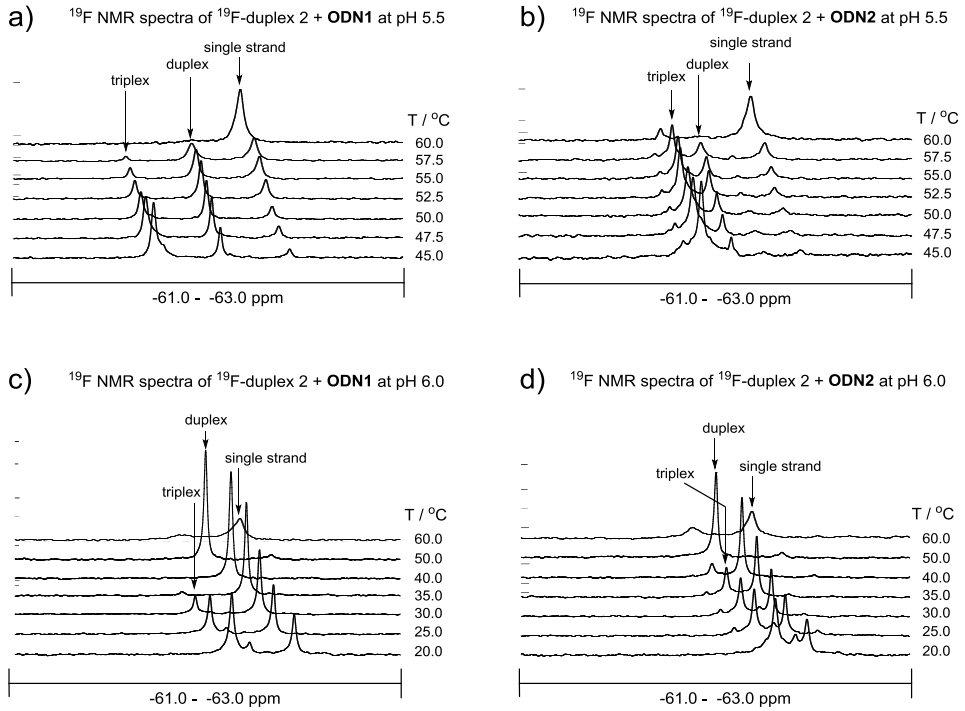


Figure 17. Triplex/duplex/single strand conversion of ^{19}F-duplex 2 + **ODN1** vs. **ODN2** followed by ^{19}F NMR spectroscopy. Conditions: 50 μmol L^{-1} of each oligonucleotide, 10 mmol L^{-1} sodium cacodylate (pH 5.5 or 6.0), 0.1 mol L^{-1} NaCl in D_2O/H_2O (1:9, *v*/*v*).

At pH 6.0, on the other hand, the triplex was not formed quantitatively after 1.0 equivalent of **ODN1** and **ODN2** was added to ^{19}F-duplex 2, but the molar fraction of the triplex was ca. 55% and 65% with **ODN1** and **ODN2**, respectively (Figure 18b). Duplex became the predominant secondary structure at temperatures over 25 °C with **ODN1** and at temperatures over 30 ° with **ODN2**. The molar fractions of

different secondary structures remain hidden in UV and CD spectrophotometric data. Therefore, at the spectrophotometry based melting points (T_m), the equilibrium constant K may be far from the general definition: $K = C_{init}/2$, where C_{init} = initial oligonucleotide concentration. For example, at the inflection points of the triplex melting curves at pH 6.0 (T_m^3, Figure 18b), the molar fraction of the duplex is much higher than the molar fraction of the triplex, and thus, $K < C_{init}/2$.

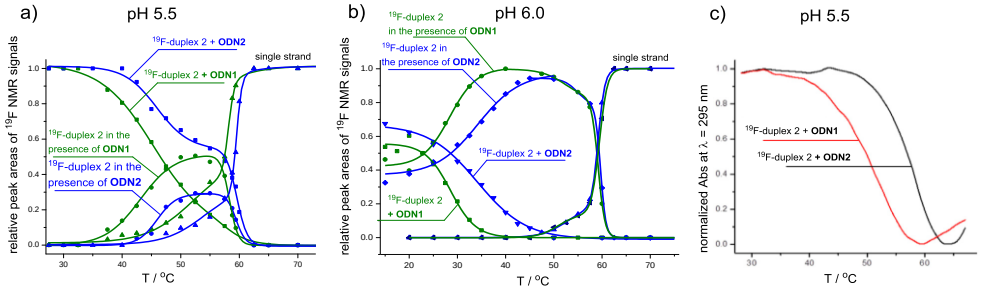


Figure 18. a–b) The relative molar fractions of triplexes, duplexes and single strands upon thermal denaturation obtained by ^{19}F NMR spectroscopy (^{19}F-duplex 2 + **ODN1** vs. **ODN2**, cf. Figure 17). c) The normalized UV melting profiles of the ^{19}F NMR samples. Conditions: 50 µmol L^{-1} of each oligonucleotide, 10 mmol L^{-1} sodium cacodylate (pH 5.5 or 6.0), 0.1 mol L^{-1} NaCl in D_2O/H_2O (1:9, *v/v*).

3.2 Recognition of the nucleobase content in the target DNA duplex by ^{19}F-labeled TFO probes

In the second section of this thesis, ^{19}F-labeled TFO sensors capable of recognizing variable nucleobases in the pyrimidine strand of the target DNA duplex were developed (**ONF 1–3**, Figure 19). Before this study, no reliable method for sensing the pyrimidine strand in double-helical DNA was reported. For the synthesis of **ONF 1–3**, phosphoramidite building blocks of **F1**–**3** (Figure 19a) were synthesized and incorporated into the TFOs by automated DNA synthesis. Sensors **F2** and **F3** are CF_3 analogs of a previously reported triplet-forming nucleobase modification 2'-deoxy-*N*^{4}-(6-aminopyridin-2-yl)cytidine,[322] whereas sensor **F1** was previously developed in our laboratory.[316] The TFO probes **ONF1–3** were targeted to the purine strand of the target duplex (**ODN11** and **ODN12**), and the nucleobase content in the pyrimidine stand of the target duplex was varied (**ODN10**, X = T, A, C or G). **ONF 1–3** were expected to give different ^{19}F NMR shifts depending on the variable nucleobases in the target duplex. In the case of **ONF1**, the plausible ^{19}F NMR shift discrimination was expected to be based on the changes in the local environment of the Watson–Hoogsteen groove. **ONF 2–3**, on the other hand, were expected to recognize the whole base pair more specifically (Figure 19b). The ^{19}F NMR shift of sensors **2** and **3** was assumed to be dependent on the spatial

orientation of the CF_3 group, the possible existence or absence of the extra hydrogen bonds, intercalation in the triplet and potential changes in the steric environment when the nucleobases in the target duplex were varied.

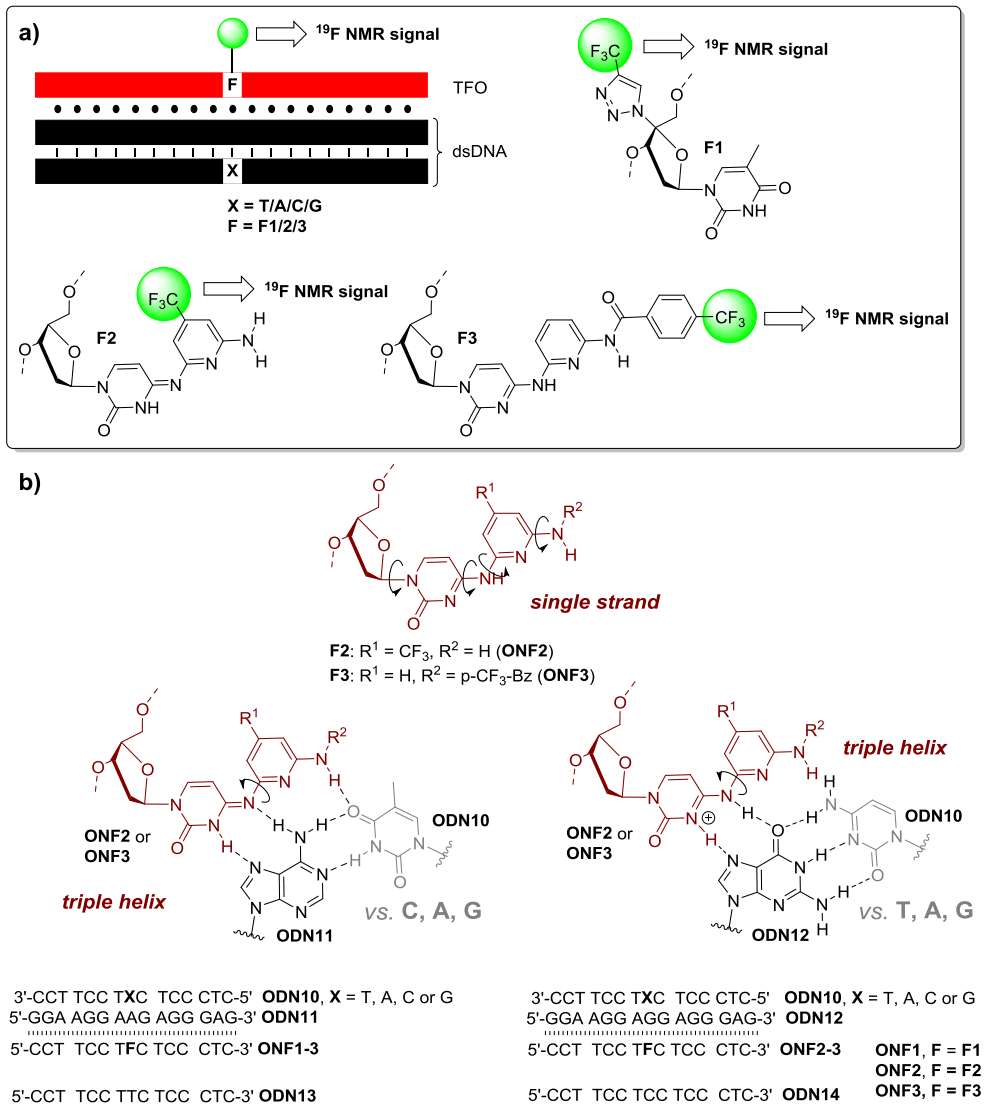
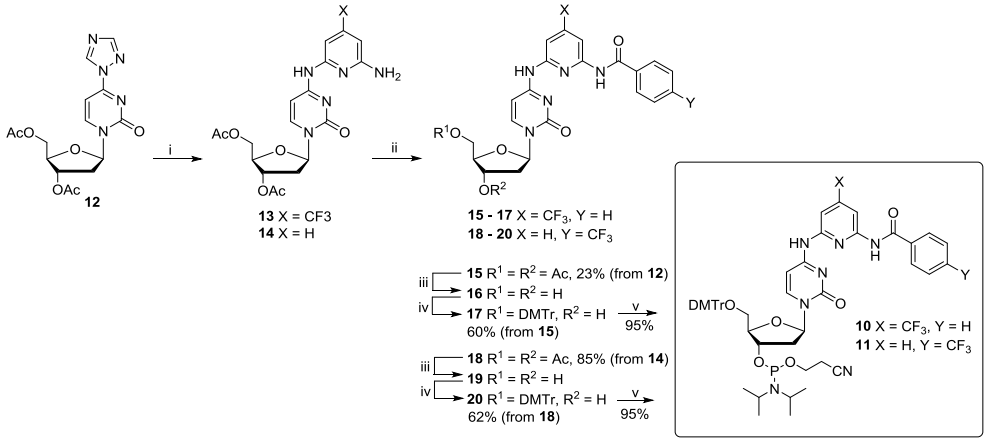


Figure 19. a) The principle of the ^{19}F-labeled TFO probes and the structures of sensors **F1–3**. b) The triple-helical models and the proposed hydrogen bonding of sensors **F2** and **F3** including tautomers and rotamers.

3.2.1 Synthesis of the ^{19}F-labeled TFO probes

3.2.1.1 Synthesis of the Phosphoramidite Building Blocks of Sensors F2 and F3

The synthesis of the phosphoramidites **10** and **11** is presented in Scheme 3. The previously reported[322] 4-triazolyl-2'-deoxyuridine (**12**) was first exposed to transamination with 2,2-diaminopyridines to yield compounds **13** and **14**. Compound **13** was then protected with benzoyl chloride to give compound **15** and compound **14** was protected with 4-trifluoromethylbenzoic acid to give compound **18**. The 3'- and 5'-O-acetyl groups were selectively removed with a gentle treatment with ammonia and the 5'-OH group was 4,4'-dimethoxytritylated to give **17** and **20**. Finally, the 3'-OH group was phosphitylated with 2-cyanoethyl *N*,*N*-diisopropylphosphoramidochloridite to give phosphoramidites **10** and **11**.



Scheme 3. (i) 4-trifluoromethyl-2,6-diaminopyridine in Py for **13** and 2,6-diaminopyridine in Py for **14**; (ii) Benzoyl chloride in Py for **15** and 4-trifluoromethylbenzoic acid, PyBOP, DIEA, DMF for **18**; (iii) NH_3 (aq), methanol; (iv) DMTrCl in Py; (v) 2-Cyanoethyl N,N-diisopropylchlorophosphoramidite, Et_3N, DCM.

3.2.1.2 Oligonucleotide synthesis

The ^{19}F-labeled oligonucleotide probes **ONF1**–**ONF3** were synthesized on solid support on a 1.0 μmol scale using an automatic DNA/RNA synthesizer. **ONF1** was synthesized by introducing the phosphoramidite building block of 4'-*C*-[(4-CF_3-1*H*-1,2,3- triazol-1-yl)methyl]thymidine (**F1**) as described previously by our group.[316] **ONF2** was synthesized on a standard dC-3'-*O*-yl succinate loaded LCAA-CPG support and **ONF3** on a dC-3'-*O*-yl hydroquinone *O*,*O*-diacetate ("Q-linker") loaded LCAA-CPG support. The phosphoramidite building blocks **10** and

11 were introduced into the oligonucleotides using coupling time of 600s. The coupling yields of the building blocks **10** and **11** were > 97%, comparable to commercial phosphoramidites. After the oligonucleotide synthesis, **ONF2** was released from the solid support by concentrated ammonia (overnight at 55 °C) and **ONF3** by 0.05 M K_2CO_3 in methanol. This gentle release from the Q-linker leaved the 4-trifluoromethylbenzoyl group of **ONF3** intact. On the other hand, the benzoyl protection of the 4-trifluoromethyl-2,6-diaminopyridine moiety of **ONF2** was surprisingly stable. To remove the benzoyl protection, **ONF2** was, after cleavage from the solid support, further treated with a mixture of 40% aqueous methylamine and concentrated ammonia (1:1, *v*/*v*, 5 h at 55 °C). Finally, **ONF2** and **ONF3** were purified by RP HPLC and lyophilized into white powders. The authenticity of the oligonucleotides was verified by MS (ESI-TOF) spectroscopy. According to UV absorbance at 260 nm, the isolated yields of **ONF2** and **ONF3** were 18% and 25%, respectively.

3.2.2 UV thermal melting studies

First, the effect of the sensors **F1–3** on the thermal stability of the triplexes (**ODN10/ODN11/ONF1–3** and **ODN10/ODN12/ONF2–3**) was studied by UV melting profile analysis. Sensors **F2** and **F3** are CF_3 analogs of previously reported 2'-deoxy-N^4-(6-aminopyridin-2-yl)cytidine[322] that was originally designed for the recognition of the C–G base pair interruption in the purine strand of dsDNA target. However, binding to A–T and, to a smaller extent, to G–C base pairs was also reported. Therefore, sensors **F2** and **F3** were expected to recognize A–T and G–C base pairs, depending on the tautomers (cf. Figure 19b). In the **ODN10/ODN11/ONF1–3** triplexes, sensors **F1–3** were aimed to bind to the adenine residue of **ODN11**, and in the **ODN10/ODN12/ONF2–3** triplexes, sensors **F2** and **F3** were aimed to bind to the guanine residue of **ODN12**. The UV thermal melting temperatures of the triplexes were determined at pH 5.5 using 2.0 μmol L^{-1} oligonucleotide concentration in 10 mmol L^{-1} sodium cacodylate and 0.1 mol L^{-1} NaCl in H_2O, detection wavelength λ = 260 nm and 295 nm. The ΔT_m^3 values given in parentheses are those compared to the T_m^3 values of the unmodified triplexes **ODN10/ODN11/ODN13** and **ODN10/ODN12/ODN14**.

As seen in Table 2, sensor **F1** (**ONF1**) reduced the triplex stability to a small extent (ΔT_m^3 = -2.4 – -3.0°C compared to the unmodified **ODN10**/**ODN11**/**ODN13** triplex), whereas sensor **F2** reduced the stability remarkably (ΔT_m^3 = -5.8 – -14.3°C compared to **ODN10**/**ODN11**/**ODN13** triplex and ΔT_m^3 = -9.6 – -16.8°C compared to **ODN10**/**ODN12**/**ODN14** triplex). Sensor **F3** also reduced the triplex stability in most cases (ΔT_m^3 = -3.6 – -7.2°C compared to **ODN10**/**ODN11**/**ODN13** and ΔT_m^3 = -1.8 – -8.1°C compared to **ODN10**/**ODN12**/**ODN14**). Interestingly, in the case of

ODN10/ODN11/ONF3 with a **F3***A-A triad (X = A in **ON10**), a small triplex stabilization was observed (ΔT_m^3 = +1.2°C). However, it may be concluded that sensors **2** and **3** are probably unable to bind through the optimal hydrogen bonding patterns suggested in Figure 19. It is worth noting that **ONF2** and **ONF3** single strands may exist as several tautomers and rotamers, but the triple-helical structure reduces the structural mobility. However, the rotation around the N^4(Cyt)–C^2(Py) σ-bond may be relatively easy, thus preventing the hydrogen bonding between the sensor and O^4(Thy) (Figure 19b), for instance. Additionally, the protonation of the sensors, particularly the protonation of the pyridine residue, may play important role in the binding. Compared to sensor **F2**, sensor **F3** caused less decrease in the triplex stability. It is possible that the 4-CF_3-benzoyl moiety of **F3** compensated the reduced triplex stability by intercalation. Sensors **F2** and **F3** showed no notable selectivity between binding to adenine and guanine bases. This observation supports the idea of the alternating tautomers of **F2** and **F3**, although the triplex stability was decreased in both cases.

Table 2. UV thermal melting temperatures of the triplexes.

duplex	**ODN13**	**ODN14**	**ONF1**	**ONF2**	**ONF3**
			T_m^3 / °C		
ODN10/ODN11					
X = T	40.8 ± 0.4		38.4 ± 0.2 (–2.4)	26.5 ± 0.4 (–14.3)	33.6 ± 1.2 (–7.2)
X = A	36.3 ± 0.2		33.3 ± 0.3 (–3.0)	30.5 ± 0.3 (–5.8)	37.5 ± 0.5 (+1.2)
X = C	37.9 ± 0.1		34.9 ± 0.1 (–3.0)	25.2 ± 0.4 (–12.7)	32.3 ± 0.7 (–5.6)
X = G	36.9 ± 0.2		33.9 ± 0.3 (–3.0)	23.3 ± 0.6 (–13.6)	34.3 ± 1.2 (–3.6)
ODN10/ODN12					
X = T		40.1 ± 0.5		29.6 ± 0.3 (–10.5)	36.3 ± 0.6 (–3.8)
X = A		40.3 ± 0.1		29.7 ± 0.2 (–10.6)	33.9 ± 0.5 (–6.4)
X = C		42.5 ± 0.4		25.7 ± 0.4 (–16.8)	34.4 ± 0.6 (–8.1)
X = G		38.8 ± 0.2		29.2 ± 0.2 (–9.6)	37.0 ± 0.6 (–1.8)

Conditions: 2.0 μmol L^{-1} of each oligonucleotide, 10 mmol L^{-1} sodium cacodylate (pH 5.5), 0.1 mol L^{-1} NaCl in H_2O, detection wavelength 295 nm. Notes: ΔT_m^3 values given in parentheses are those compared to the T_m^3 values of the unmodified triplexes **ODN10/ODN11/ODN13** and **ODN10/ODN12/ODN14**.

3.2.3 ^{19}F NMR studies

Next, the ability of sensors **F1**, **F2** and **F3** to recognize variable nucleobase content in the pyrimidine strand of the target duplexes was evaluated. The ^{19}F NMR spectroscopic measurements were performed using 10 μmol L^{-1} oligonucleotide concentration in a mixture of 10 mmol L^{-1} sodium cacodylate (pH 5.5) and

0.1 mol L^{-1} NaCl in D_2O/H_2O (1:9, *v*/*v*) at 20 °C. First, the applicability of sensor **F1** (**ONF1**) was examined (Figure 20a). **ONF1** single strand gave a relatively sharp ^{19}F resonance signal at –62.73 ppm. When 1 equivalent of the purine strand **ODN11** was added, the original signal was replaced by several scattered signals, indicating unspecific interactions between **ONF1** and **ODN11**. Adding 1 equivalent of the pyrimidine strand **ODN10** resulted in one, relatively sharp ^{19}F resonance signal. With the complementary **ODN10/ODN11/ONF1** triplex (X = T in **ODN10**, forming a **F1*A-**T triplet), the chemical shift of this signal was –61.87 ppm. The non-matching **F1*A-**C triplet (X = C in **ODN10**) gave a signal at exactly the same chemical shift. The **F1*A-**A triplet (X = A in **ODN10**) gave a signal with a slightly different chemical shift (–61.83 ppm), whereas in the case of the **F1*A-**G triplet (X = G in **ODN10**), the change in the chemical shift was notable (–61.54 ppm). Taken together, sensor **F1** was able to discriminate the adenine-pyrimidine base pairs (**A-**T and **A-**C) from adenine-purine base pairs (**A-**A and **A-**G), although the discrimination between the **A-**A base pair and adenine-pyrimidine base pairs was modest.

Next, the same measurements were performed with sensors **F2** and **F3** (**ONF2** and **ONF3**), which were expected to recognize the variable nucleobase content more specifically. **ONF2** single strand gave a ^{19}F resonance signal at –65.11 ppm (Figure 20b). This signal was considerably broader compared to **ONF1**. When 1 equivalent of **ODN11** was added, a new broad signal at –64.70 ppm was observed. Adding 1 equivalent of **ODN10** resulted in a unique chemical shift for each **ODN10/ODN11/ONF2** triplex (X = T, A, C or G in **ODN10**). All the signals were relatively broad. However, all the signals could be distinguished from a mixture containing all the four **ODN10/ODN11/ONF2** triplexes (X = T, A, C or G in **ODN10**; see Figure 22a), although the signals of **F2*A-**A and **F2*A-**G overlapped partially.

ONF3 single strand gave two distinct ^{19}F resonance signals, major one at –62.68 ppm and minor one at –62.20 ppm (Figure 20c). The distinct signals most likely originate from rotamers of **F3**. The addition of the duplex **ODN10/ODN11** resulted in unique signals with the **F3*A-**T triplet, with the **F3*A-**A triplet and with the **F3*A-**C triplet, but the **F3*A-**A triplet and the **F3*A-**G triplet gave the same chemical shift. In other words, sensor **F3** was unable to discriminate purine bases from each other. The ^{19}F NMR spectrum of a mixture containing all the four **ODN10/ODN11/ONF3** triplexes (X = T, A, C or G in **ODN10**) is shown in Figure 22b.

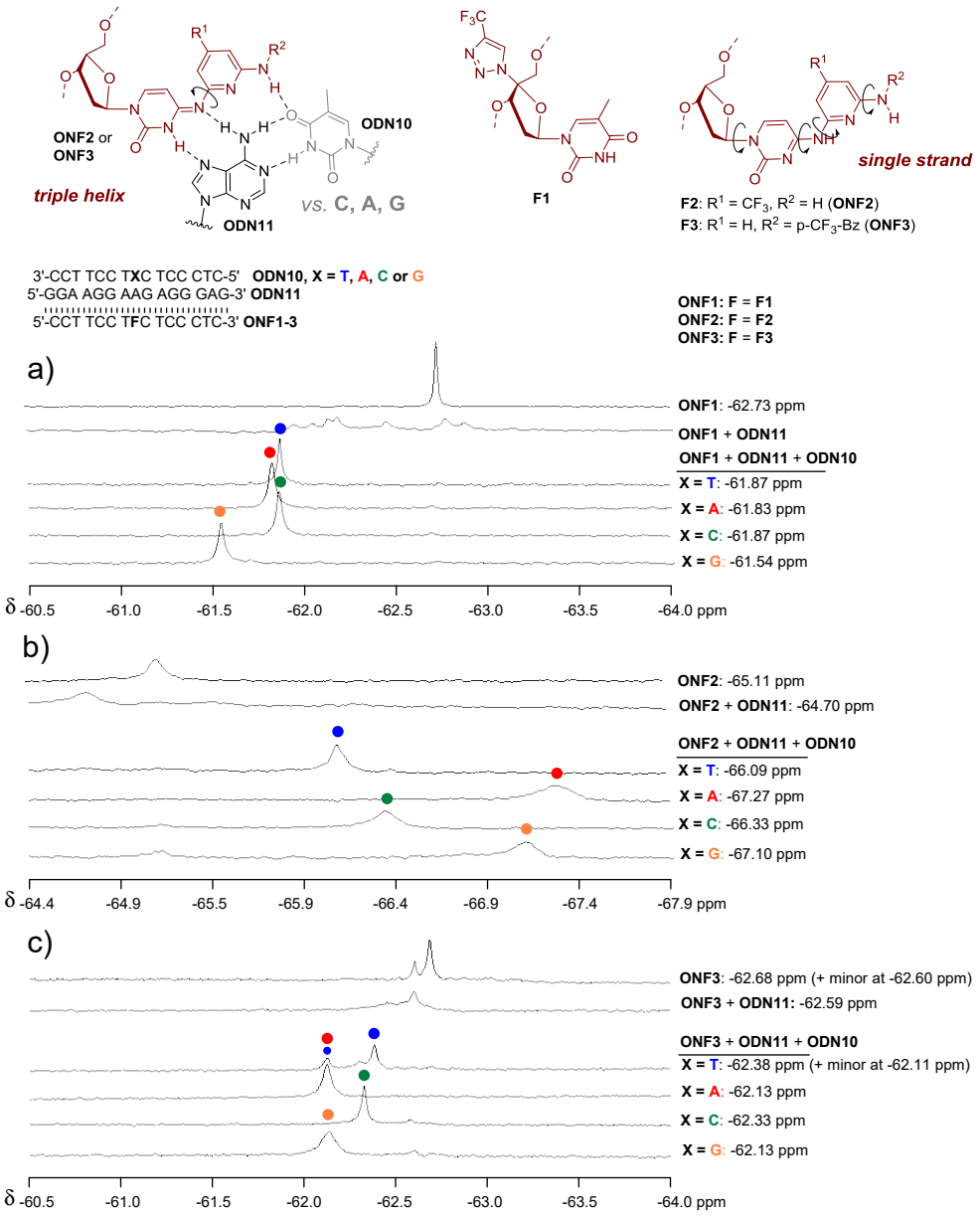


Figure 20. ^{19}F NMR spectra of the **ODN10/ODN11/ONF1–3** triplexes (**F*A**-X triplets, X = T, A, C or G). Conditions: 10 μmol L^{-1} of each ON, 10 mmol L^{-1} sodium cacodylate (pH 5.5), 0.1 mol L^{-1} NaCl in D_2O/H_2O (1:9, *v*/*v*) at 20 °C.

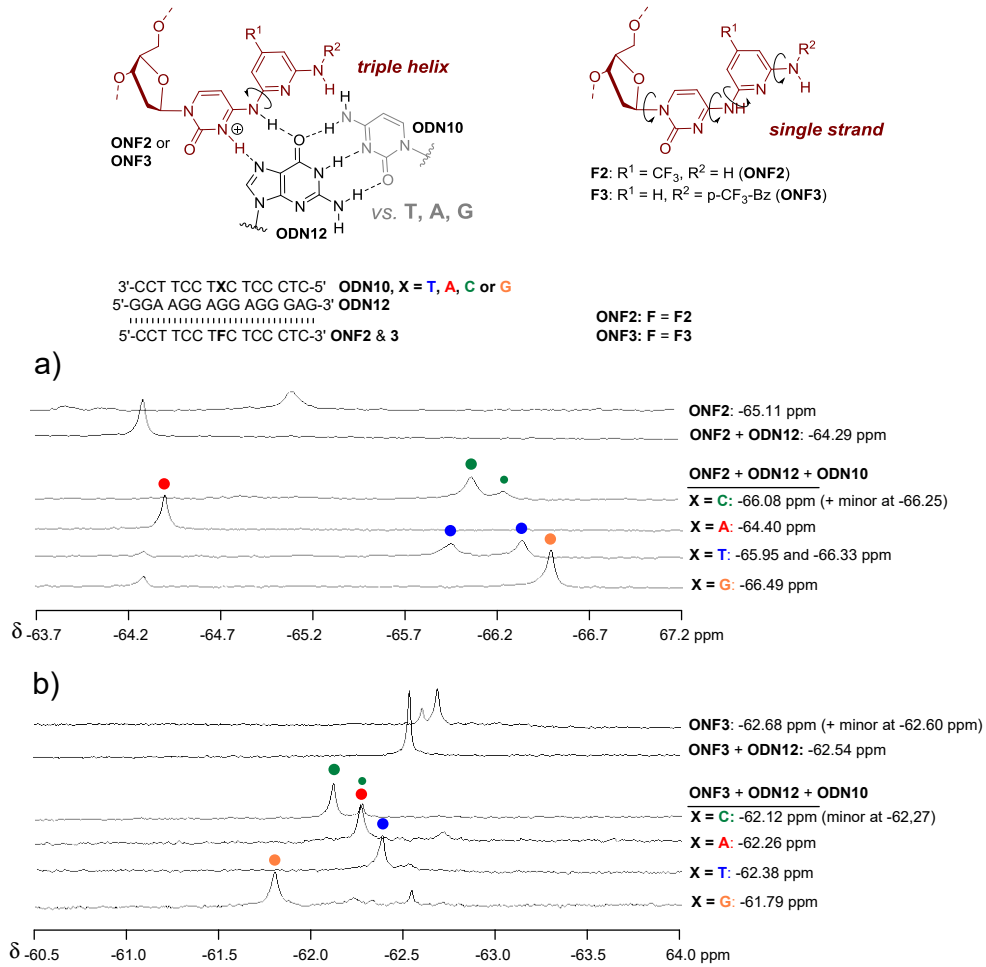


Figure 21. ^{19}F NMR spectra of the **ODN10/ODN12/ONF2–3** triplexes (**F*G**-X triplets, X = T, A, C or G). Conditions: 10 μmol L^{-1} of each ON, 10 mmol L^{-1} sodium cacodylate (pH 5.5), 0.1 mol L^{-1} NaCl in D_2O/H_2O (1:9, *v*/*v*) at 20 °C.

Depending on the tautomers, sensors **F2** and **F3** may also bind to the guanine base in the target duplex. Therefore, the ability of **ONF2** and **ONF3** to recognize variable nucleobase content in **ODN10**/**ODN12** duplex (X = T, A, C or G in **ODN10**) was next studied (Figure 21). Interestingly, both **ONF2** and **ONF3** gave a relatively sharp ^{19}F resonance signal after the addition of the purine strand **ODN12**. When **ODN10** was added to the mixture of **ONF2** and **ODN12**, unique chemical shifts were observed for each **ODN10/ODN12/ONF2** triplex (Figure 21a). As seen, the **F2*G**-purine triplets gave one, relatively sharp signal, whereas the **F2*G**-pyrimidine triplets gave two resonance signals. Consequently, sensor **F2** targeted to **A**-T and **A**-C base pairs may adopt two relatively stable

conformers, thus resulting in distinct signals. These split signals also disturbed the ^{19}F NMR analysis of a mixture containing all the four **ODN10/ODN12/ONF2** triplexes (Figure 22c). However, all the signals could be distinguished also from this mixture.

Among these experiments, sensor **F3** (**ONF3**) with **ODN10**/**ODN12** duplex showed the best ^{19}F NMR spectroscopic results. Unique chemical shifts were observed for each **ODN10/ODN12/ONF3** triplex (Figure 21b). The signals could also be distinguished from a mixture containing all the four **ODN10/ODN12/ONF3** triplexes (Figure 22d). When sensors **F2** and **F3** are compared to each other, the 4-trifluoromethylbenzoyl group of **F3** seems to be beneficial for the recognition. The 2,6-diaminopyridine moiety of sensor **F3** is probably prone to adopting alternative conformers, thus resulting in two distinct ^{19}F resonance signals when targeted to **G**-pyrimidine base pairs.

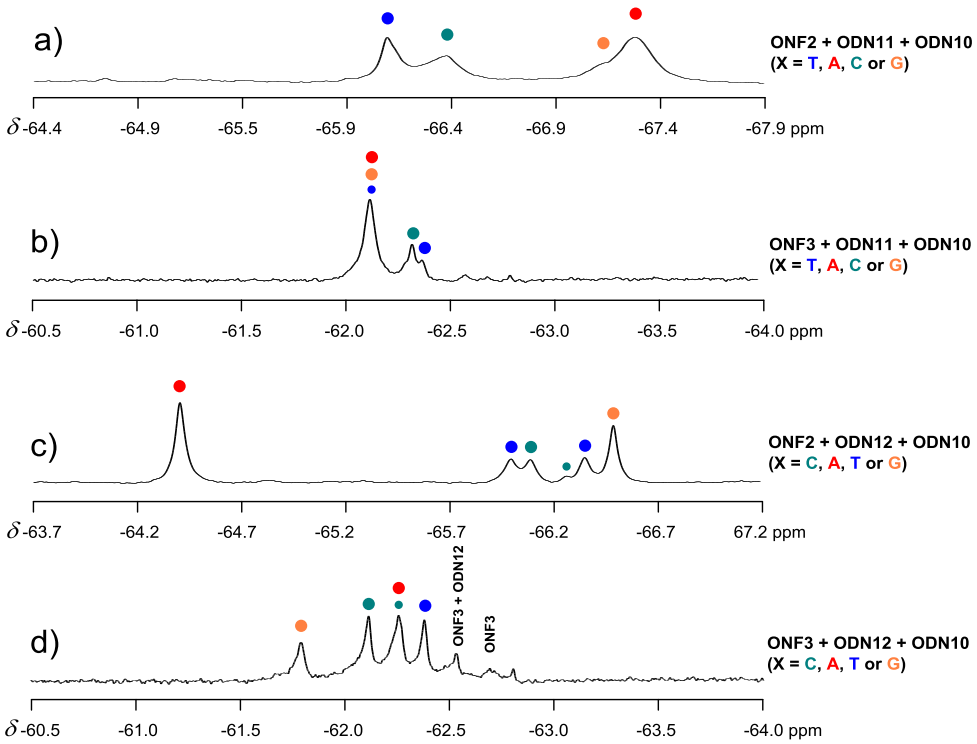


Figure 22. ^{19}F NMR spectra of the mixtures of **ODN10/ODN11/ONF2**, **ODN10/ODN11/ONF3**, **ODN10/ODN12/ONF2** and **ODN10/ODN12/ONF3** triplexes. Conditions: 10 μmol L^{-1} of each **ON10**, 40 μmol L^{-1} of **ODN11**, **ODN12**, **ONF2** and **ONF3**, 10 mmol L^{-1} sodium cacodylate (pH 5.5), 0.1 mol L^{-1} NaCl in D_2O/H_2O (1:9, *v*/*v*) at 20 °C.

3.3 Recognition of double-helical microRNA by TFPNAs

In the third section of this thesis, ^{19}F NMR spectroscopy was used to obtain new information on the different binding mechanisms by which triplex-forming PNAs (**PNAs 1**–**10,** Figure 23) bind to double-stranded regions of RNA hairpins. For this purpose, ^{19}F sensors, previously developed in our laboratory, were incorporated to two target RNA hairpins, an AU-rich duplex with a hexaethylene glycol loop and a biologically relevant miR-215 model (**ONR1** and **ORN2**, Figure 23). The ^{19}F NMR spectroscopic analysis revealed detailed information on the stoichiometry and transition between two alternative binding modes, 1:1 PNA/RNA triplex and 2:1 $(PNA)_2$/RNA triplex invasion complex. Moreover, aiming to enhance the binding affinity and selectivity for 1:1 PNA/RNA triplex formation, chiral γ-(*R*)-hydroxymethyl and γ-(*S*)-guanidinylmethyl modifications were introduced to the TFPNAs (**PNAs 2**, **3** and **5**–**9**, Figure 23).

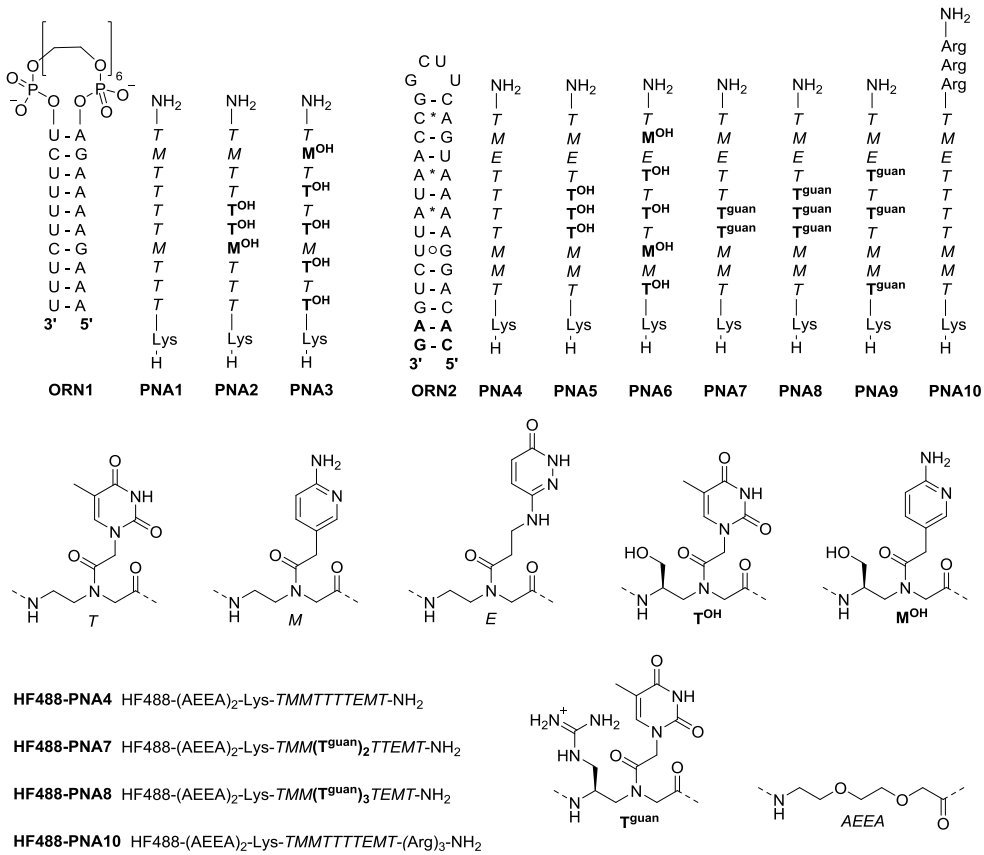


Figure 23. The RNA hairpin targets and the structures of the PNAs studied herein.

3.3.1 Synthesis of PNAs

3.3.1.1 PNA monomers

The modified PNA monomers 2**1**–**25** used for the synthesis of TFPNAs are outlined in Figure 24. 2-(6-Aminopyridin-3-yl)acetate (**21**) and 2-[(6-oxo-1,6-dihydropyridazin-3-yl)amino]acetate (**22**) derived monomers were previously described by the groups of Rozners and Nielsen, but with different protecting groups.[25,323] In the monomer **21**, trityl (Trt) was used as protecting group for the N6 side chain instead of *tert*-butyloxycarbonyl (Boc) that was applied in the report by Zengeya et al.[25] Monomer **22** differs from the previously reported E monomers[24,323] by its p-methoxybenzyl (PMB) protection, which increased the solubility in *N*-methyl-2-pyrrolidone (NMP). The syntheses of 2**1**–2**5** are presented in Scheme 4. For the synthesis of monomer **21**, 2-(6-aminopyridin-3-yl)acetic acid (**26**) was N^6-tritylated through trimethylsilyl ether (TMS) transient protection and coupled to **28**.[324] Finally, Pd^0 catalyzed allyl group removal gave monomer **21** in 62% overall yield (from **26**, 3 steps). For the synthesis of monomer **22**, compound **30** was converted into compound **31** by nucleophilic aromatic substitution using *p*-methoxybenzyl alcoholate in *p*-methoxybenzyl alcohol at elevated temperature. Compound **31** and compound **28** were then coupled using BOP as coupling reagent, and Pd^0 catalyzed allyl removal gave monomer **22** in 20 % overall yield (from **30**, 3 steps).

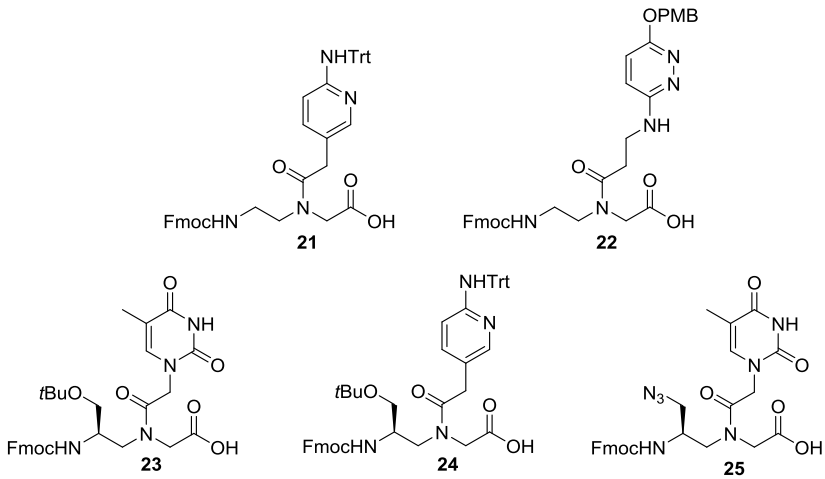
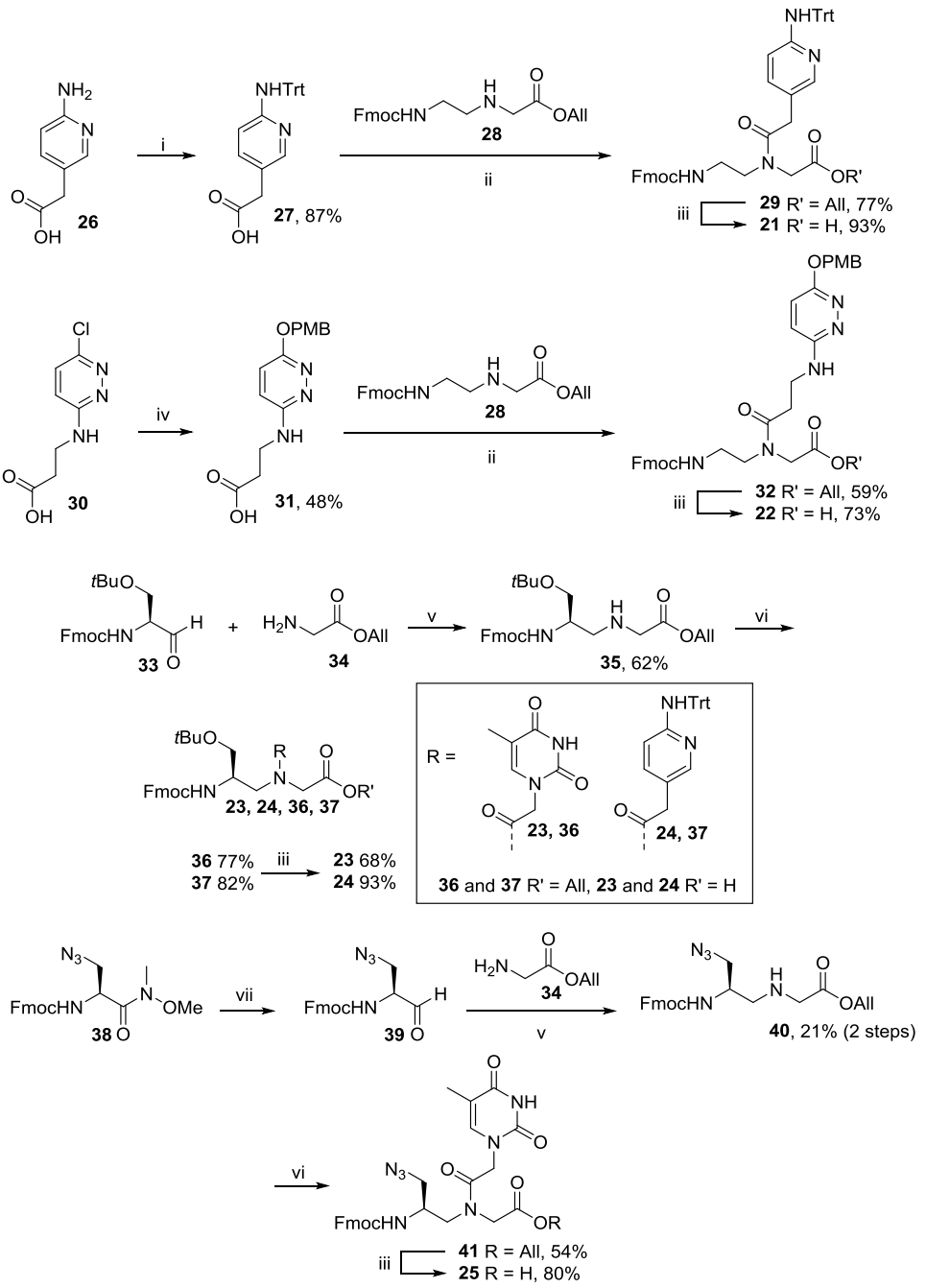


Figure 24. Modified PNA building blocks **21**–**25** used for the synthesis of TFPNAs.

For the synthesis of the γ-(*R*)-*tert*-butyloxymethyl modified monomers **23** and **24**, allyl *N*-[(*R*)-3-*tert*-butoxy-N^2-Fmoc-2-aminoprop-1-yl]glycinate (**35**) was

synthesized by reductive amination between freshly prepared (*S*)-2-amino-3-hydroxypropanal (**33**, "(*S*)-*N*-Fmoc-*O*-*t*Bu-serinaldehyde") and allyl glycinate (**34**). The optical purity of **35** was confirmed by chiral HPLC analysis. Thereafter, compound **35** was coupled with either thymine-1-yl acetic acid or 2-[6-(tritylamino)pyridine-3-yl]acetic acid (**27**) using BOP as coupling reagent, and finally Pd^0 catalyzed allyl removal gave monomers **23** and **24** in 32 and 47 % overall yields, respectively (from **35**, 3 steps). For the synthesis of the γ-(*R*)-azidomethyl modified monomer **25**, compound **39** was first synthesized from the Weinreb amide **38**[325] by reduction with $LiAlH_4$, and reductive amination between freshly prepared **39** and allyl glycinate (**34**) then yielded compound **40**. The optical purity of **40** was confirmed by chiral HPLC analysis. Thereafter, compound **40** and thymine-1-yl acetic acid were coupled using BOP as coupling reagent, and finally Pd^0 catalyzed allyl removal gave the monomer **25** in 9 % overall yield (from **38**, 4 steps). All the monomers **21**–**25** were readily soluble in NMP to provide 0.25 mol L^{-1} solutions needed for the peptide synthesizer.



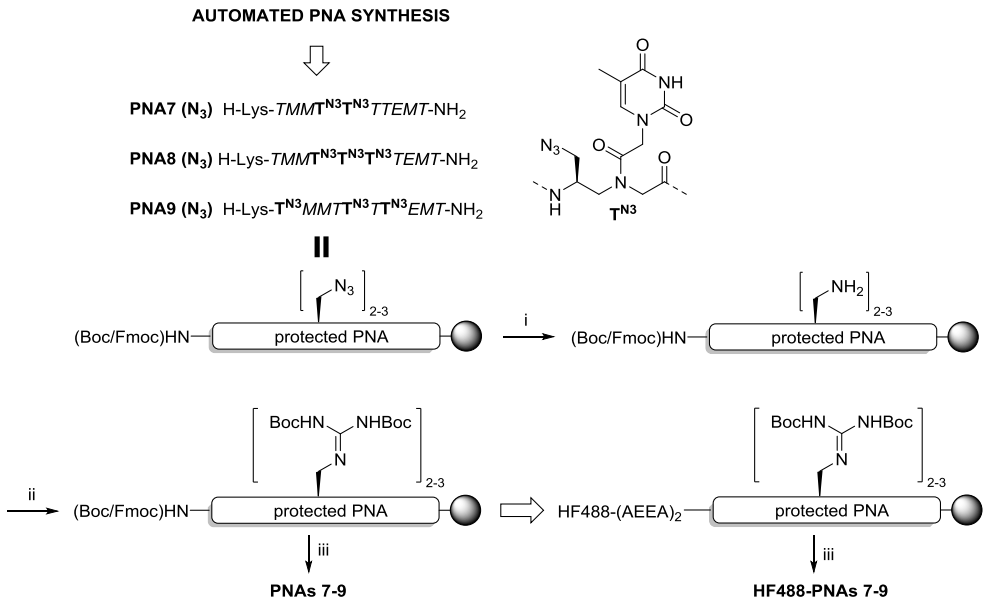
Scheme 4. (i) 1) Me_3SiCl, Py; 2) TrtCl, Py; 3) H_2O; (ii) BOP, DIEA, DMF; (iii) [$Pd(PPh_3)_4$], phenylsilane, THF; (iv) NaH, *p*-methoxybenzyl alcohol; (v) $NaBH_3CN$, AcOH/MeOH (1:99 *v*/*v*); (vi) thymine-1-yl acetic acid or 2-[6-(tritylamino)pyridine-3-yl]acetic acid (**27**), BOP, DIPEA, DMF; (vii) $LiAlH_4$, THF.

3.3.1.2 Synthesis of TFPNAs and ORNs

PNAs 1–**10** (Figure 23) were synthesized on a 10 μmol scale by automated peptide synthesizer. **PNA1** was previously described by Zengeya et al.[25] In the PNA synthesis, the modified PNA monomers 2**1**–2**5** and commercially available Fmoc-PNA-T-OH and amino acid building blocks (Boc-Lys(Fmoc)-OH, Fmoc-Arg(Pbf)-OH, Fmoc-Lys(Fmoc)-OH) were incorporated into the PNA sequences using Fmoc/*t*-Bu peptide synthesis cycle. After synthesis, the solid-supported PNAs were released with a mixture of anisole and trifluoroacetic acid (TFA), precipitated in Et_2O and purified by RP HPLC. The homogeneity of the products was confirmed by RP HPLC, and the authenticity was verified by ESI-TOF MS.

For the synthesis of γ-(*S*)-guanidinylmethyl modified **PNAs 7–9**, γ-(*R*)-azidomethyl modified **PNAs 7–9 (N_3)** were first synthesized by automated PNA synthesis (Scheme 5). The γ-(*R*)-azidomethyl groups were then reduced into γ-(*S*)-aminomethyl groups and guanidinylated on solid support. The reduction step was carried out by treatment with Me_3P and the guanidinylation step by treatment with 1,3-di-Boc-2-(trifluoromethylsulfonyl)-guanidine. For the synthesis of the HiLyte Fluor 488 (HF488) labeled PNAs (**HF488-PNA4** & **HF488-PNAs 7–10**, Figure 23), two 2-(2-aminoethoxy)ethoxy acetic acid (AEEA) spacers were attached to the amino terminus of the PNAs and the HF488 label was coupled following a previously described procedure.[326] More details on the PNA synthesis are described in the Experimental section.

The RNA hairpin targets **ORN1** and **ORN2** (Figure 23) and the ^{19}F-labeled oligoribonucleotides **ORN3–8** were synthesized on a 1.0 μmol scale by automatic DNA/RNA synthesizer. The ^{19}F probes were incorporated using 2'-O-[(4-CF_3-triazol-1-yl)methyl]uridine[317] and -adenosine[314] derived phosphoramidite building block as previously described. More details on the ORN synthesis are described in the original publications.



Scheme 5. (i) Me_3P, H_2O/dioxane; (ii) 1,3-di-Boc-2-(trifluoromethylsulfonyl)-guanidine, Et_3N, THF; (iii) anisole/TFA.

3.3.2 ^{19}F NMR spectroscopic analysis of PNA/RNA triplexes

3.3.2.1 PNA/heg hairpin triplexes

The ^{19}F NMR spectroscopic analysis was started by monitoring the thermal unwinding of the ^{19}F-labeled RNA hairpin models **ORN3** and **ORN4** without PNA (Figure 25). **ORN3** and **ORN4** contain an AU-rich duplex region connected by a hexaethylene glycol (heg) loop. The 2'-*O*-[(4-CF_3-triazol-1-yl)methyl]sensor was placed either in the middle of the pyrimidine strand (**ORN3**) or at the opposite site in the purine strand (**ORN4**). The ^{19}F NMR spectroscopic measurements were performed by using an RNA concentration of 50 μmol L^{-1} in a mixture of 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*). Upon heating, both hairpins **ORN3** and **ORN4** gave a well-behaved coalescence signal (Figure 25 a & b). To determine the ^{19}F NMR spectroscopy based thermal melting temperatures, the chemical shifts of these signals were presented as a function of temperature (after subtracting the passive temperature-dependent shift 0.0135 ppm $°C^{-1}$). The resulting melting profiles followed negative S curves, and the melting temperatures were determined based on the inflection points of the curves: (**A**/**B**, Figure 25; **ORN3**: T_m1 = 59.2 °C and **ORN4**: T_m1 = 56.9 °C; Table 3). The ^{19}F NMR spectroscopy based T_m values were consistent with the T_m values obtained by the UV melting profiles: 60.6 °C, 58.7 °C and 57.1

°C for **ORN1**, **ORN3** and **ORN4**, respectively (Table 6). Accordingly, the 2'-*O*-[(4-CF_3-triazol-1-yl)methyl]sensor destabilized the hairpin to some extent, and the destabilizing effect was higher in the purine strand than in the pyrimidine strand (**ORN3** vs. **ORN4**, ΔT_m = –1.9 °C vs. –3.5 °C).

Next, the applicability of ^{19}F NMR spectroscopy for the detection of PNA/RNA triplexes was evaluated. **PNAs 1–3** (Figure 23) were designed for targeting the RNA hairpin model **ORN1** (and the corresponding ^{19}F-labeled **ORN3** and **ORN4**). **PNAs 1–3** are decamers containing eight thymine and two 2-aminopyridine (M) nucleobases that expectedly bind to the Hoogsteen face of the purine strand of **ORN1**. One lysine residue was introduced to the amino terminus of all PNAs to improve solubility. **PNA2** contained three consecutive γ-(*R*)-hydroxymethyl units and **PNA3** contained a γ-(*R*)-hydroxymethyl unit at every other residue. ^{19}F NMR spectra of **ORN3** in the presence of **PNAs 1–3** at 25 °C are described in Figure 26a. **ORN3** alone gave a signal at δ = –62.08 ppm (**A**). When 1.0 equivalent of **PNAs 1–3** was added, the initial signal was completely replaced by a new signal, indicating the formation of PNA/RNA triplex **C**. When the temperature was increased, the chemical shift of these signals nearly followed the passive temperature-dependent shift (Figure 26 b–d). The relative peak areas of signals **C** vs. **A/B** were presented as a function of temperature (Figure 26e). The resulting melting profiles followed negative S curves, and the melting temperatures of the PNA/RNA triplexes (T_m2) were determined based on the inflection points of the curves. As seen in the melting curves and in Table 3, the γ-(*R*)-hydroxymethyl modifications reduced the affinity compared to **PNA1** by ΔT_m = –3.7 °C (**PNA2**) and –4.8 °C (**PNA3**). The thermal melting studies with **ORN4**/PNA triplexes were consistent with these results (Table 3; for the ^{19}F NMR spectra, see the original publication III).

When 2.0 equivalents of **PNA1–3** was added to **ORN3**, an additional signal was observed with all PNAs (Figure 26a). This signal was attributed to the formation of 2:1 $(PNA)_2$/RNA triplex invasion complex **D**. The intensity of this signal increased when more PNA (3.0 and 4.0 equivalents) was added and the intensity of the PNA/RNA triplex signal **C** decreased respectively (Figure 27 a–c). The γ-(*R*)-hydroxymethyl units reduced the formation of complex **D**: With **PNA1**, complex **C** was replaced by complex **D** after the addition of 3.0 equivalents of PNA, whereas 4.0 equivalents of **PNA2** were required for the quantitative formation of complex **D**. With **PNA3**, complexes **C** and **D** occurred in approximately 1:1 ratio after the addition of 4.0 equivalents of PNA. These observations and independent studies with **ORN4** (Figure 27d) supported the suggestion of $(PNA)_2$/RNA triplex invasion complex.

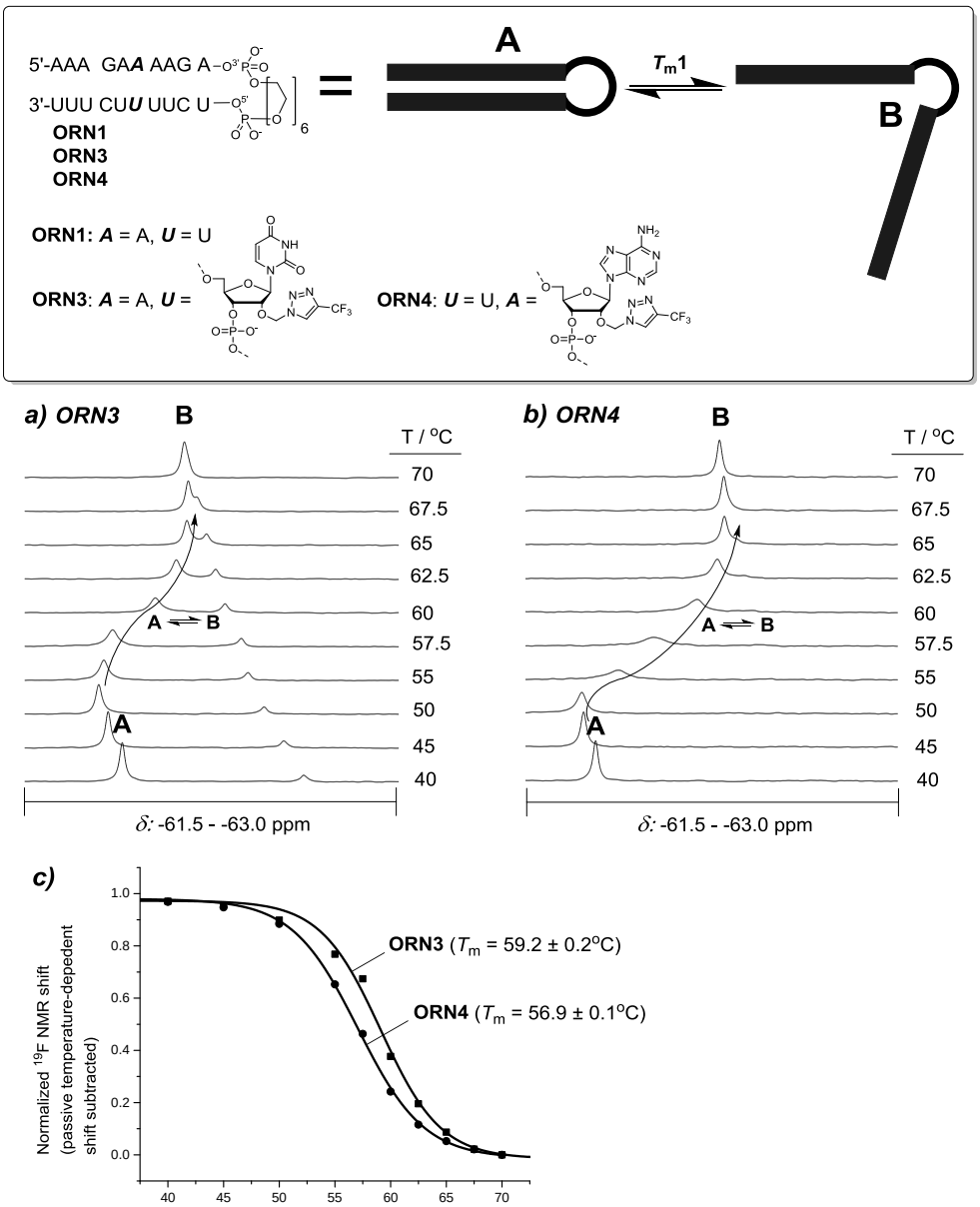


Figure 25. Thermal melting of the ^{19}F-labeled RNA hairpin models **ORN3** and **ORN4** followed by ^{19}F NMR spectroscopy. Conditions: 50 μmol L^{-1} of RNA, 10 mmol L^{-1} sodium cacodylate (pH 7.0), 0.1 mol L^{-1} NaCl in D_2O/H_2O (1:9, *v*/*v*).

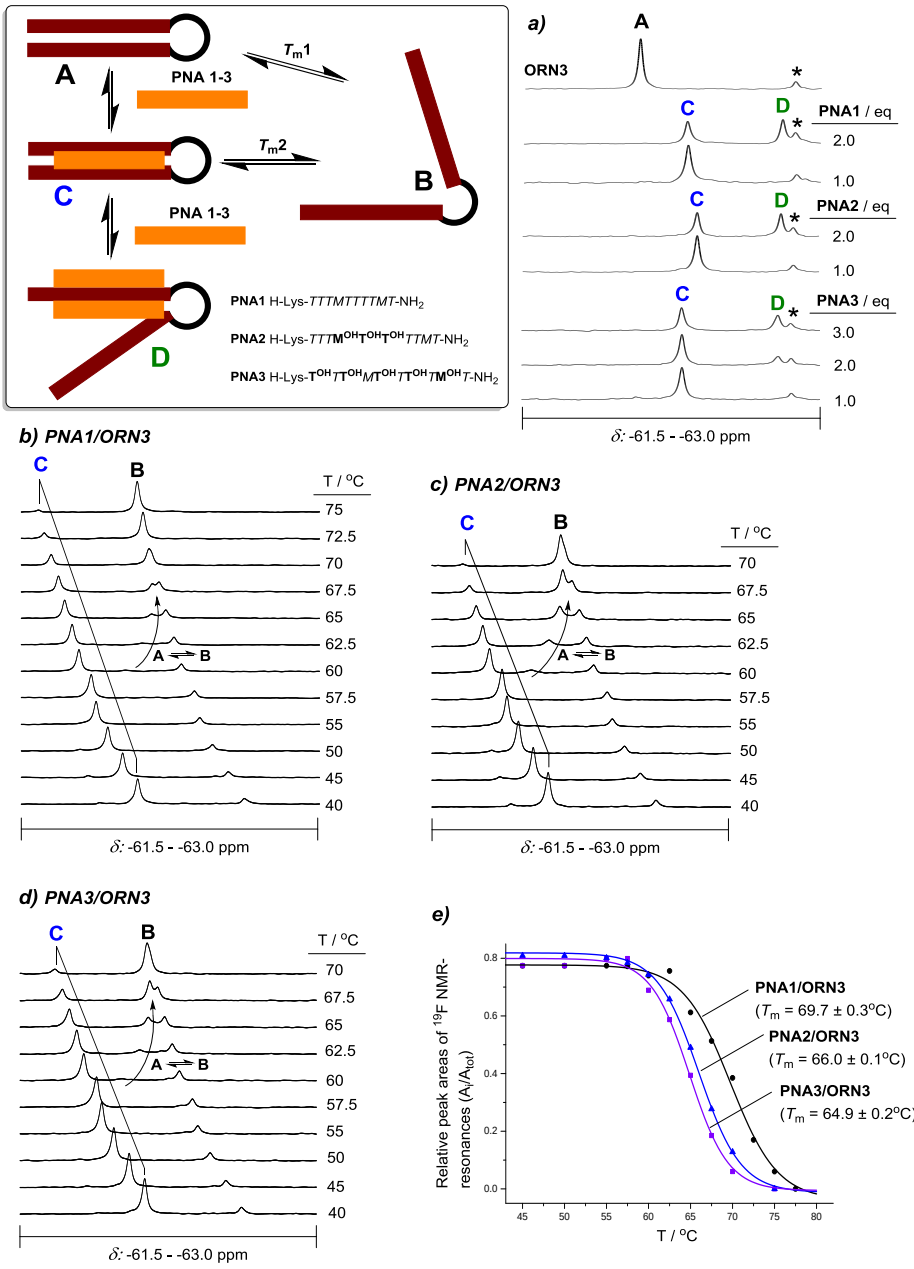


Figure 26. ^{19}F NMR spectroscopic data of PNA/**ORN3** complexes. Conditions: In each mixture: 50 µmol L^{-1} **ORN3**, 10 mmol L^{-1} sodium cacodylate (pH 7.0), 0.1 mol L^{-1} NaCl in D_2O/H_2O (1:9, *v*/*v*). a) 0–3.0 equiv. of PNA at 25 °C; b–e) 1.0 equiv. of PNA. Notes: a) δ(**A**) = –62.08 ppm; with **PNA1**, δ(**C**) = –62.32 ppm and δ(**D**) = –62.81 ppm; with **PNA2**, δ(**C**) = –62.38 ppm and δ(**D**) = –62.81 ppm; with **PNA3**, δ(**C**) = –62.26 ppm and δ(**D**) = –62.80 ppm. * = single-stranded trace, δ = –62.87 ppm.

Table 3. ^{19}F NMR spectroscopy based thermal melting temperatures [°C] of the heg hairpins in the presence of **PNAs 1–3**.

RNA	Hairpin alone T_m1	**PNA1** T_m2	**PNA2** T_m2	**PNA3** T_m2
ORN3	59.2 ± 0.2	69.7 ± 0.3	66.0 ± 0.1 (–3.7)	64.9 ± 0.2 (–4.8)
ORN4	56.9 ± 0.1	72.1 ± 0.1	67.9 ± 0.6 (–4.2)	67.9 ± 0.6 (–4.2)

Conditions: 50 μmol L^{-1} RNA + 0 (T_m1 values) or 1.0 (T_m2 values) PNA, 10 mmol L^{-1} sodium cacodylate (pH 7.0), 0.1 mol L^{-1} NaCl in D_2O/H_2O (1:9, *v*/*v*). Notes: T_m1 values are extracted from the shift vs. temperature profiles after the subtraction of the passive temperature-dependent shift 0.0135 ppm °C^{-1}. T_m2 values are extracted from the relative peak areas of the ^{19}F resonance signals (A_i/A_{tot}). ΔT_m values given in parentheses are those compared to the T_m values of the **PNA1**/RNA triplex.

It may be worth noting that the purine strand of **ORN3** and **ORN4** contains a symmetric track of eight nucleotides (AGAAAAGA), which may, in principle, form a direct antiparallel duplex with **PNAs 1–3**. In other words, the PNAs could also bind to the RNA hairpins through stoichiometric strand invasion. To test the favorability of plausible duplex formation, the binding of **PNA1** to a single-stranded RNA model **ORN5** was studied (Figure 27e). **ORN5** is the purine strand of **ORN3** and **ORN4**. **ORN5** alone gave a signal at δ = –62.90 ppm (**A'**). When **PNA1** was added to **ORN5**, only one signal at δ = –62.08 ppm was observed. The stoichiometry of the binding referred to $(PNA)_2$/RNA complex **D'** instead of PNA/RNA duplex **E'**. The absence of the duplex is probably related to the M nucleobases that, in their protonated form (pK_a = 6.7), are modest Watson–Crick binders.[25] Thus, it may be concluded that direct strand invasion to the heg hairpin models was unlikely. The PNAs bound to **ORN3** and **ORN4** through Hoogsteen base pairing as expected (**C**), and the $(PNA)_2$/RNA triplex invasion complex **D** was formed after an excess of PNA was added.

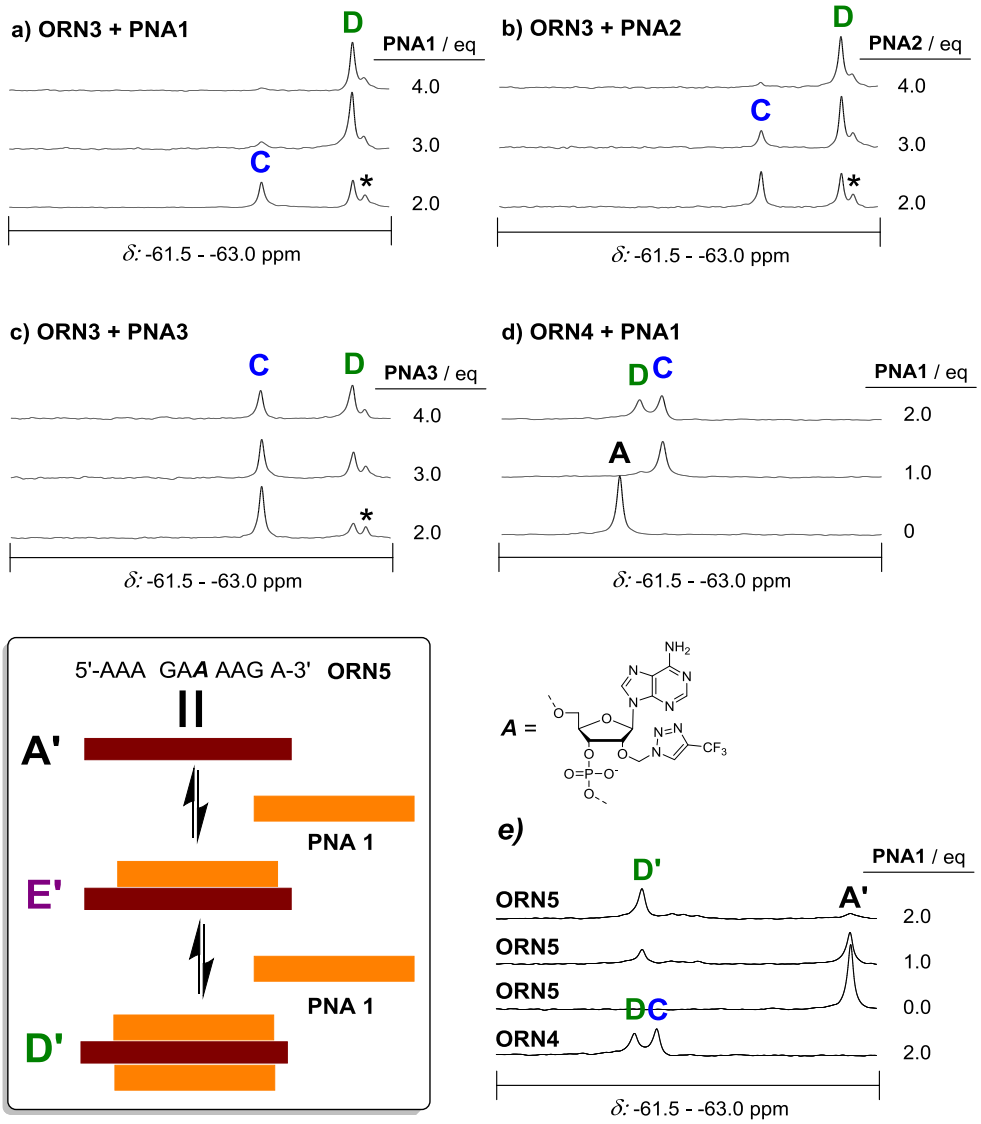


Figure 27. ^{19}F NMR spectra of **ORN3** + **PNA1–3** (2.0–4.0 equivalents) and **ORN4** + **PNA1** (0–2.0 equivalents). Conditions: 50 µmol L^{-1} RNA, 10 mmol L^{-1} sodium cacodylate (pH 7.0), 0.1 mol L^{-1} NaCl in D_2O/H_2O (1:9, *v/v*), 25 °C.

3.3.2.2 PNA/miR-215 triplexes

Next, the binding of PNAs to ^{19}F-labeled miR-215 models **ORN6** and **ORN7** (Figure 28) was analyzed. MiR-215 is related to cancer development and chemoresistance.[327,328] Zengeya et al. had previously developed **PNA4** (Figure 23) for the recognition of a miR-215 model (cf. **ORN2** in Figure 23).[25] **PNA4** contains

three M nucleobases[25] for the Hoogsteen-face recognition of guanine bases and one E nucleobase[24,323] for the recognition of the uracil interrupt. According to the ITC and UV thermal melting studies of Zengeya et al., **PNA4** bound to the miR-215 model with high affinity (K_a = 1.2×10^7) and 1:1 stoichiometry at physiologically relevant conditions.[25] We first wanted to study if the ^{19}F NMR spectroscopic analysis would reveal more details of the binding. As with the heg RNA hairpin models, the 2'-*O*-[(4-CF_3-triazol-1-yl)methyl]sensors were placed either in the middle of the pyrimidine-rich strand (**ORN6**) or at the opposite site in the purine-rich strand (**ORN7**). The ^{19}F NMR spectroscopic measurements were performed under the same conditions as with the heg hairpins. Upon heating, both **ORN6** and **ORN7** alone gave the expected **A**/**B** coalescence signal (Figure 28 a & b). The ^{19}F NMR spectroscopy based melting temperatures (T_m1) were 53.3 °C for **ORN6** and 51.6 °C for **ORN7** (Figure 28c). These T_m values were consistent with the T_m values obtained by UV spectroscopy: 54.4 °C for **ORN2**, 54.4 °C for **ORN6** and 52.5 °C for **ORN7**. As observed with the heg hairpins, the destabilizing effect of the 2'-*O*-[(4-CF_3-triazol-1-yl)methyl]sensor was higher in the purine-rich strand than in the pyrimidine-rich strand (**ORN6** vs. **ORN7**, ΔT_m = 0 vs. –1.9 °C).

To study the binding stoichiometry of **PNA4** to miR-215, **PNA4** was added gradually to **ORN6**. ^{19}F NMR spectra of **ORN6** in the presence of 0–2.0 equivalents of **PNA4** at 40 °C are shown in Figure 29a. **ORN6** alone gave a signal at δ = –61.90 ppm (**A**). Titration with **PNA4** resulted in a new signal very close to the original signal, indicating the formation of the **PNA4**/**ORN6** triplex **C**. However, another signal was observed when more than 0.8 equivalents of **PNA4** was added. This signal resulted from the formation of the (**PNA4**)$_2$/**ORN6** triplex invasion complex **D**. According to the relative peak areas of the ^{19}F resonance signals, complex **C** followed 1:1 stoichiometry and complex **D** 2:1 stoichiometry (Figure 29 d). When more than two equivalents of **PNA4** (3.0 and 4.0 equivalents) were added, no notable change in the chemical shift or intensity of the latter signal was observed (see the original publication II).

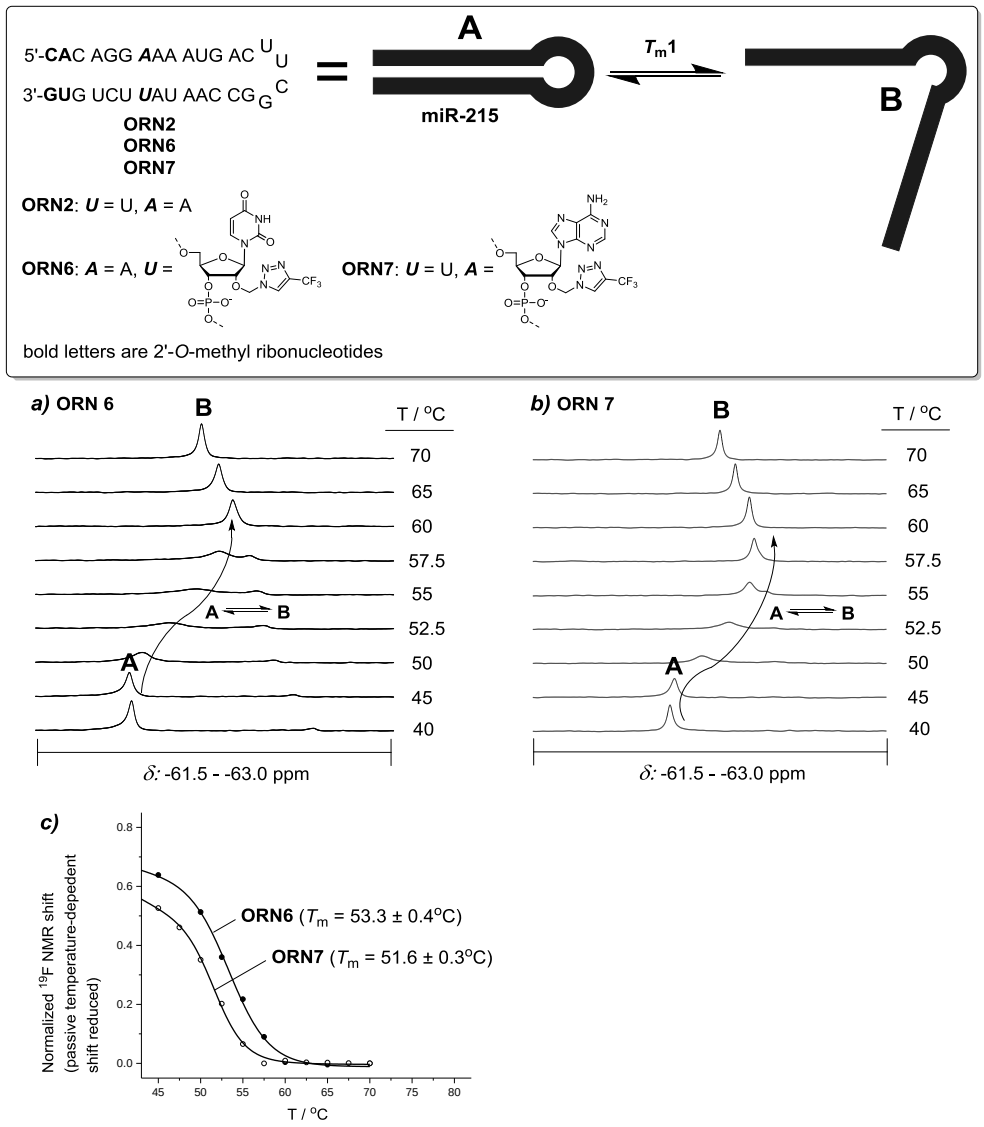


Figure 28. Thermal melting of the ^{19}F-labeled miR-215 models **ORN6** and **ORN7** followed by ^{19}F NMR spectroscopy. Conditions: 50 µmol L^{-1} of RNA, 10 mmol L^{-1} sodium cacodylate (pH 7.0), 0.1 mol L^{-1} NaCl in D_2O/H_2O (1:9, *v*/*v*).

To determine the thermal stabilities of complexes **C** and **D**, the ^{19}F NMR spectra were recorded at various temperatures. The ^{19}F NMR spectra of **ORN6** in the presence of 1.0 equivalent of **PNA4** at temperatures 40–70 °C are shown in Figure 30a, and the corresponding relative peak areas of the ^{19}F resonance signals as a function of temperature in Figure 30d. Based on these relative peak areas, thermal

denaturation profiles and distinct melting temperatures could be determined for both complexes: **C**: T_m2 = 58.9 °C and **D**: T_m3 = 69.9 °C. It is noteworthy that the molar fraction of (**PNA4**)$_2$/**ORN6** complex **D** increased at elevated temperatures and at T > 57.5 °C, **D** became the predominant complex. T_m3 was determined also separately from the melting profile of **ORN6** with two equivalents of **PNA4** added (Figure 30f). Based on this melting curve, T_m3 = 70.3 °C.

Next, the effect of the γ-(*R*)-hydroxymethyl modifications on the miR-215 recognition was studied. The ^{19}F NMR spectra of **ORN6** in the presence of 0–2.0 equivalents of **PNA5** and **PNA6** at 40 °C are shown in Figure 29. The ^{19}F NMR spectra of **ORN6** in the presence of 1.0 equivalent of PNA at temperatures 40–65 °C and the corresponding melting profiles are shown in Figure 30. The ^{19}F NMR spectroscopy based melting temperatures are listed in Table 4. **PNA5** with three consecutive γ-(*R*)-hydroxymethyl units was first examined. The addition of 1.0 equivalent of **PNA5** resulted in a signal at δ = –61.88 ppm, indicating the formation of the **PNA5**/**ORN6** triplex **C** (Figure 29b). This signal overlapped with the initial hairpin signal **A** at 40 °C, but distinct signals were observed at higher temperatures (Figure 30b). The addition of more than 1.0 equivalents of **PNA5** gave rise to broad and undefined resonances, attributed to the unspecific formation of the (**PNA5**)$_2$/**ORN6** triplex invasion complex **D**. According to the relative peak areas of the ^{19}F resonance signals, complex **C** followed 1:1 stoichiometry and complex **D** 2:1 stoichiometry (Figure 29 e). Next, ^{19}F NMR spectroscopic melting profiles for **C** and **D** were determined (Figure 30 b, e & f). The data showed that **PNA5** did not reduce the stability of the stoichiometric PNA/RNA triplex **C**, but reduced the stability of the invasion complex **D** to a great extent (**PNA5**/**ORN6** and **PNA5**/**ORN7** triplexes: ΔT_m2 = +0.4 and 0 °C; (**PNA5**)$_2$/**ORN6** invasion complex: ΔT_m3 = –6.2 °C, Table 4). The T_m3 value of (**PNA5**)$_2$/**ORN7** could not be determined because of the wide dispersion of the ^{19}F resonances. The reduced stability of complex **D** along with the wide signal dispersion may be attributed to the undefined formation of the invasion complex, or even aggregation. In addition, when 1.0 equivalent of **PNA5** was added to **ORN6**, only trace amounts of **D** were observed even at higher temperatures. Hence, **PNA5** had an increased selectivity for stoichiometric PNA/RNA triplex formation compared to **PNA4**.

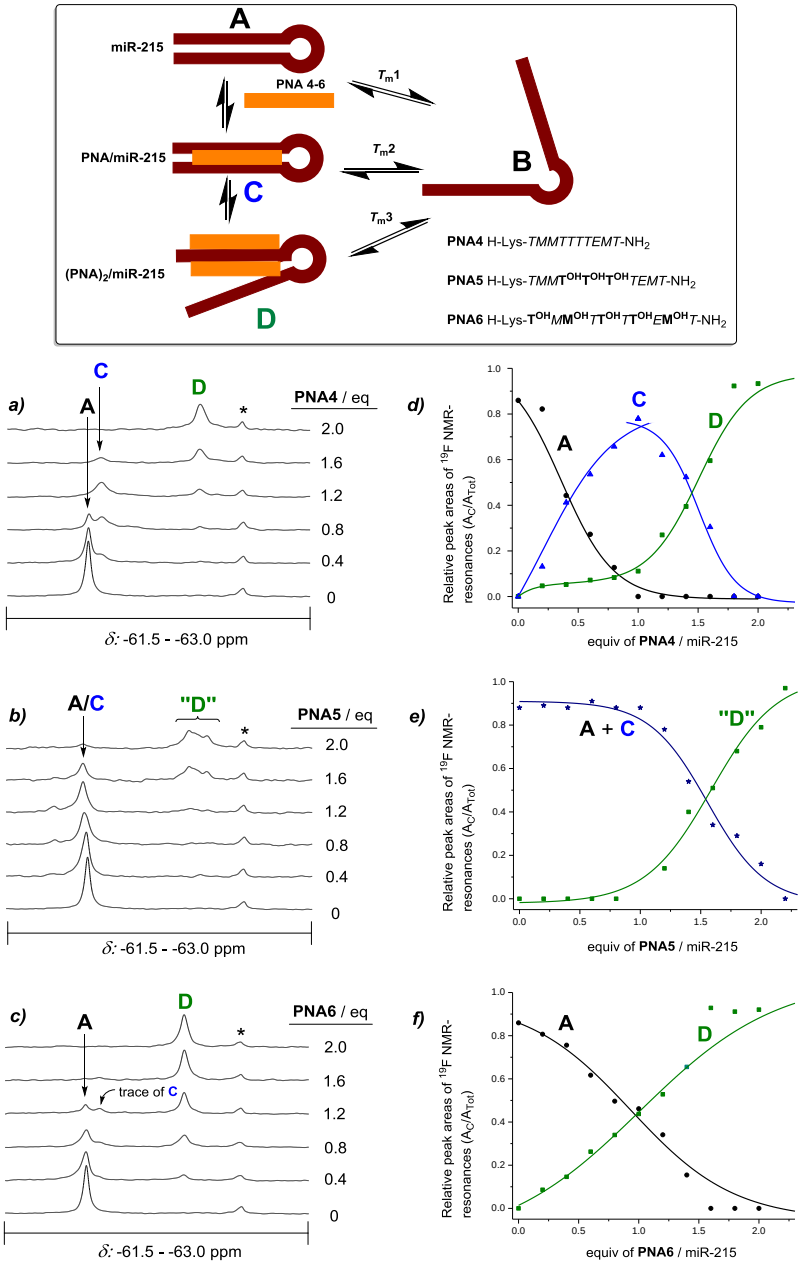


Figure 29. Titration of the miR-215 model **ORN6** with **PNAs 4–6** followed by ^{19}F NMR spectroscopy. Conditions: 50 µmol L^{-1} **ORN6** + **PNA4–6** (0–2.0 equivalents), 10 mmol L^{-1} sodium cacodylate (pH 7.0), 0.1 mol L^{-1} NaCl in D_2O/H_2O (1:9, *v*/*v*). Notes: a–c) δ(**A**) = –61.90 ppm, * = single-stranded trace, δ = –62.67 ppm.; a) δ(**C**) = –61.97 ppm and δ(**D**) = –62.56 ppm; b) δ(**C**) = –61.88 ppm and δ(**D**) = –62.20 to –62.49 ppm; c) δ(**C**) = –61.97 ppm and δ(**D**) = –62.38 ppm.

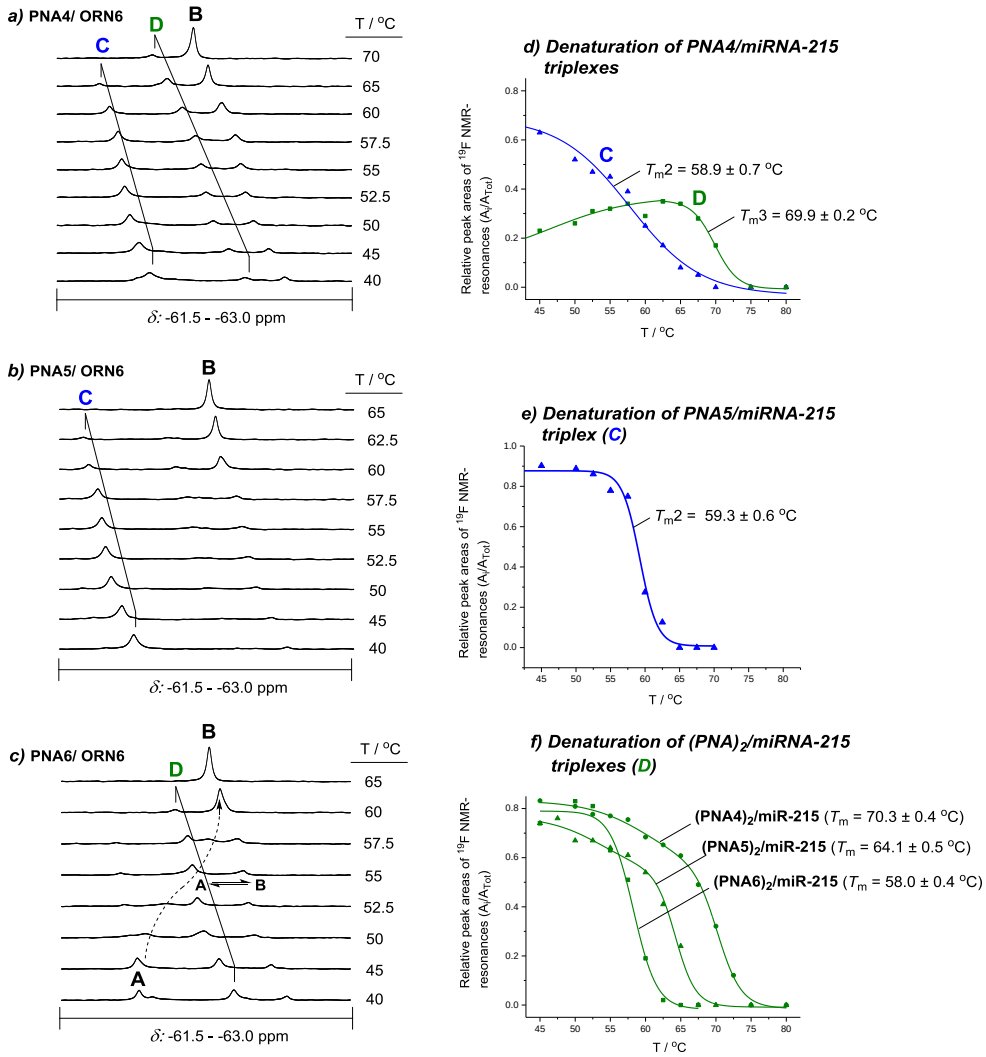


Figure 30. The thermal unwinding of PNA4–6/miR-215 complexes monitored by ^{19}F NMR spectroscopy. Conditions: 50 µmol L^{-1} **ORN6** + **PNA4–6** (a–e: 1.0 equivalent; f: 2.0 equivalents), 10 mmol L^{-1} sodium cacodylate (pH 7.0), 0.1 mol L^{-1} NaCl in D_2O/H_2O (1:9, *v*/*v*).

Thereafter, the binding of **PNA6** with five γ-(*R*)-hydroxymethyl units at every other position was studied. Interestingly, upon titration of **ORN6** with **PNA6**, the triplex invasion complex **D** was formed without notable accumulation of the PNA/RNA triplex **C** (Figure 29 c & f). However, the ^{19}F NMR spectroscopy based T_m3 value of this complex was surprisingly low: T_m3 with **ORN6** = 59.3 °C, Δ T_m3 = –11.0 °C; T_m3 with **ORN7** = 58.0 °C, Δ T_m3 = –9.8 °C (Figure 30 c & f and

Table 4). Hence, **PNA6** showed selectivity for the triplex invasion complex **D**, but the stability of the complex was even lower than with **PNA5**.

As discussed previously with the heg hairpins, the favorability of direct strand invasion to the miR-215 model was evaluated by following the binding of **PNA4** to a single-stranded model **ORN8** (Figure 31). **ORN8** is the purine-rich strand of **ORN7**. **ORN8** alone gave a signal at δ = –62.76 ppm (**A'**). The titration of **ORN8** with **PNA4** first gave a minor signal at δ = –62.39 ppm, which most likely indicated the formation of **PNA4**/**ORN6** duplex **E**. However, the predominant complex at δ = –62.31 ppm followed 2:1 stoichiometry, which referred to the formation of (**PNA4**)$_2$/**ORN6** complex **D'**. Hence, it may be reasonable to conclude that the binding of the PNAs to the miR-215 models through direct strand invasion was also unlikely, but triplex formation was required for the binding.

Table 4. ^{19}F NMR spectroscopy based thermal melting temperatures [°C] of the miR-215 models in the presence of **PNAs 4–6**.

RNA	Hairpin alone T_m1	**PNA4** T_m2[a]/T_m3[b]	**PNA5** T_m2[2]/T_m3[b]	**PNA6** T_m2[a]/T_m3[b]
ORN6	53.3 ± 0.4	58.9 ± 0.7[a]	59.3 ± 0.6[a] (+0.4)	–
		70.3 ± 0.4[b]	64.1 ± 0.5[b] (–6.2)	59.3 ± 0.5[b] (–11.0)
ORN7	51.6 ± 0.3	57.3 ± 0.6[a]	57.3 ± 0.5[a] (0)	–
		67.8 ± 0.3[b]	–	58.0 ± 0.4[b] (–9.8)

Conditions: 50 µmol L^{-1} RNA + 0 (T_m1 values), 1.0 (T_m2 values) or 2.0 (T_m3 values) of PNA, 10 mmol L^{-1} sodium cacodylate (pH 7.0), 0.1 mol L^{-1} NaCl in D_2O/H_2O (1:9, *v*/*v*). Notes: T_m1 values are extracted from the shift vs. temperature profiles after the subtraction of the passive temperature-dependent shift 0.0135 ppm °C^{-1}. T_m2 and T_m3 values are extracted from the relative peak areas of the ^{19}F resonance signals (A_i/A_{tot}). ΔT_m values for **PNA5** and **PNA6** given in parentheses are those compared to the T_m2 and T_m3 values of the **PNA4**/RNA triplex and (**PNA4**)$_2$/RNA complex.

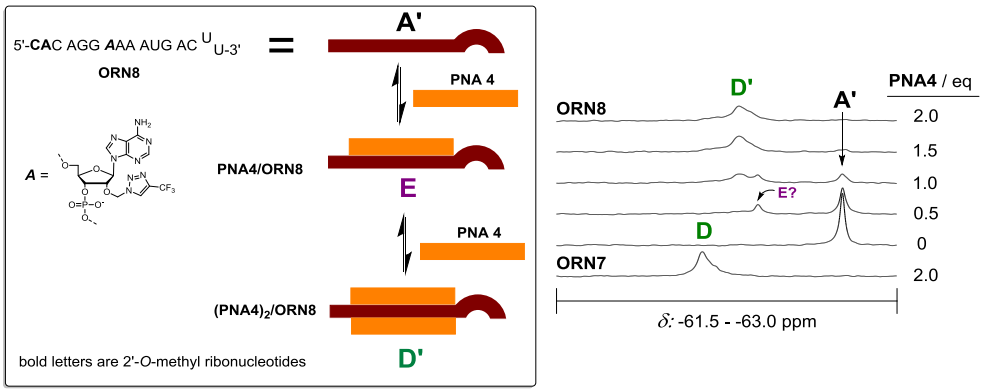


Figure 31. ^{19}F NMR spectroscopy data of **PNA4**/**ORN8** complexes. Conditions: 50 µmol L^{-1} **ORN8** + **PNA** (1.0 equivalent), 10 mmol L^{-1} sodium cacodylate (pH 7.0), 0.1 mol L^{-1} NaCl in D_2O/H_2O (1:9, *v*/*v*) at 35 °C. Notes: δ(**A'**) = –62.76 ppm, δ(**E**) = –62.39 ppm, δ(**D'**) = –62.31 ppm, δ(**D**) = –62.15 ppm.

3.3.2.2.1 γ-(S)-guanidinylmethyl-modified TFPNAs

The experiments with the γ-(*R*)-hydroxymethyl modified TFPNAs demonstrated that γ-chiral modifications affect the binding selectivity between stoichiometric PNA/miR-215 triplex and $(PNA)_2$/miR-215 triplex invasion complex. Furthermore, the positioning of the chiral units proved important; three consequent γ-(*R*)-hydroxymethyl units resulted in the best selectivity for the PNA/miR-215 triplex formation. Aiming to enhance the binding properties even more, the effect of γ-(*S*)-guanidinylmethyl modifications was next evaluated. The γ-(*S*)-guanidinylmethyl units were introduced into variable positions of the miR-215 recognizing TFPNAs (**PNAs 7–9**, Figure 23). For comparison to the guanidinylmethyl modifications, the effect of three simple arginine units at the carboxyl terminus was also evaluated (**PNA10**, Figure 23). For the ^{19}F NMR spectroscopic analysis, **ORN6** was selected as the target miR-215 model.

PNA7 and **PNA8**, containing two and three consecutive γ-(*S*)-guanidinylmethyl modifications, showed high selectivity for the stoichiometric PNA/RNA triplex formation (Figure 32). When 1.0 equivalent of **PNA7** or **PNA8** was added, **ORN6** was quite quantitatively replaced by the PNA/**ORN6** triplex **C** without notable formation of the $(PNA)_2$/**ORN6** complex. When more than 1.0 equivalent of **PNA7** or **PNA8** was added, the signal of the PNA/**ORN6** triplex was replaced by broad and undefined resonances that were very branched at higher temperatures. These resonances were attributed to the unspecific formation of the $(PNA)_2$/**ORN6** complex **D**, or even aggregation. Next, ^{19}F NMR spectroscopic melting profiles for the **PNA7**/**ORN6** and **PNA8**/**ORN6** triplexes were determined. The ^{19}F NMR spectra of **ORN6** in the presence of 1.0 equivalent of **PNA7** and

PNA8 at temperatures 40–70 °C are shown in Figure 33a & b, and the corresponding melting profiles in Figure 33d & e. The ^{19}F NMR spectroscopy based melting temperature for the **PNA7**/**ORN6** triplex was 60.5 °C, 1.6 °C higher compared to the original **PNA4**/**ORN6** triplex. The ^{19}F NMR spectroscopy based melting temperature for the **PNA8**/**ORN6** triplex was 62.8 °C, 3.9 °C higher compared to the original **PNA4**/**ORN6** triplex. Moreover, in contrast to **PNA4**, no notable formation of the triplex invasion complex was observed even at elevated temperatures.

To study the effect of the spacing between γ-(*S*)-guanidomethyl units on the miR-215 recognition, the binding of **PNA9** was examined. **PNA9** has three γ-(*S*)-guanidomethyl modifications separated by one to three PNA building blocks. In earlier studies, the spacing between arginine side chains was demonstrated to be important for the cellular uptake of peptides.[329] Interestingly, the titration of **ORN6** with **PNA9** resulted in a very wide dispersion of ^{19}F resonance signals, indicating unspecific formation of the complex **C** (Figure 32 c). When more than 1.0 equivalent of **PNA9** was added, this signal pattern was replaced by broad and undefined resonances, indicating unspecific formation of the complex **D**. Even if the formation of **C** followed approximately 1:1 stoichiometry and the formation of **D** approximately 2:1 stoichiometry, the wide dispersion of the ^{19}F resonance signals suggests unspecific and undefined binding to the **ORN6** target. Because of the wide dispersion of the signals, the melting temperatures could not be determined for **C** or **D**.

Finally, we wanted to study how simple arginine overhang affects the binding affinity and Hoogsteen-face selectivity of the TFPNA compared to γ-(*S*)-guanidomethyl units. For this purpose, the binding of **PNA10** with three arginine units at the carboxyl terminus was examined. The titration of **ORN6** with **PNA10** showed very similar results compared to **PNA4** (Figure 32 d & g). Consequently, the arginine units did not affect the Hoogsteen-face selectivity of the binding. ^{19}F NMR spectra of **ORN6** in the presence of 1.25 equivalents of **PNA10** at temperatures 40–70 °C are shown in Figure 33c. (1.25 equivalents of **PNA10** was used to get all of the hairpin **A** hybridized with **PNA10**.) The corresponding relative peak areas of the ^{19}F resonance signals are plotted as a function of temperature in Figure 33f. The melting temperature for complex **C** was 63.8 °C, 4.9 °C higher compared to the original **PNA4**/**ORN6** triplex. However, the molar fraction of the triplex invasion complex **D** increased at higher temperatures and **D** became the predominant complex at T > 63 °C. As a matter of fact, the arginine overhang proved to stabilize complex **D** even more than complex **C**: the melting temperature for complex **D** was 78.0 °C, 7.7 °C higher compared to **PNA4** (Table 5). The ^{19}F NMR spectroscopy based thermal melting temperatures of **PNA7–10**/miR-215 complexes are listed in Table 5.

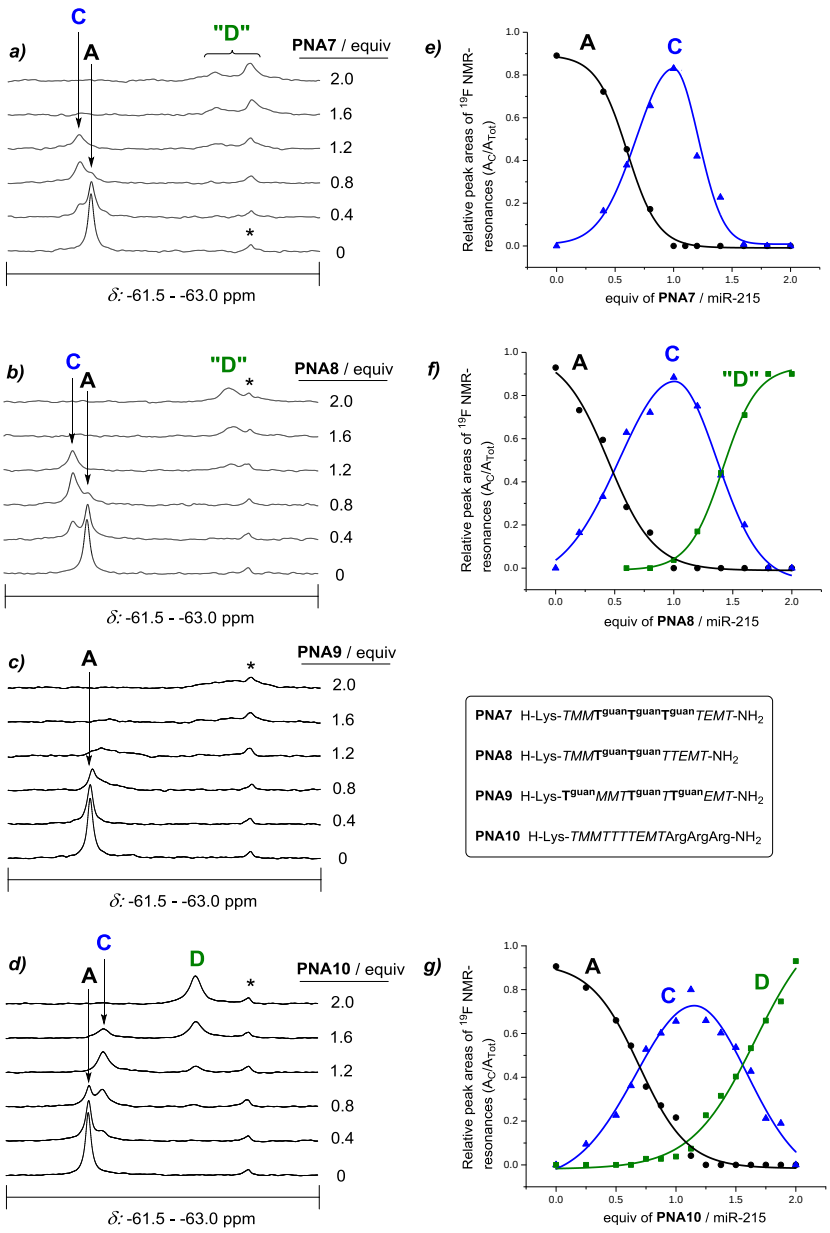


Figure 32. Titration of the miR-215 model **ORN6** with **PNAs 7–10** followed by ^{19}F NMR spectroscopy. Conditions: 50 μmol L^{-1} **ORN6** + **PNA7–10** (0–2.0 equivalents), 10 mmol L^{-1} sodium cacodylate (pH 7.0), 0.1 mol L^{-1} NaCl in D_2O/H_2O (1:9, *v*/*v*). Notes: a–d) δ(**A**) = –61.90 ppm, * = single-stranded trace, δ = –62.67 ppm.; a) δ(**C**) = –61.85 ppm and δ(**D**) = –62.34 to –62.82 ppm; b) δ(**C**) = –61.83 ppm and δ(**D**) = –62.58 ppm; d) δ(**C**) = –61.97 ppm and δ(**D**) = –62.41 ppm.

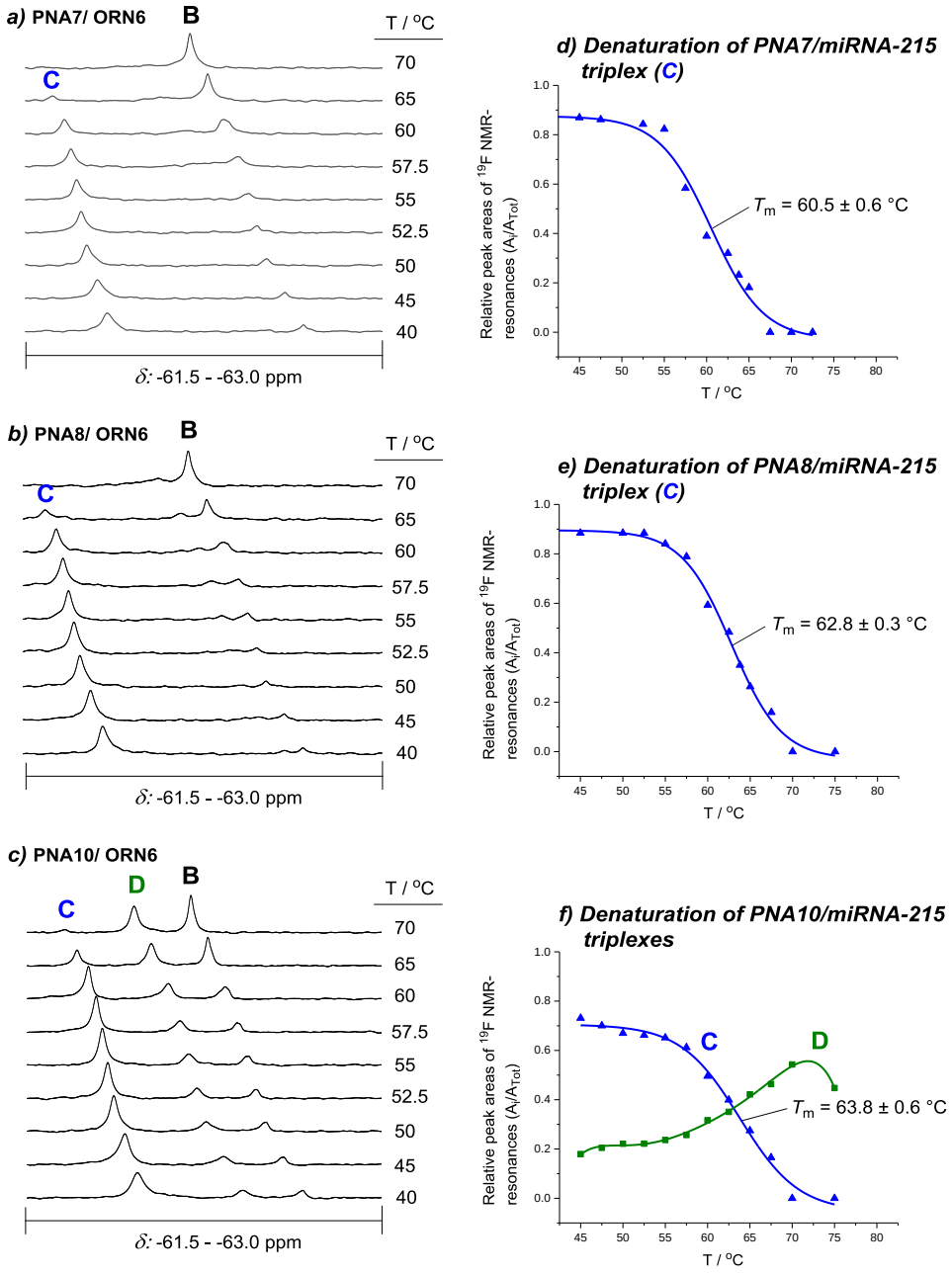


Figure 33. The thermal unwinding of **PNA7–10**/miR-215 complexes monitored by ^{19}F NMR spectroscopy. Conditions: 50 µmol L^{-1} **ORN6** + **PNA7–10** (c & f: 1.25 equivalents; others: 1.0 equivalent), 10 mmol L^{-1} sodium cacodylate (pH 7.0), 0.1 mol L^{-1} NaCl in D_2O/H_2O (1:9, *v*/*v*).

Table 5. ^{19}F NMR spectroscopy based thermal melting temperatures [°C] of the miR-215 model **ORN6** in the presence of **PNAs 7–10**.

ORN6+PNA4 $T_m2^{[a]}/T_m3^{[b]}$	**ORN6+PNA7** $T_m2^{[a]}/T_m3^{[b]}$	**ORN6+PNA8** $T_m2^{[a]}/T_m3^{[b]}$	**ORN6+PNA9** $T_m2^{[2]}/T_m3^{[b]}$	**ORN6+PNA10** $T_m2^{[a]}/T_m3^{[b]}$
58.9 ± 0.7[a]	60.5 ± 0.6 [a] (+1.6)	62.8 ± 0.3 [a] (+3.9)	–	63.8 ± 0.6 [a] (+4.9)
70.3 ± 0.4[b]	–	–	–	78.0 ± 0.4 [b] (+7.7)

Conditions: 50 μmol L^{-1} RNA + 1.0 (T_m2 values) or 2.0 (T_m3 values) of PNA, 10 mmol L^{-1} sodium cacodylate (pH 7.0), 0.1 mol L^{-1} NaCl in D_2O/H_2O (1:9, *v*/*v*). Notes: T_m2 and T_m3 values are extracted from the relative peak areas of the ^{19}F resonance signals (A_i/A_{tot}). ΔT_m values for **PNAs 7–10** given in parentheses are those compared to the T_m2 and T_m3 values of the **PNA4**/RNA triplex and (**PNA4**)$_2$/RNA complex.

3.3.3 UV and CD spectroscopic analysis of PNA/RNA triplexes

The UV spectroscopic analysis of the PNA/RNA triplexes was performed by using 2.0 μmol L^{-1} solutions of each oligonucleotide in the presence of 0, 1.0 or 2.0 equivalents of **PNA1–10** in a mixture of 10 mmol L^{-1} sodium cacodylate (pH 7.0) and 0.1 mol L^{-1} NaCl in H_2O, detection wavelength λ = 260 nm. The UV thermal melting temperatures of the heg hairpins in the presence of **PNAs 1–3** are listed in Table 6. Because of the 25-fold more diluted RNA concentration, the UV based T_m values were lower than the ^{19}F NMR spectroscopy based T_m values. The UV melting temperatures confirmed that the γ-(*R*)-hydroxymethyl modifications lowered the thermal stability of the **PNA1–3**/RNA triplexes, as observed with ^{19}F NMR spectroscopy.

The UV thermal melting temperatures of the miR-215 models in the presence of **PNAs 4–6** are listed in Table 7 and examples of the UV melting profiles of **PNA4–6**/miR-215 complexes are shown in Figure 34. The UV based melting temperatures suggested that the thermal stability of the PNA/miR-215 triplexes would be lower with both γ-(*R*)-hydroxymethyl modified PNAs **PNA5** and **PNA6**. However, the ^{19}F NMR spectroscopic analysis showed that the molar fraction of the (**PNA4**)$_2$/miR-215 complex increased at higher temperatures. This fact remains hidden in the UV thermal melting profiles. Accordingly, the observed UV melting profile with **PNA4** is a sum of the melting profiles of both complexes **C** and **D**, **D** being thermally more stable than **C**. Therefore, a falsely high UV based T_m value was observed with **PNA4**. With **PNA5**, on the other hand, the ^{19}F NMR spectroscopic studies showed that practically only PNA/miR-215 triplex is formed after the addition of 1.0 equivalent of PNA even at high temperatures. Hence, it may be reasonable to confirm that the UV melting profile with **PNA5** presents the denaturation of PNA/miR-215 triplex only. Likewise, the UV melting profile with

PNA6 presents the denaturation of (**PNA6**)$_2$/miR-215 invasion complex only. The UV experiments were also performed with 2.0 equivalents of **PNAs 4–6** and using 280 nm as the detection wavelength (see the original publication III). No significant difference from the T_m values in Table 7 was observed.

Table 6. UV thermal melting temperatures [°C] of the heg hairpins in the presence of **PNAs 1–3**.

RNA	Hairpin alone T_m1	**PNA1** T_m2	**PNA2** T_m2	**PNA3** T_m2
ORN1	60.6 ± 0.4	65.6 ± 0.5	62.1 ± 0.4	61.6 ± 0.6
ORN3	58.7 ± 0.7 (–1.9)	63.9 ± 0.7 (–1.7)	60.4 ± 0.6 (–1.7)	59.8 ± 0.5 (–0.8)
ORN4	57.1 ± 0.7 (–3.5)	63.4 ± 0.4 (–2.2)	60.2 ± 0.5 (–1.9)	58.3 ± 0.8 (–3.3)

Conditions: 2.0 μmol L^{-1} RNA + 0 (T_m1 values) or 1.0 equivalent (T_m2 values) of PNA, 10 mmol L^{-1} sodium cacodylate (pH 7.0), 0.1 mol L^{-1} NaCl in H_2O. Notes: ΔT_m values given in parentheses are those compared to the T_m values with the unlabeled miR-215 model **ORN1**.

Table 7. UV thermal melting temperatures [°C] of the miR-215 models in the presence of **PNAs 4–6**.

RNA	Hairpin alone T_m1	**PNA4** T_m2	**PNA5** T_m2	**PNA6** T_m2
ORN2	54.4 ± 0.6	60.9 ± 0.3	57.2 ± 0.5	54.5 ± 0.5
ORN6	54.4 ± 0.7 (0)	61.0 ± 0.4 (+0.1)	57.0 ± 0.6 (–0.2)	54.3 ± 0.7 (–0.2)
ORN7	52.5 ± 0.7 (–1.9)	59.6 ± 0.1 (–1.3)	55.8 ± 0.5 (–1.4)	52.4 ± 0.5 (–2.1)

Conditions: 2.0 μmol L^{-1} RNA + 0 (T_m1 values) or 1.0 equivalent (T_m2 values) of PNA, 10 mmol L^{-1} sodium cacodylate (pH 7.0), 0.1 mol L^{-1} NaCl in H_2O. Notes: ΔT_m values given in parentheses are those compared to the T_m values with the unlabeled miR-215 model **ORN2**.

The thermal melting of the **PNA4–6**/miR-215 complexes was also monitored by CD spectroscopy using 2.0 μmol L^{-1} RNA solutions with 1.0 equivalent of PNAs. Although the CD spectroscopic data certainly showed differences between different PNA/miR-215 complexes (see the original publication III), no convincing evidence of different binding modes was observed. To determine the CD based T_m values, the ΔA values at a specific wavelength were presented as a function of temperature. The T_m value was then determined from the inflection point of the resulting curve. The CD based T_m values were consistent with the T_m values obtained by UV spectroscopy.

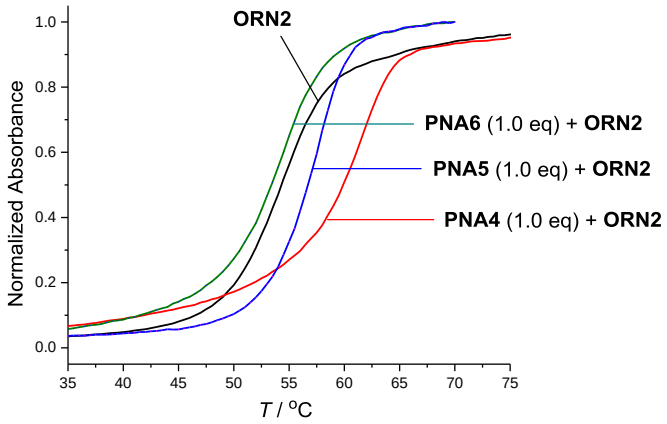


Figure 34. Normalized UV melting profiles of **PNA4–6**/miR-215 triplexes. Conditions: 2.0 μmol L^{-1} **ORN2** (black line) and 2.0 μmol L^{-1} **ORN2** + 1.0 equiv. **PNA4** (red line), **PNA5** (blue line), **PNA3** (green line), 10 mmol L^{-1} sodium cacodylate (pH 7.0), 0.1 mol L^{-1} NaCl in H_2O.

Next, the thermal melting of the γ-(*S*)-guanidylmethyl modified **PNA7–10**/miR-215 triplexes was monitored by UV spectroscopy. The UV melting profiles of the miR-215 model **ORN2** with 1.0 equivalent of **PNA7–10** are shown in Figure 35 and the corresponding thermal melting temperatures are listed in Table 8. The incorporated γ-(*S*)-guanidylmethyl units (**PNAs 7–9**) increased the observed triplex melting temperature by 2.0–2.8 °C compared to **PNA4**. As discussed previously, the observed T_m value with **PNA4** is probably higher than the T_m value of the **PNA4**/miR-215 triplex, because the observed T_m with **PNA4** is a sum of the melting profiles of both complexes **C** and **D**. With **PNA7** and **PNA8**, on the other hand, the ^{19}F NMR spectroscopic measurements showed that practically only PNA/miR-215 triplex is formed after the addition of 1.0 equivalent of PNA even at high temperatures. Therefore, it may be reasonable to conclude that the observed UV melting profiles with **PNA7** and **PNA8** present the denaturation of the PNA/miR-215 triplex only. Interestingly, the UV based T_m value with **PNA9** was very close to that with **PNA8**, even if the formation of the **PNA9**/miR-215 complexes resulted in a wide dispersion of ^{19}F resonance signals. The ^{19}F NMR spectroscopic data proposes that although the UV based T_m was relatively high, the formation of the **PNA9**/miR-215 triplex was unspecific and undefined compared to the other PNAs. **PNA10**, with three arginine units at the carboxyl terminus, gave the highest increase in the UV based T_m values, 5.4 °C. However, the ^{19}F NMR spectroscopic analysis revealed that **PNA10**, as well as **PNA4**, was not very selective for PNA/miR-215 triplex formation. After the addition of 1.0 equivalent

of **PNA10**, also (**PNA10**)$_2$/miR-215 complex was formed, and its molar fraction increased at high temperatures. Thus, the real T_m value of the **PNA10**/miR-215 triplex may be lower than the observed one. The UV experiments were also performed with 2.0 equivalents of **PNAs 7–10**.

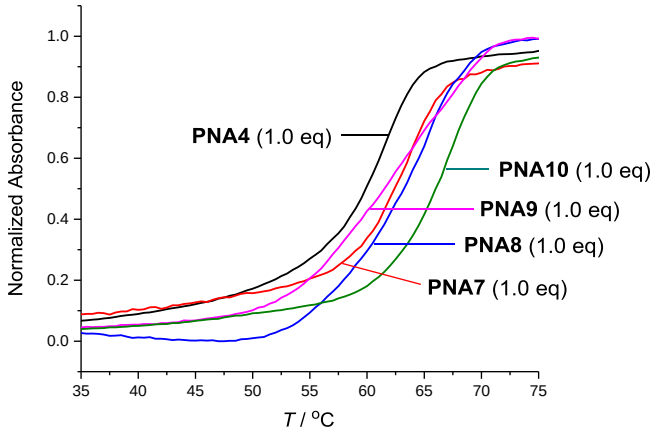


Figure 35. Normalized UV melting profiles of **PNA7–10**/miR-215 triplexes. Conditions: 2.0 μmol L^{-1} **ORN2** + 1.0 equiv. **PNA4** (black line), **PNA7** (red line), **PNA8** (blue line), **PNA9** (pink line) or **PNA10** (green line), 10 mmol L^{-1} sodium cacodylate (pH 7.0), 0.1 mol L^{-1} NaCl in H_2O.

Before the present study, M-modified PNAs had been shown to be highly selective for double-stranded RNA over DNA.[25] To confirm that the guanidine-modifications maintain the selectivity for double-helical RNA, the binding of **PNAs 4** and **7–10** to a miR-215 DNA derivative (**ODN17**), single-stranded RNA (**ORN9**) and single-stranded DNA (**ODN18**) was also studied (Figure 36). At least in theory, **ORN9** and **ODN18** are capable of forming a duplex with the TFPNAs through A-T and G-M base pairs. The binding was monitored by UV spectroscopy. Compared to the PNA/miR-215 triplexes, the observed melting temperatures of the PNA/**ODN17** triplexes were approximately 20 °C lower and the melting temperatures of the PNA/**ORN9** duplexes approximately 30 °C lower (see the original publication IV). The affinity of the TFPNAs for **ODN18** was even weaker. As may be expected because of the positive charge of guanidine groups, the γ-(*S*)-guanidomethyl and arginine modifications stabilized all the complexes to some extent. Nevertheless, the TFPNAs proved to have considerably weaker affinity for **ODN17**, **ORN9** and **ODN18** compared to miR-215. Thus, the γ-(*S*)-guanidomethyl modified TFPNAs maintained the selectivity for double-stranded RNA.

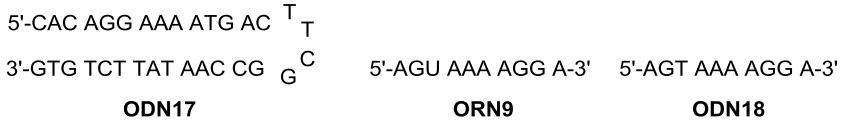


Figure 36. The sequences of a miR-215 DNA derivative (**ODN17**) and single-stranded RNA (**ORN9**) and DNA (**ODN18**) designed to form a duplex with **PNAs 4** and **7–10**.

Table 8. UV thermal melting temperatures [°C] of the miR-215 model **ORN2** in the presence of **PNAs 7–10**.

ORN2+PNA4 T_m2[a]/T_m3[b]	**ORN2+PNA7** T_m2[a]/T_m3[b]	**ORN2+PNA8** T_m2[a]/T_m3[b]	**ORN2+PNA9** T_m2[2]/T_m3[b]	**ORN2+PNA10** T_m2[a]/T_m3[b]
60.9 ± 0.3 [a]	62.9 ± 0.3 [a] (+2.0)	63.7 ± 0.6 [a] (+2.8)	63.3 ± 0.7 [a] (+2.4)	66.5 ± 0.4 [a] (+5.4)
62.3 ± 0.3[b] (+1.4)	61.9 ± 0.7 [b] (–1.0)	63.6 ± 0.8 [b] (–0.1)	66.0 ± 0.4 [b] (+2.7)	68.2 ± 0.6 [b] (+1.7)

Conditions: 2.0 μmol L^{-1} **ORN2** + 0 (T_m1 values), 1.0 (T_m2 values) or 2.0 equivavents (T_m3 values) of PNA, 10 mmol L^{-1} sodium cacodylate (pH 7.0), 0.1 mol L^{-1} NaCl in H_2O. [a] ΔT_m values given in parentheses are those compared with the T_m values of **ORN2** + **PNA1**. [b] ΔT_m values given in parentheses are those compared with the T_m values of **ORN2** + 1.0 equiv. of the corresponding PNA.

3.3.4 Cellular uptake studies of TFPNAs

To study the influence of the γ-(*S*)-guanidylmethyl modifications on the cellular uptake of PNAs, **PNA4**, **PNA7**, **PNA8** and **PNA10** were selected and labeled at the amino terminus with HiLyte Fluor 488 (HF488) dye following a previously described protocol (**HF488-PNAs**, Figure 23).[326] The cellular internalization of the PNAs to living PC-3 prostate cancer cells was followed by confocal microscopy. The cells were incubated with 5 μmol L^{-1} concentration of the HF488-labeled PNAs in OptiMEM (without serum) in the absence of transfection agents for 36–48 h at 37 °C, 5% CO_2. After incubation, the cells were washed with Opti-MEM and incubated with Hoechst 33342 to distinguish nuclei.

At 5 μmol L^{-1} concentration, **PNA7**, **PNA8** and **PNA10** were able to penetrate through cell membrane, which was detected as positive fluorescent signal in cytoplasm (Figure 37). The fluorescent intensity increased in the order **PNA7** < **PNA8** < **PNA10**. **PNA4**, on the other hand, showed no signal at 5 μmol L^{-1} concentration. The cellular uptake of **PNA4** could be detected after increasing the concentration to 10 μmol L^{-1} (data not shown). Hnedzko et al. previously demonstrated that the M nucleobase modification and arginine conjugates both enhance the cellular uptake of PNAs.[31] Our data shows that the γ-(*S*)-guanidylmethyl modifications improve the cellular delivery of PNAs, albeit the improvement was lower compared to arginine conjugates (cf. **PNA8** with three γ-

(*S*)-guanidylmethyl groups vs. **PNA10** with three arginine conjugates). The γ-(*S*)-guanidylmethyl modification and arginine both bear a guanidine head group that has been demonstrated to be the critical component for the more efficient cellular uptake of arginine oligomers compared to other polycationic oligomers.[330]

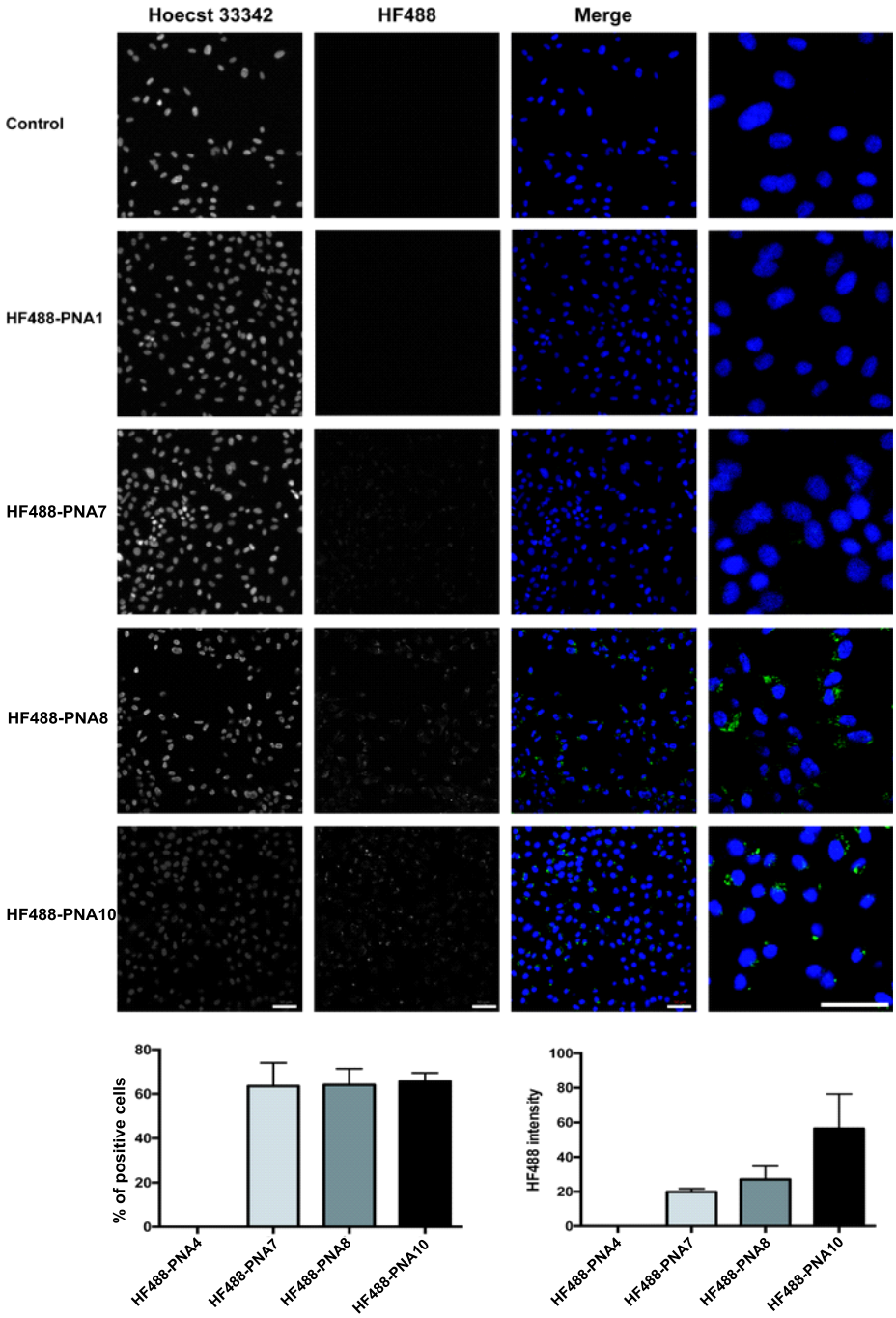


Figure 37. PNA uptake by PC-3 cells. Cells were incubated with 5 µM PNAs in Opti-MEM for 36-48 h at 37℃. Cells were washed after incubation, stained with Hoechst 33342 and detected by confocal microscope (Bar = 50 µm).

4 Summary

Aiming to increase the binding affinity of triplex-forming oligonucleotides (TFOs) for DNA duplexes, TFOs were conjugated with neomycin units. The neomycin units were conjugated either to the *C*-5 site (**ODN2**), *C*-2' site (**ODN3**) or *C*-4' site (**ODN4**) of a thymidine residue, and the site within the TFO sequence was also varied. The neomycin–TFO conjugates were synthesized on solid support using the click chemistry conditions. The effect of the neomycin conjugates on the triplex stability was evaluated by UV and CD thermal melting temperature analysis under slightly acidic conditions (pH 5.5 and 6.0). The neomycin conjugation site proved to be important for the triplex-stabilizing effect: the *C*-5 conjugates stabilized the triplexes remarkably, whereas the *C*-2' and *C*-4' conjugates caused no effect on the triplex stability. The importance of the conjugation site was expected, since the *C*-5 site is oriented toward the Watson–Hoogsteen groove, where the binding of neomycin takes place. The *C*-2' and *C*-4' sites, on the other hand, are oriented outward from the triple helix.

At pH 6.0, the *C*-5 conjugate **ODN2** increased the triplex melting temperature $T_m{}^3$ by ~ +8 °C (0.1 mol L^{-1} NaCl) and ~ +4 °C (0.3 mol L^{-1} NaCl). With two neomycin units (**ODN7**), the triplex-stabilizing effect was doubled ($\Delta T_m{}^3 \approx +16$ °C in 0.1 mol L^{-1} NaCl and $\Delta T_m{}^3 \approx +13$ °C in 0.3 mol L^{-1} NaCl). The conjugation site within the TFO sequence also affected the stabilization (**ODN2** and **ODN5** vs. **ODN6**). Compared to the triplex-stabilizing influence of discrete neomycin and structurally much simpler *C*-5-(alkyn-1-yl)-modified TFOs,[67–70] the stabilization obtained with the neomycin conjugates may be smaller than expected. However, the stabilization obtained with the bis-neomycin-conjugated **ODN7** was remarkable, and the neomycin-conjugated TFOs were able to stabilize the triplex also over a Hoogsteen mismatch. Additionally, compared to **ODN7**, a remarkable excess of discrete neomycin is required to stabilize triplexes.[115,122,133] Although the neomycin–TFO conjugates herein were studied under slightly acidic conditions, discrete neomycin has been shown to be still effective at pH 8.1.[115,125,126,134] Therefore, the binding affinity of these neomycin–TFO conjugates at physiological pH could probably be improved by replacing the cytosine residues with modified nucleobases, such as 2-aminopyridine derivatives.[61–64] In addition, ^{19}F-labeled target duplexes and ^{19}F NMR spectroscopy were used in this study to provide more detailed information on the triplex/duplex/single strand conversion. The ^{19}F

NMR spectroscopic analysis revealed the molar fractions of different secondary structures (triplex, duplex and single strand) at different temperatures, and an indication of locally varying environments in the triplex model was also observed.

Aiming to recognize variable nucleobases in the pyrimidine strand of double-helical DNA, ^{19}F-labeled TFO probes were developed. For this purpose, phosphoramidite building blocks of CF_3-modified sensors **F2** and **F3** were introduced into TFOs by automated DNA synthesizer. The sensors destabilized DNA triplexes, but the ^{19}F NMR chemical shifts of the sensors were sensitive to the nucleobase content of the target duplex. Unique fingerprints for the nucleobase content were observed, particularly with sensor **F3** that formed **F3*****G**-X triplets, where X = T, A, C or G. Unique ^{19}F resonance signals could be distinguished also from a mixture containing all the four target duplexes with variable nucleobase content. Because a cytosine-rich triplex model was used (the same model as with the neomycin–TFO conjugates), these studies were performed at pH 5.5. As discussed previously, the triplex could probably be stabilized at physiological pH by replacing the cytosine residues with modified nucleobases. Therefore, the principle of ^{19}F-labeled TFO sensors capable of recognizing nucleobase content in DNA duplexes proved promising. Compared to fluorescence-based detection methods and PCR, the sensitivity of ^{19}F NMR is modest, limiting its applicability to detect SNPs. Even if the sensitivity of ^{19}F NMR may be increased by improvements in NMR techniques, such as exploiting hyperpolarization,[331] the modest sensitivity may prevent the use of this kind of ^{19}F-labeled TFO sensors in medical diagnosis. However, considering academic interest, these ^{19}F-labeled TFOs stand as another proof of the sensitivity of ^{19}F NMR for detecting small changes in the local environments of oligonucleotides. This was the first report of ^{19}F-labeled TFOs that recognize nucleobase content in the target duplex and the first report of TFOs capable of sensing the pyrimidine strand in the target duplex. Additionally, no method for distinguishing mismatched Watson–Crick base pairs from each other was reported before this study.

Aiming to increase the binding affinity of triplex-forming peptide nucleic acids (TFPNAs) for double-helical RNA hairpins, chiral γ-(*R*)-hydroxymethyl and γ-(*S*)-guanidinylmethyl modifications were introduced to the TFPNAs (**PNAs 2**, **3** and **5**–**9**). Additionally, ^{19}F NMR spectroscopy was used to obtain new information on the different mechanisms by which TFPNAs bind to double-helical RNA targets, **ORN1** and a miR-215 model **ORN2**. For this purpose, ^{19}F sensors, previously developed in our laboratory, were incorporated to the target RNAs (**ORN 3**, **4**, **6** and **7**). The ^{19}F NMR spectroscopic analysis revealed detailed information on the stoichiometry and transition between alternative binding modes, 1:1 PNA/RNA triple helix and 2:1 $(PNA)_2$/RNA triplex invasion complex. This information remained hidden in conventional UV and CD thermal melting analyses.

TFPNAs with two or three consecutive γ-(*R*)-hydroxymethyl or γ-(*S*)-guanidinylmethyl modifications in the middle of the sequence (**PNAs 5**, **7** and **8**) reduced the stability of the $(PNA)_2$/miR-215 invasion complex and preferred stoichiometric PNA/miR-215 triplex formation. Moreover, the γ-(*S*)-guanidinylmethyl-modified **PNAs 7** and **8** increased the thermal stability of the PNA/miR-215 triplex. The increased Hoogsteen-face selectivity of **PNAs 5**, **7** and **8** is likely to originate from the γ-chiral backbone modification that provides favorable preorganization for right-handed helical conformation.[269,273,332] The increased binding affinity is likely to result from the positive charge of the γ-(*S*)-guanidinylmethyl residues at neutral pH. **PNA8** with three consecutive γ-(*S*)-guanidinylmethyl modifications showed the best binding affinity and Hoogsteen-face selectivity. According to ^{19}F NMR, the triplex melting temperature was increased by ~ 4 °C at pH 7.0, and the stabilization was observed also by UV spectroscopy. Furthermore, γ-(*S*)-guanidinylmethyl modifications enhanced the cellular uptake of TFPNAs. It should be noted that relatively high concentrations (50 μmol L^{-1}) were used in the ^{19}F NMR measurements, and the molar fractions of $(PNA)_2$/RNA triplex invasion complexes may be less prominent at lower concentrations. However, different binding modes are still likely to be present also at lower concentrations. For the potential therapeutic applications in the future, it is crucial to confirm the selective binding of TFPNAs and to reveal the equilibria between different binding modes and how they are affected by modifications.

5 Experimental

5.1 General

The syntheses and characterization of novel compounds are described in the original publications. The novel compounds were characterized by ^{1}H NMR, ^{13}C NMR, ^{31}P NMR, ^{19}F NMR and HRMS when applicable. The optical purity of compounds **35** and **40** was confirmed by chiral HPLC analysis using an analytical chiral Lux Cellulose (250 × 5 mm, 5μm) column, isocratic elution with 25% (compound **35**) or 70% (compound **40**) iPrOH in hexane, flow rate 1.0 mL min^{-1} and UV detection at 260 nm. For reference, small quantities of the *S*-isomers of compounds **35** and **40** were synthesized. The chiral HPLC profiles are provided in the original publications II and IV.

5.2 Oligonucleotide synthesis

The oligonucleotides were synthesized with an Applied Biosystems 3400 DNA/RNA synthesizer using standard protocols. More details on the oligonucleotide synthesis are provided in the Results and Discussion and in the original publications. Oligonucleotides were purified by RP-HPLC using a semipreparative column (C-18, 250 mm x 10 mm, 5 μm) with a gradient elution, typically from 0 to 40–60% acetonitrile in 0.1 mol L^{-1} aqueous triethylammonium acetate over 25 min, flow rate 3.0 mLmin^{-1}, UV detection wavelength 260 nm. The homogeneity of the purified oligonucleotides was confirmed by RP-HPLC using an analytical column (C-18, 250 mm x 5 mm, 5 μm), flow rate 1.0 mLmin^{-1}, elution conditions and detection wavelength the same as during purification. The authenticity of the oligonucleotides was verified by electrospray ionization time-of-flight mass spectrometry (ESI-TOF MS).

5.3 PNA synthesis

PNAs were synthesized on a 10 μmol scale on a Rink amide-derived Chem Matrix resin by Applied Biosystems 433A peptide synthesizer. In the PNA synthesis, the modified PNA monomers **1**–**5** and commercially available Fmoc-PNA-T-OH and amino acid building blocks (Boc-Lys(Fmoc)-OH, Fmoc-Arg(Pbf)-OH, Fmoc-

Lys(Fmoc)-OH) were incorporated into the PNA sequences using Fmoc/*t*-Bu peptide synthesis cycle. Benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate (PyBOP) was used as coupling reagent. For each coupling, PNA monomer/amino acid (50 μmol, 0.25 mol L^{-1} solution in NMP), PyBOP (5 equiv), DIEA (10 equiv) and a 30 min coupling time at rt were used. The coupling was followed by a capping step with an acetic anhydride treatment (Ac_2O, Py, NMP, 1:25:25, *v*/*v*/*v*, 1 min at rt. For the Fmoc deprotection, 20% piperidine in NMP was used (7 min at rt). Solid-supported PNAs were released with a mixture of anisole and TFA (1:10, *v*/*v*, 2 h at rt), precipitated in cold diethyl ether, dissolved in a 0.1% aqueous solution of TFA, and purified by RP HPLC. The product fractions were lyophilized to dryness to give the desired homogenized PNAs as white powders. The homogeneity of the products was confirmed by RP HPLC, and the authenticity was verified by MS (ESI-TOF). Yields of the isolated products were determined from the UV absorbance at $\lambda = 260$ nm.

5.3.1 Synthesis of guanidine-modified PNAs 7–9 and HF488-PNAs 7–9

After automated PNA synthesis, the γ-(*R*)-azidomethyl groups of **PNAs 7–9 (N_3)** were reduced into γ-(*S*)-aminomethyl groups and then guanidinylated on solid support. For the reduction step, the resin was suspended in a mixture of water and dioxane (1:4, *v*/*v*) and a solution of 1 mol L^{-1} Me_3P in toluene (24 equiv/amino group) was added under nitrogen atmosphere. The suspension was shaken for 2 h at room temperature, after which the resin was filtered and washed with DMF and CH_2Cl_2. For the guanidinylation step, the resin was suspended in dry THF and 1,3-di-Boc-2-(trifluoromethylsulfonyl)guanidine (10 equiv/amino group) and triethylamine (50 equiv/amino group) were added. Because the guanidinylation required highly basic conditions, the lysine at the amino terminus of the PNA was protected with Boc instead of Fmoc. The suspension was shaken overnight at room temperature. The guanidinylation step turned out to be sluggish, and the treatment was repeated 8 times to get all the amino groups guanidinylated. Thereafter, the resin was filtered, washed with DMF and CH_2Cl_2 and PNAs were released from the solid support as described previously.

When synthesizing γ-(*S*)-guanidinylmethyl modified **HF488-PNAs 7–9**, the synthesis was continued after the guanidinylation step. Therefore, the lysine at the amino terminus of the PNA had to be protected with Fmoc instead of Boc. Consequently, less basic conditions (1 equivalent triethylamine/amino group) were used and the treatment was repeated three times overnight. Although the guanidinylation was not fully complete after this treatment, the final products could be purified effortlessly by RP HPLC after attaching the AEEA spacers and the HF488 label.

5.4 UV and CD thermal melting studies

The UV melting curves (absorbance vs. temperature) were recorded at 260 or 295 nm on a Perkin Elmer Lambda 35 UV-vis spectrometer equipped with a multiple cell holder and a Peltier temperature controller. The T_m values were determined as the local maximum (260 nm) or minimum (295 nm) of the first derivative of the melting curve. The CD spectra were recorded on an Applied Photophysics Chirascan spectropolarimeter. With both instruments, an internal thermometer was used to confirm the accurate target temperature. The detailed measurement conditions are described in the Results and Discussion and in the original publications.

5.5 ^{19}F NMR studies

The ^{19}F NMR spectra were recorded at a frequency of 470.6 MHz on a Bruker Avance 500 MHz spectrometer (publications I, II and IV) or Bruker Avance III HD 600 MHz spectrometer equipped with a cryogenic probe (publication III). The cryogenic probe increases the sensitivity ca. tenfold compared to conventional probes that work at ambient temperature. 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 the typical number of scans 2048. The parameters were optimized to obtain the signals with the longest relaxation rate. The temperature ramps were recorded automatically using a macro command and a 20 min equilibration time for each temperature. The sample temperatures were calibrated using known chemical shifts of ethylene glycol at different temperatures. The detailed measurement conditions are described in the Results and Discussion and in the original publications.

Acknowledgements

This thesis is based on experimental work performed in the Laboratory of Organic Chemistry and Chemical Biology at the Department of Chemistry, University of Turku during the years 2014–2018. The financial support from the Academy of Finland and Turku University Foundation are gratefully acknowledged.

First and foremost, I wish to express my deepest gratitude to Professor Pasi Virta for giving me an opportunity to pursue my doctoral thesis under his guidance. I am truly impressed by his excellent knowledge in chemistry and inspired by his passion in research. Whenever I had questions or faced problems in my research, I could always turn to him and – what is more – he almost every time came up with a solution.

I wish to thank Professor Eriks Rozners and Professor Mikko Oivanen for the careful reviewing of my thesis and providing their valuable comments and Professor Sabine Müller for accepting to act as my opponent.

I am grateful to my collaborators Dr. Lotta Granqvist, Dr. Naresh Bhuma, Dr. Merita Murtola, Prof. Roger Strömberg, Alejandra Verhassel and Dr. Johanna Tuomela for their contribution in the papers included in this thesis.

I want to thank all the people in the (Royal) Bioorganic group with whom I have had the honor to work with during these years: Aapo Aho, Dr. Alejandro Gimenez Molina, Antti Äärelä, Dr. Anna Leisvuori, Dr. Anu Kiviniemi, Asmo Aro-Heinilä, Chungfeng Chen, Dattatraya Ukale, Dr. Emilia Kiuru, Professor Harri Lönnberg, Dr. Heidi Korhonen, Dr. Jianwei Li, Jinghui Yang, Lange Saleh, Dr. Luigi Lain, Madhuri Hande, Dr. Mikko Ora, Dr. Oleg Golubev, Olli Moisio, Petja Rosenqvist, Petteri Lepistö, Dr. Päivi Poijärvi-Virta, Dr. Sajal Maity, Samu Savikko, Dr. Satish Jadhav, Dr. Satu Mikkola, Dr. Sharmin Taherpour, Dr. Teija Niittymäki, Tiina Buss, Dr. Tuomas Karskela, Asst. Prof. Tuomas Lönnberg, Vijay Gulumkar, Vyacheslav Kungurtsev, Yonglei Lyu, Dr. Yu Cao. Thank you for your help and advice and for creating a pleasant working atmosphere. Thank you for the great conversations and for your company in lab, during coffee breaks, conference trips and other events. I wish to extend my thanks to Dr. Jari Sinkkonen, Dr. Kaisa Ketomäki, Kirsi Laaksonen and the "magicians" Mauri Nauma and Kari Loikas for fixing all the problems related to instruments and computers. Also, the mental support from "Joko syödään" lunch group is gratefully acknowledged.

I also wish to thank all my friends outside the world of chemistry for keeping me (more or less) sane during these years. I am lucky to have you in my life! Special thanks go to Livonsaaren Kalastusurheilun ystävät and Ski Team Garlic Bath.

My warmest thanks are devoted to my parents, Annikki and Markku, and my sister Jenni. Thank you for your love and support throughout my life!

Turku, September 2019

References

(1) Praseuth, D.; Guieysse, A. L.; Hélène, C. *Biochim. Biophys. Acta - Gene Struct. Expr.* **1999**, *1489*, 181–206.
(2) Duca, M.; Vekhoff, P.; Oussedik, K.; Halby, L.; Arimondo, P. B. *Nucleic Acids Res.* **2008**, *36*, 5123–5138.
(3) Watson, J. D.; Crick, F. H. C. *Nature* **1953**, *171*, 737–738.
(4) Wilkins, M. H. F.; Stokes, A. R.; Wilson, H. R. *Nature* **1953**, *171*, 738–740.
(5) Franklin, R. E.; Gosling, R. G. *Nature* **1953**, *172*, 156–157.
(6) Felsenfeld, G.; Davies, D. R.; Rich, A. *J. Am. Chem. Soc.* **1957**, *79*, 2023–2024.
(7) Moser, Heinz E.; Dervan, P. B. *Science* **1987**, *238*, 645–650.
(8) Le Doan, T.; Perrouault, L.; Praseuth, D.; Habhoub, N.; Decout, J. L.; Thuong, N. T.; Lhomme, J.; Hélène, C. *Nucleic Acids Res.* **1987**, *15*, 7749–7760.
(9) Broitman, S. L.; Im, D. D.; Fresco, J. R. *Proc. Natl. Acad. Sci. U. S. A.* **1987**, *84*, 5120–5124.
(10) Wells, R. D.; Collier, D. A.; Hanvey, J. C.; Shimizu, M.; Wohlrab, F. *FASEB J.* **1988**, *2*, 2939–2949.
(11) Seidman, M. M.; Glazer, P. M. *J. Clin. Invest.* **2003**, *112*, 487–494.
(12) Jain, A., Wang, G., Vasquez, K. M. *Biochimie* **2008**, *90*, 1117–1130.
(13) Mukherjee, A.; Vasquez, K. M. *Biochimie* **2011**, *93*, 1197–1208.
(14) Bartel, D. P. *Cell* **2004**, *116*, 281–297.
(15) Bartel, D. P. *Cell* **2009**, *136*, 215–233.
(16) Ling, H.; Fabbri, M.; Calin, G. A. *Nat. Rev. Drug Discov.* **2013**, *12*, 847–865.
(17) Sharp, P. A. *Cell* **2009**, *136*, 577–580.
(18) Li, M.; Zengeya, T.; Rozners, E. *J. Am. Chem. Soc.* **2010**, *132*, 17052–17052.
(19) Zengeya, T.; Li, M.; Rozners, E. *Bioorg. Med. Chem. Lett.* **2011**, *21*, 2121–2124.
(20) Devi, G.; Yuan, Z.; Lu, Y.; Zhao, Y.; Chen, G. *Nucleic Acids Res.* **2014**, *42*, 4008–4018.
(21) Sato, T.; Sato, Y.; Nishizawa, S. *J. Am. Chem. Soc.* **2016**, *138*, 9397–9400.
(22) Toh, D.-F. K.; Patil, K. M.; Chen, G. *J. Vis. Exp.* **2017**, *127*, e56221–e56221.
(23) Puah, R. Y.; Jia, H.; Maraswami, M.; Kaixin Toh, D.-F.; Ero, R.; Yang, L.; Patil, K. M.; Lerk Ong, A. A.; Krishna, M. S.; Sun, R.; Tong, C.; Huang, M.; Chen, X.; Loh, T. P.; Gao, Y.-G.; Liu, D. X.; Chen, *Biochemistry* **2018**, *57*, 149–159.
(24) Gupta, P.; Zengeya, T.; Rozners, E. *Chem. Commun.* **2011**, *47*, 11125–11127.
(25) Zengeya, T.; Gupta, P.; Rozners, E. *Angew. Chem. Int. Ed.* **2012**, *51*, 12593–12596.
(26) Gupta, P.; Muse, O.; Rozners, E. *Biochemistry* **2012**, *51*, 63–73.
(27) Muse, O.; Zengeya, T.; Mwaura, J.; Hnedzko, D.; McGee, D. W.; Grewer, C. T.; Rozners, E. *ACS Chem. Biol.* **2013**, *8*, 1683–1686.
(28) Endoh, T.; Hnedzko, D.; Rozners, E.; Sugimoto, N. *Angew. Chem. Int. Ed.* **2016**, *55*, 899–903.
(29) Cheruiyot, S. K.; Rozners, E. *ChemBioChem* **2016**, *17*, 1558–1562.
(30) Annoni, C.; Endoh, T.; Hnedzko, D.; Rozners, E.; Sugimoto, N. *Chem. Commun.* **2016**, *52*, 7935–7938.
(31) Hnedzko, D.; McGee, D. W.; Karamitas, Y. A.; Rozners, E. *RNA* **2017**, *23*, 58–69.
(32) Hoogsteen, K.; IUCr. *Acta Crystallogr.* **1959**, *12*, 822–823.
(33) Sun, J. S.; Hélène, C. *Curr. Opin. Struct. Biol.* **1993**, *3*, 345–356.

(34) Radhakrishnan, I.; Patel, D. J. *Biochemistry* **1994**, *33*, 11405–11416.
(35) Soĭfer, V.; Potaman, V. N. *Triple-Helical Nucleic Acids*; Spinger, **1996**.
(36) Thuong, N. T.; Hélène, C. *Angew. Chem. Int. Ed. English* **1993**, *32*, 666–690.
(37) Buchini, S.; Leumann, C. J. *Curr. Opin. Chem. Biol.* ***2003****, 7, 717–726.*
(38) Asensio, J. L.; Brown, T.; Lane, A. N. *Structure* **1999**, *7*, 1–11.
(39) Rinaldi, C.; Wood, M. J. A. *Nat. Rev. Neurol.* **2017**, *14*, 9–21.
(40) Sardone, V.; Zhou, H.; Muntoni, F.; Ferlini, A.; Falzarano, M.; Sardone, V.; Zhou, H.; Muntoni, F.; Ferlini, A.; Falzarano, M. S. *Molecules* **2017**, *22*, 563.
(41) Goñi, J. R.; Cruz, X. de la; Orozco, M. *Nucleic Acids Res.* **2004**, *32*, 354.
(42) Goñi, J. R.; Vaquerizas, J. M.; Dopazo, J.; Orozco, M. *BMC Genomics* **2006**, *7*, 63.
(43) Frank-Kamenetskii, M. D.; Mirkin, S. M. *Annu. Rev. Biochem.* **1995**, *64*, 65–95.
(44) Zain, R.; Sun, J.-S. *Cell. Mol. Life Sci.* **2003**, *60*, 862–870.
(45) Buske, F. A.; Mattick, J. S.; Bailey, T. L. *RNA Biol.* **2011**, *8*, 427–439.
(46) Belotserkovskii, B. P.; De Silva, E.; Tornaletti, S.; Wang, G.; Vasquez, K. M.; Hanawalt, P. C. *J. Biol. Chem.* **2007**, *282*, 32433–32441.
(47) Wang, G.; Vasquez, K. M. *Proc. Natl. Acad. Sci.* **2004**, *101*, 13448–13453.
(48) Wang, G.; Vasquez, K. M. *DNA Repair (Amst).* **2014**, *19*, 143–151.
(49) Wells, R. D. *FASEB J.* **2008**, *22*, 1625–1634.
(50) Rocher, C.; Dalibart, R.; Letellier, T.; Précigoux, G.; Lestienne, P. *Nucleic Acids Res.* **2001**, *29*, 3320–3326.
(51) Mondal, T.; Subhash, S.; Vaid, R.; Enroth, S.; Uday, S.; Reinius, B.; Mitra, S.; Mohammed, A.; James, A. R.; Hoberg, E.; Moustakas, A.; Gyllensten, U.; Jones, S. J. M.; Gustafsson, C. M.; Sims, A. H.; Westerlund, F.; Gorab, E.; Kanduri, C.; *Nat. Commun.* **2015**, *6*, 7743.
(52) Floris, R.; Scaggiante, B.; Manzini, G.; Quadrifoglio, F.; Xodo, L. E. *Eur. J. Biochem.* **1999**, *260*, 801–809.
(53) Blume, S. W.; Lebowitz, J.; Zacharias, W.; Guarcello, V.; Mayfield, C. A.; Ebbinghaus, S. W.; Bates, P.; Jones, D. E.; Trent, J.; Vigneswaran, N.; Miller, D. M. *Nucleic Acids Res.* **1999**, *27*, 695–702.
(54) Thomas, T.; Thomas, T. J. *Biochemistry* **1993**, *32*, 14068–14074.
(55) Thomas, T. J.; Kulkarni, G. D.; Greenfield, N. J.; Shirahata, A.; Thomas, T. *Biochem. J.* **1996**, *319*, 591–599.
(56) Davis, J. T. *Angew. Chem. Int. Ed.* **2004**, *43*, 668–698.
(57) Arimondo, P. B.; Barcelo, F.; Sun, J. S.; Maurizot, J. C.; Garestier, T.; Hélène, C. *Biochemistry* **1998**, *37*, 16627–16635.
(58) Mergny, J. L.; Lacroix, L.; Hélène, C.; Han, X.; Leroy, J. L. *J. Am. Chem. Soc.* **1995**, *117*, 8887–8898.
(59) Lacroix, L.; Mergny, J. L.; Leroy, J. L.; Hélène, C. *Biochemistry* **1996**, *35*, 8715–8722.
(60) Povsic, T. J.; Dervan, P. B. *J. Am. Chem. Soc.* **1989**, *111*, 3059–3061.
(61) Hildbrand, S.; Blaser, A.; Parel, S. P.; Leumann, C. J. *J. Am. Chem. Soc.* **1997**, *119*, 5499–5511.
(62) Cassidy, S.; Slickers, P.; Trent, J. O.; Capaldi, D. C.; Roselt, P. D.; Reese, C. B.; Neidle, S.; Fox, K. R. *Nucleic Acids Res.* **1997**, *25*, 4891–4898.
(63) Rusling, D. A.; Powers, V. E. C.; Ranasinghe, R. T.; Wang, Y.; Osborne, S. D.; Brown, T.; Fox, K. R. *Nucleic Acids Res.* **2005**, *33*, 3025–3032.
(64) Dong Li Chen; McLaughlin, L. W. *J. Org. Chem.* **2000**, *65*, 7468–7474.
(65) Mills, M.; Arimondo, P. B.; Lacroix, L.; Garestier, T.; Klump, H.; Mergny, J. L. *Biochemistry* **2002**, *41*, 357–366.
(66) Froehler, B. C.; Wadwani, S.; Terhorst, T. J.; Gerrard, S. R. *Tetrahedron Lett.* **1992**, *33*, 5307–5310.
(67) Bijapur, J.; Keppler, M. D.; Bergqvist, S.; Brown, T.; Fox, K. R. *Nucleic Acids Res.* **1999**, *27*, 1802–1809.
(68) Sollogoub, M.; Darby, R. A. J.; Cuenoud, B.; Brown, T.; Fox, K. R. *Biochemistry* **2002**, *41*, 7224–7231.

(69) Brazier, J. A.; Shibata, T.; Townsley, J.; Taylor, B. F.; Frary, E.; Williams, N. H.; Williams, D. M. *Nucleic Acids Res.* **2005**, *33*, 1362–1371.
(70) Rusling, D. A.; Le Strat, L.; Powers, V. E. C.; Broughton-Head, V. J.; Booth, J.; Lack, O.; Brown, T.; Fox, K. R. *FEBS Lett.* **2005**, *579*, 6616–6620.
(71) Walsh, S.; El-Sagheer, A. H.; Brown, T. *Chem. Sci.* **2018**, *9*, 7681–7687.
(72) Mergny, J. L.; Sun, J. S.; Rougée, M.; Montenay-Garestier, T.; Chomilier, J.; Hélène, C.; Barcelo, F. *Biochemistry* **1991**, *30*, 9791–9798.
(73) Greenberg, W. A.; Dervan, P. B. *J. Am. Chem. Soc.* **1995**, *117*, 5016–5022.
(74) Best, G. C.; Dervan, P. B. *J. Am. Chem. Soc.* **1995**, *117*, 1187–1193.
(75) Malnuit, V.; Duca, M.; Benhida, R. *Org. Biomol. Chem.* **2011**, *9*, 326–336.
(76) Chan, P. P.; Glazer, P. M. *J. Mol. Med.* **1997**, *75*, 267–282.
(77) Guianvarc'h, D.; Benhida, R.; Fourrey, J. L.; Maurisse, R.; Sun, J. S. *Chem. Commun.* **2001**, *1*, 1814–1815.
(78) Guianvarc'h, D.; Fourrey, J. L.; Maurisse, R.; Sun, J. S.; Benhida, R. *Org. Lett.* **2002**, *4*, 4209–4212.
(79) Guianvarc'h, D.; Fourrey, J. L.; Maurisse, R.; Sun, J. S.; Benhida, R. *Bioorg. Med. Chem.* **2003**, *11*, 2751–2759.
(80) Wang, Y.; Rusling, D. A.; Powers, V. E. C.; Lack, O.; Osborne, S. D.; Fox, K. R.; Brown, T. *Biochemistry* **2005**, *44*, 5884–5892.
(81) Hari, Y.; Obika, S.; Imanishi, T. *Eur. J. Org. Chem.* **2012**, *2012*, 2875–2887.
(82) Ohkubo, A.; Yamada, K.; Ito, Y.; Yoshimura, K.; Miyauchi, K.; Kanamori, T.; Masaki, Y.; Seio, K.; Yuasa, H.; Sekine, M. *Nucleic Acids Res.* **2015**, *43*, 5675–5686.
(83) Roberts, R. W.; Crothers, D. M. *Science* **1992**, *258*, 1463.
(84) Seidman, M. M.; Puri, N.; Majumdar, A.; Cuenoud, B.; Miller, P. S.; Alam, R. *Ann. N. Y. Acad. Sci.* **2005**, *1058*, 119–127.
(85) Petersen, M.; Wengel, J. *Trends Biotechnol.* **2003**, *21*, 74–81.
(86) Watts, J. K. *Chem. Commun.* **2013**, *49*, 5618.
(87) Lundin, K. E.; Højland, T.; Hansen, B. R.; Persson, R.; Bramsen, J. B.; Kjems, J.; Koch, T.; Wengel, J.; Smith, C. I. E. *Adv. Genet.* **2013**, *82*, 47–107.
(88) Zhou, Y.; Kierzek, E.; Loo, Z. P.; Antonio, M.; Yau, Y. H.; Chuah, Y. W.; Geifman-Shochat, S.; Kierzek, R.; Chen, G. *Nucleic Acids Res.* **2013**, *41*, 6664–6673.
(89) Kool, E. T. *Chem. Rev.* **1997**, *97*, 1473–1488.
(90) Torigoe, H.; Hari, Y.; Sekiguchi, M.; Obika, S.; Imanishi, T. *J. Biol. Chem.* **2001**, *276*, 2354–2360.
(91) Brunet, E.; Alberti, P.; Perrouault, L.; Babu, R.; Wengel, J.; Giovannangeli, C. *J. Biol. Chem.* **2005**, *280*, 20076–20085.
(92) Sun, B. W.; Babu, B. R.; Sørensen, M. D.; Zakrzewska, K.; Wengel, J.; Sun, J. S. *Biochemistry* **2004**, *43*, 4160–4169.
(93) Højland, T.; Kumar, S.; Babu, B. R.; Umemoto, T.; Albæk, N.; Sharma, P. K.; Nielsen, P.; Wengel, J. *Org. Biomol. Chem.* **2007**, *5*, 2375–2379.
(94) Alam, M. R.; Majumdar, A.; Thazhathveetil, A. K.; Liu, S. T.; Liu, J. L.; Puri, N.; Cuenoud, B.; Sasaki, S.; Miller, P. S.; Seidman, M. M. *Biochemistry* **2007**, *46*, 10222–10233.
(95) Rahman, S. M. A.; Seki, S.; Obika, S.; Yoshikawa, H.; Miyashita, K.; Imanishi, T. *J. Am. Chem. Soc.* **2008**, *130*, 4886–4896.
(96) Kim, H. G.; Reddoch, J. F.; Mayfield, C.; Ebbinghaus, S.; Vigneswaran, N.; Thomas, S.; Jones, D. E.; Miller, D. M. *Biochemistry* **1998**, *37*, 2299–2304.
(97) Michel, T.; Debart, F.; Heitz, F.; Vasseur, J.-J. *ChemBioChem* **2005**, *6*, 1254–1262.
(98) Nielsen, P. E. *Chem. Biodivers.* **2010**, *7*, 786–804.
(99) Harborth, J.; Weber, K.; Manninga, H.; Vandenburgh, K.; Tuschl, T.; Scaringe, S. A.; Elbashir, S. M. *Antisense Nucleic Acid Drug Dev.* **2003**, *13*, 83–105.
(100) Sehgal, A.; Vaishnaw, A.; Fitzgerald, K. *Journal of Hepatology*. **2013**, *59*, 1354–1359.

(101) Miller, C. M.; Donner, A. J.; Seth, P. P.; Prakash, T. P.; Swayze, E. E.; Tanowitz, M.; Harris, E. N. *Nucleic Acid Ther.* **2018**, *28*, 119–127.
(102) Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. *Science* **1991**, *254*, 1497–1500.
(103) Egholm, M.; Buchardt, O.; Christensen, L.; Behrens, C.; Freier, S. M.; Driver, D. A.; Berg, R. H.; Kim, S. K.; Norden, B.; Nielsen, P. E. *Nature* **1993**, *365*, 566–568.
(104) Kilså Jensen, K.; Ørum, H.; Nielsen, P. E.; Nordén, B. *Biochemistry* **1997**, *36*, 5072–5077.
(105) Ratilainen, T.; Holmén, A.; Tuite, E.; Nielsen, P. E.; Nordén, B. *Biochemistry* **2000**, *39*, 7781–7791.
(106) Demidov, V. V.; Potaman, V. N.; Frank-Kamenetskil, M. D.; Egholm, M.; Buchard, O.; Sönnichsen, S. H.; Nlelsen, P. E. *Biochem. Pharmacol.* **1994**, *48*, 1310–1313.
(107) Hamilton, S. E.; Iyer, M.; Norton, J. C.; Corey, D. R. *Bioorg. Med. Chem. Lett.* **1996**, *6*, 2897–2900.
(108) Wancewicz, E. V.; Maier, M. A.; Siwkowski, A. M.; Albertshofer, K.; Winger, T. M.; Berdeja, A.; Gaus, H.; Vickers, T. A.; Bennett, C. F.; Monia, B. P.; Griffey, R. H.; Nulf, C. J.; Hu, J.; Corey, D. R.; Swayze, E. E.; Kinberger, G. A. *J. Med. Chem.* **2010**, *53*, 3919–3926.
(109) Hu, J.; Corey, D. R. *Biochemistry* **2007**, *46*, 7581–7589.
(110) Hu, J.; Matsui, M.; Gagnon, K. T.; Schwartz, J. C.; Gabillet, S.; Arar, K.; Wu, J.; Bezprozvanny, I.; Corey, D. R. *Nat. Biotechnol.* **2009**, *27*, 478–484.
(111) Fabani, M. M.; Abreu-Goodger, C.; Williams, D.; Lyons, P. A.; Torres, A. G.; Smith, K. G. C.; Enright, A. J.; Gait, M. J.; Vigorito, E. *Nucleic Acids Res.* **2010**, *38*, 4466–4475.
(112) Fabani, M. M.; Gait, M. J. *RNA* **2008**, *14*, 336–346.
(113) Egholm, M.; Christensen, L.; Dueholm, K. L.; Buchardt, O.; Coull, J.; Nielsen, P. E. *Nucleic Acids Res.* **1995**, *23*, 217–222.
(114) Thadke, S. A.; Hridya, V. M.; Perera, J. D. R.; Gil, R. R.; Mukherjee, A.; Ly, D. H. *Commun. Chem.* **2018**, *1*, 79.
(115) Arya, D. P. *Acc. Chem. Res.* **2011**, *44*, 134–146.
(116) Chaires, J. B.; Ren, J.; Henary, M.; Zegrocka, O.; Reid Bishop, G.; Strekowski, L. *J. Am. Chem. Soc.* **2003**, *125*, 7272–7283.
(117) Strekowski, L.; Hojjat, M.; Wolinska, E.; Parker, A. N.; Paliakov, E.; Gorecki, T.; Tanious, F. A.; Wilson, W. D. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 1097–1100.
(118) Holt, P. A.; Ragazzon, P.; Strekowski, L.; Chaires, J. B.; Trent, J. O. *Nucleic Acids Res.* **2009**, *37*, 1280–1287.
(119) Bergquist, H.; Nikravesh, A.; Fernández, R. D.; Larsson, V.; Nguyen, C.-H.; Good, L.; Zain, R. *ChemBioChem* **2009**, *10*, 2629–2637.
(120) Lane, A.; Jenkins, T. *Curr. Org. Chem.* **2001**, *5*, 845–869.
(121) Arya, D. P.; Coffee, R. L.; Willis, B.; Abramovitch, A. I. *J. Am. Chem. Soc.* **2001**, *123*, 5385–5395.
(122) Arya, D. P.; Coffee, R. L. *Bioorg. Med. Chem. Lett.* **2000**, *10*, 1897–1899.
(123) Arya, D. P.; Coffee R.L., J.; Charles, I. *J. Am. Chem. Soc.* **2001**, *123*, 11093–11094.
(124) Arya, D. P.; Xue, L.; Willis, B. *J. Am. Chem. Soc.* **2003**, *125*, 10148–10149.
(125) Arya, D. P.; Micovic, L.; Charles, I.; Coffee, R. L.; Willis, B.; Xue, L. *J. Am. Chem. Soc.* **2003**, *125*, 3733–3744.
(126) Xi, H.; Arya, D. *Curr. Med. Chem. Agents* **2005**, *5*, 327–338.
(127) Willis, B.; Arya, D. P. *Adv. Carbohydr. Chem. Biochem.* **2006**, *60*, 251–302.
(128) Xi, H.; Kumar, S.; Dosen-Micovic, L.; Arya, D. P. *Biochimie* **2010**, *92*, 514–529.
(129) Xue, L.; Ranjan, N.; Arya, D. P. *Biochemistry* **2011**, *50*, 2838–2849.
(130) Waksman, S. A.; Lechevalier, H. A. *Science* **1949**, *109*, 305–307.
(131) Schatz, A.; Bugle, E.; Waksman, S. A. *Exp. Biol. Med.* **1944**, *55*, 66–69.
(132) Napoli, S.; Carbone, G. M.; Catapano, C. V.; Shaw, N.; Arya, D. P. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 3467–3469.
(133) Granqvist, L.; Tähtinen, V.; Virta, P.; Granqvist, L.; Tähtinen, V.; Virta, P. *Molecules* **2019**, *24*, 580.

(134) Arya, D. P.; Xue, L.; Tennant, P. *J. Am. Chem. Soc.* **2003**, *125*, 8070–8071.
(135) Zain, R.; Marchand, C.; Sun, J. S.; Nguyen, C. H.; Bisagni, E.; Garestier, T.; Hélène, C. *Chem. Biol.* **1999**, *6*, 771–777.
(136) Escudé, C.; Nguyen, C. H.; Kukreti, S.; Janin, Y.; Sun, J. S.; Bisagni, E.; Garestier, T.; Hélène, C. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 3591–3596.
(137) Xue, L.; Xi, H.; Kumar, S.; Gray, D.; Davis, E.; Hamilton, P.; Skriba, M.; Arya, D. P. *Biochemistry* **2010**, *49*, 5540–5552.
(138) Rajeev, K. G.; Jadhav, V. R.; Ganesh, K. N. *Nucleic Acids Res.* **1997**, *25*, 4187–4193.
(139) Stonehouse, T. J.; Fox, K. R. *BBA - Gene Struct. Expr.* **1994**, *1218*, 322–330.
(140) Robles, J.; McLaughlin, L. W. *J. Am. Chem. Soc.* **1997**, *119*, 6014–6021.
(141) Dempcy, R. O.; Kutyavin, I. V; Mills, A. G.; Lukhtanov, E. A.; Meyer, R. B. *Nucleic Acids Res.* **1999**, *27*, 2931–2937.
(142) Silver, G. C.; Sun, J. S.; Nguyen, C. H.; Boutorine, A. S.; Bisagni, E.; Hélène, C. *J. Am. Chem. Soc.* **1997**, *119*, 263–268.
(143) Liu, Y.; Nairn, R. S.; Vasquez, K. M. *Nucleic Acids Res.* **2009**, *37*, 6378–6388.
(144) Charles, I.; Xi, H.; Arya, D. P. *Bioconjug. Chem.* **2007**, *18*, 160–168.
(145) Kiviniemi, A.; Virta, P. *Bioconjug. Chem.* **2011**, *22*, 1559–1566.
(146) Shi, F.; Hoekstra, D. *J. Controlled Release* **2004**, *97*, 189–209.
(147) Manoharan, M. *Antisense Nucleic Acid Drug Dev.* **2002**, *12*, 103–128.
(148) Juliano, R.; Bauman, J.; Kang, H.; Ming, X. *Mol. Pharm.* **2009**, *6*, 686–695.
(149) Khalil, I. A.; Kogure, K.; Akita, H.; Harashima, H. *Pharmacol. Rev.* **2006**, *58*, 32–45.
(150) Debart, F.; Abes, S.; Deglane, G.; Moulton, H.; Clair, P.; Gait, M.; Vasseur, J.-J.; Lebleu, B. *Curr. Top. Med. Chem.* **2007**, *7*, 727–737.
(151) Stein, C. A.; Hansen, J. B.; Lai, J.; Wu, S. J.; Voskresenskiy, A.; Høg, A.; Worm, J.; Hedtjärn, M.; Souleimanian, N.; Miller, P.; Soifer, H. S.; Castanotto, D.; Benimetskaya, L.; Ørum, H.; Koch, T. *Nucleic Acids Res.* **2009**, *38*, e3.
(152) Castanotto, D.; Lin, M.; Kowolik, C.; Wang, L. A.; Ren, X. Q.; Soifer, H. S.; Koch, T.; Hansen, B. R.; Ørum, H.; Armstrong, B.; Wang, Z.; Bauer, P.; Rossi, J.; Stein, C.A. *Nucleic Acids Res.* **2015**, *43*, 9350–9361.
(153) Liang, X. H.; Shen, W.; Sun, H.; Kinberger, G. A.; Prakash, T. P.; Nichols, J. G.; Crooke, S. T. *Nucleic Acids Res.* **2016**, *44*, 3892–3907.
(154) Crooke, S. T.; Wang, S.; Vickers, T. A.; Shen, W.; Liang, X. H. *Nat. Biotechnol.* **2017**, *35*, 230–237.
(155) Dowdy, S. F. *Nat. Biotechnol.* **2017**, *35*, 222–229.
(156) Santhakumaran, L. M.; Thomas, T.; Thomas, T. J. *Nucleic Acids Res.* **2004**, *32*, 2102–2112.
(157) Cheng, K.; Ye, Z.; Guntaka, R. V; Mahato, R. I. *J. Pharmacol. Exp. Ther.* **2006**, *317*, 797–805.
(158) Ohmichi, T.; Kuwahara, M.; Sasaki, N.; Hasegawa, M.; Nishikata, T.; Sawai, H.; Sugimoto, N. *Angew. Chem. Int. Ed.* **2005**, *44*, 6682–6685.
(159) Deglane, G.; Abes, S.; Michel, T.; Prévot, P.; Vives, E.; Debart, F.; Barvik, I.; Lebleu, B.; Vasseur, J.-J. *ChemBioChem* **2006**, *7*, 684–692.
(160) Zhou, P.; Wang, M.; Du, L.; Fisher, G. W.; Waggoner, A.; Ly, D. H. *J. Am. Chem. Soc.* **2003**, *125*, 6878–6879.
(161) Zhou, P.; Dragulescu-Andrasi, A.; Bhattacharya, B.; O'Keefe, H.; Vatta, P.; Hyldig-Nielsen, J. J.; Ly, D. H. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 4931–4935.
(162) Yan, H.; Tram, K. *Glycoconj. J.* **2007**, *24*, 107–123.
(163) Deamond, S. F.; Lee, Y. C.; Flesher, J. E.; Hangeland, J. J.; Ts'o, P. O. P.; Frost, J. J. *Antisense Nucleic Acid Drug Dev.* **1997**, *7*, 141–149.
(164) Duff, R. J.; Deamond, S. F.; Roby, C.; Zhou, Y.; Ts'o, P. O. P. *Methods Enzymol.* **2000**, *313*, 297–321.
(165) Ye, Z.; Guntaka, R. V.; Mahato, R. I. *Biochemistry* **2007**, *46*, 11240–11252.
(166) Lorenz, P.; Misteli, T.; Baker, B. F.; Bennett, C. F.; Spector, D. L. *Nucleic Acids Res.* **2000**, *28*, 582–592.

(167) Giovannangeli, C. .; Hélène, C. *Antisense Nucleic Acid Drug Dev.* **2011**, *7*, 413–421.
(168) Maher, L. J. *Cancer Invest.* **1996**, *14*, 66–82.
(169) Vasquez, Karen M., Wilson, J. H. *Trends Biochem. Sci.* **1998**, *23*, 4–9.
(170) Hewett, P. W.; Daft, E. L.; Laughton, C. A.; Ahmad, S.; Ahmed, A.; Murray, J. C. *Mol. Med.* **2006**, *12*, 8–16.
(171) Svinarchuk, F.; Nagibneva, I.; Cherny, D.; Ait-Si-Ali, S.; Pritchard, L. L.; Robin, P.; Malvy, C.; Harel-Bellan, A.; Chern, D. *Nucleic Acids Res.* **1997**, *25*, 3459–3464.
(172) Karympalis, V.; Kalopita, K.; Zarros, A.; Carageorgiou, H. *Biochem.* **2004**, *69*, 855–860.
(173) Giovannangeli, C.; Hélène, C. *Curr. Opin. Mol. Ther.* **2000**, *2*, 288–296.
(174) Ebbinghaus, S. W.; Fortinberry, H.; Gamper, H. B. *Biochemistry* **1999**, *38*, 619–628.
(175) Giovannangeli, C., Perrouault, L., Escudé, C., Gryaznov, S., Hélène, C. *J. Mol. Biol.* **1996**, *261*, 386–398.
(176) Xu, X. S., Glazer, P. M., Wang, G. *Gene* **2000**, *242*, 219–228.
(177) Ghosh, M. K.; Katyal, A.; Chandra, R.; Brahmachari, V. *Mol. Cell. Biochem.* **2005**, *278*, 147–155.
(178) Faria, M.; Wood, C. D.; Perrouault, L.; Nelson, J. S.; Winter, A.; White, M. R.; Helene, C.; Giovannangeli, C. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 3862–3867.
(179) Song, J., Intody, Z., Li, M., Wilson, J. H. *Gene* **2004**, *324*, 183–190.
(180) Pesce, C. D.; Bolacchi, F.; Bongiovanni, B.; Cisotta, F.; Capozzi, M.; Diviacco, S.; Quadrifoglio, F.; Mango, R.; Novelli, G.; Mossa, G.; Esposito, C.; Ombres, D.; Rocchi, G.; Bergamini, A. *Antiviral Res.* **2005**, *66*, 13–22.
(181) Guieysse, A. L.; Praseuth, D.; Francois, J. C.; Helene, C. *Biochem. Biophys. Res. Commun.* **1995**, *217*, 186–194.
(182) Hélène, C.; Giovannangeli, C.; Rapozzi, V.; Diviacco, S.; Quadrifoglio, F.; Xodo, L. E. *FASEB J.* **2002**, *15*, 2660–2668.
(183) Nagatsugi, F.; Sasaki, S.; Miller, P. S.; Seidman, M. M. *Nucleic Acids Res.* **2003**, *31*, e31.
(184) Vasquez, K. M.; Narayanan, L.; Glazer, P. M. *Science* **2000**, *290*, 530–533.
(185) Rogers, F. A.; Lloyd, J. A.; Glazer, P. M. *Curr Med Chem Anticancer Agents* **2005**, *5*, 319–326.
(186) Perrouault, L.; Asseline, U.; Rivalle, C.; Thuong, N. T.; Bisagni, E.; Giovannangeli, C.; Doan, T. Le; Hélène, C. *Nature* **1990**, *344*, 358–360.
(187) Strobel, S. A.; Dervan, P. B. *Science* **1990**, *249*, 73–75.
(188) François, J. C.; Saison-Behmoaras, T.; Barbier, C.; Chassignol, M.; Thuong, N. T.; Hélène, C. *Proc. Natl. Acad. Sci. U. S. A.* **1989**, *86*, 9702–9706.
(189) Pei, D.; Corey, D. R.; Schultz, P. G. *Proc. Natl. Acad. Sci. U. S. A.* **1990**, *87*, 9858–9862.
(190) Eisenschmidt, K.; Lanio, T.; Simoncsits, A.; Jeltsch, A.; Pingoud, V.; Wende, W.; Pingoud, A. *Nucleic Acids Res.* **2005**, *33*, 7039–7047.
(191) Wang, G.; Levy, D. D.; Seidman, M. M.; Glazer, P. M. *Mol. Cell. Biol.* **1995**, *15*, 1759–1768.
(192) Shahid, K. A.; Majumdar, A.; Alam, R.; Liu, S. T.; Kuan, J. Y.; Sui, X.; Cuenoud, B.; Glazer, P. M.; Miller, P. S.; Seidman, M. M. *Biochemistry* **2006**, *45*, 1970–1978.
(193) Faruqi, A. F.; Datta, H. J.; Carroll, D.; Seidman, M. M.; Glazer, P. M. *Mol. Cell. Biol.* **2000**, *20*, 990–1000.
(194) Chin, J. Y.; Kuan, J. Y.; Lonkar, P. S.; Krause, D. S.; Seidman, M. M.; Peterson, K. R.; Nielsen, P. E.; Kole, R.; Glazer, P. M. *Proc. Natl. Acad. Sci.* **2008**, *105*, 13514–13519.
(195) Bahal, R.; Ali McNeer, N.; Quijano, E.; Liu, Y.; Sulkowski, P.; Turchick, A.; Lu, Y. C.; Bhunia, D. C.; Manna, A.; Greiner, D. L.; Brehm, M. A.; Cheng, C. J.; López-Giráldez, F.; Ricciardi, A.; Beloor, J.; Krause, D. S.; Kumar, P.; Gallagher, P. G.; Braddock, D. T.; Salzman, W. M.; Ly, D. L.;Glazer, P. M. *Nat. Commun.* **2016**, *7*, 13304.
(196) Matteucci, M.; Lin, K. Y.; Huang, T.; Wagner, R.; Sternbach, D. D.; Mehrotra, M.; Besterman, J. M. *J. Am. Chem. Soc.* **1997**, *119*, 6939–6940.
(197) Hausmann, M.; Winkler, R.; Hildenbrand, G.; Finsterle, J.; Weisel, A.; Rapp, A.; Schmitt, E.; Janz, S.; Cremer, C. *Biotechniques* **2003**, *35*, 564–577.
(198) Schluep, T.; Cooney, C. L. *Nucleic Acids Res.* **1998**, *26*, 4524–4528.

(199) Spitzner, J. R.; Chung, I. K.; Muller, M. T. *J. Biol. Chem.* **1995**, *270*, 5932–5943.
(200) Chandrasekaran, A. R.; Rusling, D. A. *Nucleic Acids Res.* **2018**, *46*, 1021–1037.
(201) Hu, Y.; Cecconello, A.; Idili, A.; Ricci, F.; Willner, I. *Angew. Chem. Int. Ed.* **2017**, *56*, 15210–15233.
(202) Streisinger, G.; Okada, Y.; Emrich, J.; Newton, J.; Tsugita, A.; Terzaghi, E.; Inouye, M. *Cold Spring Harb. Symp. Quant. Biol.* **1966**, *31*, 77–84.
(203) Sando, S.; Sasaki, T.; Kanatani, K.; Aoyama, Y. *J. Am. Chem. Soc.* **2003**, *125*, 15720–15721.
(204) Mei, S. H. J.; Liu, Z.; Brennan, J. D.; Li, Y. *J. Am. Chem. Soc.* **2003**, *125*, 412–420.
(205) Liu, G.; Wan, Y.; Gau, V.; Zhang, J.; Wang, L.; Song, S.; Fan, C. *J. Am. Chem. Soc.* **2008**, *130*, 6820–6825.
(206) Kolpashchikov, D. M. *J. Am. Chem. Soc.* **2005**, *127*, 12442–12443.
(207) Stojanovic, M. N.; de Prada, P.; Landry, D. W. *J. Am. Chem. Soc.* **2001**, *123*, 4928–4931.
(208) Ho, H. A.; Doré, K.; Boissinot, M.; Bergeron, M. G.; Tanguay, R. M.; Boudreau, D.; Leclerc, M. *J. Am. Chem. Soc.* **2005**, *127*, 12673–12676.
(209) Brent S. Gaylord; Alan J. Heeger, A.; Bazan, G. C. *J. Am. Chem. Soc.* **2003**, *125*, 896–900.
(210) Tang, Z.; Mallikaratchy, P.; Yang, R.; Kim, Y.; Zhu, Z.; Wang, H.; Tan, W. *J. Am. Chem. Soc.* **2008**, *130*, 11268–11269.
(211) Farjami, E.; Clima, L.; Gothelf, K.; Ferapontova, E. E. *Anal. Chem.* **2011**, *83*, 1594–1602.
(212) Okamoto, A.; Tainaka, K.; Ochi, Y.; Kanatani, K.; Saito, I. *Mol. Biosyst.* **2006**, *2*, 122.
(213) Li, X.; Wang, Y.; Guo, J.; Tang, X. *ChemBioChem* **2011**, *12*, 2863–2870.
(214) Van Daele, I.; Bomholt, N.; Filichev, V. V.; Van Calenbergh, S.; Pedersen, E. B. *ChemBioChem* **2008**, *9*, 791–801.
(215) Tanabe, K.; Tsuda, T.; Ito, T.; Nishimoto, S. *Chem. Eur. J.* **2013**, *19*, 15133–15140.
(216) Riedl, J.; Pohl, R.; Rulíšek, L.; Hocek, M. *J. Org. Chem.* **2012**, *77*, 1026–1044.
(217) Sakamoto, T.; Hasegawa, D.; Fujimoto, K. *Chem. Commun.* **2015**, *51*, 8749–8752.
(218) Sakamoto, T.; Hasegawa, D.; Fujimoto, K. *Analyst* **2016**, *141*, 1214–1217.
(219) Conrad, N. K. *Wiley Interdiscip. Rev. RNA* **2014**, *5*, 15–29.
(220) Bacolla, A.; Wang, G.; Vasquez, K. M. *PLOS Genet.* **2015**, *11*, e1005696.
(221) Butcher, S. E.; Pyle, A. M. *Acc. Chem. Res.* **2011**, *44*, 1302–1311.
(222) Devi, G.; Zhou, Y.; Zhong, Z.; Toh, D.-F. K.; Chen, G. *Wiley Interdiscip. Rev. RNA* **2015**, *6*, 111–128.
(223) Mitton-Fry, R. M.; DeGregorio, S. J.; Wang, J.; Steitz, T. A.; Steitz, J. A. *Science* **2010**, *330*, 1244–1247.
(224) Tycowski, K. T.; Shu, M.; Borah, S.; Shi, M.; Steitz, J. A. *Cell Rep.* **2012**, *2*, 26–32.
(225) Wilusz, J. E.; JnBaptiste, C. K.; Lu, L. Y.; Kuhn, C.-D.; Joshua-Tor, L.; Sharp, P. A. *Genes Dev.* **2012**, *26*, 2392–2407.
(226) Brown, J. A.; Valenstein, M. L.; Yario, T. A.; Tycowski, K. T.; Steitz, J. A. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 19202–19207.
(227) Theimer, C. A.; Blois, C. A.; Feigon, J. *Mol. Cell* **2005**, *17*, 671–682.
(228) Ulyanov, N. B.; Shefer, K.; James, T. L.; Tzfati, Y. *Nucleic Acids Res.* **2007**, *35*, 6150–6160.
(229) Shefer, K.; Brown, Y.; Gorkovoy, V.; Nussbaum, T.; Ulyanov, N. B.; Tzfati, Y. *Mol. Cell. Biol.* **2007**, *27*, 2130–2143.
(230) Kim, N. K.; Zhang, Q.; Zhou, J.; Theimer, C. A.; Peterson, R. D.; Feigon, J. *J. Mol. Biol.* **2008**, *384*, 1249–1261.
(231) Qiao, F.; Cech, T. R. *Nat. Struct. Mol. Biol.* **2008**, *15*, 634–640.
(232) Cash, D. D.; Cohen-Zontag, O.; Kim, N.-K.; Shefer, K.; Brown, Y.; Ulyanov, N. B.; Tzfati, Y.; Feigon, J. *Proc. Natl. Acad. Sci.* **2013**, *110*, 10970–10975.
(233) Appasamy, S. D.; Ramlan, E. I.; Firdaus-Raih, M. *PLoS One* **2013**, *8*, e73984.
(234) Klein, D. J.; Ferré-D'Amaré, A. R. *Science* **2006**, *313*, 1752–1756.
(235) Cochrane, J. C.; Lipchock, S. V.; Strobel, S. A. *Chem. Biol.* **2007**, *14*, 97–105.
(236) Gilbert, S. D.; Rambo, R. P.; Van Tyne, D.; Batey, R. T. *Nat. Struct. Mol. Biol.* **2008**, *15*, 177–182.

(237) Kang, M.; Peterson, R.; Feigon, J. *Mol. Cell* **2009**, *33*, 784–790.
(238) Klein, D. J.; Edwards, T. E.; Ferré-D'Amaré, A. R. *Nat. Struct. Mol. Biol.* **2009**, *16*, 343–344.
(239) Spitale, R. C.; Torelli, A. T.; Krucinska, J.; Bandarian, V.; Wedekind, J. E. *J. Biol. Chem.* **2009**, *284*, 11012–11016.
(240) Batey, R. T. *Q. Rev. Biophys.* **2012**, *45*, 345–381.
(241) Liberman, J. A.; Salim, M.; Krucinska, J.; Wedekind, J. E. *Nat. Chem. Biol.* **2013**, *9*, 353–355.
(242) Klein, D. J.; Schmeing, T. M.; Moore, P. B.; Steitz, T. A. *EMBO J.* **2001**, *20*, 4214–4221.
(243) Nissen, P.; Ippolito, J. A.; Ban, N.; Moore, P. B.; Steitz, T. A. *Proc. Natl. Acad. Sci.* **2001**, *98*, 4899–4903.
(244) Noller, H. F. *Science* **2005**, *309*, 1508–1514.
(245) Rich, A.; Su, L.; Chen, L.; Egli, M.; Berger, J. M. *Nat. Struct. Biol.* **1999**, *6*, 285–292.
(246) Michiels, P. J. A.; Versleijen, A. A. M.; Verlaan, P. W.; Pleij, C. W. A.; Hilbers, C. W.; Heus, H. A. *J. Mol. Biol.* **2001**, *310*, 1109–1123.
(247) Nixon, P. L.; Rangan, A.; Kim, Y. G.; Rich, A.; Hoffman, D. W.; Hennig, M.; Giedroc, D. P. *J. Mol. Biol.* **2002**, *322*, 621–633.
(248) Pallan, P. S.; Marshall, W. S.; Harp, J.; Jewett, F. C.; Wawrzak, Z.; Brown, B. A.; Rich, A.; Egli, M.; Rich, A.; Egli, M. *Biochemistry* **2005**, *44*, 11315–11322.
(249) Giedroc, D. P.; Cornish, P. V. *Virus Res.* **2009**, *139*, 193–208.
(250) Olsthoorn, R. C. L.; Reumerman, R.; Hilbers, C. W.; Pleij, C. W. A.; Heus, H. A. *Nucleic Acids Res.* **2010**, *38*, 7665–7672.
(251) Adams, P. L.; Stahley, M. R.; Kosek, A. B.; Wang, J.; Strobel, S. A. *Nature* **2004**, *430*, 45–50.
(252) Golden, B. L.; Kim, H.; Chase, E. *Nat. Struct. Mol. Biol.* **2005**, *12*, 82–89.
(253) Guo, F.; Gooding, A. R.; Cech, T. R. *Mol. Cell* **2004**, *16*, 351–362.
(254) Toor, N.; Keating, K. S.; Taylor, S. D.; Pyle, A. M. *Science* **2008**, *320*, 77–82.
(255) Calin, G. A.; Dumitru, C. D.; Shimizu, M.; Bichi, R.; Zupo, S.; Noch, E.; Aldler, H.; Rattan, S.; Keating, M.; Rai, K.; Rassenti, L.; Kipps, T.; Negrini, M.; Bullrich, F.; Croce, C. M. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 15524–15529.
(256) Carmona, P.; Molina, M. *Nucleic Acids Res.* **2002**, *30*, 1333–1337.
(257) Escudé, C.; François, J. C.; Sun, J. S.; Ott, G.; Sprinzl, M.; Garestier, T.; Hélène, C. *Nucleic Acids Res.* **1993**, *21*, 5547–5553.
(258) Wilds, C. J.; Damha, M. J. *Bioconjug. Chem.* **1999**, *10*, 299–305.
(259) Han, H.; Dervan, P. B. *Proc. Natl. Acad. Sci. U. S. A.* **1993**, *90*, 3806–3810.
(260) Wang, S.; Kool, E. T. *Nucleic Acids Res.* **1995**, *23*, 1157–1164.
(261) Szabat, M.; Kierzek, E.; Kierzek, R. *Sci. Rep.* **2018**, *8*, 13023.
(262) Toh, D. F. K.; Devi, G.; Patil, K. M.; Qu, Q.; Maraswami, M.; Xiao, Y.; Loh, T. P.; Zhao, Y.; Chen, G. *Nucleic Acids Res.* **2016**, *44*, 9071–9082.
(263) Ong, A. A. L.; Toh, D. F. K.; Patil, K. M.; Meng, Z.; Yuan, Z.; Krishna, M. S.; Devi, G.; Haruehanroengra, P.; Lu, Y.; Xia, K.; Okamura, K.; Sheng, J.; Chen, G. *Biochemistry* **2019**, *58*, 1319–1331.
(264) Kotikam, V.; Kennedy, S. D.; MacKay, J. A.; Rozners, E. *Chem. Eur. J.* **2019**, *25*, 4367–4372.
(265) Hansen, M. E.; Bentin, T.; Nielsen, P. E. *Nucleic Acids Res.* **2009**, *37*, 4498–4507.
(266) Buchini, S.; Leumann, C. J. *Angew. Chem. Int. Ed.* **2004**, *43*, 3925–3928.
(267) Gildea, B.; Mclaughlin, L. W. *Nucleic Acids Res.* **1989**, *17*, 2261–2281.
(268) Kumar, V. A.; Ganesh, K. N. *Acc. Chem. Res.* **2005**, *38*, 404–412.
(269) Corradini, R.; Sforza, S.; Tedeschi, T.; Totsingan, F.; Manicardi, A.; Marchelli, R. *Curr. Top. Med. Chem.* **2011**, *11*, 1535–1554.
(270) Dragulescu-Andrasi, A.; Zhou, P.; He, G.; Ly, D. H. *Chem. Commun.* **2005**, *0*, 244–246.
(271) Manicardi, A.; Fabbri, E.; Tedeschi, T.; Sforza, S.; Bianchi, N.; Brognara, E.; Gambari, R.; Marchelli, R.; Corradini, R. *ChemBioChem* **2012**, *13*, 1327–1337.
(272) Oyaghire, S. N.; Cherubim, C. J.; Telmer, C. A.; Martinez, J. A.; Bruchez, M. P.; Armitage, B. A. *Biochemistry* **2016**, *55*, 1977–1988.
(273) Sugiyama, T.; Kittaka, A. *Molecules* **2013**, *18*, 287–310.

(274) Caliskan, N.; Peske, F.; Rodnina, M. V. *Trends Biochem. Sci.* **2015**, *40*, 265–274.
(275) Garcia-Miranda, P.; Becker, J. T.; Benner, B. E.; Blume, A.; Sherer, N. M.; Butcher, S. E. *J. Virol.* **2016**, *90*, 6906–6917.
(276) Kesy, J.; Patil, K. M.; Kumar, S. R.; Shu, Z.; Yong, H. Y.; Zimmermann, L.; Ong, A. A. L.; Toh, D.-F. K.; Krishna, M. S.; Yang, L.; Decout, J.-L.; Luo, D.; Prabakaran, M.; Chen, G.; Kierzek, E. *Bioconjug. Chem.* **2019**, *30*, 931–943.
(277) Mergny, J.-L.; Lacroix, L. *Oligonucleotides* **2003**, *13*, 515–537.
(278) Mergny, J.-L.; Li, J.; Lacroix, L.; Amrane, S.; Chaires, J. B. *Nucleic Acids Res.* **2005**, *33*, e138–e138.
(279) Gray, D. M.; Hung, S. H.; Johnson, K. H. *Methods Enzymol.* **1995**, *246*, 19–34.
(280) Kypr, J.; Kejnovská, I.; Renčiuk, D.; Vorlíčková, M. *Nucleic Acids Res.* **2009**, *37*, 1713–1725.
(281) Rajagopal, P.; Feigon, J. *Nature* **1989**, *339*, 637–640.
(282) Yang, M.; Ghosh, S. S.; Millar, D. P. *Biochemistry* **1994**, *33*, 15329–15337.
(283) 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.
(284) Hoyne, P. R.; Gacy, A. M.; McMurray, C. T.; Maher, L. J. *Nucleic Acids Res.* **2000**, *28*, 770–775.
(285) Plum, G. E. *Biopolymers* **1997**, *44*, 241–256.
(286) Rozners, E.; Pilch, D. S.; Egli, M. *Curr. Protoc. Nucleic Acid Chem.* **2015**, *63*, 7.4.1-7.4.12.
(287) Chen, H.; Viel, S.; Ziarelli, F.; Peng, L. *Chem. Soc. Rev.* **2013**, *42*, 7971.
(288) Rastinejad, F.; Evilia, C.; Lu, P. *Methods Enzymol.* **1995**, *261*, 560–575.
(289) Gerig, J. T. *Prog. Nucl. Magn. Reson. Spectrosc.* **1994**, *26*, 293–370.
(290) Cobb, S. L.; Murphy, C. D. *J. Fluor. Chem.* **2009**, *130*, 132–143.
(291) Guo, F.; Li, Q.; Zhou, C. *Org. Biomol. Chem.* **2017**, *15*, 9552–9565.
(292) Kreutz, C.; Kählig, H.; Konrat, R.; Micura, R. *Angew. Chem. Int. Ed.* **2006**, *45*, 3450–3453.
(293) Sakamoto, T.; Hayakawa, H.; Fujimoto, K. *Chem. Lett.* **2011**, *40*, 720–721.
(294) Ishizuka, T.; Yamashita, A.; Asada, Y.; Xu, Y. *ACS Omega* **2017**, *2*, 8843–8848.
(295) Olejniczak, M.; Gdaniec, Z.; Fischer, A.; Grabarkiewicz, T.; Bielecki, L.; Adamiak, R. W. *Nucleic Acids Res.* **2002**, *30*, 4241–4249.
(296) Scott, L. G.; Geierstanger, B. H.; Williamson, J. R.; Hennig, M. *J. Am. Chem. Soc.* **2004**, *126*, 11776–11777.
(297) Olsen, G. L.; Edwards, T. E.; Deka, P.; Varani, G.; Sigurdsson, S. T.; Drobny, G. P. *Nucleic Acids Res.* **2005**, *33*, 3447–3454.
(298) Kiviniemi, A.; Virta, P. *J. Am. Chem. Soc.* **2010**, *132*, 8560–8562.
(299) Scott, L. G.; Hennig, M. *Methods Enzymol.* **2016**, *566*, 59–87.
(300) Hennig, M.; Scott, L. G.; Sperling, E.; Bermel, W.; Williamson, J. R. *J. Am. Chem. Soc.* **2007**, *129*, 14911–14921.
(301) Rieder, U.; Kreutz, C.; Micura, R. *Proc. Natl. Acad. Sci.* **2010**, *107*, 10804–10809.
(302) Aigner, M.; Haller, A.; Micura, R.; Rieder, U.; Blanchard, S. C. *Nat. Chem. Biol.* **2011**, *7*, 393–400.
(303) Santner, T.; Rieder, U.; Kreutz, C.; Micura, R. *J. Am. Chem. Soc.* **2012**, *134*, 11928–11931.
(304) Sochor, F.; Silvers, R.; Müller, D.; Richter, C.; Fürtig, B.; Schwalbe, H. *J. Biomol. NMR* **2016**, *64*, 63–74.
(305) Hammann, C.; Norman, D. G.; Lilley, D. M. J. *Proc. Natl. Acad. Sci.* **2002**, *98*, 5503–5508.
(306) Zhao, C.; Bachu, R.; Popović, M.; Devany, M.; Brenowitz, M.; Schlatterer, J. C.; Greenbaum, N. L. *RNA* **2013**, *19*, 561–573.
(307) Zhao, C.; Anklin, C.; Greenbaum, N. L. *Methods Enzymol.* **2014**, *549*, 267–285.
(308) Zhou, L.; Rajabzadeh, M.; Traficante, D. D.; Cho, B. P. *J. Am. Chem. Soc.* **1997**, *119*, 5384–5389.
(309) Cho, B. P.; Zhou, L. *Biochemistry* **1999**, *38*, 7572–7583.
(310) Bao, H.-L.; Ishizuka, T.; Sakamoto, T.; Fujimoto, K.; Uechi, T.; Kenmochi, N.; Xu, Y. *Nucleic Acids Res.* **2017**, *45*, 5501–5511.

(311) Bao, H. L.; Ishizuka, T.; Iwanami, A.; Oyoshi, T.; Xu, Y. *ChemistrySelect* **2017**, *2*, 4170–4175.
(312) Ishizuka, T.; Zhao, P. Y.; Bao, H. L.; Xu, Y. *Analyst* **2017**, *142*, 4083–4088.
(313) 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.
(314) Granqvist, L.; Virta, P. *Chem. Eur. J.* **2016**, *22*, 15360–15372.
(315) Tanabe, K.; Sugiura, M.; Nishimoto, S. I. *Bioorg. Med. Chem.* **2010**, *18*, 6690–6694.
(316) Granqvist, L.; Virta, P. *J. Org. Chem.* **2014**, *79*, 3529–3536.
(317) Granqvist, L.; Virta, P. *J. Org. Chem.* **2015**, *80*, 7961–7970.
(318) Kiviniemi, A.; Virta, P.; Lönnberg, H. *Bioconjug. Chem.* **2008**, *19*, 1726–1734.
(319) Kiviniemi, A.; Virta, P.; Lönnberg, H. *Bioconjug. Chem.* **2010**, *21*, 1890–1901.
(320) Abendroth, F.; Bujotzek, A.; Shan, M.; Haag, R.; Weber, M.; Seitz, O. *Angew. Chem. Int. Ed.* **2011**, *50*, 8592–8596.
(321) Seela, F.; Schweinberger, E.; Xu, K.; Sirivolu, V. R.; Rosemeyer, H.; Becker, E. M. *Tetrahedron* **2007**, *63*, 3471–3482.
(322) Huang, C.; Bi, G.; Miller, P. S. *Nucleic Acids Res.* **1996**, *24*, 2606–2613.
(323) Eldrup, A. B.; Dahl, O.; Nielsen, P. E. *J. Am. Chem. Soc.* **1997**, *119*, 11116–11117.
(324) Seitz, O.; Köhler, O. *Chem. Eur. J.* **2001**, *7*, 3911–3925.
(325) Panda, G.; Rao, N. V. *Synlett* **2004**, *2004*, 714–716.
(326) Hnedzko, D.; McGee, D. W.; Rozners, E. *Bioorg. Med. Chem.* **2016**, *24*, 4199–4205.
(327) Song, B.; Wang, Y.; Titmus, M. A.; Botchkina, G.; Formentini, A.; Kornmann, M.; Ju, J. *Mol. Cancer* **2010**, *9*, 96.
(328) Jin, Z.; Selaru, F. M.; Cheng, Y.; Kan, T.; Agarwal, R.; Mori, Y.; Olaru, A. V; Yang, J.; David, S.; Hamilton, J. P.; Duncan, M.; Montgomery, E. A.; Meltzer, S. J. *Oncogene* **2011**, *30*, 1577–1585.
(329) Rothbard, J. B.; Kreider, E.; VanDeusen, C. L.; Wright, L.; Wylie, B. L.; Wender, P. A. *J. Med. Chem.* **2002**, *45*, 3612–3618.
(330) Mitchell, D. J.; Steinman, L.; Kim, D. T.; Fathman, C. G.; Rothbard, J. B. *J. Pept. Res.* **2000**, *56*, 318–325.
(331) Lee, Y.; Zeng, H.; Ruedisser, S.; Gossert, A. D.; Hilty, C. *J. Am. Chem. Soc.* **2012**, *134*, 17448–17451.
(332) Dragulescu-Andrasi, A.; Rapireddy, S.; Frezza, B. M.; Gayathri, C.; Gil, R. R.; Ly, D. H. *J. Am. Chem. Soc.* **2006**, *128*, 10258–10267.

UNIVERSITY OF TURKU

Painosalama Oy, Turku , Finland 2019

ISBN 978-951-29-7785-7 (PRINT)
ISBN 978-951-29-7786-4 (PDF)
ISSN 0082-7002 (PRINT) ISSN 2343-3175 (ONLINE)