Downloaded from orbit.dtu.dk on: Apr 09, 2024

DTU Library

Novel tools for ultra-specific targeting of nucleic acids

Taskova, Maria

Publication date:
2019

Document Version
Publisher's PDF, also known as Version of record

Link back to DTU Orbit

Citation (APA):
Taskova, M. (2019). *Novel tools for ultra-specific targeting of nucleic acids*. Technical University of Denmark.

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Novel tools for ultra-specific targeting of nucleic acids

PhD Thesis

Maria Taskova

Supervised by Professor Kira Astakhova

Department of Chemistry

Technical University of Denmark

April, 2019

To my family,

Maria Taskova
Copenhagen, 2019

Preface

Source for this PhD thesis is the work carried out from April 2016 to April 2019 under supervision of Associate Professor Kira Astakhova. The first part of the work (April 2016 - September 2017) was conducted at the Department of Physics, Chemistry and Pharmacy, Faculty of Science, University of Southern Denmark (SDU). The second part of the work (September 2017 - April 2019) was carried out at the Department of Chemistry, Technical University of Denmark (DTU). During the PhD studies, an external research stay (January 2018-May 2018) was completed at the Beckman Research Institute, City Of Hope, Duarte, USA under supervision of Associate Professor Kevin Morris.

In the past three years, I assisted and co-supervised five bachelor and three master students in the Astakhova laboratory. Beside the research, I taught undergraduate students in Chemistry of the Elements, Organic Chemistry and Nucleic acids and medicinal chemistry. Moreover, I attended and passed the following PhD courses: Cross-institutional PhD course in Molecular Biophysics, Advanced Microscopy and Biophotonics, Biophysics PhD seminar, Power performance workshop and Responsible conduct of Research.

I gave talk for my research at international conferences and symposiums: International Research Forum (Beijing, China 2016), BIT's 7th Annual Symposium of Drug Delivery Systems (Prague, Czech Republic, 2017) and 7th EUCheMS Chemistry Congress (Liverpool, UK, 2018). I also presented scientific posters: XXII IRT of Nucleosides, Nucleotides and Nucleic acids (Paris, France, 2016) and Oligonucleotides and Peptides Therapeutics (TIDES) (San Diego, USA, 2017).

Villum Fonden (Reward nr. 13152) funded the PhD program. Otto Mønsted Fonden partially supported the external research stay.

Maria Taskova

Kongens Lyngby, 2019

Papers included in the thesis/Contribution

Paper I

Synthetic Nucleic Acid Analogues in Gene Therapy: An update for Peptide-Oligonucleotide Conjugates

Taskova M, Mantsiou A, Astakhova K., *Chembiochem.,* 2017, 17, 1676-82.

Contribution to the paper: Involvement in the planning and the design of the article structure; Conduction of literature review; Writing the first draft of a complete review article; Development of tables and figures; Active involvement into the submission and review process. Two revision and improvement cycles conducted with the journal editor.

Paper II

Antisense Oligonucleotides Internally Labeled with Peptides Show Improved Target Recognition and Stability to Enzymatic Degradation

Taskova M, Madsen CS, Jensen KJ, Hansen LH, Vester B, Astakhova K., *Bioconjug Chem.*, 2017, 3, 768-74.

Contribution to the paper: Development of specific research questions; Establishment of research methodology; Conduction of literature review; Optimization and execution of the synthesis; Characterization of all probes apart from peptide synthesis; Performance of the biophysical studies and the serum stability studies; Analysis and representation of the results; Writing the first draft of a complete article; Development of tables and figures; Active involvement into the submission and review process. One revision and enhancement cycle conducted with the journal editor.

Paper III

Studies of Impending Oligonucleotide Therapeutics in Simulated Biofluids

Domljanovic I, Hansen AH, Hansen LH, Klitgaard JK, Taskova M*, Astakhova K., *Nucleic Acid Ther.*, 2018, 28, 348-56.

Contribution to the paper: Formulation of the conceptual framework; Development of specific research questions; Establishment of research methodology; Conduction of literature review; Daily supervision under assay performance and trouble shouting of the technical issues; Analysis and interpretation of the results; Active involvement in the writing of complete article. Submission and review process. Two-enhancement cycles conducted with the journal editor.

Paper IV (In revision, *Bioconjugate Chemistry*, submitted 21th January.2019)

Fluorescent oligonucleotides with bis (prop-2-yn-1-yloxy) butane-1,3-diol scaffold allow for rapid detection of disease associated nucleic acids

Taskova M, Astakhova K.

Contribution to the paper: Development of specific research questions; Establishment of research methodology; Conduction of literature review; Synthesis of the new bis-alkyne scaffold; Optimization and execution of the synthesis of the fluorescent oligonucleotides; Performance of the biophysical studies; Performance of the solid-phase hybridization fluorescence assay; Analysis, interpretation and representation of the results; Writing the first draft of a complete manuscript; Development of tables and figures; Active involvement into the submission process and revision process.

Chapter 5

The projects described in Chapter 5 include studies carried out at the laboratory of Prof. Kevin Morris at the Center for Gene Therapy, COH, USA. (Not published work)

5.1 siRNA gene therapeutics that target the HIV 1 362 site in the 5'LTR

Input: Preparation of the nanoplexes; Involvement in the synthesis and characterization of the conjugates; Performance of the cell culture studies; Analysis of the results.

5.2 Repression of lnRNA (BGas) for Cystic Fibrosis Transmembrance Conductance Regulator expression

Input: Preparation of the nanoplexes; Involvement in the synthesis and characterization of the conjugates; Performance of the cell culture studies; Execution of the quantification assays; Analysis of the results.

Acknowledgements

I truly appreciate the support and the help that I had during performance of the projects, manuscripts writing and thesis preparation. Therefore, I wish to express my gratitude to everyone connected with this work.

Foremost, I truly thank Professor Kira Astakhova for giving me the opportunity to undertake this PhD program and for her engagement as a supervisor. I sincerely appreciate all great discussions, ideas, motivation and the endless inspiration. It was pleasure to be part of productive working environment with high collaborative feeling and appreciation.

Next, I acknowledge all present and former members in Astakhova group for the enjoyable work environment. Especially, I thank Jonas Hansen, Anders H. Hansen, and Mick Hornum for all inspirational deliberations. Credits go to Ivana Domljanovic, Mads W. Mulberg, Jesper Uhd, Sofie Slott, Petya Popova and Sangita Khatri for the wonderful lab work done together.

I would like to express my deep gratitude to Professor Kevin Morris for welcoming me in his group at the Beckman Research Institute, COH, USA. Especially, I thank Tristan Scott, Roslyn Ray and Olga Villamizar for hospitality and guidance into tissue culture work.

I acknowledge our collaborators for their support. I would like to thank the group of Professor Birte Vester. In particular, I thank Lykke H. Hansen for her assistance and support during the radioactive biostability studies. In addition, I would like to thank the group of Professor Knud J. Jensen. Charlotte S. Madsen is acknowledged for providing us with peptide sequences.

I thank Professor Mads H. Clausen and the technical stuff at the department of Chemistry DTU for their great support. Special thanks to the technical stuff at the department of Physic, Chemistry and Pharmacy, SDU and the NAC crew, Professor Jesper Wengel, Per Trolle Jørgensen, Tina G. Hansen and Joan Hansen.

Lastly, I would like to thank my Parents, my brother Vlatko, Gjorgi, Teodora and my entire Family and Friends for their endless support and love.

Abbreviations

ASO – Antisense oligonucleotide

AGO – Argonaute protein

CRISPR - Clustered regularly interspaced short palindromic repeats

CFTR - Cystic fibrosis transmembrane conductance regulator

CuAAC – Copper catalyzed azide-alkyne cycloaddition

DCC - Dicyclohexyl carbodiimide

DMT - Dimethoxytrityl

DMD – Duchene muscular dystrophy

GalNAc - N-acetylgalactosamine

GFP – Green fluorescent protein

GIT – Gastrointestinal tract

HPLC – High performance liquid chromatography

lnRNA – Long noncoding RNA

LNA – Locked nucleic acid

LTR – Long terminal repeat

MALDI - Matrix-assisted laser desorption/ionization

miRNA - microRNA

ODC – Oligonucleotide-dye conjugate

POCs – Peptide-oligonucleotide conjugate

PCR – Polymerase chain reaction

PAM – Protospacer adjacent motif

PBS – Phosphate buffered saline

PEG – Polyethylene glycol

PS - Phosphorotioate

PMO - Phosphorodiamidate morpholino

RISC - RNA-induced silencing complex

RNAi – RNA interference

RLU – Relative light units

SD – Standard deviation

STR – Short tandem repeat

SMA – Spinal muscular dystrophy

SNP – Single nucleotide polymorphism

SPAAK – Stain promoted azide-alkyne cycloaddition

TAR – Trans-activator response region

TLC – Thin layer chromatography

UNA – Unlocked nucleic acid

2'F - 2 fluoro

2'OMe - 2'-O-methyl

2'OMOE - 2'-O-(-2-methoxyethyl)

Table of Contents

Preface v
Papers included in the thesis/Contribution vii
Acknowledgements ix
Abbreviations xi
Abstract xv
Resume xvii
Chapter 1 1
Aims and Study Design 1
1.1. Research Questions and Hypothesis 1
1.2. Aims 2
1.3. Study design 3
Chapter 2 5
Intro 5
2.1 Nucleic acid discovery, structure and biological function 5
2.2 Synthetic nucleic acids – chemical synthesis development and applications 8
2.3 Synthetic oligonucleotides as therapeutics; mechanisms of action 11
2.3.1 Antisense oligonucleotides (ASOs) 13
2.3.2 siRNAs 14
2.3.3 CRISPR Cas 15
2.4 Chemical modifications of therapeutic oligonucleotides 16
2.4.1 Modifications that increase the stability towards nuclease degradation 17
2.4.2 Modifications that improve target binding 19
2.4.3 Prominent modifications headed for targeted intracellular delivery 20
2.5 Clinical status of ASOs, siRNA and CRISPR Cas9 therapeutics 21
2.6 Synthetic Nucleic Acid Analogues in Gene Therapy: An update for Peptide-Oligonucleotide Conjugates 25
Chapter 3 39
Antisense Oligonucleotides Internally Labeled with Peptides Show Improved Target Recognition and Stability to Enzymatic Degradation 39
Chapter 4 67
Studies of Impending Oligonucleotide Therapeutics in Simulated Biofluids 67
Chapter 5 85
5.1 siRNA gene therapeutics that target the HIV 1 362 site in the 5'LTR 85
5.2 Repression of lnRNA (BGas) for Cystic Fibrosis Transmembrance Conductance Regulator expression . 98

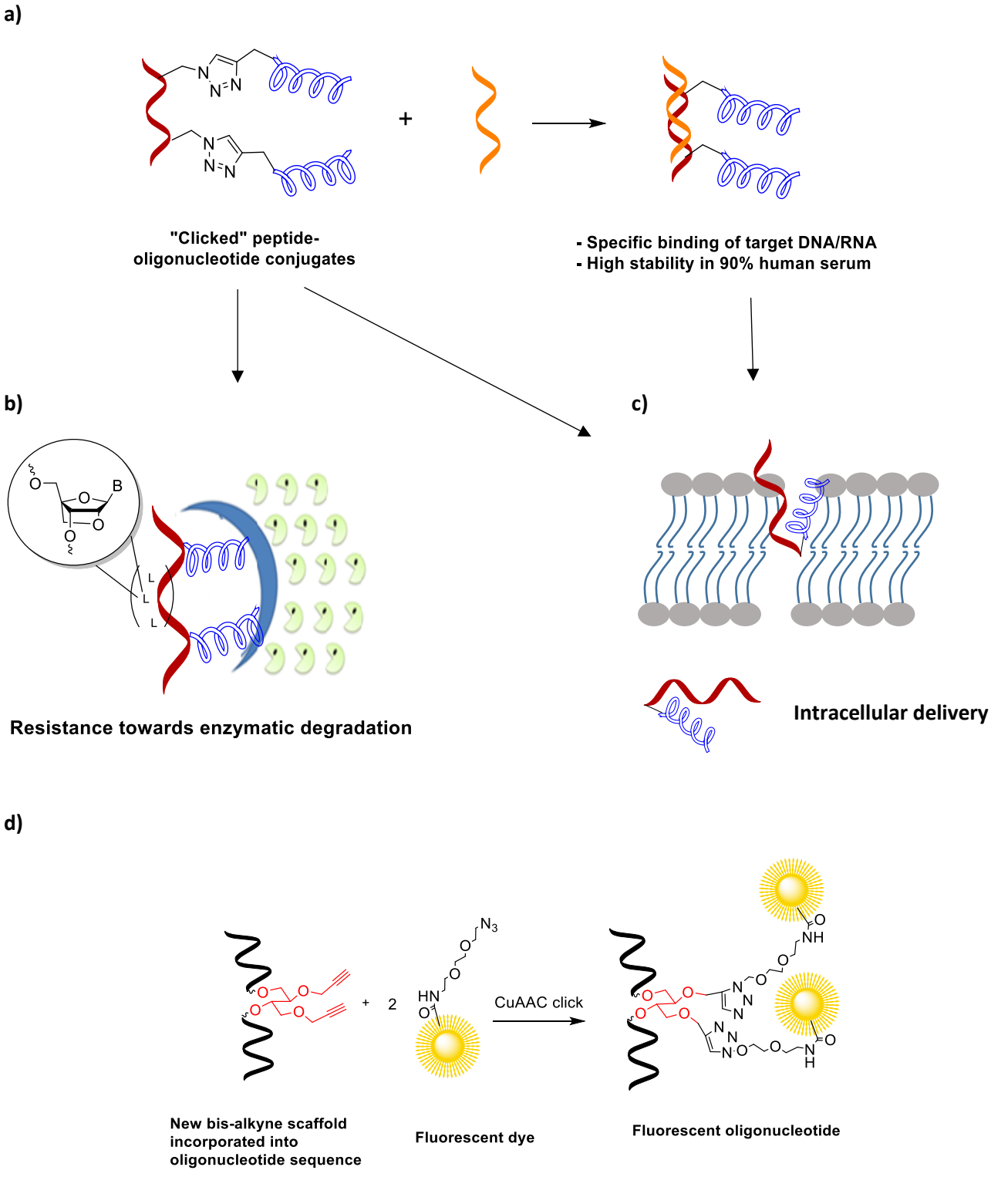
Chapter 6 107
Fluorescent oligonucleotides with bis (prop-2-yn-1-yloxy) butane-1,3-diol scaffold allow for rapid detection of disease associated nucleic acids 107
Chapter 7 139
Summary of Results and General Discussion 139
Chapter 8 145
Concluding Remarks and Perspectives 145
REFERENCES 147
Appendix A 163
Appendix B 171
Declaration of co-authorship 171

Abstract

With the development of phosphoramidite chemistry, the use of synthetic oligonucleotides grows progressively. Because of their structure, periodicity and specific interactions, short oligonucleotides find application in numerous scientific fields. Some of them include biology, biomedicine and clinical diagnostics. The modifications in their chemical structure did give rise to more stable and more efficient oligonucleotide sequences with higher potential for practical application as therapeutics and diagnostics tools.

Chapters 3, 4 and 5 in this thesis, describe the synthesis and the characterization of advanced oligonucleotide therapeutics. We successfully synthetized a library of peptide-oligonucleotide conjugates (POCs) model therapeutics and investigated their biophysical properties *in vitro*. Our findings confirmed the POCs ability for efficient target binding and single nucleotide polymorphism (SNP) discrimination. Additionally, the serum stability studies confirmed higher stability of the POCs when compared with the naked-parent oligonucleotides. Moreover, we formulated model oligonucleotides as complexes and conjugates with peptides and proteins and we investigated further their stability in various simulated bio-fluids. The data revealed that both strategies improve the stability; however, the effect of the conjugated peptide was superior over the complexed one. In essence, the nature of the peptide sequence influenced the oligonucleotide stability in the simulated gastric juice. Finally, we formulated complexed and conjugated oligonucleotide therapeutics with peptide sequences, designed to reduce the HIV1-mRNA or the BGas lnRNA expression and function. The cell culture experiments showed promising effect and cell internalization for the complexes and conjugates. Nevertheless, the variability within the quantitative results was high.

Chapter 6 covers the work on developing a simple and cost-effective strategy for modification and functionalization of synthetic oligonucleotides as fluorescent diagnostic tools. It also describes the biophysical studies and the ability for single mutation detection by fluorescence. We synthetized new bis-alkyne nucleic acids building block that afterwards was incorporated into oligonucleotide sequences. This followed successful post-synthetic conjugation with three different fluorescent dyes. According to our findings, while the oligonucleotide-dye conjugates (ODCs) with perylene intercalated within the duplex efficient and distinguished SNP by shift in the fluorescence intensities at emission maxima, the ODCs with 5JOE and PEP failed to discriminate SNP by temperature melting and fluorescence measurements experiments. In addition, the ODCs with perylene detected mutant DNA extracted from cancer cell lines.



a) Peptide-oligonucleotide conjugates show high affinity for target binding; b) Peptides and proteins conjugated and complexed to oligonucleotide improve its stability in bio-fluids; c) Improved intracellular delivery of the oligonucleotide conjugated with peptide; d) Oligonucleotides with new bis-alkyne building block conjugated with fluorescent dyes, detect single point mutation in extract from cancer cell line.

Resume

Med udviklingen af phosphoramidit syntesen blev brugen af syntetiske oligonukleotider øget markant. Korte oligonukleotider har fundet bred anvendelse i forskellige forskningsområder grundet deres regelmæssige struktur og specifikke interaktioner. Disse inkluderer kemisk biologi, biomedicin og klinisk diagnostik. Modifikationer af oligonukleotiders kemiske struktur har ledt til udviklingen af mere stabile og effektive sekvenser med større potentiale som lægemidler og diagnostiske værktøjer til klinisk brug.

Kapitel 3, 4 og 5 i denne afhandling beskriver syntesen og karakterisering af avancerede oligonukleotid-lægemidler. Vi syntetiserede et bibliotek af peptid-oligonukleotid konjugater (POC) som modellægemidler, og undersøgte deres biofysiske egenskaber *in vitro*. Vores resultater bekræftede, at konjugaterne binder effektivt til deres komplimentære sekvens og kan identificere sekvenser med mutation på én enkelt nukleobase. Ydermere viste stabilitets studier i plasma, at konjugaterne har øget stabilitet sammenlignet med de frie oligonukleotider. En række komplekser og konjugater blev derefter formuleret, hvor oligonukleotiderne blev kombineret med peptider og proteiner. Formuleringerne blev testet i forskellige simulerede biovæsker, hvor resultaterne viste, at begge strategier øger stabiliteten af oligonukleotiderne, men. Konjugering med peptiderne var mere effektivt end kompleksering. Detaljeret analyse viste hvordan sekvensen af de konjugerede peptider ændrede stabiliteten af oligonukleotiderne i simuleret mave syre. Slutteligt blev en serie af terapeutiske oligonukleotider designet til at reducere HIV1-mRNA eller BGas lnRNA ekspression og funktion. Oligonukleotiderne blev komplekseret og konjugeret med peptider og testet i cellestudier, hvilket viste lovende resultater med forbedret optag i cellerne. Der var dog stor variation i de kvantitative resultater.

Kapitel 6 omhandler udviklingen af en simpel og omkostningseffektiv metode til at funktionalisere og modificere oligonukleotider som fluorescente værktøjer til diagnostik. Ligeledes bliver de biofysiske studier samt muligheden for at identificere mutationer af enkelte nukleobaser beskrevet. Vi syntetiserede en ny bis-alkyn nukleinsyre byggeblok, som efterfølgende blev inkorporeret i en oligonukleotid sekvens. Derefter blev sekvensen konjugeret med 3 fluorescente farvestoffer. Ifølge vores resultater er oligonukleotid-farvestofkonjugatet med perylen i stand til at identificere mutationer af enkelte nukleobaser via interkalering i det resulterende duplex, hvilket ændrer intensiteten ved emissionens maksimum. Lignende effekter blev ikke observeret for konjugater med 5JOE eller PEP farvestoffer. Ydermere var perylen-konjugatet i stand til at detektere muteret DNA ekstraheret fra cancer celle linjer.

Chapter 1
Aims and Study Design

Literature emphasize nucleic acids as a very promising tools for meeting the need of advanced treatment and diagnosis of otherwise not targetable diseases. Short oligonucleotides manipulate on genetic level and have properties that are beyond the characteristics of small molecules and protein therapeutics. In addition, they are suitable tools for molecular diagnostics and superlative building blocks in nanotechnology. Nevertheless, after more than fifty years extensive research, it is evident that for better translation from bench side to the clinics, oligonucleotides as therapeutics and diagnostic tools need to overcome vast challenges.

1.1. Research Questions and Hypothesis

1) Do rationally designed oligonucleotide therapeutics conjugated with peptide sequences have optimal biophysical properties?
2) How different formulation approaches can affect the stability of oligonucleotide sequences in various bio-fluids?
3) What is the intracellular delivery efficiency and the potency of logically formulated oligonucleotide therapeutics as nanoplexes and conjugates with peptides?
4) Can we detect single point DNA mutation by simple fluorescence method using synthetic multi-fluorescent oligonucleotides?

Subsequent the research questions we give the following **Hypotheses**:

- Oligonucleotide conjugates with short peptides have improved properties as specific target binding and good stability in bio-fluids;
- Covalently conjugated peptide is superior in protecting the oligonucleotide sequence from enzymatic degradation compared to complexed one;
- Efficient intracellular delivery of therapeutic oligonucleotides is achieved by conjugation with rationally designed peptide sequence and N-acetylgalactosamine (GalNAc) molecule.
- Single gene mutations can be distinguished by multi-fluorescent oligonucleotides without the use of robust sequencing methods;

1.2. Aims

Overall aim: Development of new modified nucleic acids applicable in the discovery of next generation oligonucleotide therapeutics and in the new methods for efficient diagnosis on the molecular level.

Specific aims:

1) Synthesis of a library of peptide-oligonucleotide conjugates designed to target *BRAF V600E* oncogene *in vivo*;
   - Optimization of a method for double internal covalent conjugation of a peptide;
   - Investigation of the target binding specificity and stability in 90% human serum. (Paper II)

2) Detailed studies on stability of differently formulated oligonucleotide therapeutics in simulated bio-fluids such us human serum, gastric juice and hydrochloric acid;
   - Assessment of stability of complexed and conjugated oligonucleotides with short peptide/protein sequences. (Paper III)

3) Development of next generation oligonucleotide therapeutics for targeted intracellular delivery and estimation on their efficiency in cell culture studies. (Chapter 5)

4) Development of multi-fluorescent oligonucleotides for detection of single point mutation by fluorescence.
   - Synthesis of new bis-alkyne scaffold monomer for oligonucleotide functionalization;
   - Incorporation of the scaffold into sequence specific oligonucleotide with subsequent conjugation with three different fluorescent dyes;
   - Investigation of the biophysical properties of the fluorescent oligonucleotides and estimation of their ability to sense target binding by shift in the fluorescence intensity at emission maxima;
   - Assessment of detection of mutated DNA extracted from cancer cell lines. (Paper IV).

1.3. Study design

Preliminary data by others and us and existing literature rationalize the selected methodology. The applied methods (represented in Figure 1.1) are suitable for extensive studies of modified oligonucleotides in diagnostic and therapeutic context.

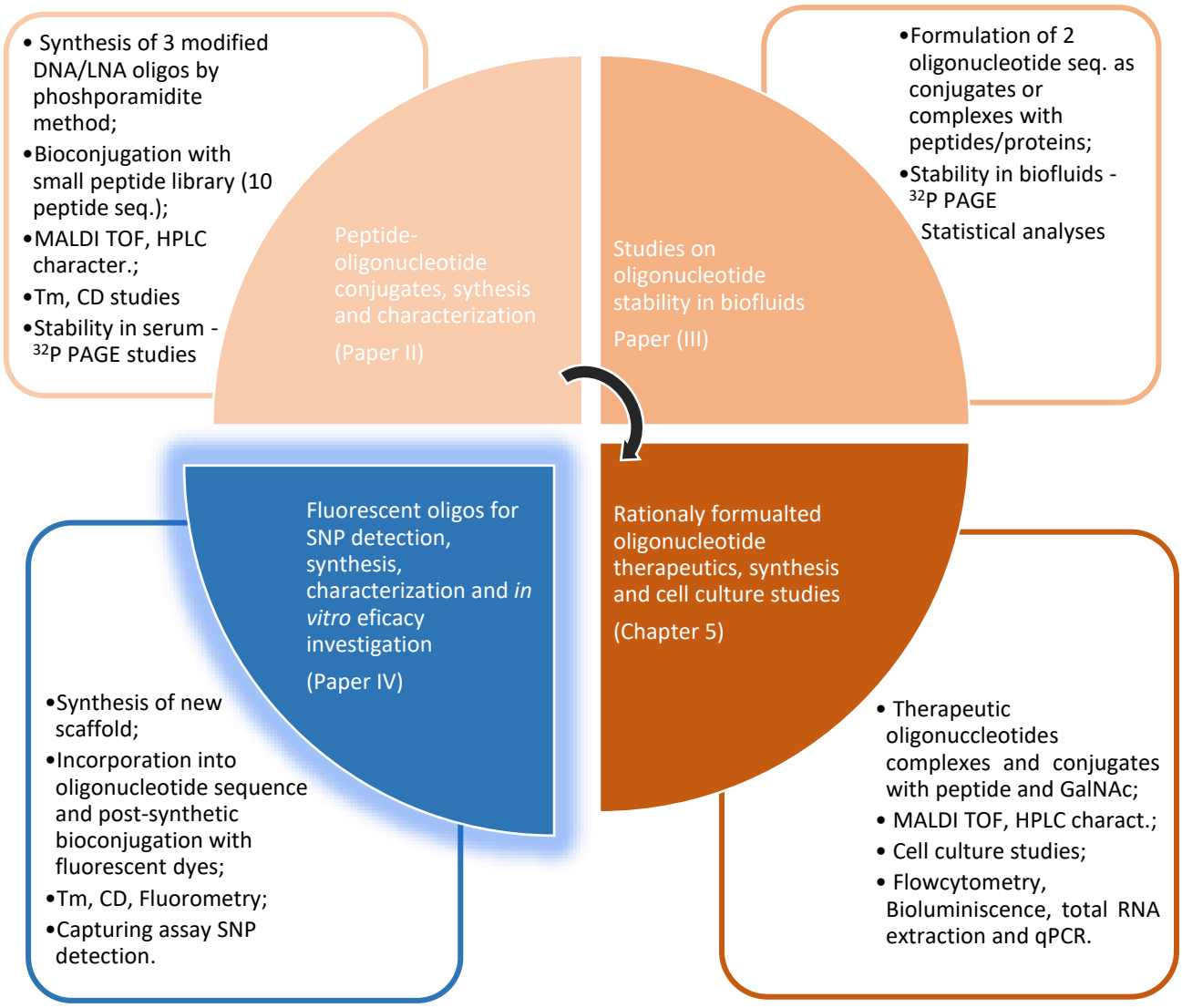


Figure 1.1. Study design summary. Schematic representation of the four different projects included into the thesis and the methods used.

Chapter 2

Intro

2.1 Nucleic acid discovery, structure and biological function

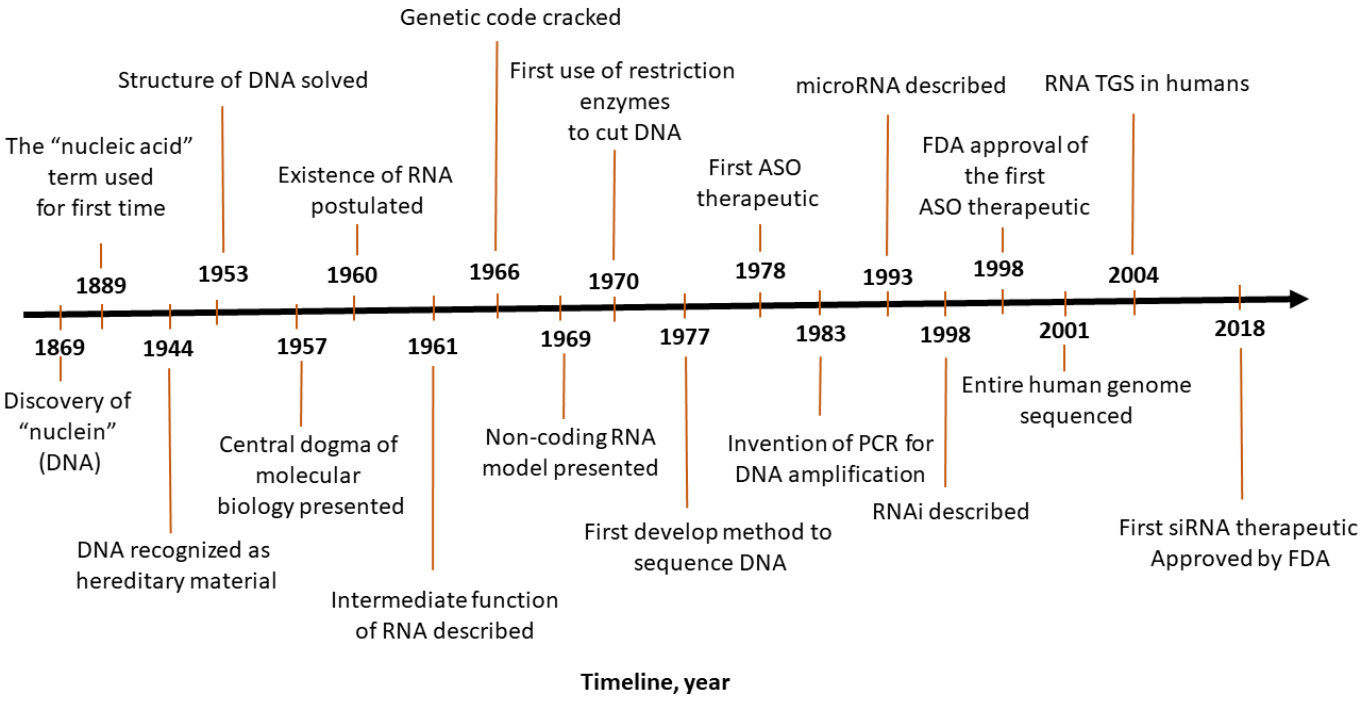


Figure 2.1. Timeline of DNA[1–4] and RNA[5–10] discovery.

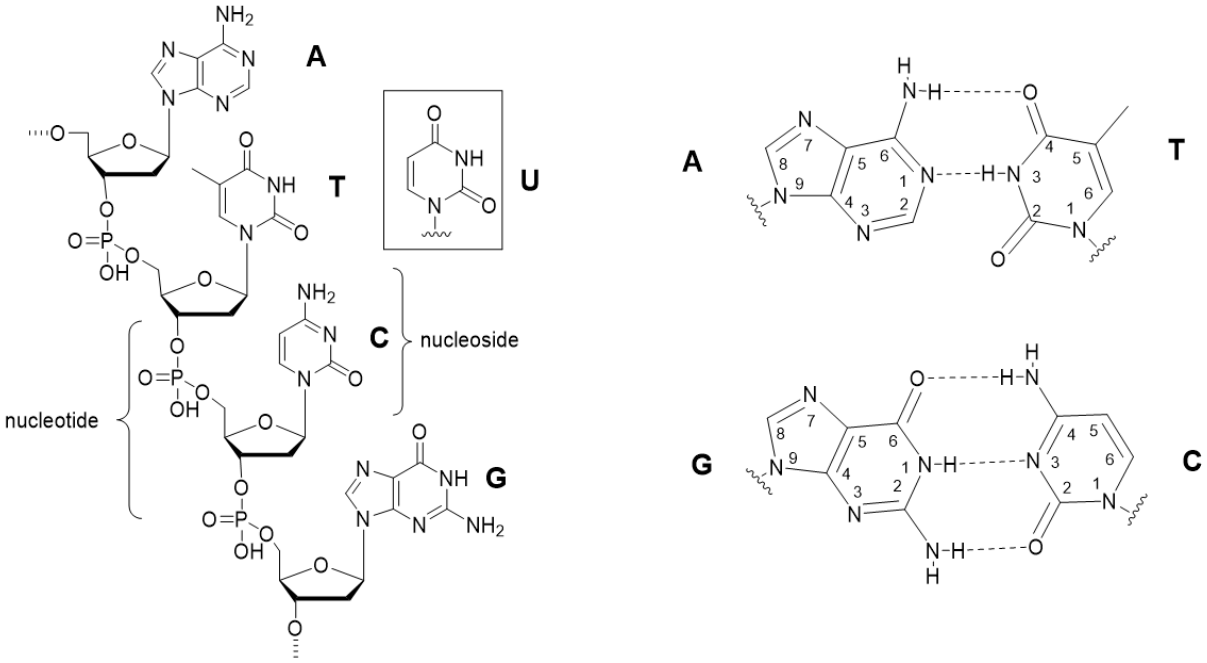
Deoxyribonucleic acid (DNA) was first isolated by Friedrich Miescher 150 years ago. Since then, it is the foundation of modern bioscience. The so-called "nuclein" was originally extracted from the nuclei of human leucocytes and clearly distinguished as a separate entity from proteins with a high phosphorus content.[1,3] Subsequently, in 1889, due to its behavior as an acid in chemical reactions, Richard Altmann named it nucleic acid.[1] Although the nucleic acids were identified as being a part of the chromosomes,[11] the importance of the DNA was recognized 77 years after this discovery. It was when Oswald Avery revealed that DNA consisting of only four subunits – thymine (T), cytosine (C), adenine (A) and guanine (G) is hereditary material.[12] By following the physicochemical studies of the Chargaff group[13,14] and the x-ray crystallographic studies by Franklin and Gosling,[15] Watson and Crick in 1953 solved the structure of DNA, the double-helix model.[16,17] Finally, in 1958, Crick presented the central principles of molecular biology, giving the mechanism of protein synthesis.[4] According to it, first the DNA is transcribed to ribonucleic

acid (RNA) in the nuclei and then translated to protein in the cytoplasm.[18] This was followed by the resolution of the genetic code indicating that the DNA codon consists of three nucleotides or letters giving one amino acid.[19] Since then, it has been possible to isolate and cleave DNA on specific locations, amplify, transfer genes from one species to another and sequence the DNA.[20] At the beginning of the 21th century, the focus in the field has been on sequencing the entire human genome.[21–23] **Figure 2.1** represents the timeline of DNA and RNA discovery To date, the knowledge of gene regulation, protein synthesis and nucleic acid-protein interactions is growing rapidly. Although complex, it is auspicious to follow the potential of gene manipulation in the nucleic acid research.

DNA and RNA are natural nucleic acids polymers made up of nucleotide monomers. Their chemical structure consists of aromatic heterocyclic nucleobase called purines (A and G) or pyrimidines (T/uracil (U) and C), sugar moiety (2'-deoxy-β-D-ribofuranose in DNA or β-D-ribofuranose in RNA) and a phosphate group. The primary structure of DNA and RNA is analogous. Two successive sugars on the 3' and 5' carbons are connected via a phosphodiester groups while the nucleobases (Nitrogen 1 for pyrimidines and Nitrogen 9 for purines) are attached to the 1' carbon of the sugar moiety. When, the G and A form Watson-Crick base pairs with C and T/U respectively, DNA:DNA or DNA/RNA duplex is formed giving the secondary structure (**Figure 2.2**). The double helix is formed by hybridization using hydrogen bonds, π-π stacking and hydrophobic interactions by the complementary nucleotide strands oriented in an antiparallel fashion. The hydrophobic nucleobases are located in the core pared up and the hydrophilic phosphates are located on the exterior of the duplex. By minimizing charge and steric repulsion and by optimizing the above-mentioned interactions, the duplexes adapt helices as tertiary structure. The most important biological active helix structures reported are helix type -A, -B and -Z.[24] The B-type duplex is right handed with wide major groove and narrow minor groove and is preferred by DNA under physiological conditions. When B-DNA is dehydrated, it adopts an A-DNA helix that is wider and has a shorter pitch. Z-DNA is promoted by alternating purine-pyrimidine sequences or by high salt concentration. The duplex is left handed with a zig-zag backbone, large pitch and short diameter of the cross sections.[25] RNA usually adopts an A-helix under physiological conditions with deeper major grove and wider minor grove. Hybrid duplexes (DNA:RNA) are also revealed in many biological processes. They adopt mixed A/B conformation, and they are structurally more similar to A-type helix.[26] Directed by the structure, the primary function of DNA is to store and pass on genetic information through RNA to proteins

– the central dogma of molecular biology. Nevertheless, today it is known that big part of the human genome contains DNA that is noncoding i.e. does not form functional protein.

I)



II)

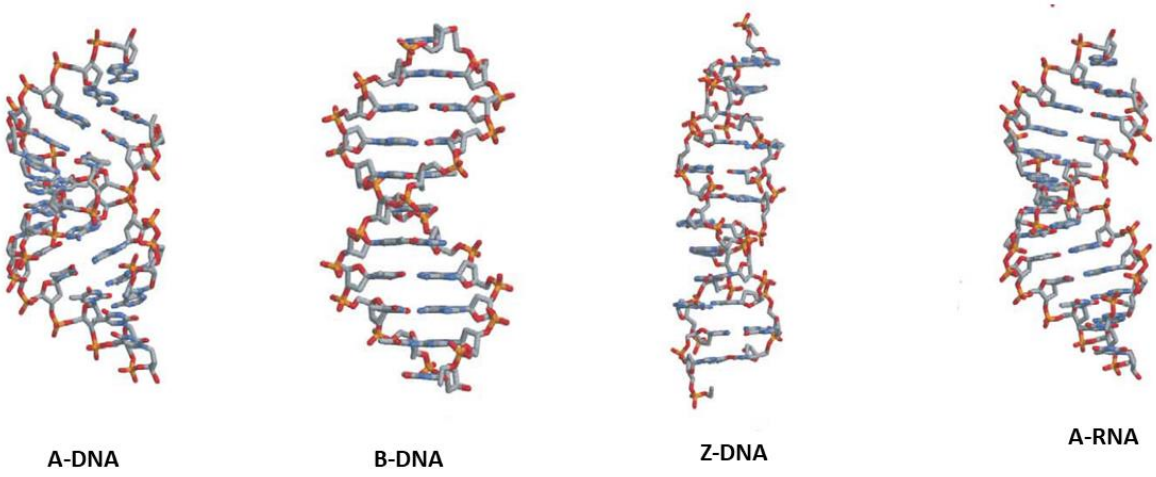


Figure 2.2. Primary, secondary (**I**) and helical-side view (**II**)[24] structures of DNA and RNA

Moreover, literature recognize that DNA and RNA influence the conformational and the dynamic properties of the nucleosomes and play an important role in many biological processes such as

replication, recombination and gene regulation.[27] In addition, DNA and RNA are part of the mechanisms for epigenetic regulation such as: DNA methylation, post-translational modification of histones, chromatin remodeling, microRNA (miRNA) and long non-coding RNA (ln RNA).[28]

2.2 Synthetic nucleic acids – chemical synthesis development and applications

Following the revolutionary article by Watson and Crick in 1953, chemists became interested in preparing new and synthetic significant bio-molecules. The main step in the chemical synthesis of oligonucleotides is generation of the internucleotide linkage using nucleoside phosphate derivatives as building blocks (**Figure 2.3**). The first dinucleotide synthesis was reported by Michelson and Todd, in 1955, containing 3'-5' phosphate link using nucleoside 3' phosphoryl chloride.[29] Three years later, since the intermediate phosphoryl chloride was susceptible to hydrolysis, Khorana introduced new concept – the phosphodiester method, using stable 3' phosphorylated nucleoside capable of coupling with dicyclohexyl carbodiimide (DCC) after activation.[30] Khorana also contributed greatly to the introduction of nucleoside protecting groups that are key to the development of sequential nucleoside coupling. He introduced the protecting group for the 5' hydroxyl nucleoside - dimethoxytrityl (DMT), and the protecting groups for the exocyclic amines of the nucleobases - isobutyryl for guanosine, benzoyl for adenosine and cytidine, and acetyl for cytidine.[31] Furthermore, in 1969 Letsinger *et. al*. introduced the phosphotriester method which improved Khorana's phosphodiester system by avoiding branching at the internucleosidic phosphate. They used β-cyanoethyl phosphate intermediates to make comparably simple and reproducible chemistry.[32] However, considering the slow reaction rate of the previous method, Letsinger and coworkers presented a new phosphite-triester methodology in 1970. Their results were based on the use of more reactive phosphorus in the P(III) oxidation state in the nucleoside phosphomonochloridite intermediate which was then efficiently oxidized to phosphotriester with aqueous iodine.[33] The prevailing contribution of Letsinger was the introduction of the first solid support polymer for oligonucleotide synthesis [34,35]. His student Caruthers developed it further by expanding the chromatography grade silica and the controlled pore glass (CPG) as supports.[36] The work of Caruthers and coworkers is regarded as crucial for the introduction of the phosphoramidite method. They succeeded by exchanging the unstable and highly reactive chloride with a

relatively controllable amine.[37] Finally, Koster *et al*. in 1983, reported successful formation on internucleotidic bonds using β-cyanoethyl 3'O-(N,N-diisopropyl) phosphoramidite.[38]

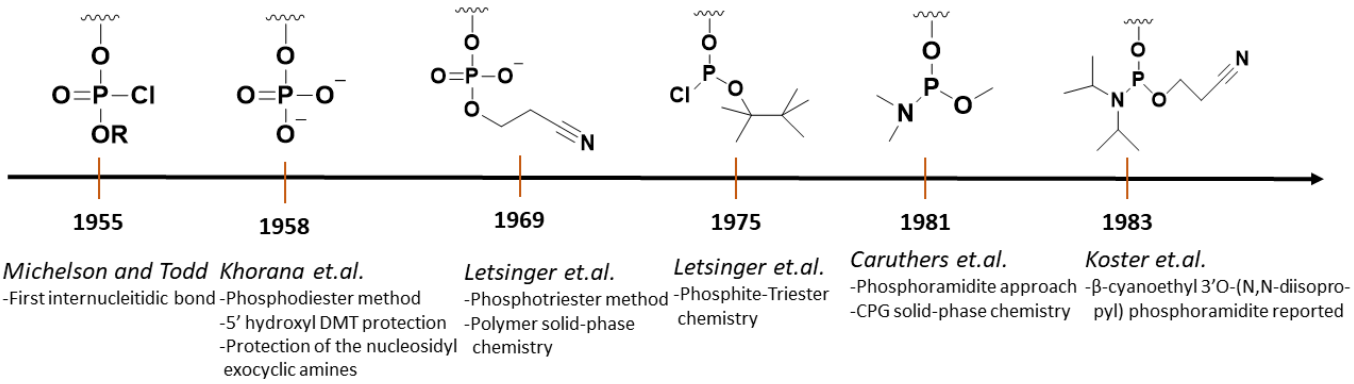


Figure 2.3. Development of phosphorous moieties and formation of inter-nucleotide phospodiester linkage

Currently the preferred method for oligonucleotide synthesis is the phosphoramidite method using β-cyanoethyl as protecting group because of its stability and easy removal. The solid-phase oligonucleotide synthesis runs as a cycle that consists of four subsequent reactions. Elongation goes from 3' to 5' direction and uses amino functionalized CPG or macroporous polystyrene supports that are rigid and low-swellable.[36,39] The amino group anchors the base-labile linker, which is also attached to the first monomer (3'-terminal). The first step in the cycle is to detritylate to remove the 5' DMT protection group of the first monomer with trichloroacetic or dichloroacetic acid in dichloromethane. Since the DMT cation is orange and has a high absorption at 495 nm, spectrometry is used to determine the efficiency of the coupling on each nucleotide.[40] Furthermore, the nucleoside phosphoramidite is activated by 1*H*-tetrazole. Consequently, it couples to the free 5'OH group of the solid–support bound first monomer and forms a phosphite triester linkage.[41] Next, the unreacted 5'OH groups are capped by acetylation with a mixture of acetic anhydride and 1-methylimidazole in tetrahydrofuran. The elongation cycle of a single nucleotide ends with oxidation of the phosphitetriester intermediate to phosphotriester with aqueous iodine in presence of pyridine.

The development of efficient methods for oligonucleotide synthesis and the recognition of the unique properties of the synthetic oligonucleotides lead to exploration for wide use in many fields. Taking into account DNA's and RNA's defined size, periodicity and programmable

features, in the field of nanotechnology they are explored as nanomachines with application in biomedicine and biotechnology.[42,43] The action and the control of these molecular machines can be achieved through the use of strand hybridization, enzymatic reactions or environmentally provoked conformational transitions.[44] Neumann and colleagues, for example, reported building tweezer like nanomachine, made of three DNA strands and fueled by another DNA strand.[45] They highlighted the DNA fueled controlled motion, indicating numerous possibilities of the use of such systems, including mediated organic synthesis[46] and cargo transport.[47]

In addition, synthetic oligonucleotides are introduced as biosensors using the controlled surface chemistry and the charge distribution. This has a significant impact on clinical diagnosis, environmental monitoring and food control.[48,49] In general, they are classified as DNA based, aptamer based and DNAzyme based biosensors. Whereas, the DNA based biosensors detection works by target hybridization and specificity;[50] aptamer based biosensors, in addition to hybridization of biosensors with DNA and RNA[51], bind and detect specific receptor molecules with high affinity. The DNAzyme based biosensors detect and cleave substrate strand in the presence of metal ions which act as cofactors.[52] In addition, many advanced methods have been introduced to transduce the resulting signal including fluorescence,[51] electrochemistry,[53] chemiluminescence[54] and colorimetry.[55] Moreover, forensic genetics and "DNA fingerprinting" were developed with the development of polymerase chain reaction (PCR)[56] and the amplification of the short tandem repeat (STR) sequences. The use of forensic analyses enabled the possibility of identification and quantification of human DNA from biological samples.[57,58]

Very recently, it was reported that accurate age and human visible characteristics can be predicted using epigenetic markers and methylation-sensitive analysis.[58] The new methods and the next-generation sequencing led to the development of new scientific disciplines in the sphere like population genetics and phylogenetics.[59] The newly developed scientific disciplines provide development of significant analytical methods to form diagram hypothesis - phylogenetic trees that are crucial for understanding the evolution of organisms, biodiversity and ecology.[60,61]

Science explore the synthetic oligonucleotides as advanced therapeutics in medicine as well. In recent decades, significant efforts have been made to develop efficient short synthetic oligonucleotide therapeutics. It was postulated that these drugs would act in the core of the disease on the genetic level by simple hybridization. Efficient targeting of any desired gene in a particular cell type can be achieved by changing the sequence of the therapeutic and by improving of the delivery vehicle.

As predicted, there are now various types of promising oligonucleotide therapeutics in development. Nevertheless, only a few of them are in clinical use. Despite of the challenges, it is clear that synthetic oligonucleotide therapeutics are important part in the development of the personalized medicine, and in future, they have high potential to become one of the major drug platforms.

2.3 Synthetic oligonucleotides as therapeutics; mechanisms of action

Many studies report oligonucleotide therapeutics as appropriate tools for managing a wide range of non - targetable diseases. Using short synthetic oligonucleotides, specific non-functional genes can be targeted, silenced and manipulated by direct interference with gene function on a cellular level. Compared to available small molecule drugs and other large biopharmaceutical products, oligonucleotide therapeutics, which mainly act by complimentary base pairing, have a much simpler design, offer more controllable action and have fewer side effects. Today, they are recognized as candidates with the highest potential to meet the medical needs for treating cancer,[62–64] neurological diseases,[65,66] cardiovascular diseases,[67,68] autoimmune diseases,[69,70] viral[71–73] and even bacterial infections.[74]

The concept of using short synthetic oligonucleotides to control gene expression dates back to 1970, when synthetic DNA oligonucleotides were proposed to replace mutated DNA.[75] However, due to their size, their negatively charged backbone and their intracellular action (**Figure 2.4**), the clinical success of oligonucleotide therapeutics in the upcoming decades was negligible. The guiding principle for small molecule drugs[76] failed to be valid and demanded new strategies of drug development.[77]

Property	**Small molecule**	**Oligonucleotide**	**Antibody**
Size	1 nm	6 nm	12 nm
Molecular weight	<500 g/mol	7000-14000 g/mol	>100000 g/mol
Site of action	Extra- and intra-cellular	Extra- and intra-cellular	Extracellular
Cell permeability	Good	Poor	Poor
Administration route	Primarily oral	IV/SC/IT/IVT	IV/SC

*IV-intravenous; SC-subcutaneous; IT-intrathecal; IVT-intravitreal

Figure 2.4. Properties of small molecule, oligonucleotide and antibody drugs[78] Figure adopted from the respective reference.

The FDA approval of the very first oligonucleotide therapeutic – ASO, Fomivirsen (Vitravene) took 20 years of extensive research and numerous clinical trials. It was first developed by Isis (Ionis) Pharmaceuticals as a 21-mer phosphorotioate ASO used to treat cytomegalovirus retinitis by intravitreal administration.[79] Macugen (Pegaptanib), a 27 nucleotide long aptamer oligonucleotide approved in 2004, was the next approved oligonucleotide drug in the line. It consists of a phosphorothioate 3'-3'deoxythymidine cap, 2'-O-methylated ribose sugars of the purines and 2'-fluorinated ribose sugars of the pyrimidines. It targets the vascular endothelial growth factor (VEGF165). From Fomivirsen's initial approval, it took another 15 years for the scientists to develop the second ASO. Kynamro (Mipomersen) indicated for familial hypercholesterolemia was approved by the FDA in 2013. It is a 20-mer phosphorotioate ASO that contains 2'-methoxyethoxy modifications on positions 1-5 and 15-20 in the sequence. Mipomersen targets the coding region of apolipoprotein B mRNA that is further cleaved by RNase H1.

In the recent years, the research circle was running faster and three new ASOs were introduced to the market in 2016. The first one, Defitelio (Defibrotide), was developed by Jazz Pharmaceuticals. It is designed to target severe hepatic veno-occlusive disease caused by chemotherapy. Its structure is a mixture of single and double stranded not modified oligonucleotides with average length of 50 nucleotides. It has a non-specific mechanism of action that most likely uses charge-charge interactions with proteins. The other two ASOs are Eteplirsen (Exondys 51)[80] and Nusinersen (Spinraza).[80] Both of them are splice switching oligonucleotides and use exon skipping and exon inclusion gene modulation, respectively. Eteplirsen is a 30-mer, phosphomorpholidate oligonucleotide. A superlative candidate that correct irregular splicing effects by skipping exon 51 in duchenne muscular dystrophy (DMD). Whereas, Nusinersen is an 18-mer modified phosphorotioate ASO with 2'-O-methoxyethoxy and methyl cytidines on position 5. It makes inclusion of exon 7 in the *survival of motor neuron 1 and 2* gene by blocking the intron 7 splice site in spinal muscular atrophy (SMA). After 15 years investigation of RNA interference, the first siRNA oligonucleotide therapeutic was approved in the summer of 2018. Alnylam Pharmaceuticals developed the Patisiran (Onpattro), and formulated it as a lipid nanoparticle that targets the liver. Its usage is important for treatment of

hereditary transthyretin-mediated amyloidosis which is characterized by deposition of amyloid in the peripheral nerves and the heart.[81]

Following literature, various types of therapeutic oligonucleotides have been developed. They work by diverse mechanisms and target different class of nucleic acids inside cells - including mRNA, pre-mRNA, miRNA, lnRNA, telomerase RNA and protein. All of these targets are structurally and functionally different, and represent different challenges in the process of drug discovery and development.[78] The following section will describe a few cutting-edge therapeutic oligonucleotide platforms and their mechanism of action.

2.3.1 Antisense oligonucleotides (ASOs)

Paterson *et al*. in 1977 introduced ASOs showing the translational blockage of mRNA by DNA sequence.[82] Following this discovery, Zamecnik and Stephenson reported the inhibition of Rous sarcoma virus replication.[83] ASOs are single stranded oligonucleotides, between 15-25 nucleotides in length that can bind to the target RNA sequence by Watson-Crick base pairing and control the gene expression.[84]

Depending on the chemical properties and the intended target, ASOs can modulate gene expression mainly by two different mechanisms. These mechanisms include either RNase H activation or steric blockage (**Figure 2.5**). When applied in gain-of function disorders, ASOs act by silencing the mutant gene expression. In this matter, "knock down" of the mRNA can be triggered either by steric block or by RNase H1 degradation of the mRNA. The steric block near the starting codon[85] or to the 5' cap region[86] prevents association with the ribosome translation machinery. Whereas the not specific endonuclease RNase H1 recognize the DNA-RNA hybrid and cleave the mRNA phosphodiester backbone. On the other hand, the modulation of alternative splicing by targeting the splice sites – exons or introns of mRNA and prevention of the recognition by the splicing factors can either restore the function of a mutant protein or can increase the expression of a protein variant in loss-of function disorders.[78,80]

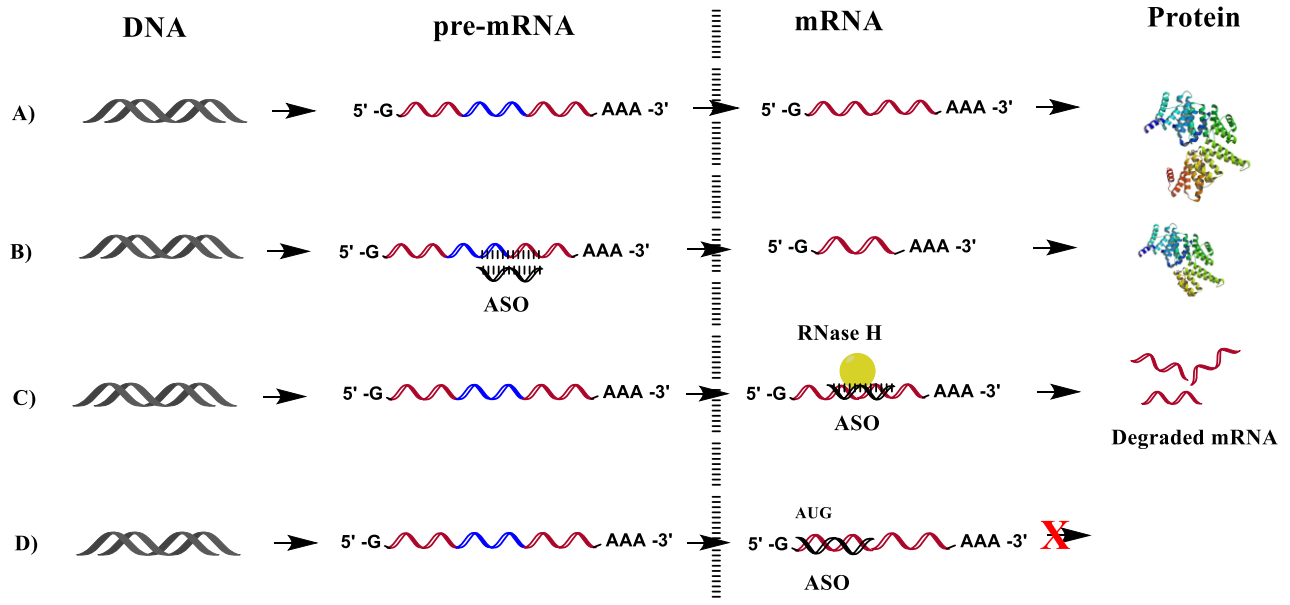


Figure 2.5. ASOs mechanism of function. A) Natural process; DNA transcription into pre-mRNA that is consequently sliced to mature mRNA in the nucleus. mRNA enters the cytoplasm and it is translated into protein by ribosomes. B) Splice exclusion by splice switching ASO binding to the pre-mRNA to skip dysfunctional exon resulting in further translation in condensed but functional protein. C) RNA "knock down" by ASO binding to mRNA and activating RNase H1 to cut the mRNA. D) Prevention of mRNA translation by steric block ASO.

2.3.2 siRNAs

The idea of RNA interference dates back to 1998 when Mello and colleagues reported that double stranded RNA, when introduced into animal can interfere with the utility of the mRNA.[9] It takes advantage of the natural defense mechanism of the cells to destroy exogenous or viral RNA using short non coding RNA duplexes, miRNA or small interfering RNA (siRNA) as templates.[87] Recognizing its potential, in 2001 Elbashir and Tuschl showed that synthetic, 21 nucleotide long, double stranded RNA is able to specifically suppress gene expression in various mammalian cell lines.[88] This finding led to the development of siRNA gene therapeutics with comparable effects to ASOs.

After entering the cell, siRNAs trigger the endogenous RNA interference machinery. It is represented by RNA-induced silencing complex (RISC) that contain members of the argonaute protein family. One of the two strands of the siRNA, usually the guide strand binds to the argonaute (AGO2) protein to form the ribonucleoprotein complex. It furthermore serves as a guide to its complementary mRNA. AGO2 protein, in the right conformation, binds the guide RNA

strand to recognize the target mRNA and then cleave the fully complementary target or steric block the target with some mismatches (**Figure 2.6**).[78]

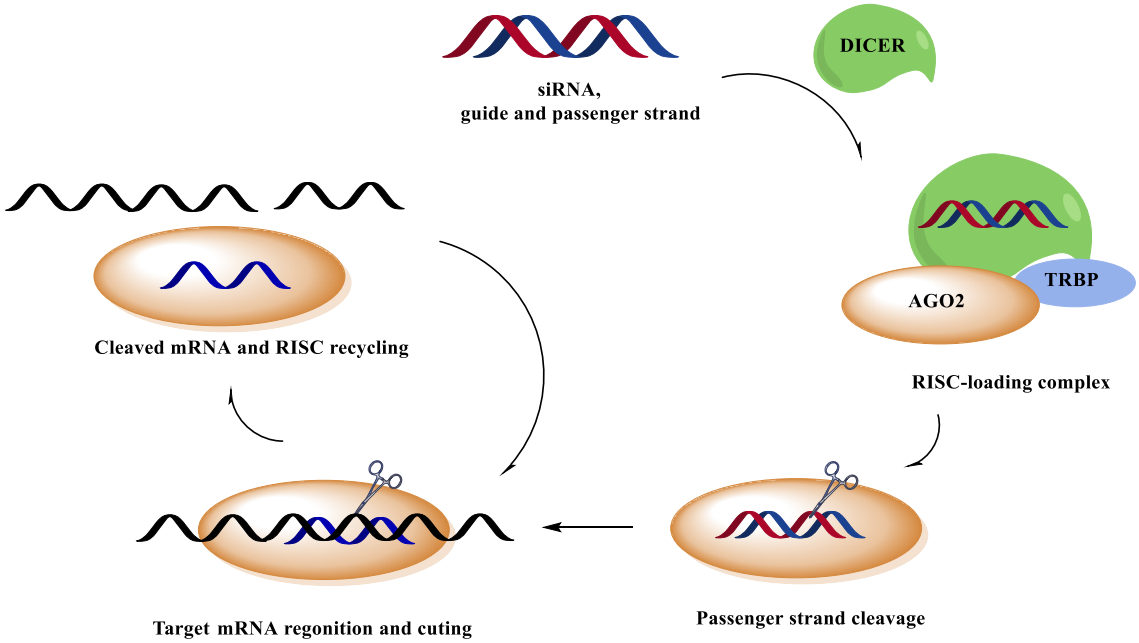


Figure 2.6. siRNA mechanism of action; First, siRNA is loaded onto argonaute (AGO2) by DICER with the help of RNA binding protein – TRBP. Then AGO2 selects the guide strand ejecting the passenger strand. Moreover, the guide strand guide AGO2 to cleave the complementary mRNA target. At the end, RISC is recycled using the same guide strand.

2.3.3 CRISPR Cas

The clustered regulatory interspaced short palindromic repeat associated 9 (CRISPR Cas9) system was discovered as an adaptive defense system in *Streptococcus pyogenes* protecting the microorganism from foreign DNA invasion.[89] After several years of extensive research on the working mechanism of the system, in 2012 Jinek *at. al.* suggested CRISPR Cas9 to be a groundbreaking tool for editing of genomic DNA. It was recognized that Cas9 is DNA endonuclease with two nuclease domains, HNH and RuvC-like, and that it is guided by two RNA sequences which interact with each other and form a duplex - tracrRNA and crRNA (**Figure 2.7**).[90–92] The DNA target is recognized by the crRNA in presence of a short PAM motif sequence.

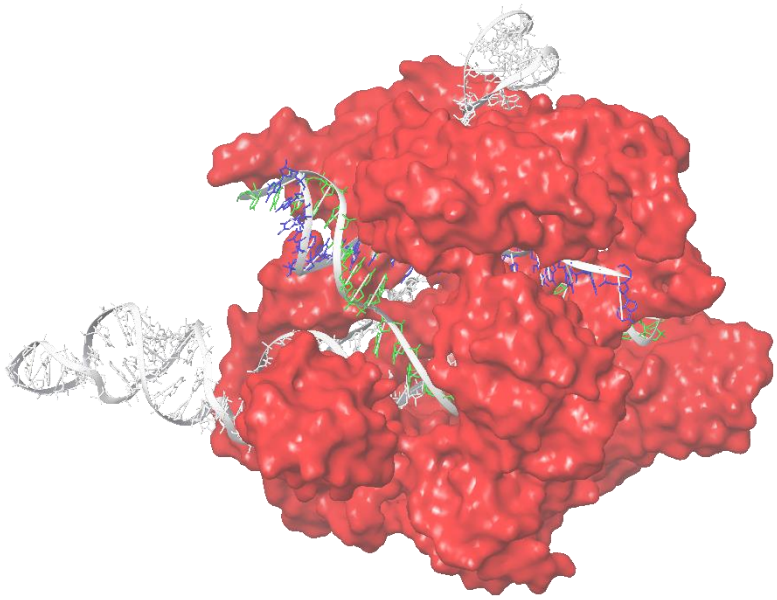


Figure 2.7. Crystal structure of CRISPR Cas9 (red) from *Streptococcus pyogenes* associated with sgRNA (green and purple) and target double stranded DNA (grey).PDB ID: 5F9R[93] Picture modified in Maestro.

Furthermore, the tracrRNA and the crRNA were combined in single guide RNA (sgRNA) which is capable of guiding the CRISPR Cas9 using Watson-Crick base pairing to any desired target DNA sequence adjusted to the protospacer adjacent motif (PAM).

Recognizing that CRISPR Cas9 has immense potential, many studies focuses on improving the Cas9 and sgRNA features.[94,95] CRISPR Cas9 is involved in functional studies concerning the molecular basis for diseases and as a therapeutic by regulation of gene expression or pathogen gene disruption.

2.4 Chemical modifications of therapeutic oligonucleotides

Although oligonucleotide therapeutics have shown auspicious and effective results in cell culture assays and animal studies, they still have a long road towards reaching the clinic. ASOs, siRNA and the other nucleic acid therapeutic platforms face big challenges preventing them from actual application. The first obstacle is the potential off target effects and immune stimulation that may lead to harmful side effects. Secondly, poor *in vivo* stability is precluding the drug efficacy by degradation and preventing systematic delivery. Finally, since oligonucleotide

therapeutics act on an intracellular level i.e. cytoplasm or nuclei; their biggest weakness is the ineffective distribution at the action site.

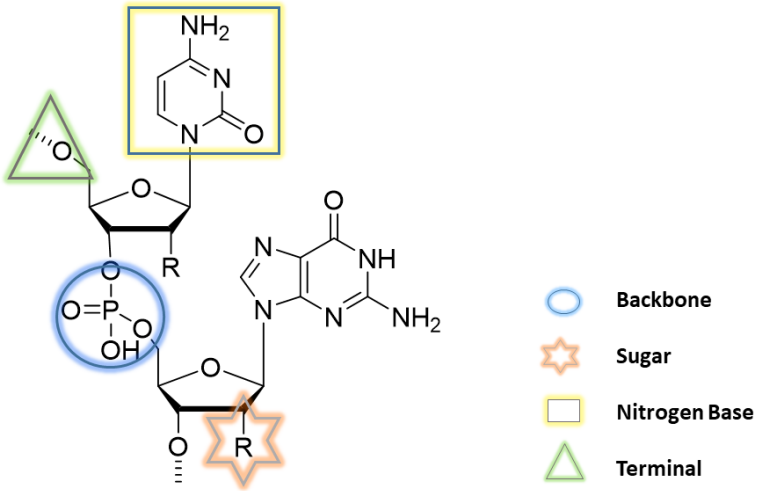


Figure 2.8. Dinucleotide representing the most common modification sites in oligonucleotide therapeutic

Although many of these obstacles are solved by a development of chemical modifications to the natural DNA and RNA, there are still many challenges that need to be resolved. Following the oligonucleotide structure, the chemical modifications were developed to deal with every challenge, and aimed to solve multiple problems simultaneously (**Figure 2.8 and Table 2.1**).[96] For instance, the inter-nucleotide linkage modifications were developed to increase nuclease resistance to endo- and exonucleases and to improve the intracellular delivery of the oligonucleotide therapeutics. The chemical modification of the sugar led to better binding affinity, which led furthermore to less off target effects. In addition to that, these chemical modifications also significantly improved the enzymatic stability. Finally, the nucleobase modifications that affect target binding and immune stimulation were also introduced, and present interesting strategy for improving oligonucleotide properties.

The following pages aim to outline the most commonly employed chemical modifications showing promising results in later clinical trials.

2.4.1 Modifications that increase the stability towards nuclease degradation

Inter-nucleotide linkage modification is alteration that protects the oligonucleotide from degradation caused by endo- and exonucleases. The most frequently used strategy in

oligonucleotide therapeutics development and biotechnology is the phosphorothioate (PS) backbone modification, where the non-bridging phosphate oxygen is replaced by sulfur.[97] It increases the half-live of DNA and RNA oligonucleotide sequences by improving the metabolic stability and by binding with serum proteins like heparin and albumin. PS oligonucleotides are mainly applicable to ASOs and have less negatively charged backbone that improve the cell internalization as well.[98] The reduced target binding ability of the PS oligonucleotides is considered an unfavorable asset. However, as oligonucleotide therapeutics, they are still able to induce RNase H mediated cleavage of mRNA. Related to that, the chemically synthetized PS oligonucleotides are not isomerically pure. They exist as a diastereomeric mixture with two stereochemical configurations, *S*p and *R*p. It is reported that stereochemistry influence the efficacy and the pharmacokinetics decreasing its activity.[98,99] Reliable synthesis and purification of oligonucleotides with defined stereochemistry of the PS linkage is still major challenge that invites further research and development.[100]

Table 2.1. Properties overview of the more frequently used oligonucleotide therapeutics modifications. – The effects of the particular assets are described as: high, moderate, fair or low.

Modification	Off target effects	Enzymatic resistance	Ability for cell penetration	Mechanism of action	Main Representatives
Backbone					
PS	High	High	Fair	RNase H +	ASOs, siRNA
PMO	Moderate	High	Fair	Steric block	ASOs,
PNA	Low	High	Fair	Steric block	ASO, siRNA
Sugar					
2'-O-Me	Low	Fair	Low	RISC /Steric block	siRNA, ASO, Aptamer
2'F	Low	Low	Low	RISC /Steric block	siRNA, ASO, Aptamer
2'-O-MOE	Low	High	Low	RNaseH /Steric block	ASO, siRNA
LNA	Very low	Fair	Low	Steric block	ASO, miRNA, siRNA
UNA	High	Fair	Low		siRNA, ASO

*PS-phosphorothioate; PMO-phosphorodiamidate morpholino; PNA-peptide nucleic acids; 2'-O-Me-2'-O-methyl RNA; 2'F – 2 fluoro; ; 2'-MOE-2'-O-(-2-methoxyethyl) RNA; LNA-locked nucleic acid; UNA-unlocked nucleic acid

Phosphorodiamidate morpholino is the second most important modification of the oligonucleotide phosphodiester backbone.[101] Compared to natural DNA and RNA, it consists of methylenemorpholine rings, which are inter connected with phosphorodiamidate groups. Following the chemical structure, the PMOs show resistance to nuclease cleavage. The PMOs

work as steric blockers and are among the most promising candidates for treatment of diseases where exon skipping is needed.[102] Nevertheless, the poor cellular uptake is the biggest obstacle that prevents the PMOs from clinical application.[103]

Peptide nucleic acids (PNA) represent a thrilling class of oligonucleotide therapeutics also. They share structural properties from both peptides and oligonucleotides, which are two key biomolecules. In the PNA structure, the N-(2-aminoethyl) glycine connects to the bases via a methylene carbonyl linker and replaces both the phosphodiester backbone and the sugar moiety. That is the main reason why they exhibit nuclease resistance.[104] Compared to its DNA variant, PNA can bind its DNA target stronger. Due to its minimized electrostatic repulsion, target recognition happens with great affinity via base pairing between the unaffected nucleobases. As such, PNA silences a particular gene by a steric block.[104,105] In addition to that, PNA is capable of forming triplex with genomic DNA. This ability is explored in gene editing which activates the endogenous DNA repair factors to make recombination of donor single stranded DNA and to fix a mutation.[106]

2.4.2 Modifications that improve target binding

In order to imitate the A-form duplex, nucleobase modifications for adjustments of the sugar conformation within the oligonucleotide sequence modulate its target binding specificity.

2'-O-methyl (2'-O-Me) modification is a RNA mimetic where 2'-O-Me replace the 2'OH. It is the most common variation for development of siRNA therapeutics, mainly because it increases the target binding affinity.[107,108] In addition, it is involved in trials developing ASOs and aptamer drugs.[109] 2'F RNA is another important imitation of RNA that increases the strength of target binding by approximately 3 ^{o}C per insert in oligo melting experiments.[110,111] Similar to 2'-O-Me, it is the main constituent in siRNA therapeutics, but is also used in the ASO and aptamer drug development platforms. The 2'-O-(-2-methoxyethyl) RNA (2'-O-MOE) is a methoxyethyl modification of the 2'OH of RNA. It increases the melting temperature and shows improved nuclease stability. When used in ASOs it can induce RNase H cleavage of mRNA.[111]

To date, the locked nucleic acid (LNA) is reported to be the leading modification candidate which can improve the target binding specificity.[112,113] It is a conformational restricted bicyclic RNA analog where the 2' OH and the C4' are linked by a methylene bridge which stabilize the duplex for 5-6 ^{o}C per insert. LNA is used in the development in almost all types' of oligonucleotide drugs. Its discovery represent the beginning of shorter and more efficient oligonucleotide sequences as therapeutics.[114] In addition, scientists developed unlocked nucleic acid (UNA) modification. UNA

lacks the C2'-C3' bond of the ribose sugar and when used it gives flexibility to the oligonucleotide. In general, UNA molecule, when incorporated into oligonucleotide lowers the affinity toward the target. Because of these characteristic, scientists combine UNA with other modifications aiming to improve the RNase H compatibility of ASOs.[115] Moreover, literature suggests that when UNA is implemented in siRNA oligonucleotides, it is capable of reducing the off target effects to the seed region, reducing siRNA toxicity in the HeLa cell line and protect siRNA from cleavage in serum.[116,117]

2.4.3 Prominent modifications headed for targeted intracellular delivery

Complexation and covalent linkage between oligonucleotides and other molecules and biomolecules open a new perspective for the development of potent, stable, and safe nucleic acid therapeutics. In particular, formulation approaches are leading to, and have already shown promising results by increasing the ability of oligonucleotide therapeutics to convey cellular uptake.

Carbohydrate – oligonucleotide conjugates are a very promising approach for efficient targeted intracellular delivery, specifically for siRNA. Scientists pay special attention to them, especially with the discovery of the asialoglycoprotein receptor expression in hepatocytes. It process *N*-acetylgalactosamine (GalNAc) residues by receptor-mediated endocytosis. To date, siRNAs covalently attached to GalNAc are extensively researched, and their efficient receptor mediated delivery to hepatocytes is confirmed.[118–120] Notably, the candidates in clinical trials, this method is a very elegant way to treat many liver conditions by oligonucleotide therapeutics. Nevertheless, they also face challenges such us suboptimal pharmacokinetics, pharmacodynamics and safety. In addition to the GalNAc conjugation, there is a need for further improvement of the siRNA oligonucleotide design. Many scientific reports highly recommend the development of new strategies similar to 2'sugar modifications and implementation on UNA for better *in vivo* performance.[119,121]

The conjugation of oligonucleotide therapeutics to lipid molecules represents another important strategy to improve the cellular uptake, to prolong the half-life in plasma and to enhance the *in vivo* potency of the ASOs and RNAi drugs.[122–124] Contrary, to cationic polymers and lipids, literature propose the more biocompatible natural lipids like cholesterol and fatty acids to be more promising candidates for bioconjugation.[125,126] A study, where a small library of 3' and 5' lipid modified ASOs were synthetized proved that the ASOs containing 5'-lipid modifications

internalize within the cell and silence the target gene. The micropinocytosis and the clathrin-mediated endocytosis were proposed to be the mechanisms for intracellular delivery.[125] In another study, when compared with its naked variant, the siRNA conjugated with cholesterol showed evidently improved pharmacological properties with fifteen times increased half-life *in vivo*.[127] Furthermore, α-tocopherol conjugated to heteroduplex oligonucleotide – gapmer DNA/complementary RNA displayed improved potency, higher tolerability and better safety compared to the parent gapmer DNA.[128]

Literature report aptamers which specifically recognize the cell surface receptor as promising tools for targeted delivery of nanoparticles, antibodies and therapeutic oligonucleotides.[129–131] They are 20-80 base pairs long DNA or RNA single stranded oligonucleotides. Upon secondary and tertiary structure formation, aptamers act as ligands and bind to a specific target molecule or receptor with high affinity. A study, report an aptamer with specific ability to deliver splice-switching ASOs to cell nuclei. This ASO-aptamer chimera is able to splice the pre-mRNA. However, the authors fail to explain the mechanism of intracellular and nuclear internalization.[132] Furthermore, literature report various aptamer-siRNA conjugates and chimeras that target cancer and HIV cells.[133,134] They are seen as powerful delivery vehicles, which are capable of reducing the treatment doses of the therapeutic oligonucleotides and the toxic side effects. Nevertheless, the low number of discovered aptamers, the low stability in serum and the potential immune stimulatory effects are among the factors limiting the wide application as delivery tools.

2.5 Clinical status of ASOs, siRNA and CRISPR Cas9 therapeutics

Nature Reviews Drug Discovery, 2018, predict the oligonucleotide therapeutics along with small molecules and biologics to become the third major platform of drug development.[135] The forty years of research on ASOs results in five ASOs approved by the FDA and at least ten others in phase III clinical trials. Furthermore, the fifteen years journey to employ siRNA as therapeutics resulted in approval of the first one – Patasiran, and six more in ongoing phase III clinical trials. Representative clinical trials for ASOs, siRNA and CRISPR Cas9 are presented in **Tables 2.2**, **2.3** and **2.4**, respectively.

Table 2.2. Status on ASO therapeutics in phase II and III clinical trials[136,137]

Name	Chemistry	Target	Indication	Phase/Stage	Trial ref.	Developer
Modified ASOs						
Isis 420915	PS, 2'-MOE	Transthyretin mRNA	Senile Systemic Amyloidosis	II withdrawn	NCT02627820	Ionis
AVI-4658	PMO	Exon 51	DMD	II completed	NCT00159250	London Imperial Coll.
ISIS 104838	PS, 2'-MOE	TNF-alpha mRNA	Active Rheumatoid Arthritis	II completed	NCT00048321	Ionis
ISIS-CRP Rx	PS, 2'-MOE	C-reactive protein	Paroxysmal atrial fibrillation	II completed	NCT01710852	Ionis
OGX-427	2'-MOE	Heat shock protein	Prostate cancer	II completed	NCT01120470	British Cancer Agency
ISIS 5132	PS	c-raf-1	Breast cancer	II completed	NCT00003236	Ionis
G3139	PS	Bcl-2	Lung cancer	II completed	NCT00005032	NCI
IONIS-STAT3Rx	PS, 2'-MOE	STAT3	Advanced cancers	II completed	NCT01563302	Ionis/AstraZeneca
IONIS DGAT2Rx	PS, 2'-MOE	Diacylglycerol O acyltransferase	Hepatic Steatosis	II completed	NCT03334214	Ionis
IONIS-HTTRx	PS, 2'-MOE	HD protein	Huntington's Disease	III started	NCT02519036	Ionis/Roche
IONIS-DMPKRx	cEt	DMPK	Myotonic Dystrophy Type 1	II competed	NCT02312011	Ionis
ISIS-GCGRRx	PS, 2'-MOE	Glucagon receptor	Type 2 Diabetes	II completed	NCT02824003	Ionis
IONIS-TTR Rx	PS, 2'-MOE	Transthyretin	Amyloid Polyneuropathy	III active	NCT02175004	Ionis
Volanesorsen	PS, 2'-MOE	Apolipoprotein C-III	Chylomicronemia Syndrome	III active	NCT02658175	Ionis/Akcea
Alicaforsen	PS	Adhesion molecule 1	Ulcerative colitis	III completed	NCT00048113	Ionis
Miravirsen	PS, LNA	MIRN122	Chronic hepatitis C	II completed	NCT01200420	Roche/Santaris
Cobomarsen	LNA	miR-155	T-cell lymphoma	II recruiting	NCT03713320	miRagen
Apatorsen	2'-MOE	HSP27 protein	Pancreatic cancer	II completed	NCT01844817	OncoGenex
Conjugated ASOs						
ISIS 678354	GalNAc	Apolipoprotein C-III	Hypertriglyceridemia	II recriuting	NCT03385239	Ionis/Akcea
Imetelstat	PS, Palmitate	Telomerase	Myelofibrosis	II completed	NCT01731951	Janssen
ISIS 681257	2'-MOE, GalNAc	Apolipoprotein A	Hyperlipoproteinaemia	II completed	NCT03070782	Ionis/Akcea
ANGPTL3-LRx	2'-MOE, GalNAc	ANGPTL3 protein	Hypertriglyceridemia	II completed	NCT02709850	Ionis/Akcea
Zimura	PEG	complement factor C5	Polypoidal vasculopathy	II withdrawn	NCT03374670	OphthoTech

Table 2.3. Satatus on siRNA therapeutics in phase II and III clinical trials[136,138]

Name	Chemistry	Target	Indication	Phase/Stage	Trial ref.	Developer
Modified siRNA						
Teprasiran	2' OMe	Protein p53	Surgery	III recruiting	NCT03510897	Quark
PF-04523655	2' OMe	RTP801/REDD1	Diabetic macular oedema	II completed	NCT01445899	Quark
ARC-520	PS, 2'OMe, 2'F	/	Hepatitis B	II terminated	NCT02604199	Arrowhead
Remlarsen	LNA	Collagen	Keloid	II recruiting	NCT03601052	miRagen
Bevasiranib	dT capp	VEGF	Macular edema	II terminated	NCT00306904	OPKO
Conjugated siRNA						
Inclisiran	GalNAc	PCSK9 protein	Hypercholesterolemia	III active	NCT03397121	Alnylam
Fitusiran	GalNAc	Antithrombin III	Haemophilia	III recruiting	NCT03549871	Alnylam
Givosiran	GalNAc	5 aminolevulinate synthetase	Intermittent porphyria	III active	NCT03338816	Alnylam
Lumasiran	GalNAc	Glycolate oxidase	Primary Hyperoxaluria	II enrolling	NCT03350451	Alnylam
ARC-AAT	Cholesterol	Hepatic Z-AAT	Alpha 1-antitrypsin deficiency	II terminated	NCT02363946	Arrowhead
STP705	Polypeptide NP	Cyclo-oxygenase 2	Hypertrophic scars	II recruiting		Sirnaomics

Table 2.4. CRISPR Cas9 involvement in clinical trials[136,139]

Name	Mechanism	Indication	Phase/Stage	Trial ref.	Developer
CTX001	Ex vivo modif. CD34 and hHSPCs	Beta-thalassaemia	II recruiting	NCT03655678	CRISPR/Verthex
UCART019	Ex vivo, Editing CAR-T Cells CD19	CD19+ leukemia and lymphoma	II recruiting	NCT03166878	Chinese PLA
CRISPR CCR5	Ex vivo, Modification of CD34+	HIV-1	I recruiting	NCT03164135	Peking University
Anti-mesothelin CAR-T cells	Ex vivo, Modification of CAR.T Cells	Solid tumors	I recruting	NCT03545815	Chinese PLA
PD-1 Knockout T Cells	Ex vivo, T cells modification	Esophageal cancer	II recruiting	NCT03081715	Anhui Kedgene
PD-1 Knockout EBV-CTLs	Ex vivo, PDCD1 gene knock out	Epstein-Barr virus	II recruiting	NCT03044743	Yang Yang

2.6 Synthetic Nucleic Acid Analogues in Gene Therapy: An update for Peptide-Oligonucleotide Conjugates

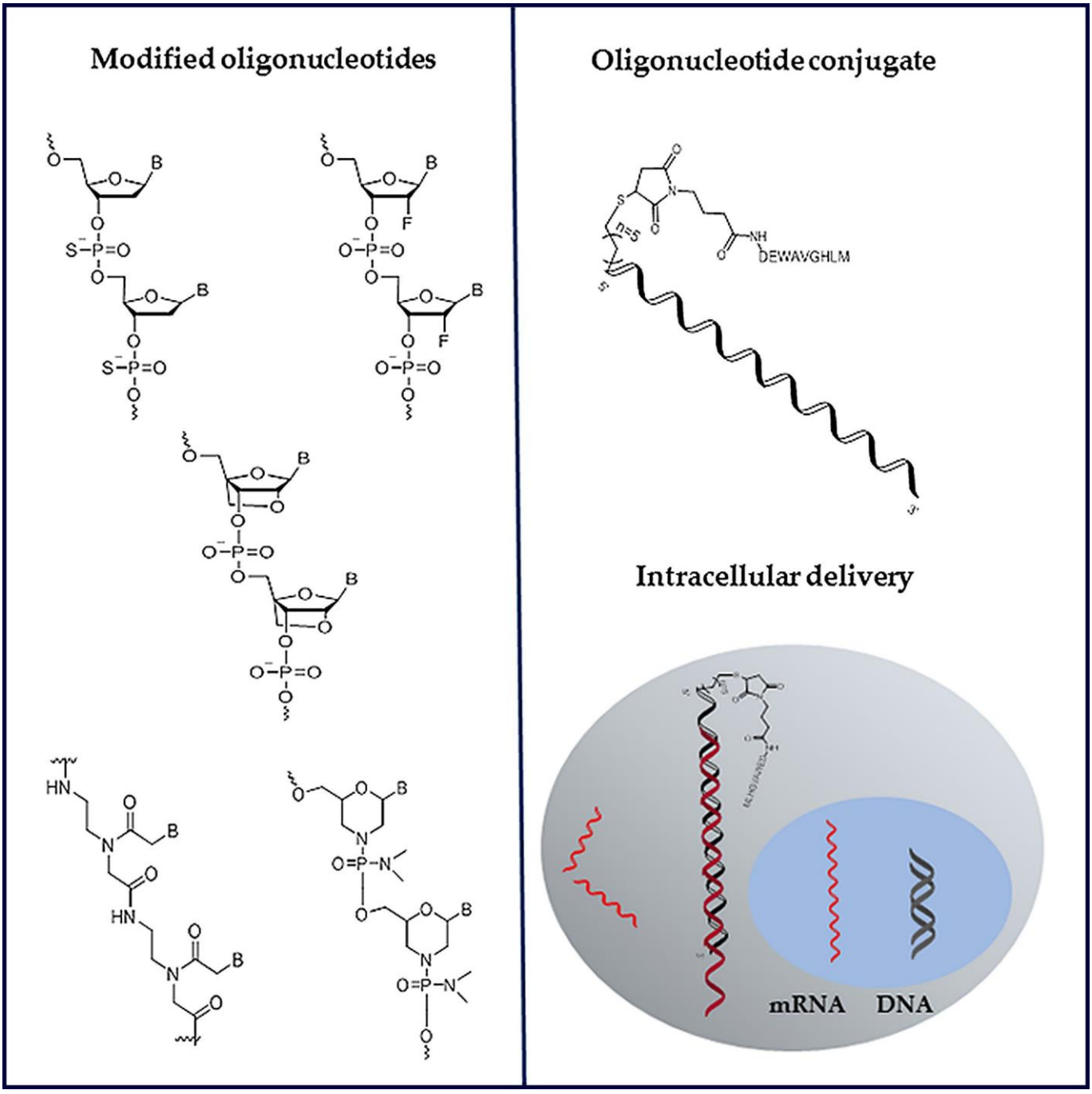
Taskova M, Mantsiou A, Astakhova K., Chembiochem., 2017, 17, 1676-82.

"Reprinted (adapted) with permission from (Taskova M, Mantsiou A, Astakhova K., Chembiochem., 2017, 17, 1676-82) Copyright (2017) John Wiley and Sons."

ChemPubSoc Europe
DOI: 10.1002/cbic.201700229
CHEMBIOCHEM Reviews

Synthetic Nucleic Acid Analogues in Gene Therapy: An Update for Peptide–Oligonucleotide Conjugates

Maria Taskova,[a] Anna Mantsiou,[a] and Kira Astakhova*[a, b]



 Wiley Online Library T 2017 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

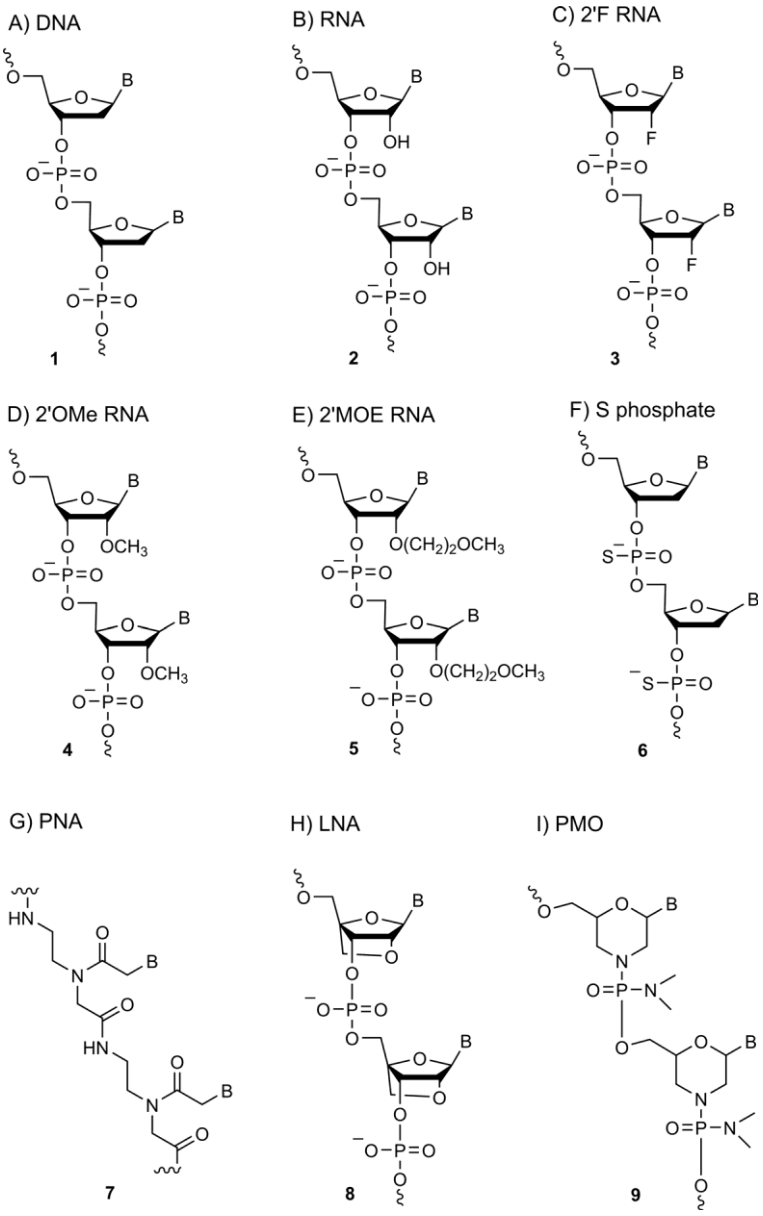
The main objective of this work is to provide an update on synthetic nucleic acid analogues and nanoassemblies as tools in gene therapy. In particular, the synthesis and properties of peptide–oligonucleotide conjugates (POCs), which have high potential in research and as therapeutics, are described in detail. The exploration of POCs has already led to fruitful results in the treatment of neurological diseases, lung disorders, cancer, leukemia, viral, and bacterial infections. However, delivery and in vivo stability are the major barriers to the clinical application of POCs and other analogues that still have to be overcome. This review summarizes recent achievements in the delivery and in vivo administration of synthetic nucleic acid analogues, focusing on POCs, and compares their efficiency.

1. Introduction

The development of the phosphoramidite chemistry of nucleosides opened up an exciting possibility for the synthesis of artificial DNA and RNA sequences.[1] By using automated oligonucleotide synthesis, synthetic nucleic acids have been widely investigated within diverse applications, such as therapeutic agents, diagnostic tools, and nanotechnology. As a result, after more than 30 years of intensive research, the high potential of synthetic oligonucleotides in personalized medicine has been confirmed. However, in order to be applied in vivo, synthetic nucleic acids still have to meet the requirements of stability in biofluids and specific delivery.[2]

This review elucidates synthetic nucleic acid analogues as potential nucleic acid therapeutics (NATs). In particular, conjugates of biomolecules with oligonucleotides are described, highlighting peptide–oligonucleotide conjugates (POCs) as a promising class of NATs. We discuss the major issues of NATs, and POCs in particular, that still need to be overcome. We comment on the delivery of NATs, which is considered to be a major barrier to their clinical application; summarize the methods used; and highlight the latest research solutions. Other issues related to NATs, such as in vivo stability, toxicity, and specific and selective therapeutic properties, are also discussed.

We outline several methods for overcoming these issues based on recent successful studies by multiple research groups. One exciting approach that we discuss is delivery that uses nucleic acid nanostructures, which can simultaneously aid uptake and monitor NAT performance in cells. As a final aspect, we deliberate chemical strategies for the preparation of NATs and compare their efficiency.



Scheme 1. Structures of modified second - and third-generation NATs. B: nucleobase.

[a] M. Taskova, A. Mantsiou, Prof. Dr. K. Astakhova
Nucleic Acid Center, Department of Physics
Chemistry and Pharmacy, University of Southern Denmark
Campusvej 55, 5230 Odense M (Denmark)

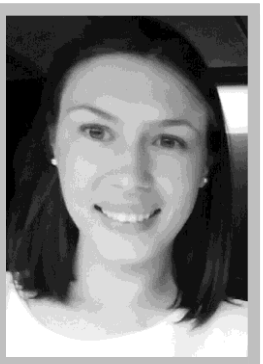
[b] Prof. Dr. K. Astakhova
Technical University of Denmark, Department of Chemistry
Kemitorvet, 2800 Kongens Lyngby (Denmark)
E-mail: kiraas@kemi.dtu.dk

The ORCID identification numbers for the authors of this article can be found underhttps://doi.org/10.1002/cbic.201700229.

2. Synthetic Nucleic Acid Analogues for Next-Generation Therapeutics

Synthetic oligonucleotides containing modified nucleotide analogues have already found multiple applications in clinical research, diagnostics, and life sciences.[3] For therapeutic applications, synthetic oligonucleotides have to meet several require-

ments; most importantly, they have to effectively target specific tissues and have selective therapeutic action within cells. Major concerns related to unmodified oligonucleotides used in vivo include the inability to cross biological membranes, accumulation in endosomes, poor pharmacokinetics, and biodistribution profiles.[4] Some of these problems were effectively solved by the synthesis of new-generation oligonucleotides; some still remain a big challenge. Useful strategies for the preparation and chemical modification of therapeutic oligonucleotides are discussed below.



Maria Taskova studied pharmacy at the Saint Cyril and Methodius University, Macedonia, and obtained a Master's degree in pharmacy in 2013. Afterwards, she carried out her Master's project in chemistry by studying novel peptide–oligonucleotide conjugates under the supervision of Prof. Kira Astakhova and gained her second M.Sc. in 2016 from the University of Southern Denmark. Currently, she is a Ph.D. student in the same laboratory. Her research interest is in the development of novel tools for ultraspecific targeting of nucleic acids.

Anna Mantsiou has a B.Sc. degree in biomolecular sciences and biotechnology. Having a passion for interdisciplinary research and a particular interest in nano-biotechnology, she obtained her M.Sc. degree in nano-bioscience and biophysics in 2016 from the University of Copenhagen. During this time, she worked on investigating interactions between cell membranes and transmembrane proteins, and she gained advanced skills in live-cell imaging. Currently, she is a researcher in the group of Prof. Kira Astakhova at the University of Southern Denmark, working on siRNA interactions and trafficking in living cells, with the aim of developing novel siRNA therapeutic tools.

Kira Astakhova obtained her Ph.D. in bioorganic chemistry in 2009. She continued to work as a postdoctoral researcher in the group of Prof. Jesper Wengel at the University of Southern Denmark, where she became Associate Professor in 2012. Until 2015, she was also a visiting professor at the University of California Santa Barbara, Department of Chemistry and Biochemistry, and at Stanford University, USA. She was awarded with several research prizes and fellowships, including a Marie Curie Early Stage Research Training Fellowship and Lundbeck Foundation research prize. Her research interests are interdisciplinary, combining organic chemistry, biomedicine, biophysics, and nano-bioscience.

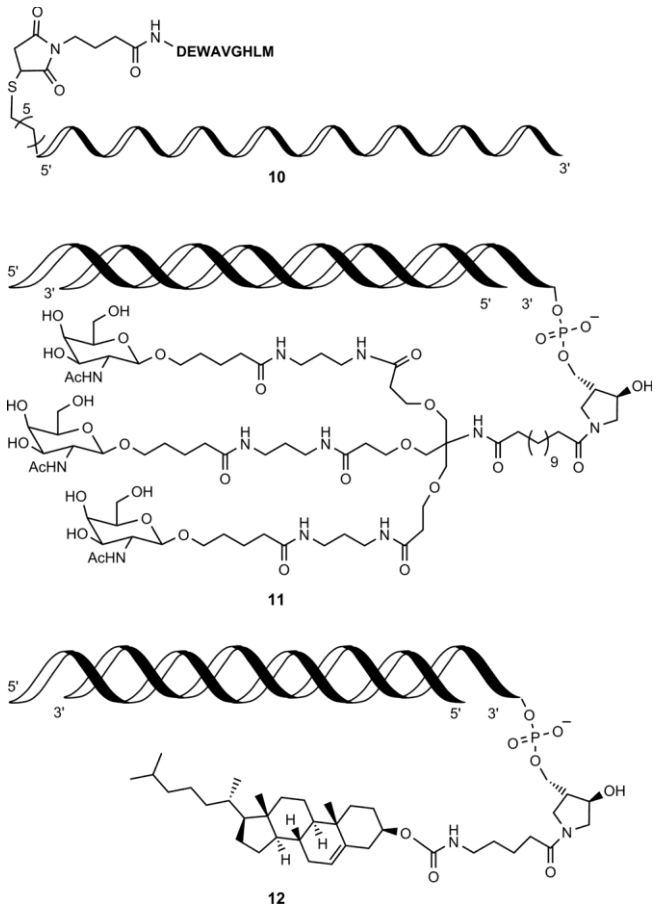
2.1. Towards improved gene therapy by using synthetic nucleic acid analogues

The implementation of various synthetic nucleic acid analogues leads to biomolecules and biomolecular assemblies with unique properties.[3, 5] Second-generation therapeutic oligonucleotides contain various modifications of the ribose sugar at position 2'.[6] Examples include fluorine (2'F), methoxy (2'OMe), and *O*-methoxyethyl (2'MOE) analogues of RNA nucleotides (Scheme 1). These modifications protect oligonucleotides from degradation by endonucleases.[4] Also, the replacement of oxygen by sulfur in the phosphate backbone increases the binding of serum proteins and makes phosphothioates more resistant to degradation.[7]

The third generation of modified oligonucleotides includes replacement of the phosphodiester backbone with peptide linkages that result in peptide nucleic acids (PNAs), replacement of the backbone with methylenemorpholine rings and phosphorodiamidate linkages that give rise to phosphorodiamidate morpholino oligomers (PMOs), and incorporation of locked nucleic acid (LNA) into the oligonucleotide sequence (Scheme 1).[6] PNAs and PMOs show good binding affinity to DNA/RNA targets and fair cellular uptake.[4, 8] Oligonucleotides that contain LNA analogues show excellent affinity for binding to complementary DNA/RNA, good mismatch discrimination, and high enzymatic stability.[3]

Overall, the second and third generations of oligonucleotides have made a big step towards overcoming the multiple disadvantages of natural DNA and RNA strands.[4] However, there is a need for new strategies to overcome the biggest challenge that faces gene therapy, that is, targeted intracellular delivery. In addition to chemical modification at the monomer level, bioconjugation of therapeutic oligonucleotides with different biomolecules is considered to be an excellent strategy that leads to modern, improved, fourth-generation NATs.

Bioconjugation can be conducted during automated solid-phase synthesis or after it by solution-phase conjugation. Strategies such as click chemistry and phosphoramidation reactions can be used to conjugate peptides, lipids, carbohydrates, or other nucleic acids to oligonucleotides.[2b] Recent examples of chemical conjugation with improved properties of NATs are shown in Scheme 2. In particular, cell-penetrating peptides (CPPs) are reported to be promising for conjugation with oligonucleotides, resulting in improved cellular specificity and intracellular delivery.[4] Boeneman et al. highlighted the use of functional peptide analogues as tools for cellular drug delivery that offered efficient uptake and endosomal escape.[9] The authors defined the key elements within the peptide sequence that had to be taken into account for such activity.[9] In another example, a peptide complex of poly(ethylene glycol)-*block*-polyethylenimine as a potential carrier of small interfering RNA (siRNA) to tumors in vivo was reported.[10] This system selec-



Scheme 2. Selected oligonucleotide conjugates that show improved properties.

tively targets CD44-overexpressed cancer cells and accumulates to significant levels in solid tumors. Thus, it inhibits primary tumor growth and prolongs survival in two different mouse models.

Receptor-mediated targeting is reported as a potential strategy for directed oligonucleotide delivery.[11] In one work, antisense oligonucleotide was conjugated to bombesin (BBN) peptide to investigate delivery and its mechanism in PC3 cells expressing the gastrin-releasing peptide receptor (10; Scheme 2). It was seen that the conjugate entered PC3 cells by receptor-mediated endocytosis. However, further trafficking the conjugate inside of the cell was not achieved.[11] This work highlights the complexity of intracellular delivery and highlights the need for further research on this topic.

Numerous successful approaches for the in vivo delivery of siRNA include the use of lipids either in lipid nanoparticle formulation or as lipid covalent conjugates. Recent studies demonstrated that siRNAs encapsulated into lipid nanoparticles successfully inhibited the expression of disease-relevant genes in humans.[12] Alternatively, the covalent conjugation of cholesterol, palmitic acid, and other lipophilic moieties directly to siRNA molecules has been shown; this results improved in vivo pharmacokinetics and increased cellular uptake, resulting to a great extent of gene silencing in multiple tissues.[13] siRNA conjugates with cholesterol are most extensively used (12; Scheme 2). Cholesterol forms conjugates with lipoproteins in the bloodstream; this leads to an increased circulation time of siRNA, which is further inserted into hepatocytes though lipo-protein receptors.[13a] siRNAs have recently also been modified with other lipids, such as *a*-tocopherol or palmitic acid.[14]

There are cell-surface receptors that are expressed or overexpressed only in certain tissues either naturally or as a result of a disease condition.[15] These receptors are potential targets for NAT delivery.[15] For example, the covalent attachment of the highly efficient ligand *N*-acetylgalactosamine (GalNAc) to siRNA has been recently reported to robustly silence gene expression in liver cells (11; Scheme 2).[16] GalNAc is an asialoglycoprotein receptor ligand expressed mainly in liver hepatocytes that mediates clearance of circulating glycoproteins through the clathrin endocytic pathway.[17] More specifically, bis- or tris-antennary GalNAcs attached to chemically modified siRNAs[16] have proved to improve siRNA stability against nucleases, as well as improve its pharmacokinetics; this results in potent silencing of the targeted gene in primary mouse hepatocytes.[16] GalNAc is a high-affinity ligand that exhibits a high internalization rate and it is easily attached to oligonucleotides by solid-phase synthesis; this makes it an advantageous candidate for siRNA-targeted delivery.[16] Alnylam Pharmaceuticals are currently performing clinical trials on three GalNAc–siRNA conjugates for the treatment of various diseases, such as hemophilia and amyloidosis.

3. Synthetic Peptides and Their Analogues for NAT Bioconjugation

Chemical conjugates of NATs with bioactive peptides have advantageous pharmacological properties, cellular uptake, and therapeutic efficacy.[18] By using advanced methods for peptide synthesis, pure, biologically active peptides can now be prepared. In particular, solid-phase automated peptide synthesis (SPPS),[19] artificial cellular factories, and advanced purification strategies find application in the field of synthetic peptides.[20]

The development of disease-specific and nontoxic peptides is required and their toxicity is an issue that has been studied. Table 1 lists antimicrobial peptides (AMPs)[21] and CPPs[22] in various stages of development, from preliminary analysis to clinical testing, and their reported toxicity. Small peptides, such as AMPs, CPPs, and tumor-targeting peptides, do not accumulate in specific organs. This minimizes their toxicity compared with that of large proteins.[23] Moreover, they are less immunogenic compared with antibodies and proteins.[23] On the other hand, the *trans*-activator of transcription (TAT), which is one of the most common CPPs, recently showed undesirable biological effects. It was reported that TAT induced apoptosis of human lung epithelial BEAS-2B cells and an inflammatory response.[24] Prior to stimulation with protein kinase C activator (PKC) of BEAS-B cells, they were pretreated with TAT. The results were inhibition of activity of the PKC. Conspicuously, TAT inhibited the production of multiple cytokines induced by PCK and the PCK activation-induced degradation of IkB*a* (alpha-nuclear factor of kappa light polypeptide gene enhancer in B-cell inhibitor).[24]

Table 1. Representative AMPs and CPPs with sequences and toxicity reports.

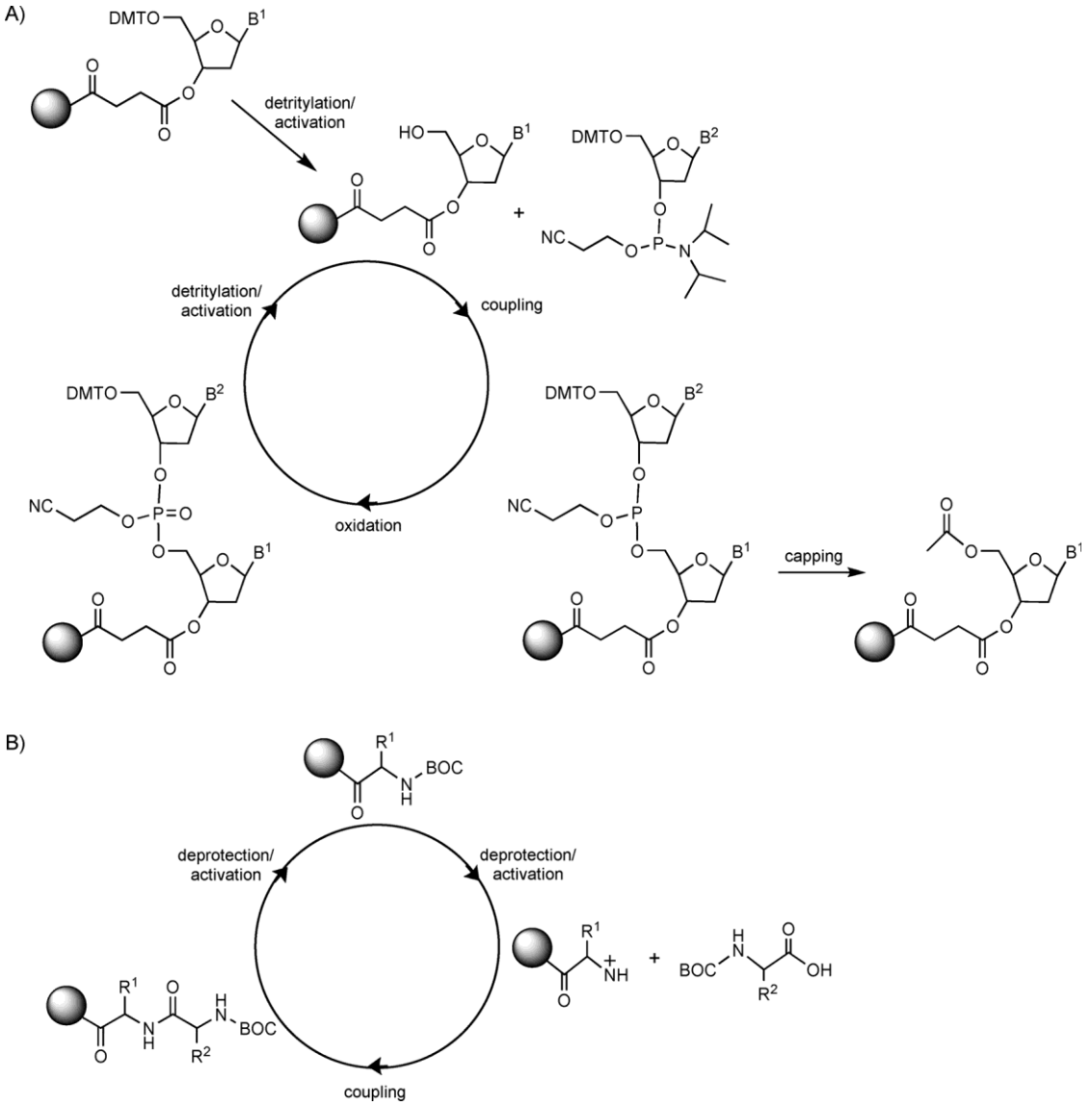
Peptide	Target	Sequence	Synthetic method	Investigated toxicity
kappacin[28]	Gram-positive/-negative bacteria	MAIPPKKNQDKTEIPTINTIASGEPTSTPT-TEAVESTVATLEDSPEVIESPPEINTVQVT-STAV	SPPS	not reported
esculentin-1a(1–21)[29]	Gram-negative bacteria	GIFSKLAGKKIKNLLISGLKG-NH_2	SPPS	not toxic against human erythrocytes and lung epithelial cells
Phd1[30]	*Candida albicans*	ACCCPIFTKIQGT–YRGKAK–CCCCCCK	SPPS, disulfide linkages	not reported
Phd2[31]	*Escherichia coli* and *Staphylococcus aureus*	FCCCPRRYKQIGT–GLPGTK–CCCCCCK	SPPS, disulfide linkages	not reported
temporin A[32]	*Leishmania*	FLPLIGRVLSGIL-NH_2	isolated from skin of *Rana temporaria*	not reported
defensin A[33]	malaria parasites	ATCDLLSGFGVGDSACAAHCIARGNRGG-YCNSKKVCVCRN	isolated from seeds of *Proteus vulgaris* L.	not reported
HβD3[34]	Gram-positive/-negative bacteria	GIINTLQKYYC–RVRGGRCAVLSCLPKEEQ-IGKCSTRGRKCCRRKK	SPPS, disulfide linkages	not toxic against human erythrocytes
BR2[26]	cancer, mutated K-ras	RAGLQFPVGRLLRRLLR	SPPS	not toxic against HaCat human keratinocytes and BJ human fibroblasts
LMWP[27]	antagonist of heparin/CPP	VSRRRRRRGGRRRR	enzymatic digestion of protamine	not toxic against white blood cells and platelets
[KW]4-E4[35]	human breast adeno-carcinoma	KWKWKWKWEEEE	SPPS	not reported
magainin II[36]	bladder cancer cell lines (RT4,647V,486P)	GIGKFLHSAKKFGKAFVGEIMNS	SPPS	not toxic against human fibroblast cell lines
pleurocidin[37]	*E. coli*	GWGSFFKKAAHVGKHVGKAALTHYL	isolated from *Pleuronectes americanus*	not toxic against human dermal fibroblast
NRC-3[38]	breast carcinoma cells	GRRKRKWLRRIGKGVKIIGGAALDHL	SPPS	not toxic against human dermal fibroblast

Evidently, the amino acid sequence, length, and various modifications can have a huge influence on the performance of the peptides, including on the toxicity profile.[25] Lim et al. reported the design of a new 17 amino acid long CPP called BR2 through the elimination of the C-terminal *a*-helical motif from their previous model CPP. This peptide was selectively cytotoxic against HeLa cells and HCT116 human colon cancer cells, but not to HaCat human keratinocytes and BJ human fibroblasts.[26] Next, low-molecular-weight protamine (LMWP), a 14-residue fragment CPP, achieved by enzymatic digestion of native protamine was reported not to be toxic, immunogenic, or antigenic.[27] The immunogenicity of LMWP in mice was examined, and a 46% reduction in production of antibodies compared with that of protamine was observed.[27] Importantly, LMWP showed decreased hemodynamic and hematological toxicity than that of protamine.[27]

SPPS is a powerful method for synthesizing homogeneous protein segments with the possibility of post-translational modification (Scheme 3 B).[39] Automated solid-phase oligonucleotide synthesis (Scheme 3 A) and SPPS have several key principles in common, namely, solid-support attachment of the first nucleotide/amino acid, automation, the use of an excess of reagents, and a series of deprotection–coupling reactions to grow the chain.[20a] Nevertheless, SPPS chemistry and oligonucleotide synthesis are not compatible due to the incompatibility of the deprotection chemistries. For cleavage from the support and the side-chain deprotection step, acidic conditions are used in peptide synthesis.[40] Conversely, in DNA/RNA oligonucleotide synthesis, basic conditions are necessary. Therefore, the syntheses cannot be combined in a single instrumental setting.[41]

Although SPPS is an effective method for preparing short peptides, such as AMPs, it is not useful for preparing longer peptides (the typical limit of SPPS is 50 amino acids).[20a] To overcome this limitation, chemical ligation of two or more short sequences can be used.[39] Recent developments related to these methods have been reviewed by Harmand et al.[39] *a*-Ketoacid–hydroxylamine (KAHA) ligation, serine/threonine ligation, alkyne–azide ligation, and amide-forming ligations are among the most effective chemical ligation methods available to date.[39] More specifically, KAHA ligation with 5-oxaproline can be used to make depsipeptides and overcome the problem of low solubility of hydrophobic sequences.[39] Serine/threonine ligation through native amide-bond formation is performed at the N-terminal serine and threonine residues.[39] Moreover, alkyne–azide click ligation produces a stable triazole group in the peptide backbone. The main advantages of this method include regioselectivity, high yields, and compatibility with a diverse range of aqueous buffers at different pH values. Similarly, the amide-forming ligation of potassium acyltrifluoroborates (KATs) and *O*-carbamoylhydroxylamines is a chemoselective assembly reaction.[39]

Alternatives to chemical methods for the synthesis of peptides are chemoenzymatic methods. The chemoenzymatic approach is a mild procedure that can provide peptides in high yields. It uses substrate-specific enzymes to form stereoselective peptide bonds.[42] As discussed by Numata and Yazawa, cysteine proteases, serine proteases, and esterases are often used in peptide synthesis.[42] In addition, engineered enzymes created by either random or side-directed mutagenesis can also be used.[42, 43] One example is nonribosomal peptide synthetases (NRPS), which can be re-engineered at the genetic



Scheme 3. General chemistries of A) solid-support oligonucleotide synthesis and B) solid-phase peptide synthesis. DMT: 4,4'-dimethoxytrityl, BOC: *tert*-butyloxycarbonyl.

level to generate specific modified bioactive peptides.[43] However, developing standard protocols and rules for the recombination of diverse NRPS modules and domains is a challenging task.[43] Calcott and Ackerley summarized the three general strategies employed by researchers to achieve this goal, that is, substitution of the A domain, target alteration of the substrate-binding pocket of the A domain, and substitution of inseparable C–A domains.[43]

Similar to oligonucleotides, generally, there are two methods for the modification of peptides: 1) SPPS modification, and 2) solution-phase modification of presynthesized peptides or bioconjugation.[44] Algar et al. reported methods for modifying peptides with 4-formylbenzoyl and 2-hydrazinonicotinoyl groups.[44] Furthermore, they described the aniline-catalyzed ligation of both functional groups, as well as subsequent conjugation of the ligated peptides to CdSe/ZnS quantum dots. The conjugated product can be used for diverse applications from in vivo and in vitro imaging and sensing to drug delivery. Another method for solid-phase N-methylation of peptides was proposed by Jensen and Roodbeen.[45] They suggested that N-methylation of peptides improved target affinity and specificity. Moreover, the oral availability and in vivo half-life were positively influenced.[45]

Protein ubiquitination is a post-translational modification, in which ubiquitin is attached to a substrate protein.[46] Abeywardana and Pratt reported the successful SPPS of ubiquitin, a 76 amino acid long protein that regulated several intracellular recognition pathways.[47] Thus, synthetic ubiquitin analogues are exciting objects of study and promising reagents for bioconjugation with other peptides, proteins and, oligonucleotides.[47]

Relaxin-2 and -3 are peptides with vital biological functions that have recently been synthesized.[48] There are two known synthetic strategies for their synthesis: SPPS and recombinantDNA synthesis. In SPPS, the two chains are synthesized separately,

followed by sequential disulfide-bond formation. In contrast, recombinant DNA synthesis is a combination of the synthesis of the propeptide, oxidative folding, and cleavage of the C-peptide. It is reported that synthetic relaxins show excellent receptor specificity and in vivo stability.[48]

4. Synthesis and Properties of POCs

4.1. Peptides for modern NATs

The fact that some peptides significantly enhance the cellular delivery of oligonucleotides and lower the dose required for treatment inspire further syntheses and studies of POCs as NATs.[49] For example, recently, Fan et al. reported a bis-peptide–siRNA conjugate as a promising therapeutic agent for cancer treatment.[50] The bis-peptide conjugated to siRNA at the 3' termini of both strands inhibited the translation of target genes for a longer period than that with the unconjugated reference. These conjugates showed stability in 80% fetal bovine serum for up to 24 h. The conjugate showed better inhibition of xenograft tumor growth in athymic mice compared with that of the reference. The circulation time in vivo in the mice was also prolonged.[50] In another case, a chimeric peptide–siRNA conjugate designed to inhibit HIV-1 infection was reported to have efficient transduction in macrophages and astrocytes.[51] The chimeric peptide was constructed from the protein transduction domain to bind with siRNA and the cell-fusing domain to enter the cell.[51] In the study, it was observed that the silencing effect did not correlate with efficient transduction; this approach may be suitable for targeting residual HIV-1 infection, but ineffective against latent viral infection.[51]

CPPs with short protein transduction domains are able to effectively penetrate cells.[52] The arginine amino acid is reported to be essential to the ability of the peptide to enter the cell. Thus, CPPs conjugated to other biomolecules enhance intracellular delivery.[53] Polyarginines and transportin were the first CPP sequences used.[54] Gait et al. proved that CPPs were taken up into the cell through endocytosis.[41, 53a] The major endocytic pathway for CPPs is either macropinocytosis or clathrin-mediated uptake.[54] It is also proven that this cellular uptake depends on the temperature and cellular adenosine triphosphate (ATP) pool.[53a] More recently, a non-endocytotic pathway for CPP–siRNA conjugates was reported.[55] In recent reviews, it has also been discussed that the mechanism of uptake for siRNA–CPPs depends on the cell-line type and can be affected by cell fixation reagents. Therefore, in vivo monitoring of uptake is crucial for making final conclusions on the intracellular faith of CPP-derived POCs.[56]

Pseudocomplementary peptide nucleic acids (pcPNAs) can form a hybrid G-quadruplex with double-stranded DNA (dsDNA). Kameshima et al. reported that, if a pyrrole/imidazole polyamide was conjugated to a pcPNA, hybridization with dsDNA was efficient under physiological conditions.[57] Specific cleavage of long dsDNA has been observed upon the application of pcPNAs. The authors prove that the pcPNA and pyrrole/imidazole polyamide worked together in a cooperative matter to enable more efficient targeting of dsDNA.[57] Table 2 provides a summary of representative conjugated sequences of CPPs with NATs and their application for disease treatment. Some interesting approaches use two elements of the POC simultaneously: the oligonucleotide and the peptide. In recent work, these elements have targeted the recognition sites of the c-Jun transcription factor.[64] Given the ability of oligonucleotides to control peptide conformation and regulate protein activity, the combination of peptides and oligonucleotides as POCs is a highly effective approach when the two parts of the POC work together.[64] In addition, the potential of multivalent designs that are capable of modulating more functions of a target protein has been highlighted by Dupont et al.[65] This approach is a good alternative to functional knockout. Such a multivalent molecule was synthesized by combining three artificial inhibitors, two aptamers, and a peptide. The conjugated moieties bind the serine protease urokinase-type plasminogen activator at three different sites and simultaneously prevent all functional interaction of the same.[65]

Table 2. Representative CPPs, their conjugates, and therapeutic applications.[a]

CPP	Sequence	Conjugated ON	Application, model
Pip5e[58]	RXRRBRRXR–ILFQY–RXRBRXRB	PMO	DMD, *mdx* mice, and patients
Pip6a[59]	RXRRBRRXR–YQFLI–RXRBRXRB	PMO	DMD, mouse model
CPP[60]	KFFKFFKFFKFFK	PNA	antibacterial against *E. coli*, BALB/c mice
CPP[61]	RFFRFFRFFRXB	PMO	antibacterial against *E. coli*, mice
CPPP7[62]	RXRRXRRXRRXRXB	PMO	antiviral to RSV, BALB/c mice
NLS[63]	PKKKRKV	PNA	SCID, mice inoculated with human BL cell lines
Antp[49a]	KKKKKWKMRRNQ-FWIKIQR	PNA	hematologic disorder, mice

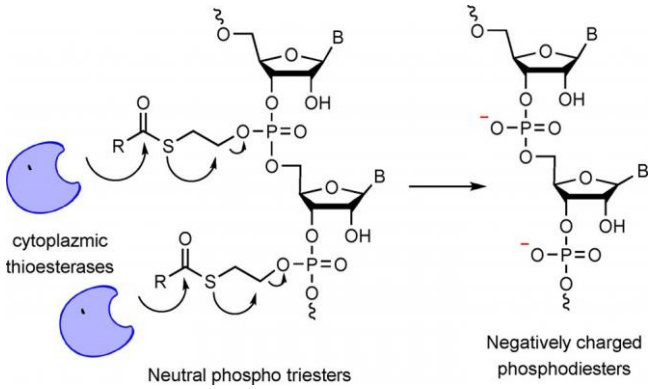
[a] B: β-alanine, DMD: Duchenne muscular dystrophy, ON: oligonucleotide, RSV: respiratory syncytial virus, SCID: severe combined immunodeficiency, X: 6-aminohexanoic acid.

Taking into account all of the aforementioned results, CPPs and their conjugates are promising candidates for conjugation to NATs. However, despite knowledge gained of CPPs in recent decades, there is still no unified CPP design that can internalize various cargos into all cell types. This obstacle creates a demand for efficient methods to develop new CPPs. Peptide libraries and quick screening assays represent a possible solution to this problem. One such method is the selection of peptide conjugates (SELPEPCON).[66] It is practiced as a strategy for the high-throughput synthesis of conjugates. Afterwards, it is used for screening, which eliminates the need for HPLC purification. Through the use of SELPEPCON, different CPPs were investigated, and the best candidates for delivery into the cells were discovered. Accordingly, 78 peptides were developed and their biological activities were investigated.[66] The splicing redirection of PNA705 in a HeLa cell assay was affected by the number of positive charges in the peptide, particularly arginine residues. CPPs with good properties have peptides with at

least eight charged arginines. Also, it was concluded that effective cell penetration of peptides was dependent on both the cargo and cell.[66]

For the molecular basis of POC structure and function, molecular charges are of high importance. Whereas CPPs are cationic, oligonucleotides have a negative charge. Charge difference also affects the synthesis of POCs. Making conjugates of peptides such as CPPs with canonical oligonucleotides can be challenging because these molecules tend to form aggregates.[54] In addition, the use of long peptide sequences that form secondary structures creates a problem in the conjugation reaction. However, optimization of the reaction conditions and the use of a denaturing agent could be a solution.[66]

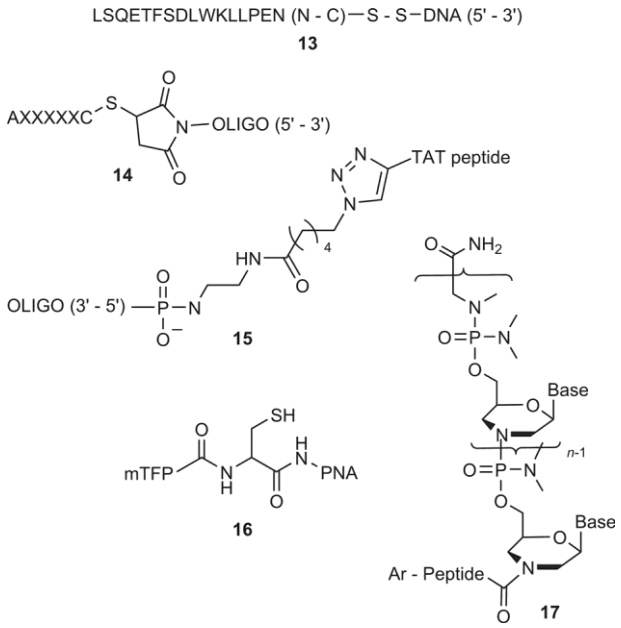
An alternative approach to the charge problem is the use of short interfering ribonucleic neutrals (siRNNs), as reported by Meade et al. (Scheme 4).[67] The siRNNs contain neutralizing bioreversible phosphotriester groups that allow their effective intracellular uptake and, excitingly, in vivo delivery. Tracking of siRNNs shows that, once inside the cell, they are converted into a native, charged phosphate backbone by cytoplasmic thioesterases. Thus, the therapeutic activity of the siRNN is regained. As a proof of concept, siRNNs conjugated to a hepatocyte-specific targeting domain were applied for systematic delivery into mice. The authors reported prolonged, dose-dependent in vivo responses in mice upon treatment with siRNNs; this makes them potent for further trials.[67]



Scheme 4. Bioreversible cytoplasmic transformation of siRNNs.

4.2. Synthetic strategies for POCs

POCs can be synthesized by biological or chemical methods.[40, 68] The latter are considered to be more useful for the synthesis of short oligonucleotides and peptides, and for the incorporation of other functionalities into the product conjugates. Several chemistries have been explored for conjugation reactions and they may affect the biological properties of POCs (Scheme 5).[40, 68a] For example, conjugates with a disulfide bond are reduced to sulfides in the cell-reducing environment; conversely, conjugates with amide bonds are more stable.[40] Chemically, POCs can be synthesized by one of two methods: in-line synthesis and fragment conjugation.[40, 68a, b] In the first method, both the peptide and oligonucleotide chains are assembled on a solid support by stepwise coupling reactions.[40]



Scheme 5. Representative chemistries of linkers for making POCs. mTFP: monomeric teal fluorescent protein.

However, due to the requirement for matching chemistries, this method is restricted to PNA conjugates with peptides. By employing the second method, the two fragments (peptide and oligonucleotide) are synthesized separately, cleaved from the support, purified, and conjugated by the reaction of choice (Scheme 3).

Site-specific modification of DNA/RNA oligonucleotides can be performed during solid-phase synthesis or postsynthetically.[53a] Insertion of various chemical moieties into the oligonucleotide during solid-phase synthesis has some limitations. Key limitations are the length of the oligonucleotide that can be synthesized and low yields of the product. To overcome this challenge, approaches for postsynthetic modifications are important. Lee et al. reported the preparation of a peptide–DNA conjugate in which an 18-nt oligodeoxyribonucleotide with 5'-thiol modification was conjugated with a fluorescein amidite (FAM) peptide through a disulfide bond (13; Scheme 5).[40] The product conjugate was efficiently internalized in HeLa cells by endocytosis. Circular dichroism (CD) spectroscopy showed that the optical and conformational characteristics of both macromolecules were affected after conjugation.[40] Other POCs were prepared by coupling a 5'-maleimide-modified oligonucleotide and a monomethoxytrityl-protected cysteine incorporated at the N terminus of the peptide (14).[69] By temperature melting (T_m) experiments, it was shown that the POCs bound the target more strongly than that of the unlabeled reference.[69]

A universal method employs a two-step phosphoramidation reaction to facilitate the conversion of azido or alkynyl groups into oligonucleotide precursors primed with a phosphate at the 5' termini. The precursor is then conjugated with a peptide by copper-catalyzed 1,3-dipolar azide–alkyne cycloaddition (CuAAC) to yield the desired POCs.[49b] CuAAC results in the formation of a chemically and biologically stable 1,2,3-triazole linker.[2b] In particular, DNA functionalized with azido groups

was successfully conjugated to an alkynyl-containing Tat peptide in a yield of 85 % (15). The cytotoxicity of the POC produced was investigated by using human A549 cells. Upon CuAAC-mediated synthesis, some cytotoxicity of the product was connected to residual copper. By using copper-free stain-promoted 1,3 dipolar azide–alkyne cycloaddition (SPAAC), cytotoxicity was eliminated. In the SPAAC reaction, the cyclooctyne alkyne reacts with the azido group without the use of copper(I); the reaction is stain-promoted.[70]

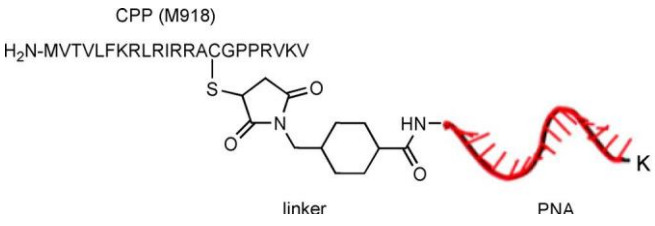
Although PNA–peptide conjugates can be made in a straightforward manner during SPPS,[40] fragment conjugation is reported to be a more convenient strategy.[71] Different methods for bioconjugation aimed at PNA–peptide conjugates and their analogues are reported: 1) conjugation through a disulfide bond: both peptide and PNA have thiol moieties, usually through the incorporation of cysteine at the C or N terminus; 2) thioether linkage: a thiol-functionalized segment reacts with a maleimide functionality; 3) oxime: the ketone functionality of one segment reacts with the aminooxy group of the other; and 4) *trans*-thioesterification: unprotected fragments containing C-terminal thioesters react with unprotected N-terminal cysteine.[71]

Additionally, expressed protein ligation (EPL) can be performed between the C-terminal thioesters of a protein and the N-terminal cysteine of a PNA, resulting in a native peptide bond.[72] Gholami et al. employed EPL to conjugate mTFP to the PNA (16; Scheme 5).[72] Using photophysical techniques, they visualized the behavior of the protein in cells, and demonstrated that this was a smart way to follow and control protein assembly.

Recently, Pip5e-PMO was created through an amide linkage (17; Scheme 5). In this synthesis, the 3'-morpholino group of the PMO was coupled to the C-terminal carboxylic acid of the peptide.[59] For direct amide conjugation, it is not necessary to use protecting groups for arginine-containing peptides;[73] this simplifies the synthesis and reduces its cost.

4.3. Specific case 1: A CPP–PNA conjugate that silences luciferase in colon adenocarcinoma cells (HT-29-luc)

Lee et al. synthesized CPP–PNA conjugates for specific delivery into the colon.[40] First, CPP-M918 was conjugated to the PNA through a maleimide–thiol coupling reaction. Then, the CPP–PNA conjugate was protected from nonspecific biological interactions by attaching poly(ethyleneglycol) (PEG) chains to the lysine residues. This was followed by reduction of the azobenzene bond, so that the CPP–PNA remained active



Scheme 6. Preparation of the CPP–PNA active conjugate.[40]

(Scheme 6).[40] This activation is expected to be performed in the colon by microflora. The authors showed that the conjugate achieved sequence-specific silencing of luciferase in cells. The developed NAT delivery could also be a new promising approach for the delivery of NATs into the colon.

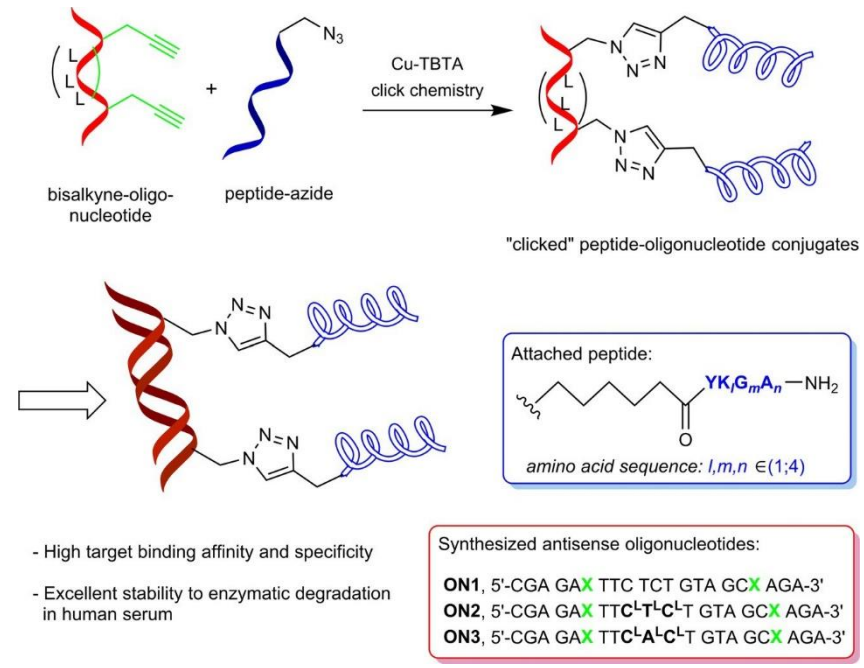
4.4. Specific case 2: New POCs targeting the V600E mutation in the BRAF oncogene

Recently, we synthesized doubly labeled 21-mer DNA/LNA oligonucleotides that targeted the *BRAF* oncogene containing the V600E mutation.[74] Through CuAAC click chemistry, the precursors were labeled internally with a library of rationally designed peptides (Scheme 7). Our synthetic approach provided 30 POCs in a time- and cost-effective manner, giving good yields (43–83%) and high purity of the products (90%). Our aim was to investigate the influence of different peptide sequences and LNA on the structure and properties of the synthetic oligonucleotides and to design POCs that were stable in biological media and had the potential to specifically bind the target in vivo. To evaluate this, spectroscopic characterization and examination of the stability of the resulting POCs to enzymatic degradation were performed.

The POCs displayed great T_m stability (Figure 1) and the peptide sequence had an influence on the specificity and affinity of target DNA/RNA binding. Lysine residues improve binding of POCs to target DNA and RNA, whereas the distance to lysine is associated with a decrease in binding of mismatched RNA targets. Glycine and tyrosine residues affect target binding as well. The resistance of POCs to enzymatic degradation is improved by the internal attachment of peptides, but not by LNA alone. Independently of the peptide sequence, POCs show stability up to 24 h in 90% human serum and POC duplexes with complementary DNA for up to 160 h in 90% human serum. Moreover, using competitive UV melting experiments, we demonstrated that POCs preferred to bind to RNA over DNA when designed as a POC:DNA duplex. We speculate that the antisense POC designed as a duplex with DNA, which is extremely stable enzymatically, has potential as a new type of antisense double-strand oligonucleotide therapeutic agent.[74]

5. Conclusions

The field of synthetic peptides, oligonucleotides, and their conjugates is growing constantly, providing new tools and exciting opportunities in life sciences, clinical research, nanotechnology, and industry. Although numerous chemistries for making NATs and successful proof-of-concept applications are reported, more studies are needed to improve in vivo performance. According to our recent research, internal incorporation of short peptides is a promising approach for NATs. Nevertheless, reports on the internal attachment of peptides to oligonucleotides are still scarce, and there is still an unmet need for extensive pharmacological studies of these conjugates. Moreover, there are numerous issues related to the efficiency, safety, and toxicity of the linkers and peptides applied for bioconjugation.[70]



Scheme 7. Synthesis of POC by means of CuAAC click chemistry.[74] TBTA: tris(benzyltriazolylmethyl)amine.

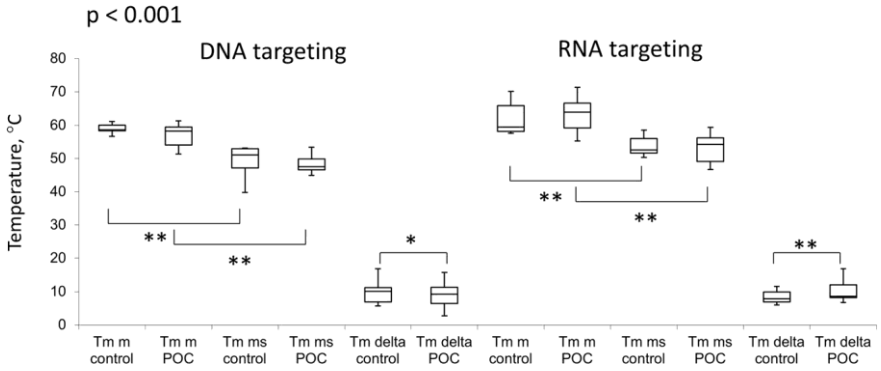


Figure 1. Box-and-whisker plot with outliers. The arms on each boxplot are values Q1-1.5xIQR and Q3+1.5xIQR. Q: quartile, IQR: interquartile range. Data points for each subject are means of two independent measurements. DNA/RNA targets: m: matched; ms: mismatched. DT_m is the difference between T_m for fully matched and mismatched duplexes. Results of the *t* test assuming unequal variances: the difference in means is statistically significant at the 90 (*) or 95% (**) significance level. Adapted with permission from ref. [74]. Copyright: American Chemical Society, 2017.

In general, the use of the aforementioned click chemistry (CuAAC and SPAAC) holds promise for further developments in POC synthesis. For the issue of delivery, it has been confirmed that decorating NATs with peptides, such as CPP and NLS, as well as incorporation of fatty acids, aids efficient uptake.[49a, 51–52] This implies that, in spite of sometimes being challenging, bioconjugation of NATs opens up exciting opportunities for a new class of therapeutics with previously unachievable activity in vivo.

Acknowledgements

Prof. Dr. Knud J. Jensen (University of Copenhagen, Denmark), Dr. Charlotte S. Madsen (University of Copenhagen), Dr. Birte Vester (University of Southern Denmark), and Lykke Hansen (University of Copenhagen) are acknowledged for valuable comments on this work.

Conflict of Interest

The authors declare no conflict of interest.

Keywords: click chemistry gene therapy oligonucleotides peptides synthesis design

[1] a) S. L. Beaucage, M. H. Caruthers, *Tetrahedron Lett.* 1981, *22*, 1859 – 1862; b) S. L. Beaucage, *Tetrahedron* 1992, *48*, 2223 –2311.
[2] a) M. A. Kay, *Nat. Rev. Genet.* 2011, *12*, 316 –328; b) I. K. Astakhova, L. H. Hansen, B. Vester, J. Wengel, *Org. Biomol. Chem.* 2013, *11*, 4240 –4249.
[3] I. K. Astakhova, J. Wengel, *Acc. Chem. Res.* 2014, *47*, 1768 –1777.

[4] R. L. Juliano, X. Ming, O. Nakagawa, *Acc. Chem. Res.* 2012, *45*, 1067 – 1076.
[5] J. Wengel, *Org. Biomol. Chem.* 2004, *2*, 277 –280.
[6] K. Sridharan, N. J. Gogtay, *British J. Clin. Pharmacol.* 2016, *82*, 659 –672.
[7] S. Kruspe, F. Mittelberger, K. Szameit, U. Hahn, *ChemMedChem* 2014, *9*, 1998 –2011.
[8] S. Paul, M. H. Caruthers, *J. Am. Chem. Soc.* 2016, *138*, 15663– 15672.
[9] K. Boeneman, J. B. Delehanty, J. B. Blanco-Canosa, K. Susumu, M. H. Stewart, E. Oh, A. L. Huston, G. Dawson, S. Ingale, R. Walters, M. Domowicz, J. R. Deschamps, W. R. Algar, S. Dimaggio, J. Manono, C. M. Spill- mann, D. Thompson, T. L. Jennings, P. E. Dawson, I. L. Medintz, *ACS Nano* 2013, *7*, 3778 –3796.
[10] M. Golan, V. Feinshtein, D. Polyak, A. Scomparin, R. Satchi-Fainaro, A. David, *Bioconjugate Chem.* 2016, *27*, 947 –960.
[11] X. Ming, M. R. Alam, M. Fisher, Y. Yan, X. Chen, R. L. Juliano, *Nucleic Acids Res.* 2010, *38*, 6567 –6576.
[12] M. A. Maier, M. Jayaraman, S. Matsuda, J. Liu, S. Barros, W. Querbes, Y. K. Tam, S. M. Ansell, V. Kumar, J. Qin, X. Zhang, Q. Wang, S. Panesar, R. Hutabarat, M. Carioto, J. Hettinger, P. Kandasamy, D. Butler, K. G. Rajeev, B. Pang, K. Charisse, K. Fitzgerald, B. L. Mui, X. Du, P. Cullis, T. D. Madden, M. J. Hope, M. Manoharan, A. Akinc, *Mol. Ther.* 2013, *21*, 1570 –1578.
[13] a) C. Wolfrum, S. Shi, K. N. Jayaprakash, M. Jayaraman, G. Wang, R. K. Pandey, K. G. Rajeev, T. Nakayama, K. Charrise, E. M. Ndungo, T. Zimmermann, V. Koteliansky, M. Manoharan, M. Stoffel, *Nat. Biotechnol.* 2007, *25*, 1149 – 1157; b) J. Soutschek, A. Akinc, B. Bramlage, K. Charisse, R. Constien, M. Donoghue, S. Elbashir, A. Geick, P. Hadwiger, J. Harborth, M. John, V. Kesavan, G. Lavine, R. K. Pandey, T. Racie, K. G. Rajeev, I. Rohl, I. Toudjarska, G. Wang, S. Wuschko, D. Bumcrot, V. Koteliansky, S. Limmer, M. Manoharan, H. P. Vornlocher, *Nature* 2004, *432*, 173 –178; c) C. Lorenz, P. Hadwiger, M. John, H. P. Vornlocher, C. Unverzagt, *Bioorg. Med. Chem. Lett.* 2004, *14*, 4975 –4977.
[14] K. Nishina, T. Unno, Y. Uno, T. Kubodera, T. Kanouchi, H. Mizusawa, T. Yokota, *Mol. Ther.* 2008, *16*, 734 –740.
[15] a) Y. Wang, H. Du, G. Zhai, *Curr. Drug Targets* 2014, *15*, 573 –599; b) N. K. Mehra, V. Mishra, N. K. Jain, *Ther. Delivery* 2013, *4*, 369 –394.
[16] J. K. Nair, J. L. Willoughby, A. Chan, K. Charisse, M. R. Alam, Q. Wang, M. Hoekstra, P. Kandasamy, A. V. Kel'in, S. Milstein, N. Taneja, J. O'Shea, S. Shaikh, L. Zhang, R. J. van der Sluis, M. E. Jung, A. Akinc, R. Hutabarat, S. Kuchimanchi, K. Fitzgerald, T. Zimmermann, T. J. van Berkel, M. A. Maier, K. G. Rajeev, M. Manoharan, *J. Am. Chem. Soc.* 2014, *136*, 16958– 16961.
[17] R. J. Stockert, *Physiol. Rev.* 1995, *75*, 591 –609.
[18] T. Tashima, *Bioorg. Med. Chem. Lett.* 2017, *27*, 121 –130.
[19] a) L. Menzker, A. Oddo, P. R. Hansen, *Methods Mol. Biol.* 2017, *1548*, 35– 49 ; b) B. K. W. Chung, C. J. White, A. K. Yudin, *Nat. Protoc.* 2017, *12*, 1277– 1287.
[20] a) P. T. Shelton, K. J. Jensen, *Methods Mol. Biol.* 2013, *1047*, 23– 41; b) T. Kimmerlin, D. Seebach, *J. Pept. Res.* 2005, *65*, 229 –260.
[21] a) J. Lakshmaiah Narayana, J. Y. Chen, *Peptides* 2015, *72*, 88– 94; b) M. A. Dawgul, K. E. Greber, W. Sawicki, W. Kamysz, *Curr. Med. Chem.* 2016, *24*, 654 –672.
[22] a) G. Guidotti, L. Brambilla, D. Rossi, *Trends Pharmacol. Sci.* 2017, *38*, 406 –424; b) S. Dissanayake, W. A. Denny, S. Gamage, V. Sarojini, *J. Con- trolled Release* 2017, *250*, 62– 76.
[23] S. Marqus, E. Pirogova, T. J. Piva, *J. Biomed. Sci.* 2017, *24*, 21.
[24] H. Kim, S. Moodley, M. Liu, *Drug Delivery Transl. Res.* 2015, *5*, 275–278.
[25] C. Chen, Y. Chen, C. Yang, P. Zeng, H. Xu, F. Pan, J. R. Lu, *ACS Appl. Mater. Interfaces* 2015, *7*, 17346– 17355.
[26] K. J. Lim, B. H. Sung, J. R. Shin, Y. W. Lee, D. J. Kim, K. S. Yang, S. C. Kim, *PLoS One* 2013, *8*, 66084 –67784.
[27] H. He, J. Ye, E. Liu, Q. Liang, Q. Liu, V. C. Yang, *J. Controlled Release* 2014, *193*, 63– 73.
[28] S. G. Dashper, N. M. O'Brien-Simpson, K. J. Cross, R. A. Paolini, B. Hoffmann, D. V. Catmull, M. Malkoski, E. C. Reynolds, *Antimicrob. Agents Chemother.* 2005, *49*, 2322 –2328.
[29] V. Luca, A. Stringaro, M. Colone, A. Pini, M. L. Mangoni, *Cell. Mol. Life Sci.* 2013, *70*, 2773 –2786.
[30] V. Krishnakumari, S. Singh, R. Nagaraj, *Peptides* 2006, *27*, 2607 –2613.
[31] V. Krishnakumari, N. Rangaraj, R. Nagaraj, *Antimicrob. Agents Chemother.* 2009, *53*, 256 –260.
[32] M. L. Mangoni, *Cell. Mol. Life Sci.* 2006, *63*, 1060 –1069.
[33] V. V. do Nascimento, O. Mello Ede, L. P. Carvalho, E. J. de Melo, O. Carvalho Ade, K. V. Fernandes, V. M. Gomes, *Biosci. Rep.* 2015, *35*, 248 –255.
[34] V. Dhople, A. Krukemeyer, A. Ramamoorthy, *Biochim. Biophys. Acta Biomembranes* 2006, *1758*, 1499 –1512.
[35] S. Mohammadi, J. Shahbazi Mojarrad, P. Zakeri-Milani, A. Shirani, S. Mussa Farkhani, N. Samadi, H. Valizadeh, *Adv. Pharm. Bull.* 2015, *5*, 41– 49.
[36] J. Lehmann, M. Retz, S. S. Sidhu, H. Suttmann, M. Sell, F. Paulsen, J. Harder, G. Unteregger, M. Stockle, *Eur. Urol.* 2006, *50*, 141 –147.
[37] A. M. Cole, P. Weis, G. Diamond, *J. Biol. Chem.* 1997, *272*, 12008 –12013.
[38] A. L. Hilchie, C. D. Doucette, D. M. Pinto, A. Patrzykat, S. Douglas, D. W. Hoskin, *Breast Cancer Res.* 2011, *13*, 102 – 118.
[39] T. J. Harmand, C. E. Murar, J. W. Bode, *Curr. Opin. Chem. Biol.* 2014, *22*, 115– 121.
[40] S. H. Lee, E. Moroz, B. Castagner, J. C. Leroux, *J. Am. Chem. Soc.* 2014, *136*, 12868 –12871.
[41] M. K. Lee, Y. B. Lim, *Bioorg. Med. Chem.* 2014, *22*, 4204 –4209.
[42] K. Yazawa, K. Numata, *Molecules* 2014, *19*, 13755– 13774.
[43] M. J. Calcott, D. F. Ackerley, *Biotechnol. Lett.* 2014, *36*, 2407 –2416.
[44] W. R. Algar, J. B. Blanco-Canosa, R. L. Manthe, K. Susumu, M. H. Stewart, P. E. Dawson, I. L. Medintz, *Methods Mol. Biol.* 2013, *1025*, 47– 73.
[45] R. Roodbeen, K. J. Jensen, *Methods Mol. Biol.* 2013, *1047*, 141 –149.
[46] D. Morimoto, E. Walinda, K. Sugase, M. Shirakawa, *Int. J. Mol. Sci.* 2017, *18*, 1145 – 1158.
[47] T. Abeywardana, M. R. Pratt, *ChemBioChem* 2014, *15*, 1547 –1554.
[48] M. A. Hossain, J. D. Wade, *Curr. Opin. Chem. Biol.* 2014, *22*, 47– 55.
[49] a) H. M. Moulton, *Methods Mol. Biol.* 2012, *867*, 407 –414 ; b) Y. C. Su, Y. L. Lo, C. C. Hwang, L. F. Wang, M. H. Wu, E. C. Wang, Y. M. Wang, T. P. Wang, *Org. Biomol. Chem.* 2014, *12*, 6624 –6633.
[50] X. Fan, Y. Zhang, X. Liu, H. He, Y. Ma, J. Sun, Y. Huang, X. Wang, Y. Wu, L. Zhang, Z. Yang, *Bioconjugate Chem.* 2016, *27*, 1131 –1142.
[51] S. Bivalkar-Mehla, R. Mehla, A. Chauhan, *J. Drug Targeting* 2017, *25*, 307 –319.
[52] E. Vives, P. Brodin, B. Lebleu, *J. Biol. Chem.* 1997, *272*, 16010 –16017.
[53] a) J. P. Richard, K. Melikov, E. Vives, C. Ramos, B. Verbeure, M. J. Gait, L. V. Chernomordik, B. Lebleu, *J. Biol. Chem.* 2003, *278*, 585 –590; b) A. As- triab-Fisher, D. Sergueev, M. Fisher, B. R. Shaw, R. L. Juliano, *Pharm. Res.* 2002, *19*, 744 –754.
[54] F. Marlin, P. Simon, T. Saison-Behmoaras, C. Giovannangeli, *ChemBioChem* 2010, *11*, 1493 –1500.
[55] T. Endoh, T. Ohtsuki, *Adv. Drug Delivery Rev.* 2009, *61*, 704 –709.
[56] L. P-rnaste, P. Arukuusk, E. Zagato, K. Braeckmans, U. Langel, *J. Drug Targeting* 2016, *24*, 508 –519.
[57] W. Kameshima, T. Ishizuka, M. Minoshima, M. Yamamoto, H. Sugiyama, Y. Xu, M. Komiyama, *Angew. Chem. Int. Ed.* 2013, *52*, 13681 –13684; *Angew. Chem.* 2013, *125*, 13926 –13929.
[58] H. Yin, A. F. Saleh, C. Betts, P. Camelliti, Y. Seow, S. Ashraf, A. Arzumanov, S. Hammond, T. Merritt, M. J. Gait, M. J. Wood, *Mol. Ther.* 2011, *19*, 1295 –1303.
[59] C. Betts, A. F. Saleh, A. A. Arzumanov, S. M. Hammond, C. Godfrey, T. Coursindel, M. J. Gait, M. J. Wood, *Mol. Ther. Nucleic Acids* 2012, *1*, 38– 51.
[60] X. X. Tan, J. K. Actor, Y. Chen, *Antimicrob. Agents Chemother.* 2005, *49*, 3203 –3207.
[61] L. D. Tilley, B. L. Mellbye, S. E. Puckett, P. L. Iversen, B. L. Geller, *J. Antimicrob. Chemother.* 2007, *59*, 66– 73.
[62] S. H. Lai, D. A. Stein, A. Guerrero-Plata, S. L. Liao, T. Ivanciuc, C. Hong, P. L. Iversen, A. Casola, R. P. Garofalo, *Mol. Ther.* 2008, *16*, 1120 – 1128.
[63] L. C. Boffa, G. Cutrona, M. Cilli, S. Matis, G. Damonte, M. R. Mariani, E. Millo, M. Moroni, S. Roncella, F. Fedeli, M. Ferrarini, *Cancer Gene Ther.* 2007, *14*, 220 –226.
[64] L. Rçglin, O. Seitz, *Org. Biomol. Chem.* 2008, *6*, 3881 –3887.
[65] D. M. Dupont, N. Bjerregaard, B. Verpaalen, P. A. Andreasen, J. K. Jensen, *Bioconjugate Chem.* 2016, *27*, 918 –926.
[66] P. J. Deuss, A. A. Arzumanov, D. L. Williams, M. J. Gait, *Org. Biomol. Chem.* 2013, *11*, 7621 –7630.
[67] B. R. Meade, K. Gogoi, A. S. Hamil, C. Palm-Apergi, A. van den Berg, J. C. Hagopian, A. D. Springer, A. Eguchi, A. D. Kacsinta, C. F. Dowdy, A. Pre- sente, P. Lonn, M. Kaulich, N. Yoshioka, E. Gros, X. S. Cui, S. F. Dowdy, *Nat. Biotechnol.* 2014, *32*, 1256 –1261.
[68] a) D. A. Stetsenko, M. J. Gait, *Methods Mol. Biol.* 2005, *288*, 205 –224; b) S. M. Hammond, G. Hazell, F. Shabanpoor, A. F. Saleh, M. Bowerman,

J. N. Sleigh, K. E. Meijboom, H. Zhou, F. Muntoni, K. Talbot, M. J. Gait, M. J. Wood, *Proc. Natl. Acad. Sci. USA* 2016, *113*, 10962 –10967; c) C. H. Tung, S. Stein, *Bioconjugate Chem.* 2000, *11*, 605 –618; d) K. Lu, Q. P. Duan, L. Ma, D. X. Zhao, *Bioconjugate Chem.* 2010, *21*, 187 –202.

[69] J. G. Harrison, S. Balasubramanian, *Nucleic Acids Res.* 1998, *26*, 3136 – 3145.

[70] I. S. Marks, J. S. Kang, B. T. Jones, K. J. Landmark, A. J. Cleland, T. A. Taton, *Bioconjugate Chem.* 2011, *22*, 1259–1263.

[71] M. C. de Koning, G. A. van der Marel, M. Overhand, *Curr. Opin. Chem. Biol.* 2003, *7*, 734 –740.

[72] Z. Gholami, L. Brunsveld, Q. Hanley, *Bioconjugate Chem.* 2013, *24*, 1378 – 1386.

[73] P. Boisgu8rin, S. Deshayes, M. J. Gait, L. O'Donovan, C. Godfrey, C. A. Betts, M. J. Wood, B. Lebleu, *Adv. Drug Delivery Rev.* 2015, *87*, 52– 67.

[74] M. Taskova, C. S. Madsen, K. J. Jensen, L. H. Hansen, B. Vester, K. Astakhova, *Bioconjugate Chem.* 2017, *28*, 768 –774.

Manuscript received: April 26, 2017

Accepted manuscript online: June 14, 2017

Version of record online: July 24, 2017

Chapter 3

Paper II

Antisense Oligonucleotides Internally Labeled with Peptides Show Improved Target Recognition and Stability to Enzymatic Degradation

Taskova M, Madsen CS, Jensen KJ, Hansen LH, Vester B, Astakhova K., *Bioconjug Chem.*, 2017, 3, 768-74.

"Reprinted (adapted) with permission from (Taskova M, Madsen CS, Jensen KJ, Hansen LH, Vester B, Astakhova K., *Bioconjug Chem.*, 2017, 3, 768-74.). Copyright (2017) American Chemical Society."

Antisense oligonucleotides internally labeled with peptides show improved target recognition and stability to enzymatic degradation

Maria Taskova,[1]* Charlotte S. Madsen,[2] Knud J. Jensen,[2] Lykke Haastrup Hansen,[3] Birte Vester,[3] Kira Astakhova[1]*

1 Nucleic Acid Center, Department of Physics, Chemistry and Pharmacy, University of Southern Denmark, Campusvej 55, 5230 Odense M, Denmark

2 Department of Chemistry, University of Copenhagen, Thorvaldsensvej 40, 1871 Frederiksberg, Denmark

3 Department of Biochemistry and Molecular Biology, University of Southern Denmark, Campusvej 55, 5230

Odense M, Denmark

E-mail: taskova@sdu.dk; ias@sdu.dk

ABSTRACT: Specific target binding and stability in diverse biological media is of crucial importance for applications of synthetic oligonucleotides as diagnostic and therapeutic tools. So far, these issues have been addressed by chemical modification of oligonucleotides and by conjugation with a peptide, most often at the terminal position of the oligonucleotide. Herein, we for the first time systematically investigate the influence of internally attached 8 amino acid-long peptides on the properties of antisense oligonucleotides. We report the synthesis and internal double labeling of 21-mer oligonucleotides that target the *BRAF* V600E oncogene, with a library of rationally designed peptides employing CuAAC "click" chemistry. The peptide sequence has an influence on the specificity and affinity of target DNA/RNA binding. We also investigated the impact of locked nucleic acids (LNAs) on the latter. Lysine residues improve binding of POCs to target DNA and RNA, whereas the distance to lysine correlates exclusively with a decrease in binding of mismatched RNA targets. Glycine and tyrosine residues affect target binding as well. Importantly, the resistance of POCs to enzymatic degradation is dramatically improved by the internal attachment of peptides but not by LNA alone. Independently of the peptide sequence, the conjugates are stable for up to 24 h in 90% human serum and duplexes of POCs with complementary DNA for up to 160 h in 90% human serum. Such excellent stability has not been previously reported for DNA and makes internally labeled POCs an exciting object of study, i.e. showing high target specificity and simultaneously stability in biological media.

As a result of recent technical advances and clinical success in the treatment of genetic diseases, gene therapy is a rapidly evolving approach in research and industry.[1] There are several successful examples of treating specific diseases with gene therapy, including neurological diseases,[2] lung disorders,[3] viral infections,[4] leukemia[5] and cancer.[6] In particular for cancer treatment, nucleic acid therapeutics (NATs) are able to directly target the gene

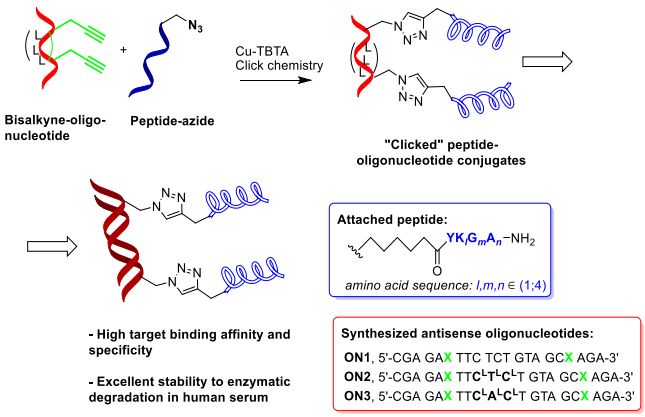


Figure 1. Design of highly stable POCs for oncogene targeting. l, m, n = number of K, G and A amino acids in the attached peptide. L = locked nucleic acid (LNA).

fragments responsible for developing tumorigenesis without causing systematic toxicity.[7] This is in contrast to the numerous toxic effects of conventional therapy.[8-10]

In spite of considerable progress in the field of diagnostic and therapeutic oligonucleotides, there are several issues that still need to be addressed. These include the *in vivo* stability, specificity and delivery of NATs. In particular, the stability of NATs in human serum is of crucial significance for their application as therapeutics.[11] Unmodified oligonucleotides are rapidly degraded by nucleases which are present in human serum. This can be improved by chemical modification such as phosphothioates,[12] locked nucleic acids (LNAs),[13-14] peptide nucleic acids (PNA)[15-16], unlocked nucleic acids (UNA)[17-18] or others.[19] The specificity of target binding can be improved by the incorporation of modifications such as LNA and UNA.[14, 17] In addition, peptides are used to enhance delivery.[20] In particular, cell-penetrating peptides with the ability to cross various epithelia and even the blood-brain barrier have shown efficient intracellular delivery.[21-22]

Simply mixed oligonucleotide-transfection reagents/peptides are applied as an economical gene therapy method.[23] However, the interaction between them is not strong, and binding specificity is

not improved.[24] There is growing evidence that by cross-linking a peptide with modified therapeutic oligonucleotide, all the aforementioned requirements to NATs can be met.[25] Chemical conjugation of peptides to oligonucleotides can be performed internally, using modified nucleotide scaffolds, as well as terminally. So far, peptide-oligonucleotide conjugates (POCs) have been most often prepared by terminal conjugation of oligonucleotides.[12] In a recent work a successful representative of internally labelled POCs was synthesized. Peptides were attached to the internal positions of oligonucleotides via a 2'-alkyne-2'-amino-LNA scaffold, and these bioconjugates showed highly improved biomedical properties compared with the unlabeled oligonucleotides.[14] In this article, we present novel POCs that target the *BRAF* V600E oncogene and studied their fundamental properties *in vitro*. The design and synthesis of the oligonucleotides and the peptides, their attachment by CuAAC "click" chemistry, spectroscopic characterization and investigation of the stability of the resulting POCs to enzymatic degradation are reported. In the present work, we aimed to investigate the influence of different peptide sequences on the structure and properties of synthetic oligonucleotides.

First, we designed three 21-mer oligonucleotide sequences, one of them pure DNA, **ON1**, and two DNA/LNA, **ON2** and **ON3;** see **Fig. 1** and SI, Table S1. The oligonucleotides **ON2** and **ON3** contain three sequential LNA monomers. It has been shown that LNAs incorporated into the oligonucleotide sequence improve the binding affinity to complementary target sequences, specificity and enzymatic stability.[13, 26-27] On the other hand, according to the literature, a 21mer probe and a GC content 43% are sufficient for unique interaction with the target.[28] The commercially available phosphoramidite **x** (SI, Chart S1), which contains alkyne functionality, was incorporated into the oligonucleotide sequences. The distance between the two modifications is eleven nucleosides to avoid the issue of steric hindrance between the peptides.[14] While **ON1** and **ON2** target the mutation *BRAF* V600E, **ON3** targets the wild type gene. Afterwards, ten peptide sequences were designed to contain none, one or more positively charged lysine amino acids. The optimal position for the lysine(s) to interaction with the oligonucleotide backbone was investigated by systematically varying this position in the peptide. The peptides were modified with an azide functionality at the N-termini to enable labelling of the ON via the CuAAC "click" reaction. This, resulting in peptides **PEP1-PEP10** (Table 1). Even though the arginine containing peptides have good transfection properties, the lysine residues were the first choice due to their good binding to oligonucleotides.[29] The length of the peptides is limited to 8 amino acids in order to prevent potential toxicity.[30] It has been reported that the sequence of conjugated peptides can affect the duplex stability of oligonucleotides. Thus, when a positively charged amino acid is present in the peptide sequence, the thermal stability of the oligonucleotide is increased.[27, 31] Taking this into account, we additionally varied the peptide sequence to find the optimal amino acid content and positions to improve the properties of POCs.

Next, the ten peptides were attached to the three oligonucleotides, each of them twice in each oligonucleotide sequence. This resulted in thirty different POCs, **POC1-POC30** (**Fig. 1**, Table 1). Conjugation of **ON1-ON3** with the peptides **PEP1-PEP10** was performed using CuAAC "click" chemistry. Microwave irradiation was used to improve the yield. The samples were heated for 15 min to 60 °C, followed by 24 h at room temperature and protected from light.[14, 32] The resulting reaction mixtures were purified on an Illustra NAP-5 column, giving the desired products **POC1-POC30** in high purity and yields of 43-83%. We observed that the yields of conjugation were not influenced by the peptide sequence. Nevertheless, POCs with higher yields were obtained when using the pure DNA oligonucleotide sequence. The identity and purity of the POCs were confirmed by ion-exchange HPLC and MALDI-TOF mass spectrometry (SI, Table S2/S3, Fig. S1/S2).

Table 1. POC composition and yields of synthesis

Pep nr	Peptide structure	Oligonucleotide nr/ POC product (yield, %)		
		ON1	ON2	ON3
1	$N_3(CH_2)_5C(O)$YKGAAGGA-NH_2	POC1 (49.4%)	POC11 (47.6%)	POC21 (44.2%)
2	$N_3(CH_2)_5C(O)$YGGKAGGA-NH_2	POC2 (75.3%)	POC12 (57.2%)	POC22 (48.8%)
3	$N_3(CH_2)_5C(O)$YGGAAKGA-NH_2	POC3 (69.1%)	POC13 (52.3%)	POC23 (44.2%)
4	$N_3(CH_2)_5C(O)$YGGAAGGK-NH_2	POC4 (69.1%)	POC14 (42.3%)	POC24 (62.8%)
5	$N_3(CH_2)_5C(O)$YKKAAGGA-NH_2	POC5 (64.3%)	POC15 (64.3%)	POC25 (74.4%)
6	$N_3(CH_2)_5C(O)$YKGAKGGA-NH_2	POC6 (83.3%)	POC16 (64.3%)	POC26 (53.5%)
7	$N_3(CH_2)_5C(O)$YKGAAGGK-NH_2	POC7 (66.6%)	POC17 (52.4%)	POC27 (44.2%)
8	$N_3(CH_2)_5C(O)$YKGAKGGK-NH_2	POC8 (42.9%)	POC18 (78.3%)	POC28 (65.1%)
9	$N_3(CH_2)_5C(O)$YGGAAGGA-NH_2	POC9 (66.6%)	POC19 (52.4%)	POC29 (58.2%)
10	$N_3(CH_2)_5C(O)$GGAAGGAY-NH_2	POC10 (54.8%)	POC20 (45.3%)	POC30 (58.2%)

* K is lysine - a positively charged amino acid; A, G and Y are alanine, glycine and tyrosine respectively. The yield of the POCs was calculated measuring the absorbance at 260 nm and comparing the amount of POC product to the corresponding starting oligonucleotide.

The hybridization properties of the POCs towards complementary and mismatched DNA/RNA were studied by the use of thermal denaturation (T_m) and circular dichroism (CD) analyses. All experiments were performed in a phosphate buffered saline medium (pH 7.4), using unmodified DNA references and precursors **ON1-3** (SI, Table S1). The POC duplexes showed CD profiles of an intermediate A/B geometry, which is characteristic

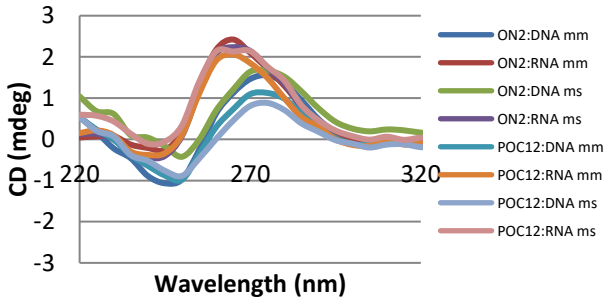


Figure 2. Representative CD spectra of POCs and unlabeled oligonucleotide duplexes. The spectra were recorded in a medium salt buffer using 2.0 μM concentration of complementary strands; mm and ms are matching and mismatching strand, respectively. **ON2:** 5'-CGA GAX TTC^{L}T^{L}C^{L}T GTA GCX AGA-3'; **POC12:** 5'-CGA GAX TTC^{L}T^{L}C^{L}T GTA GCX AGA-3' + 2x N_3-$(CH_2)_5$C(O)YGGKAGGA-NH_2

of unlabeled LNA/DNA oligonucleotide duplexes (**Fig. 2** and SI, Fig. S3).[14] Compared with the unmodified reference duplexes, the POCs displayed stronger peak intensities. According to the literature,[14] this confirms the successful adoption of peptides within the duplexes and excludes the possibility of structural changes to the oligonucleotides that contain LNA and the oligonucleotides conjugated with peptides. Furthermore, all POCs displayed *S*-shaped melting curves that are similar to the T_m curves of the unmodified reference (SI, Fig. S4). In agreement with the literature on LNA/DNA, the duplexes of the prepared POCs and the three oligonucleotides **ON1**, **ON2** and **ON3** showed higher T_m values when bound to the matching RNA compared to DNA (SI, Table S4/Fig. S5).[14, 33]

We investigated the influence of peptides on target binding to POCs by comparison of their T_m values to those of precursor strands (SI, Table S1). The best stability (T_m = 71.3 ºC) was shown by **POC15** bound to complementary RNA. However, when bound to complementary DNA, **POC12**, **POC13**, **POC25**, **POC28** and **POC29** displayed the highest stabilities (**Fig. 3** and SI, Fig. S5-S6; T_m = 60.5 ºC, 61.0 ºC, 60.9 ºC, 61.1 ºC and 61.3 ºC, respectively). When analyzing grouped T_m data and the difference between fully matched and mismatched binding (ΔT_m), it is also clear that the attachment of peptides affects RNA target binding more than the interaction with DNA (**Fig. 3**). The difference in T_m data for each target is statistically significant with $p < 0.001$.[34]

We further rationalized the T_m using descriptive statistics and linear regression.[35] First, we compared T_m values for POCs and their DNA and LNA/DNA analogues. As mentioned above and in the literature,[30] RNA binding is more stable than DNA, and for POCs this effect is more pronounced then for reference DNA and LNA/DNA (for example, median T_m values 58.6 °C and 59.5 °C for DNA and RNA complexes of reference probes compared to 58.4 °C and 63.9 °C for POCs). Looking at median values for the T_m of peptide-free vs peptide containing probes (SI, Table S7), discrimination of RNA and DNA mismatches is similar, and in both cases it is improved by the attachment of the peptides (Table S7, increase in median ΔT_m value 3.1 °C for binding mismatched RNA by POCs vs peptide-free probes).

Statistical analyses also confirmed that the substitution with the peptide did alter target binding at a statistically significant level, $p < 0.05$ (SI, Table S5).[31] By using a two-tailed *t*-test with unequal variances,[36-37] the effect of each peptide on the T_m of the three different oligonucleotides was evaluated (SI, Table S6). Based on this, **PEP10 (**p=0.4) had the highest and **PEP5 (**p=0.9**)** the lowest influence on T_m among all the peptides Furthermore, when comparing the two neutral peptide sequences, **PEP9 (**p=0.6**)** showed superior stabilization compared to **PEP10 (**p=0.4**)**, most likely due to the presence of tyrosine near the oligonucleotide sequence.[33] Notably, all of the LNA/DNA POCs of **ON2** and **ON3** demonstrated higher mismatch discrimination for both DNA and RNA compared with the LNA-free POCs of **ON1** (SI, Chart 2). LNA helps the mismatch discrimination for RNA, while the effects for DNA are moderate. This is in good agreement with previous data on the ability of LNA to stabilize the interaction with target RNA.[34]

Another way to analyze the influence of peptide on the binding of target by POC is to plot the difference between highest and lowest T_m obtained for each particular oligonucleotide when different peptides are being attached (data dispersity analysis shown in **Fig. 4**). Interestingly, when using LNA in ON2 and ON3, the peptide sequence had less effect on the dispersity in binding affinity than for LNA-free strand ON1, while the overall dispersity values were higher in the presence of LNA. This means that there is a bigger variation in the T_m data for LNA containing POCs then for LNA free analogue, but at the same time with a smaller difference among the peptides in LNA-rich POCs

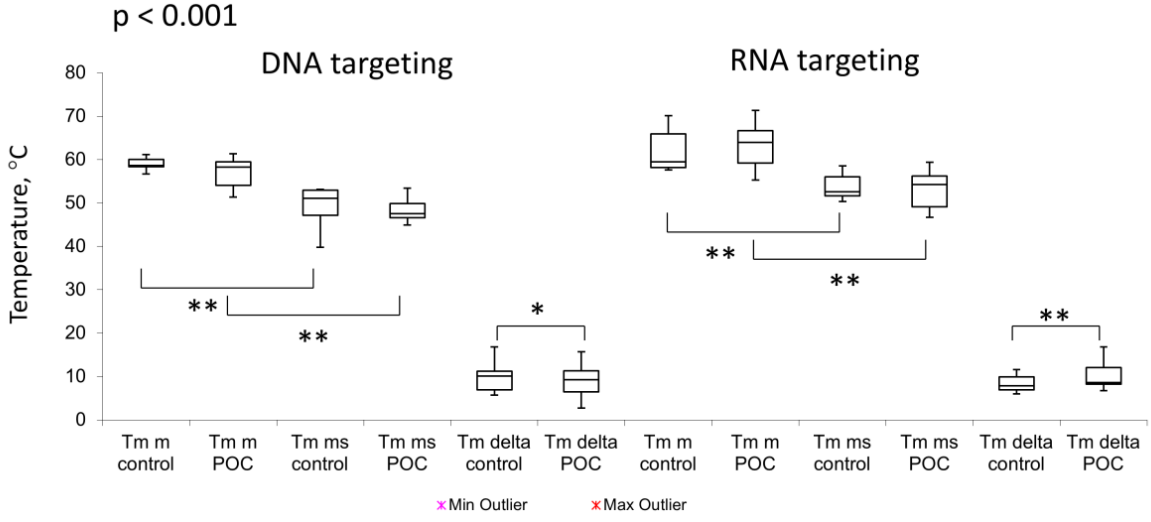


Figure 3. Box-and-whisker plot with outliers. The arms on each boxplot are values Q1 – 1,5×IQR and Q3 + 1,5×IQR. Data points for each subject are means for two independent measurements (SI, Table S13). DNA/RNA targets: m = matched; ms = mismatched. ΔT_m is the difference between T_m for fully matched and mismatched duplexes. Results of t-test assuming unequal variances: the difference in means is statistically significant at 90% (*) or 95% (**) significance level.

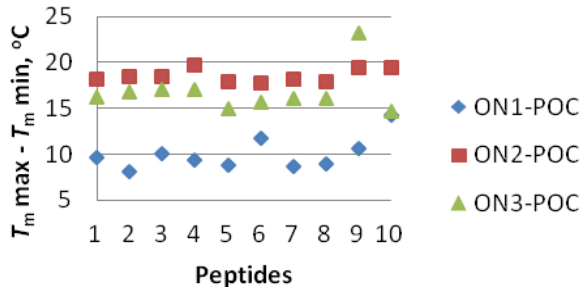


Figure 4. Data dispersity analyses of the displayed T_m values of the POCs. The data was analyzed taking the difference of T_m max and T_m min (when bound with DNA and RNA match and mismatch) plotted for all peptides and the three oligonucleotides.

Next, we studied sensitivity of POCs to mismatched targets. Among the POCs, **POC12** and **POC13** showed superior mismatch discrimination when binding RNA targets, with ΔT_m values of 16.7 °C and 16.8 °C, respectively. Consequently, these two POCs can be considered as the most specific candidates for targeting RNA. Mean ΔT_m value for ON2POC-DNA complexes is 8.8 ° C, and for ON3POC-DNA is 11.8 ° C. However, when ON2POC and ON3POC bind to RNA, the situation changes. Mean ΔT_m values are 12.5 ° C for ON2POC-RNA and 8.5 ° C for ON3POC-RNA (SI, Chart S2). This implies that +A-rA mismatch of ON3POC with RNA target is stabilized compared to +T-rU mismatch in ON2POC. One possible reason for the stabilization of +A-rA mismatch in POC-RNA duplexes is a formation of stable secondary structure. According to the literature, +A often base pairs with rA by unusual stacking.[38-39] In that way, +A-rA pair makes a unique surface that could stabilize mismatched duplexes and in that way increase ΔT_m value of ON3-POC:RNA series.

Our next goal was to better understand the correlations between structural features of the peptides and the binding by POCs of the complementary and mismatched targets. To address this, we defined a set of variables for POCs that takes into account structural aspects of the attached peptides and correlated them with T_m values for fully matched and mismatched DNA and RNA targets. This was done using linear regression in R software. Analyzed variables were: number of lysine residues (0-3), number of amino acids between the oligonucleotide and lysine (0-7), number of glycine and tyrosine residues (0-4 and 0-1), and number of nucleic acids between the oligonucleotide and tyrosine (0-7). Resulting data is presented in SI, Tables S8-S12. Based on the analysis, the number of lysine amino acids in the peptide has a stronger effect on DNA then RNA binding. Moreover, this effect is stronger on mismatched then matched interaction with F values of 2.77 and 0.97 and p values 0.08 and 0.12 for mismatched DNA and RNA, respectively.

Confidence intervals at 90% significance level are similar for the correlations of T_m points of matched and mismatched duplexes, and higher for the binding of RNA targets compared to DNA. Nevertheless, when analyzing the change in binding, ΔT_m, confidence intervals for the correlations are similar for DNA and RNA binding: 1.91 and 1.54, respectively. As to the distance between the oligonucleotide and lysine residue, it has only a minor effect on the T_m values but affects the discrimination of mismatched RNA (ΔT_m for RNA, F value 0.93). RNA target discrimination is also affected by the number of glycine residues in the peptide (SI Table S10; F value 1.79). Interestingly, the number of glycine amino acids also positively correlated with T_m values of the probes with DNA but not RNA targets (SI Table S10; F values 1.25 and 2.63 vs 0.06 and 0.26 for the complexes with mmDNA, msDNA vs. mmRNA, msRNA, respectively). Confidence intervals at 90% significance levels are rather similar for the correlated variables and lie within region 1.1-2.3.

We studied the influence of the presence of tyrosine and number of amino acids between it and the oligonucleotide on the T_m data (SI, Tables S11,S12). As can be seen tyrosine has a positive effect on T_m of POCs vs DNA match target and on the discrimination of DNA and RNA targets (F values 0.75-1.65). This is accompanied by high confidence intervals for the correlations at 90% significance level (up to 4.98). The number of amino acids to the tyrosine correlates with T_m of match and mismatch DNA and less with RNA. However, this parameter has weak correlation with ΔT_m values.

To evaluate the potential of using POC:DNA duplexes as therapeutics, we performed competitive T_m experiments. First, POC12 was annealed with DNA target. Next, the product duplex was incubated with 1 eq and 2 eq RNA target for 3h at 37 ° C. The conditions of competitive experiments were adapted from incubation protocols of synthetic oligonucleotide in cell culture and *in vivo*, and also mimic physiological conditions.[40] Resulting mixtures were analyzed by UV melting experiments (SI, Fig. S7). The measured T_m values were 67.9 °C and 68.2 °C for POC12:DNA+1eq RNA and POC12:DNA+2eq RNA respectively. Comparing these values with the T_m values of POC12:DNA and POC12:RNA, 60.5 °C and 70.5 °C, we confirm the preference of POC12 for RNA target over DNA target strand.

Serum stability studies of the POCs were performed on a few selected POCs to evaluate their potential for *in vivo* applications. The best LNA/DNA (**ON2)**-POCs and representative LNA-free (**ON1)** POCs were 5'-end labeled with ^{32}P and incubated in 90% human serum (HS) for 24 h. The oligonucleotides **ON2** and **ON1** were used as references. Samples were withdrawn at various time points and analyzed on 13% denaturing polyacrylamide gels (7 M urea, 1x TBE). The degradation was visualized by autoradiography. Notably, **ON2** was degraded within 1 h in HS. However, the partial products of degradation showed stability for more than 8 h. The POCs of **ON2** demonstrated significant resistance to degradation after 24 h in 90% HS. Furthermore, while **ON1** was degraded in 20 min, the POCs of **ON1** showed stability for 8 h; see **Fig. 5** (SI, Fig. S8). In agreement with previous work, sequential incorporation of three LNA monomers in the oligonucleotide sequence improved the serum stability of the oligonucleotide.[13-14] These results are comparable with a previously reported study on the 24 h serum stability of Tat peptide-oligonucleotides with phosphorothioate backbone conjugates.[12] Moreover, the POCs of DNA/LNA oligonucleotides showed significantly higher resistance to degradation in HS compared with the POCs of LNA-free oligonucleotides. However, the POCs of the LNA-free oligonucleotides showed significant stability compared with unlabeled oligonucleotides. This implies that the peptides attached to the oligonucleotide has a significant positive impact on serum stability. Finally, the combination of an internally incorporated LNA monomer into the oligonucleotide sequence and the attachment of two peptide residues had an additive effect on improving the stability of oligonucleotides in HS.[14]

As a final aspect, the stability of **ON1** and two representative POCs duplexes with a DNA complementary strand was investigated in 90% HS. Higher stability of ds vs ss oligonucleotides is already being utilized in RNAi[41-42]. We hypothesized that similar effects can benefit the stability of antisense oligonucleotides and potentially can be used in delivery. The duplexes were 5'-labeled with ^{32}P and incubated in HS for 160 h. Samples were withdrawn at various time points and analyzed as

described above; see **Fig. 5** (SI, Fig. S9). As evident from SI, Fig. S9, the duplexes showed stability in 90% HS for up to 160 h. These results are superior to previously reported results regarding the stability of duplexes of modified oligonucleotides in complex biological media.[43-44]

Notably, the duplex of unmodified DNA showed high stability in 90% HS as well (160 h). Nevertheless, the peptide-conjugated antisense strands have an advantage of improved target recognition and therefore have higher potential as therapeutics.[1,11b]

In conclusion, we rationally designed diverse POCs and investigated how the peptide sequence influences the stability of the oligonucleotide. While all POCs bound the target more strongly and more selectively compared to the unconjugated oligonucleotide and the unmodified DNA reference, some of them showed greater stabilization than others.

The most important finding of this work is that the amino acid composition of covalently linked POCs has an effect on the target recognition. Lysine residues are known to improve binding of POCs to target DNA and RNA. Based on our data and statistical analyses the number of lysines affects binding to DNA and RNA mismatched targets, whereas the distance to lysine correlates exclusively with the decrease in T_m for RNA mismatch compared to complementary RNA binding. In addition, we observe that other amino acids in the POC (glycine, tyrosine) affect the T_m and mismatch discrimination as well. In particular, T_m values with match and mismatch DNA and the decrease in T_m for RNA recognition are affected.

The POCs were stable in 90% human serum for up to 24 h. The duplexes of the POCs with complementary DNA were stable for up to 160 h. The best POC candidates described herein have considerable potential for selective and efficient binding to the *BRAF* V600E oncogene *in vivo*. Consequently, the DNA based POCs can be used in further studies to investigate the cellular uptake and the delivery *in vivo*.

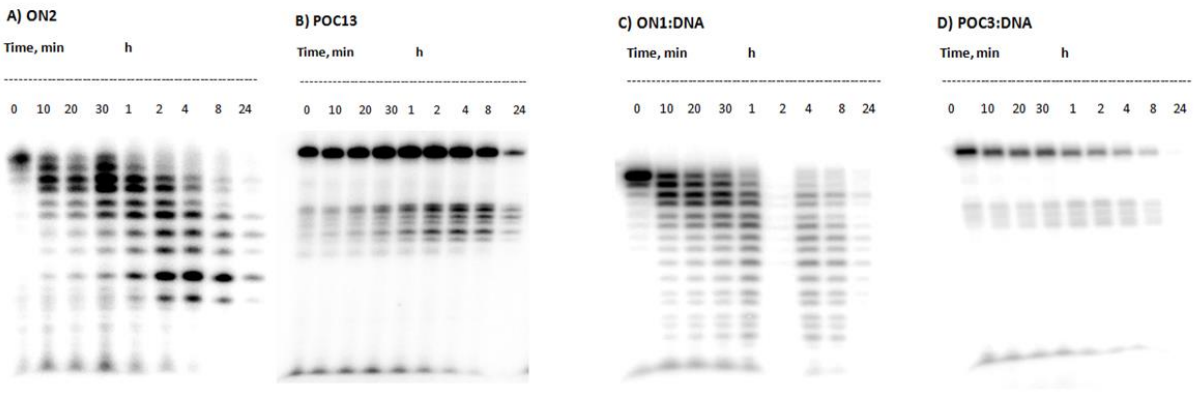


Figure 5. Gel electrophoresis of the 5'-^{32}P labeled oligonucleotides, POCs and their duplexes after incubation in 90% human serum.

Supporting Information

Experimental details, design, synthesis and analyses of the oligonucleotides and the POCs, CD and T_m experiments and results, HS experiments, gel electrophoresis details and statistical analyses details are included in the supporting information.

AUTHOR INFORMATION

Corresponding Author

* M. Taskova and K. Astakhova. E-mail: taskova@sdu.dk; ias@sdu.dk

Author Contributions

The manuscript was written through contributions of all authors. / All authors have given approval to the final version of the manuscript.

Funding Sources

This work has been supported by the Young Investigator Programme Grant from Villum Foundation, Denmark (grant no. 90-7315).

ACKNOWLEDGMENT

This work was supported by the Villum Fonden and Lumiprobe LLC.

ABBREVIATIONS

NAT, nucleic acid therapy/therapeutics; ON, oligonucleotide; PEP, peptide; POC, peptide-oligonucleotide conjugate; T_m, melting themperature; CD, circular dichroism.

REFERENCES

1. Kay, M. A. (2011) State-of-the-art gene-based therapies: the road ahead. *Nat. Rev. Genet. 12* (5), 316-328.
2. Greenberg, D. S., Soreq, H. (2014) MicroRNA therapeutics in neurological disease. *Curr. Pharm. Des. 20* (38), 6022-6027.
3. Koli, U., Krishnan, R. A., Pofali, P., Jain, R., Dandekar, P. (2014) SiRNA-based therapies for pulmonary diseases. *J. Biomed. Nanotechnol. 10* (9), 1953-1997.
4. Gebbing, M., Bergmann, T., Schulz, E., Ehrhardt, A. (2015) Gene therapeutic approaches to inhibit hepatitis B virus replication. *World J. Hepatol. 7* (2), 150-164.
5. Landry, B., Valencia-Serna, J., Gul-Uludag, H., Jiang, X., Janowska-Wieczorek, A., Brandwein, J., Uludag, H. (2015) Progress in RNAi-mediated Molecular Therapy of Acute and Chronic Myeloid Leukemia. *Mol. Ther. Nucleic Acids 4*, e240.
6. Kaboli, P. J., Rahmat, A., Ismail, P., Ling, K. H. (2015) MicroRNA-based therapy and breast cancer: A comprehensive review of novel therapeutic strategies from diagnosis to treatment. *Pharmacol. Res. 97*, 104-121.

7. Chen, Y., Stamatoyannopoulos, G., Song, C. Z. (2003) Down-regulation of CXCR4 by inducible small interfering RNA inhibits breast cancer cell invasion in vitro. *Cancer Res. 63* (16), 4801-4804.
8. Gulec, S. A. (2016) Y-90 Radiomicrosphere Therapy for Colorectal Cancer Liver Metastases. *Semin. Nucl. Med. 46* (2), 126-134.
9. Coleman, R. (2016) Treatment of Metastatic Bone Disease and the Emerging Role of Radium-223. *Semin. Nucl. Med. 46* (2), 99-104.
10. Zarei, M., Shivanandappa, T. (2016) Neuroprotective effect of Decalepis hamiltonii on cyclophosphamide-induced oxidative stress in the mouse brain. *Journal of basic and clinical physiology and pharmacology* 27(4), 341-348.
11. Wengel, J., Vester, B., Lundberg, L. B., Douthwaite, S., Sorensen, M. D., Babu, B. R., Gait, M. J., Arzumanov, A., Petersen, M., Nielsen, J. T. (2003) LNA and alpha-L-LNA: towards therapeutic applications. *Nucleosides, nucleotides & nucleic acids 22* (5-8), 601-604.
12. Maegawa, Y., Mochizuki, S., Miyamoto, N., Sakurai, K. (2016) Gene silencing using a conjugate comprising Tat peptide and antisense oligonucleotide with phosphorothioate backbones. *Bioorg. Med. Chem. Lett.* 26(4), 1276-1278.
13. Crouzier, L., Dubois, C., Wengel, J., Veedu, R. N. (2012) Cleavage and protection of locked nucleic acid-modified DNA by restriction endonucleases. *Bioorg. Med. Chem. Lett. 22* (14), 4836-4838.
14. Astakhova, I. K., Hansen, L. H., Vester, B., Wengel, J. (2013) Peptide-LNA oligonucleotide conjugates. *Org. Biomol. Chem. 11* (25), 4240-4249.
15. Accetta, A., Petrovic, A. G., Marchelli, R., Berova, N., Corradini, R. (2015) Structural Studies on Porphyrin-PNA Conjugates in Parallel PNA:PNA Duplexes: Effect of Stacking Interactions on Helicity. *Chirality 27* (12), 864-874.
16. Igloi, G. L. (1998) Variability in the stability of DNA-peptide nucleic acid (PNA) single-base mismatched duplexes: real-time hybridization during affinity electrophoresis in PNA-containing gels. *Proc. Natl. Acad. Sci. U S A 95* (15), 8562-8567.
17. Pasternak, A., Wengel, J. (2010) Thermodynamics of RNA duplexes modified with unlocked nucleic acid nucleotides. *Nucleic. Acids. Res. 38* (19), 6697-6706.
18. Perlikova, P., Karlsen, K. K., Pedersen, E. B., Wengel, J. (2014) Unlocked nucleic acids with a pyrene-modified uracil: synthesis, hybridization studies, fluorescent properties and i-motif stability. *ChemBioChem 15* (1), 146-156.
19. Prakash, T. P., Manoharan, M., Fraser, A. S., Kawasaki, A. M., Lesnik, E. A., Owens, S. R. (2000) Zwitterionic oligonucleotides with 2′-O-[3-(N,N-dimethylamino)propyl]-RNA modification: synthesis and properties. *Tetrahedron Lett. 41* (25), 4855-4859.
20. Zanta, M. A., Belguise-Valladier, P., Behr, J. P. (1999) Gene delivery: a single nuclear localization signal peptide is sufficient to carry DNA to the cell nucleus. *Proc. Natl. Acad. Sci. U S A 96* (1), 91-96.
21. Kristensen, M., Birch, D., Morck Nielsen, H. (2016) Applications and Challenges for Use of Cell-Penetrating Peptides as Delivery Vectors for Peptide and Protein Cargos. *Int. J. Mol. Sci. 17* (2), 185-202.
22. Liu, H., Zeng, F., Zhang, M., Huang, F., Wang, J., Guo, J., Liu, C., Wang, H. (2016) Emerging landscape of cell penetrating peptide in reprogramming and gene editing. *J. Control Release* 226, 124-137.
23. Berens, C., Porschke, D. (2013) Recognition of operator DNA by Tet repressor. *The journal of physical chemistry. B 117* (6), 1880-1885.
24. Porschke, D. (1978) Thermodynamic and kinetic parameters of oligonucleotide--oligopeptide interactions. Specificity of arginine . inosine association. *European journal of biochemistry / FEBS 86* (1), 291-299.
25. Williams, B. A., Chaput, J. C. (2010) Synthesis of peptide-oligonucleotide conjugates using a heterobifunctional crosslinker. *Current protocols in nucleic acid chemistry / edited by Serge L. Beaucage ... [et al.] Chapter 4*, Unit4.41.
26. Singh, S. K., Kumar, R., Wengel, J. (1998) Synthesis of Novel Bicyclo[2.2.1] Ribonucleosides: 2'-Amino- and 2'-Thio-LNA Monomeric Nucleosides. *The Journal of organic chemistry 63* (18), 6078-6079.
27. Johannsen, M. W., Crispino, L., Wamberg, M. C., Kalra, N., Wengel, J. (2011) Amino acids attached to 2'-amino-LNA: synthesis and excellent duplex stability. *Org. Biomol. Chem. 9* (1), 243-252.
28. Yuan, B., Latek, R., Hossbach, M., Tuschl, T., Lewitter, F. (2004) siRNA Selection Server: an automated siRNA oligonucleotide prediction server. *Nucleic Acids Res. 32* (Web Server issue), W130-134.
29. Fedoreyeva, L. I., Kireev, II, Khavinson, V., Vanyushin, B. F. (2011) Penetration of short fluorescence-labeled peptides into the nucleus in HeLa cells and in vitro specific interaction of the peptides with deoxyribooligonucleotides and DNA. *Biochemistry. Biokhimiia 76* (11), 1210-1219.
30. Ramesh, S., Govender, T., Kruger, H. G., de la Torre, B. G., Albericio, F. (2016) Short AntiMicrobial Peptides (SAMPs) as a class of extraordinary promising therapeutic agents. *Journal of peptide science : an official publication of the European Peptide Society 22* (7), 438-451.
31. Noir, R., Kotera, M., Pons, B., Remy, J. S., Behr, J. P. (2008) Oligonucleotide-oligospermine conjugates (zip nucleic acids): a convenient means of finely tuning hybridization temperatures. *J. Am. Chem. Soc. 130* (40), 13500-13505.
32. Okholm, A., Kjems, J., Astakhova, K. (2014) Fluorescence detection of natural RNA using rationally designed "clickable" oligonucleotide probes. *RSC Advances 4* (86), 45653-45656.
33. Andersen, N. K., Dossing, H.; Jensen, F., Vester, B., Nielsen, P. (2011) Duplex and triplex formation of mixed pyrimidine oligonucleotides with stacking of phenyl-triazole moieties in the major groove. *The Journal of organic chemistry 76* (15), 6177-6187.
34. Harrison, J. G., Balasubramanian, S. (1998) Synthesis and hybridization analysis of a small library of peptide-oligonucleotide conjugates. *Nucleic Acids Res. 26* (13), 3136-3145.
35. Seber, G. A. F., Lee, A. J. (2003) Linear Regression: Estimation and Distribution Theory, in *Linear Regression Analysis*, *Second Edition*, pp 35-95, John Wiley & Sons, Inc., Hoboken, NJ, USA.
36. Huang, D. W., Sherman, B. T., Lempicki, R. A. (2008) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protocols 4* (1), 44-57.
37. Parsons, D. W., Jones, S., Zhang, X., Lin, J. C.-H., Leary, R. J., Angenendt, P., Mankoo, P., Carter, H., Siu, I.-M., Gallia, G. L., Olivi, A., McLendon, R. et al. (2008) An Integrated Genomic Analysis of Human Glioblastoma Multiforme. *Science 321* (5897), 1807-1812.
38. Baeyens, K. J., De Bondt, H. L., Pardi, A., Holbrook, S. R. (1996) A curved RNA helix incorporating an

internal loop with G.A and A.A non-Watson-Crick base pairing. *Proc. Natl. Acad. Sci. U S A 93* (23), 12851-12855.

39. Owczarzy, R., You, Y., Moreira, B. G., Manthey, J. A., Huang, L., Behlke, M. A., Walder, J. A. (2004) Effects of sodium ions on DNA duplex oligomers: improved predictions of melting temperatures. *Biochemistry 43* (12), 3537-3554.

40. Zhao, X. L., Chen, B. C., Han, J. C., Wei, L., Pan, X. B. (2015) Delivery of cell-penetrating peptide-peptide nucleic acid conjugates by assembly on an oligonucleotide scaffold. *Scientific reports 5*, 17640.

41. Hemmings-Mieszczak, M., Dorn, G., Natt, F. J., Hall, J., Wishart, W. L. (2003) Independent combinatorial effect of antisense oligonucleotides and RNAi-mediated specific inhibition of the recombinant rat P2X3 receptor. *Nucleic Acids Res. 31* (8), 2117-2126.

42. Astriab-Fisher, A., Fisher, M. H., Juliano, R., Herdewijn, P. (2004) Increased uptake of antisense oligonucleotides by delivery as double stranded complexes. *Biochemical Pharmacology 68* (3), 403-407.

43. Deleavey, G. F., Watts, J. K., Alain, T., Robert, F., Kalota, A., Aishwarya, V., Pelletier, J., Gewirtz, A. M., Sonenberg, N., Damha, M. J. (2010) Synergistic effects between analogs of DNA and RNA improve the potency of siRNA-mediated gene silencing. *Nucleic Acids Res. 38* (13), 4547-4557.

44. Petrova Kruglova, N. S., Meschaninova, M. I., Venyaminova, A. G., Zenkova, M. A., Vlassov, V. V., Chernolovskaya, E. L. (2010) 2'-O-methyl-modified anti-MDR1 fork-siRNA duplexes exhibiting high nuclease resistance and prolonged silencing activity. *Oligonucleotides 20* (6), 297-308.

Supporting Information For

Antisense oligonucleotides internally labeled with peptides show improved target recognition and stability to enzymatic degradation

Maria Taskova,[1*] Charlotte S. Madsen,[2] Knud J. Jensen,[2] Lykke Haastrup Hansen,[3] Birte Vester,[3] Kira Astakhova[1]*

1 Nucleic Acid Center, Department of Physics, Chemistry and Pharmacy, University of Southern Denmark, Campusvej 55, 5230 Odense M, Denmark

2 Department of Chemistry, University of Copenhagen, Thorvaldsensvej 40, 1871 Frederiksberg, Denmark

3 Department of Biochemistry and Molecular Biology, University of Southern Denmark, Campusvej 55, 5230

Odense M, Denmark

I. General Experimental Details

All reagents and solvents were used as received. 2'-*O*-propargyl-3'-CEP uridine was supplied from Jena Bioscience. TBTA for click chemistry was obtained from Lumiprobe LLC. Stock solutions for click chemistry were prepared as described.[1] Click reactions were performed in 1 mL reactor glass tubes under argon stirring in Emrys Creator (Personal Chemistry). Human serum and Hank's buffered salt solution were purchased from Sigma-Aldrich.

Unmodified reference DNA/RNA strands were obtained from Integrated DNA Technologies and used without further purification. The oligonucleotides (**ON1**,**ON2** and **ON3**) were synthesized on an automated DNA synthesizer - PerSpective Bio-systems Expedite 8909. MALDI-TOF mass spectrometry was performed on a Ultraflex II TOF/TOF instrument, Bruker, using 3-hydroxypicolinic acid matrix (10 mg 3-hydroxypicolinic acid, 50mM ammonium citrate in 70% aqueous acetonitrile). IE HPLC was performed using a Merck Hitachi LaChrom instrument equipped with a Dionex DNAPac Pa-100 column (250mm x 4 mm). The T_m studies were performed on a DU800 UV/VIS spectrophotometer equipped with a Beckman Coulter Performance Temperature Controller. The CD spectra were recorded on a JASCO J-815 CD spectrometer equipped with CDF 4265/15 temperature controller.

Oligonucleotides **ON1**, **ON2** and **ON3** were synthesized on an automated DNA synthesizer PerSpective Bio-systems Expedite 8909. Synthesis was carried out in the DMT-ON mode, in 0.2/1.0 μmol scale on a CPG support type using the standard protocol. For incorporation of LNA monomers and the phosphoramidite (2'-*O*-propargyl-3'-CEP uridine) a hand-coupling procedure was applied. For the hand coupling procedure, they were dissolved in acetonitrile and coupled for 18 min using tetrazole as an activator. The coupling efficiencies were followed measuring the absorbance of the dimethoxytrityl cation realized after each coupling. Cleavage of the oligonucleotides from the solid support was performed with ammonia solution (28-30%) at 55 ^{o}C, overnight. The oligonucleotides were subjected to detritylation with 80% acetic acid, 30 min, followed by addition of water, sodium acetate 3M and sodium perchlorate 5M and precipitation with cold acetone (-20 ^{o}C, 10 min), the procedure was finished with washing two times with acetone. Then the identity and purity of the oligonucleotides were verified by MALDI-TOF mass spectrometry and IE HPLC, respectively.

All amino acids, HBTU, and HOBt were purchased from GL Biochem (Shanghai) Ltd. 6-azido hexanoic acid, Fmoc-Rink amide AM polystyrene resin, DIPEA, piperidine and TFA were purchased from Iris Biotech GmbH. TES and DODT were purchased from Sigma-Aldrich (Denmark) and TentaGel S RAM resins were purchased from Rapp Polymere. All chemicals were used as received and without further purification. Peptide synthesis was carried out using an automated peptide synthezier (Biotage SYRO II). Analytical HPLC was performed on a Dionex Ultimate 3000 system with Phenomenex Germini C18 (3 μm, 50 × 4.6 mm) and a linear gradient flow of (0.1% formic acid), column oven thermostated to 42 °C, connected to an ESI-MS (MSO Plus Mass Spectrometer, Dionex). Purification of the peptide was performed on a preparative Dionex Ultimate 3000 HPLC with a C18 column from FeF chemicals (Denmark) (5μm, 250 × 21.2 mm, 300Å) or Phenomenex Gemini Axia (5μm, 100 × 21.2 mm, 110Å). Unless otherwise stated (0.1 % TFA) was used as eluent with a flow of 15 ml/min. High-resolution mass spectrometry (HR-MS) was performed on a SolariX XR instrument from Bruker by direct injection with ionization in positive electrospray mode.

The peptides were synthesized using Tentagel S RAM (400 mg; 0.1 mmol; loading 0.25 mmol/g) resins as solid support that initially was swelled in CH_2Cl_2. Coupling of the amino acids (5.2 eq) and 6-azidohexanoic acid (2 eq) were utilized via standard

Fmoc solid phase peptide synthesis using HBTU (5 eq), HOAt (5.2 eq) and DIPEA (9.4 eq) as coupling reagents. Coupling time used was 2 × 1.5 h with a washing step (NMP) in between. After the coupling step the resins were washed (2 × NMP, 1 × DCM, 3 × NMP). Fmoc deprotection was carried out using 40% piperidine in DMF for 3 min followed by 20% piperidine in DMF for 2 × 15 min with a washing step (NMP) in between. Subsequently, the resins were washed (3 × NMP, 1 × DCM, 3 × NMP). The peptides were cleaved/deprotected by adding 95:2.5:2.5 TFA:TES:H_2O for 2 hours. After removing TFA under nitrogen flow, the peptides were precipitated with cold diethylether giving a white powder. The crude product was purified by preparative RP-HPLC (gradient 5-65% over 25 min) and then analyzed by ESI mass spectrometry.

MS results for the peptides, calc./found: PEP1, 832.5/832.3; PEP2, 818.4/818.5; PEP3, 832.5/832.4; PEP4, 818.4/818.5; PEP5, 903.5/903.4; PEP6, 889.5/889.6; PEP7, 889.5/889.4; PEP8, 946.5/946.4; PEP9, 761.4/761.4; PEP10, 761.4/761.4.

CuAAC "click" reactions were performed for oligonucleotides **ON1-ON3**, each labeled with two alkyne groups (2'-*O*-propargyl-3'-CEP uridine), with the peptides **PEP1-PEP10** that have azide group attached to the N-terminus. Concentrations of the oligonucleotides were calculated measuring the absorbance at 260 nm and using that nmol/OD_{260} is 4.7; 4.7; and 4.6 for **ON1**, **ON2** and **ON3** respectively. Corresponding starting oligonucleotide (20 nmol in MQ water) was mixed with triethylammonium acetate buffer (20 μL, 1 M in MQ water), amminoguanidine hydrochloride (4μL, 50 mM in MQ water), DMSO, peptide (120 nM, 12x concentration of the oligonucleotide, in DMSO), MQ water, Cu (II) TBTA (10 mM in MQ water:DMSO = 1:1) and ascorbic acid solution (50 mM in MQ water) in final volume of 200 μL, with 50% DMSO and 50% MQ water. The reaction mixture was deaerated using argon, transferred to microwave glass tube (0.2 - 0.5 mL), closed tightly, vortexed and put in a microwave reactor for 15 min, 60 °C. Next, the reaction mixture was stored at room temperature and dark place overnight. Afterwards, the reaction was filtrated using Illustra NAP-5 columns. The filtration, DNA purification by the process of gel filtration (Illustra NAP-5 column) was proceed using the protocol from the manufacturer (GE Healthcare). The resulting solution was dried under N_2 flow without heat and dissolved in MQ water. Absorbance was measured at 260nm to find out the concentration of the POCs. All the products were analyzed by MALDI-TOF mass spectrometry and IE HPLC.

Hybridization studies were performed for **POC1-POC30**, **ON1-ON3** and the two commercially obtained reference DNA sequences, **D1** and **D2**. The mixed sequences were annealed with DNA and RNA matching and mismatching strands. Complementary strands (1 μM of each strand), in a medium salt phosphate buffer (8 mM Na_2HPO_4, 2 mM NaH_2PO_4, 100 mM NaCl, 0.1 mM EDTA, pH 7.0) were mixed, denatured 10 min at 95 °C and subsequently cooled to 15 °C, the temperature on which the experiment was started. Thermal denaturation temperatures (T_m values, °C), were determined for all samples using the thermal denaturation curve (A_{260} *vs*. temperature). Reported T_m values present the maximum of the first derivative of the curve and are an average of the two measurements.

The CD studies were performed for **POC12**, **POC13**, **ON2** and **D1**. The samples were annealed with DNA and RNA matching and mismatching strands. Complementary strands (0.2 μM of each strand), in a medium salt phosphate buffer (8 mM Na_2HPO_4, 2 mM NaH_2PO_4, 100 mM NaCl, 0.1 mM EDTA, pH 7.0) were mixed, denatured 10 min at 95 °C and subsequently cooled to 15 °C, the temperature on which the experiment was started. For the measurements were used quartz optical cells with a patch-length of 0.5 cm.

The serum stability assay was made for the oligonucleotides **ON2**, **ON1**, and the POCs **POC12**, **POC13**, **POC14**, **POC15**, **POC18**, **POC3** and **POC8**. Oligonucleotides and POCs were ^{32}P labeled at the 5'-end by the following standard procedure:

1µL oligonucleotide/POC in concentration 3 pmol/µL, 1 µL Kinase buffer 10x, 3 µL ATP Gamma ^{32}P, 5µL water and 0.2 µL PNK enzyme in a PCR tube in a total volume of 10 µL. The samples were incubated in a Eppendrof Mastercycler personal of 1:30 h at 37 °C followed by 15 min at 75 °C, and then stored at -20 °C. Kinase buffer 10x was 10 M Tris HCl, pH 8.0, 10 mM $MgCl_2$ and 70 mM DTT. The labeled **ON1**, **POC3** and **POC8** were annealed with the complementary DNA strand, DT1 in Hank's buffered salt solution in total concentration of 0.1115 pmol. 0.1115 pmol radioactive oligonucleotide mixed with 4.485 pmol unlabeled oligonucleotide were added to human serum (18.4 µL, 90%, in HBSS buffer, pH = 7.4) in a PCR tube in a total volume of 23 µL. 1.6 µL, 50 mM EDTA added. The samples were incubated in an Eppendorf Mastercycler personal at 37 °C. Aliquots (2 µL in 1 µL loading dye) were withdrawn at time points: 0 min, 10 min, 20 min, 30 min, 1 h, 2 h, 4 h, 8 h, and 24 h (plus 48 h, 72 h, 96 h, and 160 h for the duplexes). The samples were then resolved on 13% standard denaturing polyacrylamide gels (7 M urea, 1x TBE) with TBE as running buffer. The gels were run for 1 hr at 13 W and dried under vacuum (1 h, 80 °C), exposed overnight and visualized by autoradiography on a Typhoon Trio Variable Mode Imager (Amersham Biosciences).

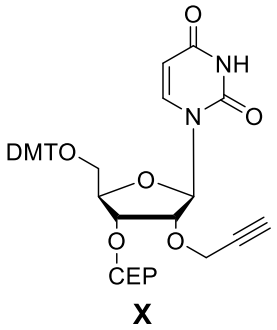
II. Oligonucleotides used in this study

Table S1. The design of oligonucleotides used in this study

Name	Sequence
ON1	5'-CGA GAX TTC TCT GTA GCX AGA-3'
ON2	5'-CGA GAX TTC^{L}T^{L}C^{L}T GTA GCX AGA-3'
ON3	5'-CGA GAX TTC ^{L}A^{L}C^{L}T GTA GCX AGA -3'
D1	5'CGA GAT TTC TCT GTA GCT AGA -3'
D2	5'CGA GAT TTC ACT GTA GCT AGA-3'
DT1	5'-TCT AGC TAC AGA GAA ATC TCG-3'
RT1	5'-UCU AGC UAC A GA GAA AUC UCG -3'
DT2	5'-TCT AGC TAC AGT GAA ATC TCG-3'
RT2	5'-UCU AGC UAC AGU GAA AUC UCG -3'

**X is 2'-O-propargyl-3'-CEP uridine (CEP phosphoramidite); C^{L}, T^{L}, and A^{L} are C-LNA, T-LNA and A-LNA monomers respectively.*

Chart S1. 2'-*O*-Propargyl-3'-CEP uridine (CEP phosphoramidite, **x)**



X

III. Analysis of oligonucleotides ON1-ON3 and representative POCs

Table S2. IE HPLC retention times and MALDI-MS of the oligonucleotides

#	IE HPLC	MALDI-MS	
	Ret. time, min	Found m/z $[M-H]^-$	Calc. m/z $[M-H]^-$
ON1	17.28	6511	6514
ON2	16.82	6629	6626
ON3	16.98	6630	6635

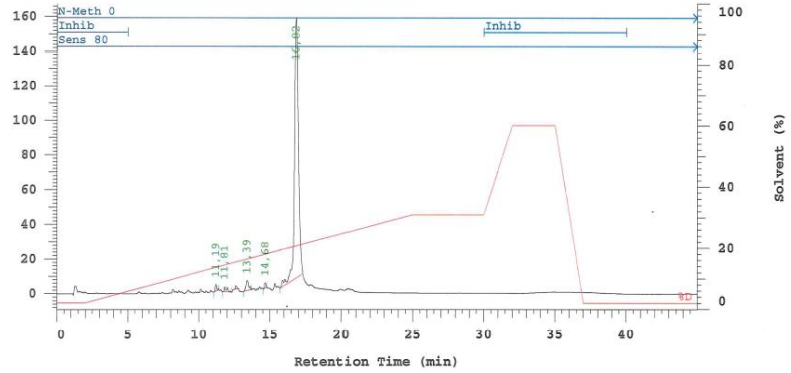
Table S3. IE HPLC retention times and MALDI-MS of the representative POCs

#	IE HPLC	MALDI-MS	
	Ret. time, min	Found m/z $[M-H]^-$	Calc. m/z $[M-H]^-$
POC3	16.89	8178	8177
POC8	15.85	8407	8406
POC14	16.26	8262	8261
POC15	15.62	8435	8431
POC24	16.61	8273	8270
POC25	15.39	8444	8440

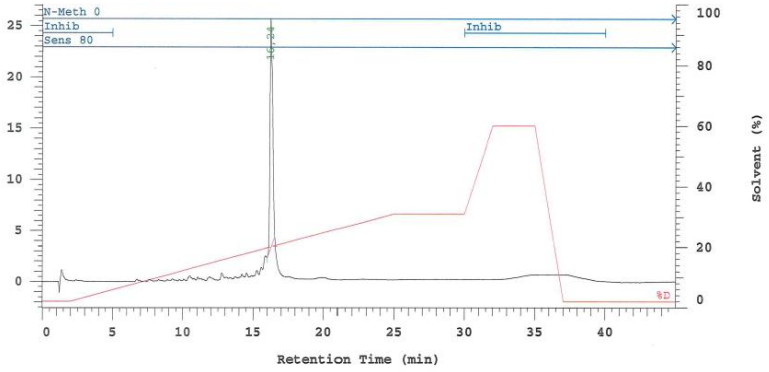
IV. IE HPLC retention times and MALDI-MS of the POCs

Figure S1. Representative IC HPLC spectra of **ON2** and the POCs

ON2



POC14



POC15

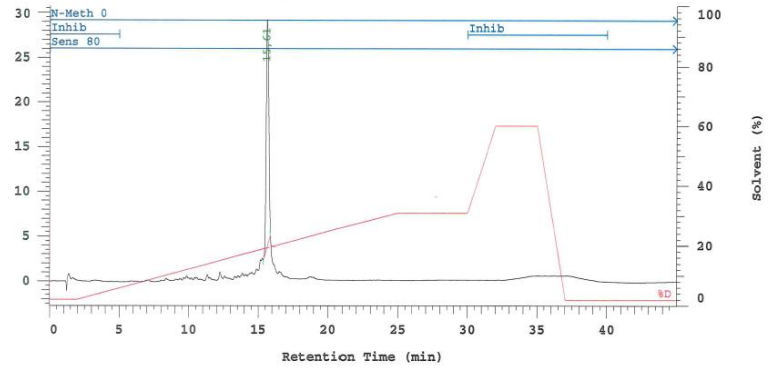
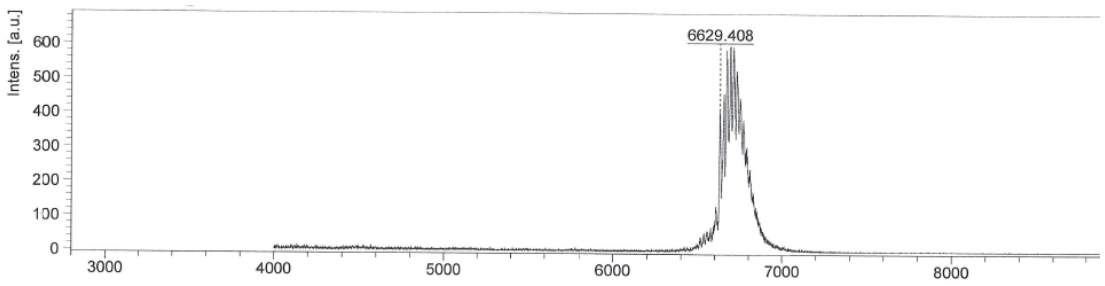
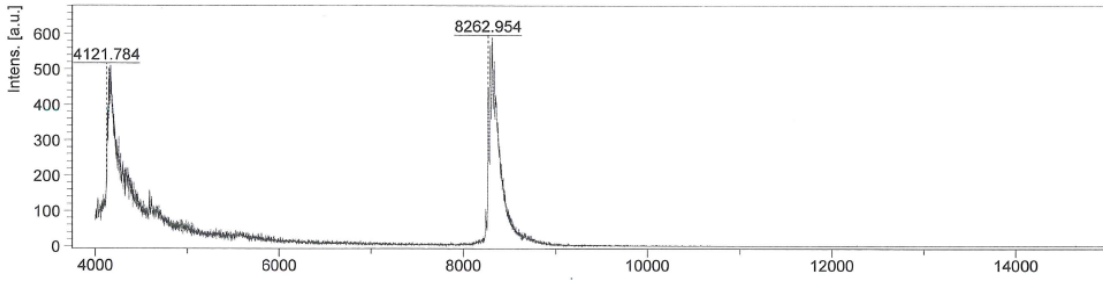


Figure S2. Representative MALDI-MS spectrum of **ON2** and the POCs

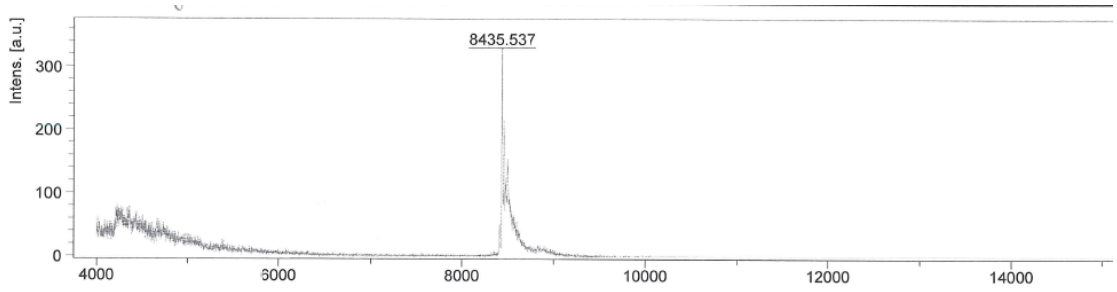
ON2



POC14



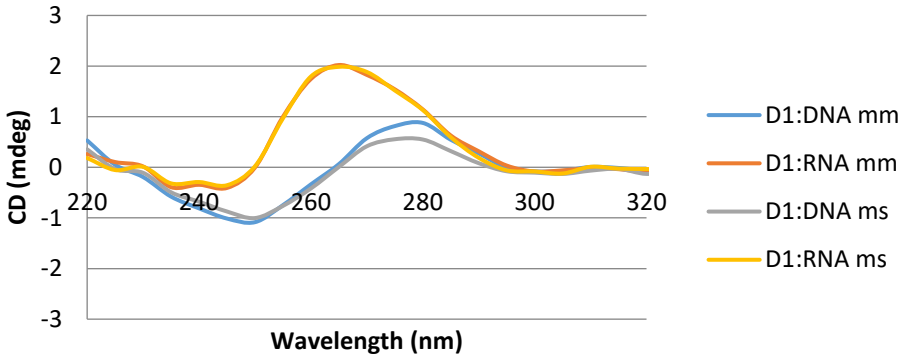
POC15



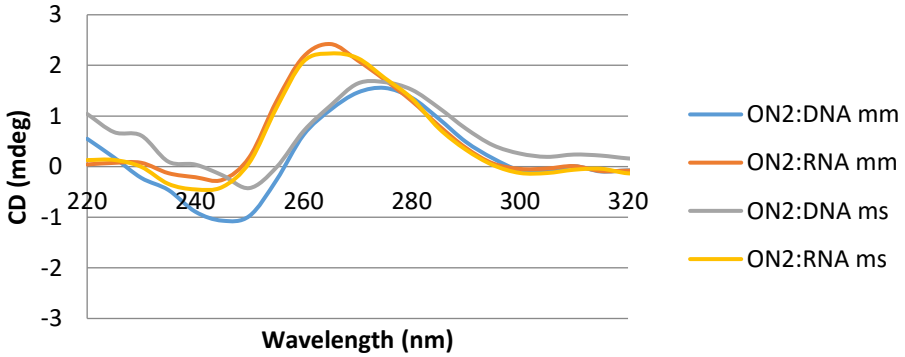
V. Representative CD curves

Figure S3. Representative CD spectra of the duplexes of **D1**, **ON2** and **POC13** with complementary DNA/RNA match and mismatch sequence

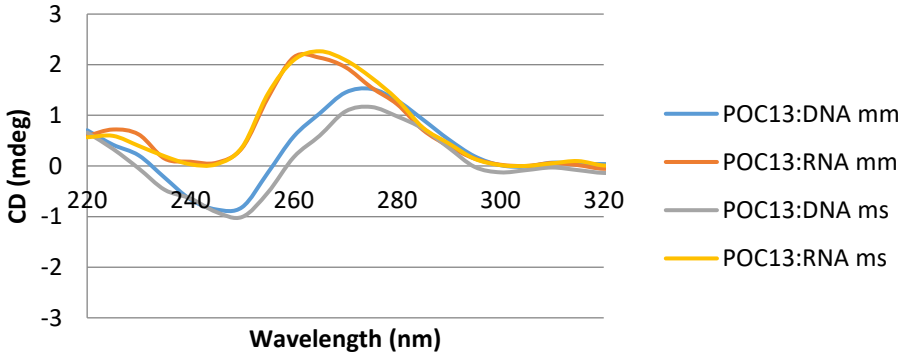
D1



ON2



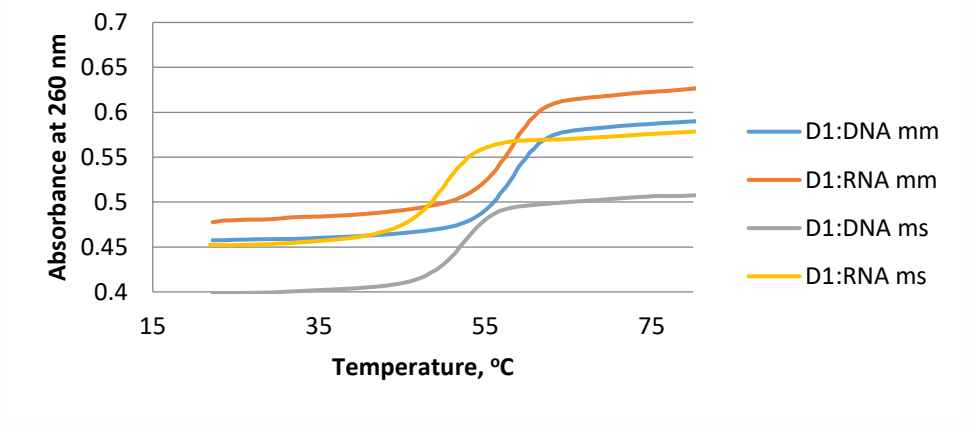
POC13



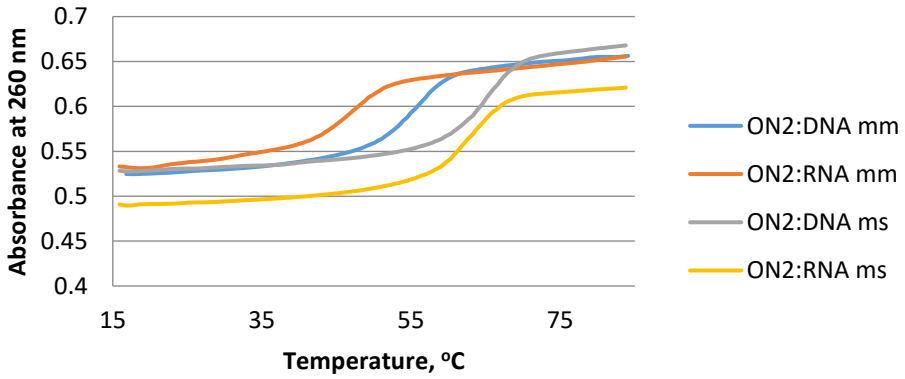
VI. Representative T_m curves, T_m results

Figure S4. Representative T_m curves

D1



ON2



POC13/POC18

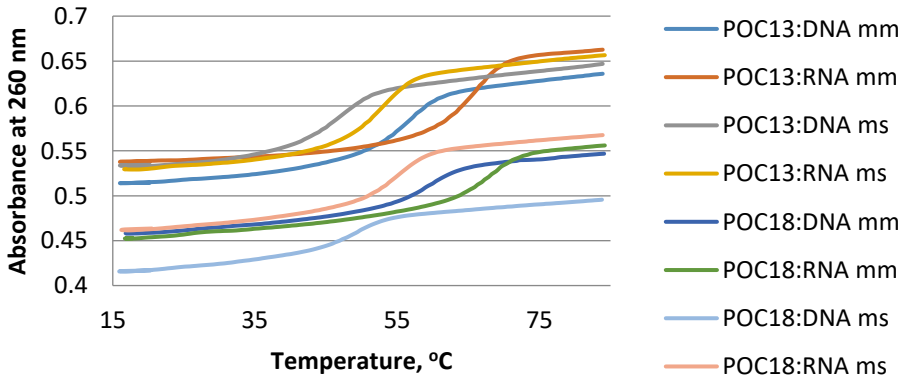
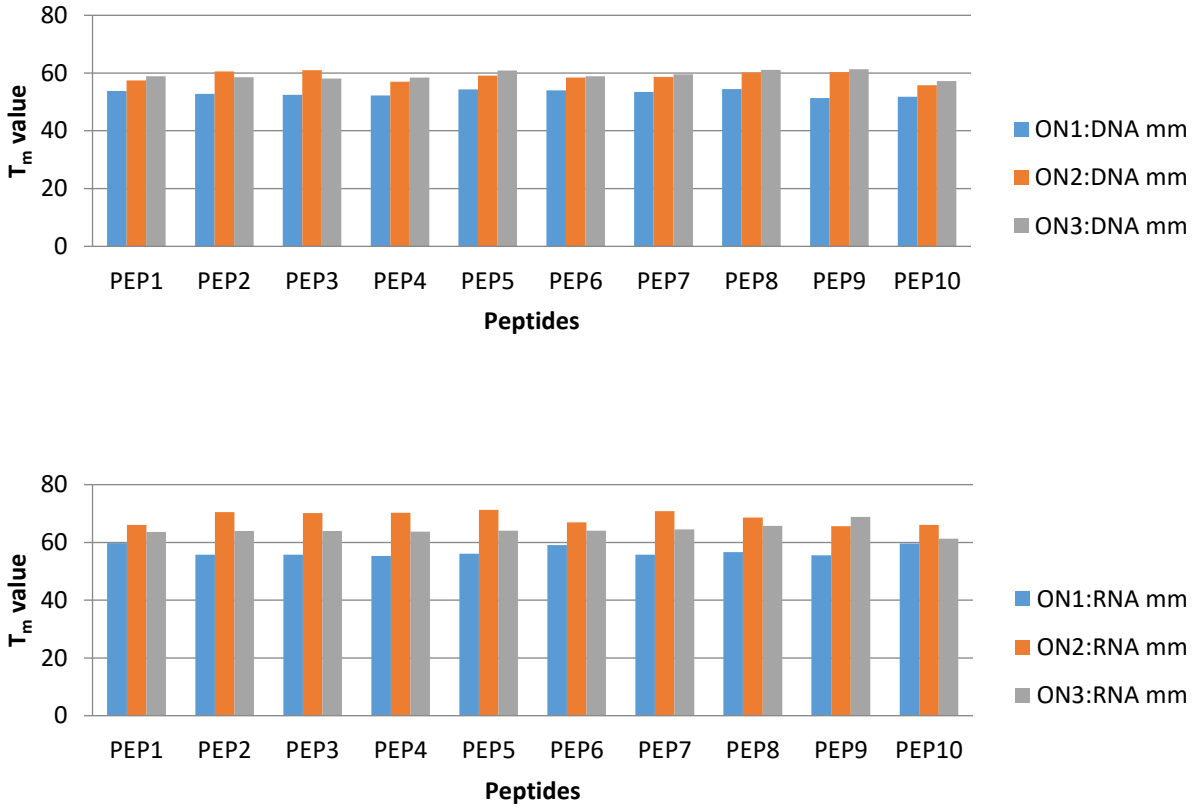


Table S4. Thermal denaturation temperatures (T_m) of the duplexes to DNA/RNA matching and mismatching strands in a medium salt phosphate buffer.

	DNA/RNA match		DNA/RNA mismatch	
Sample	DT1	RT1	DT2	RT2
D1	58.6	58.2	52.9	50.3
ON1	56.65	59.5	39.8	52.6
ON2	61.20	70.1	51.1	58.5
POC1	53.8	59.7	50.0	51.4
POC2	52.8	55.7	50.1	47.6
POC3	52.5	55.7	45.7	47.1
POC4	52.2	55.3	45.9	47.1
POC5	54.3	56.1	47.3	47.6
POC6	54.0	59.1	47.3	51.3
POC7	53.5	55.7	47.1	47.8
POC8	54.5	56.6	47.7	48.4
POC9	51.4	55.5	44.9	46.7
POC10	51.8	59.6	45.4	47.3
POC11	57.4	66.0	47.8	53.7
POC12	60.5	70.5	52.0	53.8
POC13	61.0	70.1	51.6	53.3
POC14	57.0	70.3	50.6	58.5
POC15	59.1	71.3	53.4	58.5
POC16	58.5	66.9	49.1	58.3
POC17	58.7	70.8	52.6	58.5
POC18	60.2	68.6	50.7	59.3
POC19	60.3	65.6	46.2	53.2
POC20	55.8	66.0	46.6	53.9
	DNA/RNA match		DNA/RNA mismatch	
	DT2	RT2	DT1	RT1
D2	60.0	57.6	53.1	51.6
ON3	58.4	65.9	47.2	56.0
POC21	58.9	63.6	47.4	55.2
POC22	58.6	63.9	47.1	55.1
POC23	58.1	63.9	46.8	55.3

POC24	58.4	63.7	46.7	54.9
POC25	60.9	64.1	49.1	56.2
POC26	58.9	64.1	48.4	56.1
POC27	59.6	64.5	48.4	56.2
POC28	61.1	65.7	49.6	57.4
POC29	61.3	68.8	45.6	59.4
POC30	57.2	61.3	46.6	54.6

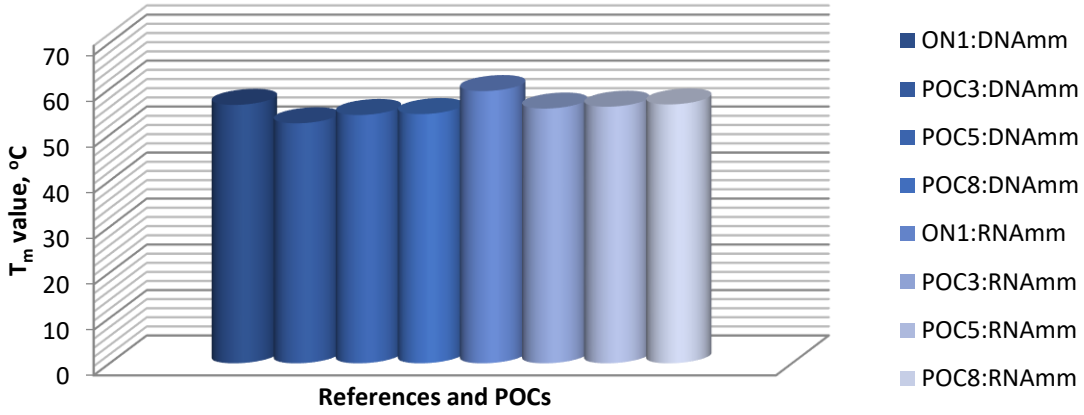
Figure S5. Presentation of T_m values of the POCs when bound to DNA and RNA matching strand



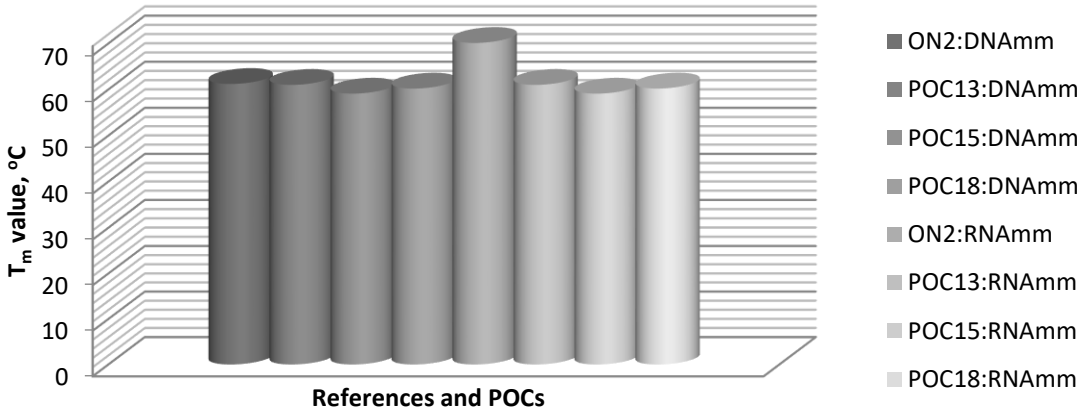
** For the abbreviations see Table S1 and Table 1*

Figure S6. Presentation of T_m values of the reference oligonucleotides and POCs when bound to DNA and RNA matching strand

ON1



ON2



ON3

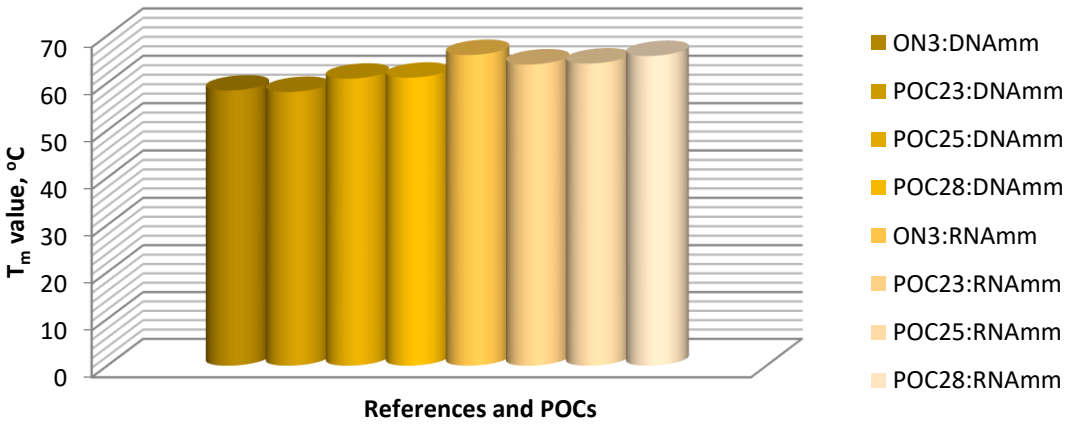
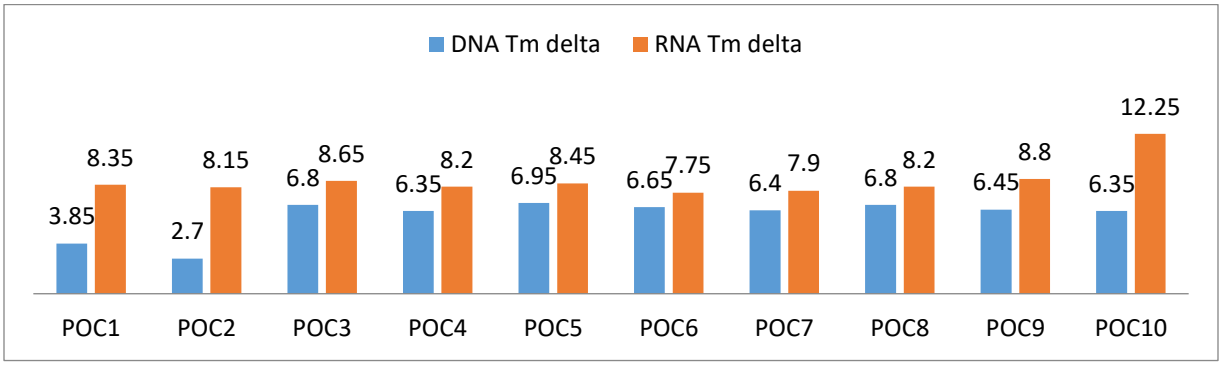
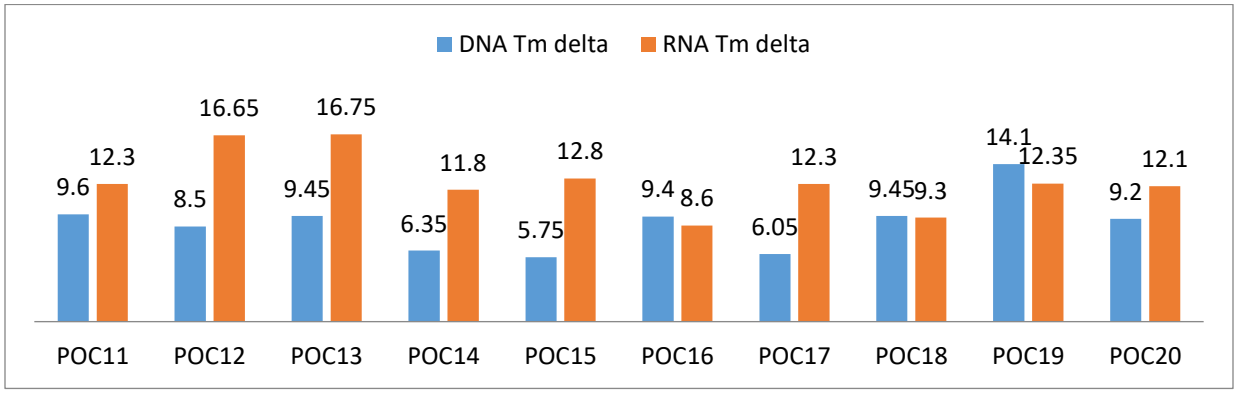


Chart S2. Mismatch discrimination (delta *T*m values) displayed by the POCs*

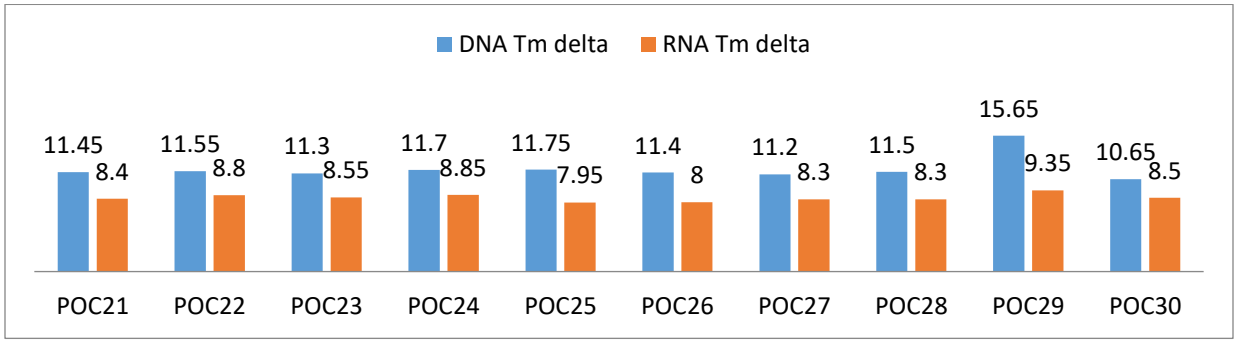
ON1



ON2

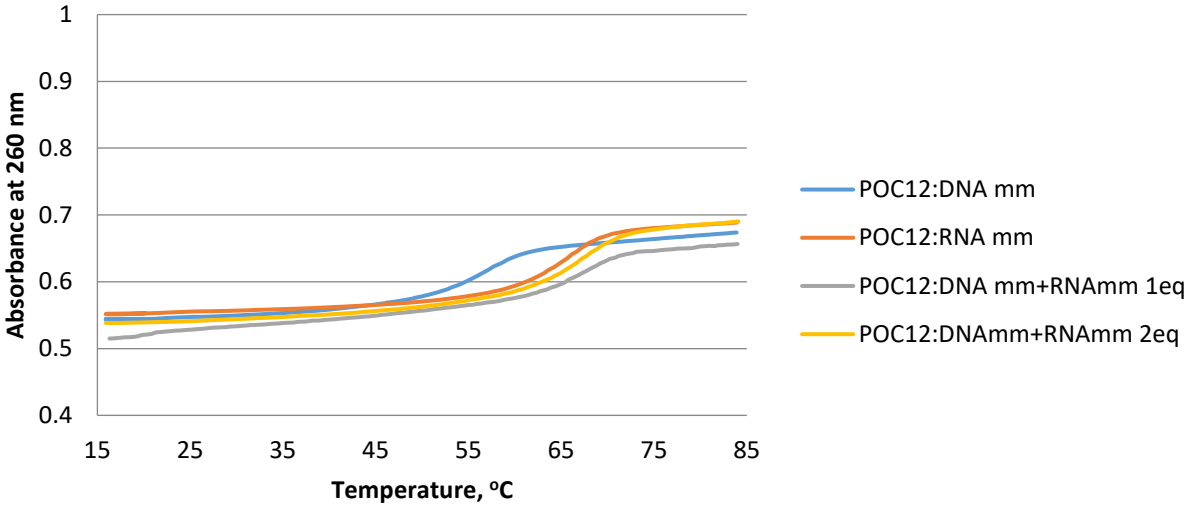


ON3



***** *DNA T*m *delta (blue color) is the difference between T*m *values when the POC is bound to DNA mm and DNA ms. RNA T*m *delta (red color) is the difference between the T*m *values when POC is bound to RNA mm and RNA ms. T*m *delta is used to show the sensitivity of the POCs towards single nucleotide mismatch.*[1]

Figure S7. T_m curves for competitive UV melting experiments

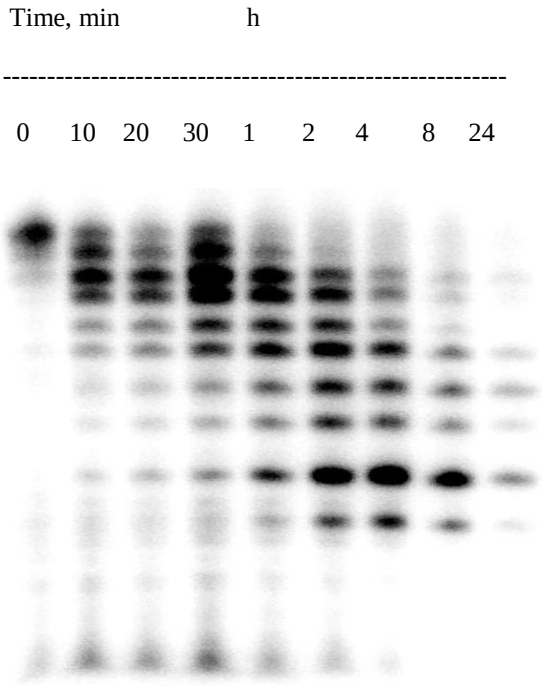


**Absorbance values for the melting curves have been normalized.*

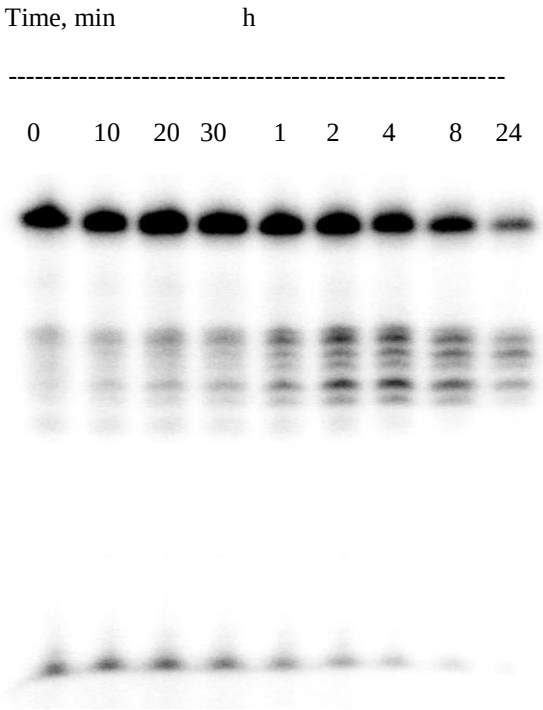
VII. Gel electrophoresis of 5'-^{32}P-labeled oligonucleotides and POCs incubated in 90% HS

Figure S8. Gel electrophoresis of 5'-^{32}P- labeled oligonucleotides and POCs incubated in HS

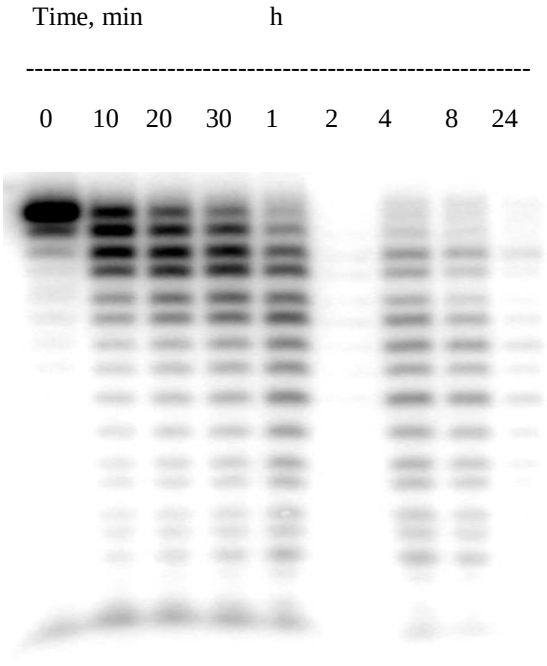
a) **ON2**



b) **POC12**



c) **ON1**



d) **POC8**

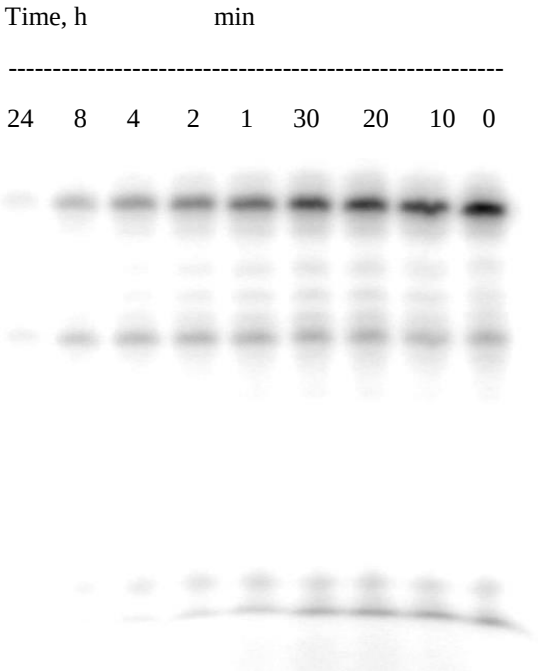
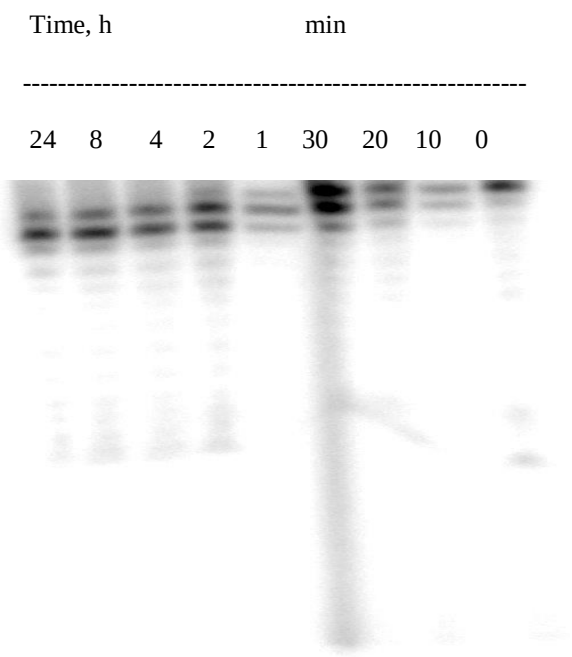
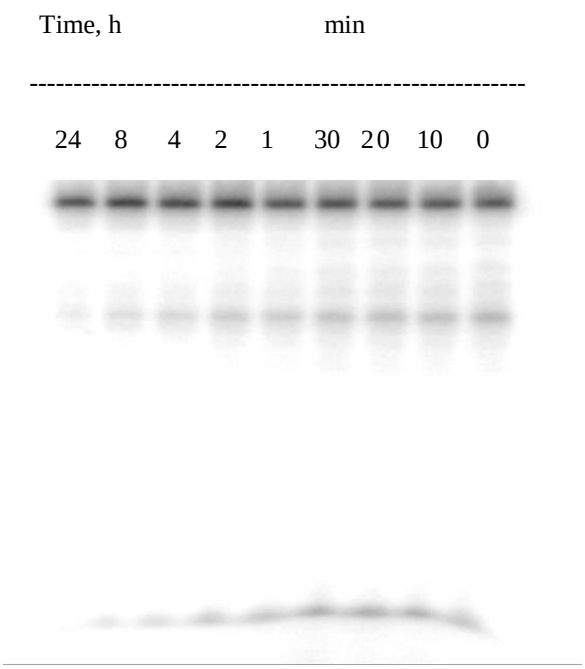


Figure S9. Gel electrophoresis of 5'-^{32}P- labeled duplexes of **ON1** and POCs with a complementary DNA strand incubated in HS

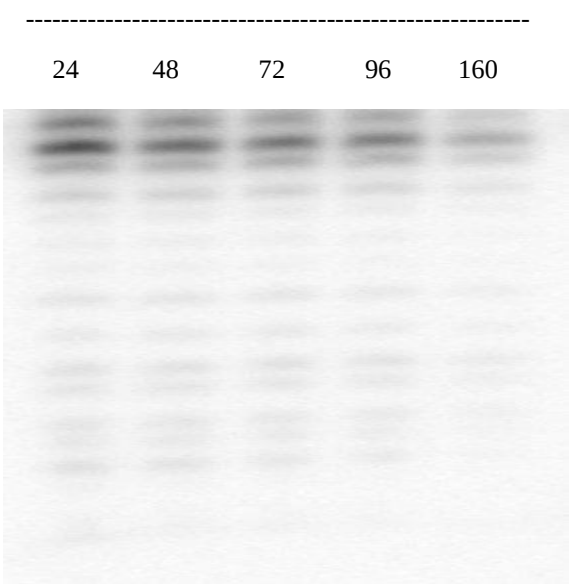
a) **ON1: DNA**



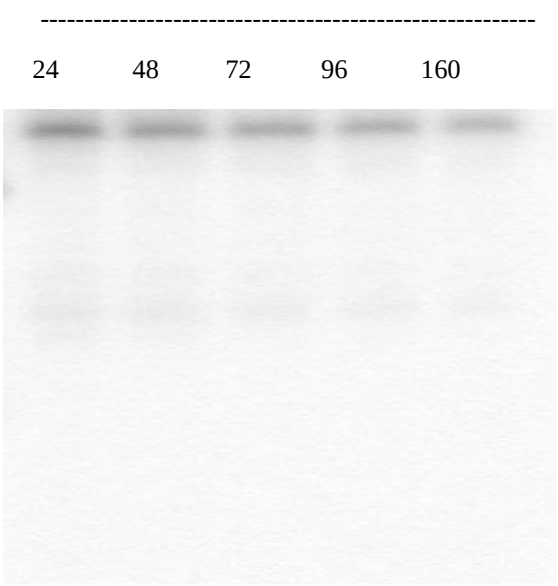
b) **POC8:DNA**



c) **ON1:DNA**



d) **POC8:DNA**



VIII. Statistical analyses

All statistical analyses were performed in R software. One-way ANOVA was used to compare the means of the twelve groups from one variable, T_m. The test statistic used for ANOVA was the F-statistic and it was calculated by taking the mean square for the variable divided by the mean square of the error. The F-statistic can be positive or zero. The P-value is the probability of being greater than the F-statistic or simply the area to the right of the F-statistic. This P-value was used to test the null hypothesis that all the group population means are equal versus the alternative that at least one is not equal. The two tailed *t*-test with unequal variances was run to compare the T_m values of the three oligonucleotides within all groups with the T_m values of each peptide separately within all groups. The *t*-test was performed in Excel and the R- and the *p*-values were calculated using the Data analysis tool.

Linear regression analysis[2] was run for the following variables: number of lysine residues (0-3), number of amino acids between the oligonucleotide and lysine (0-7), number of glycine and tyrosine residues (0-4 and 0-1), and number of nucleic acids between the oligonucleotide and tyrosine (0-7). Resulting data is presented in Tables S8-S12.

P values below 0.1 were considered significant at 90% confidence level. Confidence intervals (CIs) for estimated parameters were calculated using the corresponding probability distributions as described.[3]

Table S5. ANOVA results

Source of Variation	SS	*Df*	MS	F	P-value	F crit
Between Groups	4997.184	11	454.2895	131.6540757	2.94455^{-57}	1.878388
Within Groups	372.668	108	3.45063			
Total	5369.852	119				

Table S6. *t*-test and linear refression results

Peptide	**PEP1**	**PEP2**	**PEP3**	**PEP4**	**PEP5**	**PEP6**	**PEP7**	**PEP8**	**PEP9**	**PEP10**
p-value	0.732	0.803	0.679	0.668	0.979	0.889	0.924	0.938	0.656	0.402
R^2	0.885	0.865	0.922	0.923	0.912	0.955	0.919	0.923	0.876	0.922

Table S7. Descriptive statistics result for the T_m data of POC complexes.*

	T_m Values					
	mmDNA	**mmRNA**	**msDNA**	**msRNA**	**Δ DNA**	**Δ RNA**
MEDIAN ref	58,6	59,5	51,1	52,6	**7,5**	**6,9**
MEDIAN POC	58,4	63,9	47,7	53,9	**10,7**	**10**

**mmDNA/mmRNA = duplexes made with complementary match DNA/RNA strand; msDNA/msRNA = duplexes made with complementary mismatch DNA/RNA strand; MEDIAN ref is a median T_m value for unmodified DNA strand; MEDIAN POC is a median T_m value for all POCs.*

Table S8. Results from a linear regression of T_m values on the number of lysine residues

Variable: number of lysine residues						
	T_m Values					
	mmDNA	**mmRNA**	**msDNA**	**msRNA**	**Delta DNA**	**Delta RNA**
F	0,16	0,21	**2,77**	**0,97**	**1,13**	0,38
Df	34	34	34	34	34	34
P-value	0,69	0,65	**0,08**	**0,12**	**0,29**	0,54
CI lower 90%	-0,7154	-1,16459	-0,01396	-0,51334	-1,5538	-1,05309
CI upper 90%	1,163137	2,035056	1,659436	1,94486	0,356064	0,492039
CI 90%	1,878538	3,199641	1,673401	2,458198	1,909864	1,545129

Table S9. Results from a linear regression of T_m values on the distance to first lysine residue

Variable: distance to first lysine residue						
	T_m Value					
	mmDNA	**mmRNA**	**msDNA**	**msRNA**	**Delta DNA**	**Delta RNA**
F	0,60	0,01	0,03	0,19	0,03	**0,93**
Df	34	34	34	34	34	34
P-value	0,442904	0,905014	0,867418	0,664252	0,874888	**0,342794**
CI lower 90%	-0,60228	-0,65948	-0,34694	-0,69251	-0,37444	-0,14621
CI upper 90%	0,223378	0,760367	0,423534	0,407763	0,451926	0,531849
CI 90%	0,825657	1,419847	0,770472	1,100278	0,826369	0,67806

Table S10. Results from a linear regression of T_m values on the number of glycine residues.

Variable: number of glycines						
	T_m Value					
	mmDNA	**mmRNA**	**msDNA**	**msRNA**	**Delta DNA**	**Delta RNA**
F	**2,63**	0,06	**1,25**	0,26	0,28	**1,79**
Df	34	34	34	34	34	34
P-value	**0,114217**	0,814656	**0,27134**	0,612982	0,598845	**0,189692**
CI lower 90%	-1,29805	-1,00904	-1,03871	-1,18259	-0,92917	-0,11552
CI upper 90%	0,027363	1,336561	0,211966	0,634333	0,485227	0,991295
CI 90%	1,325415	2,345602	1,250676	1,816923	1,414398	1,106815

Table S11. Results from a linear regression of T_m values on the presence of tyrosine residues.

Variable: presence of tyrosine						
	T_m Value					
	mmDNA	**mmRNA**	**msDNA**	**msRNA**	**Delta DNA**	**Delta RNA**
F	**1,65**	0,16	0,16	0,03	0,75	**1,26**
Df	34	34	34	34	34	34
P-value	**0,207608**	0,690774	0,688519	0,864627	0,393837	**0,26933**
CI lower 90%	-4,38614	-3,31354	-2,92073	-3,725	-3,93521	-0,69535
CI upper 90%	0,599473	5,373536	1,794068	3,038333	1,275209	3,442013
CI 90%	4,985612	8,687072	4,714802	6,763332	5,210418	4,137359

Table S12. Results from a linear regression of T_m values on the number of amino acids to tyrosine

Variable: number of amino acids to tyrosine						
	T_m value					
	mmDNA	**mmRNA**	**msDNA**	**msRNA**	**Delta DNA**	**Delta RNA**
F	**2,08**	0,08	**1,90**	0,48	0,02	0,26
Df	34	34	**34**	34	34	34
P-value	**0,158518**	0,7759	**0,177756**	0,49302	0,877322	0,610294
CI lower 90%	**-0,81957**	-0,90825	**-0,74431**	-0,84547	-0,5135	-0,26092
CI upper 90%	**0,065254**	0,644857	**0,076451**	0,354103	0,427046	0,488893
CI 90%	**0,884824**	1,553106	**0,820759**	1,199575	0,940551	0,74981

Table S13. Input data for boxplot shown in Figure 3.

T_m mmDNA	T_m mmDNA	T_m msDNA	T_m msDNA	DNA T_m delta	DNA T_m delta	T_m mmRNA	T_m mmRNA	T_m msRNA	T_m msRNA	RNA T_m delta	RNA T_m delta
58,6	53,8	52,9	50	5,7	3,8	58,2	59,7	50,3	51,4	7,9	8,3
56,65	52,8	39,8	50,1	16,85	2,7	59,5	55,7	52,6	47,6	6,9	8,1
61,2	52,5	51,1	45,7	10,1	6,8	70,1	55,7	58,5	47,1	11,6	8,6
60	52,2	53,1	45,9	6,9	6,3	57,6	55,3	51,6	47,1	6	8,2
58,4	54,3	47,2	47,3	11,2	7	65,9	56,1	56	47,6	9,9	8,5
	54		47,3		6,7		59,1		51,3		7,8
	53,5		47,1		6,4		55,7		47,8		7,9
	54,5		47,7		6,8		56,6		48,4		8,2
	51,4		44,9		6,5		55,5		46,7		8,8
	51,8		45,4		6,4		59,6		47,3		12,3
	57,4		47,8		9,6		66		53,7		12,3
	60,5		52		8,5		70,5		53,8		16,7
	61		51,6		9,4		70,1		53,3		16,8
	57		50,6		6,4		70,3		58,5		11,8
	59,1		53,4		5,7		71,3		58,5		12,8
	58,5		49,1		9,4		66,9		58,3		8,6
	58,7		52,6		6,1		70,8		58,5		12,3
	60,2		50,7		9,5		68,6		59,3		9,3
	60,3		46,2		14,1		65,6		53,2		12,4
	55,8		46,6		9,2		66		53,9		12,1
	58,9		47,4		11,5		63,6		55,2		8,4
	58,6		47,1		11,5		63,9		55,1		8,8
	58,1		46,8		11,3		63,9		55,3		8,6
	58,4		46,7		11,7		63,7		54,9		8,8
	60,9		49,1		11,8		64,1		56,2		7,9
	58,9		48,4		10,5		64,1		56,1		8
	59,6		48,4		11,2		64,5		56,2		8,3
	61,1		49,6		11,5		65,7		57,4		8,3
	61,3		45,6		15,7		68,8		59,4		9,4
	57,2		46,6		10,6		61,3		54,6		6,7

References

1. Astakhova, I.K., Hansen, L.H., Vester, B., and Wengel, J. (2013). Peptide-LNA oligonucleotide conjugates. Org Biomol Chem 11, 4240-4249.
2. Seber, G.A.F., and Lee, A.J. (2003). Linear Regression: Estimation and Distribution Theory. In Linear Regression Analysis. (John Wiley & Sons, Inc.), pp 35-95.
3. Zar, J.H. (1999). Biostatistical Analysis.(Prentice Hall).

Chapter 4

Paper III

Studies of Impending Oligonucleotide Therapeutics in Simulated Biofluids

Domljanovic I, Hansen AH, Hansen LH, Klitgaard JK, Taskova M*, Astakhova K., *Nucleic Acid Ther.*, 2018, doi: 10.1089/nat.2017.0704

"Reprinted (adapted) with permission from (Domljanovic I, Hansen AH, Hansen LH, Klitgaard JK, Taskova M*, Astakhova K., *Nucleic Acid Ther.*, 2018, doi: 10.1089/nat.2017.0704) Copyright (2018) Mary Ann Liebert, Inc."

Studies of Impending Oligonucleotide Therapeutics in Simulated Biofluids

Ivana Domljanovic,[1] Anders Højgaard Hansen,[1] Lykke H. Hansen,[3] Janne Kudsk Klitgaard,[2,3] Maria Taskova,[1]* and Kira Astakhova[1]

Synthetic oligonucleotides, their complexes and conjugates with other biomolecules represent valuable research tools and therapeutic agents. In spite of growing applications in basic research and clinical science, only few studies have addressed the issue of such compounds' stability in biological media. Herein, we studied the stability of two therapeutically relevant oligonucleotide probes in simulated biofluids; the 21 nucleotide long DNA/LNA oligonucleotide ON targeted towards cancer associated *BRAF* V600E mutation, and a longer DNA analogue (TTC) originating from *BRAF* gene. We found that stability of peptide-oligonucleotide conjugates (POCs) in human serum was superior compared to the naked or complexed 21mer oligonucleotide, whereas stability of POCs in simulated gastric juice was dependent on the peptide sequence. Addition of pepstatin A in general increased the stability of oligonucleotides after 24 h digestion in human serum and simulated gastric juice. Similarly, complexation with optimal amounts of histone proteins was found to rescue oligonucleotide stability after 24 h digestion in hydrochloric acid.

Keywords: stability, oligonucleotide, biofluids, peptide, gene therapy

Introduction

Antisense oligonucleotides (ASOs) comprise short synthetic DNA oligonucleotides (~15–25 nucleotides long) designed to hybridize to a unique RNA target. Upon target recognition, ASOs modulate gene expression by different mechanisms, such as RNase H-induced silencing or steric blocking of the mRNA, in a sequence-dependent manner [1–3]. Various modifications have been introduced into ASOs to improve their pharmacokinetics and biodistribution profiles, some of which include incorporation of backbone phosphorothioates, modification of the ribose sugar at the 2' position, i.e. installation of fluorine (2'-F) or methoxy (2'-OMe) groups, and replacement of natural ribose-based scaffold by a locked nucleic acid (LNA) [4-6]. Overall, it is recognized that chemical modification help to improve stability, bioavailability and target binding, including single nucleotide polymorphism discrimination by ASOs [5–7]. Formulation strategies that enable the protection of oligonucleotides from degradation by enzymes and improve cellular delivery are highly warranted for development of next generation therapeutic oligonucleotides. In literature, various approaches have been proposed to enhance *in vitro* and *in vivo* stability and to improve targeted delivery, such as conjugation of oligonucleotides to different biomolecules [8–10]. Therapeutic oligonucleotides have previously been covalently and non-covalently conjugated to lipids, proteins, and carbohydrates [8]. One example reported by Prakash et al. describes ASOs conjugated to tri-antennary *N*-acetylgalactosamine (GalNAc), which were efficiently internalized into hepatocytes by the use of the hepatocyte-specific asialoglycoprotein receptor. Upon cellular uptake, ASO-GalNAc conjugates were able to modulate gene expression and prolong the duration of action compared to unconjugated ASOs [11].

Using a small library of POCs, we recently reported on a structure−property relationship demonstrating how peptide sequences and individual LNA incorporations affect target binding ability and stability in human serum [12]. The ability of POCs to bind target strands was found to be dependent on peptide sequence and on the positioning of LNA within the oligonucleotide sequence. Moreover, POCs showed high stability in 90% human serum after 24 h incubation, an effect that was ascribed to the presence of peptide (independent of the sequence) and not LNA [12].

To further promote synthetic oligonucleotides as drugs, oral delivery is highly desired and is already explored for several ASOs [13]. However, only a few studies have systematically

1. Department of Chemistry, Technical University of Denmark, Kemitorvet 206-207, 2800 Kgs. Lyngby, Denmark
2. Institute of Clinical Research, University of Southern Denmark, J.B. Winsløws Vej 21, 2 55, 5000 Odense C, Denmark
3. Department of Biochemistry and Molecular Biology, University of Southern Denmark, Campusvej 55, 5230 Odense M, Denmark

addressed oligonucleotide stability in acidic biofluids [14,15]. Oligonucleotides are generally not considered to be chemically stable under the harsh conditions of the gastrointestinal tract, wherefore oral administration of oligonucleotide therapeutics typically is not considered a viable route of administration [15,16]. A small number of reports on dietary oligonucleotides, i.e. nucleic acids ingested from foods, is raising awareness [15]. It has been shown that orally ingested nucleic acids have important effects on vital processes in mammalian animals, such as growth performance, immune responses and antibacterial resistance [17,18]. Moreover, the notion that histones and other proteins are separated from the dietary nucleic acids only in the stomach, and that actual digestion of oligonucleotides into mononucleotides happens further on in the intestinal tract was recently challenged [17,19]. Efficient digestion of dietary nucleic acids by pepsin in gastric juice mixtures and in the stomach of two types of fish, i.e. *Channa argus* and *Epinephelus awoara*, was recently reported [15,21]. Even though enzymes of the bovine stomach showed no detectable digestion of nucleic acids, it was concluded that oligonucleotide breakdown in the stomach may further accelerate digestion in the intestine and that this effect may be dependent on the exact genotype [20].

A small number of oligonucleotide-based therapeutics already exists on the marked, and others are under development [21]. However, more formulation tools enabling effective and safe administration of oligonucleotide-based therapeutics are highly warranted [22]. Rather than focusing on targeted delivery of oligonucleotides, this study aimed at investigating how different formulation approaches in general affected the stability of two therapeutically relevant oligonucleotides in simulated biofluids. Two oligonucleotides, one short (ON) that targets cancer associated *BRAF* V600E mutation, and one long (TTC) representing a *BRAF* genomic fragment previously identified in various cancer forms most predominantly in malignant melanomas (66%), were included [23].

Materials and Methods

Chemicals

2'-*O*-propargyl-3'-CEP uridine was purchased from Jena Bioscience. TBTA (Tris[(1-benzyl-1*H*-1,2,3-triazol-4-yl)methyl]amine) was obtained from Lumiprobe LLC. Human serum (HS, Immunovision, cat. no. HNP 0000, USA), Hank's balanced salt solution, pepsin from porcine gastric (EC number 3.4.23.1), α-amylase from porcine pancreas (EC number 3.2.1.1), mucin from porcine stomach (EC number, 282-010-7), pepstatin A – microbial, ≥ 90% (HPLC), poly-L-lysine – mol wt > 30,000 and hydrochloric acid (HCl) (37%) were purchased from Sigma-Aldrich. Histone proteins H4 and H2B, and T4 polynucleotide kinase (PNK) were obtained from New England BioLabs. ATP-γ-^{32}P was purchased from Perkin Elmer. Peptides PEP3 and PEP8 were obtained as described previously [12]. Commercially available TTC oligonucleotide was purchased from Integrated DNA Technologies. All reagents were used without further purification.

Oligonucleotide synthesis

The 21mer oligonucleotide ON (**Fig. 1A**) was synthesized on an automated DNA synthesizer PerSpective Bio-systems Expedite 8909 using a standard protocol in a DMT-ON mode (1.0 μmol scale, CPG support). LNA monomers and monomer X (2'-*O*-propargyl-3'-CEP uridine, **Fig. 1B**) were incorporated using hand-coupling procedure [12]. Coupling efficiencies were monitored by measuring the absorbance of the dimethoxytrityl cation obtained after each coupling step. Cleavage from the CPG solid support was performed with 30% aqueous ammonia at 55 °C overnight. The identity of ON was confirmed by MALDI-TOF using Ultraflex II TOF/TOF instrument from Bruker and 3-hydroxypicolinic acid matrix (10 mg/mL 3-hydroxypicolinic acid, 50mM ammonium citrate in 70% aqueous acetonitrile). Purity was assessed by IE HPLC using a Merck Hitachi LaChrom instrument equipped with a Dionex DNAPac Pa-100 column (250 mm x 4 mm).

Covalent conjugation

POC13 and POC18 (**Fig. 1C**) were synthesized by CuAAC "click" reaction between internal double alkyne-functionalized ON and N-terminal azide-functionalized PEP3 and PEP8 (Supplementary, **Table S1**). In 200 μL total reaction volume of DMSO:water (1:1), 20 nmol oligonucleotide ON was mixed with 120 nmol peptide (PEP3/PEP8), 4 μL (50 mM) aminoguanidine hydrochloride, 10 μL (10 mM) Cu(II)-TBTA and 4 μL (50 mM) ascorbic acid. The reaction mixture was degassed, placed in a microwave reactor for 15 min at 60 °C, and stored at room temperature overnight. Next, POC products were purified on an Illustra NAP-5 column (GE Healthcare) using the protocol from the manufacturer. Product conjugates were analyzed by MALDI-TOF mass spectrometry and IE HPLC (**FIG. S1** and **FIG. S2**, **Fig. 1E)**.

Non-covalent complexation

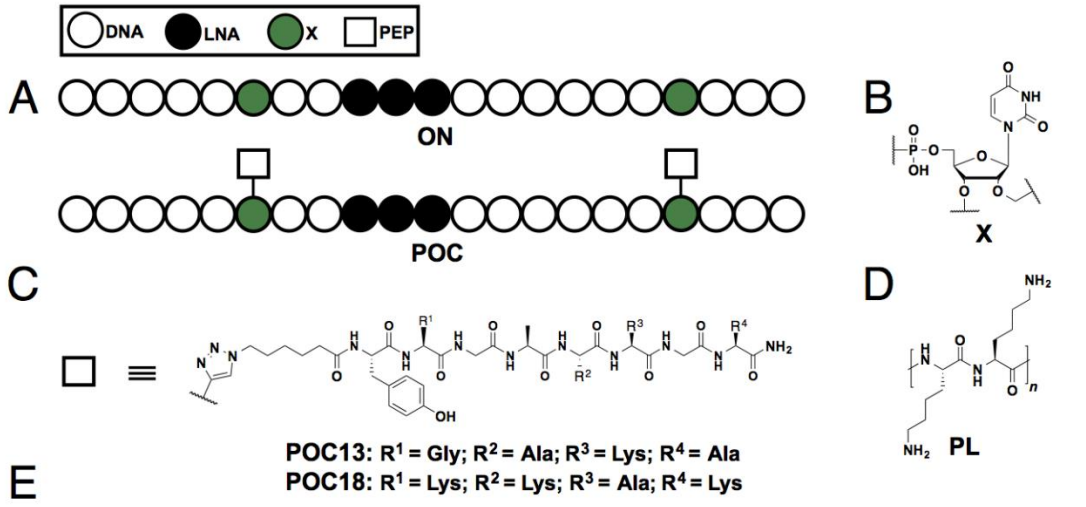
ON-poly-L-lysine (PL) complexes were prepared by mixing 4 pmol ON with 8 pmol PL in a total volume of 15 μL in water at room temperature for 15 min. TTC-histone protein H2B complexes (TTC-H2B, 1:2 and 1:4) were prepared by mixing 4.5 pmol TTC with 9 pmol/18 pmol H2B in a total volume of 14 μL/15 μL water at room temperature for 15 min (Supplementary, **Table S1**). Same protocol was adapted to prepare TTC-H4 (1:2 and 1:4).

^{32}P labeling

Oligonucleotides, POCs and peptide-oligonucleotide complexes were ^{32}P-labeled at the 5'-end using the following procedure: 1 μL oligonucleotide or POC at a concentration of 3 pmol/μL; 1 μL 10x kinase buffer; 3 μL, 10 mCi/mL ATP-γ-^{32}P blue (502Z 1mCi, 6000 Ci/mmol); 5μL water and 0.2 μL T4 PNK enzyme (10 U/mL) were mixed in a PCR tube (total volume of 10.2 μL). The samples were incubated for 1.5 h at 37 °C followed by 15 min at 75 °C, and then stored at -20 °C.

Digestion in 90% human serum (HS) – General procedure

Radioactive oligonucleotide or POC (0.11 pmol) and unlabeled oligonucleotide or POC (4.5 pmol) were added to 18.4 μL human serum (90%, in Hank's Balanced Salt Solution (HBSS) buffer, pH = 7.4) in a PCR tube to give a total volume of 23 μL. Then, 1.6 μL 50 mM EDTA was added and samples were incubated in an Eppendorf personal



ON/POC	Sequence (5′ to 3′)	Mass Calcd.	Mass Obs.	%UV Purity	IE HPLC Ret. time (min)
ON	CGA GAX TTC T C T GTA GCX AGA	6626.0	6629.4	92.4	16.82
POC13	CGA GAX[-PEP3] TTC T C T GTA GCX[-PEP3] AGA	8291.0	8293.9	100	15.63
POC18	CGA GAX[-PEP8] TTC T C T GTA GCX[-PEP8] AGA	8519.0	8521.0	100	16.43

FIG. 1. (A) Structures of 21mer oligonucleotide (ON) and POCs used in this study. LNA, locked nucleic acid; (B) X: 2-O Propargy l-2'-deoxyuridine derivative; (C) PEP, peptide; (D) PL, poly-l-lysine; (E) POCs, peptide–oligonucleotide conjugates; POC13, POC where the peptide PEP3 is covalently conjugated to ON; POC18, POC where the peptide PEP8 is covalently conjugated to ON.

Mastercycler at 37 °C. Aliquots (2 µL in 1 µL loading dye) were withdrawn at the following time points: 0 min, 10 min, 20 min, 30 min, 1 h, 2 h, 4 h, 8 h, and 24 h. The samples were then resolved on 13% denaturing polyacrylamide/bisacrylamide gels (30:1, 7 M urea, 1xTBE, 0,3 mm thick) with 1x TBE as a running buffer. The gels were run for 1 h at 13 W and dried under vacuum (1 h, 80 °C), exposed overnight and visualized by autoradiography on a Typhoon Trio Variable Mode Imager (Amersham Biosciences). ON, POC13, POC18, ON-PL, TTC, TTC-H2B, and TTC-H4 were digested in 90% human serum.

Digestion in 90% human serum with inhibitor pepstatin A (HSi)

The HSi mixture was prepared to contain 0.5 µg/µL pepstatin A (10% (v/v) acetic acid in methanol) followed by incubation for 30 min at 37 °C. Samples (ON, ON-PL, POC13 and POC18) were digested as described in the General Procedure.

Digestion in simulated gastric juice mixture (GJ)

The simulated gastric juice mixture was prepared to contain 2 mg/mL mucin, 4 mg/mL pepsin, 4 mg/mL α-amylase, and 50 mM HCl in 50 mM PBS buffer. The pH was adjusted to 3.5 at 37 °C. Samples (ON, ON-PL, POC13, and POC18) were digested as described in the General Procedure.

Digestion in gastric juice with inhibitor pepstatin A (GJi)

The mixture was prepared to contain 2.5 µg/µL pepstatin A (10% (v/v) acetic acid in methanol) followed by incubation for 30 min at 37 °C. Samples (ON, ON-PL, POC13, and POC18) were digested as described in the General Procedure.

Digestion in HCl

Samples (ON, POC13, TTC, TTC-H2B, and TTC-H4) were digested in 0.5% or 1% HCl containing 0.05 mM KCl according to the General Procedure.

Radioactive gels analysis

Typhoon Trio scan data of radioactive gels were analyzed using GelEval (Version 1.37). For each gel, band intensities at different time points (I_t, for t = 10, 20, 30 min, and 1, 2, 4, 8 and 24 h) corresponding to remaining non-digested oligonucleotide starting material were normalized with respect to the initial band intensity ($I_{t=0}$) thereby providing the estimated remaining fraction of oligonucleotide (RF_{oligo}) at the indicated time points (Equation 1).

$$RF_{oligo} = \frac{I_t}{I_{t=0}} \quad (1)$$

Digestions exhibiting a clear trend towards exponential decay were assumed to follow first order kinetics and fitted to an exponential one-phase decay model (goodness of fit criterion of $R^2 > 0.8$), from which the rate constants (k) and half-lives ($t_{1/2}$) were estimated (Equation 2).

$$t_{1/2} = \frac{\ln 2}{k} \quad (2)$$

Digestions were performed in at least duplicates on two different days. Also, using GraphPad Prism (version 5.0c) software, all time points for digestion experiments were presented as the average mean of repeated experiments, and error bars represent the standard deviation (except for rate constants, which are presented as means ± S.E.M.).

Results

NAT design

In this study, we designed a 21mer DNA/LNA oligonucleotide (ON) that contains two internal alkyne functionalities and target the *BRAF* V600E mutation *in vivo* (**Fig. 1A** and **Fig. 1E**). Next, we synthesized POC13 and POC18 (**Fig. 1C** and **Fig. 1E**) composed of ON covalently conjugated to two peptides (POC13: ON+2xPEP3; POC18: ON+2xPEP8) each containing either one (PEP3) or three (PEP8) positively charged lysines. Moreover, ON was formulated as a non-covalent complex with PL (**Fig. 1D**), a cationic polymer able to facilitate delivery of oligonucleotides [24]. It has previously been reported that short peptides containing positively charged residues exhibit enzyme-protective properties, while also promoting intracellular delivery of the corresponding POC [12,25]. In addition, a longer 50 nt synthetic DNA analogue of the *BRAF* genomic fragment (TTC, Supplementary, **Table S1**) was complexed to histone proteins H2B or H4 (Supplementary, **Table S1**), which facilitate the packaging of DNA in eukaryotic cells [26,27]. In this study, we focused solely on the ability of said peptides, PL and histone proteins to modulate stability of oligonucleotides in simulated biofluids rather than on their potential as moieties for promoting intercellular delivery.

Selected simulated biofluids

Various administration routes for therapeutic oligonucleotides have been reported in literature, such as intravitreous administration of formvirsen for treatment of cytomegalovirus retinitis [28] and intramuscular administration of ASO PRO051 in patients with Duchenne's muscular dystrophy [29]. It was also shown that ASO PRO051 successfully could be applied *via* subcutaneous administration [30]. Further, *in vivo* studies in mouse models have demonstrated that dynamin2 (*Dnm2*) ASO could be administered by intraperitoneal injections leading to DNM2 gene knockdown and the reversal of muscle pathology within 2 weeks [31].

In the present study, we investigated how different formulation strategies affected stability of oligonucleotides in simulated biofluids important to consider in the context of oral administration (simulated gastric juice and HCl assays), and systemic circulation (human serum assay). To study stability of oligonucleotide probes in different simulated biofluids, we resolved the ^{32}P 5'-end labeled ON and POCs on denaturing polyacrylamide gels that were found to be sensitive and convenient for low (pM) concentration oligonucleotides and POCs. Oligonucleotides were labeled with the ^{32}P isotope at the 5'-termini, digested in biofluids for various time points (0 min - 24 h) and subsequently resolved on 13% denaturing polyacrylamide gels (**FIG. S3-5**).

Oligonucleotide and peptide digestion in simulated biofluids

First, 21mers ON, ON-PL, POC13 and POC18 were digested in human serum and mean remaining fraction of the oligonucleotide (RF_{oligo}) were estimated after 2 h, 8 h, and 24 h incubation (**FIG. 2A**). While roughly 20% of naked oligonucleotide (ON) remained intact after 24 h digestion, the non-degraded portions estimated for ON-PL, POC13 and

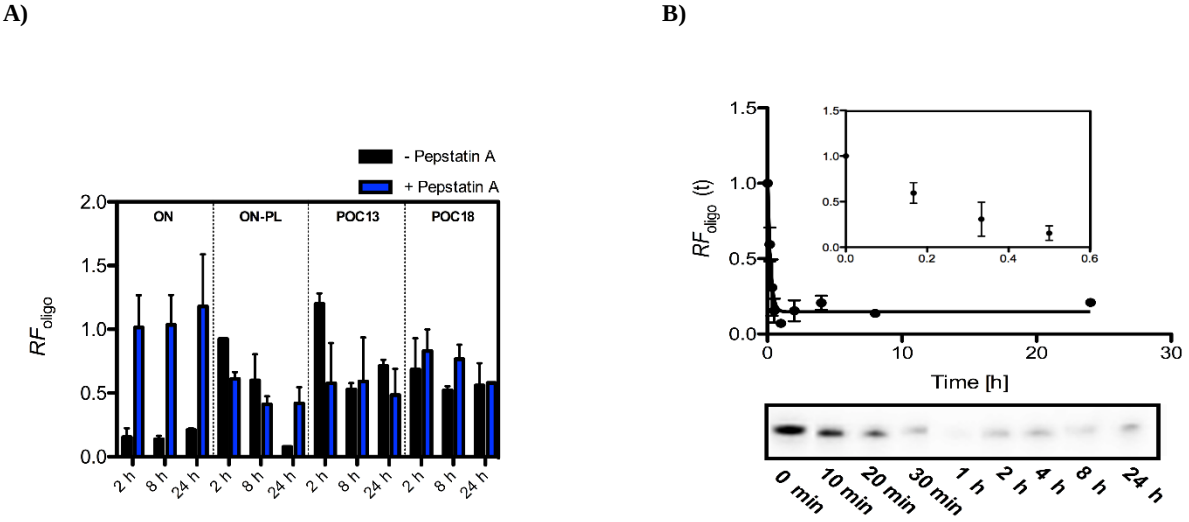


FIG. 2. (A) Estimated remaining fraction of oligonucleotide after 2, 4, and 8 h digestion in HS. Data represent mean ± SD. (B) Oligonucleotide (ON) digestion in HS; ON levels estimated over the course of 24 h at different time points fitted to an exponential one-phase decay model. Representative gel displaying ON main bands is depicted below the graph. Data represent mean ± SD. HS, human serum.

POC18 were 10%, 70% and 55%, respectively. Notably, as observed in **FIG. 2B** ON exhibits fast degradation kinetics in human serum ($t_{1/2}$ of 8.9 min and k of 4.6 ± 0.8 h^{-1}) whereas ON-PL, POC13 and POC18 showed higher degree stability with visible band intensities at 8 h time point (Supplementary, **FIG. S3**). Next, human serum was pretreated with pepstatin A (HSi) to potentially inhibit proteases such as endopeptidases. Remaining fractions of the probes ON and ON-PL after 24 h incubation were observed to further increase to roughly 100% and 50%, respectively, while pepstatin A did not markedly affect the stability of POC13 and POC18 (**FIG. 2A**).

Digestion in simulated gastric juice revealed that around 80% of naked 21mer ON had degraded after 24 h incubation and that the complexation to PL extended RF_{oligo} value by roughly two-fold (RF_{oligo} of ~ 30%). Co-addition of pestatin A clearly impeded the degradation of oligonucleotide and its complex with PL (RF_{oligo} of ~ 60% and 70%, respectively), as observed in **FIG. 3C**. POC13 and POC18 were subjected to the same conditions (RF_{oligo} as a function of time is depicted in **FIG. 3A**, **FIG. 3B**), and while POC13 exhibited a marked increase in RF_{oligo} (~ 45% at 24 h time point) when compared to ON and ON-PL complex, POC18 was almost fully degraded after 24 h incubation (RF_{oligo} ~ 5%). In addition, this set of experiments was fitted to an exponential one-phase decay model in order to estimate first order rate constants (k, expressed as h^{-1}) and half-lives ($t_{1/2}$, expressed in min) of the differently formulated ON probes (**FIG. 3D**). Of the two POCs tested in simulated gastric juice assay, POC13 did exhibit the shortest half-life ($t_{1/2}$ of 11 min) while the rate of degradation was lower for POC18 ($t_{1/2}$ of 31 min). Despite the shorter half-life of POC13, degradation of

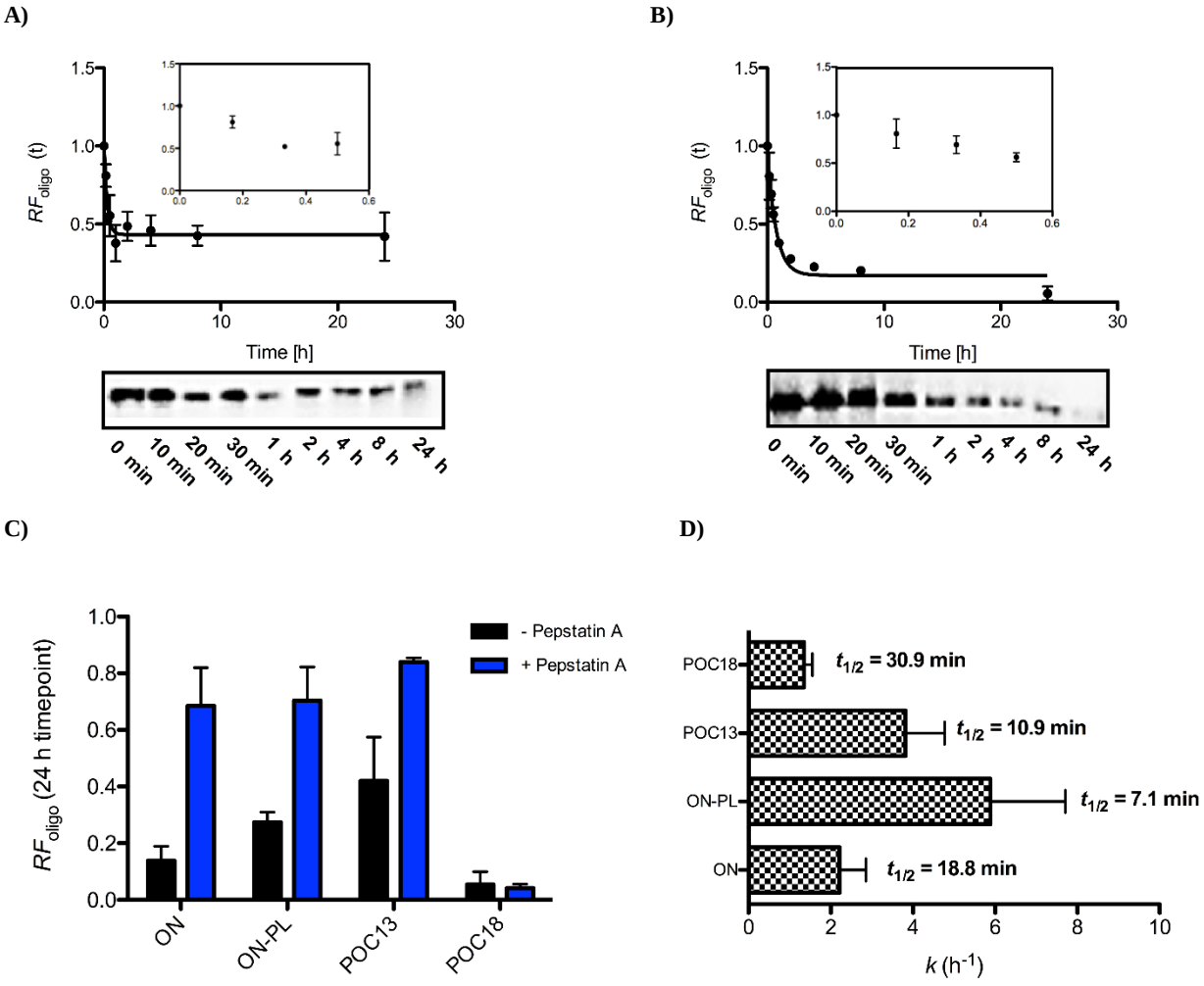


FIG. 3. (A) POC13 digestion in simulated GJ; POC13 levels estimated over the course of 24 h at different time points and fitted to an exponential one-phase decay model. Representative gel displaying POC13 main bands is depicted below the graph. Data represent mean ± SD. (B) POC18 digestion in simulated GJ; POC18 levels estimated over the course of 24 h at different time points and fitted to an exponential one-phase decay model. Representative gel displaying POC18 main bands is depicted below the graph. Data represent mean ± SD. (C) *RF*oligo values of ON, ON–PL, POC13, and POC18 after 24 h incubation in simulated GJ. Data represent mean ± SD. (D) Kinetic profile of ON, ON–PL, POC13, and POC18 in simulated gastric juice. Best-fit values of k represent mean ± SEM. GJ, gastric juice.

POC13 reached a plateau after 2 h incubation time (RF_{oligo} ~ 40%) after which further degradation was impeded, whereas degradation of POC18 in simulated gastric juice continued until nearly full consumption after 24 h. Similarly, comparing ON ($t_{1/2}$ of 19 min) and ON-PL ($t_{1/2}$ of 7 min), naked oligonucleotide (ON) exhibited the slowest initial degradation kinetics, however over a 24 h incubation period PL did to some extent impede degradation of parent ON (ΔRF_{oligo} ~ 13%) in simulated gastric juice (**FIG. 3C**). Inhibition of pepsin by pepstatin A significantly increased remaining band intensities of ON, ON-PL and POC13 when evaluated after 24 h digestion (**FIG. 3C**), while POC18 appeared to be unaffected by the presence of pepstatin A.

Altogether, POC13 exhibited the highest stability in simulated gastric juice at the 24 h time point and was further tested in HCl assay together the corresponding free oligonucleotide (ON). Both compounds exhibited high stability in 0.5% HCl at the 8 h time point (RF_{oligo} ~ 90%), but efficiently degraded when subjected to 1% HCl (**FIG. 4**) (radioactive gels, Supplementary, **FIG. S4**).

We further examined the stability of TTC and its complexes with histone proteins, TTC-H2B and TTC-H4. In human serum, TTC and the corresponding complexes TTC-H2B and TTC-H4 exhibited moderate to high stability at the 24 h time point (digestion analysis on radioactive gels, Supplementary **FIG. S5**). When subjected to acidic conditions, TTC was found only to be stable for 2 h in 1% HCl (RF_{oligo} of ~ 8%), whereas considerable amounts of TTC remained after 8 h incubation in 0.5% HCl (RF_{oligo} of ~ 40%) (**FIG. 4**). Next, the ability of histone proteins to protect TTC from degradation in HCl was investigated. Briefly, the addition of 2 equivalents of histone proteins caused a slight decrease in half-life, and despite degradation of TTC to some extent halted at the 2 h time point when compared to TTC without H2B and H4, histone proteins did not significantly prevent TTC degradation in 1% HCl (green bars, **FIG. 4**, Supplementary **FIG. S6**). Interestingly, 4 equivalents of H2B or H4 was found to cause a marked increase in stability of TTC in 1% HCl even though this effect was more pronounced for TTC-H2B than for TTC-H4 when evaluated at the 4 h time point (blue bars, **FIG. 4**). The decent stability of POC13 in simulated gastric juice (**FIG. 3C**) encouraged further evaluation of this POC and ON (control) in acidic media. In short, both POC13 and ON exhibited a high degree of stability in 0.5% HCl at the 8 h time point (**FIG. 4**), and despite rapid degradation of both analogs within the first hour of incubation in 1% HCl to lower RF_{oligo} values, appreciable band intensity of POC13 remained after 2 h incubation period (mean RF_{oligo} of ~ 30%).

Finally, we studied two aspects that additionally could affect the stability of oligonucleotides in human serum and in HCl, namely peptide degradation and oligonucleotide depurination. We used the precursor of POC18 (PEP8) for digestion experiments in human serum and 0.5% HCl, which were monitored by reverse phase HPLC (Supplementary, **FIG. S7-S8**). PEP8 was found to degrade over time and after 24 h incubation at 37 °C the amount of peptide decreased by roughly 50% and 70% for digestion in human serum and 0.5% HCl, respectively.

Discussion

Antisense therapeutics is an exciting class of emerging pharmaceuticals for treatment of a variety of diseases. One of the most important properties therapeutic oligonucleotides need to fulfill is to exhibit resistance towards enzymatic degradation. In order to survive systemic circulation, oligonucleotides have to be stabile towards enzymes in human serum, such as exo- and endonucleases and possibly also endopeptidases, although the exact role of the latter and other proteases in the breakdown of oligonucleotides remains unclear [32–34]. All nucleases have the ability to cleave the phosphodiester bond of oligonucleotides, where 3'-exonuclease predominantly cleave 3' non-modified oligonucleotides [35].

We investigated digestions of 21mer oligonucleotide probes in human serum. POC13 and POC18 displayed approximately threefold higher stability at the 24 h time point when compared to ON and ON-PL. These results are in agreement with previous reports that peptides conjugated covalently to oligonucleotides increase the resistance towards enzymatic degradation of the parent oligonucleotide in human serum [12]. Furthermore, POC13 (RF_{oligo} of ~ 70%, 24 time point) and POC18 (~ 55%) exhibited higher degree of stability compared to ON-PL (~ 8%), indicating that merely stoichiometric amounts of peptide is needed in order to achieve moderate oligonucleotide stability in human serum over 24 h. In addition, a non-covalent formulation strategy using PL was found to be suboptimal, presumably due to uncontrolled dissociation of PL from oligonucleotide and from of binding of PL to negatively charged human serum albumin [36]. The short ON quickly degraded in human serum reaching an RF_{oligo} value of approximately 20% after 24 h incubation, whereas the longer TTC sequence was significantly more stable under these conditions (RF_{oligo} of approximately 80%). This suggests that stability of naked oligonucleotides is dependent on the nature and the length of the sequence and the apparent higher stability of TTC may be due to the possible formation of secondary structures such as hairpins or dumbbells [37].

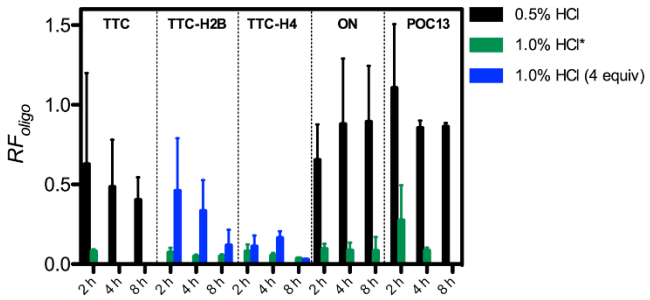


FIG. 4. Estimated remaining fraction of oligonucleotide after 2-, 4-, and 8-h digestion in HS, 0.5% HCl (pH of 1.04), and 1% HCl (pH of 0.65). *2 equivalents of histone protein added to TTC (TTC–H2B and TTC–H4). HCl, hydrochloric acid

The finding that inhibition of pepsin by pepstatin A impedes the breakdown of nucleic acids in stomach [20] prompted us to perform control experiments in human serum of oligonucleotide degradation in the presence of pepstatin A. At the 24 h time point, both 21mer oligonucleotide and its complex with PL were characterized by increased band intensities and RF_{oligo} values, whereas degradation of POC18 and POC13 was not significantly affected by pepstatin A. These results suggest that certain proteolytic enzymes such as endopeptides could have a role in the degradation of oligonucleotides in human serum [38], and that the breakdown of POC13 and POC18 in this media most likely occurs via a different mechanism. Ruling out pepsin as main actor in oligonucleotide degradation in human serum is further substantiated by the facts that its proenzyme pepsinogen A is present at low concentration in human serum [39], that pepsinogen A is converted to pepsin only at low pH [40], and that pepsin exhibits low catalytic activity at pH 7.4 (maximal catalytic activity of porcine pepsin occurs at pH 3.0 [41]).

21Mer oligonucleotide probes were also examined in simulated gastric juice. Incubation of reference oligonucleotide ON and ON-PL led to a substantial loss of oligonucleotide after 24 h digestion - an effect that was efficiently hampered by the addition of pepstatin A (approximate four-fold and three-fold increased band intensity, respectively). This demonstrates that inhibition of porcine pepsin by pepstatin A to a large extent improves oligonucleotide stability in simulated gastric juice. Again, this finding is in line with previous reports that nucleic acid degradation is inhibited by pepstatin A [14]. Further, relative to ON and ON–PL, POC13 displayed a pronounced increase in RF_{oligo} (~ 40% at 24 h time point) in simulated gastric juice, whereas POC18 was almost fully degraded after 24 h incubation (RF_{oligo} ~ 5%). Despite the longer half-life of POC18 ($t_{1/2}$ of 31 min) in simulated gastric juice when compared to POC13 ($t_{1/2}$ of 11 min), the ability of the slightly less cationic peptide (PEP3) was found to be superior at protecting ON against degradation by porcine pepsin when compared to PEP8. Notably, ON-PL and POC18, both containing higher net positive charge than POC13, were found to have lower stability compared to POC13 when estimating RF_{oligo} values at the 24 h time point. This observation is in agreement with previous reports describing how attractive forces can help recruit the substrate to the enzyme interface [20]. The isoelectric point of pepsin is 2.2 [42], and at pH 3.5 porcine pepsin is negatively charged, most likely causing it to bind more efficiently to POC18 and ON-PL than to POC13 [43]. Further experiments are needed to fully understand why POC13, despite its shorter half-life of 11 min, exhibits higher stability over 24 h digestion in simulated gastric juice.

Addition of pepstatin A was found to efficiently inhibit degradation of ON, its complex with PL, and POC13 (**FIG. 3C**) thereby indicating that these compounds can act as substrates for pepsin [14,20,44]. Contrarily, since POC18 stability was unaffected by the presence of pepstatin A, POC18 presumably degrades in simulated gastric juice via a different mechanism.

Genomic DNA wraps around positively charged histone proteins rendering a condensed form of DNA [45,46] that likely contribute the relatively high stability of DNA in acidic environments [27]. We hypothesized that similar interactions might arise in upon complex formation between histone proteins (H2B and H4) and the 50 bp long TTC oligonucleotide. Depurination of nucleic acids is known to occur under harsh acidic conditions (pH < 2), such as in gastric juice [14,44]. To the best of our knowledge, there are only few reports on the stability of encapsulated synthetic oligonucleotides in hydrochloric acid [47], and none regarding the stability of naked oligonucleotides. The stability of single-stranded (ss)DNA encapsulated into a spherical hydrogel matrix of polyacrylamide and covered by a cross-linked polystyrene shell was previously investigated by others. The half-life of ssDNA digested in acidic Theorell–Stenhagen buffer (pH 2.0) was 9.6 h, whereas the half-life of encapsulated ssDNA was 75 h [47]. We observed no signs of depurination of a DNA/LNA ASO in 0.5% HCl over 12 h incubation period at 37 °C. However, treatment with human serum led to loss of adenine after 1 h incubation followed by complete degradation at the 4 h time point (data not shown). 21Mer oligonucleotides ON and POC13 were found to be equally stabile in 0.5% HCl (pH ~ 1, black bars, **FIG. 4**) with decent amounts of remaining oligonucleotide after 24 h incubation. However, increasing the concentration to 1% HCl rendered fast degradation of both ON and POC13 (RF_{oligo} values of ~ 8%, 4 h time point), effects that can be solely ascribed to chemical degradation at very low pH (pH < 1).

The 50 bp TTC oligonucleotide displayed moderate stability in 0.5% HCl when evaluated after 24 h incubation, whereas TTC stability was highly compromised in 1% HCl ($t_{1/2}$ of 26 min, no visible bands after 2 h incubation). Unexpectedly, naked oligonucleotide (ON) exhibited higher stability than TTC in 0.5% HCl, which might be explained by its shorter length and different nucleotide content [48]. Complexing TTC to 4 equivalents of either H2B or H4 was found to markedly improve TTC stability in 1% HCl (**FIG. 4**), while the addition of 2 equivalents only to some extent impeded further breakdown of TTC. These results suggest that proteins such as H2B and H4 can protect oligonucleotides towards degradation in acidic media when added at optimal ratios.

In conclusion, POCs and ON-PL showed improved stability in 90% human serum compared to the parent oligonucleotide ON, and covalently attached peptides (PEP3 and PEP8) were found to be more efficient in protecting ON from degradation in human serum than PL in complex with ON. Furthermore, POC18 efficiently degraded when incubated in gastric juice whereas POC13, containing a slightly less cationic peptide, exhibited moderate stability after 24 h incubation period. Inhibition of pepsin by pepstatin A in general caused a boost in stability of oligonucleotide probes. Altogether, POC13 displayed the most promising properties when evaluated in human serum and simulated gastric juice. Lastly, non-covalent complexation of histone proteins (H2B and H4) markedly enhanced stability of TTC in acidic environments, and this effect was dependent on the amount of added protein.

Supplementary Information. IE HPLC, MALDI-TOF MS, and representative radioactive gels are presented in Supplementary Information.

Corresponding Author

*Maria Taskova, e-mail: marita@kemi.dtu.dk

Funding Sources

We appreciate the financial support from the Villum Foundation Young Investigator Programme (grant no. 13152).

Acknowledgment

Prof. Dr. Knud J. Jensen and Dr. Charlotte S. Madsen (University of Copenhagen, Denmark) are acknowledged for

providing the peptide sequences PEP3 and PEP8. Dr. Birte Vester (University of Southern Denmark, Denmark) is acknowledged for the support regarding serum stability studies. Tina Hansen and Prof. Dr. Jesper Wengel (University of Southern Denmark) are acknowledged for assistance with MALDI MS analyses.

Author disclosure statement

The manuscript was written through contributions of all authors. All authors have no conflicts of interests to declare.

References

1. Wang S, N Allen, TA Vickers, AS Revenko, H Sun, X Liang and ST Crooke. (2018). Cellular uptake mediated by epidermal growth factor receptor facilitates the intracellular activity of phosphorothioate-modified antisense oligonucleotides. Nucleic Acids Res 46:3579-3594.
2. Veltrop M, L van Vliet, M Hulsker, J Claassens, C Brouwers, C Breukel, J van der Kaa, MM Linssen, JT den Dunnen, S Verbeek, A Aartsma-Rus and M van Putten. (2018). A dystrophic Duchenne mouse model for testing human antisense oligonucleotides.PLoS One 13:1-18.
3. Paton DM. (2017). Nusinersen: antisense oligonucleotide to increase SMN protein production in spinal muscular atrophy. Drugs of Today 53:327-337.
4. Geary RS, D Norris, R Yu and CF Bennett. (2015). Pharmacokinetics, biodistribution and cell uptake of antisense oligonucleotides. Adv Drug Deliv Rev 87:46–51.
5. Prakash TP. (2011). An Overview of Sugar-Modified Oligonucleotides for Antisense Therapeutics. Chem Biodivers 8:1616–1641.
6. Chan JHP, S Lim and WSF Wong. (2006). Antisense oligonucleotides: from design to therapeutic application. Clin Exp Pharmacol Physiol 33:533–540.
7. Kool ET. (2002). Replacing the Nucleobases in DNA with Designer Molecules. Acc Chem Res 35:936–943.
8. Xu C and J Wang. (2015). Delivery systems for siRNA drug development in cancer therapy. Asian J Pharm Sci 10:1–12.
9. Abes R, HM Moulton, P Clair, ST Yang, S Abes, K Melikov, P Prevot, DS Youngblood, PL Iversen, LV Chernomordik and B Lebleu. (2008). Delivery of steric block morpholino oligomers by (R-X-R)4 peptides: structure–activity studies. Nucleic Acids Res 36:6343–6354.
10. Taskova M, A Mantsiou and K Astakhova. (2017). Synthetic Nucleic Acid Analogues in Gene Therapy: An Update for Peptide-Oligonucleotide Conjugates. ChemBioChem 18:1671-1682.
11. Prakash TP, MJ Graham, J Yu, R Carty, A Low, A Chappell, K Schmidt, C Zhao, M Aghajan, HF Murray, S Riney, SL Booten, SF Murray, H Gaus, J Crosby, WF Lima, S Guo, BP Monia, EE Swayze and PP Seth. (2014). Targeted delivery of antisense oligonucleotides to hepatocytes using triantennary N-acetyl galactosamine improves potency 10-fold in mice. Nucleic Acids Res 42:8796–8807.
12. Taskova M, CS Madsen, KJ Jensen, LH Hansen, B Vester and K Astakhova. (2017). Antisense Oligonucleotides Internally Labeled with Peptides Show Improved Target Recognition and Stability to Enzymatic Degradation. Bioconjug Chem 28:768–774.
13. Liu Y, Y Zhang, H Guo, W Wu, P Dong and X Liang. (2016). Accelerated digestion of nucleic acids by pepsin from the stomach of chicken. Br Poult Sci 57:1–8.
14. Liu Y, Y Zhang, P Dong, R An, C Xue, Y Ge, L Wei and X Liang. (2015). Digestion of Nucleic Acids Starts in the Stomach. Sci Rep 5:674-681.
15. Attarwala H, M Han, J Kim and M Amiji. (2018). Oral nucleic acid therapy using multicompartmental delivery systems. Wiley Interdiscip Rev Nanomedicine Nanobiotechnology 10:1-20.
16. Forbes DC and Peppas NA. (2012). Oral delivery of small RNA and DNA. J Control. Release 162:438–445.
17. Carver JD ans W Allan Walker. (1995). The role of nucleotides in human nutrition. J Nutr Biochem 6:58–72.
18. Tahmasebi-Kohyani A, S Keyvanshokooh, A Nematollahi, N Mahmoudi and H Pasha-Zanoosi, H. (2011). Dietary administration of nucleotides to enhance growth, humoral immune responses, and disease resistance of the rainbow trout (Oncorhynchus mykiss) fingerlings. Fish Shellfish Immunol 30:189–193.
19. McAllan AB. (1980). The degradation of nucleic acids in, and the removal of breakdown products from the small intestines of steers. Br J Nutr 44:99-112.
20. Liu Y, Y Zhang, W Jiang, J Wang, X Pan, W Wu, M Cao, P Dong and X Liang. (2017). Nucleic acids digestion by enzymes in the stomach of snakehead (Channa argus) and banded grouper (Epinephelus awoara). Fish Physiol Biochem 43:127–136.
21. Stein CA and D Castanotto. (2017). FDA-Approved Oligonucleotide Therapies in 2017. Mol Ther 25:1069–1075.
22. Vickers TA and ST Crooke. (2016). Development of a Quantitative BRET Affinity Assay for Nucleic Acid-Protein Interactions. PLoS One 11:1-17.
23. Mihic-Probst D, A Perren, S Schmid, P Saremaslani, P Komminoth and PU Heitz. (2004). Absence of BRAF Gene Mutations Differentiates Spitz Nevi from Malignant Melanoma Anticancer Res 24:2415–2418.
24. Lochmann D, E Jauk and A Zimmer. (2004). Drug delivery of oligonucleotides by peptides. Eur J Pharm Biopharm 58:237–251.
25. Abes R, HM Moulton, P Clair, ST Yang, S Abes, K Melikov, P Prevot, DS Youngblood, PL Iversen, LV Chernomordik, B Lebleu. (2008). Delivery of steric block morpholino oligomers by (R-X-R)4 peptides: structure-activity studies. Nucleic Acids Res 36:6343–6354.
26. Toncheva V, MA Wolfert, PR Dash, D Oupicky, K Ulbrich, LW Seymour and EH Schacht. (1998). Novel vectors for gene delivery formed by self-assembly of DNA with poly(l-lysine) grafted with hydrophilic polymers. Biochim Biophys Acta 1380:354–368.
27. Marino-Ramirez L, KM Levine, M Morales, S Zhang, RT Moreland, AD Baxevanis and D Landsman. (2011). The Histone Database: an integrated resource for histones and histone fold-containing proteins. Database 2011:1-7.
28. Vitravene Study Group. (2002). A randomized controlled clinical trial of intravitreous fomivirsen for treatment of newly diagnosed peripheral cytomegalovirus retinitis in patients with AIDS. Am J Ophthalmol 133:467–474.

29. van Deutekom JC, AA Janson, IB Ginjaar, WS Frankhuizen, A Aartsma-Rus, M Bremmer-Bout, JT den Dunnen, K Koop, AJ van der Kooi, NM Goemans, SJ de Kimpe, PF Ekhart, EH Venneker, GJ Platenburg, JJ Verschuuren and GJB van Ommen. (2007). Local dystrophin restoration with antisense oligonucleotide PRO051. N. Engl. J. Med 357:2677–2686.
30. Goemans NM, M Tulinius, JT van den Akker, BE Burm, PF Ekhart, N Heuvelmans, T Holling, AA Janson, GJ Platenburg, JA Sipkens, JMA Sitsen, A Aartsma-Rus,GJB van Ommen, G Buyse, N Darin, JJ Verschuuren, GV Campion, SJ de Kimpe, JC van Deutekom. (2011). Systemic administration of PRO051 in Duchenne's muscular dystrophy. N Engl J Med 364:1513–1522.
31. Tasfaout H, S Buono, S Guo, C Kretz, N Messaddeq, S Booten, S Greenlee, BP Monia, BS Cowling and J Laporte. (2017). Antisense oligonucleotide-mediated Dnm2 knockdown prevents and reverts myotubular myopathy in mice. Nat Commun 8:1-13.
32. Spillantini MG, A Panconesi, PL Del Bianco and F Sicuteri. (1986). Enkephalinase and angiotensin converting enzyme activities in human venous and arterial plasma. Neuropeptides 8:111–117.
33. Spillantini MG, F Sicuteri, S Salmon and B Malfroy. (1990). Characterization of endopeptidase 3.4.24.11 ("enkephalinase") activity in human plasma and cerebrospinal fluid. Biochem Pharmacol 39:1353–1356.
34. Yi J, C Kim and CA Gelfand. (2007). Inhibition of Intrinsic Proteolytic Activities Moderates Preanalytical Variability and Instability of Human Plasma. J Proteome Res 6:1768–1781.
35. Shaw JP, K Kent, J Bird, J Fishback and B Froehler. (1991). Modified deoxyoligonucleotides stable to exonuclease degradation in serum. Nucleic Acids Res 19:747–750.
36. Sisavath N, L Leclercq, T Le Saux, F Oukacine and H Cottet. (2013). Study of interactions between oppositely charged dendrigraft poly-L-lysine and human serum albumin by continuous frontal analysis capillary electrophoresis and fluorescence spectroscopy. J Chromatogr A 1289: 127–132.
37. Chu BC and LE Orgel.(1992). The stability of different forms of double-stranded decoy DNA in serum and nuclear extracts. Nucleic Acids Res 20:5857–5858.
38. Demidov VV, VN Potaman, MD Frank-Kamenetskii, M Egholm, O Buchard, SH Sönnichsen and PE Nielsen. (1994). Stability of peptide nucleic acids in human serum and cellular extracts. Biochem Pharmacol 48:1310–1313.
39. Zwiers A, B Crusius, G Pals, AJ Donker, SG Meuwissen and RW ten Kate. (1990). Human Pepsinogen A Isozymogen Patterns in Serum and Gastric Mucosa. Gastroenterology 99:1576–1580.
40. Sanny CG, JA Hartsuck and JJ Tang. (1975). Conversion of Pepsinogen to Pepsin. Biol Chem 250:2635–2639.
41. Cornish-Bowden AJ and JR Knowles. (1969). The pH-dependence of pepsin-catalysed reactions. Biochem J 113:369–375.
42. Malamud D and JW Drysdale. (1978). Isoelectric points of proteins: a table. Anal Biochem 86:620–647.
43. Andreeva NS and MN James. (1991). Why does pepsin have a negative charge at very low pH? An analysis of conserved charged residues in aspartic proteinases. Adv Exp Med Biol 306:39–45.
44. Zhang Y, C Li,Y Liu, X Wang, P Dong and X Liang. (2016). Mechanism of extraordinary DNA digestion by pepsin. Biochem Biophys Res Commun 472:101–107.
45. Hammond CM,CB Strømme, H Huang, DJ Patel and A Groth. (2017). Histone chaperone networks shaping chromatin function. Nat Rev Mol Cell Biol 18:141–158.
46. Kamakaka RT and S Biggins. (2005). Histone variants: deviants? Genes Dev 19:295–310.
47. Stenzel S, J Bohrisch and MJ Meyer. (2014). Enhancing ssDNA stability at acidic pH by encapsulation for the usage as DNA marking system. Appl Polym Sci 132:1-7.
48. LeProust EM, BJ Peck, K Spirin, HB McCuen, B Moore, E Namsaraev and MH Caruthers. (2010). Synthesis of high-quality libraries of long (150mer) oligonucleotides by a novel depurination controlled process. Nucleic Acids Res 38:2522–2540.

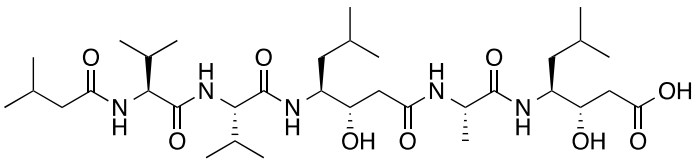
Supplementary Data

Sequences of PEP3, PEP8, TTC, histone proteins, and structure of pepstatin A

Table S1. Sequences of PEP3, PEP8, TTC and histone proteins.

Probe	Sequence
PEP3	$N_3(CH_2)_5C(O)YGGAAKGA\text{-}NH_2$
PEP8	$N_3(CH_2)_5C(O)YKGAKGGK\text{-}NH_2$
TTC	5'-AAT GTT AAT CAC AGC TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT GTG AG-3'
H2B	PEPAK SAPAP KKGSK KAVTK AQKKD GKKRK RSRKE SYSIY VYKVL KQVHP DTGIS SKAMG IMNSF VNDIF ERIAG EASRL AHYNK RSTIT SREIQ TAVRL LLPGE LAKHA VSEGT KAVTK YTSSK
H4	SGRGK GGKGL GKGGA KRHRK VLRDN IQGIT KPAIR RLARR GGVKR ISGLI YEETR GVLKV FLENV IRDAV TYTEH AKRKT VTAMD VVYAL KRQGR TLYGF GG

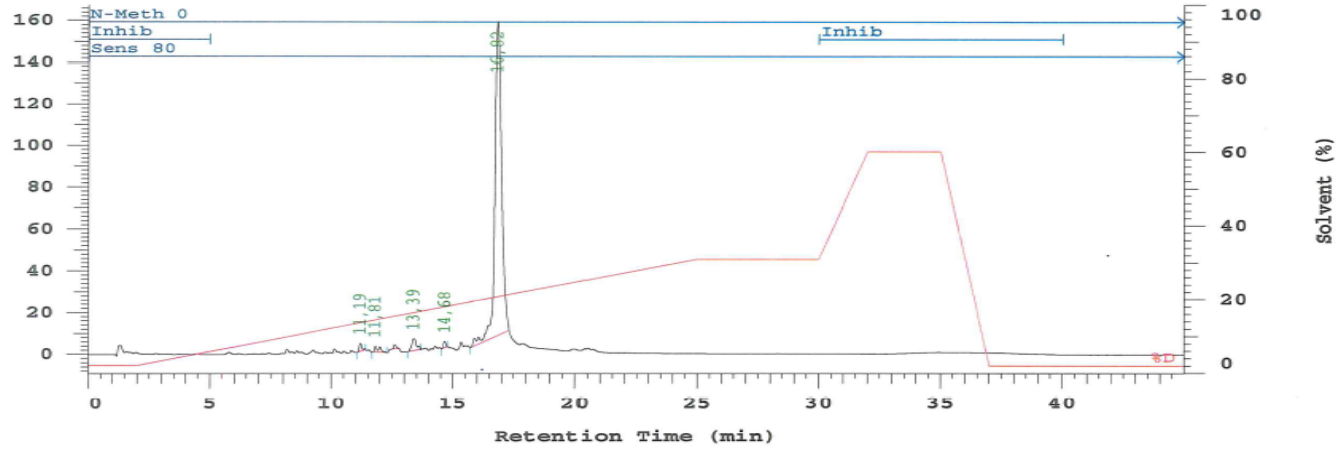
TTC, 50 bp synthetic oligonucleotide-representing BRAF gene fragment.



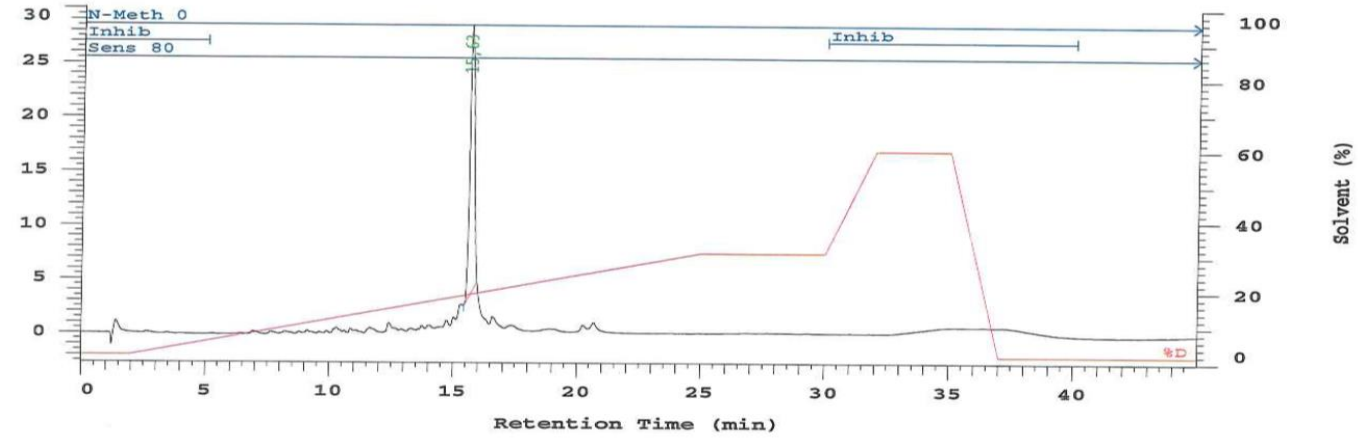
Pepstatin A

Ion-Exchange High-Performance Liquid Chromatography (IE HPLC) Retention Times and Matrix Assisted Laser Desorption Ionization-Time of Flight Mass Spectrometry (MALDI-TOF MS) of the Peptide-Oligonucleotide Conjugates

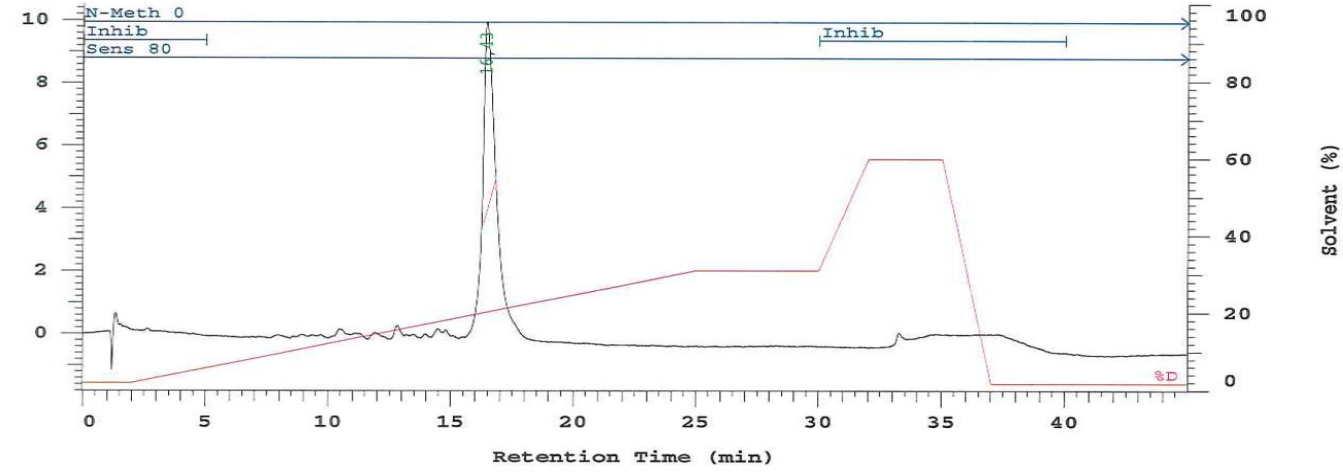
ON



POC13

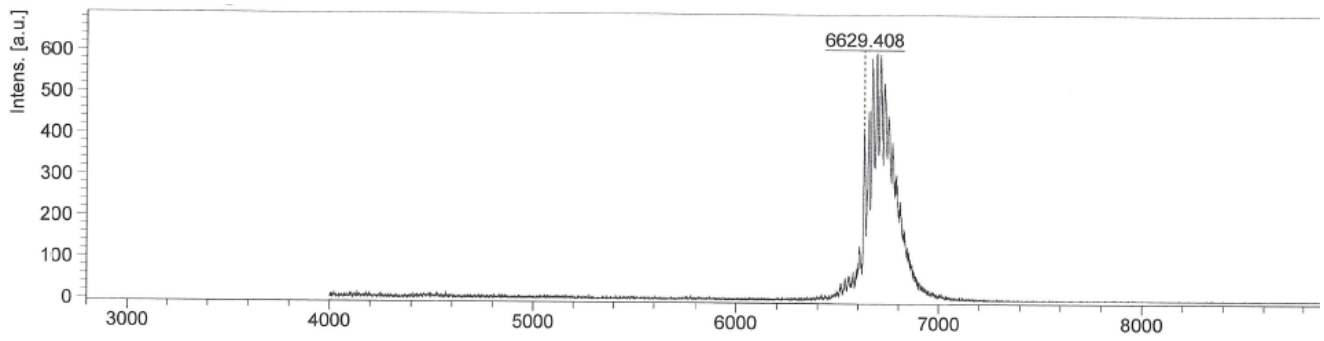


POC18

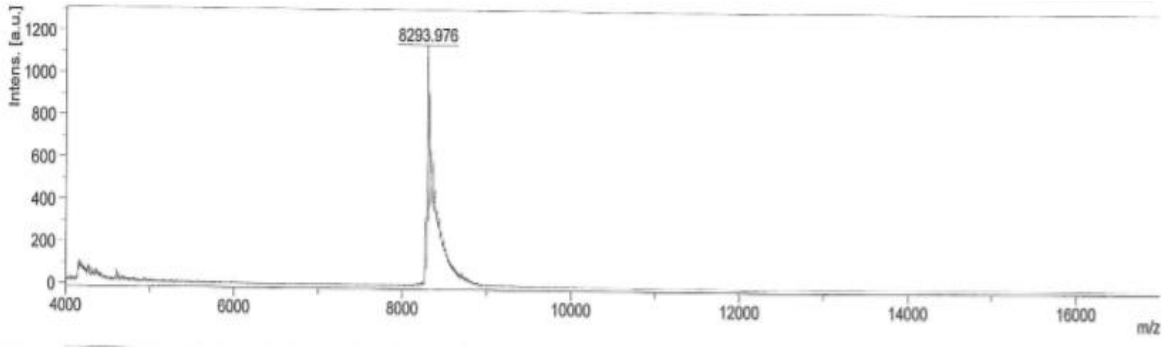


SUPPLEMENTARY FIG. S1. Ion-Exchange High-Performance Liquid Chromatography (IE HPLC) of ON, POC13, and POC18. IE HPLC; ON, 21mer DNA/LNA oligonucleotide that target BRAF V600E mutation in vivo (Fig. 1E); POC, peptide–oligonucleotide conjugate; POC13, POC where the peptide PEP3 is covalently conjugated to ON; POC18, POC where the peptide PEP8 is covalently conjugated to ON.

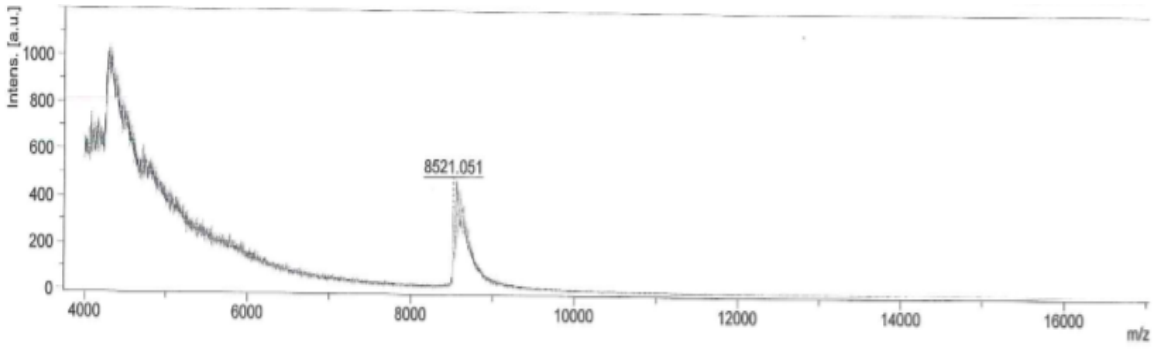
ON



POC13

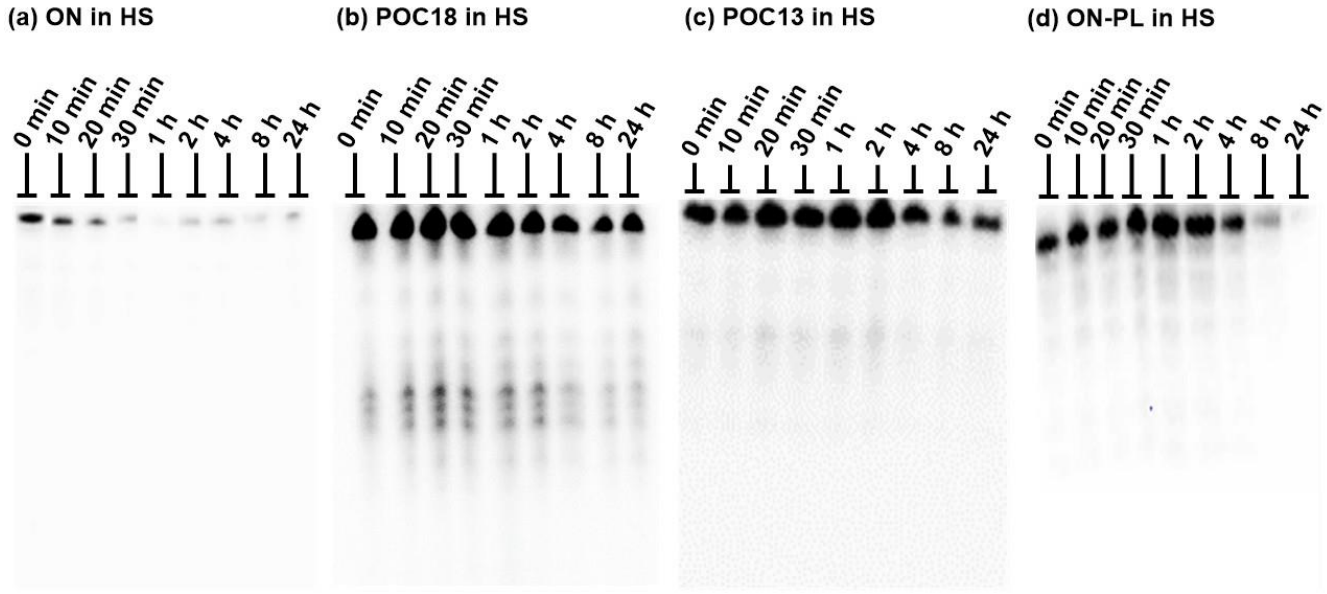


POC18

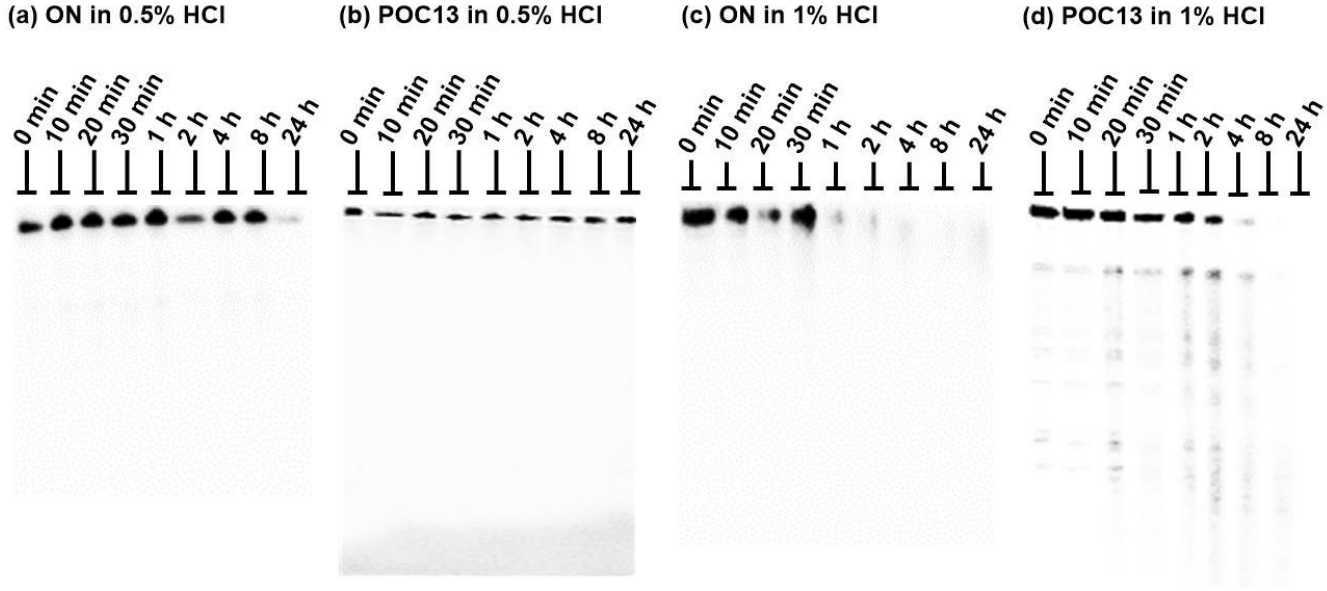


SUPPLEMENTARY FIG. S2. MALDI-TOF MS spectrum of ON, POC13 and POC18. MALDI-TOF MS, matrix assisted laser desorption ionization-time of flight mass spectrometry..

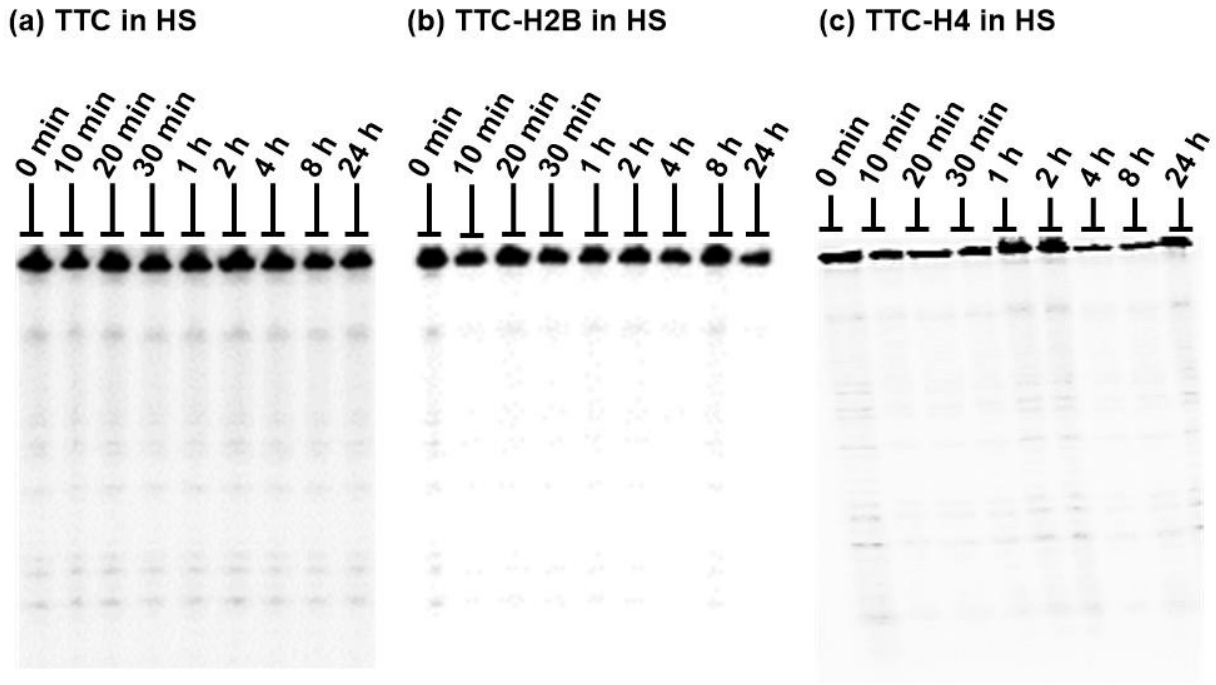
Gel electrophoresis



SUPPLEMENTARY FIG. S3. Representative gel pictures of (A) ON, (B) POC18, (C) POC13, and (D) ON-PL in HS. HS, human serum; PL, poly-l-lysine.

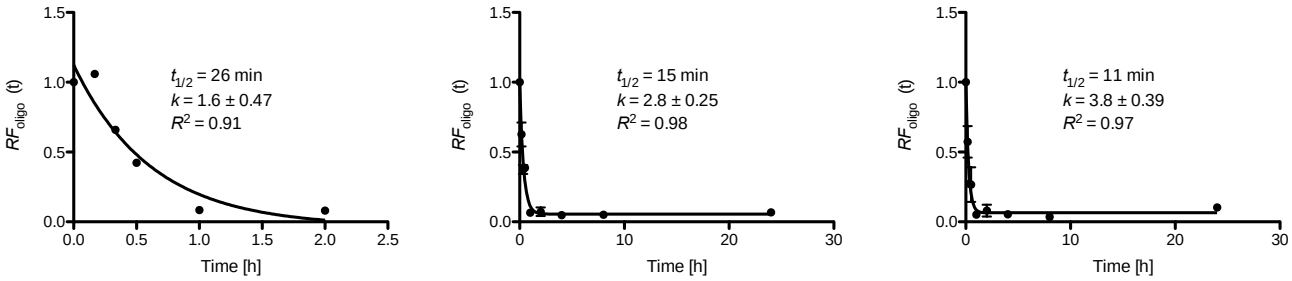


SUPPLEMENTARY FIG. S4. Representative gel pictures of ON and POC13 in (A, B) 0.5% and (C, D) 1% HCl. HCl,hydrochloric acid.



SUPPLEMENTARY FIG. S5. Representative gel pictures of (A) TTC, (B) TTC-H2B and (C) TTC-H4 in HS.

Oligonucleotide digestions in simulated biofluids.

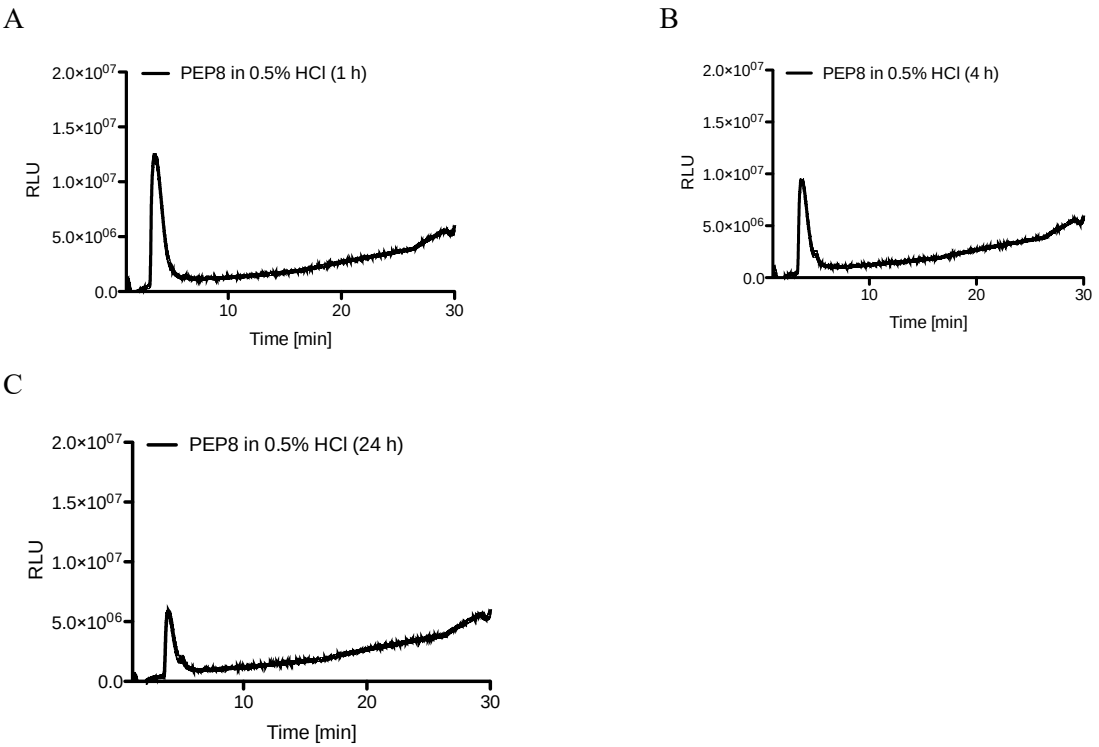


SUPPLEMENTARY FIG. S6. (A) TTC digestion in 1% HCl; TTC levels measured over the course of 2 h were fitted to an exponential one-phase decay model. Data represent mean – SD. (B) TTC-H2B (2 equivalents) digestion in 1% HCl; TTC levels measured over the course of 24 h were fitted to an exponential one-phase decay model. Data represent mean – SD. (C) TTC-H4 (2 equivalents) digestion in 1% HCl; TTC levels measured over the course of 24 h were fitted to an exponential one-phase decay model. Data represent mean – SD.

PEP8 digestion study

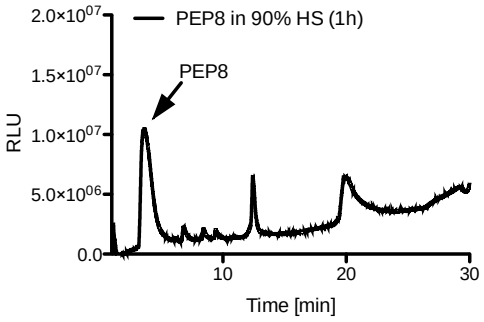
General procedure

Incubation in 90% HS/HEPES and 0.5% HCl was carried out as described in the Material and Methods section in the main article. The samples were taken at certain time points and analyzed by reverse phase HPLC. PEP8 stability studies and control experiments were performed using a reverse phase Waters Alliance 2695 HPLC system using Symmetry® C-18 column 3.5 µm, 4.6 x 75 mm, column temp; 25 °C, (1 ml/min.) with detection at 254 nm using a diode array detector. Eluent A (0.1% formic acid in H_2O) and eluent B (0.1% formic acid in MeCN) were applied in a linear gradient (95% A to 100% B) with a run time of 32 min.

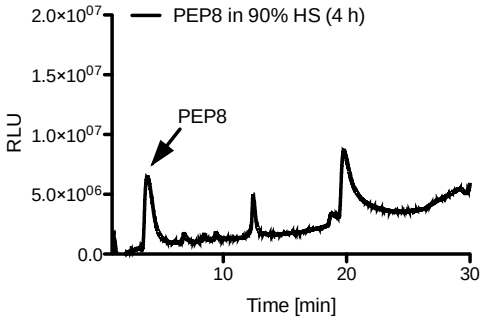


SUPPLEMENTARY FIG. S7. HPLC chromatograms of PEP8 incubated for (A) 1 h, (B) 4 h, and (C) and 24 h at 37 °C..

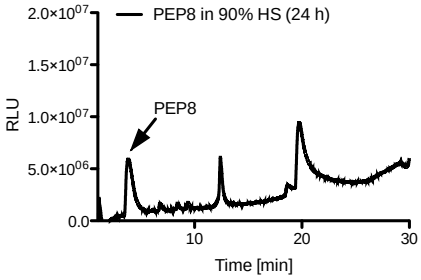
A



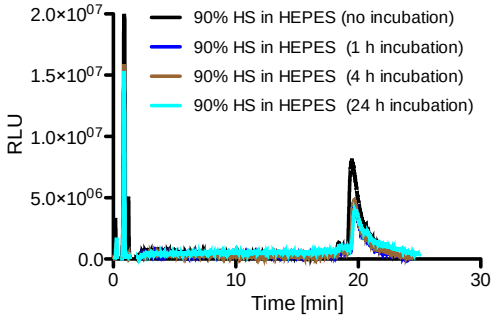
B



C



D



SUPPLEMENTARY FIG. S8. HPLC chromatograms of PEP8 incubated in HS at 37_C for (A) 1 h, (B) 4 h, and (C) 24 h. (D) HPLC chromatograms of HS incubated in HEPES (control) at 37_C for 0 h (black), 1 h (blue), 4 h (brown), and 24 h (cyan) at 37 ºC. RLU, relative light units..

Chapter 5

5.1 siRNA gene therapeutics that target the HIV 1 362 site in the 5'LTR

The Encode project aims to characterize and annotate the human genome. It reveals that more than 80% of the human genome transcribes to RNA. From those genes, only 2% encode for proteins. Encode classifies the other 78% of the RNAs, different from the transfer RNA and the ribosomal RNA as family of noncoding RNA (ncRNAs). They involve microRNAs, small interfering RNAs, PIWI-interacting RNAs and long noncoding RNAs (lnRNAs).[140]

Numerous data report antisense lnRNAs as epigenetic regulators in controlling gene expression.[141] They guide silencing complexes to target loci in the promoters of coding genes resulting in longer duration of gene silencing.[142] This discovery lead to the idea that lnRNAs may be suitable targets to mediate changes in protein coding gene expression.[2]

The number of discovered lnRNAs grows in parallel with the understanding of their involvement in various pathological conditions.[143–145] Following literature, viral latency involves epigenetic mechanisms as well.[146] For instance, ncRNAs from the HIV-1 genome suppress HIV-1 gene expression and replication.

Morris *et al*. confirmed the existence of an HIV-1 encoded lnRNA. They described its location in the nuclei of infected cells in antisense orientation, overlapping the 5' long-term repeat (LTR) and viral mRNA.[142] They also confirmed that this lnRNA transcript is epigenetic regulator of HIV-1 and transcriptional blockage. It acts by guiding chromatin modifying proteins such as, DNA methyltransferase 3a, polycomb protein enhancer of zeste homolog 2 and histone deacetylase 1 to the viral 5'LTR, resulting in histone modifications and formation of heterochromatin. Moreover, Morris *et al*. found that suppression of the HIV-1 antisense ncRNA resulted in activation of latent HIV genes in cell lines and primary human CD4+ T cells.[142]

While antiretroviral therapy successfully increases the lifespan of HIV-1 positive individuals, studies show that HIV-1 positive individuals on antiretroviral therapy are still at risk of developing AIDS.[147,148] This is due to the persistence of infected reservoir cells (resting CD4+ T cells) leading to non-eradicable latent infection. Literature propose therapeutic reactivation of the proviral gene expression, combined with antiretroviral treatments –"shock and kill" strategy as promising strategy to eliminate latent HIV-1 infection.[148]

Weinberg and Morris engineered a CRISPR/Cas9-VPR system guided by several different single guide RNAs (sgRNA) towards the 5'LTR HIV-1 promotor with aim to determine loci that resulted

in activation of HIV-1 gene expression.[149] Using LTRmCherry-IRES-Tat (LChIT) reporter system and set of five sgRNA candidates targeting different region of LTR, they assessed the ability of the system in activation of viral transcription. The sgRNA362f, showed the best properties and was the most promising candidate in activation of latent HIV-1 in seven different *in vitro* models.[149] While this model utilized activation of HIV-1 transcription as its marker, this study together with couple more studies also suggested that this site in the LTR was critical for HIV-1 transcription.[150,151]

Viscidi *et al*. in 1989 described the short cell penetrating peptides (CPPs) and their ability to penetrate cell membranes.[152] Since the discovery, they are among the most explored biomolecules used to enhance the intracellular delivery of various cargoes including therapeutic oligonucleotides. Reports in literature state that the delivery efficacy depends of the nature of the peptide, the amino acid sequence, the size of cargo-peptide probe, peptide concentration and lipid components in the membrane.[153–155]

Based on the physicochemical properties, literature classifies CPPs as cationic, amphipathic and hydrophobic. These differences direct to different mechanism for cell internalization via both energy-dependent and energy independent processes.[156] The mechanisms include direct penetration, endocytosis pathway and translocation by formation of invert micelle[157] (**Figure 5.1**).

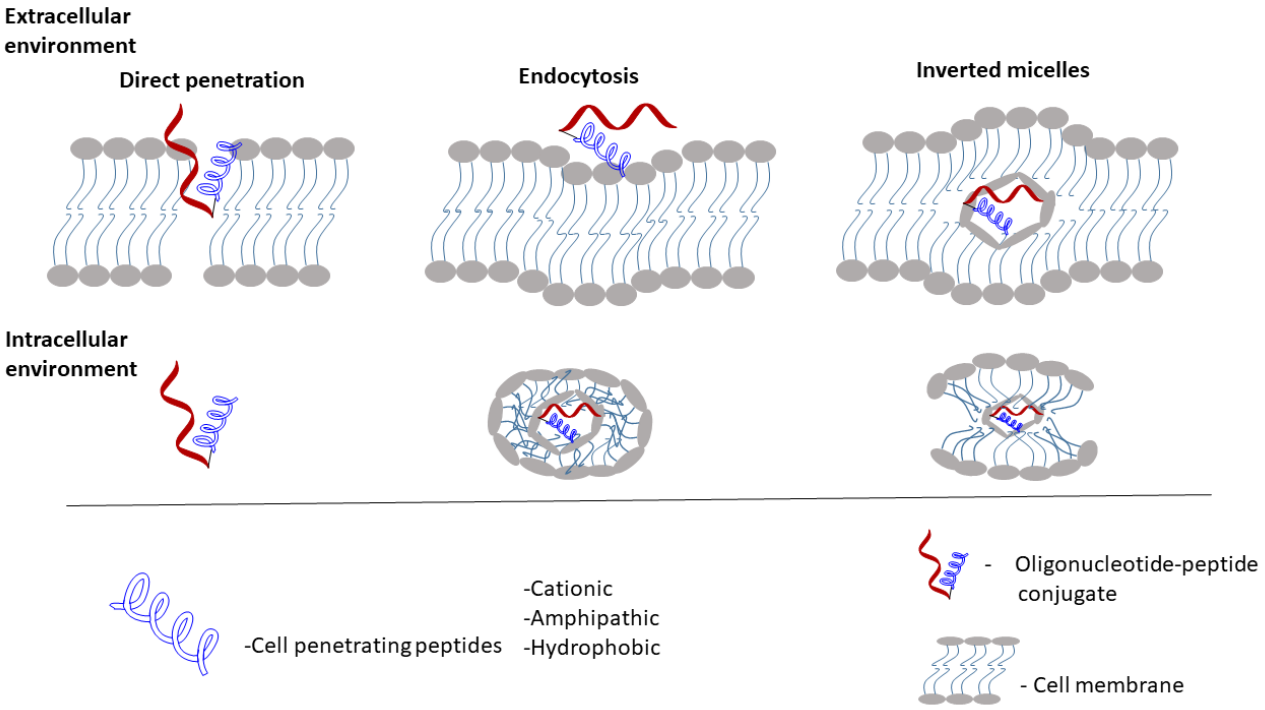


Figure 5.1. Proposed mechanism of intracellular delivery of oligonucleotide-peptide conjugates through cell membrane. Figure inspired from Derakhshankhah and Jafari[155]

Since, just small number peptide sequences show the same features, different peptide sequences display different performance and toxicity levels when engaged with therapeutic oligonucleotides and used in *in vitro* and *in vivo* studies .[158] Therefore, selecting the right amino acid composition within the peptide sequence to be associated with the oligonucleotide, might result in efficient ASOs and siRNA therapeutics.[159]

Scientists explore CPPs to achieve cellular internalization for oligonucleotides as cargos using covalent and non-covalent interactions. The non-covalent approach uses direct mixing of the two biomolecules, whereas, the covalent attachment employs certain type of chemical reaction to make covalent bond. [155]

In the present work, we formulate siRNA gene therapeutics that target the HIV-1 362 site in the 5'LTR. We develop a siRNA therapeutic formulation, complex and conjugate with peptides with aim to enable efficient intracellular and intranuclear delivery without the use of toxic transfection reagents. Moreover, we investigate the therapeutic efficiency of the formulations siRNAs to inhibit active HIV-1 transcription in a model cell lines.

Materials and Methods

siRNA nanoplexes formulation

The RNA oligonucleotides were purchased from Integrated DNA Technologies (IDT). siRNAs were prepared by annealing the sense (s) and antisense (as) strands (20 nmol of each) in 1xPBS buffer. First, we activated both sequences, s/as at 85 ^{o}C for 10 min with subsequent graduate cooling down to room temperature for 1 h for hybridization and duplex formation followed by duplex stabilization at +4 ^{o}C for 1 h. Next, the peptide/GalNAc (30 nmol or 60 nmol) was added and the mixture was incubated at rt for 1 h.

siRNA covalent conjugation and characterization

The RNA oligonucleotides were obtained from IDT. The sense strands of the three-siRNA probes were designed to contain alkyne functionality on the 5' termini. Next, all three strands (S2, S4 and SCS2) were covalently conjugated to azide modified Tat1a peptide (**Table 5.1**). The CuAAC "click" reaction was performed under inert atmosphere using 20 nmol of the alkyne modified

sense strand (S2,S4 or SCS2), 4 eq or 80 nmol azide Tat1a peptide, 1M TEAA buffer pH 7.3, aminoguanidine hydrochloride (50 mM 4 μM), formamide (7%, 14 μL), CuTHPTA (10 mM; 5 μL) and a fresh solution of ascorbic acid (50 mM; 2 μL) in total reaction volume of 200 μL. The reaction mixture was kept under steering for 12 h with subsequent desalting on a NAP-5 Sephadex column (Illustra; GE Lifescience).

Purification and analysis were done by ion exchange (IE) HPLC, on a Dionex UltiMate 3000 LC System using 0.25 M $NaClO_4$ and 0.25 M Tris in 20 % MeOH in MilliQ water. The purity of S2-Tat1a, S4-Tat1a and SCS2-Tat1a after purification has been confirmed by the analytical IE HPLC, and was found to be 96.34%-99.88% (Appendix A, Fig. A2).

The identity of oligonucleotides have been confirmed by MALDI-TOF mass spectrometry, performed on a Ultraflex II TOF/TOF instrument, Bruker, using 3-hydroxy picolinic acid matrix (10 mg 3-hydroxypicolinic acid, 50mM ammonium citrate in 70% aqueous acetonitrile). MS calcd./MS found values ($[M-H]^+$) were 7423.2/7426.0, S2-Tat1a; 8035.5/8037.0, S4-Tat1a; 7423.2/7429.0, SCS2-Tat1a (Appendix A, Fig. A1).

Cell lines

The acute lymphoblastic leukemia cell line, CEM (ATCC VA, USA) and LChIT (CEM cell based reporter system consist of single copy of HIV that encodes mCherry IRES-Tat from the full-length HIV LTR)[149] cell lines were maintained in RPMI-1640 media (ATCC VA, USA) that contain 10% fetal bovine serum (GeminiBio CA, USA) at 37°C in a water-jacket incubator, passaged twice weekly. Cell viability was checked by the use of Trypan blue ensuring viability was above 90% before starting the experiment. The cells were plated at density of 100000 cells per well/48 well plate in total volume of 250 μL. Positive controls were transfected by electroporation using Neon Transfection Device (Life Technologies) using one pulse with voltage-1230 V and pulse width-40 ms.

The human embryonic kidney (HEK-pMO-GFP) cell line (ATCC VA, USA) was maintained in Eagle's Minimum Essential Medium(ATCC VA, USA) supplemented with 10% fetal bovine serum (GeminiBio CA, USA) at 37°C in a water-jacket incubator, passaged twice weekly. The cells were plated at density of 75000 cells per well/24 well plate in a total volume of 250 μL. Positive controls were transfected using 3 μL RNAiMAX per condition.

The epithelial cells (TZMbl) were cultured in dulbecco modified eagle medium (ATCC VA, USA) supplemented with 10% fetal bovine serum (GeminiBio CA, USA) at 37°C in a water-jacket incubator, passaged twice weekly. The cells were plated at density of 50000 cells per well/12 well plate in a total volume of 500 µL. Cells were transfected the next day using Lipofectamine 3000 (Thermo Fisher Scientific, according to manufacturer's instructions). The plasmid pcDNA-TAT-dsRED was transfected at 100 ng/well. Expression of the HIV-1 TAT protein leads to activation and sustained HIV-1 transcription, by binding to the TAR region in the 5'LTR of HIV-1. For the dual luciferase experiments, additional transfection of a control plasmid, pcDNA-Renilla at 100 ng/well, was used to measure transfection efficiency per well.

Flow cytometry, luciferase assay and dual luciferase assay

72 h/ 48 h after siRNA transfection the media of the cells was replaced with 1xPBS and mCherry/GFP expression was analyzed by flow cytometry (FACSCalibur II, Becton, Dickinson and Company, East Rutherford, NJ), with 25000 events counted per sample. Data was analyzed using FlowJo, LLC (version 10.1, Ashland, OR), where percentage GFP expression was determined.

Luciferase reporter assay (48 h after transfection) was performed to measure luminescence on a luminometer (GloMax Explorer, Promega). The media of the cells was discarded followed by a subsequent wash with 1xPBS and lidding the adherent cells with 120 µL luciferase assay reagent (LAR) by Promega. The measurements were performed in a 96 well plate using 100 µL of the lysed cell mixture.

For the dual luciferase assay (48 h after transfection) cells were washed with 1xPBS before the addition of 100 µL 1x Passive Lysis buffer. Samples were incubated for 20 min at rt on a shaker. Thereafter, 25 µL of each well was transferred to a 96 well plate, followed by addition of 50 µL luciferase assay reagent II and reading on a luminmoeter (GLoMax Explorer, Promega) of the firefly luciferase (fLUC). Next, 50 µL 1x Stop&Glo reagent was added and samples were read again for the renilla luciferase (rLUC). The results were obtained dividing the fLUC by their corresponding rLUC values. This was done to correct for the transfection efficiency in each well.

Design and synthesis

We designed two siRNAs (S1/AS1 and S3/AS2) targeting the HIV-1 LTR 362 site and a control scrambled siRNA (SCS1/SCAS) - **Table 5.1**.

Table 5.1. System siRNAs designed; Sequences of siRNA sense and antisense strands and peptide sequences.

Name	Sequence
S1	5' -r(CUU UCC GCU GGG GAC UUU C)
AS1	5' -r(GAA AGU CCC CAG CGG AAA G)
S2	5' -Hexynyl-r(CUU UCC GCU GGG GAC UUU C)
S3	5'-r(UUC UUU CCG CUG GGG ACU UUC)
AS2	5'-r(GAA AGU CCC CAG CGG AAA G UU)
S4	5'-Hexynyl-r(UUC UUU CCG CUG GGG ACU UUC)
SCS1	5'-r(UCG CUU AGC GUU CUC GCG U)
SCAS	5'-r(CGC GAG AAC GCU AAG CGA)
SCS2	5'-Hexynyl-r(UCG CUU AGC GUU CUC GCG U)
Peptides/GalNAc:	
TatO	AEDGY
Tat1	GRKKRRQRRR
Tat1a	5-azidopentanoic acid-GGKKRRQKGR
GalNAc	α-GalNAc-PEG3-Azide

First, we formulated the siRNAs and the scrambled control siRNA as nanoplexes. All three siRNA probes were complexed with TatO, Tat1 peptides and GalNAc molecule as well. Using the electrostatic or columbic interactions between them, we made complexation by direct mixing of both biomolecules (**Figure 5.2**).

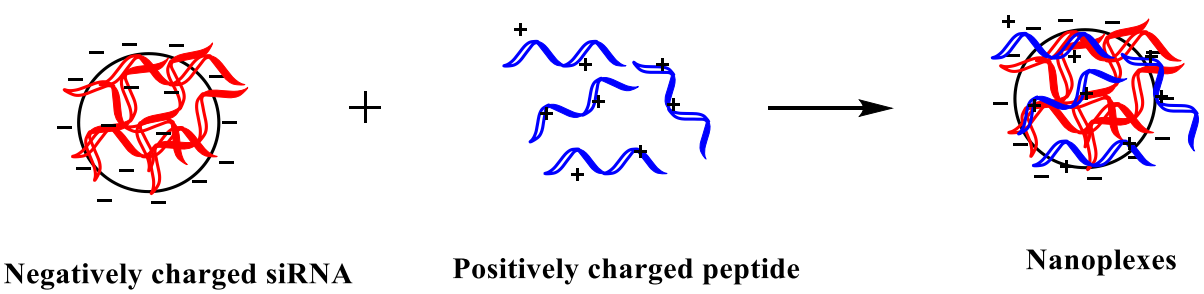


Figure 5.2. Formation of nanoplexes by two polyelectrolytes, anionic siRNA and cationic peptide

We formulated set of 21 probes (**Table 5.2**) using two different concentrations of the peptides/GalNAc molecules. We tested the probes in cell culture studies at the Center for Gene Therapy, COH, USA.

Table 5.2. Nanoplexes targeting LTR 362 HIV1

Name	**siRNA, 20 nmol**	**Peptide/GalNAc**
L1	S1+AS1	/
L2	S3+AS2	/
L3	SCS1+SCAS	/
		30 nmol
L4	S1+AS1	TatO
L5	S3+AS2	TatO
L6	SCS1+SCAS	TatO
L7	S1+AS1	Tat1
L8	S3+AS2	Tat1
L9	SCS1+SCAS	Tat1
L10	S1+AS1	GalNAc
L11	S3+AS2	GalNAc
L12	SCS1+SCAS	GalNAc
		60 nmol
L13	S1+AS1	Tat0
L14	S3+AS2	Tat0
L15	SCS1+SCAS	Tat0
L16	S1+AS1	Tat1
L17	S3+AS2	Tat1

L18	SCS1+SCAS	Tat1
L19	S1+AS1	GalNAc
L20	S3+AS2	GalNAc
L21	SCS1+SCAS	GalNAc

Results and discussion

Based on the preliminary data from Morris *et al*. described above, we design the S1/AS1 and S3/AS2 siRNAs targeting the HIV-1 LTR 362 site and control scrambled siRNA SCS1/SCAS. Both, S1/AS1 and S3/AS2 target the same region of the LTR 362 with a difference that S3/AS2 contains UU overhangs on the both, 3' and 5' termini. Literature describes that siRNA having two UU overhangs on the sense strand silence the genes more effective *in vivo* and in tissue culture studies.[160] Accordingly, we set an objective to investigate if the UU overhangs will show an effect in our model system. Next, literature propose short arginine and lysine peptides as one of the most fortunate tools for intracellular delivery of large cargoes such as biomolecules. They enter the cell by energy independent (membrane destabilization) or clathrin-mediated endocytosis.[161,162] We use the peptide sequences, TatO and Tat1 in two distinct quantities (30 nmol and 60 nmol) to prepare a nanoplexes and to investigate their performance in three different cell lines. We also include a mixture of the siRNAs with GalNAc molecules.

First, we examined the delivery and the efficacy of the nanoplexes in CEM and LChIT cell lines. LChIT is reporter cell line (developed from a CEM parental cell line) that stably expresses an LTR driven mCherry gene. This reporter system works in a way that mCherry gene expression will lead to mCherry fluorescent protein manifestation, and contrary, silencing of the gene will decrease the presence of the mCherry fluorescent protein. We assessed the efficiency of the siRNA drug candidates to enter the cells and to block the mCherry expression by fluorescence-activated cell sorting (FACS). We used CEM cell line as a negative control. CEM do not express the LTR driven mCherry fluorescent protein expression.

We performed a dose response with several siRNA nanoplexes, L1, L3, L4 and L6 (from **Table 5.2**), each in seven different concentrations (15 nM, 25 nM, 50 nM, 100 nM, 200 nM, 300 nM and 500 nM concentration), in duplicates. After 72 h, we observed that none of the nanoplexes at any of the used concentrations reduces the expression of mCherry. The CEM cell line showed

negative expression and LChIT cell line showed positive. As seen in **Figure 5.3**, the mCherry expression remained unchanged in the control versus the treated samples.

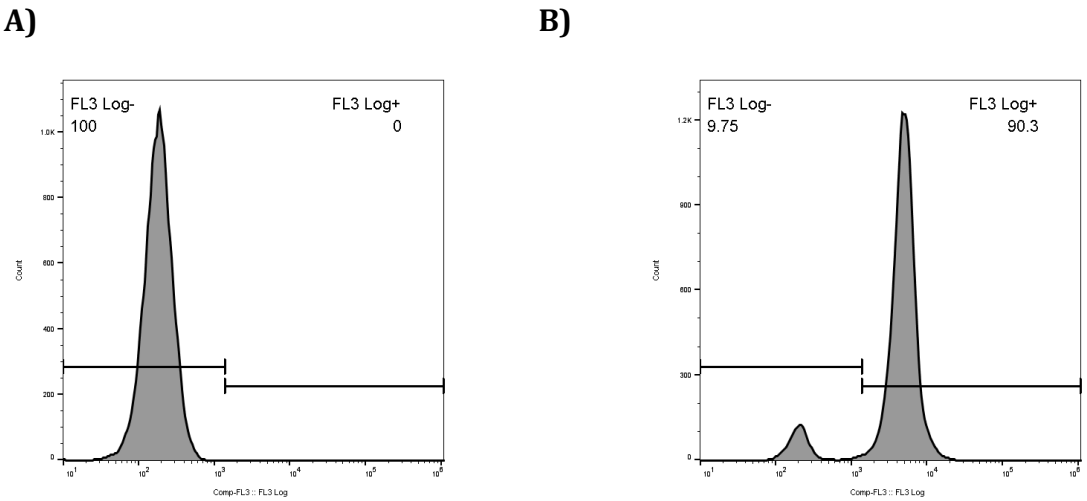


Figure 5.3. CEM cell line as negative control that do not express mCherry (A) and LChIT cell line that express mCherry (B). Nanoplexes have no effect in LChIT cells. 400000 cells/ml CEM/LCHITs were seeded and 15-300 nM nanoplexes L1, L3, L4 and L6 were dropped onto the cells. 72 h later, mCherry expression was measured by FACs.

Next, using the same experimental set up, we tested all 21 nanoplexes using 500 nM and 1000nM concentrations. L1-naked siRNA probe (50 nM concentration) was additionally transfected by electroporation. Assessing the flow cytometry measurements, the nanoplexes failed to reduce mCherry expression even for the positive – L1 control. Finally, we tested just the naked siRNA samples L1-L3 in three different concentration (500 nM, 1000 nM and 2000 nM) transfected by electroporation using Neon Transfection Device. Unfortunately, the probes showed no reduction of the mCherry fluorescence signal. We terminated this experimental setting noticing that the designed siRNAs fail to block the LTR driven mCherry gene expression in the developed LChIT cell line system.

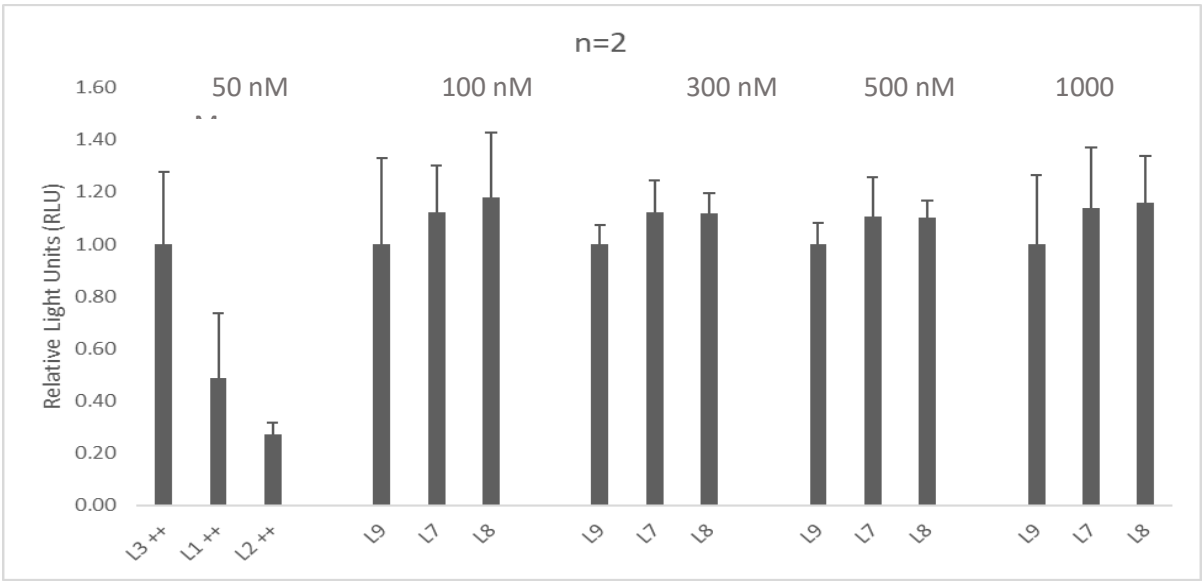
Further, we tested the nanoplexes using HEK-pMO-GFP cell line. They stably express LTR driven GFP reporter. Performing flow cytometry measurement, we observed just 8% reduction of the expression of the GFP (green fluorescent protein – associated with the LTR gene in the Hek-pMo-GFP cells) for the L1 and L2 positive controls and no significant reduction from the nanoplex probes L13-L18. We ended this experimental set up opining that the designed siRNAs do not significant silence the GFP gene in Hek-pMo-GFP cells.

Next, we tested the nanoplexes using epithelial cells (TZMbl). They contain a stably integrated reporter gene that expresses luciferase when activated with TAT. The activation of HIV-1 mRNA expression by the HIV-1 TAT protein, works through the TAT protein binding to the TAR loop of the LTR promoter and activates gene expression.[163] The luciferase assay uses the oxidative enzyme, luciferase, that upon a highly energetic chemical reaction with the substrate, luciferin, releases energy in form of light.[164]

First, we screened all nanoplexes L1-L21 in 50 nM concentration measuring the luminescence. (**Appendix A, Chart A1-a**). Notably there was no significant reduction of the luciferase expression. In the next round experiments, we included L1 and L3 naked as positive controls and their nanoplexes with TatO and Tat1 in three different concentrations (50 nM, 100 nM and 200 nM) – (**Appendix A, Chart A1-b**). The data was normalized according the scrambled control L3. Again, we observed slight and variable reduction of the luciferase activity.

Further, we examined L1-L3 (50 nM) as positive controls and nanoplexes L7-L9 in 100-1000 nM concentration - **Chart 5.1**. In this stage, we discarded the mixtures with GalNAc from further examination. Notably, the positive controls L1 and L2 on 50 nM concentration showed a 50% and 70% reduction of the luciferase expression, respectively. Nanoplexes L7-L9 (complexes with 30 nM Tat1) failed to show reduction in the luciferase activity.

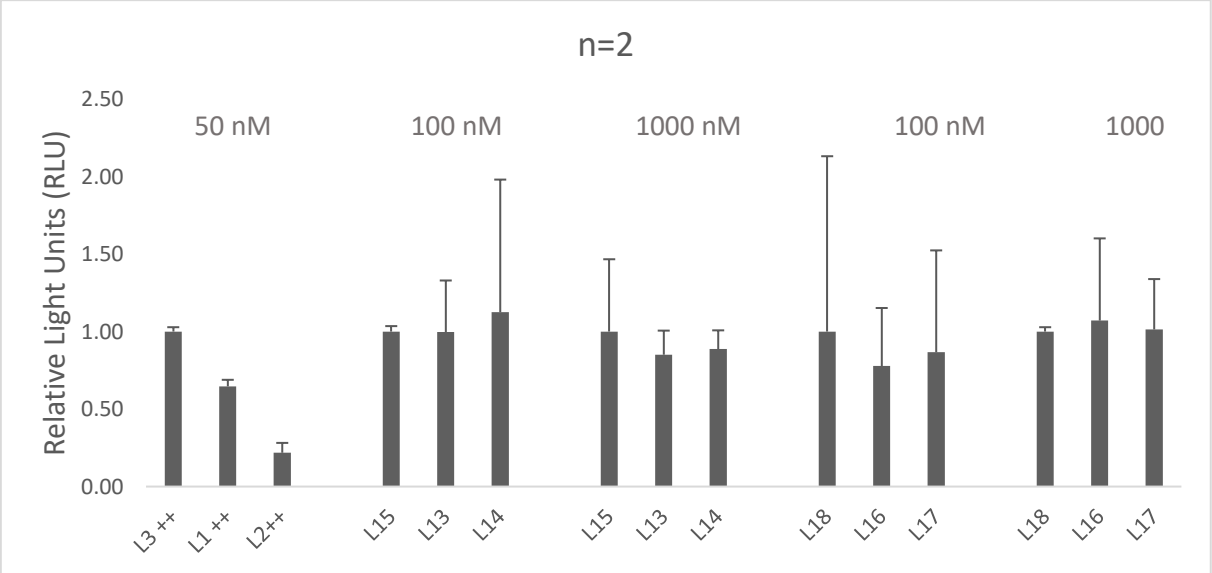
Chart 5.1. Luciferase reporter assay results; RLU representing the expression of the luciferase gene in epithelial cells TZMbl cells



* 100000 cells/ml TZMbl were seeded and 24 h latter, 100 ng/well pcDNA-TAT-dsRED was transfected to activate HIV-1 transcription. 24 h after, 50 nM nanoplexes L1-L2 were transfected using Lipofectamine 3000 (++). In addition, 100-1000 nM nanoplexes L7-L9 were dropped onto the cells. 48 h later, luminescence was determined.

Next, we utilized a new cell stock of TZMbl cell line, as a control to make sure the results observed were not due to changes in the cell line itself due to over passaging. First, we screened nanoplexes L13-L18 at three different concentrations (100 nM, 500 nM and 1000 nM). We transfected L1-L3 at 50 nM concentration as the positive controls (**Appendix A, Chart A1-c**). We again observed, promising results from the L2 probe with a 70% reduction in luciferase activity. However, the L1 probe did not show any effect in this experiment. From the nanoplexes, L13-1000 nM and L16-100nM and -1000 nM concentration showed promising results reducing the luciferase activity for approx. 50% in this screening experiment. Both probes (L13 and L16) are equivalents of L1 complexed with TatO and Tat1 (60 nmol), respectively. Further, we repeated these measurements using the same probes L13-L18 just in two concentrations (100 nM and 1000 nM), in duplicates. Notably, **Chart 5.2**, we replicated the previous results for the positive controls L1 and L2. However, the nanoplexes showed significant standard deviation (SD) and CV values (ratio of the standard deviation to the mean) of 2.8% - 113.1% between the duplicates.

Chart 5.2. Luciferase reporter assay results; RLU representing the expression of the luciferase gene in epithelial cells TZMbl cells



* 100000 cells/ml TZMbl were seeded and 24 h latter, 100 ng/well pcDNA-TAT-dsRED was transfected to activate HIV-1 transcription. 24 h after, 50 nM nanoplexes L1-L3 were transfected using Lipofectamine 3000 (++). In addition, 100-1000 nM nanoplexes L13-L18 were dropped onto the cells. 48 h later, luminescence was determined.

Next, because of the high degrees of variation between the data, we proceeded with dual luciferase experiments. In this assay, we measured the activities of firefly (Photinus Pyralis) and

Renilla (Renilla Reniformis, also known as sea pansy) luciferases. The renilla luciferase measurements indicate the transfection efficiency for each sample, and thus provide a measurement of normalization for each sample. We performed the assay two distinct times, adding each nanoplex probe (L13-L18, 1000 nM) and positive controls (L1-L3, 50 nM) in duplicates/triplicates (**Appendix A, Chart A2-a** and **-b**). Notably the results, L1 and L2 positive controls reduce the luciferase signal for approximately 50% and 70%, respectively. Nanoplexes again show variable and not reproducible results with high CV values.

Conclusion

We designed and prepared nanoplex formulations of siRNAs that target HIV1 LTR 362 with two different types of peptides and GalNAc. The efficiency of all nanoplexes was examined in three different cell lines LChIT, Hek-pMoGFP and TZMbl using appropriate assays for quantification of the signal/effect. The designed siRNA probes showed inefficient blockage of expression of the LTR gene in the LChIT cells and insignificant effect in Hek-pMoGFP cells as well. However, in TZMbl cells the siRNAs showed promising results reducing the expression of the LTR gene for 50% and 70% for the L1 and L2 (positive controls) siRNAs, respectively. On the other hand, the delivery of the siRNAs as nanoplexes with the peptides was inconstant and with high variability within the data. Even though L13 and L16 - siRNA probes complexed with 60 nmol TatO or Tat1 were among the most promising candidates, the results were not reproducible. Overall, the complexation of therapeutic siRNA with peptide sequences have effect on the intracellular delivery to some extent. Nevertheless, this effect may highly depend on the uniformity and the stability of the nanoplexes. In addition, many factors as peptide sequence and concentration, storage, temperature variations and mechanical stress most likely affect the formulation. Finally, the efficiency for intracellular delivery of the nanoplexes may be highly dependent on the cell type.

Further work, Synthesis - Covalent Conjugation

To prepare more auspicious set of samples, we designed the sense strands of the three siRNAs to be alkyne functionalized and subsequently covalently conjugated with the azide cell penetrating-Tat1a peptide – **Table 5.1**. We performed CuAAC reaction, previously optimized by us for DNA oligonucleotides. However, working with pure RNA sequences was more challenging.

We did not observed degradation of the RNA in the reaction conditions; however, we observed collecting lower amounts of RNA after IC HPLC purification. Usage of 7% formamide was crucial for the reaction to happen preventing precipitation of the Tat1a peptide. The conjugates were characterized by Maldi-TOF and IC HPLC (**Appendix A, Figure A1 and A2**)

We speculate that the conjugated siRNA probes having well defined structure will perform better activity in cell culture studies and *in vivo* compared with the nanoplexes. Roslyn Ray, PhD, will test the probes in the Morris laboratory at COH following the optimized conditions for the cell lines - TZMbL previously described.

5.2 Repression of lnRNA (BGas) for Cystic Fibrosis Transmembrance Conductance Regulator expression

Cystic fibrosis is a life threatening disease. The cystic fibrosis foundation patient's registry contains more than 70000 patients registered worldwide.

It is a multifaceted disorder associated with mutation in the cystic fibrosis transmembrane conductance regulator (*CFTR*) gene that also affect the activity of the encoded ion channels on the epithelial cell surface.[165–167] Morris *et al*. discovered new lnRNA named BGas, antisense to the protein coding sense strand in relationship with gene suppression. They indicated that blockage of the BGas lead to *CFTR* expression and restoration of the chloride ion function.[168]

In the present work, we design siRNA and ASO gene therapeutics directed to silence the BGas. We aim to investigate their potential to restore the *CFTR* expression. In addition, we enhance the gene therapeutics modifying them as a conjugates with peptide sequence or making formulation as nanoplexes with peptide sequence. We investigate the efficiency for intracellular and intranuclear delivery in adenocarcinomic human alveolar basal epithelial (A549) cells.

Materials and methods

siRNA covalent conjugation

The 5' hexynyl and the 5' amino modified RNA oligonucleotides were obtained from IDT. The two RNA sense strands (BGas s and BGas SC s) and the three DNA ASO sequences (Gap, Gap A and Gap SC) were covalently conjugated to azide modified Tat1a peptide (**Table 5.3**). The CuAAC "click" reaction was performed under inert atmosphere using 20 nmol of the alkyne modified oligonucleotides, 4 eq or 80 nmol azide Tat1a peptide, 1M TEAA buffer pH 7.3, aminoguanidine hydrochloride (50 mM 4 μM), formamide (7%, 14 μL), CuTHPTA (10 mM; 5 μL) and a fresh solution of ascorbic acid (50 mM; 2 μL) in total reaction volume of 200 μL. The reaction mixture was steered overnight at room temperature with subsequent desalting on a NAP-5 Sephadex column (Illustra; GE Lifescience). For the stain-promoted azide-alkyne cycloaddition (SPAAC) conjugated Gap A, first the amino oligonucleotide (20 nmol dissolved in 180 μL, 0,1 M PBS, pH = 8.5) was reacted with dibenzocyclooctyne (DBCO) NHS ester – Lumiprobe (160 nmol dissolved in 20 μL DMSO) for 4 h at rt followed by +4 ^{o}C overnight. Next, the product was desalted on a NAP-5 Sephadex column and used as that in the SPAAC reaction. The Gap A-DBCO (20 nmol) was added into solution of azide Tat1a peptide (25 nmol) in total volume of 200 μL (DMSO:H_2O, 1:1).[169] The reaction was carried out under microwave irradiation using Biotage initiator

microwave at 60 °C for 3 h followed by subsequent desalting on a NAP-5 Sephadex column run twice.

Purification and analysis were done by ion exchange (IE) HPLC, on a Dionex UltiMate 3000 LC System using 0.25 M $NaClO_4$ and 0.25 M Tris in 20 % MeOH in MilliQ water. The purity of BGas s-Tat1a, BGas SC s-Tat1a, Gap-Tat1a, Gap A-Tat1a and Gap SC-Tat1a after purification has been confirmed by the analytical IE HPLC, and was found to be 97.98-99.86% (Appendix A, Fig. A4).

The identity of oligonucleotides have been confirmed by MALDI-TOF mass spectrometry, performed on a Ultraflex II TOF/TOF instrument, Bruker, using 3-hydroxy picolinic acid matrix (10 mg 3-hydroxypicolinic acid, 50mM ammonium citrate in 70% aqueous acetonitrile). MS calcd./MS found values ($[M-H]^+$) were 8810.2/8803.0, BGas s-Tat1a; 8810.2/8812.0, BGas SC s-Tat1a; 7658.8/7653.0, Gap-Tat1a; 8048/8044, Gap A-Tat1a 7658.8/7657.0, Gap SC-Tat1a (Appendix A, Fig. A3).

siRNA nanoplexes formulation

The siRNAs were prepared by annealing the sense and antisense strands (20 nmol of each) in 1xPBS buffer. They were first activated at 85 °C for 10 min with subsequent graduate cooling down to room temperature for 1 h followed by stabilization of the duplex siRNA at +4 °C for 1 h. Lastly, the peptide in certain amount (100 nmol, 30 nmol or 60 nmol) was added and the mixture was incubated at rt for 1h.

Cell line culture

We tested the probes in adenocarcinomic human alveolar basal epithelial (A549) cell line (ATCC – CCL 185, USA). Cells were cultured in Kaighn's Modification of Ham's F-12 Medium (ATCC VA, USA) supplemented with 10% fetal bovine serum (GeminiBio CA, USA) at 37 °C in a water-jacket incubator, passaged twice weekly. The cells were plated at density of 100000 cells/48 well plate in a total volume of 250 µL and supplemented the next day with the Gap probes (G1-G4).

RNA isolation, reverse transcription and qPCR

72 h after transfection, the total RNA was isolated with Maxwell® RSC simply RNA Cells Kit (Promega) following the standard protocol from the manufacturer. The amount and purity of total RNA were measured with NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies Inc.,Wilmington, DE, USA).

Next, 0.1-1 g of the RNA was reverse transcribed into cDNA using QuantiTect® Reverse Transcription kit (Qiagen, Hilden, Germany). In particular, the genomic DNA was eliminated using gDNA wipeout buffer 1 µL, template RNA 1µL and RNase free water 5 µL. The mixture was incubated for 3 min at 42 ^{o}C and placed immediately on ice. Next reverse-transcription mix was prepared to contain: quantiscript reverse transcriptase 0.5 µL, quantiscript RT buffer 5x 2 µL, RT primer mix 0.5µL and template RNA (entire genomic DNA elimination reaction, from step one) 7 µL. The mixture was incubated for 30 min at 42 ^{o}C, followed by incubation for 3 min at 95 ^{o}C to inactivate the transcriptase.

Quantitative measurements of RNA levels were done by real time PCR using KAPA SYBR® FAST qPCR Master Mix (2X) Kit (Kapa Biosystems, Wilmington, MA, USA). The PCR mix contained KAPA SYBR FAST qPCR Master Mix (2X) 10 µL, forward CFTR primer (5'-TGCAAACGTAACAGGAACCCGACT) 4 µM/mL, 2.5 µL, reverse CFTR primer (5'-TCTTTAGGTCCAGTTGGCAACGCT) 4 µM/mL, 2.5 µL, PCR-grade water 3 µL and cDNA template 2 µL (from the 10 µL step above). The qPCR was executed in a 96 well plate with the Roche Lightcycler instrument (Roche, Basel, Switzerland). Thermal cycling parameters started with 3 minutes at 95 °C, followed by 40 cycles of 95 °C for 3 seconds and 60 °C for 30 seconds. Specificity of the PCR products was verified by melting curve visualization. All data was analyzed using the delta-delta cT method calibrated to a reference gene (Beta-Actin, forward primer (5'-CACCAACTGGGACGACAT), reverse primer (5'- ACAGCCTGGATAGCAACG) and normalized to the scrambled control or control treated sample set to 1.

Design and Synthesis

Based on the preliminary data described above, we together with Morris group at COH, designed siRNAs (BGas s/as) targeting the nascent RNA from the promoter sequence of BGas and control scrambled siRNA (BGas SC s/as). We also designed ASO oligonucleotide (Gap) targeting the intron one of the BGas and control scrambled Gap SC as well. - **Table 5.3**. The BGas siRNA sense strands and the Gap sequences contain alkyne/amino functionality at the 5' termini.

Table 5.3. System siRNAs and Gap DNA oligonucleotides designed; Sequences of siRNA sense and antisense strands, Gap and peptide sequences.

Name	Sequence
BGas s	5'-Hexenyl-r(UUG AAC AAA AAC ACA AAA AUG UU)
BGas as	5'-r(CAU UUU UGU GUU UUU GUU CAA UU)
BGas SC s	5'-Hexenyl-r(AUU AAC CAA UAA GAA AGA AAC UU)
BGas SC as	5'-r(GUU UCU UUC UUA UUG GUU AAU UU)
Gap	5'-Hexynyl-GTG GTA TAA AAG ATA ATT AT
Gap A	5'-AminoC12-GTG GTA TAA AAG ATA ATT AT
Gap SC	5'-Hexenyl-AGT AAT GTA GGA TAA TTA TA
Peptides:	
TatO	AEDGY
Tat1	GRKKRRQRRR
Tat1a	5-azidopentanoic acid-GGKKRRQKGR

First, we used the ASO Gap sequences for preliminary studies to investigate their effect on the *CFTR* expression. We prepared four probes, G1, control naked Gap sequence as shown in **Table 5.3**; G2, Gap sequence complexed with 100 nmol Tat1a peptide; G3, Gap covalently conjugated with Tat1a peptide using CuAAC chemistry; and G4, Gap A covalently conjugated with Tat1a using SPAAC chemistry.

Results and Discussion

Cell line studies

We tested the probes in A549 cell line plated at density of 100000 cells per well with subsequent addition, the next day of the Gap probes, G1-G4, in three different concentrations (100 nM, 200 nM and 400 nM).

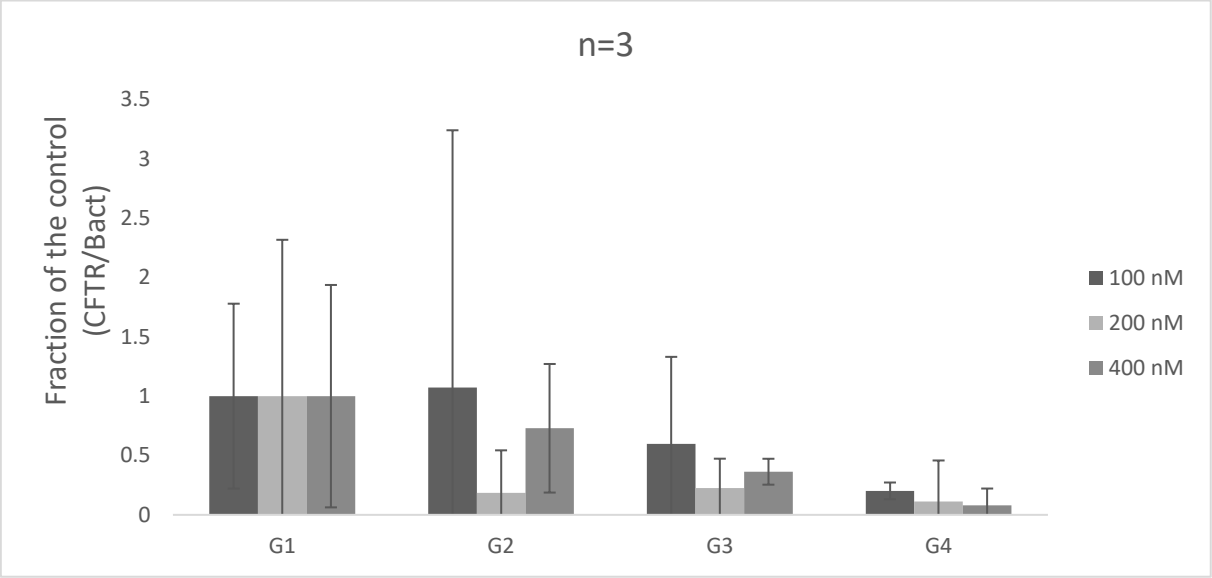
72 h after transfection, cells were collected and the total RNA isolated as described in the material and method section. Next, the RNA was reverse transcribed to cDNA followed by qPCR

to quantify the *CFTR* mRNA levels. Notably **Chart 5.3-a**, we observed big variation within the data with CV values >30%.

In the next set experiment, we added the probes G1-G4 to the plated cells using lower concentrations (10 nM, 20 nM and 50 nM). Even though, the complexed G2 probe showed slight increase of the mRNA levels when used in concentrations of 10 nM and 50 nM (**Chart 5.3-b**), the variability within the data was big. Finally, we screened the probes using lower concentrations for the G3 and G4 probe (2 nM and 4 nM). Notably the qPCR results – **Chart 5.3-c**, G3 and G4 probes transfected in 2 nM concentration give results with more than two fold increase of the mRNA levels. G3 probe showed similar results when transfected in 4 nM concentration, whereas the efficiency of the G4 probe 4 nM was lowered, increasing of the mRNA levels only for 10%. In this experiment set up we were not able to replicate the results for probe G2 when transfected in 50 nM concentration. Worth noticing is the big variation within the data - CV (%) values presented in Appendix A, **Table A1**.

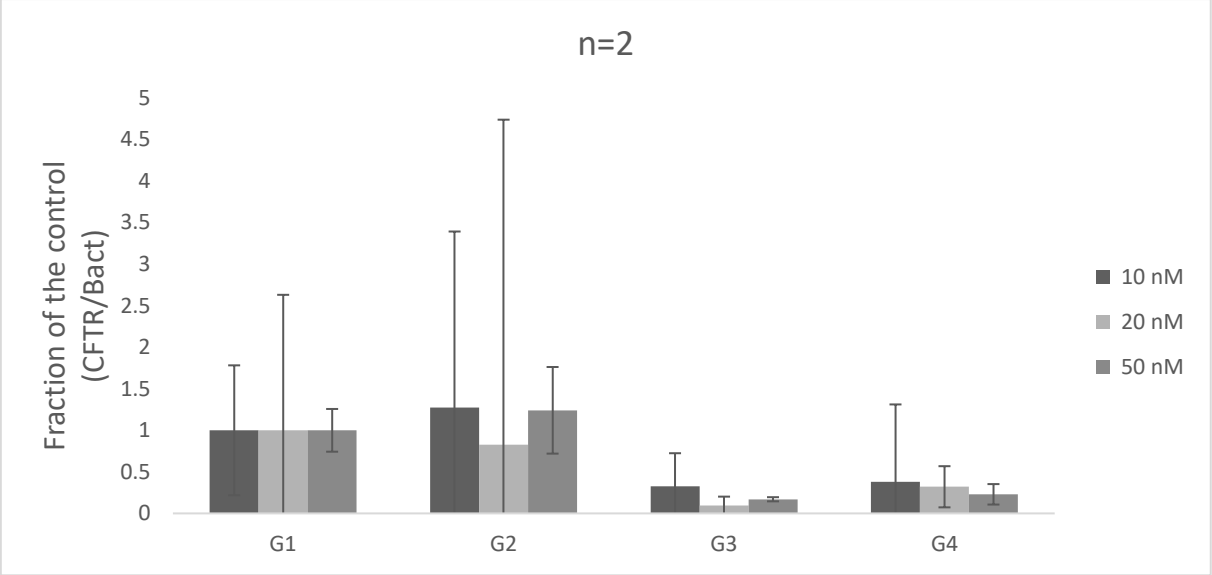
Chart 5.3. Repression of BGas using Gap (G1-G4) in A549 cell line

a)



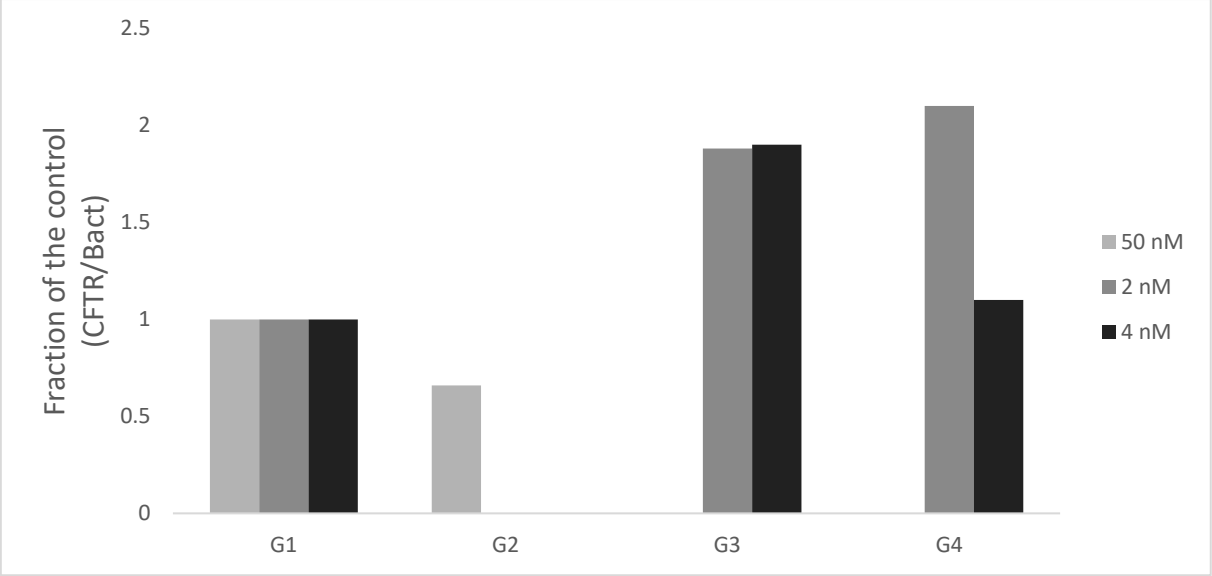
*400000 cells/ml A549 were seeded and 24 h latter, probes G1-G4, 100-400 nM concentrations were dropped onto the cells. 74 h later the total RNA was isolated and quantified by qPCR.

b)



*400000 cells/ml A549 were seeded and 24 h latter, probes G1-G4, 10-50 nM concentrations were dropped onto the cells. 74 h later the total RNA was isolated and quantified by qPCR.

c)



*400000 cells/ml A549 were seeded and 24 h latter, probes G1-G4, 2-50 nM concentrations were dropped onto the cells. 74 h later the total RNA was isolated and quantified by qPCR. SD not shown because of big data variation. For CV (%) values see Table A1, Appendix A.

Conclusion

We performed preliminary set of screening studies for the Gap ASO probe complexed or conjugated with Tat1a peptide in A549 cells for their ability to suppress lnRNA BGas function. In

general, the conjugated G3 and G4 probes showed minimal effects increasing the *CFTR* mRNA levels but with big variability within the data (**Table A1**). Additional positive control experiments are needed to assess the Gap efficacy and activity towards the BGas function in the A549 cells.

Further work, Nanoplexes formulation

Next, to examine the efficacy of the siBGas complexed with peptides, we formulated the siBGas and the scrambled control siBGas SC as nanoplexes with peptides TatO and Tat1 (**Table 5.3**). The complexation was done by mixing and incubation of both biomolecules using the electrostatic interactions between them (**Figure 5.2**).

We ended up with set of 10 probes (**Table 5.4**) that next will be tested in cell culture studies (A549 cell line) at the Center for Gene Therapy, COH by Olga Villamizar, PhD.

Table 5.4. siRNA nanoplexes targeting BGas intron

Name	**siRNA, 20 nmol**	**Peptide**
B1	BGas s+BGas as	/
B2	BGas SC s+BGas SC as	/
		30 nmol
B3	BGas s+BGas as	TatO
B4	BGas SC s+BGas SC as	TatO
B5	BGas s+BGas as	Tat1
B6	BGas SC s+BGas SC as	Tat1
		60 nmol
B7	BGas s+BGas as	Tat0
B8	BGas SC s+BGas SC as	Tat0
B9	BGas s+BGas as	Tat1
B10	BGas SC s+BGas SC as	Tat1

Further work, Covalent Conjugation

Lastly, we prepared covalently conjugates of BGas s/as and BGas SC s/as to Tat1a and Gap and Gap SC as well to further investigate and confirm the data discussed above. First the alkyne sense strands and the alkyne Gap sequences were covalently conjugated to the azide Tat1a peptide sequence. The reaction was done using CuAAC chemistry and already optimized method for clicking the RNA described above. Next, the sense and antisense strands were annealed to make the siRNA. The samples will be tested in A549 cell line studies at COH by Olga Villamizar, PhD.

Chapter 6

Paper IV

Fluorescent oligonucleotides with bis (prop-2-yn-1-yloxy) butane-1,3-diol scaffold allow for rapid detection of disease associated nucleic acids

Taskova M, Astakhova K.

Fluorescent oligonucleotides with bis(prop-2-yn-1-yloxy)butane-1,3-diol scaffold rapidly detect disease-associated nucleic acids

Maria Taskova and Kira Astakhova*

Department of Chemistry, Technical University of Denmark, Kemitorvet 206, 2800 Kgs. Lyngby, Denmark

Abstract: Biomedical research and clinical work demand rapid and reliable detection of disease-associated nucleic acids. Fluorescent oligonucleotides that bind precisely and sense target DNA or RNA are useful tools for simple hybridization-based assays. Although the literature reports a plethora of oligonucleotide modifications, they often result in poor yields and are very expensive. We describe a simple and cost-effective synthesis of a new bisalkyne butane-1,3-diol scaffold and its efficient coupling into oligonucleotide sequences. In addition, we post-synthetically bioconjugated the oligonucleotides (double and quaternary) with azide dyes perylene, 5JOE and (phenylethynyl)pyrene. We investigated the biophysical and photophysical properties of the oligonucleotide–dye conjugates and confirmed a "light up" fluorescent sensing mechanism of the probes consequent on target binding. The data verify the ability of the oligonucleotide–perylene conjugates in efficient target binding and detection of single-nucleotide polymorphisms in human oncogenes in a convenient and fast (2.5 h) bead-based assay.

In the past two decades, modified oligonucleotides have gained immense interest in a broad spectrum of research fields and technology, including medicinal chemistry, nanotechnology and biomedicine.[1,2] To address a rapidly growing demand for modified oligonucleotide probes, nucleic acid chemistry proposes numerous modifications to the nucleic acid backbone,[3] the sugar[4] and the nucleobase.[5,6] Moreover, the literature describes bioconjugation of the modified oligonucleotides with various fluorescent dyes and other performance-improving tags as a promising strategy to develop effective tools for clinical diagnostics, imaging etc.[7–9]

Clinical diagnosis is a field where fluorescent oligonucleotides have wide application.[10] Taking advantage of developed optical methods and improved fluorescent dyes, hybridization-based platforms allow for rapid detection and studies of human genes and of their disease-related variants.[11] However, this requires the design and synthesis of stable and specific fluorescent oligonucleotide probes.[11–13]

A plethora of studies report strategies for labeling oligonucleotides with fluorescent molecules.[14] One of those includes enzymatic incorporation of fluorescent nucleotides into DNA probes during a PCR reaction.[15] Nevertheless, the polymerase-mediated strategy is often inefficient. This is due to the enzymatic dependency and the bulky nature of the fluorescent dyes, which often dramatically reduce labeling yields.[9]

Another approach for preparing fluorescent oligonucleotide probes couples a fluorescently labeled nucleotide monomer into a sequence-specific probe during solid-phase oligonucleotide synthesis.[16] The efficiency of this method depends strongly on the size and geometry of the labeled nucleotide molecule.[16,17] Although this method is habitually used, it faces challenges such as low coupling efficiency, low yields and insufficient purity of the product. Alternatively, post-synthetic bioconjugation strategies propose introduction of a functional group into the oligonucleotide sequence that serves as a substrate for post-synthetic labeling.[18–20] Efficient coupling of the functionalized oligonucleotide with different molecules using standard reagents and the achieving of high product yields are the major improvements of this method.[20]

Besides these benefits, the post-synthetic bioconjugation method faces concerns that need consideration. For instance, preparation of the reported functionalized nucleotide monomers takes a long time and involves complicated synthetic routes.[21,22] Furthermore, the commercially available functionalized nucleic acid scaffolds are expensive and remote. This implies a growing need for the development of simpler synthetic routes for novel functionalized nucleotide monomers and the use of economical

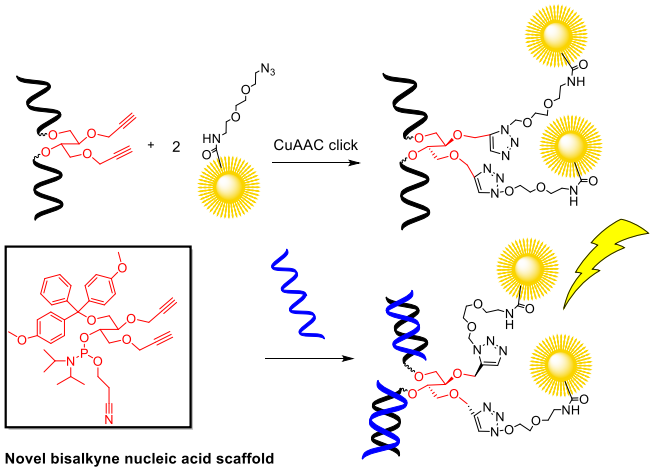


Figure 1. Chemical structure of bisalkyne butane-1,3-diol developed herein, and post-synthetic bioconjugation of bisalkyne-functionalized oligonucleotide with azide dye.

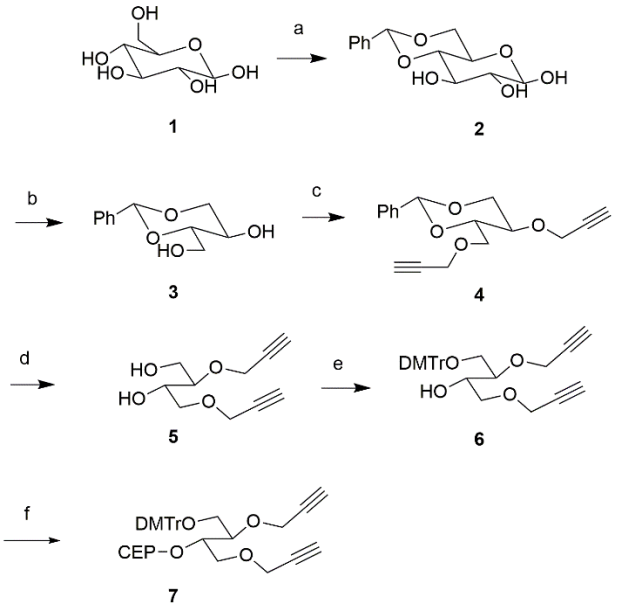
starting materials. Moreover, fluorescence detection of DNA and RNA faces major issues, such as low abundance of the target and quenching of the dye. This causes weak fluorescence signals and uncertain detection.[23–25] The literature proposes the introduction of high density fluorophores into the oligonucleotide probe as a potential solution to this problem.[25]

Herein, we synthesized a new bisalkyne butane-1,3-diol scaffold and incorporated it into short oncogene-targeting oligonucleotides (Figure 1). Next, we double-bioconjugated the

oligonucleotide with different dye molecules. After characterization of the oligonucleotide–dye conjugates (ODCs) that target *BRAF V600E* mutation, we studied the biophysical properties toward mutated (mismatched) and complementary (matched) DNA and RNA targets. The data showed the most potent probes, which we then applied in a bead-based genotyping of human oncogenes *BRAF* and *EGFR.*

First, the bisalkyne butane-1,3-diol scaffold **7** was prepared by a new, simple and cost-effective protocol (Scheme 1) using glucose **1** as a starting material. Initially, by the use of $ZnCl_2$ as a Lewis acid of moderate strength and PhCHO, both C4 and C6 OH groups were protected by the formation of acetal. Next, by oxidation with $NaIO_4$ followed by reduction with $NaBH_4$, the glucose ring of **2** was opened and two new OH groups were formed. Treatment of **3** with NaH, followed by addition of propargyl bromide C_3H_3Br resulted in the formation of product **4** with two alkyne functionalities. The acetal was then hydrolyzed using trifluoroacetic acid (TFA) in dichloromethane, yielding product **5**. Next, the primary OH group of **5** was selectively DMTr-protected by the use of 4,4-dimethoxytrityl chloride (DMTrCl) in pyridine (**6**). Finally, by phosphitylation with diisopropylammonium tetrazolide (DIAT) and 2-cyanoethyl *N,N,N',N'*-tetraisopropylphosphoroamidite, the phosphoramidite monomer **7** was synthesized.

Scheme 1. Synthesis of bis(prop-2-yn-1-yloxy)butane-1,3-diol nucleic acid scaffold



*Reagents, conditions and yields: (a) $ZnCl_2$, PhCHO, rt, 5h, 9%; (b) $NaIO_4$, $NaHCO_3$, rt, 1 h; $NaBH_4$, rt, 1 h, 70%; (c) NaH, 120 °C, 1 h; C_3H_3Br, rt, 24 h, 65%; (d) TFA, DCM 0 °C, 2–3 h, 43%; (e) DMTrCl, pyridine, rt, 24 h, 57%; (f) DIAT, $\{[(CH_3)_2CH]_2N\}_2POCH_2CH_2CN$, rt, 16 h, 82%.

Next, we synthesized two oligonucleotide sequences K1 and K2, a 21mer and a 15mer, respectively, (Table 1). Following their design, the sequences targeted the *BRAF V600E* mutation (the target sequence was obtained from the NCBI Pubmed). While K2 is a pure DNA sequence that contains the new bisalkyne scaffold **7** incorporated once, K1 is a mixed DNA/locked nucleic acid (LNA) sequence that contains the bisalkyne scaffold **7** twice, in positions 6 and 18. The installed three sequential LNA monomers (+C+T+C) in K1 improve the specificity and the binding affinity to the complementary target oligonucleotide.[26,27] Concerning the length, the literature reports that 13mer–25mer oligonucleotides as acceptable for sufficient target interaction and suggest a 21mer oligonucleotide probe for ensuring unique interaction with the target.[28,29] We included the shorter 15mer, K2 oligonucleotide with the aim of investigating its performance and comparing it with the standard 21mer. The oligonucleotide synthesis of K1 and K2 was executed automatically and the bisalkyne scaffold was incorporated into the desired position by hand coupling (SI, Scheme S1).

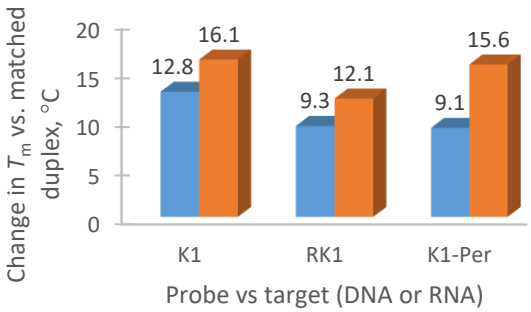
The oligonucleotides were purified and characterized by ion-exchange HPLC and MALDI-TOF mass spectrometry (SI, Figures S1 and S2). Subsequent, post-synthetic covalent conjugation of the bisalkyne-functionalized oligonucleotides K1 and K2 with azide derivatives of perylene (K1/K2–Per), 5JOE (K1/K2–5JOE) and (phenylethynyl)pyrene (PEP) (K1/K2–PEP) dyes (Scheme 1) was performed by copper(I)-catalyzed alkyne-azide

Table 1. Oligonucleotide sequences

#	Sequence	Mw calc.	Mw exp.
K1	5'-CGAGAXTT+C+T+CTGTAGCXAGA	6457.0	6457.4
RK1	5'-CGAGAdUTT+C+T+CTGTAGCdUAGA	6520.3	6520.3
K2	5'-GATTTCXCTGTAGCT	4503.0	4503.8
RK2	5'-GATTTCdUCTGTAGCT	4535.0	4535.0

*+C and +T are LNA nucleotides.

Chart 1. Single-nucleotide polymorphism (SNP) discrimination in DNA and RNA targets by probes K1, RK1 and K1–Per.



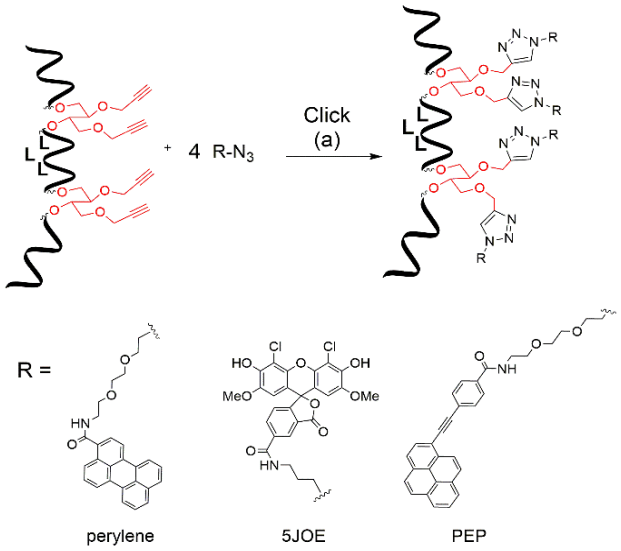
*ΔT_m values for DNA and RNA targets represented in blue and red, respectively. The values were calculated by subtracting the T_m values for the complementary and mismatched target DNA/RNA. K1 is the oligonucleotide that contains the new bisalkyne scaffold incorporated twice, K1–Per is K1 quaternary-conjugated with perylene dyes and RK1 is the unmodified reference sequence.

cycloaddition chemistry (CuAAC; Scheme 2). Then, the conjugated products were refined on an Illustra NAP-5 column, giving products with high purity in yields of 32–80%. We observed that the yields of the products were affected by the type of dye. The lowest yields were observed for rapidly precipitating PEP-azide dye. The identity and purity of the ODCs were confirmed by ion-exchange HPLC and MALDI-TOF mass spectrometry (SI, Figures S1 and S2, Table S2).

The temperature-melting UV experiments (SI, Table S3) were performed in a phosphate-buffered saline (100 mM, pH 7.4). Unmodified DNA sequences RK1 and RK2, which contained 2'-deoxyuridine in the position on the bisalkyne scaffold, were used as controls (Table 1). For K1 and K2, we designed complementary and mismatched (in position 10) DNA and RNA targets (21mers). For K2, we also included short complementary 15mers and short variable mismatched (A, U or C, in position 8) targets along the bisalkyne modification – Table S1. The duplexes of modified oligonucleotides K1, K2 and ODCs with perylene – with all complementary and mismatched

DNA/RNA long sequences – showed S-shape melting curves, similar to the unmodified controls (SI, Figure S3). However, the ODCs of K1 and K2 with 5JOE and PEP gave no clear transition in the melting curves. When examining the ODC duplexes of K2 with short DNA and RNA targets, we observed S-shaped melting curves and clear T_m transitions solely for the conjugates with perylene with the complementary targets. All the duplexes of ODCs derived from K2 with short mismatched targets showed no clear transition, independent of the variable mismatches (A, U or C; SI, Table S3).

Scheme 2. Bioconjugation of oligonucleotide K1 with azide dyes



*Reagents and conditions: (a) 0.1 M Triethylammonium acetate buffer, pH 7.3; Cu(II):TBTA complex 10 mM, ascorbic acid aq. 50 mM, Mili Q water/dimethyl sulfoxide 0.8:1.2, v/v/v, 75 °C 10 min → 60 °C Mw 45 min → rt overnight.

In general, the oligonucleotide–perylene probes showed similar strength of binding with DNA and RNA targets. The duplexes of modified oligonucleotides K1, K2 and their conjugates displayed lower T_m values compared to the duplexes of unmodified oligonucleotides RK1 and RK2. However, the ODCs K1–Per and K2–Per stabilized the duplex with DNA and RNA targets. In particular, the decrease in T_m values for K1–Per compared with RK1 were –3.7 °C and –5.8 °C for DNA and RNA targets, respectively. In turn, the values for K2–Per vs. RK2 were –0.2 °C and –6.0 °C for DNA and RNA targets, respectively. The results were similar for K2–Per binding the short targets, with one even more pronounced effect where K2–Per was binding the short-matched DNA target at 1.2 °C stronger compared to RK2 (SI, Table S3). We also estimated the single-nucleotide polymorphism (SNP) discrimination by calculating ΔT_m values, the difference between DNA/RNA matching and mismatching sequences. RK1 discriminated mismatches with ΔT_m values of 9.3 °C and 12.1 °C for DNA and RNA, respectively. K1–Per showed similar DNA mismatch sensitivity, with ΔT_m 9.1 °C, and improved RNA mismatch discrimination, with ΔT_m 15.6 °C (Chart 1). Notably, the data derived for RK2 and K2 – variable nucleobase mismatched (A, U or C) on the same position – gave differences in T_m values ranging from 0.4 to 3 °C. Even though the duplexes of K1/K2–Per probes with DNA and RNA showed equal strength, in agreement with literature, the K1–Per could discriminate the mismatch in RNA better compared with the mismatch in DNA.[30]

With regard to the flanking ends of the K2 targets (DT1/RT1 vs. DT1 short/RT1 short), our data showed an increase in the T_m values when the target was longer (21mer) and contained three LNAs, compared with the short one (15mer). This effect was more pronounced for the modified K2 sequence compared to the reference RK2 (3.5 °C and 2.8 °C for K2 with DNA and RNA respectively vs. 0.9 °C and 1.6 °C with DNA and RNA, respectively). This confirmed the stabilizing effect of the long targets on probe binding, without affecting the excellent SNP discrimination by the perylene-containing ODCs.

To investigate the structure of ODC duplexes with complementary and mismatched DNA and RNA targets (15mer targets for K2) we recorded CD spectra in a phosphate-buffered saline (pH 7.4). For comparison, we measured spectra for the K1, K2 and the unmodified RK1 and RK2 reference sequences (Figure 2 and SI, Figure S4). The data confirmed no variation in the shapes of the spectra between ODCs, the non-conjugated precursors K1/K2 and the references RK1/RK2. In all cases, the CD spectra followed a similar pattern, i.e. a negative peak at 205–210 nm, a positive peak at 220–225 nm, followed by more negative and positive peaks at 240–245 nm and 265–275 nm, receptively (Figure 2). The spectra of the duplexes with complementary and mismatched targets were overlapping. These results were similar to the CD spectra observed previously for fluorescent oligonucleotides with single-dye labeling.[19] However, K1/K2 and the references RK1/RK2 showed higher peak intensities than the ODCs in all cases. The numerical sum of the results (Tables S4/S5) confirmed that the three dyes related and intercalated within the duplexes to different degrees. In particular, the ODCs with perylene showed the biggest alteration in the duplex geometry.

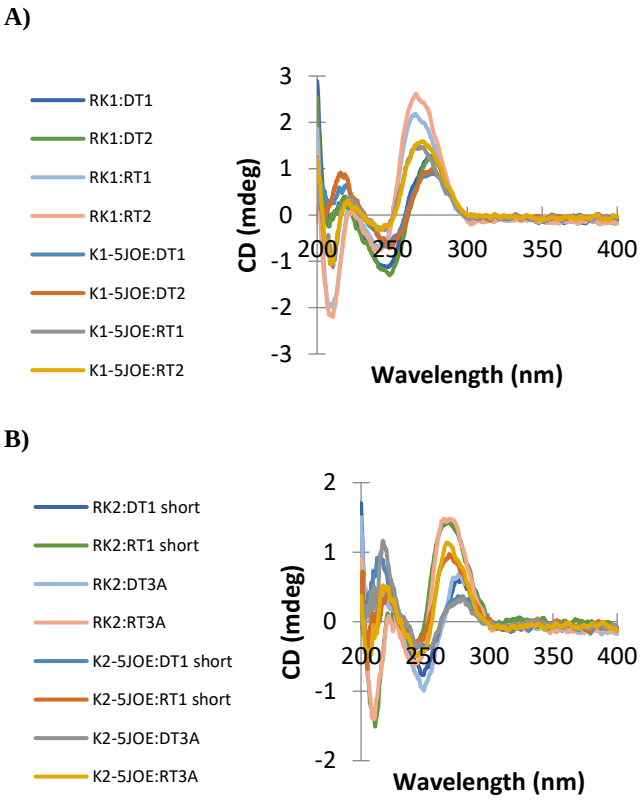


Figure 2. Representative CD spectra of ODC duplexes and controls: A) K1 oligonucleotide conjugated with four 5JOE dyes (K1–5JOE) and control unmodified sequence RK1; B) K2 oligonucleotide conjugated with two 5JOE dyes (K2–5JOE) and control unmodified sequence RK2. The spectra were recorded in 100 mM phosphate buffer, pH 7.4, using 2 µM concentration of each strand.

Next, we studied the target DNA/RNA binding and SNP sensitivity by fluorometry. The measurements were performed in a phosphate-buffered saline (pH 7.4) using 1 µM concentration of the perylene (excited at 425 nm) and the PEP (excited at 375 nm) ODCs and 0.5 µM concentration of the 5JOE (excited at 495 nm) ODCs.

The fluorescence emission spectra showed hypochromic shifts for the K1 ODCs. Conversely, the K2 ODCs showed bathochromic shifts. (Figure 3) While the conjugates with perylene and PEP dyes displayed compelling emission maxima shifts (>2 nm), the conjugates with 5JOE showed minor shifts. These results, together with the CD data, suggested that the perylene and PEP dyes intercalated into the DNA and RNA duplex more efficiently than did 5JOE. Significantly and in agreement with the literature, they also intercalated more efficiently into the DNA duplex compared to the duplex with RNA.[30]

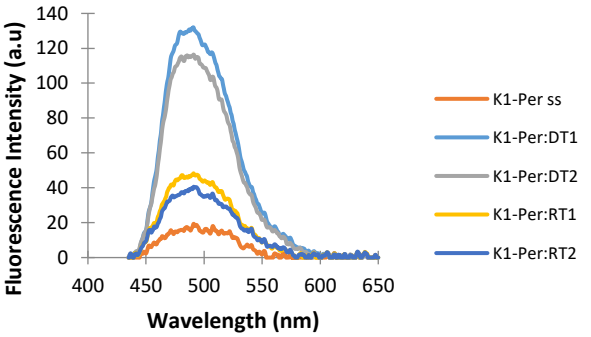
In addition, the ODCs with perylene and PEP discriminated the single-strand probes and the duplexes by alterations in the fluorescence intensity at emission maxima at 490 nm and 400 nm, respectively (Figure 3 and SI, Charts S2/S3). The ODCs with perylene showed pronounced effect with ratios of fluorescence intensities between the duplex with DNA and the single strand (I_{ds}/I_{ss}) of 7.3 and 3.6, for K1–Per and K2–Per, respectively. In addition, the ODCs with PEP dye showed fluorescence intensity ratios of 2.4 and 2.6 for K1–PEP and K2–PEP, respectively. ODCs with 5JOE failed to detect hybridization with target DNA and RNA sequences by fluorescence.

Notably, as seen in SI Chart S1/S2, the K1–Per probe efficiently discriminated SNP in DNA and RNA target sequences by fluorescence intensity alteration. The probe exhibited higher fluorescent intensity at emission maxima, binding the wild-type targets, compared with the mutant DNA and RNA sequences. In particular, the difference in fluorescence intensity binding match and mismatch DNA was 15.5 a.u., whereas this value for binding match and mismatch RNA was 8.2 a.u. K2–Per showed satisfactory SNP discrimination just for the DNA duplex, exhibiting differences in fluorescent intensities binding the wild-type and mutant DNA of 22.7 a.u. The K1- and K2–5JOE did not discriminate SNP.

Among the synthesized probes, the ODCs with perylene displayed desirable properties for homogenous fluorescence assay.[31] They clearly showed increased fluorescence intensity emission upon transition from single strength to DNA and RNA target binding. Moreover, the ODCs with perylene efficiently distinguished DNA and RNA sequences with single-point mutation from the wild-type target by alteration in the fluorescence intensity.

These results agree with the literature reports on ODCs that show a “light up effect”, differentiating single-stranded fluorescent probe vs. hybridized one.[31]

A)



B)

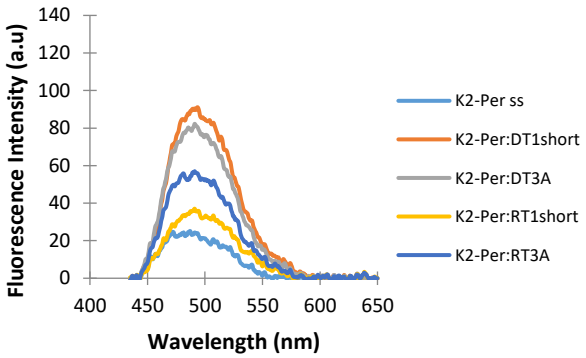


Figure 3. Representative fluorescence emission spectra of ODCs: A) single-stranded and duplexes of K1–Per (A) and K2–Per (B) with complementary and mismatched DNA and RNA targets.

Therefore, the data suggest that the K1–Per and K2–Per probes distinguish single-stranded status vs. target binding by positioning of the dyes within the duplex. Single-stranded probes expose the perylene dyes to the solvent environment, which results in fluorescence quenching and low fluorescence emission intensity. Conversely, oligonucleotide–perylene duplexes intercalate the dyes, preventing dye quenching and showing increased fluorescence emission. The same is also valid for SNP discrimination. The locally distorted duplex of the single-point-mutated target fails the decisive prevention of the dye from environment quenching. This causes a degree of decrease in the fluorescence intensity.

Lastly, considering the promising photophysical properties of the K1– and K2–Per probes, we developed an assay to detect clinically relevant mutated oncogenes, *BRAF* and *EGFR*. For these studies, genomic DNA targets were purified from commercial cancer cell lines and digested with EcoRI. As a negative, i.e. mutation-free, control we used DNA extracted from healthy cell line HMC. For the assay, we additionally designed and synthesized the probe E1–Per, a 21mer DNA/LNA (sequence shown in the SI) that targets *EGFR* mutation. The E1 sequence was first synthesized and the bisalkyne scaffold **7** was incorporated by hand coupling at the 5’ termini, followed by CuAAC click conjugation with perylene azide dye. Moreover, we designed *BRAF* and *EGFR* 5’-biotinylated-capturing probes (SI).

The capturing assay was performed using streptavidin-coated magnetic dynabeads, biotin-conjugated capturing probes and our enriched fluorescent perylene–oligonucleotides (K1–Per, K2–Per or E1–Per). Repeating washing, incubation and annealing-dehybridizing steps were performed to ensure the required high specificity of the assay (SI). Quantitative results were gained by fluorometry. Overall, the assay took only 2.5 h, which is comparable to or shorter than existing PCR and sequencing assays.[4–7]

Notably, all three probes performed well concerning the target specificity (see fluorescence intensities for mutated DNA vs. HMC in SI, Chart S3). For K2–Per, the signal in HMC negative control was elevated, but it was still 50% lower compared to its cancer cell line intensity. K1–Per showed double the fluorescence intensity compared to the K2–Per and E1–Per probes. These results agreed with the composition of the K1–Per probe containing four molecules of perylene dye vs. the other two probes that contain two molecules of conjugated dyes. The estimated limit of detection (LOD) of the capturing assay was 300 ± 5 pM mutated DNA; the error of two independent measurements (CV) was 4.5%. These metrics are similar to or lower than existing genotyping methods.[4–7] Given the high speed and low cost of our method, we believe it has the potential to contribute to current genotyping methods as a

convenient and reliable alternative to amplification-based techniques.

In summary, we developed a convenient synthetic route for a new bis(prop-2-yn-1-yloxy)butane-1,3-diol nucleic acid scaffold. The synthesized bisalkyne scaffold coupled effectively to oligonucleotide probes, followed by post-synthetic double and quaternary bioconjugation with fluorophores: perylene, PEP and 5JOE. We use the multi-fluorophore oligonucleotides for biophysical studies of the duplexes with complementary DNA/RNA *in vitro*. CD spectroscopy indicated that the prepared conjugates adopt a different duplex geometry when compared to the unmodified references. While the perylene–oligonucleotide conjugates showed successful adaptation within a duplex structure with DNA and RNA, the 5JOE and PEP dyes failed to make duplex intercalation. Moreover, the oligonucleotide–perylene conjugates (K1/K2–Per) showed sufficient target binding and excellent sensitivity toward the mismatched DNA and RNA targets by T_m and fluorometry studies. In addition, fluorometry studies in a model system and in cancer DNA confirmed that these conjugates discriminated well between the single strand and the duplex, and effectively detected single point mutated DNA in as little as a 300 pM sample. Taking into account these optimal properties, the oligonucleotide–perylene conjugates developed herein have a high potential as effective fluorescent tools for studies of nucleic acids and the detection of disease-related gene mutations.

ASSOCIATED CONTENT

Supporting Information
Detailed experimental procedures and additional figures; Commercial DNA and RNA sequences; HPLC and MALDI spectra; T_m numerical results and melting curves; CD spectra; Fluorescence results; NMR spectra.

AUTOR INFORMATION

Corresponding Author
E-mail: kiraas@kemi.dtu.dk

Keywords: nucleic acid scaffold, oligonucleotides, bioconjugation, fluorescent dyes, oncogene

1. Kratschmer, C. and Levy, M. (2017) Effect of Chemical Modifications on Aptamer Stability in Serum. *Nucleic Acid Ther.* 27, 335–344.

2. Schoch, K. M. & Miller, T. M. (2017) Antisense Oligonucleotides: Translation from Mouse Models to Human Neurodegenerative Diseases. *Neuron* 94, 1056–1070.

3. Chan, J. H., Lim, S. and Wong, W. F. (2006) Antisense oligonucleotides from design to therapeutic application. *Clin. Exp. Pharmacol. Physiol.* 33, 533–540.

4. Prakash, T. P. (2011) An Overview of Sugar-Modified Oligonucleotides for Antisense Therapeutics. *Chem. Biodivers.* 8, 1616–1641.

5. Wilson, J. N. & Kool, E. T. (2006) Fluorescent DNA base replacements: Reporters and sensors for biological systems. *Org. Biomol. Chem.* 4, 4265–4274.

6. Kool, E. T. (2002) Replacing the nucleobases in DNA with designer molecules. *Acc. Chem. Res.* 35**,** 936–943.

7. Astakhova, I. K., Samokhina, E., Babu, B. R. & Wengel, J. (2012) Novel (phenylethynyl)pyrene-LNA constructs for fluorescence SNP sensing in polymorphic nucleic acid targets. *Chembiochem* 13**,** 1509–1519.

8. Huang, Y. (2017) Preclinical and Clinical Advances of GalNAc-Decorated Nucleic Acid Therapeutics. *Mol. Ther. Nucleic Acids* 6, 116–132.

9. Dai, N. & Kool, E. T. (2011) Fluorescent DNA-based enzyme sensors. *Chem. Soc. Rev.* 40**,** 5756–5770

10. Miotke, L., Maity, A., Ji, H., Brewer, J. & Astakhova, K. (2015) Enzyme-Free Detection of Mutations in Cancer DNA Using Synthetic Oligonucleotide Probes and Fluorescence Microscopy. *PLoS One* 10**,** 1371-1382.

11. Suzuki, A., Yanagi, M., Takeda, T., Hudson, R. H. E. & Saito, Y. (2017) The fluorescently responsive 3-(naphthalen-1-ylethynyl)-3-deaza-2′-deoxyguanosine discriminates cytidine via the DNA minor groove. *Org. Biomol. Chem.* 15**,** 7853–7859.

12. Vietz, C., Lalkens, B., Acuna, G. P. & Tinnefeld, P. (2017) Synergistic Combination of Unquenching and Plasmonic Fluorescence Enhancement in Fluorogenic Nucleic Acid Hybridization Probes. *Nano Lett.* 17**,** 6496–6500.

13. Su, X., Xiao, X., Zhang, C. & Zhao, M. (2012) Nucleic Acid Fluorescent Probes for Biological Sensing. *Appl. Spectrosc.* 66**,** 1249–1261..

14. van der Velde, J. H. M., Oelerich J., Huang J., Smit J.H., Jazi A.A,. Galiani S., Kolmakov K., Gouridis G., Eggeling C., Herrmann A. *et al.* (2016) A simple and versatile design concept for fluorophore derivatives with intramolecular photostabilization. *Nat. Commun.* 7**,** 10144-10159.

15. Tasara, T., Angerer B., Damond M., Winter H., Dörhöfer S., Hübscher U., Amacker M. (2003) Incorporation of reporter molecule-labeled nucleotides by DNA polymerases. II. High-density labeling of natural DNA. *Nucleic Acids Res.* 31**,** 2636–2646.

16. Kölmel, D. K., Barandun, L. J. & Kool, E. T. (2016) Efficient synthesis of fluorescent alkynyl C-nucleosides via Sonogashira coupling for the preparation of DNA-based polyfluorophores. *Org. Biomol. Chem.* 14**,** 6407–6412.

17. Segal, M. & Fischer, B. (2012) Analogues of uracil nucleosides with intrinsic fluorescence (NIF-analogues): synthesis and photophysical properties. *Org. Biomol. Chem.* 10**,** 1571-1580.

18. Walter, H.-K., Olshausen, B., Schepers, U. and Wagenknecht, H.-A. (2017) A postsynthetically 2'-"clickable" uridine with arabino configuration and its application for fluorescent labeling and imaging of DNA. *Beilstein J. Org. Chem.* 13**,** 127–137.

19. Taskova, M., Barducci, M. C. & Astakhova, K. (2017) Environmentally sensitive molecular probes reveal mutations and epigenetic 5-methyl cytosine in human oncogenes. *Org. Biomol. Chem.* 15**,** 5680–5684.

20. Taskova, M., Madsen C.S., Jensen K.J., Hansen L.H., Vester B., Astakhova K. (2017) Antisense Oligonucleotides Internally Labeled with Peptides Show Improved Target Recognition and Stability to Enzymatic Degradation. *Bioconjug. Chem.* 28**,** 768–774..

21. Ries, A., Kumar R., Lou C., Kosbar T., Vengut-Climent E., Jørgensen P.T., Morales J.C., Wengel J. (2016) Synthesis and Biophysical Investigations of Oligonucleotides Containing Galactose-Modified DNA, LNA, and 2′-Amino-LNA Monomers. *J. Org. Chem.* 81**,** 10845–10856.

22. Kumar, P., Sharma P.K., Hansen J., Jedinak L,Reslow-Jacobsen C,Hornum M,Nielsen P. (2016) Three new double-headed nucleotides with additional nucleobases connected to C-5 of pyrimidines; synthesis, duplex and triplex studies. *Bioorg. Med. Chem.* 24**,** 742–749

23. Astakhova, K. (2014) Toward Non-Enzymatic Ultrasensitive Identification of Single Nucleotide Polymorphisms by Optical Methods. *Chemosensors* 2**,** 193–206.

24. Wilson, J. N., Cho, Y., Tan, S., Cuppoletti, A. and Kool, E. T. (2008) Quenching of fluorescent nucleobases by neighboring DNA: the "insulator" concept. *Chembiochem* 9**,** 279–285.

25. Ren, X., El-Sagheer, A. H. and Brown, T. (2016) Efficient enzymatic synthesis and dual-colour fluorescent labelling of DNA probes using long chain azido-dUTP and BCN dyes. *Nucleic Acids Res.* 44**,** 79-90.

26. Astakhova, I. V*.,* Malakhov A.D., Stepanova I.A., Ustinov A.V., Bondarev S.L., Paramonov A.S., Korshun V.A. - Phenylethynylpyrene (1-PEPy) as Refined Excimer Forming Alternative to Pyrene: Case of DNA Major Groove Excimer. 18, 1972-1980.

27. Astakhova, I. K., Pasternak, K., Campbell, M. A., Gupta, P. & Wengel, J. (2013) A locked nucleic acid-based nanocrawler: designed and reversible movement detected by multicolor fluorescence. *J. Am. Chem. Soc.* 135**,** 2423–2426.

28. Yuan, B., Latek, R., Hossbach, M., Tuschl, T. & Lewitter, F. (2004) siRNA Selection Server: an automated siRNA oligonucleotide prediction server. *Nucleic Acids Res.* 32**,** 130-134.

29. Dias, N. & Stein, C. A. (2002) Antisense oligonucleotides: basic concepts and mechanisms. *Mol. Cancer Ther.* 1**,** 347–355.

30. Perlíková, P., Ejlersen, M., Langkjaer, N. & Wengel, J. (2014) Bis-Pyrene-Modified Unlocked Nucleic Acids: Synthesis, Hybridization Studies, and Fluorescent Properties. *ChemMedChem* 9, 2120–2127.

31. Astakhova, I. K. & Wengel, J. (2013) Interfacing click chemistry with automated oligonucleotide synthesis for the preparation of fluorescent DNA probes containing internal xanthene and cyanine dyes. *Chemistry* 19**,** 1112–1122 .

Supporting Information for

Fluorescent oligonucleotides with bis (prop-2-yn-1-yloxy) butane-1,3-diol scaffold rapidly detect disease associated nucleic acids

Maria Taskova, Kira Astakhova*

Department of Chemistry, Technical University of Denmark, Kemitorvet 206, 2800 Kgs. Lyngby, Denmark

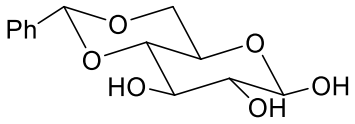
Experimental Procedures

1. Synthesis

General synthetic methods

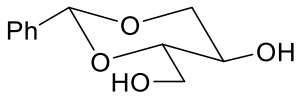
All commercial reagents and solvents were used as received. The reactions were monitored by TLC using Merck silica 60 F_{254} aluminum sheets. The TLC plates were visualized by UV-light at 254 nm and/or by staining by an appropriate staining agent, either vanillin in ethanol and/or 5% sulfuric acid in ethanol, followed by heating. Column chromatography was performed using silica gel (0.040-0.063 mm), Merck. ^{1}H, ^{13}C and ^{31}P NMR were recorded on Bruker Avance III 400 spectrometer at 400.12 MHz, 100.62 MHz and 161.97 MHz, respectively. Chemical shifts (δ) are reported in ppm relative to tetramethylsilane ($\delta_{H,C}$ 0 ppm) or the solvents (DMSO-d_6 δ_C 39.5 ppm, $CDCl_3$ δ_C 77.16 ppm) The coupling constants (*J*) are reported in Hz. As external standard for ^{31}P NMR spectra, 85% H_3PO_4 was used. 2DNMR spectra (HSQC, COSY) have been used in assigning the ^{1}H and ^{13}C NMR signals. High resolution MS-ESI was recorded on a Bruker microTOF-Q II in a positive ion mode with an accuracy of ±5 ppm. Oligonucleotides were synthesized on an automated DNA synthesizer - PerSpective Bio-systems Expedite 8909. MALDI-TOF mass spectrometry was performed on a Ultraflex II TOF/TOF instrument, Bruker, using 3-hydroxypicolinic acid matrix (10 mg/mL 3-hydroxypicolinic acid, 50mM ammonium citrate in 70% aqueous acetonitrile). IE HPLC was performed using a Merck Hitachi LaChrom instrument equipped with a Dionex DNAPac Pa-100 column (250mm x 4 mm).

Preparation of 4,6-O-benzylidene-D-glucose (2)



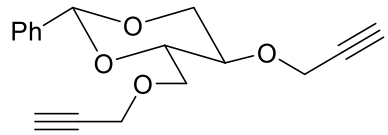
Anhydrous glucose (67.5 g, 374 mmol) and 270 g of dry zinc chloride were stirred in 675 mL of benzaldehyde at room temperature (rt) for 5 h. Then the reaction mixture was cooled to 0 °C and 1350 mL of water were added. The solution was stored at 4 °C for 24 hours. The resulting white precipitate was filtrated, washed with cold water and dried under high vacuum overnight to obtain **2** as a white solid (9 g, 33.5 mmol).
Yield: 9%. ***Rf*** = 0.30 (ethyl acetate).
HRMS-ESI ($M+Na^+$): found m/z: 291.0845, calcd: 291.0845.

Preparation of 1,3-O-benzylidene-L-erythriol (3)



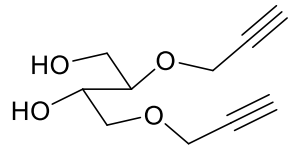
Compound **2** (9 g, 33.5 mmol) was dissolved in 180 mL of water and 67 mL of tetrahydrofuran. Sodium metaperiodate (14.3 g) and sodium hydrogen carbonate (7.4 g) were added and the reaction mixture was stirred at rt for 1 h. Then the solution, that presented a white precipitate, was cooled to 0°C and sodium borohydride (2.7 g) was added. The resulting mixture was stirred for 1 h. The precipitate was filtered and the product was extracted with ethyl acetate. The extract was washed with a solution of sodium thiosulfate and a solution of sodium hydrogen carbonate. The organic fraction was concentrated under reduced pressure and the crude material was purified by column chromatography (40→100% Ethyl acetate in Petroleum ether) to give a white compound **3** (4.9 g, 23.3 mmol).
Yield: 69%. ***Rf*** = 0.40 (ethyl acetate).
HRMS-ESI ($M+Na^+$): found m/z: 233.0795, calcd: 233.0892.
^{1}H NMR (400 MHz, DMSO): δ 7.46-7.43 (m, 2H), 7.38-7.33 (m, 3H), 5.50 (s, 1H), 5.18 (d, *J* = 5.0 Hz, 1H), 4.65 (t, *J* = 5.7 Hz, 1H), 4.10 (dd, *J* = 9.3 Hz, 6.4 Hz, 1H), 4.08 (dd, *J* = 9.3 Hz, 6.4 Hz, 1H), 3.75 (td, *J* = 9.8 Hz, 4.2 Hz, 1H), 3.57-3.46 (m, 3H); **^{13}C NMR** (101 MHz, DMSO): δ 138.3, 128.6, 127.8, 126.3, 100.1, 83.1, 70.7, 60.8.

Preparation of (2R,4S,5R)-2-phenyl-5-(prop-2-yn-1-yloxy)-4-((prop-2-yn-1-yloxy)methyl)-1,3-dioxane (4)



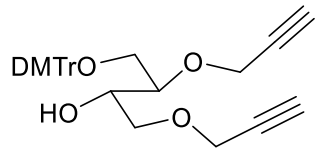
Compound 3 (4.5 g, 21.4 mmol) was dissolved in 112 mL dimethylformamide and 6 g of sodium hydride (60%) was added. The solution was stirred for 1 h under reflux and then was cooled down to rt before adding 13.7 mL propargyl bromide (80%). The solution was left under stirring overnight. The resulted brown precipitate was disolved after addition of 100 mL of ethyl acetate and 100 ml water. The yellowish organic phase collected with two more times of extraction with ethyl acetate was concentrated under reduced pressure and the crude material was purified by silica gel column chromatography (40→100% Ethyl acetate in Petroleum ether) affording compound **4** (4 g, 13.9 mmol) as a light orange oil.
Yield: 65%. ***Rf*** = 0.44 (ethyl acetate).
HRMS-ESI (M+Na$^+$): found m/z: 309.1099, calcd: 309.1103.
^{1}H NMR (400 MHz, DMSO): δ 7.43-7.40 (m, 2H), 7.40-7.36 (m, 3H), 5.56 (s, 1H), 4.46-4.38 (m, 1H), 4.27 (d, *J* = 2.4 Hz, 2H), 4.20 (d, *J* = 2.3 Hz, 2H), 3.85 (m, 1H), 3.76 (dd, *J* = 11.0 Hz, 1.7 Hz, 1H), 3.64 (dd, *J* = 11.0 Hz, 5.7 Hz, 1H), 3.60-3.56 (m, 1H), 3.51 (t, *J* = 2.3 Hz, 1H), 3.44 (t, *J* = 2.3 Hz, 1H), 3.41-3.38 (m, 1H); **^{13}C NMR** (101 MHz, DMSO): δ 137.8, 128.8, 128.0, 126.2, 100.1, 80.3, 80.1, 79.0, 77.4, 77.3, 68.5, 68.2, 68.0, 57.8, 57.1.

Preparation of 2,4-bis(prop-2-yn-1-yloxy)butane-1,3-diol (5)



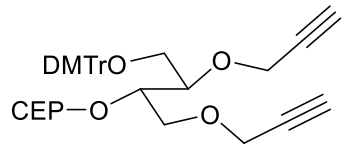
Compound **4** (0.4 g, 1.39 mmol) was dissolved in 12 mL dry dichloromethane and cooled to 0 ^{o}C. 3 mL trifluoroacetic acid diluted in 12 mL dry dichloromethane was added drop-wise over 1 h. The reaction mixture was left under stirring for 16 h, The reaction was quenched with 8 mL (2.5 eq) dry pyridine. It was concentrated under reduced pressure and the crude material was purified by flash chromatography (0→10% methanol in dichloromethane) to give a light yellow oil (0.12 g, 0.60 mmol)
Yield: 43%. ***Rf*** = 0.41 (ethyl acetate).
HRMS-ESI (M+Na$^+$): found m/z: 221.0774, calcd: 221.0790.
^{1}H NMR (400 MHz, DMSO): δ 4.91 (dd, *J* = 5.6, 2.9 Hz, 1H), 4.56 (t, *J* = 5.5 Hz, 1H), 4.22 (dd, *J* = 15.5, 2,4 Hz, 2H), 4.15 (d, *J* = 2.4 Hz, 2H), 3.70-3.55-3.65 (m, 4H, H4), 3.50-3.35 (m, 4H); **^{13}C NMR** (101 MHz, DMSO): δ 80.9, 80.4, 80.1, 76.9, 76.5, 71.3, 68.9, 60.3, 57.8, 57.7.

Preparation of 2,4-bis(prop-2-yn-1-yloxy)butane-1-O-dimetoxytrityl-3-ol (6)



Compound **5** (0.12 g, 0.060 mmol) was co-evaporated with anhydrous pyridine (2 x 10 mL) and dissolved in 4 mL anhydrous pyridine. Dimethoxytrityl chloride (0.45 mg) was added and the reaction mixture was stirred at rt for 24 h. Ethanol (99.6%, 2 mL) was added and it was stirred for 15 min and evaporated to dryness under reduced pressure. The crude mixture was dissolved in dichloromethane 20 mL and washed with water (2 x 20 mL) and a solution of saturated sodium hydrogen carbonate (2 x 20 ml). The resulting organic phase was dried over sodium sulfate and concentrated under reduced pressure The crude material was purified by column chromatography (0→30% Ethyl acetate in Petroleum ether) affording compound **6** (0.17 g, 0.034 mmol) as a light yellow oil.
Yield: 57%. ***Rf*** = 0.24 (40% ethyl acetate in petroleum ether).
HRMS-ESI (M+Na$^+$): found: m/z 523.2089 , calcd: 523.2100.
^{1}H NMR (400 MHz, $CDCl_3$): δ 7.46-7.43 (m, 2H), 7.34 (dt, J = 8.9 Hz, 3,1 Hz 4H), 7.30 – 7.20 (m, 3H), 6.83 (dt, *J* = 8.9 Hz, 2.6 Hz, 4H), 4.33 (dd, *J* = 15.7 Hz, 2.4 Hz, 1H), 4.25 (dd, *J* = 15.7 Hz, 2.4 Hz, 1H), 4.16 (d, *J* = 2.4 Hz, 2H), 3.95 (td, *J* = 6.1 Hz, 3.7 Hz, 1H), 3.78 (s, 6H), 3.75 – 3.71 (m, 1H), 3.64 (dd, *J* = 9.8 Hz, 3.8 Hz, 1H), 3.59 (dd, *J* = 9.8 Hz, 6.0 Hz, 1H), 3.40 (dd, *J* = 10.3 Hz, 3.9 Hz, 1H), 3.29 (d, *J* = 10.3 Hz, 5.3 Hz, 1H); **^{13}C NMR** (101 MHz, $CDCl_3$): δ 158.5, 144.8, 135.9, 135.9, 130.1, 128.2, 127.9, 126.8, 113.2, 86.5, 80.0, 79.5, 78.0, 77.3, 77.0, 76.7, 74.7, 74.4, 70.8, 70.7, 63.2, 58.6, 57.9, 55.2.

Preparation of 4-(bis(4-methoxyphenyl)(phenyl)methoxy)-1,3-bis(prop-2-yn-1-yloxy)butan-2-yl(2-cyanoethyl) diisopropylphosphoramidite (7)

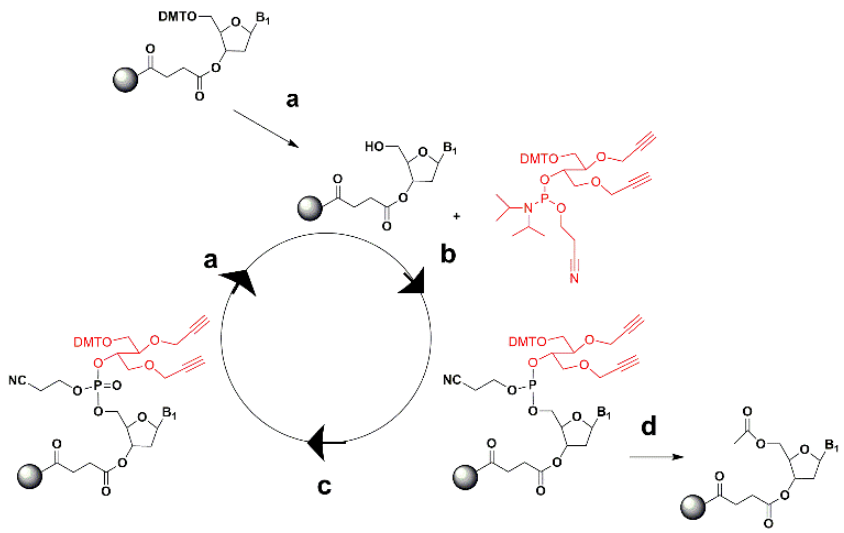


Compound **6** (0.17 g, 0.034 mmol) was coevaporated with dry dichloroethane (2x10 mL) and dissolved in anhydrous dichloromethane 16 mL. Diisopropylammonium tetrazolide (60 mg) and 2-cyanoethyl-N,N,N',N'-tetraisopropylphosphoroamidite (110 µL) were added and the reaction mixture was stirred under argon for 16 h. The mixture was concentrated under reduced pressure and the crude material was purified by flash column chromatography (20→50% ethyl acetate in petroleum ether) to give compound **7** (0.2 g, 0.028 mmol) as a light yellow oil.
Yield: 82%. ***Rf*** = 0.71 (50% ethyl acetate in petroleum ether).
HRMS-ESI (M+Na$^+$): found: m/z: 723.3170 , calcd: 723.3200.
^{31}P NMR (162 MHz, $CDCl_3$): δ 150.15, 150.03.

Oligonucleotide synthesis
Oligonucleotides K1 and K2 were synthesized on automated DNA synthesizer PerSpective Bio-systems Expedite 8909. Synthesis was done using standard protocol in a DMT-ON mode, scale of 1.0 µmol on 500 Å CPG support. LNA monomers and new bisalkyne nucleic acid scaffold were incorporated using hand-coupling procedure. These reagents were dissolved in acetonitrile and coupled for 18 min using 1*H*-tetrazole as an activator. Coupling efficiencies were measured using absorbance of the dimethoxytrityl cation realized after each coupling. Cleavage of the oligonucleotides from the CPG solid support was performed with 30% aqueous ammonia solution at 55 °C, overnight. Detritylation of the oligonucleotides was done with 80% acetic acid, 30 min, followed by addition of water, sodium acetate 3M, sodium perchlorate 5M and precipitation with cold acetone (-20 °C, 10 min). Work up was finished by washing two times with cold acetone. The identity and purity of ONs were verified by MALDI-TOF mass spectrometry and IE HPLC, respectively.

Bioconjugation
The oligonucleotide-dye conjugates (ODCs) were synthetized by CuAAC "click" reaction between the K1/K2 oligonucleotide modified with the bisalkyne nucleic acid scaffold and azide reagents: perylene, 5JOE and phenylethynylpyrene (Scheme 3). 10 nmol Oligonucleotide was mixed with 1 M triethylammonium acetate buffer, 200/100 nmol azide dye for K1 and K2, respectively, 10 µL 10 mM Cu (II):TBTA and 4 µL 50 mM ascorbic acid solution in a final volume of 200 µL, DMSO:MQ water, 0.8:1.2. The reaction mixture was deaerated, heated on 75 °C for 10 min, placed in a microwave reactor for 45 min, 60 °C and afterwards kept at rt overnight. The products were purified on a Illustra NAP-5 column (GE Healthcare), using the protocol from the manufacturer. The obtained solution was dried under N_2 flow and dissolved in Mili-Q water. Concentration of ODCs was calculated by measuring the absorbance at 260 nm. The products were analyzed by MALDI-TOF mass spectrometry and IE HPLC.

Scheme S1. Oligonucleotide synthesis: main steps and coupling of the bisalkyne scaffold



*Reagents and conditions for solid-phase oligonucleotide synthesis: (a) CCl_3COOH, DCM (3/100, v/v); (b) 1H-tetrazole, CH_3CN, DCM 18 min; (c) THF, H_2O, pyridine, I_2, (90.54/9.05/0.41/0.43, v/v/v/v); (d) $(CH_3CO)_2O$, THF (Cap A, 9.1/90.9, v/v), THF, $C_4H_6N_2$, pyridine, (Cap B, 8/1/1, v/v/v).

Table S1. Commercial DNA and RNA targets used in this study*

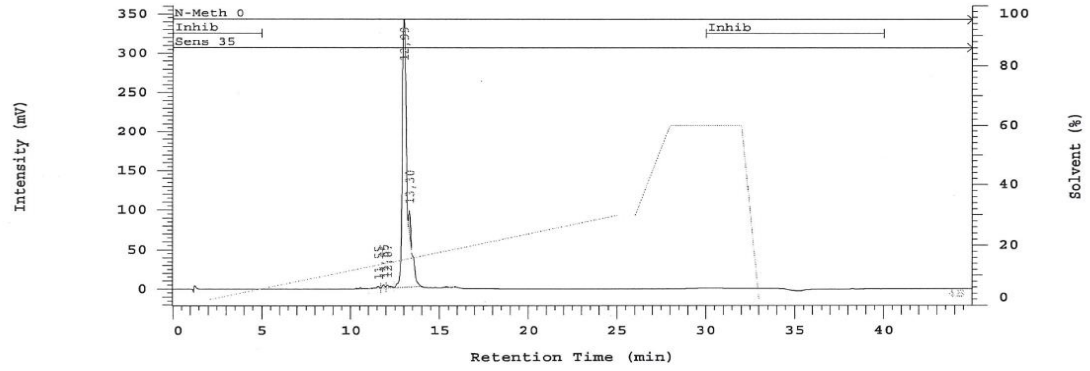
#	Sequence
DT1	5'- TCT AGC TAC AGA GAA ATC TCG-3'
DT2	5'-TCT AGC TAC AGT GAA ATC TCG-3'
RT1	5'-r(UCU AGC UAC AGA GAA AUC UCG)-3'
RT2	5'-r(UCU AGC UAC AGU GAA AUC UCG)-3'
DT1sh	5'- AGC TAC AGA GAA ATC-3'
DT2sh	5'- AGC TAC AGT GAA ATC-3'
DT3A	5'- AGC TAC AAA GAA ATC-3
DT3U	5'- AGC TAC AUA GAA ATC-3'
DT3C	5'- AGC TAC ACA GAA ATC-3'
RT1sh	5'- r(AGC UAC AGA GAA AUC)-3'
RT2sh	5'- r(AGC UAC AGU GAA AUC)-3'
RT3A	5'- r(AGC UAC AAA GAA AUC)-3
RT3U	5'- r(AGC UAC AUA GAA AUC)-3'
RT3C	5'-r(AGC UAC ACA GAA AUC)-3'

* Mismatched nucleotides are shown in red.

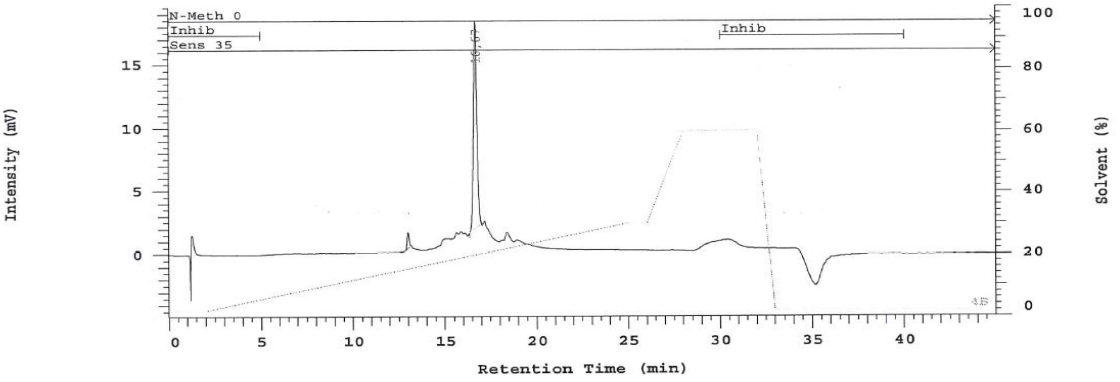
2. Characterization of oligonucleotide precursors and oligonucleotide-dye conjugates

Figure S1. Representative IC HPLC of the bisalkyne-modified oligonucleotides and ODCs

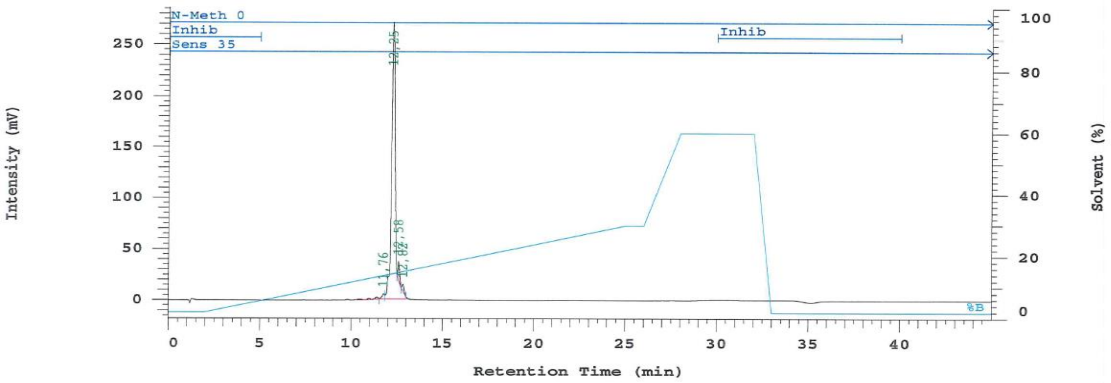
K1



K1-Per



K2



K2-Per

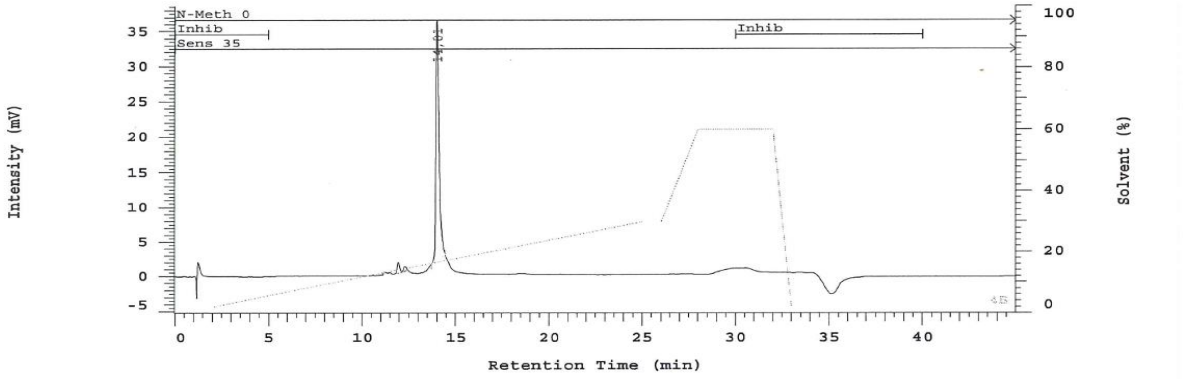
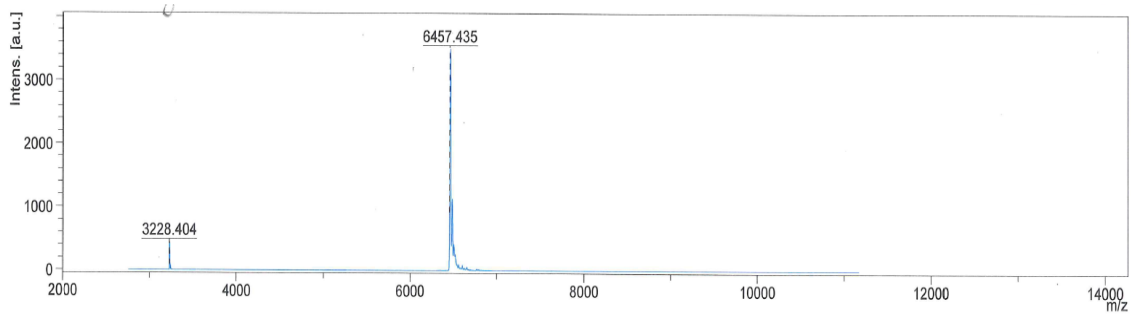
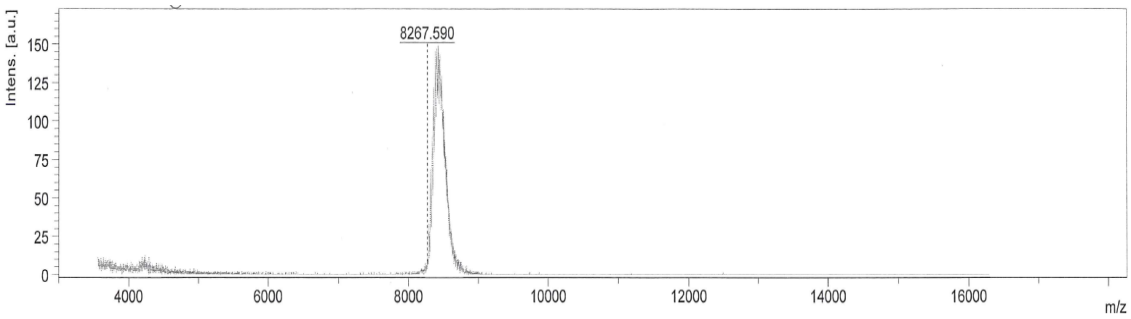


Figure S2. Representative Maldi-TOF MS of the bisalkyne modified oligonucleotides and ODCs

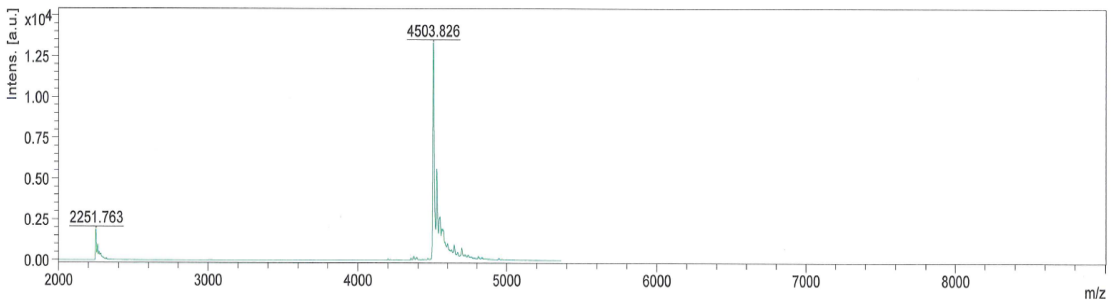
K1



K1-Per



K2



K2-Per

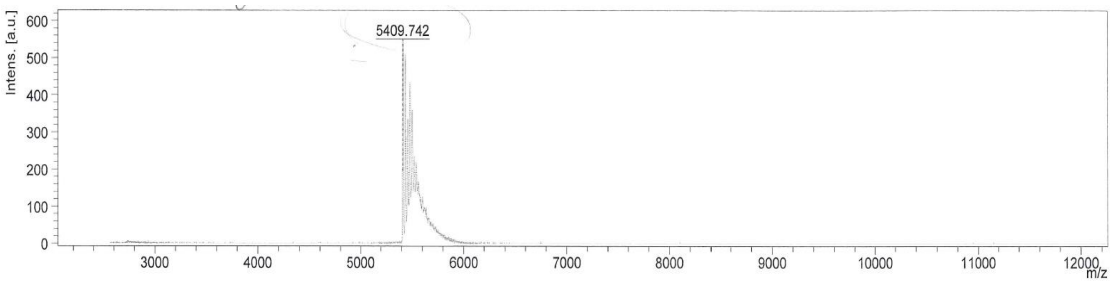


Table S2. Oligonucleotide dyes conjugates MALDI results

#	Mw calc.	Mw exp.
K1-Per	8267.0	8267.6
K1-5JOE	8806.2	8805.9
K1-PEP	8635.4	8637.3
K2-Per	5408.0	5409.7
K2-5JOE	5677.6	5677.8
K2-PEP	5592.2	5595.1

3. Thermal denaturation results (T_m) and representative curves for modified duplexes

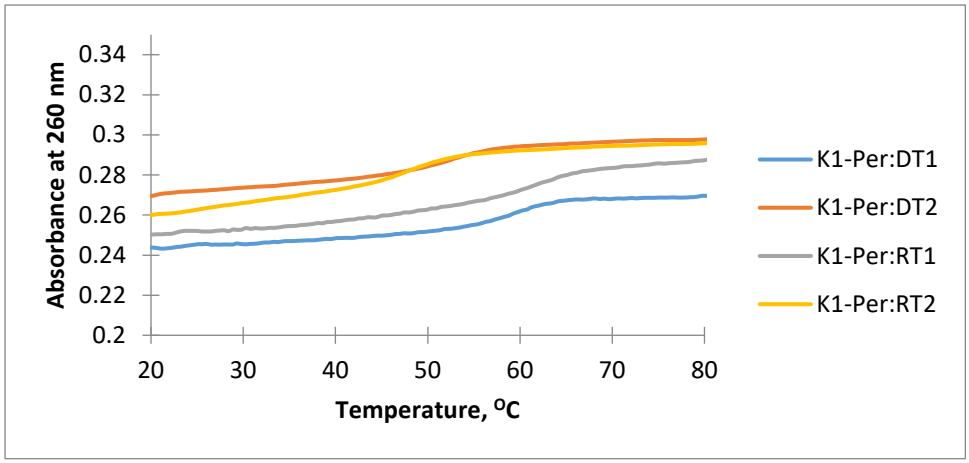
Table S3. T_m values*

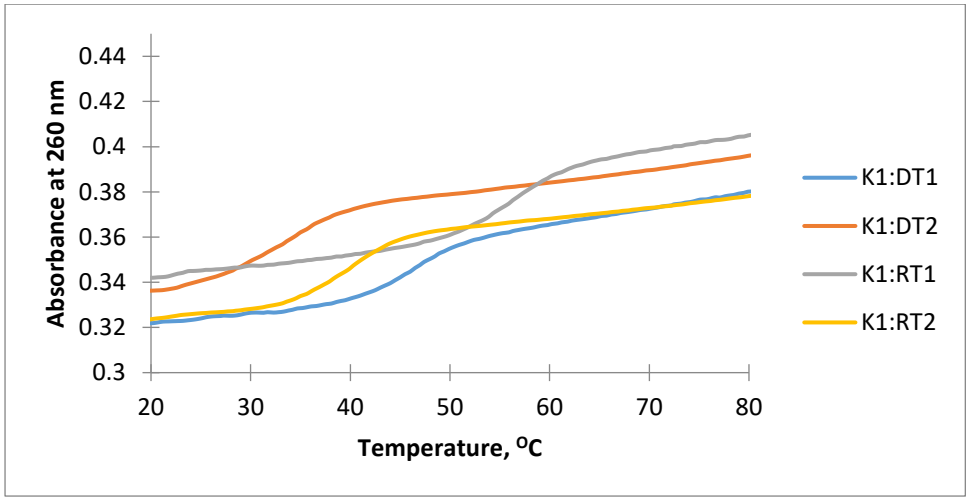
Probe/Target	DT1	DT2	RT1	RT2
K1	46.6	33.8	56.2	40.1
K1-Per	59.0	49.9	61.5	45.9
K1-5JOE	nct	nct	nct	nct
K1-PEP	nct	nct	nct	nct
RK1	62.7	53.4	67.3	55.2
K2	38.1	37.1	39.7	35.7
K2-Per	49.3	49.1	46.2	45.2
K2-5JOE	nct	nct	nct	nct
K2-PEP	nct	nct	nct	nct
RK2	49.5	42.8	52.2	41.9

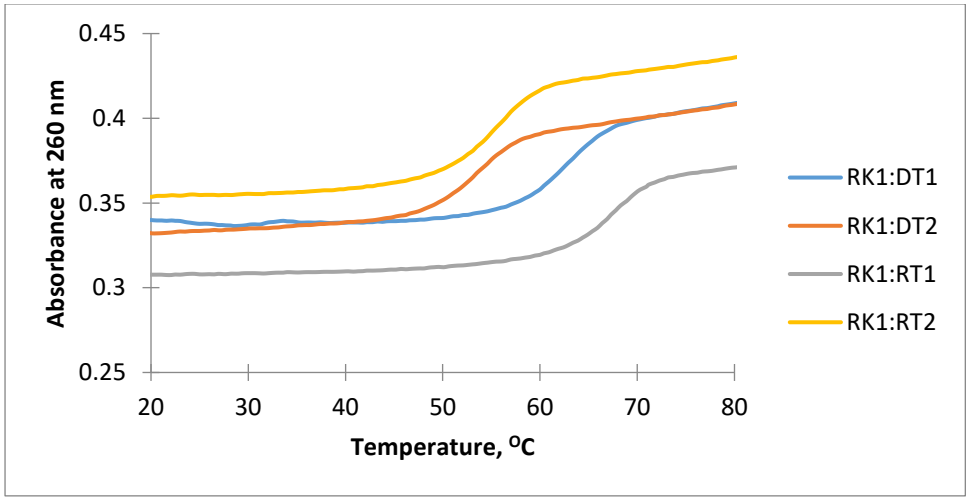
Probe/Target	DT1sh	DT2sh	DT3A	DT3U	DT3C	RT1sh	RT2sh	RT3A	RT3U	RT3C
K2	34.6	32.5	25.6	26.6	25.2	36.9	31.9	24.2	23.6	22.9
K2-Per	49.8	49.4	nct	nct	nct	42.5	43.1	nct	nct	nct
K2-5JOE	nct	nct	nct	nct	nct	nct	nct	nct	nct	nct
K2-PEP	nct	nct	nct	nct	nct	nct	nct	nct	nct	nct
RK2	48.6	39.7	33.6	35.6	32.6	50.6	39.6	35.4	34.0	32.4

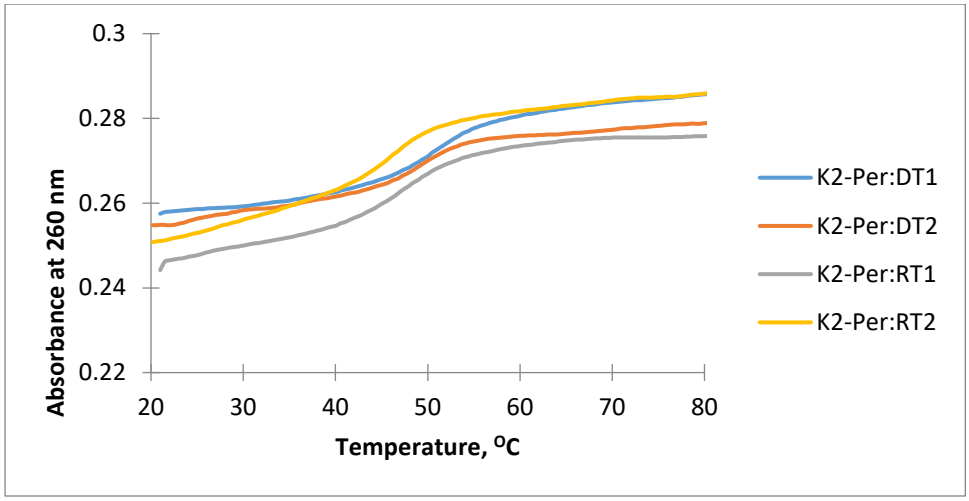
* nct = no clear transition

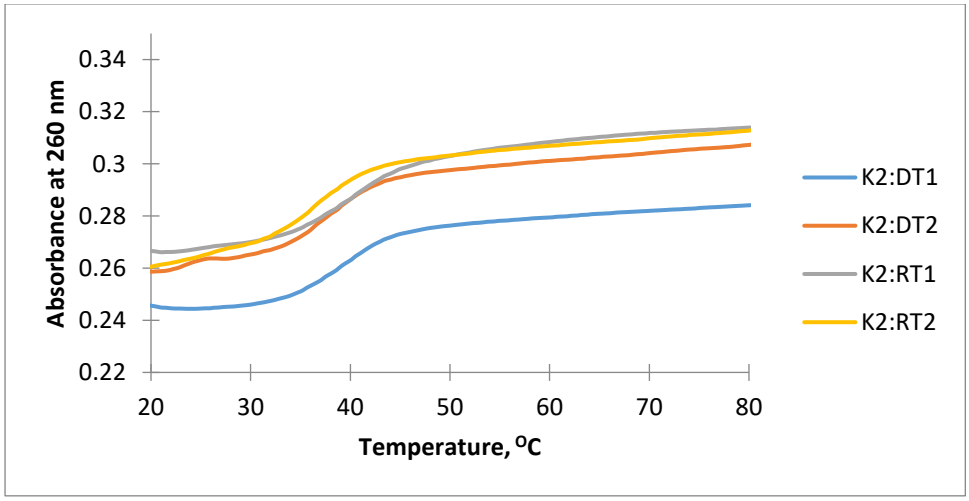
Figure S3. Representative T_m graphs

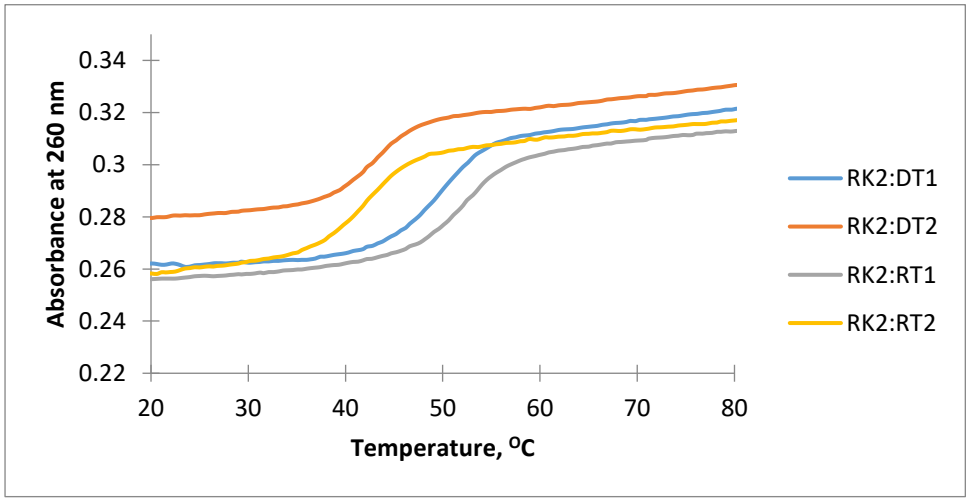






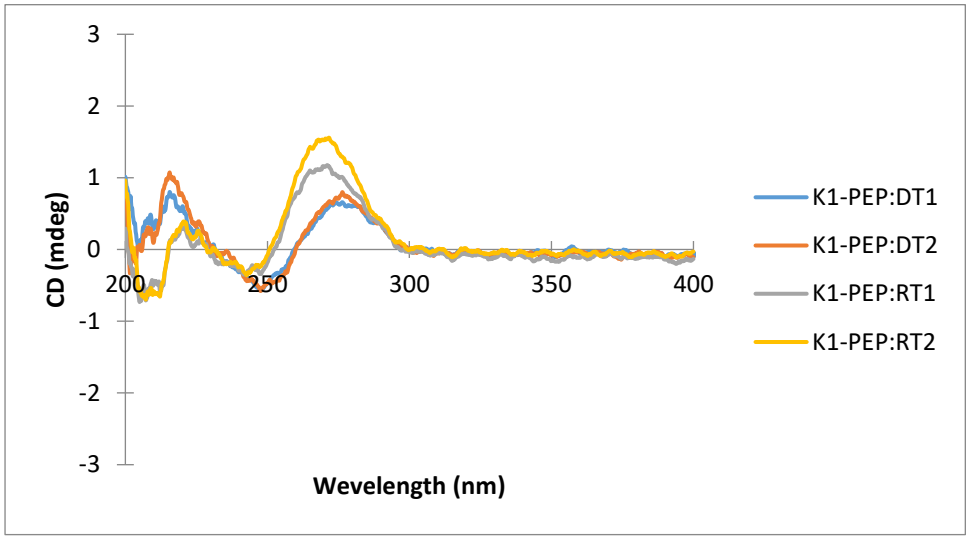

0.3
0.28
0.26
0.24
0.22
Absorbance at 260 nm
20
30
40
50
60
70
80
Temperature, °C
K2-Per:DT1
K2-Per:DT2
K2-Per:RT1
K2-Per:RT2

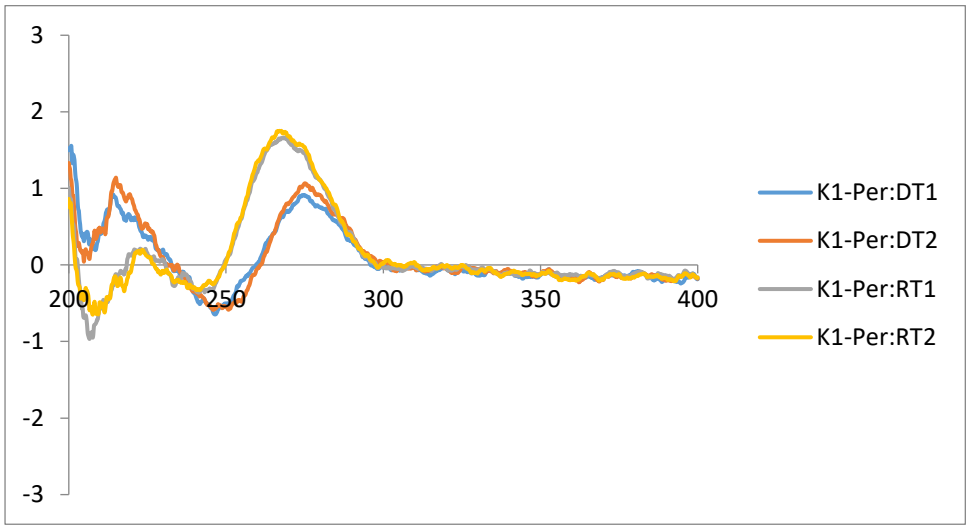

0.34
0.32
0.3
0.28
0.26
0.24
0.22
Absorbance at 260 nm
20
30
40
50
60
70
80
Temperature, °C
K2:DT1
K2:DT2
K2:RT1
K2:RT2

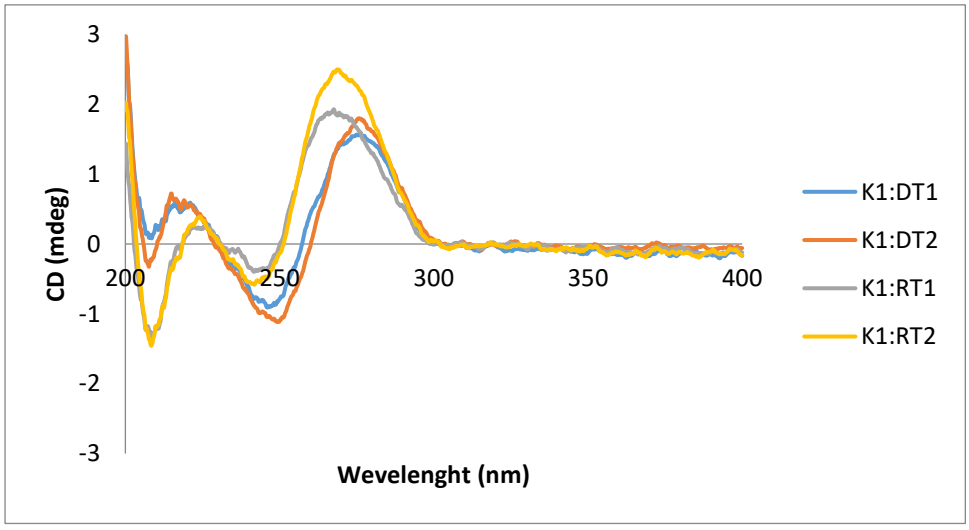

0.34
0.32
0.3
0.28
0.26
0.24
0.22
Absorbance at 260 nm
20
30
40
50
60
70
80
Temperature, °C
RK2:DT1
RK2:DT2
RK2:RT1
RK2:RT2

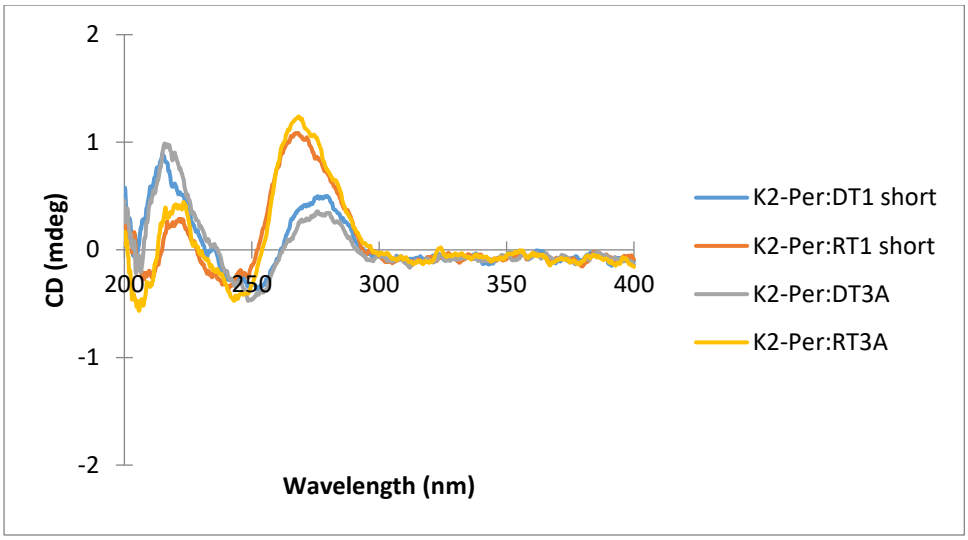
4. Circular Dichroism (CD) results and representative graphs

Figure S4. Representative CD graphs









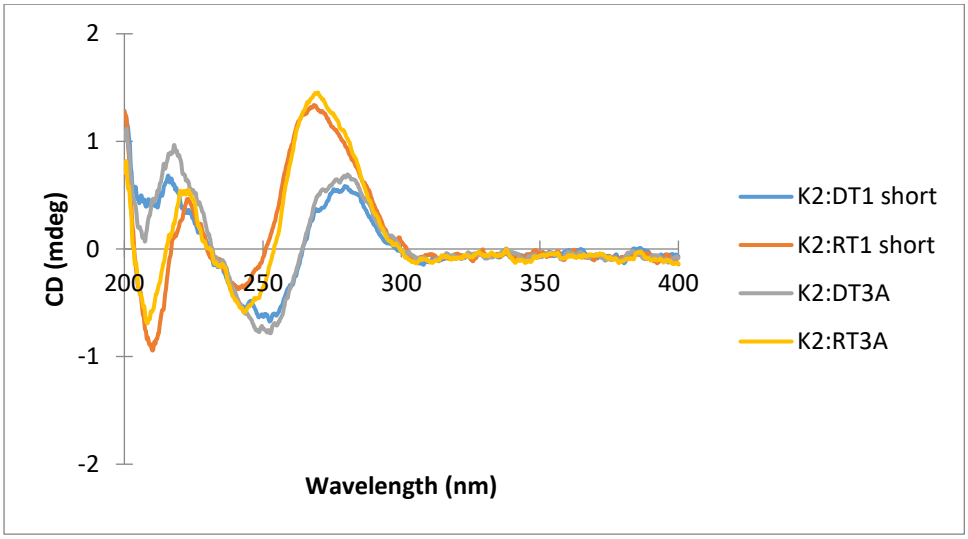


Table S4. CD results: numerical intensities at selected wavelengths.

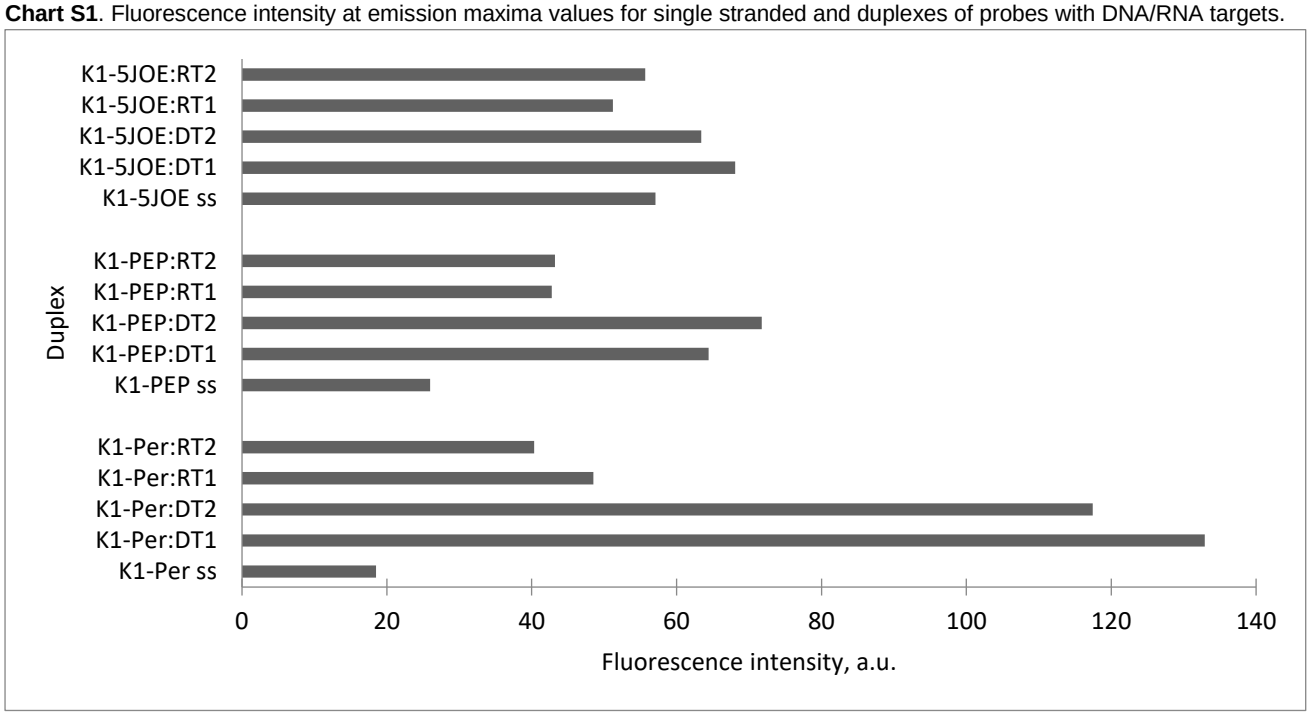
Probe	Target: DNA					RNA				
	Wavelength, nm:									
	205	220	245	275	**SUM**	210nm	225nm	240nm	265nm	**SUM**
RK1	0.00	0.25	-1.25	1,25	**0.25**	-2.25	0.00	-0.75	2.25	**-0.75**
K1	-0.25	0.50	-1.25	1.75	**0.30**	-1.50	0.25	-0.50	2.00	**0.25**
K1-Per	0.00	1.00	-0.5	1.00	**1.50**	-0.75	0.25	-0.25	1.75	**1.00**
K1-PEP	0.00	1.00	-0.5	0.75	**1.25**	-0.5	0.5	-0.50	1.25	**0.75**
K1-5JOE	0.00	0.75	-0.75	1.00	**1.00**	-1	0.25	-0.25	1.50	**0.50**

Table S5. CD results, numerical intensities at certain wavelength

Probe	Target: DNA					RNA				
	Wavelength, nm:									
	210	220	245	275	**SUM**	210	225	240	270	**SUM**
RK2	0.00	0.50	-0.75	0.75	**0.50**	-1.50	0.00	-0.25	1.50	**-0.25**
K2	0.25	0.75	-0.75	0.75	**1.00**	-0.75	0.50	-0.50	1.50	**0.75**
K2-Per	-0.25	1.00	-0.50	+0.5	**0.75**	-0.50	0.50	-0.50	1.25	**0.75**
K2-5JOE	0.00	1.00	-0.50	0.25	**0.75**	-0.50	0.50	-0.50	1.00	**0.50**

5. Fluorescence results

Chart S1. Fluorescence intensity at emission maxima values for single stranded and duplexes of probes with DNA/RNA targets.



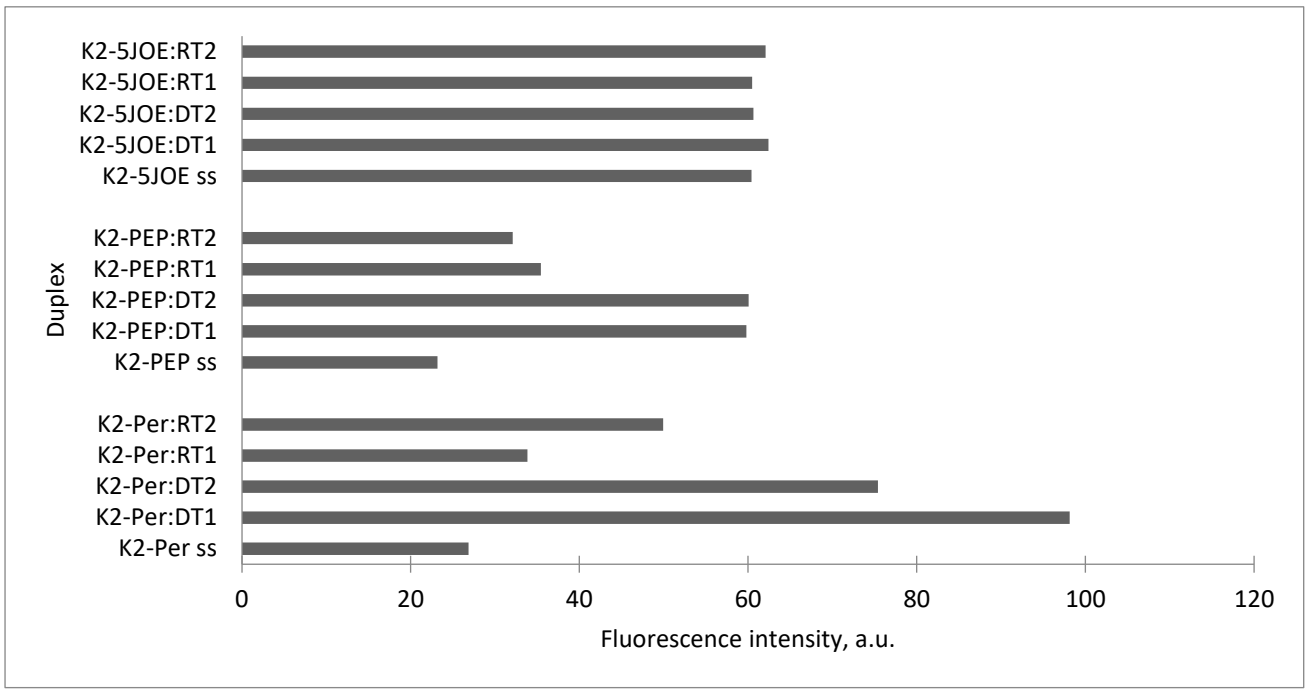
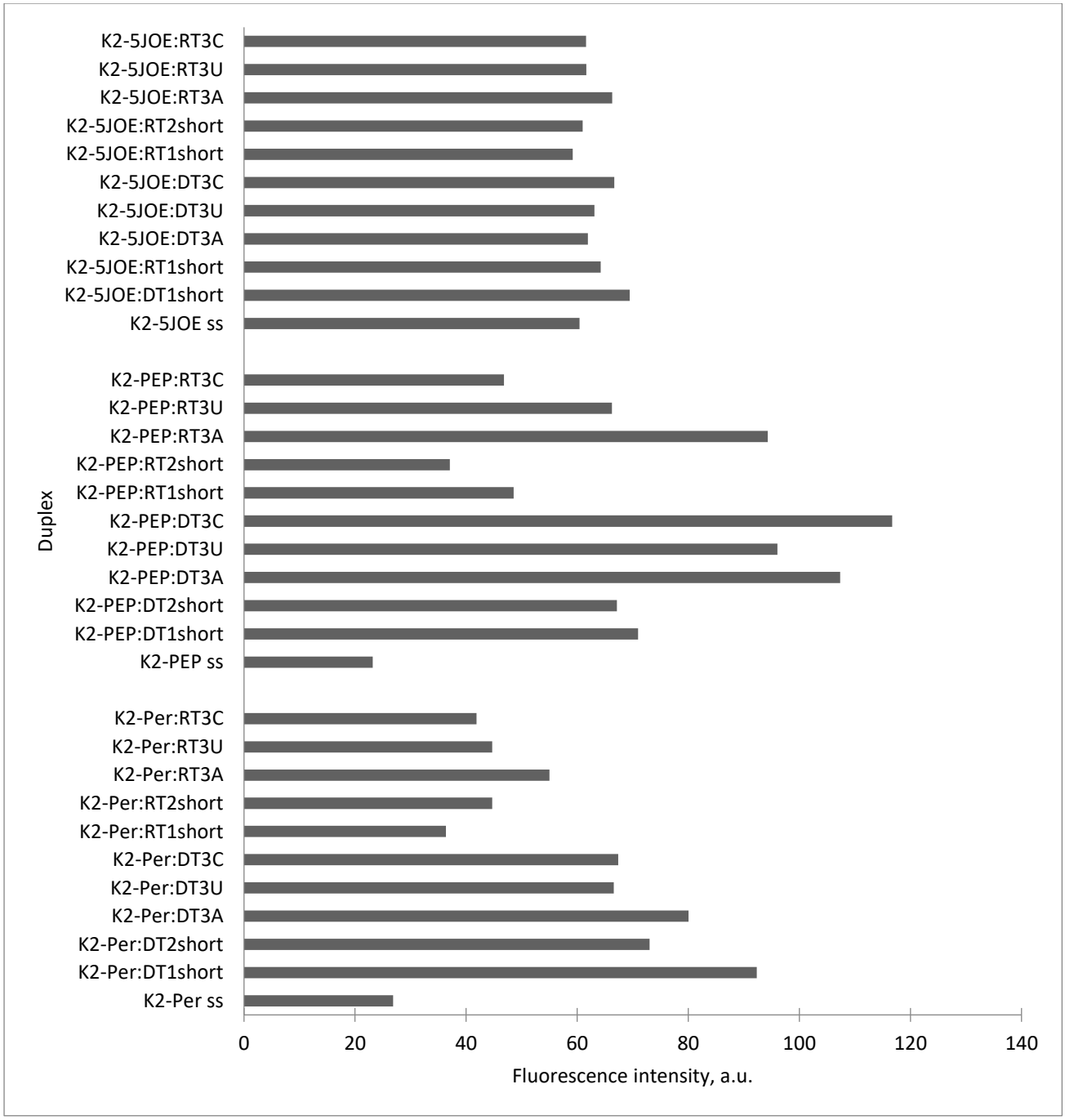


Chart S2. Fluorescence intensity at emission maxima values for single stranded and duplexes of probes with DNA/RNA targets.



6. Capturing assay

Initially, we designed target specific capturing and derection probes to *BRAF* V600E and *EGFR* L858R regions of human genome, following previously described strategy.[1]
The resulting probes were as follows:

BRAF capturing probe:
5'-/Biosg/ GAA AAT ACT ATA GTT GAG ACC TTC AAT GAC TTT CTA GTA ACT CAG CAG CAT CTC AGG GCC AAA AAT TTA ATC AGT GGA AAA ATA GCC TCA ATT CTT ACC ATC CAC AAA ATG GAT CCA GAC

EGFR capturing probe:
5'-/Biosg/ CTG GAG AGC ATC CTC CCC TGC ATG TGT TAA ACA ATA CAG CTA GTG GGA AGG CAG CCT GGT CCC TGG TGT CAG GAA AAT GCT GGC TGA CCT AAA GCC ACC TCC TTA CTT TGC CTC CTT CTG

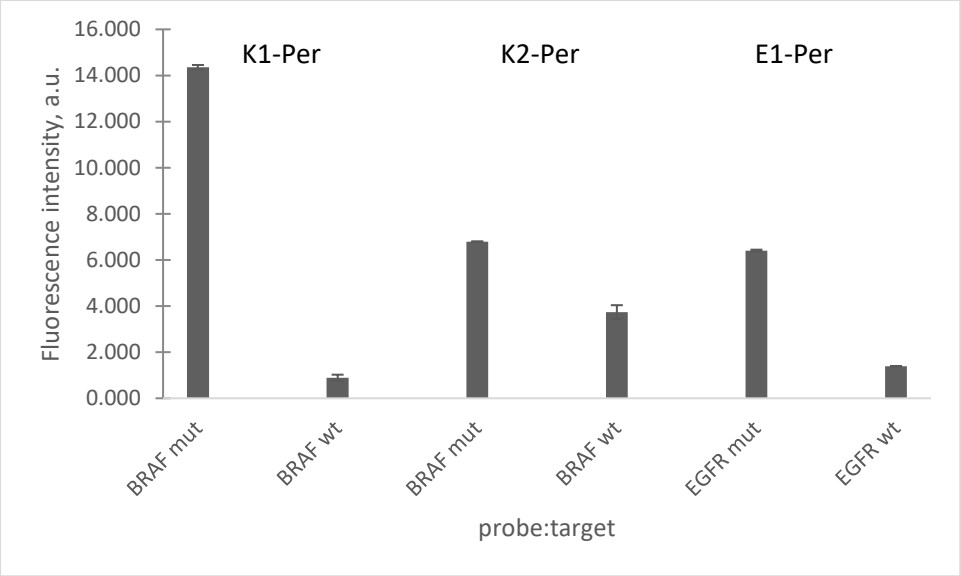
E1-Per:
5'-/Perylene/ GCA GTT TGG C+C+C +GCC CAA AAT Mw calc: 7668,5 g/mol; Mw exp:7665.0 g/mol

The sequences of *BRAF* V600E specific probes K1 and K2 are given in Table 1 (main article).

For the assay, we used purified genomic DNA from A101D – human skin malignant melanoma cell line (mutated *BRAF*, ATCC) and NCI-H1975 – human long non-small cell line carcinoma (mutant *EGFR*, ATCC). As a negative control we used DNA extracted from healthy cell line HMC-1 (ATCC). The capturing probes were purchesed by Qiagen and the E1-Per probe was synthetised on a Expedite as described above (handcoupled with the bisalkyne scaffold) and bioconjugated with the perylene-azide using CuAAC click chemistry (described above). For fluorescence quantification we used LightCycler® 480 Real-Time PCR System (Roche), LightCycler® Cyan 500 chanel (440 nm-488 nm).

[1] Miotke, Astakhova et al. PLoS ONE 2015.

Chart S3. Capturing assay results, K1-Per, K2-Per, E1-Per and the HMCs respectively.*

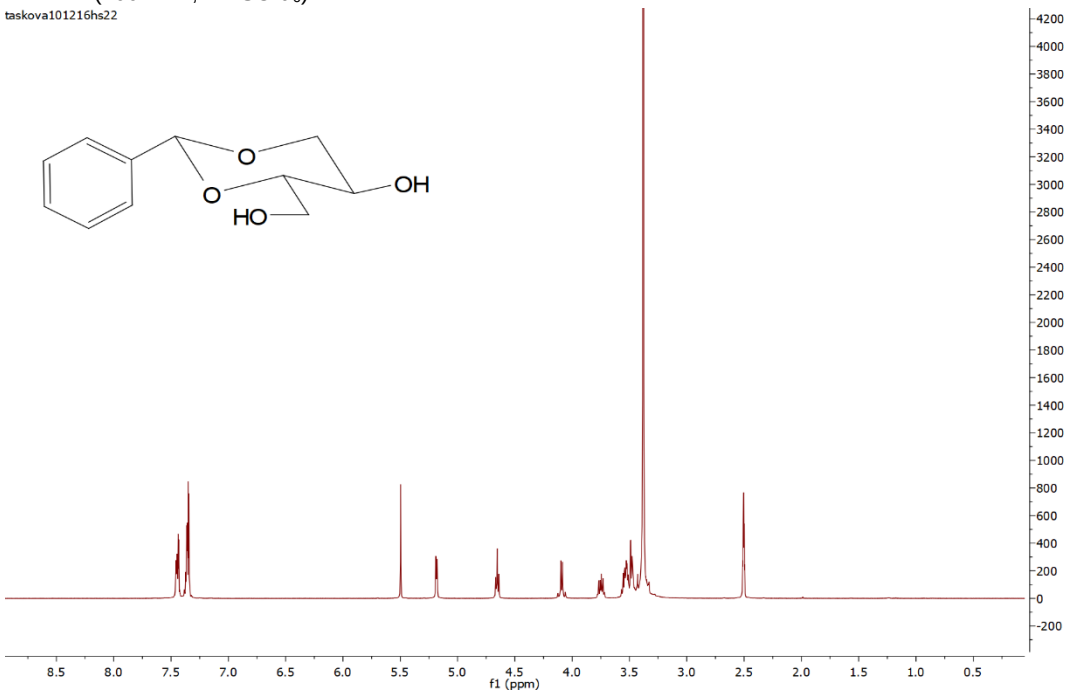


* Each data point is an average of two independent measurements with CV < 5%.

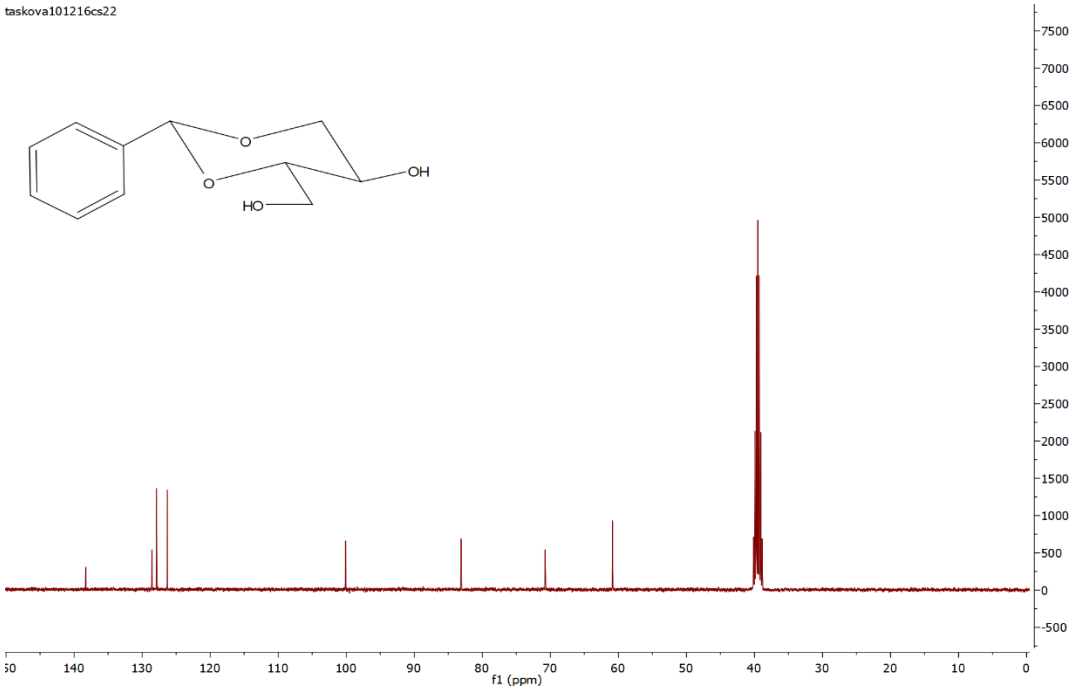
7. ^{1}H and ^{13}C NMR spectra

Preparation of 1,3-O-benzylidene-L-erythriol (3)

^{1}H NMR (400 MHz, DMSO-d_6)

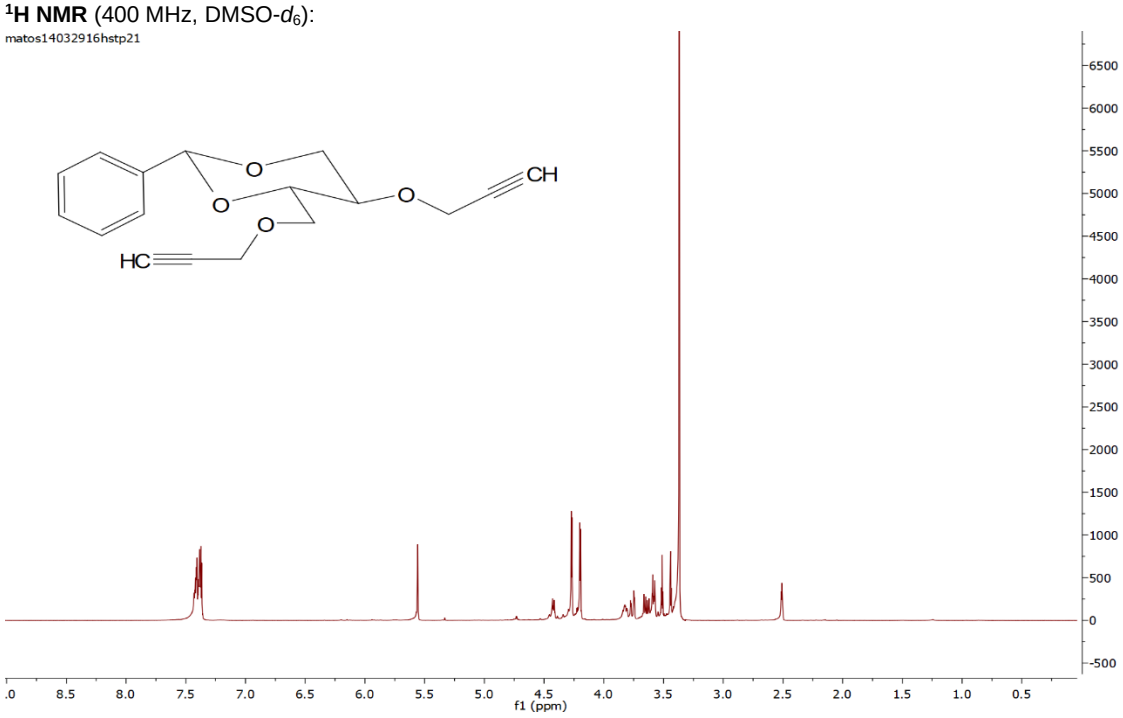


^{13}C NMR (101 MHz, DMSO-d_6)

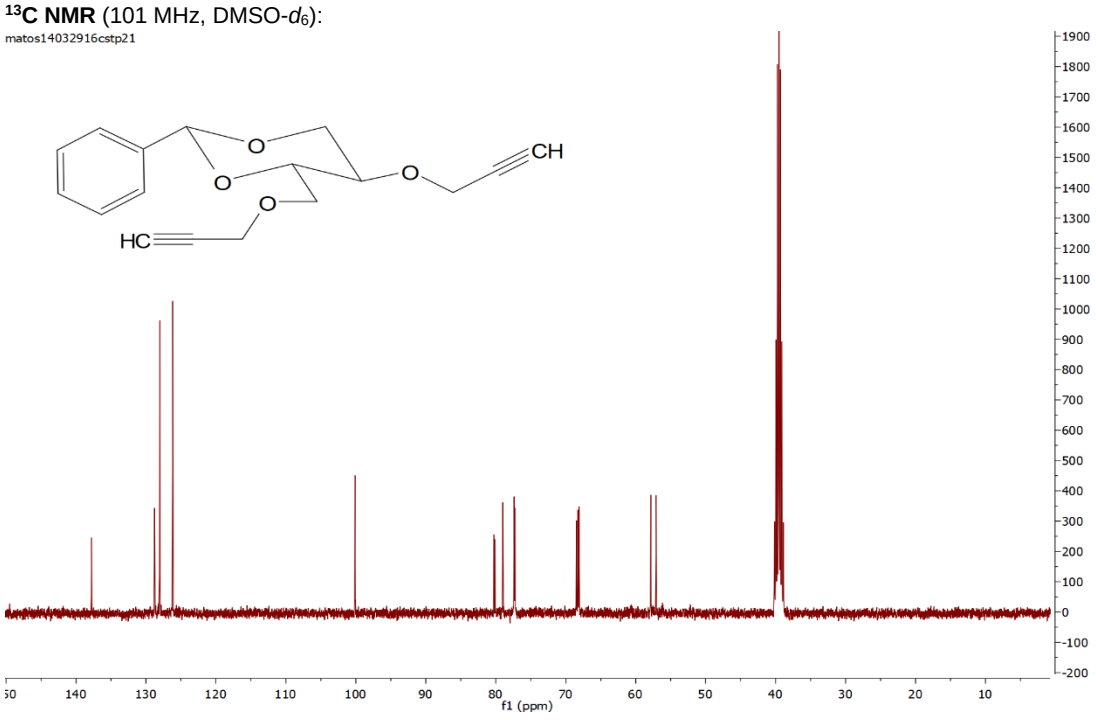


Preparation of (2R,4S,5R)-2-phenyl-5-(prop-2-yn-1-yloxy)-4-((prop-2-yn-1-yloxy)methyl)-1,3-dioxane (4)

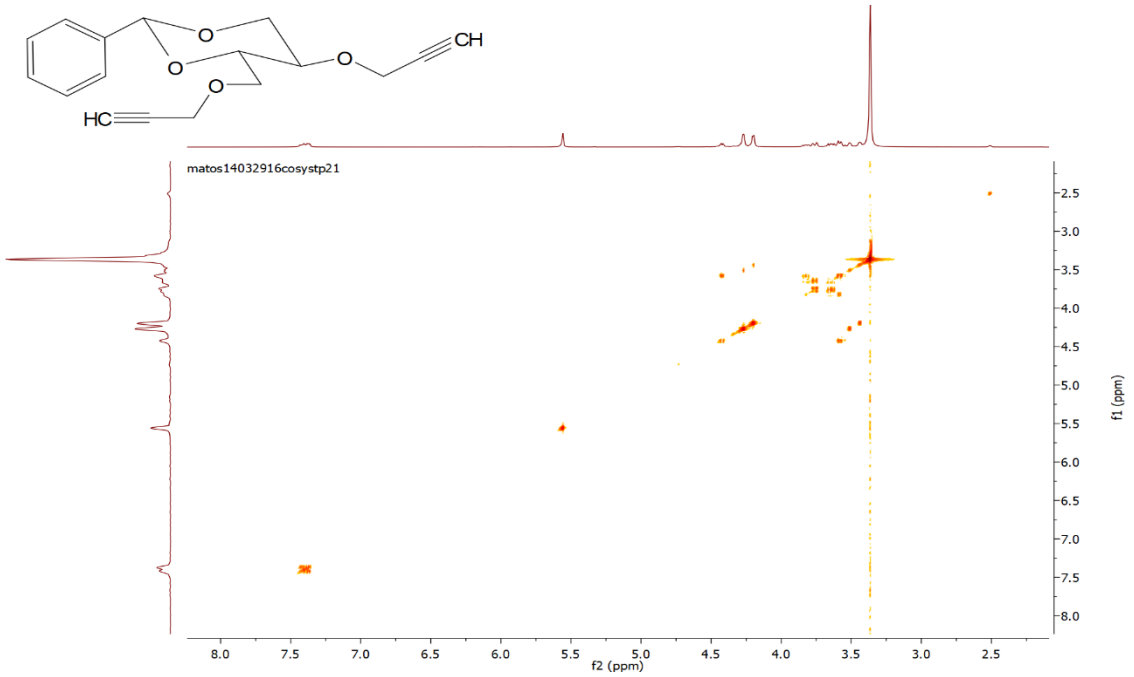
^{1}H NMR (400 MHz, DMSO-d_6):



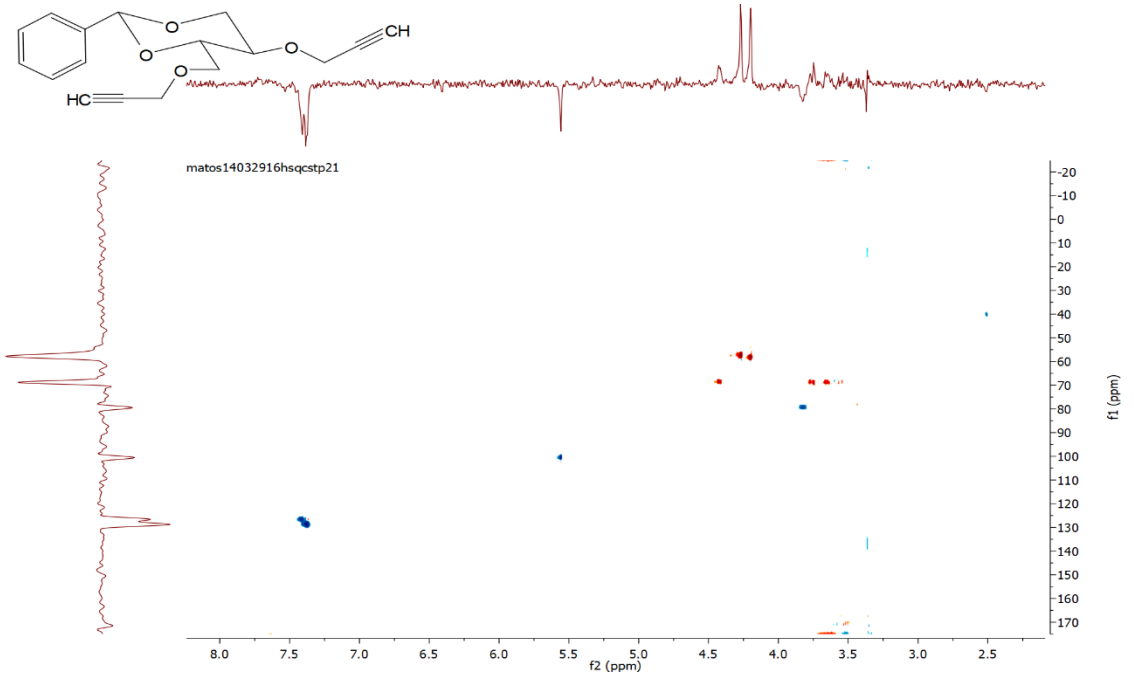
^{13}C NMR (101 MHz, DMSO-d_6):



COSY (400 MHz, DMSO-d_6):

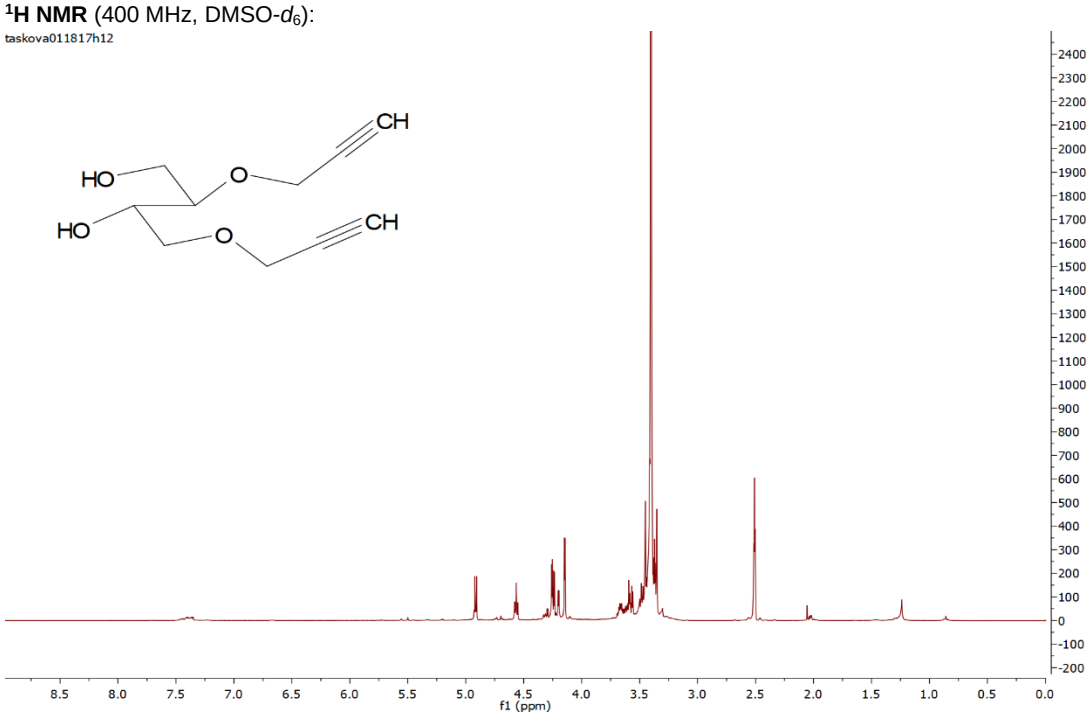


HSQC (400 MHz, DMSO-d_6):

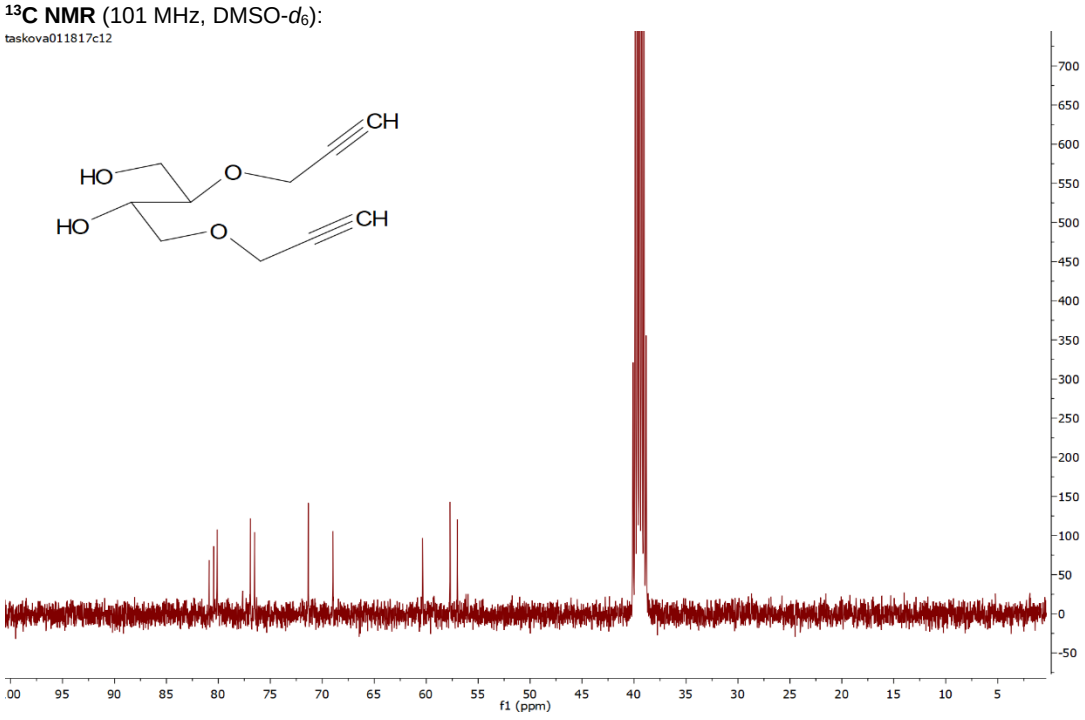


Preparation of 2,4-bis(prop-2-yn-1-yloxy)butane-1,3-diol (5)

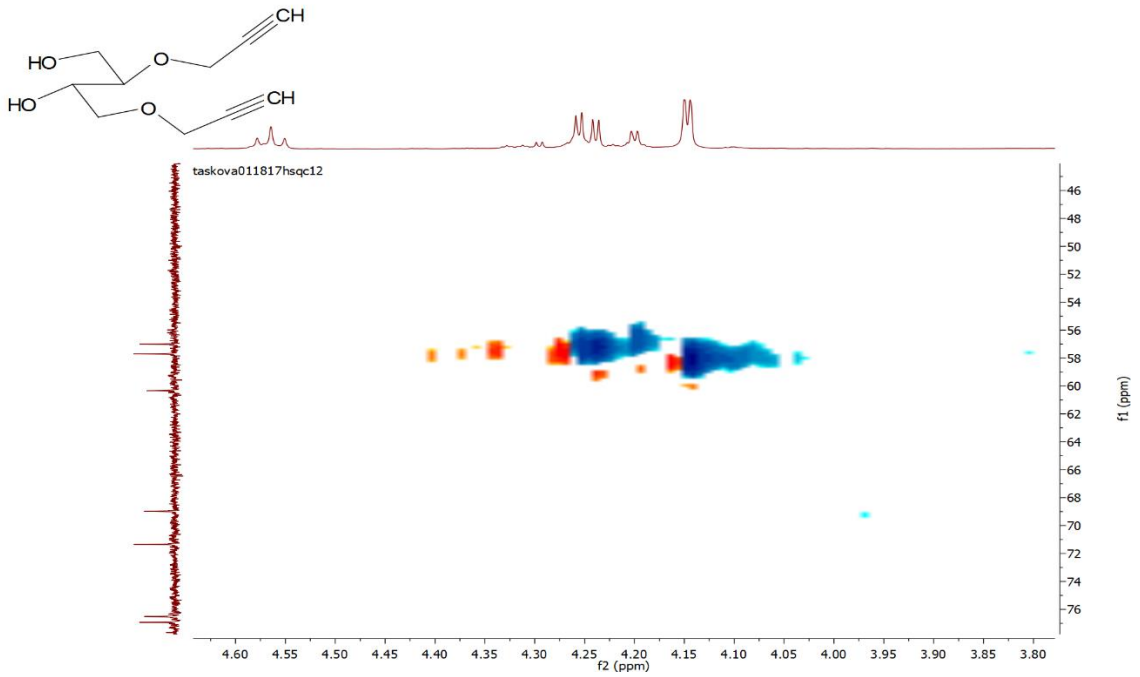
^{1}H NMR (400 MHz, DMSO-d_6):



^{13}C NMR (101 MHz, DMSO-d_6):

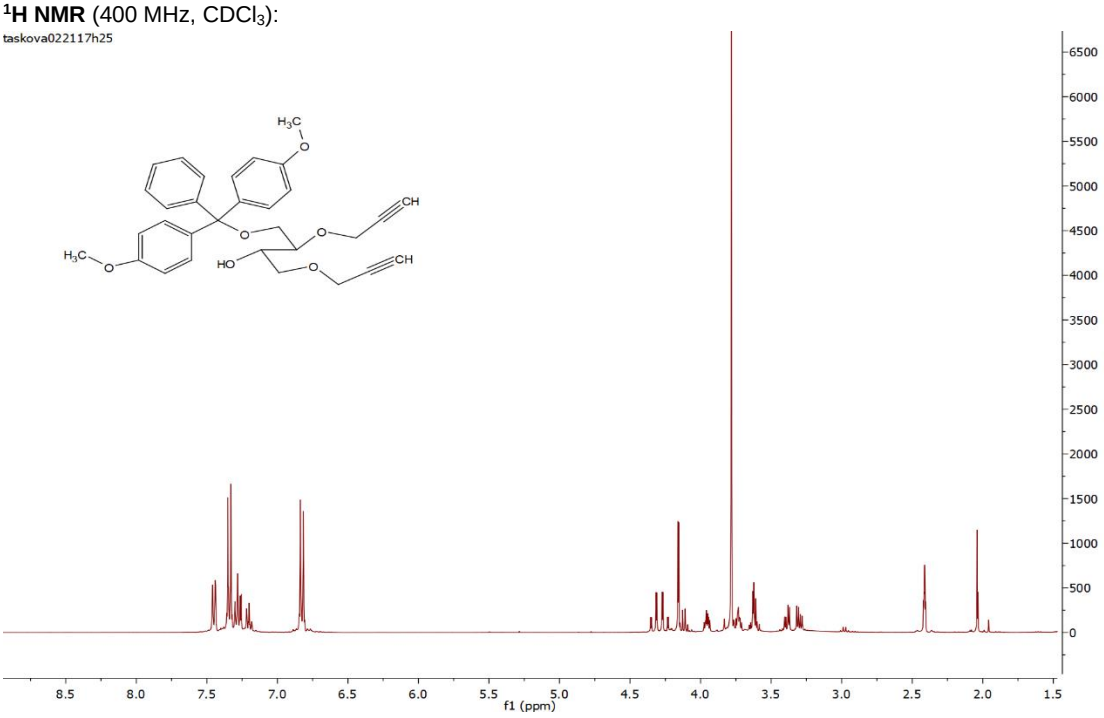


HSQC (400 MHz, DMSO-d_6):

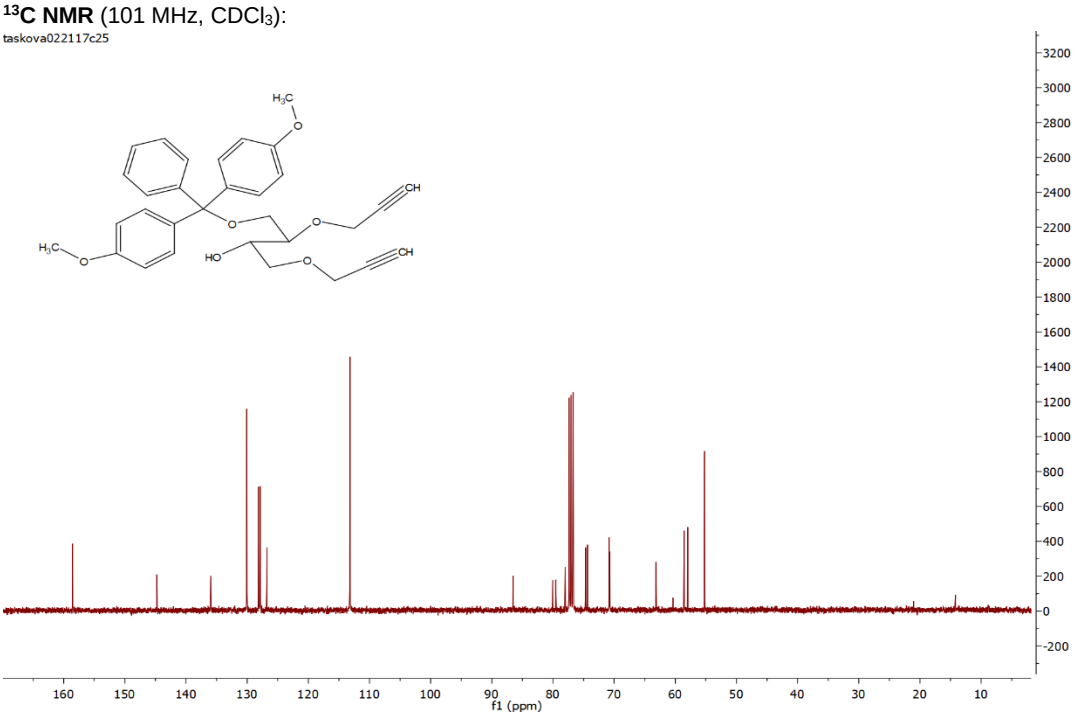


Preparation of 2,4-bis(prop-2-yn-1-yloxy)butane-1-O-dimetoxytrityl-3-ol (6)

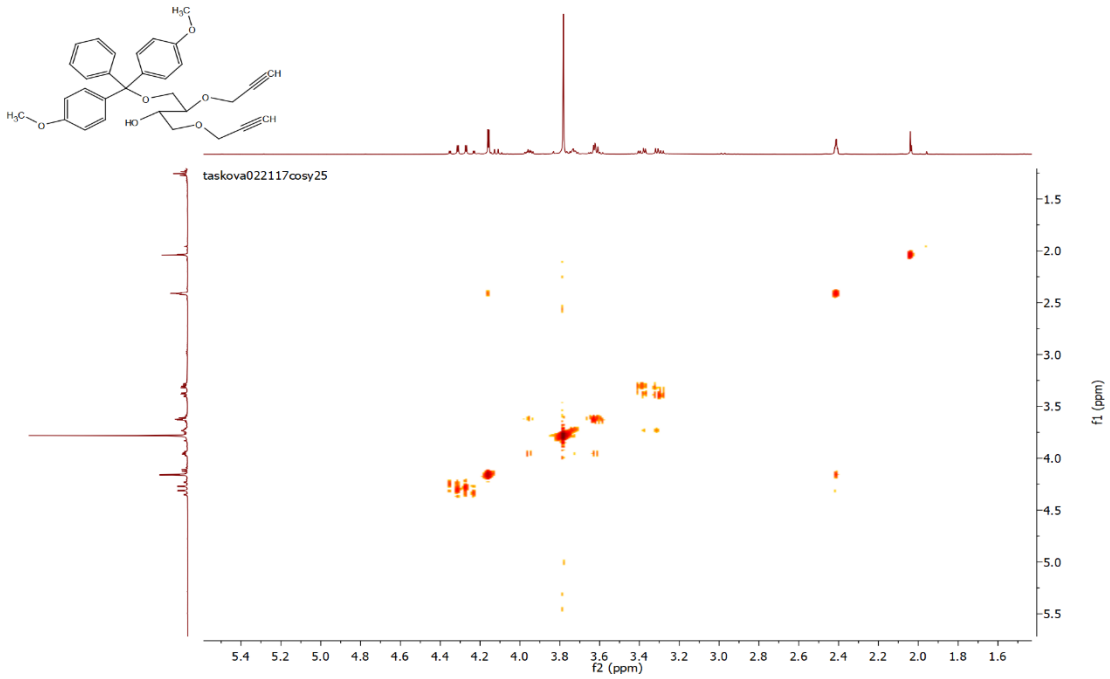
^{1}H NMR (400 MHz, $CDCl_3$):



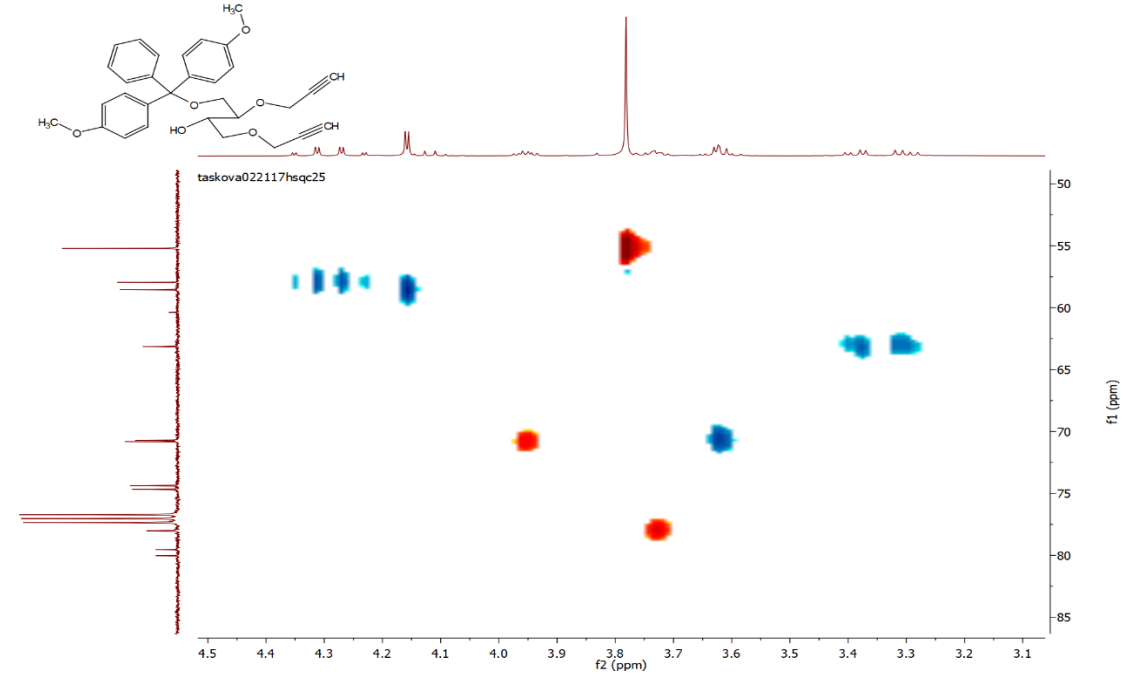
^{13}C NMR (101 MHz, $CDCl_3$):



COSY (400 MHz, $CDCl_3$):

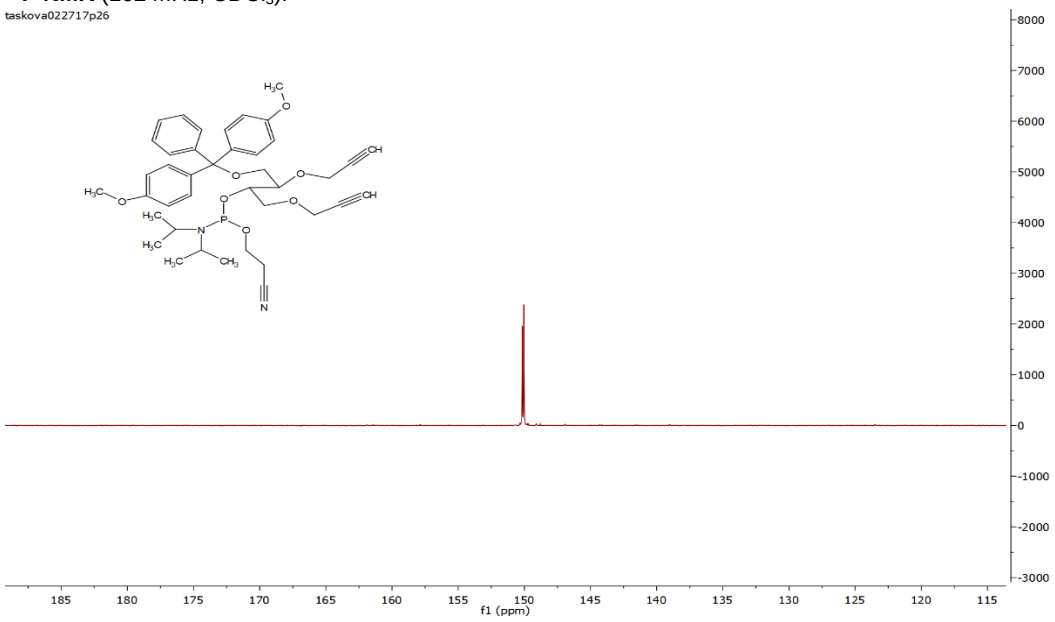


HSQC (400 MHz, $CDCl_3$):



Preparation of 4-(bis(4-methoxyphenyl)(phenyl)methoxy)-1,3-bis(prop-2-yn-1-yloxy)butan-2-yl(2-cyanoethyl) diisopropylphosphoramidite (7)

^{31}P NMR (162 MHz, $CDCl_3$):



Chapter 7

Summary of Results and General Discussion

The studies presented in this thesis come across the main objective to develop advanced short synthetic oligonucleotide sequences with high potential for enhanced performance as therapeutics and as diagnostics.

Realizing the potential of synthetic oligonucleotides as therapeutics, the interest of new modified oligonucleotide sequences and their properties is constantly growing. While, almost all FDA approved oligonucleotide therapeutics contain certain backbone or sugar modification, bioconjugates of the therapeutic oligonucleotides are amplifying its presence in the pharmaceutical development pipelines. Ideal biomolecule, when conjugated to an oligonucleotide, should enhance its properties such as efficient target binding, improved serum stability, targeted cell internalization and enhanced pharmacokinetics.

First, we synthetized a small library of POCs – Paper II. We aimed to characterize and investigate the influence of different peptide sequences on the oligonucleotide properties *in vitro*. Since the conjugation between negatively charged DNA and cationic peptides can be challenging,[158] we optimized the method for successful synthesis of a library of thirty POCs. First, we synthetized the LNA modified and alkyne functionalized DNA oligonucleotide sequences, designed to target the *BRAF V600E* oncogene. This was followed by "CuAAC" internal conjugation with ten different peptide sequences to create POCs, each containing the same peptide sequence twice. Next, we performed a set of biophysical studies (T_m and CD) to investigate the specificity towards target binding and the structural properties of the POCs duplexes with DNA and RNA matching and mismatching targets. We finalized the studies by performing serum stability investigation for the POCs and their duplexes with DNA.

The results confirmed efficient duplex formation of the POCs with the DNA and the RNA targets. According the CD investigation, all of the POCs showed mixed A/B geometry correlated to the not conjugated DNA/LNA starting oligonucleotide. In addition, the T_m experiments revealed melting curves with *S*-shape, characteristic for the unmodified reference sequence. Next, the statistical analyses of the T_m values confirmed that LNA have important role into SNP discrimination. In addition, the nature of the attached peptide influenced the target binding (DNA and RNA) and the ability for mismatched discrimination. The data demonstrated that the

peptide sequence feature is important factor that affect the target binding. Chapter 3 (Paper II) contain detailed discussion on the effect from particular amino acid presence and the distance between the oligonucleotide and the particular amino acid. Furthermore, the conjugation with peptides showed superior improvement of the stability of the oligonucleotides in 90% human serum when compared with just the presence of LNA.

Our study adds to the basic understanding on the structure and the properties of POCs *in vitro*. While literature contain numerous reports on the delivery performance of different peptides conjugated with ASOs and siRNAs, not many elaborate on the target-binding efficacy and the resistance towards enzymatic degradation. For example, Victor C. Yang and coworkers, report on the synthesis and intracellular delivery on siRNA conjugated with low molecular weight protamine peptides.[170] They demonstrate that usage of polyethylene glycol (PEG) spacer and cleavable disulfide linker is improving the cellular uptake of siRNA. However, they point out that new siRNA design with optimized stability may advance the system even more.[170] We speculate that the best performing POCs in the *in vitro* assays may give superior results in *in vivo* studies. The most promising POCs should be further tested for *in vivo* distribution and performance.

The most common routes for administration of oligonucleotide therapeutics are injection based. Since, the oral delivery route is very challenging, in clinical trials oligonucleotide therapeutics are most often administrated by intravenous, subcutaneous, intravitreal or intratumoral pathway. However, considering the advantages of oral drug delivery such as large surface area of absorption and high patient comfort, the scientific community starts to consider oral delivery systems for ASOs and siRNAs as prerequisite. An orally formulated oligonucleotide drug, should be stable in the acidic environment of the gastrointestinal tract (GIT) and resistant to nuclease degradation until it reach the intracellular compartment.[171] Moreover, the used additional excipients should be inert, biocompatible and nontoxic.

Attempts to orally deliver siRNA and ASO therapeutics were done just for GIT diseases such as intestinal inflammation and colorectal liver.[172–174] Following the need of stable oligonucleotide drugs, we investigated the stability of differently formulated, therapeutic relevant oligonucleotides in simulated bio-fluids (Paper III, Chapter 4). In this study, we examined the ability of short peptides (covalently attached), poly-L-lysine and histone proteins (complexed) to alleviate the oligonucleotides stability in simulated bio-fluids such as human serum, gastric juice and hydrochloric acid. The probes were first radiolabeled with ^{32}P and subsequently incubated in the bio-fluids for 24 h deducting small volume portion at particular time points.

Next, the fractions were resolved on PAGE gels and imaged. For quantitative comparison, we estimated the remaining fractions of the naked oligonucleotides, the conjugates and the complexes computing the band intensities.

Incubation in 90% human serum revealed that stability of naked oligonucleotide depend on the length and the nature of the sequence. However, formulation of oligonucleotide-peptide conjugates and complexes, in general improved the stability. In addition, the covalent complexes with short peptides showed superior protection ability compared with the noncovalent complex with poly-L-lysine. Interesting, in gastric juice the net positive charge of the conjugated peptide sequence influenced the stability of the oligonucleotide. The less positively charged peptide protected the oligonucleotide from degradation more efficient. However, blockage of the pepsin action by a specific inhibitor - pepstatin A, in general enhanced the stability of the oligonucleotides. In addition, complexation of histone proteins to the longer oligonucleotide sequence displayed increased stability in hydrochloric acid. This study, together with small number of other studies of stability of nucleic acids in bio-fluids indicate that some proteolytic enzymes may play role into the degradation of the nucleic acids.[175,176] To meet the requirements for orally available oligonucleotide therapeutics, further studies on the structure-formulation-stability relationship will be necessary. Weakness of this study is the use of simulated bio-fluids. Detail investigation of the stability of various formulated oligonucleotide therapeutics in freshly extracted human serum and gastric juice from human stomach lumen are necessary to confirm the indications and to give more details on the degradation mechanism.

Chapter 5 covers two research projects where we designed and synthetized siRNA and ASOs therapeutics formulated as complexes and conjugates with peptide sequences to investigate their performance in cell culture studies. In the first project, we designed two siRNA therapeutics that target HIV 1 LTR 362 with the aim to inhibit the translation of the viral RNA. First we complexed the siRNAs with two peptide sequences and a GalNAc molecule. We examined the intracellular delivery and efficacy in three model cell lines. While, the siRNAs were not able to reduce the expression of the model target genes in LChIT and HEK-pMO-GFP cell lines, we observed reduction of expression of the model gene in TZMbl cells. In regards to the delivery, the complexes with higher concentration peptide (60 nmol) showed promising results. Nevertheless, these results exhibited high SD and high variability between the data of the repetitive experiments.

The siRNAs showed cell dependent efficacy and intracellular delivery. Moreover, the complexes with peptides showed not efficient and variable cellular internalization probably because of the lower complex stability. In addition, we prepared conjugates of the siRNAs with the Tat1a peptide and its performance in TZMbl cells will be investigated in the near future. We speculate that the conjugates having well defined structure will perform better in *in vivo* studies compared with the nanoplexes.

In the second project, we designed and synthetized ASOs - Gap and Gap A targeting the intron of BGas and siRNAs targeting the nascent RNA from the promoter sequence of BGas. BGas is lnRNA antisense to the *CFTR* protein coding sense strand. We aimed to suppress the BGas binding to *CFTR* and increase its expression. Preliminary, we screened the effectiveness of the ASOs as complexes and conjugates (two distinct chemistries – CuAAC and SPAAK) with Tat1a peptide in A549 cells. Analyzing the data, the performance in increasing the *CFTR* expression varied and was highly concentration dependent. While in general, the complexes needed higher concentration, conjugates, needed approximately 15-25 fold less concentrated probes to see an effect. Nevertheless, we were not able to reproduce this preliminary data and further detailed examination will be necessary to provide stronger conclusions.

The final project, described in this thesis (Chapter 6, Paper IV), includes synthesis of a novel multi-fluorescent oligonucleotides as a tools for detection of single gene mutation. Using cheap and affordable starting materials, we synthetized new bis-alkyne nucleic acids building block that was afterwards successfully integrated into three oligonucleotide sequences. The synthesis of the bis (prop-2-yn-1-yloxy) butane-1,3-diol scaffold was achieved in six synthetic steps giving sufficient amount of the final product to be hand coupled into the oligonucleotide sequences. The yields of each reaction step were decent except for the first step – synthesis of 4.6-O-benzylidene-D-glucopyranose. For its synthesis we followed synthetic strategy by Lindhorst.[177] The reaction achieved the desired product after filtration without further purification. Even though the authors reported 50% yield, we achieved 9% yield of the product. Previous data following the same procedure also reported no more than 10% yield.[178] In addition, a report using tosic acid catalyzed transacetalation of D-glucose stated a 15% yield of product. [179] However, taking into account the simplicity of the Lindhorst strategy, we scaled up this initial step and after getting a reasonable amount of product, we continued carrying out the next steps. We performed the second step as described in literature without changing the method achieving 70% product (3).[179,180] Next, product (4) was achieved following a protocol for making

polyalkyne scaffolds gaining 65% yield.[181] We found the removal of the benzylidene group in step four challenging. We first performed the reaction in 80% acetic acid as described by the literature.[182] However, even though we observed formation of a product by TLC, after wash with saturated solution of sodium bicarbonate, water and brine, followed flash chromatography, we failed to gain the product. We started optimizing the reaction first using the same concentration acetic acid using higher temperature (80 ^{o}C) for shorter time. Figuring out that this was not positive, we also tried using 50% and 100% acetic acid, but without success. We failed to obtain product when using 60% acetic acid in acetonitrile as well. Then we used a strategy using phosphomolybdic acid supported on silica gel for removing acetals.[183] Performing the reaction in ethyl acetate at r.t., we observed only slight conversion of the starting material by TLC. Next, following the Carrigan *et al*. report on deprotection of acetals by triflates, we performed the reaction with copper triflate (Cu(OTf)3) in THF and water at rt.[184] Unfortunately this strategy was not successful failing to convert the starting material. Finally, we performed the reaction with TFA in DCM to achieve product (5) in 43% yield. Next, the primary alcohol was selectively protected with a DMT group,[185] followed by phosphitylation of the secondary alcohol to gain the phosphoramidite building block.[186]

After synthesis of the scaffold, it was successfully incorporated (internal and terminal) into the oligonucleotide sequences by hand coupling. Next, the double and quaternary alkyne functionalized oligonucleotides were conjugated with the three-different azide fluorescent dyes. Satisfactory yields of the ODCs products were achieved after performing the CuAAC reaction assisted by high temperature activation for short time, followed by microwave reactor treatment and r.t. reaction overnight. Afterwards, the ODCs were used in experiments that investigated their ability and specificity for target binding. Next, the fluorescence intensity shift at emission maxima after binding target and mismatched DNA and RNA strands was investigated. Finally, the most promising ODCs were used for detection of cancer related DNA from cell extract in a bead assay.

The data showed that the target binding specificity and the duplex geometry with DNA and RNA target sequences depend from the attached dye. The ODCs with perylene dyes showed efficient intercalation within the duplex and ability to sense target hybridization by shift in the fluorescence intensity at emission maxima. Whereas the ODCs with 5JOE and PEP dyes failed to show clear transition in the T_m melting curves and fluorescence intensity shift trend. Moreover, both the K1-Per and K2-Per conjugates, showed fluorescent intensity alteration at emission

maxima binding the mutant compared with the wild target. Following this data, we used the ODCs with perylene for the capturing (bead) assay, to detect single mutated regions in genes extracted from cancer cell lines.

.

Chapter 8

Concluding Remarks and Perspectives

Advancement of the properties of the synthetic oligonucleotides will lead to efficient probes applicable as therapeutic and as diagnostic tools. The required properties depend on the application, but primary requirements for oligonucleotide probes are descent stability in the media used and selectivity towards target binding.

Synthetic oligonucleotides as therapeutics should be stable with prolonged half-life in blood and in serum of a minimum of 12 h, bind the target specifically without off-target interactions, to be safe for *in vivo* administration and affordable to produce. Moreover, as oligonucleotides are affecting intracellular targets, effective methods for cell internalization are required.

The data presented in chapters 3, 4, and 5 show that short peptides advance the therapeutic oligonucleotide performance in regards to target binding efficiency, stability and intracellular delivery to some extent. In particular, the T_m experiments show that the specificity for target binding depend on the nature of the peptide sequences. In addition, the amino acid content of the peptide affect oligonucleotides stability in bio-fluids, especially in simulated gastric juice. The cell culture experiment data suggest that the peptide concentration used and the cell type may affect the intracellular delivery of the oligonucleotide nanoplexes.

While the strategy using short peptide sequences is advantageous for efficient target binding and protection towards enzyme degradation of the oligonucleotide, it also face with drawbacks. The lack of cell-type selectivity, incompatibility with oral administration and renal filtration are the biggest problems that scientist will need to overcome in future.

The focus will be on development of new systems for efficient delivery on target side, new modification chemistries and mechanistic studies.

Fluorescent oligonucleotide hybridization-based diagnostic tools should show efficient target binding, discrimination between matched and mismatched target, bright fluorescence signal and alteration of the fluorescent emission signal upon mismatched target binding. If designed for *in vivo* application further enzymatic stability and targeted intracellular delivery are required.

The results described in chapter 6, confirm that the new synthetized bis-alkyne nucleic acid building block successfully couple into oligonucleotide sequences and conjugate double and quaternary with different fluorescence dyes. Oligonucleotides conjugated with perylene dyes show the most promising assets in the T_m, CD and fluorescence studies, when hybridized with

mismatched and target DNA and RNA sequences. Consequently, when applied in the capturing (bead) assay, they successfully detect mutated gene in cell extract.
In the future, the bis-alkyne scaffold can be explored for conjugation with excimer probes, for example with two molecules of pyrene dye. The formation of excimer may show a longer emission wavelength and longer fluorescence lifetime.

Higher selectivity and detection of ultra-low abundant single point gene mutations remain scientific and technical challenge. Scientist will focus on development of new probe designs with multiple chemical functionalities, new signal generation strategies and advanced measurement techniques. Moreover, the interest in development of high-throughput automatic platforms such as microfluidic chips will greatly develop the field.

REFERENCES

1. Dahm, R. Friedrich Miescher and the discovery of DNA. *Dev. Biol.* **278,** 274–88 (2005).

2. Weinberg, M. S. & Morris, K. V. Transcriptional gene silencing in humans. *Nucleic Acids Res.* **44,** 6505–6517 (2016).

3. Portin, P. The birth and development of the DNA theory of inheritance: sixty years since the discovery of the structure of DNA. *J Genet*, **93**, 293-302 (2014).

4. Crick, F. H. On protein synthesis. *Symp. Soc. Exp. Biol.* **12,** 138–63 (1958).

5. Riley, M., Pardee, A. B., Jacob, F. & Monad, J. On the Expression of a Structural Gene. *Journal of Molecular Biology* **2,** 423-432 (1960).

6. Wightman, B., Ha, I. & Ruvkun, G. Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans. *Cell* **75,** 855–862 (1993).

7. Britten, R. J. & Davidson, E. H. Gene regulation for higher cells: a theory. *Science* **165,** 349–357 (1969).

8. Berget, S. M., Moore, C. & Sharp, P. A. Spliced segments at the 5’ terminus of adenovirus 2 late mRNA. *Proc. Natl. Acad. Sci. U. S. A.* **74,** 3171–3175 (1977).

9. Fire, A. *et al.* Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature* **391,** 806–811 (1998).

10. Morris, K. V., Chan, S. W.-L., Jacobsen, S. E. & Looney, D. J. Small Interfering RNA-Induced Transcriptional Gene Silencing in Human Cells. *Science* **305,** 1289–1292 (2004).

11. Mirsky, A. E. The discovery of DNA. *Sci. Am.* **218,** 78–88 (1968).

12. Avery, O. T., Macleod, C. M. & Mccarty, M. Studies on the chemical nature of the substance inducing transformation of pneumococcal types induction of transformation by a desoxyribonucleic acid fraction isolated from Pneumococcus type III. **4**, 344-365 (1995).

13. Chargaff, E. Structure and function of nucleic acids as cell constituents. *Fed. Proc.* **10,** 654–659 (1951).

14. Chargaff, E. Chemical specificity of nucleic acids and mechanism of their enzymatic

degradation. *Experientia* **6,** 201–209 (1950).

15. Franklin, R. E. & Gosling, R. G. Molecular structure of nucleic acids. Molecular configuration in sodium thymonucleate. 1953. *Ann. N. Y. Acad. Sci.* **758,** 16–17 (1995).

16. Watson, J. D. & Crick, F. H. Genetical implications of the structure of deoxyribonucleic acid. *Nature* **171,** 964–967 (1953).

17. Watson, J. D. & Crick, F. H. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature* **171,** 737–738 (1953).

18. Crick, F. Central dogma of molecular biology. *Nature* **227,** 561–563 (1970).

19. Crick, F. H., Barnett, L., Brenner, S. & Watts-Tobin, R. J. General nature of the genetic code for proteins. *Nature* **192,** 1227–1232 (1961).

20. Portin, P. The concept of the gene: short history and present status. *Q. Rev. Biol.* **68,** 173–223 (1993).

21. Sanger, F., Nicklen, S. & Coulson, A. R. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. U. S. A.* **74,** 5463–5467 (1977).

22. Venter, J. C. *et al.* The sequence of the human genome. *Science* **291,** 1304–1351 (2001).

23. Maxam, A. M. & Gilbert, W. A method for determining DNA sequence by labeling the end of the molecule and cleaving at the base. Isolation of DNA fragments, end-labeling, cleavage, electrophoresis in polyacrylamide gel and analysis of results. *Mol. Biol. (Mosk).* **20,** 581–638 (1986)

24. Arnott, S. Historical article: DNA polymorphism and the early history of the double helix. *Trends Biochem Sci*. **6**, 349-354 (2006)

25. Wang, G. & Vasquez, K. M. Z-DNA, an active element in the genome. *Front. Biosci.* **12,** 4424–4438 (2007).

26. Gyi, J. I., Lane, A. N., Conn, G. L. & Brown, T. Solution Structures of DNA,RNA Hybrids with Purine-Rich and Pyrimidine-Rich Strands: Comparison with the Homologous DNA and RNA. *Biochemistry* **1**, 73-80 (1998).

27. Culkin, J., de Bruin, L., Tompitak, M., Phillips, R. & Schiessel, H. The role of DNA sequence in nucleosome breathing. *Eur. Phys. J.* **40,** 106-118 (2017).

28. Crider, K. S., Yang, T. P., Berry, R. J. & Bailey, L. B. Folate and DNA Methylation: A Review of Molecular Mechanisms and the Evidence for Folate's Role. *Adv. Nutr.* **3,** 21–38 (2012).

29. Michelson, A. M. & Todd, A. R. Nucleotides part XXXII. Synthesis of a dithymidine dinucleotide containing a 3′: 5′-internucleotidic linkage. *J. Chem. Soc.* **0,** 2632–2638 (1955).

30. Gilham, P. T. & Khorana, H. G. Studies on Polynucleotides. I. A New and General Method for the Chemical Synthesis of the C_{5}''-C_{3}'' Internucleotidic Linkage. Syntheses of Deoxyribodinucleotides. *J. Am. Chem. Soc.* **80,** 6212–6222 (1958).

31. Schaller, H., Weimann, G., Lerch, B. & Khorana, H. G. Syntheses of Deoxyribonucleoside-3' Phosphates 3821 Experimental35 Studies on Polynucleotides. XXIV.1 The Stepwise Synthesis of Specific Deoxyribopolynucleotides (4).2 Protected Derivatives of Deoxyribonucleosides and New Syntheses of Deoxyribonucleoside-3' Phosphates3. *J. Am. Chem. Soc.* **5, 3**821-3827 (1963).

32. Letsinger, R. L. & Ogilvie, K. K. Nucleotide chemistry. XIII. Synthesis of oligothymidylates via phosphotriester intermediates. *J. Am. Chem. Soc.* **91,** 3350–3355 (1969).

33. Letsinger, R. L., Finnan, J. L., Heavner, G. A. & Lunsford, W. B. Nucleotide chemistry. XX. Phosphite coupling procedure for generating internucleotide links. *J. Am. Chem. Soc.* **97,** 3278–3279 (1975).

34. Letsinger, R. L. & Mahadevan, V. Oligonucleotide Synthesis on a Polymer Support [1,2]. *J. Am. Chem. Soc.* **87,** 3526–3527 (1965).

35. Merrifield R B, Letsinger, L. L. & Hamilton, S. B. *Reactions on Polymer Supports. E. H. Immergut, Makromol. Chem* **24,** 3045-3046 (1962).

36. Matteucci, M. D. & Caruthers, M. H. Synthesis of deoxyoligonucleotides on a polymer support. *J. Am. Chem. Soc.* **103,** 3185–3191 (1981).

37. Beaucage, S. L. & Caruthers, M. H. Deoxynucleoside phosphoramidites—A new class of key intermediates for deoxypolynucleotide synthesis. *Tetrahedron Lett.* **22,** 1859–1862 (1981).

38. Sinha, N. D., Biernat, J. & Kuster, H.β-Cyanoethyl N,N dialkylamino/N morpholinomonochloro phosphoamidites new phosphitylating agents fascilitating ease of

deprotection and work-up of synthetised oligonucleotides. *Tetrahedron Letters* **24,** 5843-5846 (1983).

39. McCollum, C. & Andrus, A. An optimized polystyrene support for rapid, efficient oligonucleotide synthesis. *Tetrahedron Lett.* **32,** 4069–4072 (1991).

40. Ellington, A. & Pollard, J. D. Synthesis and purification of oligonucleotides. *Curr. Protoc. Mol. Biol.* **Chapter 2,** Unit2.11 (2001).

41. Berner, S., Mühlegger, K. & Seliger, H. Studies on the role of tetrazole in the activation of phosphoramidites. *Nucleic Acids Res.* **17,** 853–864 (1989).

42. Shibata, T. *et al.* Protein-driven RNA nanostructured devices that function in vitro and control mammalian cell fate. *Nat. Commun.* **8,** 540-553 (2017).

43. Chen, Y.-J. *et al.* Programmable chemical controllers made from DNA. *Nat. Nanotechnol.* **8,** 755–762 (2013).

44. Pan, J., Li, F., Cha, T.-G., Chen, H. & Choi, J. H. Recent progress on DNA based walkers. *Curr. Opin. Biotechnol.* **34,** 56–64 (2015).

45. Yurke, B., Turberfield, A. J., Mills, A. P., Simmel, F. C. & Neumann, J. L. A DNA-fuelled molecular machine made of DNA. *Nature* **406,** 605–608 (2000).

46. He, Y. & Liu, D. R. Autonomous multistep organic synthesis in a single isothermal solution mediated by a DNA walker. *Nat. Nanotechnol.* **5,** 778–782 (2010).

47. Cha, T.-G. *et al.* A synthetic DNA motor that transports nanoparticles along carbon nanotubes. *Nat. Nanotechnol.* **9,** 39–43 (2014).

48. Jolly, P., Estrela, P. & Ladomery, M. Oligonucleotide-based systems: DNA, microRNAs, DNA/RNA aptamers. *Essays Biochem.* **60,** 27–35 (2016).

49. Prado, M. *et al.* Advanced DNA- and Protein-based Methods for the Detection and Investigation of Food Allergens. *Crit. Rev. Food Sci. Nutr.* **56,** 2511–2542 (2016).

50. Li, D., Song, S. & Fan, C. Target-responsive structural switching for nucleic acid-based sensors. *Acc. Chem. Res.* **43,** 631–41 (2010).

51. Hashemian, Z., Khayamian, T., Saraji, M. & Shirani, M. P. Aptasensor based on fluorescence resonance energy transfer for the analysis of adenosine in urine samples of lung cancer

patients. *Biosens. Bioelectron.* **79,** 334–340 (2016).

52. Gong, L. *et al.* DNAzyme-based biosensors and nanodevices. *Chem. Commun.* **51,** 979–995 (2015).

53. Lin, M. *et al.* Electrochemical detection of nucleic acids, proteins, small molecules and cells using a DNA-nanostructure-based universal biosensing platform. *Nat. Protoc.* **11,** 1244–1263 (2016).

54. Chen, Z. *et al.* Stimulus-response mesoporous silica nanoparticle-based chemiluminescence biosensor for cocaine determination. *Biosens. Bioelectron.* **75,** 8–14 (2016).

55. Wu, C., Fan, D., Zhou, C., Liu, Y. & Wang, E. Colorimetric Strategy for Highly Sensitive and Selective Simultaneous Detection of Histidine and Cysteine Based on G-Quadruplex-Cu(II) Metalloenzyme. *Anal. Chem.* **88,** 2899–2903 (2016).

56. Saiki, R. K. *et al.* Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* **239,** 487–491 (1988).

57. Lee, S. B., McCord, B. & Buel, E. Advances in forensic DNA quantification: A review. *Electrophoresis* **35,** 3044–3052 (2014).

58. Parson, W. Age Estimation with DNA: From Forensic DNA Fingerprinting to Forensic (Epi)Genomics: A Mini-Review. *Gerontology* **64,** 326–332 (2018).

59. Haubold, B. Alignment-free phylogenetics and population genetics. *Brief. Bioinform.* **15,** 407–418 (2014).

60. Klemm, E. & Dougan, G. Advances in Understanding Bacterial Pathogenesis Gained from Whole-Genome Sequencing and Phylogenetics. *Cell Host Microbe* **19,** 599–610 (2016).

61. Buckley, T. R. Applications of phylogenetics to solve practical problems in insect conservation. *Curr. Opin. Insect Sci.* **18,** 35–39 (2016).

62. Xue, J., Yang, J., Luo, M., Cho, W. C. & Liu, X. MicroRNA-targeted therapeutics for lung cancer treatment. *Expert Opin. Drug Discov.* **12,** 141–157 (2017).

63. Wang, T. *et al.* Challenges and opportunities for siRNA-based cancer treatment. *Cancer Lett.* **387,** 77–83 (2017).

64. Young, S. W. S., Stenzel, M. & Jia-Lin, Y. Nanoparticle-siRNA: A potential cancer therapy? *Crit. Rev. Oncol. Hematol.* **98,** 159–169 (2016).

65. Khorkova, O. & Wahlestedt, C. Oligonucleotide therapies for disorders of the nervous system. *Nat. Biotechnol.* **35,** 249–263 (2017).

66. Choong, C.-J., Baba, K. & Mochizuki, H. Gene therapy for neurological disorders. *Expert Opin. Biol. Ther.* **16,** 143–159 (2016).

67. Sperry, B. W. & Tang, W. H. W. Amyloid heart disease: genetics translated into disease-modifying therapy. *Heart* **103,** 812–817 (2017).

68. Sayed, D., Yang, Z., He, M., Pfleger, J. M. & Abdellatif, M. Acute Targeting of General Transcription Factor IIB Restricts Cardiac Hypertrophy via Selective Inhibition of Gene Transcription. *Circ. Hear. Fail.* **8,** 138–148 (2015).

69. Mourich, D. V. *et al.* Alternative Splice Forms of CTLA-4 Induced by Antisense Mediated Splice-Switching Influences Autoimmune Diabetes Susceptibility in NOD Mice. *Nucleic Acid Ther.* **24,** 114–126 (2014).

70. Li, W. & Lan, X. Aptamer Oligonucleotides: Novel Potential Therapeutic Agents in Autoimmune Disease. *Nucleic Acid Ther.* **25,** 173–179 (2015).

71. Fraietta, J. A. *et al.* Abasic Phosphorothioate Oligomers Inhibit HIV-1 Reverse Transcription and Block Virus Transmission across Polarized Ectocervical Organ Cultures. *Antimicrob. Agents Chemother.* **58,** 7056–7071 (2014).

72. Asha, K. *et al.* Advancements in Nucleic Acid Based Therapeutics against Respiratory Viral Infections. *J. Clin. Med.* **8,** 6-30 (2018).

73. Sully, E. K. & Geller, B. L. Antisense antimicrobial therapeutics. *Curr. Opin. Microbiol.* **33,** 47–55 (2016).

74. Bennett, C. F., Baker, B. F., Pham, N., Swayze, E. & Geary, R. S. Pharmacology of Antisense Drugs. *Annu. Rev. Pharmacol. Toxicol.* **57,** 81–105 (2017).

75. Friedmann, T. & Roblin, R. Gene therapy for human genetic disease? *Science* **175,** 949–955 (1972).

76. Lipinski, C. A., Lombardo, F., Dominy, B. W. & Feeney, P. J. Experimental and computational

approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.* **46,** 3–26 (2001).

77. Crooke, S. T. Molecular Mechanisms of Antisense Oligonucleotides. *Nucleic Acid Ther.* **27,** 70–77 (2017).

78. Shen, X. & Corey, D. R. Chemistry, mechanism and clinical status of antisense oligonucleotides and duplex RNAs. *Nucleic Acids Res.* **46,** 1584–1600 (2018).

79. Geary, R. S., Henry, S. P. & Grillone, L. R. Fomivirsen. *Clin. Pharmacokinet.* **41,** 255–260 (2002).

80. Lim, K. R. Q. & Yokota, T. Invention and Early History of Exon Skipping and Splice Modulation. *Methods. Mol. Biol*. **1828**, 3-30 (2018)

81. Stein, C. A. & Castanotto, D. FDA-Approved Oligonucleotide Therapies in 2017. *Mol. Ther.* **25,** 1069–1075 (2017).

82. Paterson, B. M., Roberts, B. E. & Kuff, E. L. Structural gene identification and mapping by DNA-mRNA hybrid-arrested cell-free translation. *Proc. Natl. Acad. Sci. U. S. A.* **74,** 4370–4374 (1977).

83. Zamecnik, P. C. & Stephenson, M. L. Inhibition of Rous sarcoma virus replication and cell transformation by a specific oligodeoxynucleotide. *Proc. Natl. Acad. Sci. U. S. A.* **75,** 280–284 (1978).

84. Goyal, N. & Narayanaswami, P. Making sense of antisense oligonucleotides: A narrative review. *Muscle Nerve* **57,** 356–370 (2018).

85. Faria, M. *et al.* Phosphoramidate oligonucleotides as potent antisense molecules in cells and in vivo. *Nat. Biotechnol.* **19,** 40–44 (2001).

86. Baker, B. F. *et al.* 2'-O-(2-Methoxy)ethyl-modified anti-intercellular adhesion molecule 1 (ICAM-1) oligonucleotides selectively increase the ICAM-1 mRNA level and inhibit formation of the ICAM-1 translation initiation complex in human umbilical vein endothelial cells. *J. Biol. Chem.* **272,** 11994–12000 (1997).

87. Hannon, G. J. RNA interference. *Nature* **11***,* 244-251 (2002).

88. Elbashir, S. M. *et al.* Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured

mammalian cells. *Nature* **411,** 494–498 (2001).

89. Deltcheva, E. *et al.* CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* **471,** 602–607 (2011).

90. Doudna, J. A. & Charpentier, E. The new frontier of genome engineering with CRISPR-Cas9. *Science* **346,** 1258096–1258096 (2014).

91. Haft, D. H., Selengut, J., Mongodin, E. F. & Nelson, K. E. A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/Cas subtypes exist in prokaryotic genomes. *PLoS Comput. Biol.* **1,** 474-484 (2005).

92. Makarova, K. S., Grishin, N. V, Shabalina, S. A., Wolf, Y. I. & Koonin, E. V. A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. *Biol. Direct* **1,** 1-26 (2006).

93. Jiang, F. *et al.* Structures of a CRISPR-Cas9 R-loop complex primed for DNA cleavage. *Science* **351,** 867–871 (2016).

94. Yin, H. *et al.* Structure-guided chemical modification of guide RNA enables potent non-viral in vivo genome editing. *Nat. Biotechnol.* **35,** 1179–1187 (2017).

95. Ran, F. A. Adaptation of CRISPR nucleases for eukaryotic applications. *Anal. Biochem.* **532,** 90–94 (2017).

96. Schmidt, K. *et al.* Characterizing the effect of GalNAc and phosphorothioate backbone on binding of antisense oligonucleotides to the asialoglycoprotein receptor. *Nucleic Acids Res.* **45,** 2294–2306 (2017).

97. Wang, S., Sun, H., Tanowitz, M., Liang, X. & Crooke, S. T. Intra-endosomal trafficking mediated by lysobisphosphatidic acid contributes to intracellular release of phosphorothioate-modified antisense oligonucleotides. *Nucleic Acids Res.* **45,** 5309–5322 (2017).

98. Iwamoto, N. *et al.* Control of phosphorothioate stereochemistry substantially increases the efficacy of antisense oligonucleotides. *Nat. Biotechnol.* **35,** 845–851 (2017).

99. Wan, W. B. *et al.* Synthesis, biophysical properties and biological activity of second generation antisense oligonucleotides containing chiral phosphorothioate linkages.

Nucleic Acids Res. **42,** 13456–13468 (2014).

100. Li, M. *et al.* Synthesis and cellular activity of stereochemically-pure 2'-O-(2-methoxyethyl)-phosphorothioate oligonucleotides. *Chem. Commun. (Camb).* **53,** 541–544 (2017).

101. Summerton, J. E. Invention and Early History of Morpholinos: From Pipe Dream to Practical Products. in *Methods in molecular biology (Clifton, N.J.)* **1565,** 1–15 (2017).

102. Blum, M., De Robertis, E. M., Wallingford, J. B. & Niehrs, C. Morpholinos: Antisense and Sensibility. *Dev. Cell* **35,** 145–149 (2015).

103. Akpulat, U. *et al.* Shorter Phosphorodiamidate Morpholino Splice-Switching Oligonucleotides May Increase Exon-Skipping Efficacy in DMD. *Mol. Ther. - Nucleic Acids* **13,** 534–542 (2018).

104. Montazersaheb, S., Hejazi, M. S. & Nozad Charoudeh, H. Potential of Peptide Nucleic Acids in Future Therapeutic Applications. *Adv. Pharm. Bull.* **8,** 551–563 (2018).

105. Narenji, H. *et al.* Peptide nucleic acids (PNAs): currently potential bactericidal agents. *Biomed. Pharmacother.* **93,** 580–588 (2017).

106. Ricciardi, A., Quijano, E., Putman, R., Saltzman, W. & Glazer, P. Peptide Nucleic Acids as a Tool for Site-Specific Gene Editing. *Molecules* **23,** 632-647 (2018).

107. Nishizaki, T., Iwai, S., Ohtsuka, E. & Nakamura, H. Solution Structure of an RNA·2'- *O* -Methylated RNA Hybrid Duplex Containing an RNA·DNA Hybrid Segment at the Center. *Biochemistry* **36,** 2577–2585 (1997).

108. Deleavey, G. F., Watts, J. K. & Damha, M. J. Chemical modification of siRNA. *Curr. Protoc. nucleic acid Chem.* **Chapter 16,** Unit 16.3 (2009).

109. Kratschmer, C. & Levy, M. Effect of Chemical Modifications on Aptamer Stability in Serum. *Nucleic Acid Ther.* **6**, 335-344 (2017).

110. Pallan, P. S. *et al.* Unexpected origins of the enhanced pairing affinity of 2'-fluoro-modified RNA. *Nucleic Acids Res.* **39,** 3482–3495 (2011).

111. Bennett, C. F. & Swayze, E. E. RNA targeting therapeutics: molecular mechanisms of antisense oligonucleotides as a therapeutic platform. *Annu. Rev. Pharmacol. Toxicol.* **50,** 259–293 (2010).

112. Birte Vester & Wengel, J. LNA (Locked Nucleic Acid): High-Affinity Targeting of Complementary RNA and DNA. **26**, 13233-13241 (2004).

113. Petersen, M. & Wengel, J. LNA: a versatile tool for therapeutics and genomics. *Trends Biotechnol.* **21,** 74–81 (2003).

114. Lundin, K. E. *et al.* Biological Activity and Biotechnological Aspects of Locked Nucleic Acids. in *Advances in genetics* **82,** 47–107 (2013).

115. Campbell, M. A. & Wengel, J. Locked vs. unlocked nucleic acids (LNA vs. UNA): contrasting structures work towards common therapeutic goals. *Chem. Soc. Rev.* **40,** 5680-5689 (2011).

116. Snead, N. M., Escamilla-Powers, J. R., Rossi, J. J. & McCaffrey, A. P. 5′ Unlocked Nucleic Acid Modification Improves siRNA Targeting. *Mol. Ther. - Nucleic Acids* **2,** 103-110 (2013).

117. Pasternak, A. & Wengel, J. Unlocked nucleic acid – an RNA modification with broad potential. *Org. Biomol. Chem.* **9,** 3591-3597 (2011).

118. Springer, A. D. & Dowdy, S. F. GalNAc-siRNA Conjugates: Leading the Way for Delivery of RNAi Therapeutics. *Nucleic Acid Ther.* **28,** 109–118 (2018).

119. Willoughby, J. L. S. *et al.* Evaluation of GalNAc-siRNA Conjugate Activity in Pre-clinical Animal Models with Reduced Asialoglycoprotein Receptor Expression. *Mol. Ther.* **26,** 105–114 (2018).

120. Haussecker, D. Current issues of RNAi therapeutics delivery and development. *J. Control. Release* **195,** 49–54 (2014).

121. Foster, D. J. *et al.* Advanced siRNA Designs Further Improve In Vivo Performance of GalNAc-siRNA Conjugates. *Mol. Ther.* **26,** 708–717 (2018).

122. Krützfeldt, J. *et al.* Silencing of microRNAs in vivo with 'antagomirs'. *Nature* **438,** 685–689 (2005).

123. Karaki, S. *et al.* Lipid-oligonucleotide conjugates improve cellular uptake and efficiency of TCTP-antisense in castration-resistant prostate cancer. *J. Control. Release* **258,** 1–9 (2017).

124. Lorenz, C., Hadwiger, P., John, M., Vornlocher, H.-P. & Unverzagt, C. Steroid and lipid

conjugates of siRNAs to enhance cellular uptake and gene silencing in liver cells. *Bioorg. Med. Chem. Lett.* **14,** 4975–4977 (2004).

125. Ugarte-Uribe, B. *et al.* Lipid-modified oligonucleotide conjugates: Insights into gene silencing, interaction with model membranes and cellular uptake mechanisms. *Bioorg. Med. Chem.* **25,** 175–186 (2017).

126. Raouane, M., Desmaële, D., Urbinati, G., Massaad-Massade, L. & Couvreur, P. Lipid Conjugated Oligonucleotides: A Useful Strategy for Delivery. *Bioconjug. Chem.* **23,** 1091–1104 (2012).

127. Soutschek, J. *et al.* Therapeutic silencing of an endogenous gene by systemic administration of modified siRNAs. *Nature* **432,** 173–178 (2004).

128. Nishina, K. *et al.* DNA/RNA heteroduplex oligonucleotide for highly efficient gene silencing. *Nat. Commun.* **6,** 7969-7982 (2015).

129. Chandola, C., Kalme, S., Casteleijn, M. G., Urtti, A. & Neerathilingam, M. Application of aptamers in diagnostics, drug-delivery and imaging. *J. Biosci.* **41,** 535–561 (2016).

130. Röthlisberger, P. & Hollenstein, M. Aptamer chemistry. *Adv. Drug Deliv. Rev.* **134,** 3–21 (2018).

131. Resende, R. R., Torres, H. A. M., Yuahasi, K. K., P, M. & H, U. Delivery systems for in vivo use of nucleic Acid drugs. *Drug Target Insights* **2,** 183–196 (2007).

132. Kotula, J. W. *et al.* Aptamer-Mediated Delivery of Splice-Switching Oligonucleotides to the Nuclei of Cancer Cells. *Nucleic Acid Ther.* **22,** 187–195 (2012).

133. Sun, H. *et al.* Oligonucleotide aptamers: new tools for targeted cancer therapy. *Mol. Ther. Nucleic Acids* **3,** 182-196 (2014).

134. Thiel, K. W. & Giangrande, P. H. Intracellular delivery of RNA-based therapeutics using aptamers. *Ther. Deliv.* **1,** 849–861 (2010).

135. Morrison, C. Alnylam prepares to land first RNAi drug approval. *Nat. Rev. Drug Discov.* **17,** 156–157 (2018).

136. Home - ClinicalTrials.gov. Available at: https://clinicaltrials.gov/. (Accessed: 10th April 2019)

137. Stein, C. A. & Castanotto, D. FDA-Approved Oligonucleotide Therapies in 2017. *Mol. Ther.* **25,** 1069–1075 (2017).

138. Chakraborty, C., Sharma, A. R., Sharma, G., Doss, C. G. P. & Lee, S.-S. Therapeutic miRNA and siRNA: Moving from Bench to Clinic as Next Generation Medicine. *Mol. Ther. Nucleic Acids* **8,** 132–143 (2017).

139. Baylis, F. & McLeod, M. First-in-human Phase 1 CRISPR Gene Editing Cancer Trials: Are We Ready? *Curr. Gene Ther.* **17,** 309–319 (2017).

140. Vadaie, N. & Morris, K. V. Long antisense non-coding RNAs and the epigenetic regulation of gene expression. *Biomol. Concepts* **4,** 411–415 (2013).

141. Mercer, T. R. & Mattick, J. S. Structure and function of long noncoding RNAs in epigenetic regulation. *Nat. Struct. Mol. Biol.* **20,** 300–307 (2013).

142. Saayman, S. *et al.* An HIV-encoded antisense long noncoding RNA epigenetically regulates viral transcription. *Mol. Ther.* **22,** 1164–1175 (2014).

143. Yarani, R., Mirza, A. H., Kaur, S. & Pociot, F. The emerging role of lncRNAs in inflammatory bowel disease. *Exp. Mol. Med.* **50,** 161-175 (2018).

144. Tang, Y., Zhou, T., Yu, X., Xue, Z. & Shen, N. The role of long non-coding RNAs in rheumatic diseases. *Nat. Rev. Rheumatol.* **13,** 657–669 (2017).

145. Gong, P. *et al.* LncRNA UCA1 promotes tumor metastasis by inducing miR-203/ZEB2 axis in gastric cancer. *Cell Death Dis.* **9,** 1158-1172 (2018).

146. Sampey, G. C. *et al.* Transcriptional Gene Silencing (TGS) via the RNAi Machinery in HIV-1 Infections. *Biology (Basel).* **1,** 339–369 (2012).

147. Burnett, J. C. *et al.* Combinatorial Latency Reactivation for HIV-1 Subtypes and Variants † Downloaded from. *J. Virol.* **84,** 5958–5974 (2010).

148. Datta, P. K. *et al.* HIV-1 Latency and Eradication: Past, Present and Future. *Cur. HIV Res*. **5**, 431-441 (2016)

149. Saayman, S. M. *et al.* Potent and Targeted Activation of Latent HIV-1 Using the CRISPR/dCas9 Activator Complex. *Mol. Ther.* **24,** 488–498 (2016).

150. Turner, A.-M. W., Ackley, A. M., Matrone, M. A. & Morris, K. V. Characterization of an HIV-

targeted transcriptional gene-silencing RNA in primary cells. *Hum. Gene Ther.* **23,** 473–483 (2012).

151. Zhou, J. *et al.* Receptor-targeted aptamer-siRNA conjugate-directed transcriptional regulation of HIV-1. *Theranostics* **8,** 1575–1590 (2018).

152. Viscidi, R. P., Mayur, K., Lederman, H. M. & Frankel, A. D. Inhibition of antigen-induced lymphocyte proliferation by Tat protein from HIV-1. *Science* **246,** 1606–1608 (1989).

153. Maeda, H. Tumor-selective delivery of macromolecular drugs via the EPR effect: background and future prospects. *Bioconjug. Chem.* **21,** 797–802 (2010).

154. Andaloussi, S. E. L. *et al.* Design of a peptide-based vector, PepFect6, for efficient delivery of siRNA in cell culture and systemically in vivo. *Nucleic Acids Res.* **39,** 3972–3987 (2011).

155. Derakhshankhah, H. & Jafari, S. Cell penetrating peptides: A concise review with emphasis on biomedical applications. *Biomed. Pharmacother.* **108,** 1090–1096 (2018).

156. Böhmová, E. *et al.* Cell-penetrating peptides: a useful tool for the delivery of various cargoes into cells. *Physiol. Res.* **67,** 267–279 (2018).

157. Zorko, M. & Langel, U. Cell-penetrating peptides: mechanism and kinetics of cargo delivery. *Adv. Drug Deliv. Rev.* **57,** 529–545 (2005).

158. Benizri, S. *et al.* Bioconjugated Oligonucleotides: Recent Developments and Therapeutic Applications. *Bioconjug. Chem.* **30,** 366–383 (2019).

159. Park, J. W., Bang, E.-K., Jeon, E. M. & Kim, B. H. Complexation and conjugation approaches to evaluate siRNA delivery using cationic, hydrophobic and amphiphilic peptides. *Org. Biomol. Chem.* **10,** 96–102 (2012).

160. Strapps, W. R. *et al.* The siRNA sequence and guide strand overhangs are determinants of in vivo duration of silencing. *Nucleic Acids Res.* **38,** 4788–4797 (2010).

161. Trabulo, S., Cardoso, A. L., Mano, M. & De Lima, M. C. P. Cell-Penetrating Peptides-Mechanisms of Cellular Uptake and Generation of Delivery Systems. *Pharmaceuticals (Basel).* **3,** 961–993 (2010).

162. Madani, F., Lindberg, S., Langel, U., Futaki, S. & Gräslund, A. Mechanisms of cellular uptake of cell-penetrating peptides. *J. Biophys.* **2011,** 414729-414739 (2011).

163. Charnay, N. *et al.* Mechanism of HIV-1 Tat RNA translation and its activation by the Tat protein. *Retrovirology* **6,** 74-92 (2009).

164. Brasier, A. R., Tate, J. E. & Habener, J. F. Optimized use of the firefly luciferase assay as a reporter gene in mammalian cell lines. *Biotechniques* **7,** 1116–1122 (1989).

165. Corvol, H. *et al.* Genome-wide association meta-analysis identifies five modifier loci of lung disease severity in cystic fibrosis. *Nature Comm.* **13,** 23-31 (2015).

166. Rafeeq, M. M. & Murad, H. A. S. Cystic fibrosis: current therapeutic targets and future approaches. *J. Transl. Med.* **15,** 84-93 (2017).

167. Elborn, J. S. Cystic fibrosis. *Lancet* **388,** 2519–2531 (2016).

168. Saayman, S. M. *et al.* Long Non-coding RNA BGas Regulates the Cystic Fibrosis Transmembrane Conductance Regulator. *Mol. Ther.* **8**, 1351-1358 (2016).

169. Lou, C. *et al.* Peptide–oligonucleotide conjugates as nanoscale building blocks for assembly of an artificial three-helix protein mimic. *Nat. Commun.* **7,** 12294-12303 (2016).

170. Ye, J. *et al.* High-Yield Synthesis of Monomeric LMWP(CPP)-siRNA Covalent Conjugate for Effective Cytosolic Delivery of siRNA. *Theranostics* **7,** 2495–2508 (2017).

171. Forbes, D. C. & Peppas, N. A. Oral delivery of small RNA and DNA. *J. Control. Release* **162,** 438–445 (2012).

172. Duan, B., Li, M., Sun, Y., Zou, S. & Xu, X. Orally Delivered Antisense Oligodeoxyribonucleotides of TNF-α via Polysaccharide-Based Nanocomposites Targeting Intestinal Inflammation. *Adv. Healthc. Mater.* **5**, 1801389-12 (2019).

173. Hyun, E.-J., Hasan, M. N., Kang, S. H., Cho, S. & Lee, Y.-K. Oral siRNA delivery using dual transporting systems to efficiently treat colorectal liver metastasis. *Int. J. Pharm.* **555,** 250–258 (2019).

174. Ball, R. L., Bajaj, P. & Whitehead, K. A. Oral delivery of siRNA lipid nanoparticles: Fate in the GI tract. *Sci. Rep.* **8,** 2178-2190 (2018).

175. Liu, Y. *et al.* Accelerated digestion of nucleic acids by pepsin from the stomach of chicken. *Br. Poult. Sci.* **57,** 1–8 (2016).

176. Liu, Y. *et al.* Digestion of Nucleic Acids Starts in the Stomach. *Sci. Rep.* **5,** 11936-11947

(2015).

177. Lindhorst, T. K. *Essential of carbohydrate chemistry and biochemistry*. (Wiley-VCH, 2007).

178. Sowden, J. C. & Kuenne, D. J. 1,2-Benzylidene-D-glucofuranose [1]. *J. Am. Chem. Soc.* **74,** 686–688 (1952).

179. Tanabe, G. *et al.* Biological evaluation of de-O-sulfonated analogs of salacinol, the role of sulfate anion in the side chain on the α-glucosidase inhibitory activity. *Bioorg. Med. Chem.* **15,** 3926–3937 (2007).

180. Ahmad Ghavami, Blair D. Johnston & Pinto, B. M. A New Class of Glycosidase Inhibitor: Synthesis of Salacinol and Its Stereoisomers. *The Journal of Org.Chem*. **7**, 2312-2317 (2001).

181. Perez-Balderas, F., Morales-Sanfrutos, J., Hernandez-Mateo, F., Isac-García, J. & Santoyo-Gonzalez, F. Click Multivalent Homogeneous Neoglycoconjugates - Synthesis and Evaluation of Their Binding Affinities. *European J. Org. Chem.* **2009,** 2441–2453 (2009).

182. Tomabechi, Y., Odate, Y., Izumi, R., Haneda, K. & Inazu, T. Acceptor specificity in the transglycosylation reaction using Endo-M. *Carbohydr. Res.* **345,** 2458–2463 (2010).

183. Kumar, P. S., Kumar, G. D. K. & Baskaran, S. Truly Catalytic and Chemoselective Cleavage of Benzylidene Acetal with Phosphomolybdic Acid Supported on Silica Gel. *European J. Org. Chem.* **2008,** 6063–6067 (2008).

184. Marc D. Carrigan, Dusan Sarapa, Russell C. Smith, Laura C. Wieland & Mohan, R. S. A Simple and Efficient Chemoselective Method for the Catalytic Deprotection of Acetals and Ketals Using Bismuth Triflate. *The Journal of Org. Chem*, **3**, 1027-1030 (2002).

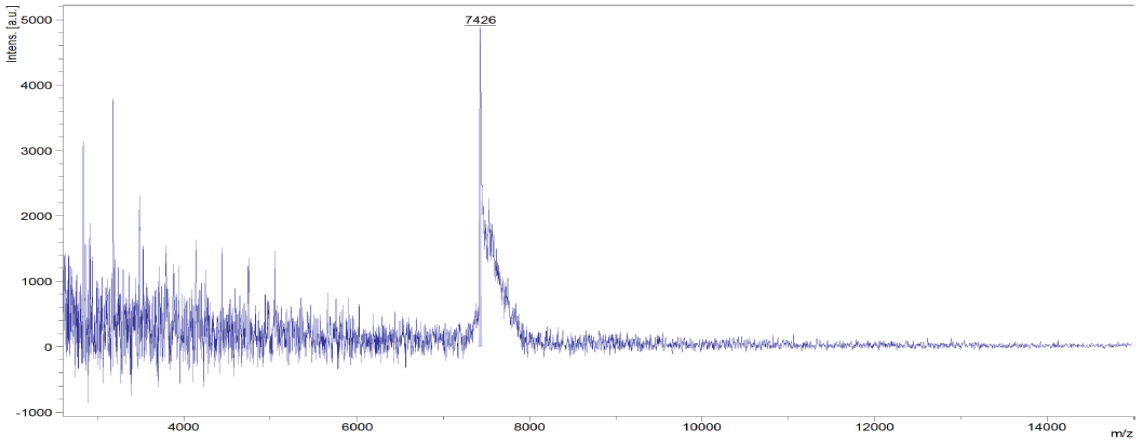
185. Catry, M. A. & Madder, A. Synthesis of functionalised nucleosides for incorporation into nucleic acid-based serine protease mimics. *Molecules* **12,** 114–129 (2007).

186. Kumar, P. *et al.* Three new double-headed nucleotides with additional nucleobases connected to C-5 of pyrimidines; synthesis, duplex and triplex studies. *Bioorg. Med. Chem.* **24,** 742–749 (2016).

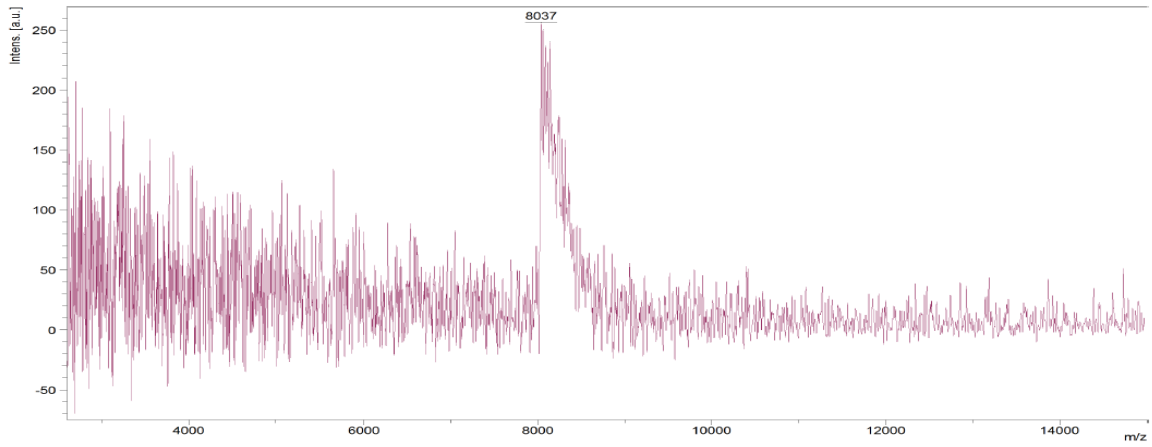
Appendix A

Figure A1. Maldi-TOF spectra of the siRNA sense strands-Tat1a conjugates

S2-Tat1a



S4-Tat1a



SCS2-Tat1a

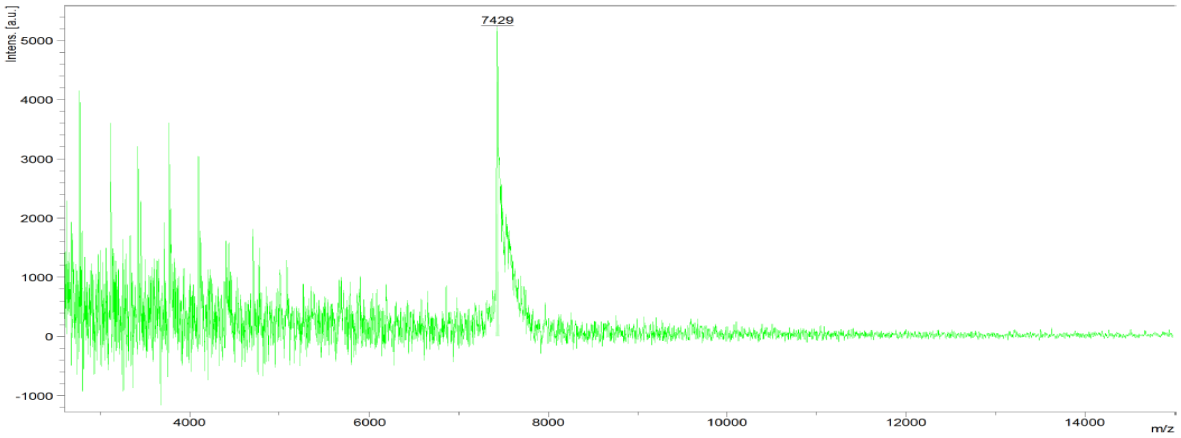
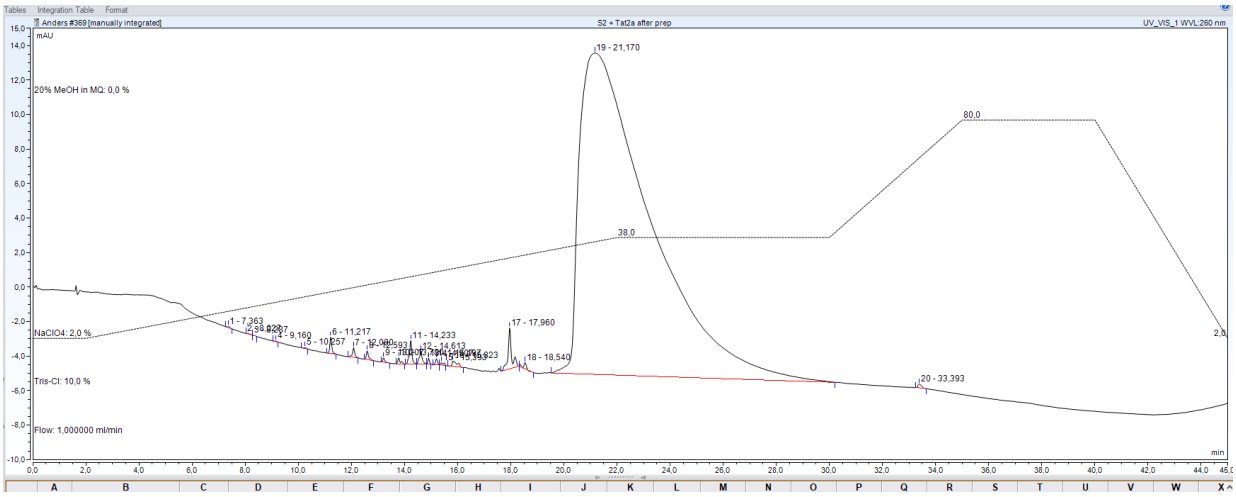
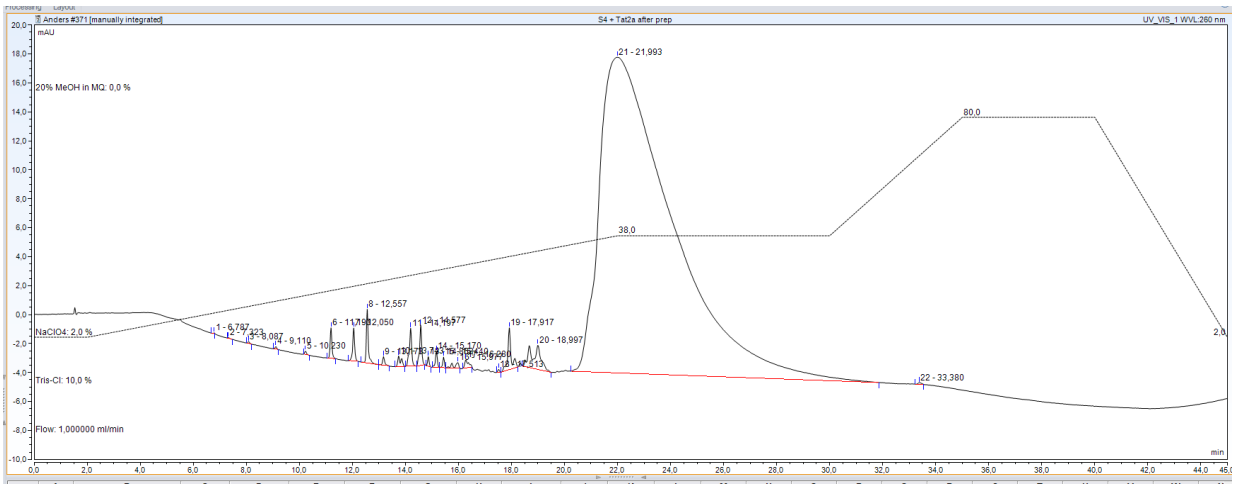


Figure A2. IC HPLC spectra of the siRNA sense strands-Tat1a conjugates

S2-Tat1a



S4-Tat1a



SCS2-Tat1a

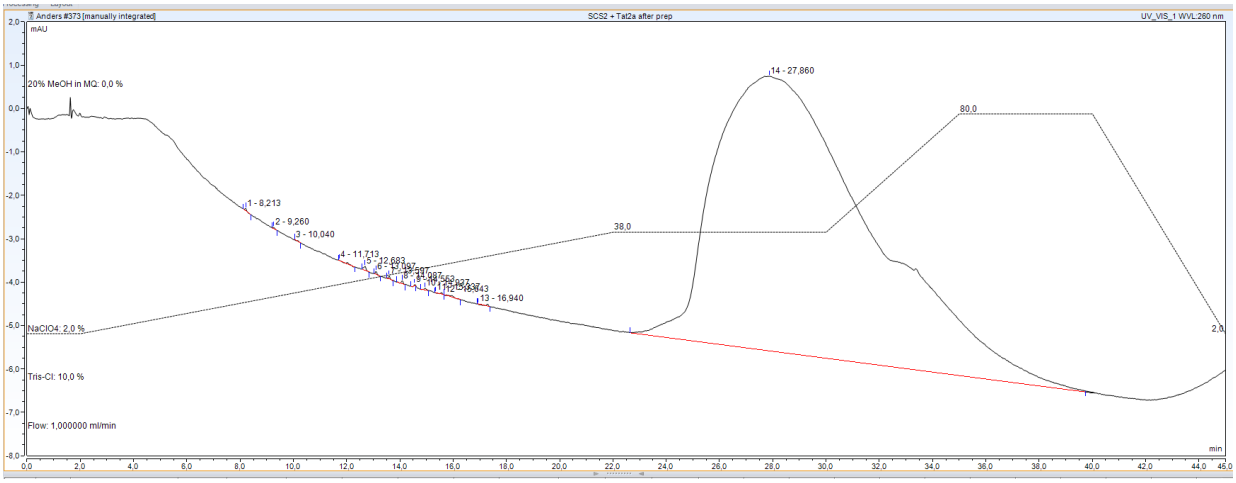
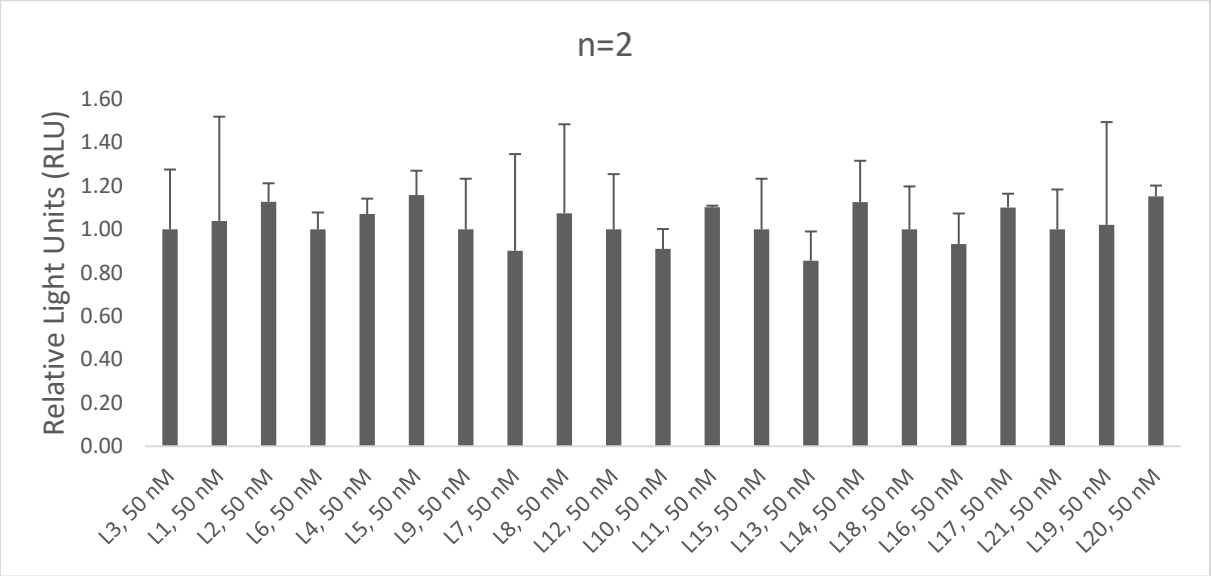


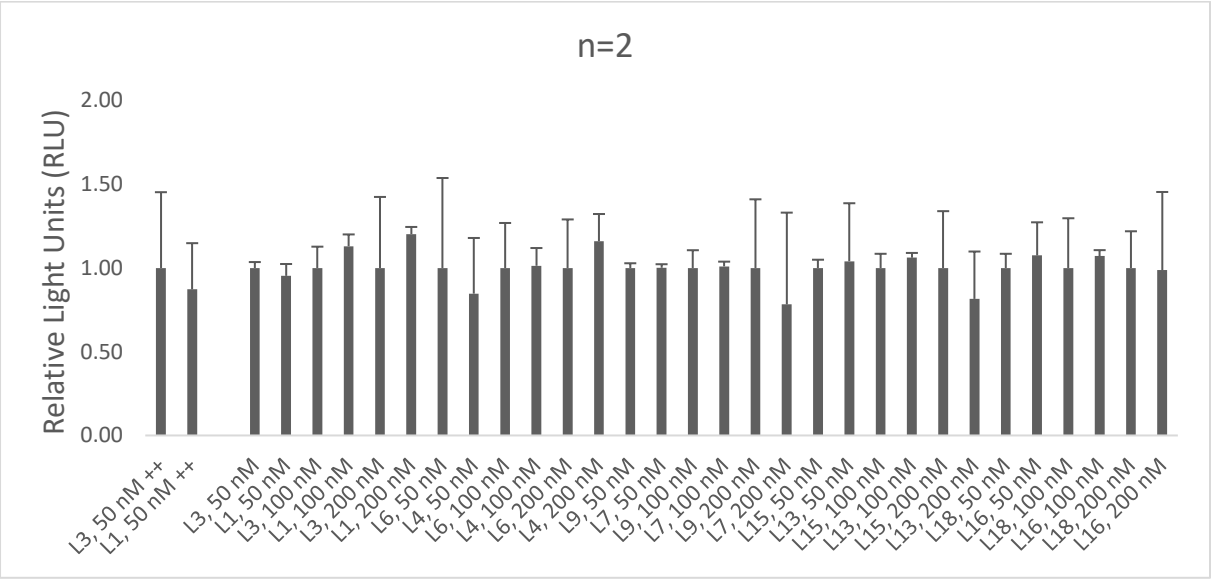
Chart A1. Luciferase reporter assay results; RLU representing the expression of the luciferase gene in epithelial cells TZMbl cells

a)



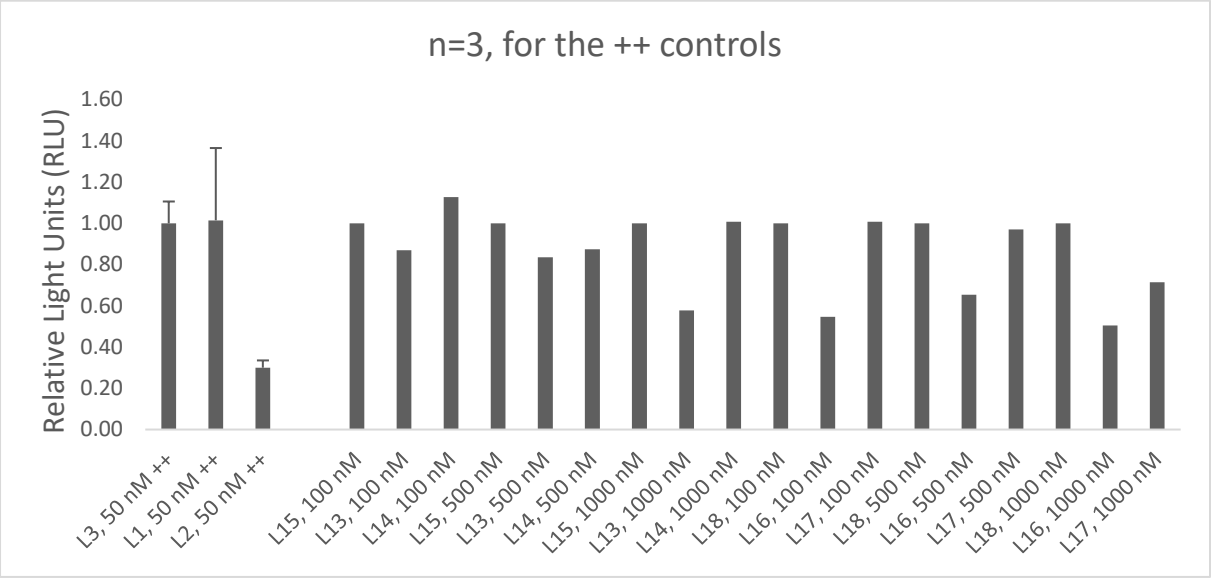
*Nanoplexes show no significant effect in TZMbl cells on the used concentration. 100000 cells/ml TZMbl were seeded and 24 h latter, 100 ng/well pcDNA-TAT-dsRED was transfected to activate HIV-1 transcription. 24 h after, 50 nM nanoplexes L1-L21 were dropped onto the cells. 48 h later, luminescence was determined.

b)



* 100000 cells/ml TZMbl were seeded and 24 h latter, 100 ng/well pcDNA-TAT-dsRED was transfected to activate HIV-1 transcription. 24 h after, 50 nM nanoplexes L1 and L3 were transfected using Lipofectamine 3000. In addition 50-200 nM nanoplexes were dropped onto the cells. 48 h later, luminescence was determined.

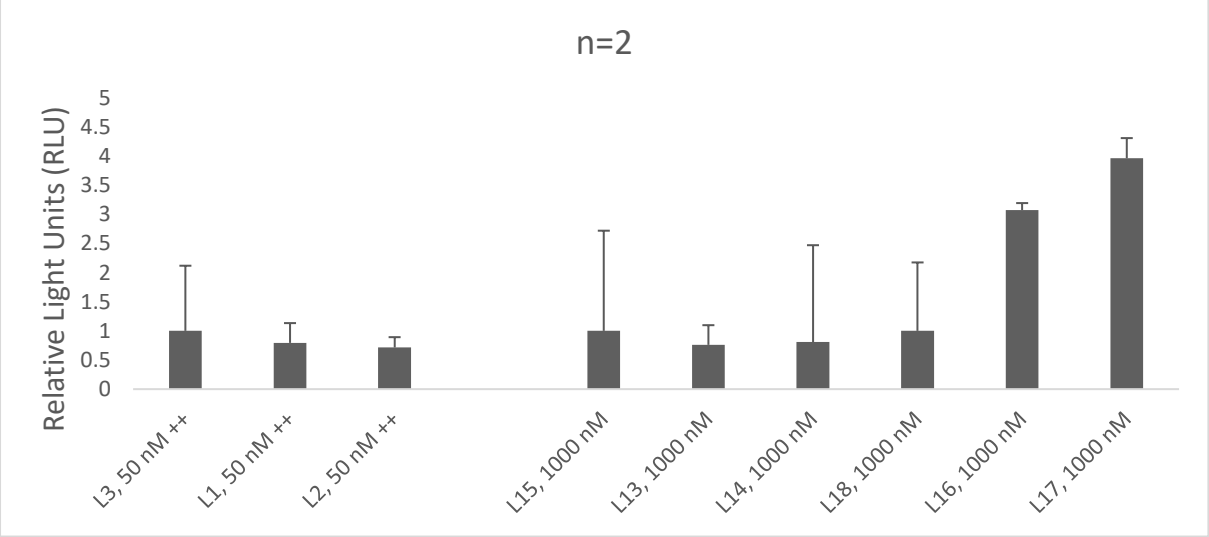
c)



* 100000 cells/ml TZMbl were seeded and 24 h latter, 100 ng/well pcDNA-TAT-dsRED was transfected to activate HIV-1 transcription. 24 h after, 50 nM nanoplexes L1 and L3 were transfected using Lipofectamine 3000. In addition 100-1000 nM nanoplexes were dropped onto the cells. 48 h later, luminescence was determined.

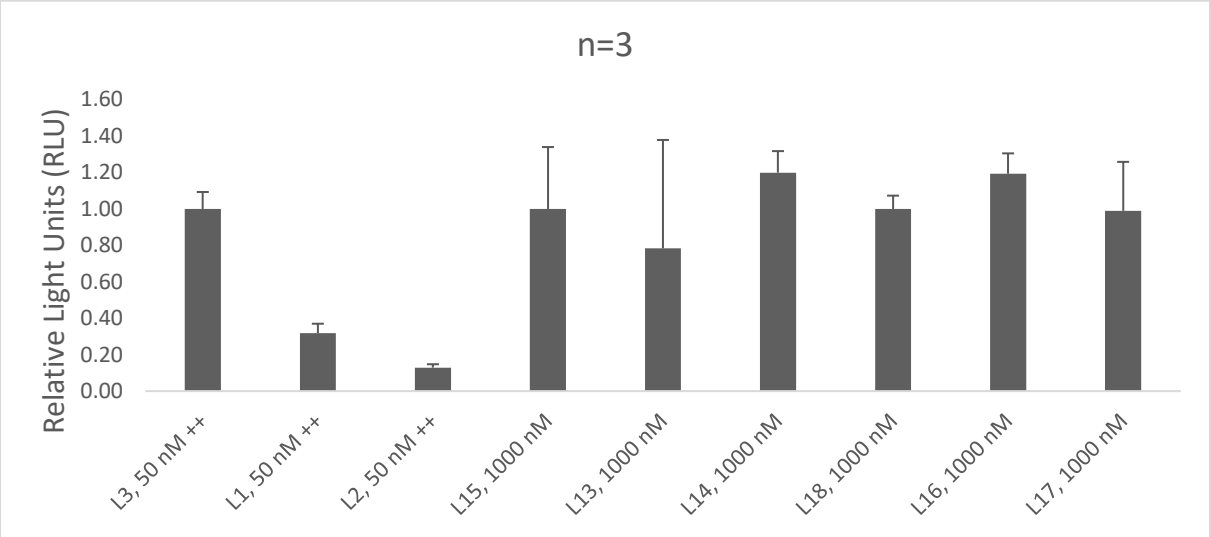
Chart A2. Dual luciferase reporter assay results; RLU representing the expression of the luciferase gene in epithelial cells TZMbl cells

a)



* 100000 cells/ml TZMbl were seeded and 24 h latter, 100 ng/well pcDNA-TAT-dsRED and 100 mg/well pcDNA-Renilla were transfected. After 24 h, 50 nM nanoplexes L1-L3 were transfected using Lipofectamine 3000 (++). In addition, 1000 nM nanoplexes L13-L18 were dropped onto the cells. 48 h later, luminescence was determined

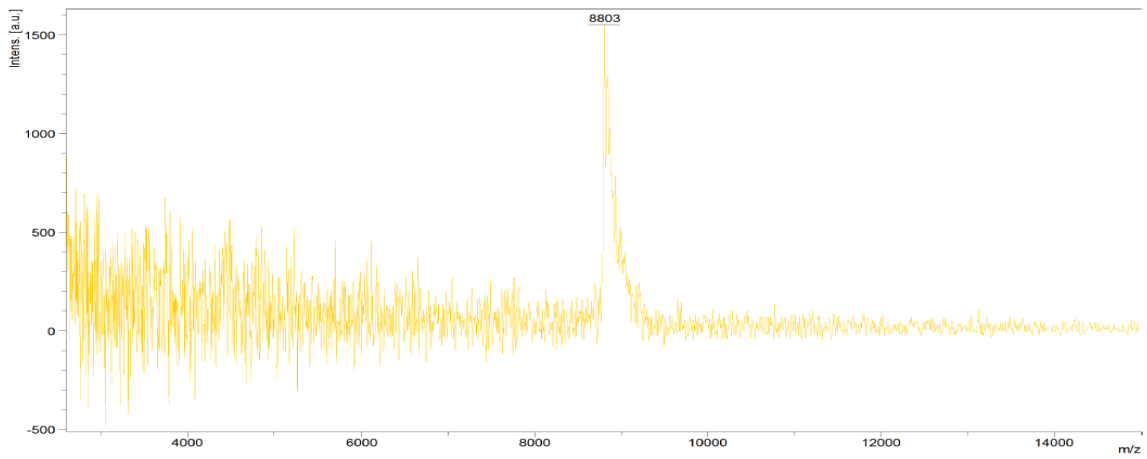
b)



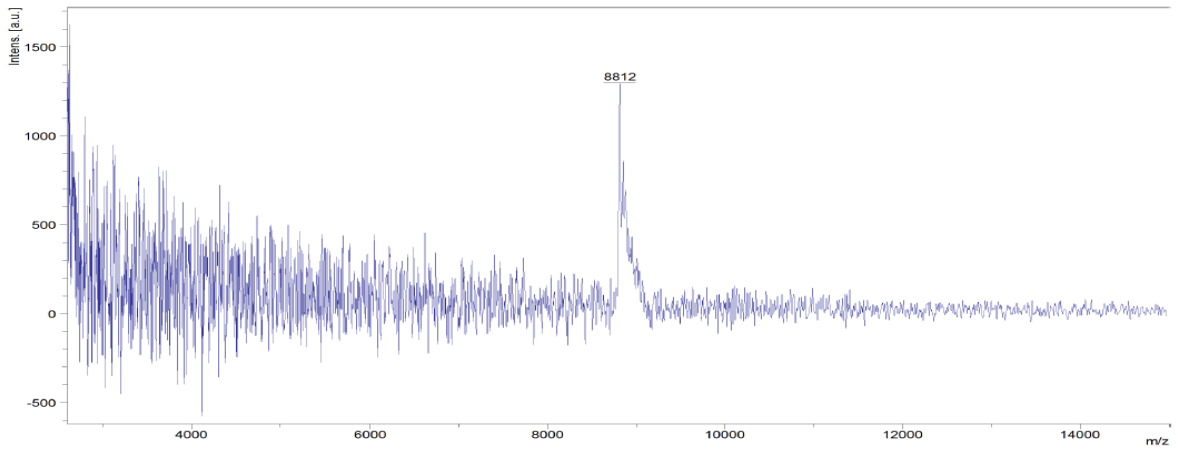
* 100000 cells/ml TZMbl were seeded and 24 h latter, 100 ng/well pcDNA-TAT-dsRED and 100 mg/well pcDNA-Renilla were transfected. After 24 h, 50 nM nanoplexes L1-L3 were transfected using Lipofectamine 3000 (++). In addition, 1000 nM nanoplexes L13-L18 were dropped onto the cells. 48 h later, luminescence was determined

Figure A3. Maldi-TOF spectra of the siBGas sense strands-Tat1a conjugates

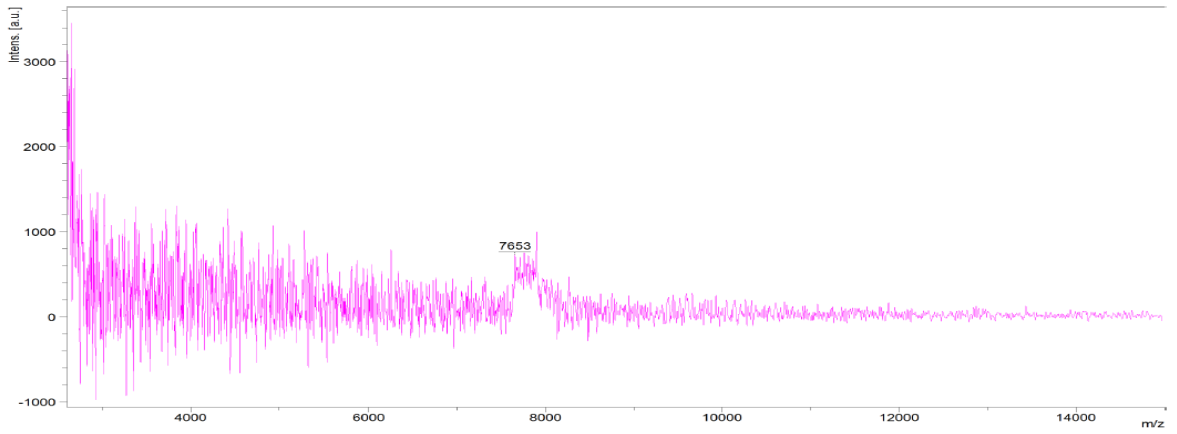
siBGas s-Tat1a



siBGas SC s–Tat1a



Gap-Tat1a



Gap SC-Tat1a

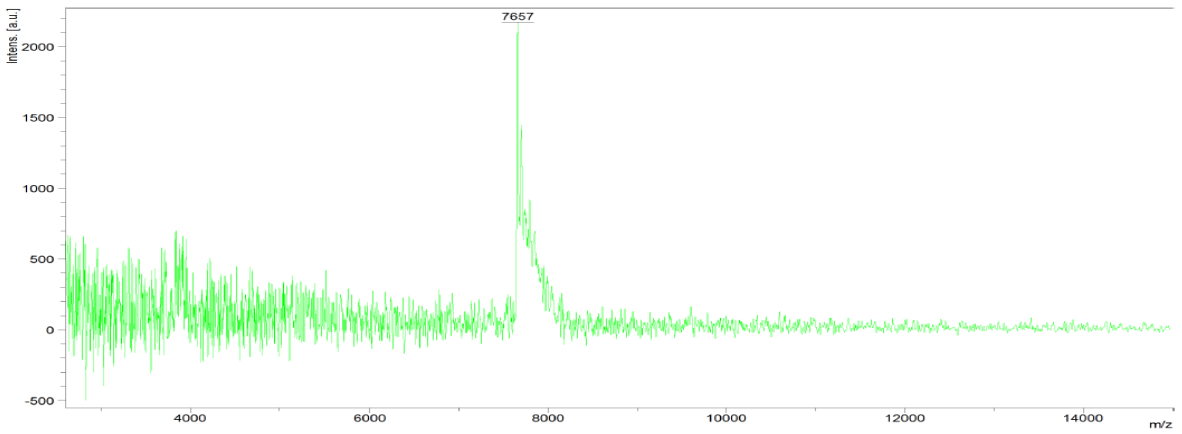
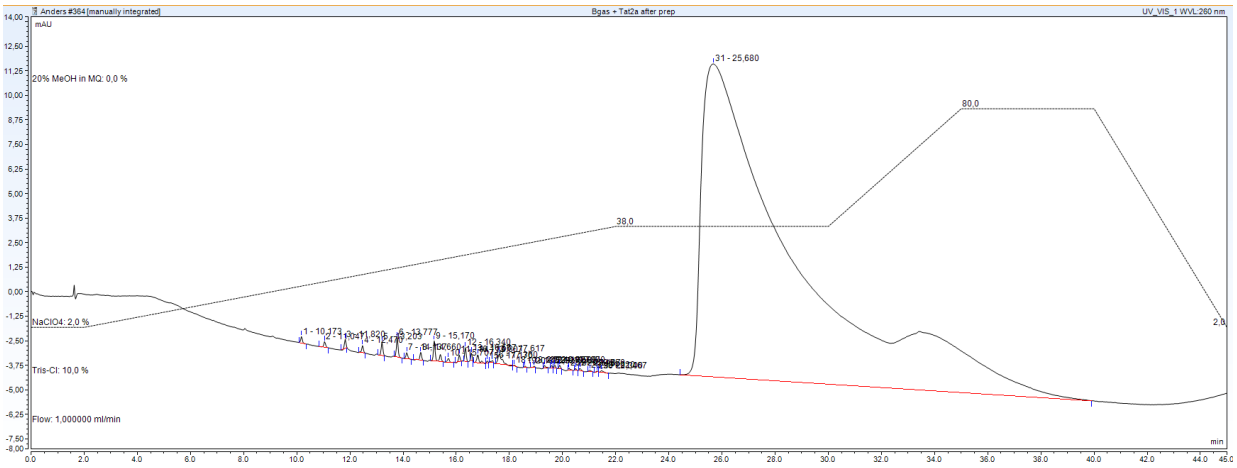
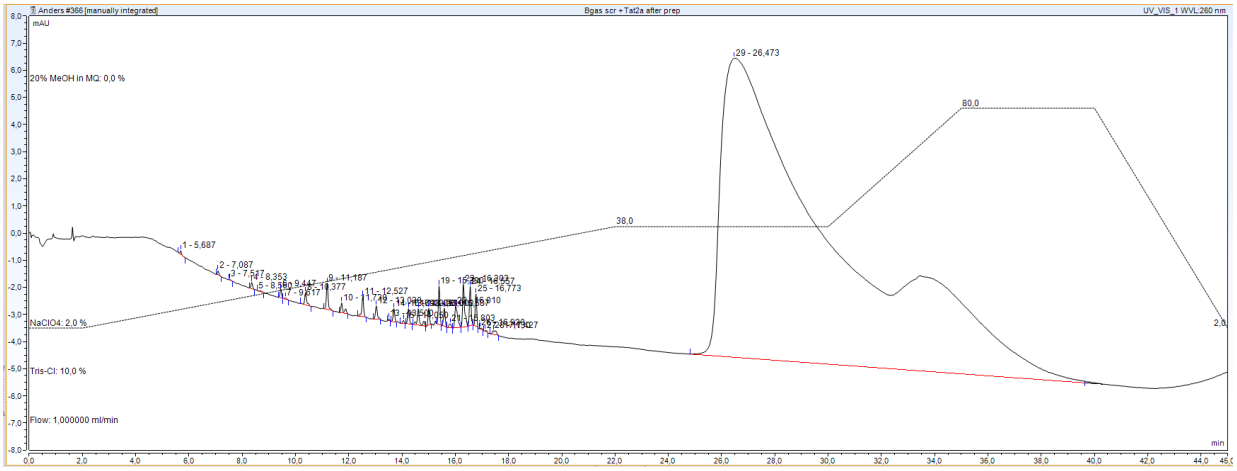


Figure A4. IC HPLC spectra of the siBGas sense strands/Gap-Tat1a conjugates

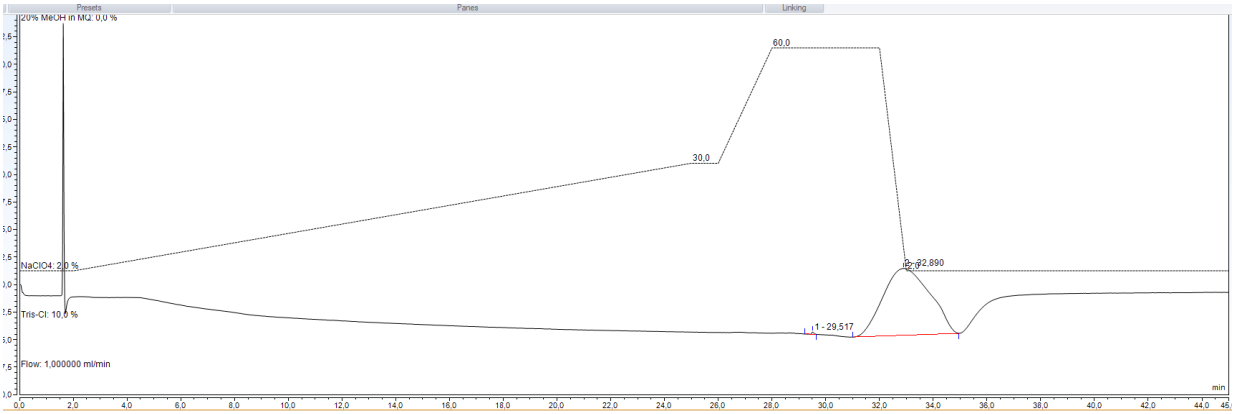
siBGas s-Tat1a



BGas SC s-Tat1a



Gap-Tat1a



Gap SC-Tat1a

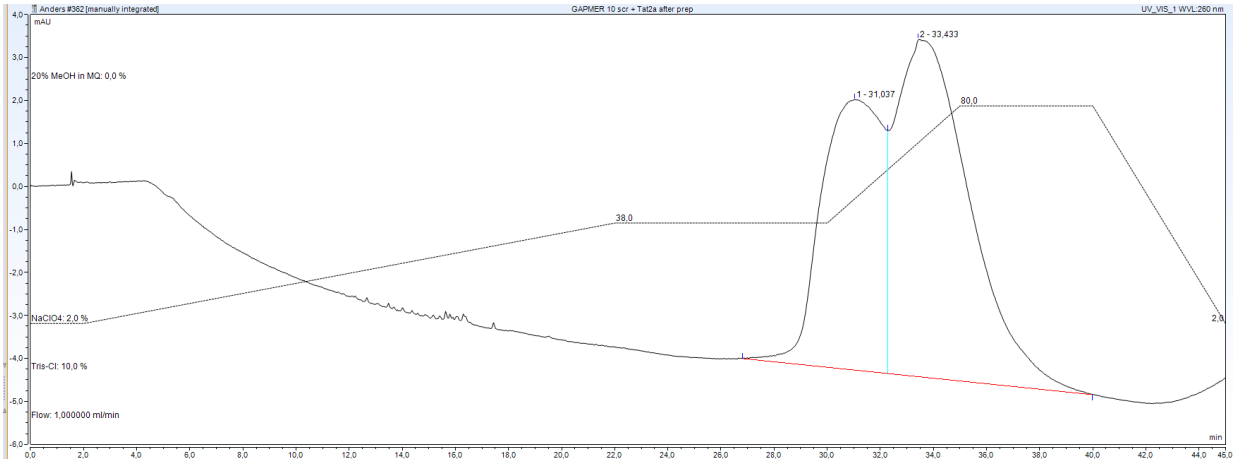


Table A1. Calculated CV (%) values for the data presented in Chart 5.3 (Chapter 5)

a)

	G1 CV (%)	G2	G3	G4
100 nM	77.9199	201.9721	122.8566	35.12727
200 nM	131.7461	194.6217	110.6064	310.2921
400 nM	93.6966	74.26713	30.01816	177.8161

b)

	G1 CV (%)	G2	G3	G4
10 nM	78.2249	165.9155	121.3167	245.5039
20 nM	163.1232	472.6099	111.4126	77.21052
50 nM	25.7082	42.01205	14.7213	53.59579

c)

	G1 CV (%)	G2	G3	G4
50 nM	59	1372.727		
2 nM	884		840.4255	9.52381
4 nM	472		2952.632	545.4545

Appendix B

Declaration of co-authorship

Paper I

Synthetic Nucleic Acid Analogues in Gene Therapy: An update for Peptide-Oligonucleotide Conjugates

Taskova M, Mantsiou A, Astakhova K., *Chembiochem.,* 2017, 17, 1676-82.

Paper II

Antisense Oligonucleotides Internally Labeled with Peptides Show Improved Target Recognition and Stability to Enzymatic Degradation

Taskova M, Madsen CS, Jensen KJ, Hansen LH, Vester B, Astakhova K., *Bioconjug Chem.*, 2017, 3, 768-74.

Paper III

Studies of Impending Oligonucleotide Therapeutics in Simulated Biofluids

Domljanovic I, Hansen AH, Hansen LH, Klitgaard JK, Taskova M*, Astakhova K., *Nucleic Acid Ther.*, 2018, doi: 10.1089/nat.2017.0704

Paper IV (In revision, submitted to Bioconjugate Chemistry)

Fluorescent oligonucleotides with bis (prop-2-yn-1-yloxy) butane-1,3-diol scaffold allow for rapid detection of disease associated nucleic acids

Taskova M, Astakhova K.

Declaration of co-authorship at DTU

If a PhD thesis contains articles (i.e. published journal and conference articles, unpublished manuscripts, chapters, etc.) written in collaboration with other researchers, a co-author statement verifying the PhD student's contribution to each article should be made.

If an article is written in collaboration with three or less researchers (including the PhD student), all researchers must sign the statement. However, if an article has more than three authors the statement may be signed by a representative sample, cf. article 12, section 4 and 5 of the Ministerial Order No. 1039, 27 August 2013. A representative sample consists of minimum three authors, which is comprised of the first author, the corresponding author, the senior author, and 1-2 authors (preferably international/non-supervisor authors).

DTU has implemented the Danish Code of Conduct for Research Integrity, which states the following regarding attribution of authorship:

"Attribution of authorship should in general be based on criteria a-d adopted from the Vancouver guidelines[1], and all individuals who meet these criteria should be recognized as authors:

a. Substantial contributions to the conception or design of the work, or the acquisition, analysis, or interpretation of data for the work, *and*
b. drafting the work or revising it critically for important intellectual content, *and*
c. final approval of the version to be published, *and*
d. agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved."[2]

For more information regarding definition of co-authorship and examples of authorship conflicts, we refer to DTU Code of Conduct for Research Integrity (pp. 19-22).

By signing this declaration, co-authors permit the PhD student to reuse whole or parts of co-authored articles in their PhD thesis, under the condition that co-authors are acknowledged in connection with the reused text or figure.

It is **important** to note that it is the responsibility of the PhD student to obtain permission from the publisher to use the article in the PhD thesis[3]

[1] International Committee of Medical Journal Editors – Recommendations for the Conduct, Reporting, Editing, and Publication of Scholarly Work in Medical Journals, updated December 2016
[2] DTU Code of Conduct for Research Integrity (E-book p. 19)
[3] Many journals will allow you to use only the post-print version of your article, meaning the accepted version of the article, without the publisher's final formatting. In the event that your article is submitted, but still under review, you should of course use the latest submitted version of your article in your thesis. Always remember to check your publisher's guidelines on reuse of published articles. Most journals, unless open access, have an embargo period on published articles, meaning that within this period you cannot freely use the article. Check your publisher's rules on this issue.

Title of article

Synthetic Nucleic Acid Analogues in Gene Therapy: An update for Peptide-Oligonucleotide Conjugates

Journal/conference

ChembioChem

Author(s)

Taskova M, Mantsiou A, Astakhova K

Name (capital letters) and signature of PhD student

Maria Taskova

PhD student's date of birth

11.08.1990

Declaration of the PhD student's contribution

For each type of work, please specify below the contribution as appropriate

	Minor contribution to the work *(please specify)*	**Substantial contribution to the work** *(please specify)*
Formulation of the conceptual framework and/or planning of the design of the study including scientific questions		Ms. Taskova outlined the paper, made literature mining, discussed the key points of the paper with other authors, clarified scientific questions
Carrying out of experiments/data collection and analysis/interpretation of results		Ms. Taskova analyzed the literature and selected the appropriate information that was included into the review
Writing of the article/revising the manuscript for intellectual content		Ms. Taskova wrote the initial draft of the paper, and upon receiving comments and corrections from co-authors and Referee's, conducted proof reading, prepared and submitetd a final lay out.

Signatures

Title of article
Synthetic Nucleic Acid Analogues in Gene Therapy: An update for Peptide-Oligonucleotide Conjugates
Journal/conference
ChembioChem
Author(s)
Taskova M, Mantsiou A, Astakhova K
Name (capital letters) and signature of PhD student
Maria Taskova
PhD student's date of birth
11.08.1990

Date	Name	Title	Signature
7/3-19	Kira Astakhova	Associate Professor	
01.03.2019	Anna Mantsiou	Master student	
11.03.2019	Maria Taskova	PhD student	

Please note that by signing this declaration, co-authors permit the PhD student to reuse whole or parts of co-authored articles in their PhD thesis, under the condition that co-authors are acknowledged in connection with the reused text or figure.

Declaration of co-authorship at DTU

If a PhD thesis contains articles (i.e. published journal and conference articles, unpublished manuscripts, chapters, etc.) written in collaboration with other researchers, a co-author statement verifying the PhD student's contribution to each article should be made.

If an article is written in collaboration with three or less researchers (including the PhD student), all researchers must sign the statement. However, if an article has more than three authors the statement may be signed by a representative sample, cf. article 12, section 4 and 5 of the Ministerial Order No. 1039, 27 August 2013. A representative sample consists of minimum three authors, which is comprised of the first author, the corresponding author, the senior author, and 1-2 authors (preferably international/non-supervisor authors).

DTU has implemented the Danish Code of Conduct for Research Integrity, which states the following regarding attribution of authorship:

"Attribution of authorship should in general be based on criteria a-d adopted from the Vancouver guidelines[1], and all individuals who meet these criteria should be recognized as authors:

a. Substantial contributions to the conception or design of the work, or the acquisition, analysis, or interpretation of data for the work, *and*
b. drafting the work or revising it critically for important intellectual content, *and*
c. final approval of the version to be published, *and*
d. agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved."[2]

For more information regarding definition of co-authorship and examples of authorship conflicts, we refer to DTU Code of Conduct for Research Integrity (pp. 19-22).

By signing this declaration, co-authors permit the PhD student to reuse whole or parts of co-authored articles in their PhD thesis, under the condition that co-authors are acknowledged in connection with the reused text or figure.

It is **important** to note that it is the responsibility of the PhD student to obtain permission from the publisher to use the article in the PhD thesis[3]

[1] International Committee of Medical Journal Editors – Recommendations for the Conduct, Reporting, Editing, and Publication of Scholarly Work in Medical Journals, updated December 2016
[2] DTU Code of Conduct for Research Integrity (E-book p. 19)
[3] Many journals will allow you to use only the post-print version of your article, meaning the accepted version of the article, without the publisher's final formatting. In the event that your article is submitted, but still under review, you should of course use the latest submitted version of your article in your thesis. Always remember to check your publisher's guidelines on reuse of published articles. Most journals, unless open access, have an embargo period on published articles, meaning that within this period you cannot freely use the article. Check your publisher's rules on this issue.

Title of article
Antisense Oligonucleotides Internally Labeled with Peptides Show Improved Target Recognition and Stability to Enzymatic Degradation
Journal/conference
Bioconjugate Chemistry
Author(s)
Taskova M, Madsen CS, Jensen KJ, Hansen LH, Vester B, Astakhova K
Name (capital letters) and signature of PhD student
Maria Taskova
PhD student's date of birth
11.08.1990
Declaration of the PhD student's contribution
For each type of work, please specify below the contribution as appropriate

	Minor contribution to the work *(please specify)*	**Substantial contribution to the work** *(please specify)*
Formulation of the conceptual framework and/or planning of the design of the study including scientific questions		Ms Taskova contributed significantly to study design including design of peptide and oligonucleotide sequences, planning assays, trouble shooting and data handling
Carrying out of experiments/data collection and analysis/interpretation of results		Ms Taskova executed the oligonucleotide synthesis, purification, characterization, conjugation with peptides, and studies the UV melting, CD and enzymatic stability of POCs.
Writing of the article/revising the manuscript for intellectual content		Ms Taskova wrote the initial draf tand participated in each revision round of the paper.

Signatures

Title of article
Antisense Oligonucleotides Internally Labeled with Peptides Show Improved Target Recognition and Stability to Enzymatic Degradation
Journal/conference
Bioconjugate Chemistry
Author(s)
Taskova M, Madsen CS, Jensen KJ, Hansen LH, Vester B, Astakhova K
Name (capital letters) and signature of PhD student
Maria Taskova
PhD student's date of birth
11.08.1990

Date	Name	Title	Signature
19.3.2019	Kira Astakhova	Associate Professor	
3/15-2019	Jensen Knud	Professor	
17.3.2019	Hansen Lykke	Researcher	
19.3.2019	Maria Taskova	PhD student	

Please note that by signing this declaration, co-authors permit the PhD student to reuse whole or parts of co-authored articles in their PhD thesis, under the condition that co-authors are acknowledged in connection with the reused text or figure.

Declaration of co-authorship at DTU

If a PhD thesis contains articles (i.e. published journal and conference articles, unpublished manuscripts, chapters, etc.) written in collaboration with other researchers, a co-author statement verifying the PhD student's contribution to each article should be made.

If an article is written in collaboration with three or less researchers (including the PhD student), all researchers must sign the statement. However, if an article has more than three authors the statement may be signed by a representative sample, cf. article 12, section 4 and 5 of the Ministerial Order No. 1039, 27 August 2013. A representative sample consists of minimum three authors, which is comprised of the first author, the corresponding author, the senior author, and 1-2 authors (preferably international/non-supervisor authors).

DTU has implemented the Danish Code of Conduct for Research Integrity, which states the following regarding attribution of authorship:

"Attribution of authorship should in general be based on criteria a-d adopted from the Vancouver guidelines[1], and all individuals who meet these criteria should be recognized as authors:

a. Substantial contributions to the conception or design of the work, or the acquisition, analysis, or interpretation of data for the work, *and*
b. drafting the work or revising it critically for important intellectual content, *and*
c. final approval of the version to be published, *and*
d. agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved."[2]

For more information regarding definition of co-authorship and examples of authorship conflicts, we refer to DTU Code of Conduct for Research Integrity (pp. 19-22).

By signing this declaration, co-authors permit the PhD student to reuse whole or parts of co-authored articles in their PhD thesis, under the condition that co-authors are acknowledged in connection with the reused text or figure.

It is **important** to note that it is the responsibility of the PhD student to obtain permission from the publisher to use the article in the PhD thesis[3]

[1] International Committee of Medical Journal Editors – Recommendations for the Conduct, Reporting, Editing, and Publication of Scholarly Work in Medical Journals, updated December 2016
[2] DTU Code of Conduct for Research Integrity (E-book p. 19)
[3] Many journals will allow you to use only the post-print version of your article, meaning the accepted version of the article, without the publisher's final formatting. In the event that your article is submitted, but still under review, you should of course use the latest submitted version of your article in your thesis. Always remember to check your publisher's guidelines on reuse of published articles. Most journals, unless open access, have an embargo period on published articles, meaning that within this period you cannot freely use the article. Check your publisher's rules on this issue.

Title of article
Studies of Impending Oligonucleotide Therapeutics in Simulated Biofluids
Journal/conference
Nucleic acid therapeutics
Author(s)
Domljanovic I, Hansen AH, Hansen LH, Klitgaard JK, Taskova M, Astakhova K
Name (capital letters) and signature of PhD student
Maria Taskova
PhD student's date of birth
11.08.1990
Declaration of the PhD student's contribution
For each type of work, please specify below the contribution as appropriate

	Minor contribution to the work *(please specify)*	**Substantial contribution to the work** *(please specify)*
Formulation of the conceptual framework and/or planning of the design of the study including scientific questions		Ms. Taskova conceptualized the study, selected test systems and conditions, planned the scientific questions to be addressed and navigated the selecting of the scientific questions while the study has been conducted.
Carrying out of experiments/data collection and analysis/interpretation of results		Ms. Taskova independently conducted the digestion studies of multiple conjugates, did data handling for all test systems, and contributed significantly to interpretation of the results
Writing of the article/revising the manuscript for intellectual content		Ms. Taskova proof read an initial draft of the paper, and revised it continuously after receiving the commments by the co-authors and Referees.

Signatures

Title of article

Studies of Impending Oligonucleotide Therapeutics in Simulated Biofluids

Journal/conference

Nucleic acid therapeutics

Author(s)

Domljanovic I, Hansen AH, Hansen LH, Klitgaard JK, Taskova M, Astakhova K

Name (capital letters) and signature of PhD student

Maria Taskova

PhD student's date of birth

11.08.1990

Date	Name	Title	Signature
19.3.2019	Kira Astakhova	Associate Professor	Kira Astakhova
11/3/2019	Anders H. Hansen	Postdoctoral Researcher	
7/3/2019	Ivana Domljanovic	PhD student	Ivana Domljanovic
11.3.2019	Maria Taskova	PhD student	

Please note that by signing this declaration, co-authors permit the PhD student to reuse whole or parts of co-authored articles in their PhD thesis, under the condition that co-authors are acknowledged in connection with the reused text or figure.

Declaration of co-authorship at DTU

If a PhD thesis contains articles (i.e. published journal and conference articles, unpublished manuscripts, chapters, etc.) written in collaboration with other researchers, a co-author statement verifying the PhD student's contribution to each article should be made.

If an article is written in collaboration with three or less researchers (including the PhD student), all researchers must sign the statement. However, if an article has more than three authors the statement may be signed by a representative sample, cf. article 12, section 4 and 5 of the Ministerial Order No. 1039, 27 August 2013. A representative sample consists of minimum three authors, which is comprised of the first author, the corresponding author, the senior author, and 1-2 authors (preferably international/non-supervisor authors).

DTU has implemented the Danish Code of Conduct for Research Integrity, which states the following regarding attribution of authorship:

"Attribution of authorship should in general be based on criteria a-d adopted from the Vancouver guidelines[1], and all individuals who meet these criteria should be recognized as authors:

a. Substantial contributions to the conception or design of the work, or the acquisition, analysis, or interpretation of data for the work, *and*
b. drafting the work or revising it critically for important intellectual content, *and*
c. final approval of the version to be published, *and*
d. agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved."[2]

For more information regarding definition of co-authorship and examples of authorship conflicts, we refer to DTU Code of Conduct for Research Integrity (pp. 19-22).

By signing this declaration, co-authors permit the PhD student to reuse whole or parts of co-authored articles in their PhD thesis, under the condition that co-authors are acknowledged in connection with the reused text or figure.

It is **important** to note that it is the responsibility of the PhD student to obtain permission from the publisher to use the article in the PhD thesis[3]

[1] International Committee of Medical Journal Editors – Recommendations for the Conduct, Reporting, Editing, and Publication of Scholarly Work in Medical Journals, updated December 2016
[2] DTU Code of Conduct for Research Integrity (E-book p. 19)
[3] Many journals will allow you to use only the post-print version of your article, meaning the accepted version of the article, without the publisher's final formatting. In the event that your article is submitted, but still under review, you should of course use the latest submitted version of your article in your thesis. Always remember to check your publisher's guidelines on reuse of published articles. Most journals, unless open access, have an embargo period on published articles, meaning that within this period you cannot freely use the article. Check your publisher's rules on this issue.

Title of article

Fluorescent oligonucleotides with bis (prop-2-yn-1-yloxy) butane-1,3-diol scaffold allow for rapid detection of disease associated nucleic acids

Journal/conference

Bioconjugate Chemistry

Author(s)

Taskova M, Astakhova K

Name (capital letters) and signature of PhD student

Maria Taskova

PhD student's date of birth

11.08.1990

Declaration of the PhD student's contribution

For each type of work, please specify below the contribution as appropriate

	Minor contribution to the work *(please specify)*	**Substantial contribution to the work** *(please specify)*
Formulation of the conceptual framework and/or planning of the design of the study including scientific questions		Ms. Taskova contributed to conceptual framework of the project, planed the syntesis and testing, specified the scientific questions of the work
Carrying out of experiments/data collection and analysis/interpretation of results		Ms. Taskova independently synthesized and characterized all the compounds and oligonucleotides included in this paper
Writing of the article/revising the manuscript for intellectual content		Ms. Taskova independently wrote the manuscript that has been afterwards revised by a co-author

Signatures

Title of article
Fluorescent oligonucleotides with bis (prop-2-yn-1-yloxy) butane-1,3-diol scaffold allow for rapid detection of disease associated nucleic acids
Journal/conference
Bioconjugate Chemistry
Author(s)
Taskova M, Astakhova K
Name (capital letters) and signature of PhD student
Maria Taskova
PhD student's date of birth
11.08.1990

Date	Name	Title	Signature
7/3-19	Kira Astakhova	Associate Professor	
11.3.2019	Maria Taskova	PhD student	

Please note that by signing this declaration, co-authors permit the PhD student to reuse whole or parts of co-authored articles in their PhD thesis, under the condition that co-authors are acknowledged in connection with the reused text or figure.