From Department of Laboratory Medicine
Karolinska Institutet, Stockholm, Sweden

ANTI-GENE OLIGONUCLEOTIDES
DNA binding and therapeutic application

Tea Umek

Karolinska Institutet

Stockholm 2021

Published by Karolinska Institutet.
Printed by Universitetsservice US-AB, 2021

ISBN 978-91-8016-346-0

Anti-gene oligonucleotides: DNA binding and therapeutic application

THESIS FOR DOCTORAL DEGREE (Ph.D.)

By

Tea Umek

The thesis will be defended in public at Birkeaulan F51 Karolinska Universitetssjukhuset Huddinge, 22nd October 2021, 10:00

Principal Supervisor:
Associate Professor Rula Zain
Karolinska Institutet
Department of Laboratory Medicine
Division of Biomolecular and Cellular Medicine

Co-supervisor(s):
Professor C.I. Edvard Smith
Karolinska Institutet
Department of Laboratory Medicine
Division of Biomolecular and Cellular Medicine

Ph.D. Karin Lundin
Karolinska Institutet
Department of Laboratory Medicine
Division of Biomolecular and Cellular Medicine

Opponent:
Associate Professor Eulàlia Martí Puig
Universitat de Barcelona
Department of Biomedical Sciences

Examination Board:
Professor Magnus Nordenskjöld
Karolinska Institutet
Department of Molecular Medicine and Surgery

Professor Matti Sällberg
Karolinska Institutet
Department of Laboratory Medicine
Division of Clinical Microbiology

Associate Professor Oommen Varghese
Uppsala Universitet
Department of Chemistry

To my family

ABSTRACT

The field of nucleic acid therapeutics has in the last decade experienced substantial growth. This is evident from a multitude of new publications, an increasing number of clinical trials and several approved therapeutics on the market. Therapeutic oligonucleotides (ONs) are designed to treat mostly genetic disorders, caused by the expression of non-functional or toxic ribonucleic acids (RNA) or proteins. They can be divided into protein-, RNA- and DNA targeting ONs. While the first two have already had approved therapies, the last is still in its infancy.

This thesis focuses on the improvement and assessment of DNA targeting ONs, also termed anti-gene ONs. Paper I evaluates a strand-invading anti-gene ON (AGO) for the treatment of Huntington's disease (HD) and its mRNA down-regulating effect during *in vitro* differentiation from induced pluripotent stem cells to neurons. It shows that the locked nucleic acid (LNA)/DNA mixmer ON with a phosphorothioate (PS) backbone directed against the repeat region of the *huntingtin (HTT)* gene downregulates *HTT* mRNA and protein without affecting the process of differentiation.

In Paper II, anti-gene LNA/DNA mixmer clamp-type ONs, which hybridize to the target by forming both Watson-Crick (WC) and Hoogsteen hydrogen bonds, are optimized to achieve improved invasion into double stranded DNA. By positioning an intercalating moiety between or on the WC and Triplex Forming ON (TFO) arms, the AGOs achieve efficient invasion at nanomolar concentrations *in vitro*. Furthermore, the corresponding PS modified ONs are tested in HD cell model for their effect on mRNA expression, where they cause significant downregulation of *HTT* mRNA.

In Paper III, a non-B-DNA structure formed by a sequence in the *MYC* gene promoter is analyzed *in vitro*. Furthermore, the effect of this structure, more specifically an H-DNA, on the strand-invasion efficiency of LNA/DNA mixmer tail-clamp AGOs was evaluated. The results show that the invasion of the ON can be positively influenced by an H-DNA if the ON does not have an intercalating moiety.

LIST OF SCIENTIFIC PAPERS

The following papers are included in the thesis.

I. **Tea Umek***, Thomas Olsson*, Olof Gissberg, Osama Saher, Eman M. Zaghloul, Karin E. Lundin, Jesper Wengel, Eric Hanse, Henrik Zetterberg, Dzeneta Vizlin-Hodzic, C. I. Edvard Smith and Rula Zain. **Oligonucleotides targeting DNA repeats down-regulate *Huntingtin* gene expression in Huntington's patient-derived neural model system**
* these authors contributet equally

II. **Tea Umek**, Karin Lundin, Per T. Jørgensen, Jesper Wengel, Rula Zain, C. I. Edvard Smith. **Anti-gene oligonucleotide clamps strand-invade into dsDNA down-regulating huntingtin expression**. (Manuscript 2021)

III. **Tea Umek**, Karin Sollander, Helen Bergquist, Jesper Wengel, Karin E Lundin, C. I. Edvard Smith, Rula Zain. **Oligonucleotide Binding to Non-B-DNA in *MYC.*** Molecules. 2019 Mar 12;24(5):1000.

Scientific paper and book chapter not included in the thesis

I. Negin Mozafari, **Tea Umek**. **Assessing Oligonucleotide Binding to Double-Stranded DNA**. Methods Mol Biol. 2019;2036:91-112. doi: 10.1007/978-1-4939-9670-4_5. PMID: 31410792

II. Osama Saher, Eman Zaghloul, **Tea Umek**, Karin E. Lundin, Jesper Wengel, C.I. Edvard Smith, Rula Zain. **Parameters influencing activity of anti-gene CAG oligonucleotides in Huntigton's disease patient cells**. (Manuscript 2021)

.

CONTENTS

LIST OF ABBREVIATIONS

2’-F	2’-fluoro
2’-*O*-Me	2’-*O*-methyl
2’-MOE	2’-*O*-methoxyethyl
A	Adenine
Ab	Antibody
AGO	Anti-gene oligonucleotide
AR	Androgen receptor
ASO	Antisense oligonucleotide
ATN1	Atrophin-1
ATXN2	Ataxin-2
BDNF	Brain-derived neurotrophic factor
bp	Base pair
BQQ	Benzoquinoquinoxaline
BRI3BP	BRI3 binding protein
C	Cytosine
DM1	Myotonic dystrophy type 1
DMD	Duchenne muscular dystrophy
DMEM	Dulbecco’s Modified Eagle Medium
DMPK	DM1 protein kinase
DNA	Deoxyribonucleic acid
dsDNA	Double stranded DNA
dsRNA	Double stranded RNA
EB	Embryoid body
EMSA	Electrophoretic Mobility Shift Assay
FBS	Fetal bovine serum
FDA	Food and Drug Administration
FRDA	Friedreich’s ataxia
G	Guanine
GalNAc	N-acetylgalactosamine
GDNF	Glial-derived neurotrophic factor

GOI	Gene of interest
H-bond	Hydrogen bond
HD	Huntington's disease
HTT	Huntingtin
iPSC	induced Pluripotent Stem Cell
Ki67	Proliferation marker protein Ki-67
LNA	Locked Nucleic Acid
lncRNA	Long noncoding RNA
MAP2ab	Mature neuron-specific microtubule-associated protein 2, isoforms a and b
miRNA	MicroRNAs
mRNA	messenger RNA
mt	Mutant
MYC	MYC Proto-Oncogene
$NHEIII_1$	Nuclease hypersensitivity element III_1
NSC	Neuronal stem cells
nt	Nucleotide
NT	Non-treated
OCT4	Octamer-binding transcription factor 4
ON	Oligonucleotide
OP	Ortho-phenanthroline
ORF	Open reading frame
PAGE	Polyacrylamide gel electrophoresis
PAX6	Paired box protein Pax-6
PCR	Polymerase chain reaction
PMO	Phosphorodiamidate morpholino oligomer
PNA	Peptide nucleic acid
PO	Phosphodiester
PolyA	Polyadenylation
POU3F2	POU-homeodomain transcription factor BRN2
PS	Phosphorothioate
Pu	Polypurine

PVDF	Polyvinylidene difluoride
Py	Polypyrimidine
qPCR	Quantitative PCR
RAN	Repeat associated non-AUG
RISC	RNA induced silencing complex
RNA	Ribonucleic acid
RNAi	RNA interference
RNase H1	Ribonuclease H1
RT	Reverse transcription
S100	Astrocyte-specific S100 calcium-binding protein
siRNA	Small interfering RNA
SMA	Spinal muscular atrophy
ssDNA	Single stranded DNA
SSEA4	Stage-specific embryonic antigen 4
SSO	Splice switching oligonucleotide
T	Thymine
TFO	Triplex-forming oligonucleotide
*T*m	Melting temperature
UTR	Untranslated region
WB	Western blot
WC	Watson-Crick
wt	Wild-type

1 INTRODUCTION

Therapeutic oligonucleotides (ONs) fall under the group of "therapeutic nucleic acids". If until recently the general population was mostly unaware of the therapeutic potential of nucleic acids, it embraced it in 2021 in the form of new mRNA vaccines against SARS-CoV-2. Generally, therapeutic nucleic acids encompass synthetic oligomers corresponding to protein-coding deoxyribonucleic acid (DNA) sequences which can execute various cellular functions. Their sequence can also correspond to non-coding DNA or ribonucleic acid (RNA) possessing a regulatory function. Function-bearing RNAs can additionally be expressed inside the cell after the transfection of corresponding genes. Therapeutic ONs are generally divided into subcategories based on the intended target, which could be DNA, RNA, or protein.

The foundations for the therapeutic use of nucleic acids and gene therapy in general, as we understand it today, were laid through the 20th century. Some of the important milestones are the discovery of bacterial transformation, conjugation and transduction [1-4], uncovering the structure of B-DNA double-helix [5], discovery of messenger RNA (mRNA) [6-8], first transfer of genetic material into mammalian cells [9] and viral transfection and integration of viral DNA into genomic DNA within cells [10]. Furthermore, fundamental achievements include the first report of inhibition of translation, replication and cell transformation using oligonucleotides [11, 12], the discovery of Ribonuclease H1 (RNase H1) mediated cleavage of RNA:DNA duplexes [13] and RNA-induced silencing complex (RICS) mediated cleavage of RNA:RNA duplexes [14, 15]. Since the therapeutic effect of nucleic acids is sequence dependant, one of the key moments was also determining the sequence of the human genome [16, 17].

The first gene therapy clinical trials began from 1970 to 1990 and used viral vectors carrying genetic information for the synthesis of proteins, for the treatment of urea cycle disorder [18], ß-thalassemia [19, 20], advanced melanoma [21], and adenosine deaminase deficiency (the first-ever gene therapy clinical trial approved by the FDA) [22]. Regarding the therapeutic ONs, in 1998 Fomiversen (Vitravene), an antisense ON (ASO), was the first approved drug on the market, used for the treatment of cytomegalovirus retinitis [23]. The second drug approved was Pegaptanib (Macugen), an aptamer ON, for the treatment of neovascular age-related macular degeneration. Both of them are now rarely used due to development of better medicines. Since the early 2000 and initially with negligible success, the field started growing exponentially in the last decade, with thirteen ON drugs being approved since 2013 and many more in the pipeline (see Table 1 for detailed list of approved drugs on the market) [24].

Despite the newfound success, the drawbacks of using ONs as a therapeutic strategy, such as limited tissue distribution, fast renal clearance, cytotoxicity, and low stability, limit its broad utility. Unmodified nucleic acids are quickly degraded by plasma and cellular nucleases upon entering the extra- and subsequently, intracellular environment [25, 26]. To improve the pharmacokinetic and pharmacodynamic profile, many chemical modifications of each of the three components of the nucleotide (nt) – the nucleobase, the pentose, and the phosphodiester

backbone – have been proposed and will be discussed further on. Furthermore, limitations of the pharmaceutical industry encompass the limited scale of synthesis of chemically modified ON and cost-effectiveness of treatment [27-29].

Table 1: Therapeutic oligonucleotides on the market listed in chronological order from the first approval.

Active compound (trade name)	Company	Mode of action and administration	Indication	Year and market of approval
Mipomersen (Kynamro)	Ionis Pharmaceuticals Genzyme	Antisense (gapmer), subcutaneous injection	homozygous familial hypercholesterolaemia	2013, USA
Defitelio (Defibrotide)	Jazz Pharmaceuticals	Unknown, IV infusion	Hepatic veno-occlusive disease	2016, USA, EU
Eteplirsen (Exondys 51)	Sarepta Therapeutics	Antisense (SSO), IV injection	Duchenne muscular dystrophy	2016, USA
Nusinersen (Spinraza)	Ionis pharmaceuticals Biogen	Antisense (SSO), intrathecal injection	Spinal muscular atrophy	2016, USA 2017, EU
Inotersen (Tegsedi)	Ionis pharmaceuticals Akcea Therapeutics PTC Therapeutics	Antisense (gapmer), subcutaneous injection	Hereditary transthyretin amyloidosis	2018, USA, EU, Canada 2019, Brazil
Patisiran (Onpattro)	Alnylam	siRNA, IV injection	Hereditary transthyretin amyloidosis	2018, USA, EU 2019, Canada, Japan 2020, Brazil
Volanesorsen (Waylivra)	Akcea Therapeutics	Antisense (gapmer), subcutaneous injection	Familial chylomicronaemia syndrome	2019, EU
Givosiran (Givlaari)	Alnylam	siRNA, subcutaneous injection	Acute hepatic porphyria	2019, USA 2020, EU
Golodirsen (Vyondys 53)	Sarepta Therapeutics	Antisense (SSO), IV injection	Duchenne muscular dystrophy	2019, USA
Lumasiran (Oxlumo)	Alnylam	siRNA, subcutaneous injection	Primary hyperoxaluria type 1	2020, USA, EU
Viltolarsen (Viltepso)	NS Pharma	Antisense (SSO) IV injection	Duchenne muscular dystrophy	2020, USA
Inclisiran (Leqvio)	Novartis Europharm Limited	siRNA, subcutaneous injection	Primary hypercholesterolaemia, Mixed dyslipidaemia	2020, EU
Casimersen (Amondys 45)	Sarepta Therapeutics	Antisense (SSO) IV injection	Duchenne muscular dystrophy	2020, FDA

1.1 CHEMICAL MODIFICATIONS OF NUCLEIC ACIDS

1.1.1 Modifications of the phosphodiester backbone

Modifications of the phosphodiester backbone (PO), which is depicted in Figure 1, are primarily used to enhance stability towards endo- and exo-nucleases. Most commonly used is the phosphorothioate backbone (PS). Additional modifications include N3'->P5' phosphoramidates, N3'->P5' thio-phosphoramidates [30], and synthetic cationic linkages, such as aminoalkylated phosphoramidates [31].

In PS, one of the non-bridging oxygen atoms in the phosphodiester group is replaced by sulphur (Figure 1). This replacement results in two potential planar stereochemical configurations, termed *S*p or *R*p [32, 33]. Since in the conventional chemical synthesis the chirality is not controlled, it produces 2^n of different isomers, where n is the number of PS linkages. It has been shown that the stereochemistry influences the stability and efficacy of the ON. More specifically, *R*p provides higher thermal stability, while *S*p increases lipophilicity and stability against nucleases [34, 35]. The PS modification is usually placed between the first 3-5 bases at the 3'- and 5'-end of the ON to prevent degradation by exonucleases [36] or along the whole length of the ON to increase stability towards endonucleases [37]. Compared to PO, ONs carrying PS backbone have lower affinity towards RNA or DNA, however this can be counteracted with modifications on either the sugar or the nucleobase. More importantly, PS improves bioavailability and cellular uptake *in vivo* by binding to plasma-, membrane- and intracellular proteins [26, 38, 39]. Over the years, several safety concerns have been raised, since it has been observed that PS modified ONs may cause inflammation, thrombocytopenia, nephrotoxicity and hepatotoxicity [40]. This can potentially be averted, by controlling the stereochemistry or adding modifications such as 2'-*O*-methoxyethyl (2'MOE) or 2'-*O*-Methyl (2'-*O*-Me) [41-43]. Despite certain drawbacks, it is currently the most commonly used backbone modification, both in ASOs and small interfering RNAs (siRNAs), due to relatively simple synthesis and low cost. It is used in the approved medicines, Spinraza, Tegsedi, Givlaari, Onpattro and Oxlumo.

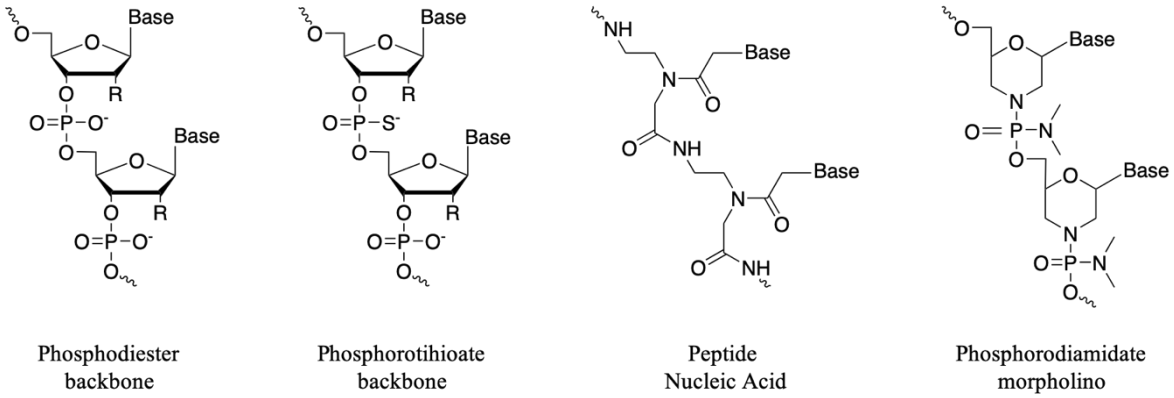

Figure 1: Chemical structure of the phosphodiester backbone and three modifications, phosphorothioate backbone, peptide nucleic acid and phosphorodiamidate morpholino.

Moreover, both the PO and the pentose can be substituted resulting in oligomers, mimicking nucleic acids. Phosphorodiamidate morpholino oligomer (PMO), is a nucleic acid mimic, where the ribose and PO backbone are replaced by a morpholine and phosphorodiamidate linkage (Figure 1) [44]. Compared to the PS backbone, the PMO is not compatible with RNase H1 cleavage but rather sterically blocks the mRNA translation [45]. The modification exhibits resistance to nucleases, esterases, and hydrolases in plasma and serum [46]. It is used in Exondys 51, Vyondys 53, Amondys 45 and Viltolarsen for the treatment of Duchenne muscular dystrophy (DMD).

Peptide nucleic acid (PNA) is an uncharged nucleic acid mimic as well. It is a combination of a pseudopeptide backbone and a nucleobase (Figure 1). It hybridizes in a sequence-specific manner, with high affinity, to both RNA and DNA via Watson-Crick (WC) base pairing [47, 48]. In low salt conditions, PNA is capable of strand displacement mode of binding to double stranded DNA (dsDNA) [47]. Furthermore, it can function as a triplex-forming ON (TFO), binding to the DNA major groove in a sequence-specific manner at polypurine•polypyrimidine (Pu•Py) stretches [49]. Due to the neutral charge of the backbone, PNA oligomers are not easily taken up by the cell and are rapidly removed from the circulation *in vivo* [50]. However, conjugation to short peptides greatly improves cellular uptake and bioavailability in cell cultures and *in vivo* [51, 52].

1.1.2 Modifications of the ribose and deoxyribose

When discussing chemical modification of ribose or deoxyribose, mostly 2’ position is altered. Modifications generally improve binding affinity towards RNA and DNA, and additionally contribute to stability [53]. They are usually combined with a backbone modification. Discussed chemical structures are depicted in Figure 2.

Substitution with 2’-fluoro (2’-F) locks the ribose in C3’ endo conformation, which increases ON:RNA duplex stability while maintaining base-pairing specificity [54]. Mixmers and gapmers support RNase H1 mediated cleavage [55] and fully modified ONs can induce steric blocking of translation and exon skipping [56, 57]. This modification is included in Lumasiran and Inclisiran which are used for the treatment of Primary hyperoxaluria type 1 and Primary hypercholesterolaemia or Mixed dyslipidaemia, respectively. 2’-*O*-Me is a naturally occurring RNA analogue. It provides stability toward nucleases, protection against general base hydrolysis [58, 59], and increases duplex melting temperature (T_m) [53]. As with 2’-F, fully modified 2’-*O*-Me inhibits RNase H1 mediated cleavage [55], however it can be incorporated in the ends of gapmers [60] and in siRNA [61], as in Onpattro, Givlaari, Lumasiran and Inclisiran. 2’-MOE is another variant of 2’ modification. It improves nuclease resistance, enhances affinity towards RNA, cellular uptake and intestinal absorption. Additionally, it reduces toxicity and immune stimulation [62, 63]. It is included in Kynamro, for the treatment of homozygous familial hypercholesterolaemia, in Tegsedi for Hereditary transthyretin amyloidosis, and in Spinraza for spinal muscular atrophy (SMA).

Locked Nucleic Acid (LNA) is a bicyclic nucleic acid analogue structurally resembling RNA. A 2'-C,4'-C-oxymethylene linker locks the sugar moiety in a C3' endo (north) conformation, resulting in a higher affinity to both RNA and DNA [64-66]. The binding affinity increases between 3-9 °C per modification, with higher melting temperature (T_m) observed in RNA duplexes [67]. LNA also provides resistance to endonucleases and 3'exonucleases, which makes it suitable for *in vivo* application [68]. Additionally, it provides mismatch discrimination and low toxicity [69, 70]. Furthermore, incorporating it into detection probes or PCR primers improves the assay sensitivity. LNA can be further modified to 2'-amino-LNA, to which additional moieties, such as norspermidine or fatty acids, can be introduced. These can contribute to higher stability of duplexes and triplexes [71, 72], or to improved biodistribution and cell uptake, respectively [73].

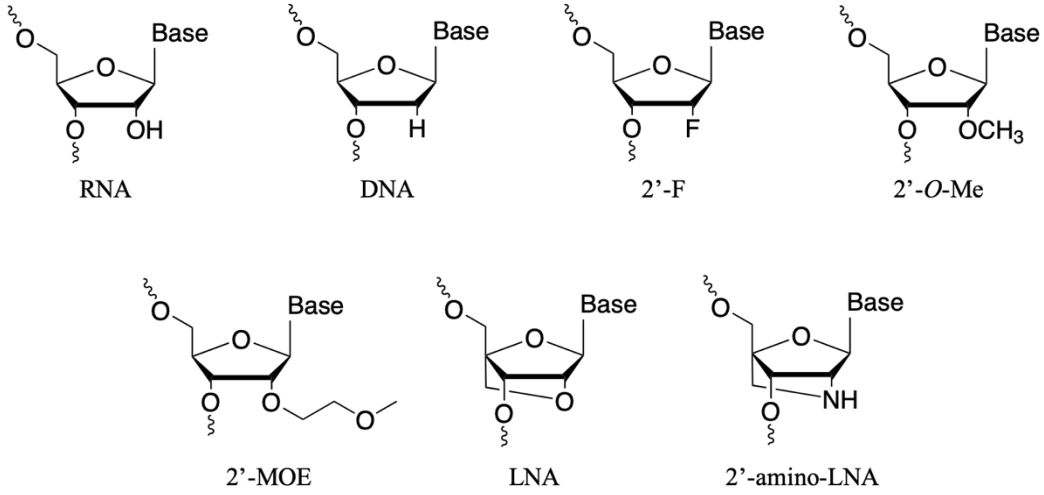

Figure 2: Chemical structure of ribose and deoxyribose and selected chemical modifications used in therapeutic oligonucleotides. 2'-fluoro (2'-F), 2'-O-Methyl (2'-O-Me), 2'-O-methoxyethyl (2'-MOE), locked nucleic acid (LNA), 2'-amino-locked nucleic acid (2'-amino-LNA).

1.1.3 Modifications of the nucleobase

Modified nucleobases first appeared within anticancer and antiviral drugs, as inhibitors of DNA replication and RNA synthesis [74, 75]. In the field of ONs, the modified bases improve duplex affinity and stability, by affecting stacking interactions or hydrogen bonds between base pairs. The prevalent site of alteration is C5 in pyrimidine with a methyl group (-CH3) applied most commonly. 5-methyl pyrimidine stacks between heterocyclic nucleobases, increasing the T_m of the duplex [76]. Additionally, it can be used to reduce the immunostimulatory effect and toxicity of the PS modified ONs [43]. 5-methylcytidine and 5-methyluridine are included in Nusinersen, Mipomersen and Inotersen. Other C5 modifications include 5-propynyl, 5-thiazolyl, 5-alkynyl, and 5-heteroaryl [77, 78]. Additionally, modifications have been reported at C2 and C4 positions of pyrimidines and C2 and C6 positions of purines [79].

1.2 RNA TARGETING OLIGONUCLEOTIDES

RNA targeting ONs include ASOs, siRNAs, microRNAs (miRNA), and DNAzymes. Each utilizes a different approach to influence gene expression either directly or indirectly. Only ASOs and RNA interference ONs (RNAi) will be discussed further, as they are currently the only ones regularly used in therapy.

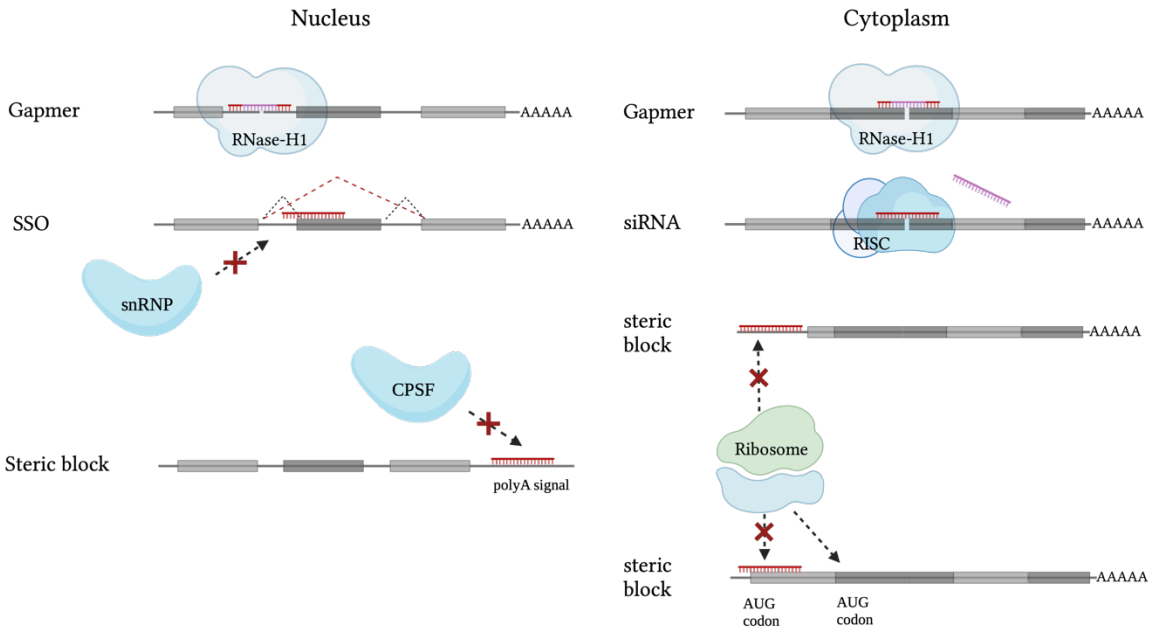

Figure 3: Mechanism of action of selected RNA targeting ONs acting in the nucleus or in the cytoplasm. ONs are coloured in red. small nuclear ribonucleoproteins (snRNP), Cleavage and polyadenylation specificity factor (CPSF), RNA-induced silencing complex (RISC) (Created with Biorender.com).

1.2.1 Antisense oligonucleotides

ASOs are single stranded chemically modified nucleic acids with a typical length between 15 to 30 nt. They bind the protein-coding mRNA, pre-mRNA or long noncoding RNA (lncRNA) via WC base pairing. Subcategories are RNase H1-dependant ASOs, and steric blockers.

1.2.1.1 RNase H1-Dependant ASO

RNase H1 belongs to the family of non-sequence-specific endonuclease enzymes. It recognizes the RNA:DNA duplex and cleaves the RNA backbone in the presence of divalent cations and protein P32 via hydrolysis [80, 81]. In eukaryotic cells, it localizes in the nucleus, cytoplasm and in mitochondria. It is responsible for the removal of initiator RNA primers of the Okazaki fragments during DNA replication and for R loop clearance that forms during RNA synthesis in the nucleus and in the mitochondria [82, 83].

The most common RNase H1 dependant ASOs are gapmers, where an 8-10 nt long DNA stretch is surrounded by 3-5 modified nts on each side. The gap supports RNase H1 mediated degradation of the target and the modified flanks contribute to increased duplex stability and resistance to cleavage by nucleases [58, 60]. The RNase H1 recognizes both mRNA- and pre-mRNA:DNA duplexes, which enables gapmer utilization for both transcripts and leads to down-regulation of target expression [84]. Successful therapeutics with the gapmer design are

Kynamro, Tegsedi and Waylivra, used for homozygous familial hypercholesterolaemia, Hereditary transthyretin amyloidosis and Familial chylomicronaemia syndrome, respectively. Common modifications are PS backbone, 2'-*O*-Me, 2'-MOE and LNA.

1.2.1.2 Steric blockers

Opposite to RNase H1 dependant ASOs, steric blockers do not necessarily induce target degradation. They are usually designed as mixmers, with chemical modifications spread across an entire sequence of the ON. By binding with high affinity and masking regulatory sequences in the pre-mRNA or mRNA they prevent regulatory proteins or ribosomes to carry out their function. Through binding to the AUG codon or the 5'cap site region, ASO inhibits ribosome access to the mRNA. This can lead to either translation downregulation or redirecting the translation machinery from an upstream open reading frame (ORF) to the primary ORF and increase of the wild-type (wt) protein [85, 86]. Increase in protein expression can additionally be achieved by targeting a negative regulatory elements in the 5' untranslated region (UTR) [87]. In the 3' end of the transcript, ASOs can be designed to target the polyadenylation (polyA) sites, which can lead to target degradation [88]. By targeting the coding region, they can be used to disrupt non-canonical RNA structures. This leads to increased protein translation or reduced RNA mediated toxicity [89, 90]. Since 1990, the steric block ONs are also frequently mentioned as potential new antiviral drugs inhibiting virus replication; however, none has made it to the late stages of a clinical trial [91-94].

Splice switching ONs (SSOs) are a subcategory of steric blockers. They bind to pre-mRNA, most commonly to donor and acceptor splice sites [95] and exonic splice enhancer motifs [96]. By restricting access to splice modulators, such as RNA-binding proteins and small nuclear RNAs [97], they can influence splicing to either exclude or include an exon [98, 99], restore the reading frame [100], switch between different mRNA isoforms [101], and even achieve gene knockdown [102]. Exondys 51, Vyondys 53 and Amondys 45 are PMO modified SSOs and are used for the treatment of DMD. Spinraza, used for Spinal muscular atrophy, has a combination of 2'MOE and PS backbone.

1.2.2 Oligonucleotides utilizing RNAi mechanism

RNAi is a cellular process where short non-coding RNA molecules hybridize to the cognate mRNA, which is then cleaved by the RNA induced silencing complex (RISC) and finally degraded [14]. The main function of the RNAi is the downregulation of endogenous gene expression and protection against viruses [103, 104]. Key molecules in the process are siRNA and miRNA, which have now been utilized by the field to regulate the expression of disease-causing genes. Both miRNA and siRNA are a 21-25 nt long double stranded RNA (dsRNA) molecules composed of a guide and a passenger strand with a 2-nt 3' overhang [105, 106]. Similar, to other ONs, unmodified siRNA and miRNA are rapidly degraded and cleared out of the circulation. Many chemical modifications have been used, with PS backbone, 2'-F and 2'-*O*-Me being the most common. Additionally, siRNA can be conjugated to N-acetylgalactosamine (GalNAc) [107] or encapsulated in lipid or polymer vesicles [108, 109] to

enhance or direct tissue distribution. siRNAs are normally designed to target one specific mRNA, while a single miRNA can affect many by targeting common, non-coding regions of the mRNA. siRNA therapeutics Onpattro, Givlaari and Inclisiran have been approved for the treatment of Hereditary transthyretin amyloidosis, Acute hepatic porphyria and Primary hypercholesterolaemia, Mixed dyslipidaemia, respectively.

1.3 ANTI-GENE OLIGONUCLEOTIDES

Compared to RNA targeting ONs, DNA targeting ONs modulate gene expression on a DNA level. Also termed anti-gene ON (AGO), they can be grouped into TFO and strand-invading ONs. In addition to possessing nuclease resistance, these ONs must withstand interference by regulatory factors and proteins, RNA polymerase, and DNA reformation, when bound to DNA. However, the desired effect can potentially be achieved with a lower number of molecules as compared to RNA targeting ONs. Currently, there are no published pre-clinical or clinical studies involving DNA targeting ONs.

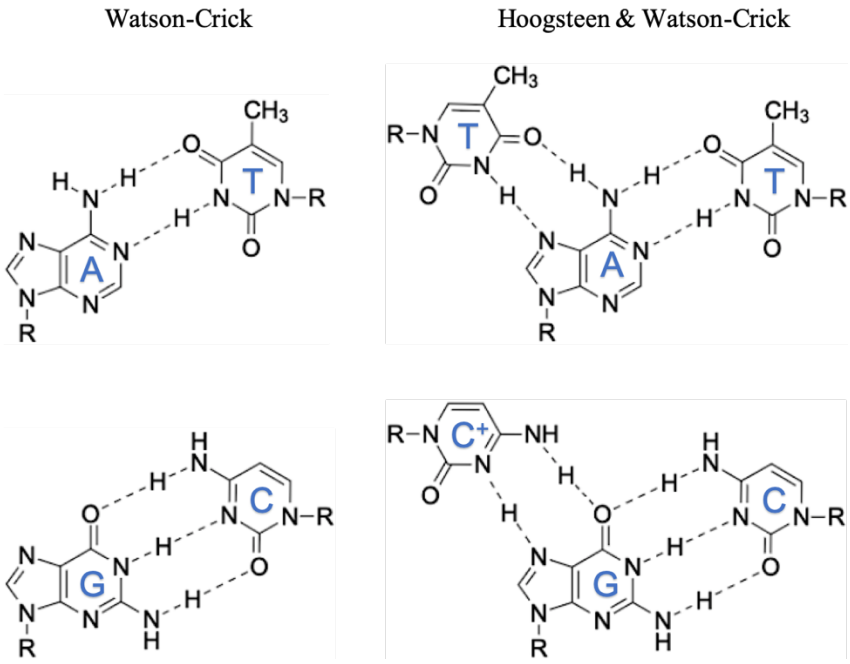

Figure 4: Watson-Crick base pairing and combined Hoogsteen and Watson-Crick base pairing. Two modes of interaction used by anti-gene oligonucleotides.

1.3.1 Triplex-forming oligonucleotides

TFOs bind to the major groove of the DNA double-helix in a sequence-specific manner. They hybridize to the purine strand of 10-30 nt long Pu•Py stretches, via Hoogsteen or via reverse Hoogsteen base pairing [110-112]. Targets include regions upstream of gene promoters, transcription start sites, and the open reading frame [113-115] . Bound to DNA they prevent the access of transcription factors to genomic DNA [116] or stall the RNA polymerization [117], reducing the level of gene expression. Induction and reactivation of gene expression have also been reported [118, 119]. Additionally, binding of modified TFOs can induce mutagenesis [120] or recruit DNA repair enzymes, stimulating either recombination or gene conversion [121, 122]. The majority of TFOs have been designed to target oncogenes, such as *MYC* [113, 114] and *HMGB1* [123]. Initially, one of the common modifications used was PNA, however due to already mentioned disadvantages new modifications needed to be tried. LNA is one chemical modification that enables TFO binding at physiological conditions. In recent years, new nucleotide analogues have been developed to enhance the binding to dsDNA, however only some have shown an effect on gene expression [124-126].

1.3.2 Strand-invading oligonucleotides

Strand-invading ONs disrupt the hydrogen bonds of the double-helix and hybridize via WC base-pairing leading to duplex formation and a displaced strand, as depicted in Figure 5. ONs can bind only through WC interaction to one [127] or both strands [128-130], or in the form of clamps bind to one strand with both WC and Hoogsteen bonds (Figure 4) [131-133]. PNA oligomers were developed first and possessed higher stability and affinity than unmodified ONs [133, 134]. Compared to PNA, LNA/DNA mixmers with a PS backbone, invade dsDNA more efficiently under physiological conditions. They were successfully used *in vitro* in the context of trinucleotide repeat disorders for the downregulation of *Huntingtin* (*HTT*) expression [127] and destabilization of non-B-DNA structures in *Frataxin* (*FXN*) [135].

The efficiency of strand-invasion under physiological conditions was improved by bisPNAs – clamps [49, 136] and tail-clamps [137]. They bind to the DNA first with the Hoogsteen binding strand and later strand-invade with the WC strand [138]. This was later supported by an LNA modified tail-clamp (bisLNA) [132]. BisPNAs have been used in genetic editing, with low off-target effects (reviewed in Economos *et al.*[139]).

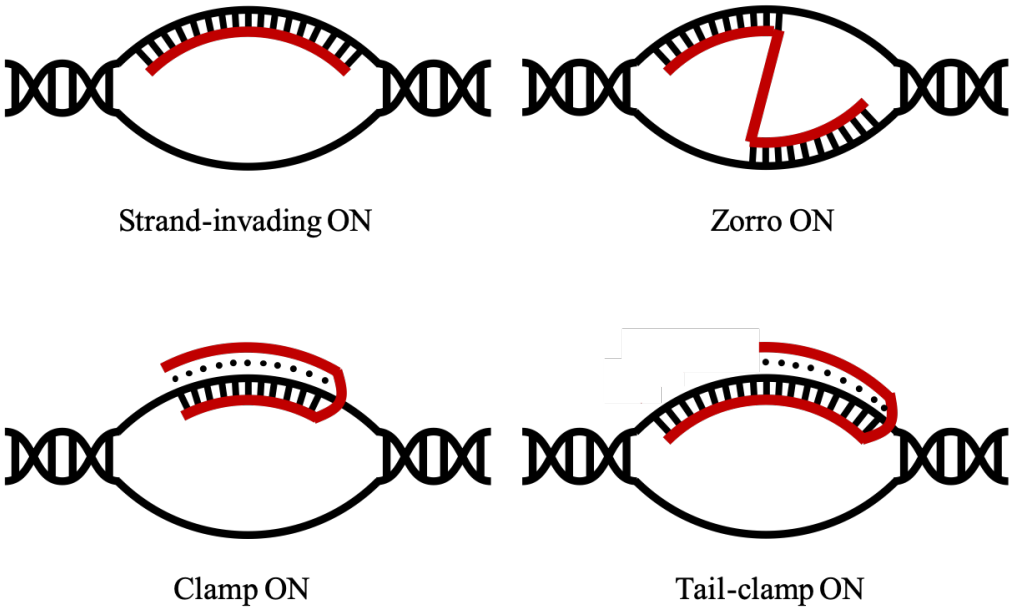

Figure 5: Types of strand-invading oligonucleotides. The AGO disrupts the H-Bonds between the two strands of the DNA duplex and binds via WC base pairing (lines) only or additionally with Hoogsteen base pairing (dots).

Under physiological pH and salt conditions, bisLNAs recognize Pu•Py sequences with high specificity. By the combination of the two modes of binding mentioned before, they form stable triplex structures able to withstand DNA relaxation. The efficiency of strand-invasion is influenced by the length of the TFO arm, the tail – duplex forming extension in the WC arm, and the number and position of the LNA nucleotides. Furthermore, a 5-6 nt long linker, which is non-complementary to the target, allows favourable stacking and base-pairing interactions

of the strand-invading WC arm [131]. Next-generation ONs were synthesized with linkers able to stack to adjacent bases, and with 2-glycylamino-LNA modifications in the arms to improve strand-invasion [132].

LNA has also been used in zorro-LNA, which hybridizes to both strands of the DNA via WC interaction. It can be synthesized as two separate arms connected via a linker [130] or as single stranded zorro-LNA synthesized in 3’5’-5’3’ orientation [128, 129].

1.4 NON-B-DNA STRUCTURES

The canonical secondary structure of the dsDNA is the right-handed B-DNA double helix imaged by Rosalind Franklin and described by James Watson and Francis Crick [5]. The average helix is 2 nm wide, with the pitch of 3.57 nm, and a minor and a major groove repeating every 10.5 base pair (bp) or one helical turn. The shape is a result of the H-bonds between A:T and G:C of the two antiparallel nucleic acids, hydrophilic backbone (sugar and phosphate) and hydrophobic bases, which stack and twist to maximise the hydrophobic environment. Right-handed A-DNA and left-handed Z-DNA have been described as well. The A-DNA is formed under dehydrating conditions [140]. The formation of Z-DNA is sequence dependant, forming within alternating purines and pyrimidines [141]. These two structures are not the only non-canonical conformations of the DNA, jointly termed the non-B-DNA structures [142]. Since the 1950's scientists have determined the presence of a hairpin, a cruciform, a triple helix (also termed H-DNA), a G-quadruplex, an I-motif and a DNA unwinding element to name a few.

The regions of genomic DNA that can fold into the non-B-DNA conformations are characterized by a non-random distribution of nucleotides, such as direct, mirror or inverted repeats. These sequences are generally present in gene regulatory regions, such as promoters, untranslated regions, origins of replications and telomeres [143, 144]. *In vivo*, the formation depends on the localized negative supercoiling or on the competing non-B-DNA structures [145, 146]. These structures generally occur during transcription, replication, repair or recombination and influence these biological processes [147-149]. The role of these motifs has also been associated with telomere maintenance and hypomethylation of CpG islands [150, 151]. Despite being functional genomic elements, they have been implicated in causing genetic instability and leading to diseases, such as cancer and repeat expansion disorders [152, 153]. The formation of the non-B-DNA structure during normal biological processes can cause DNA polymerase stalling during replication and ultimately lead to double strand breaks and chromosomal rearrangements [154]. Additionally, transcription-induced structures that generally form on the exposed ssDNA coding strand, can attract DNA repair machinery which potentially results in genetic instability [155, 156].

1.4.1 H-DNA

Intramolecular triple helix or H-DNA forms by Pu•Py mirror repeat sequences [157, 158]. The structure is characterized by the separation of the dsDNA in one part, looping of the ssDNA and hybridization of the ssDNA via hydrogen bonds in a sequence specific manner to the major groove of the dsDNA (Figure 6). When the third strand is composed by the pyrimidines, it binds to purines (T•A:T or C*•G:C) in parallel orientation and the base pairing is Hoogsteen [159]. When the third strand is purine rich it also hybridizes to the purine strand in the dsDNA (A•AT or G•GC), however in the antiparallel orientation, therefore the base pairing is reverse Hoogsteen. T•A:T, A•A:T and G•G:C form at physiological pH. The C*•G:C formation is favoured in acidic conditions as the triplex forming C needs to be protonated. *In vivo* the formation of the H-DNA is associated with negative supercoiling following replication and transcription [160]. Additionally, the stability of the triplex depends on the presence of mono-

and divalent cations, the amount of G:C in the repeat sequence, the sequence of the flanking region, sequence similarity of the mirror repeat and the length of the Pu•Py sequence [161-163].

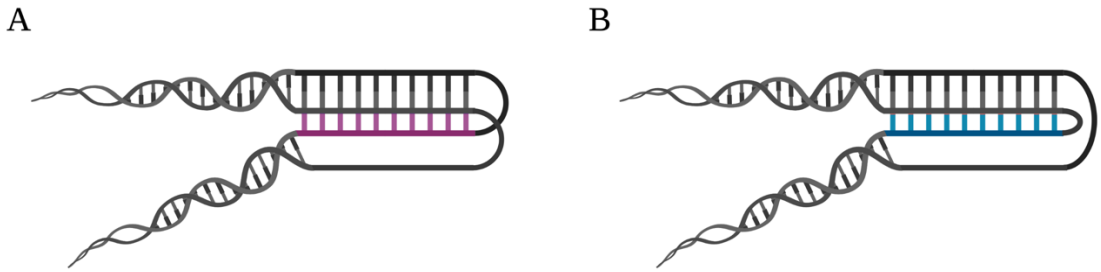

Figure 6. Illustration of A) a parallel and B) an antiparallel H-DNA. In parallel H-DNA the third strand binds with Hoogsteen base pairing and has the same 5'-3'orientation as the purine strand. In antiparallel H-DNA the third strand binds with reverse Hoogsteen base pairing and has the opposite 5'-3' orientation as the purine strand (created with Biorender.com).

In the mammalian genome, mirror repeat sequences are abundant. They are mostly present in the promoters and exons which indicates their role as *cis*-acting regulatory elements of transcription [164-166]. These sequences have also been associated with translocation and deletion hot spots and are connected to many cancers [152, 167]. Furthermore, in repeat expansion diseases, they have been associated with exacerbating the mutation and lowering the transcription rate [168, 169].

1.4.2 G-quadruplex

Intramolecular G-quadruplex is formed by a ssDNA with a sequence of at least four separate G stretches [170]. The structure is characterized by the stacking of two or more planar G quartets, where each G quartet is a plane of four G bases connected by Hoogsteen base pairing. It is further stabilized by various metal ions that fit within the anionic space of the structure, preferentially K^+ *in vivo*. Based on the 5'-3' orientation of the connected guanines the structure adopts either the parallel, antiparallel or hybrid topology [171-173]. In the parallel orientation all connecting strands have the same 5'-3' orientation. In the antiparallel topology, the 5'-3' orientation of the connecting strands changes and the hybrid topology contains both parallel and antiparallel orientation of the connecting strands as shown in Figure 7.

Different bioinformatic and sequencing studies of the human genome found between 376,000 and 700,000 of potential G-quadruplex forming sequences, and immunoprecipitation techniques revealed 10,000 that can form under cellular conditions [174-177]. In eukaryotic genomes, they are enriched in telomeres, promoters, first introns and 5'UTRs [143]. For example, over 40 % of the human gene promoters have at least one such motif [178]. They were also found in origins of replication in human genome [179]. Furthermore, these sequences are generally absent in the coding strands of exons [174]. The combination of these discoveries implies that G-quadruplex plays a role in regulation of replication, transcription and telomere

maintenance [151, 180-182]. Especially, since it has been detected in cell models predominantly during the S phase (the phase of DNA replication) and in transcriptionally active euchromatin [177, 183]. Like other non-B-DNAs, the G-quadruplex is also connected to cancer. It has been shown that the DNA breakpoints and these motifs are associated and the number of G-quadruplex positive nuclei is higher in liver and stomach cancer cells compared to healthy cells [184, 185].

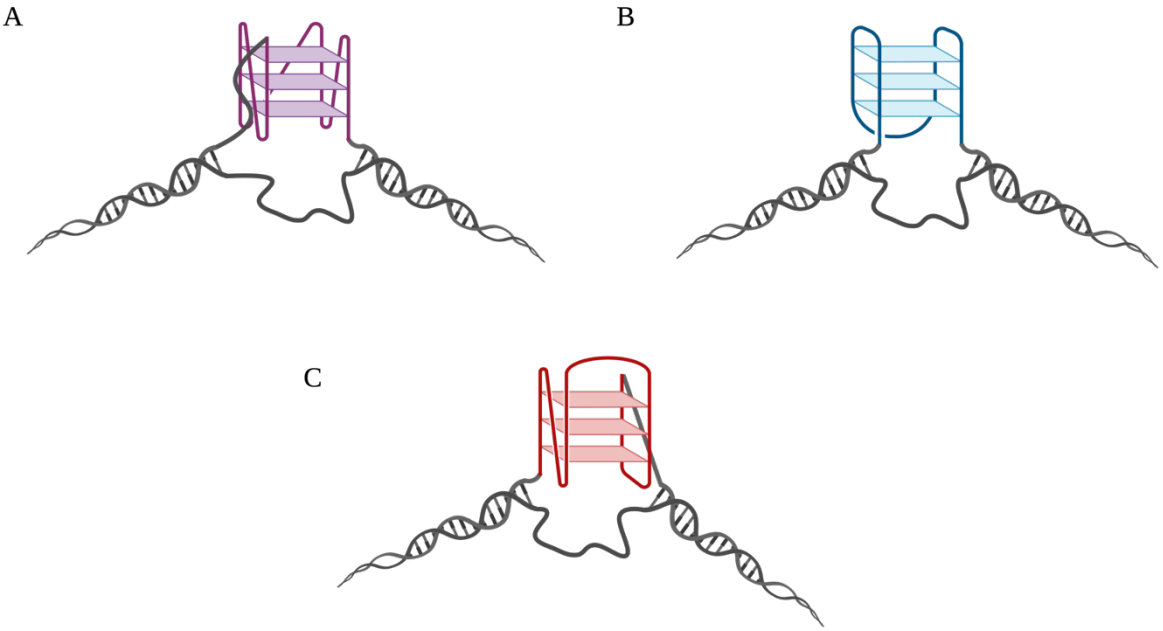

Figure 7. Illustrations of A) a parallel G-quadruplex, B) an antiparallel G-quadruplex and C) a hybrid G-quadruplex. In a parallel G-quadruplex all connecting strands have the same 5'-3' orientation. In the antiparallel G-quadruplex the 5'-3' orientation of the connecting strands changes. The hybrid G-quadruplex has both parallel and antiparallel 5'-3' orientation. (created with Biorender.com).

1.4.3 Hairpin and cruciform

Hairpins or hairpin motif-containing structures are formed in DNA with a sequence of direct repeats, inverted repeats or with a palindromic sequence [186, 187]. The ssDNA self-folds into a structure with a WC base-paired stem and a single stranded loop of varying size [188]. If two hairpins are formed opposite each other on both strands of the DNA with a four-way junction in between, this structure is then called the cruciform [189]. As with other structures, the hairpin is formed when the DNA duplex is unwound during replication or transcription [190, 191]. Under negative supercoiling the minimal number of base pairs to form a stem is seven [192]. The role of the hairpin and cruciform *in vivo* is connected to regulation of DNA replication and transcription [193-197]. Cruciform structures are specifically recognized by architectural, regulatory and DNA binding proteins [198].

Longer inverted repeats with a spacer are most abundant on the human X chromosome, mostly in genes expressed in testes [199]. Additionally, inverted repeats longer than 7 bp are non-

randomly distributed in promoters upstream from transcription start sites [200]. Palindromes are present in the entire genome, with an overrepresentation in introns and upstream regions, while they are underrepresented in exons [201, 202]. Direct tandem repeats are present in centromeres and functional regions of genes [203].

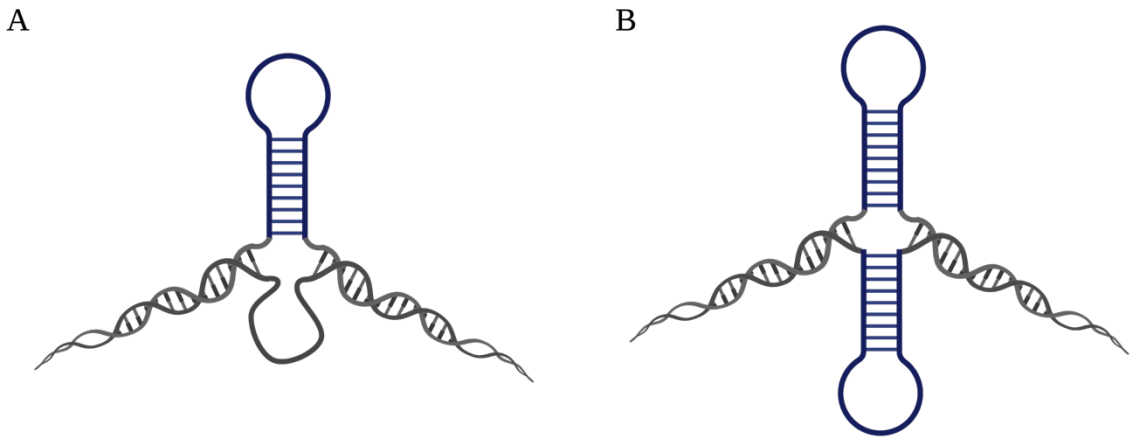

Figure 8. Illustrations of A) a hairpin and B) a cruciform (created with Biorender.com).

Inverted repeats and direct repeats are also enriched at hotspots of chromosomal rearrangements, deletions, recombination and gene amplifications [204-206]. Hairpin formation leads to replication stalling and double strand breaks [207]. The structure characteristics that increase genomic instability are short spacer, long stem length and high complementarity of the stem forming sequence [208]. In direct tandem triplet repeat sequences, the most stable hairpin is formed in the order CGG, CTG, CAG and least by CCG. The stability is influenced by mismatches in the stem and prevalence of G:C base pairing. It has been shown that tandem repeats associated with certain pathologies have a higher copy number variability among individuals [209].

1.5 TRINUCLEOTIDE REPEAT DISORDERS

It is estimated that tandem repeats cover approximately 6 % of the human genome, with the di-nucleotide motif being the most prevalent. They are present in both coding and non-coding regions, mostly in upstream and 5'UTR regions [210]. Tandem repeats are the biggest source of *de novo* mutations in humans, with the highest variability observed in disease causing genes [209, 211]. It has been suggested that their variability in copy numbers plays a role in morphological variation in dogs, in phenotype alterations of *Saccharomyces cerevisiae*, and in human traits, such as height, body mass index and white blood count, as well as type 1 diabetes, multiple sclerosis, and cancer [212-214].

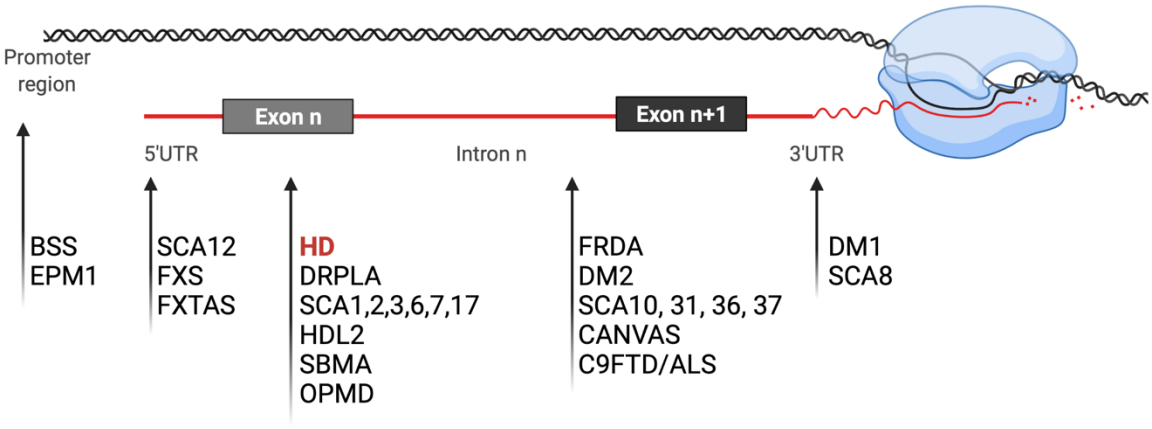

Figure 9. Selected expansion repeat disorders and position of the repeats in the coding or non-coding region. Bernard-Soulier syndrome (BSS), Progressive myoclonic epilepsy type 1 (EPM1), Spinocerebellar ataxia (SCA), Fragile X syndrome (FXS), Fragile X-associated tremor/ataxia syndrome (FXTAS), Huntington's disease (HD), Dentatorubral-pallidoluysian atrophy (DRPLA), Huntington disease-like 2 (HDL2), Spinal-bulbar muscular atrophy (SBMA), Oculopharyngeal muscular dystrophy (OPMD), Friedreich's ataxia (FRDA), Myotonic dystrophy type 1 and 2 (DM1 & 2), Cerebellar ataxia, neuropathy, and vestibular areflexia syndrome (CANVAS), C9orf72-associated diseases with clinical features of FTD, ALS or both (C9FTD/ALS) (created with Biorender.com).

Expansion repeat disorders include all genetic diseases caused by expansion of tandem repeats above a specific threshold, which varies between diseases. The number of the repeats is characterized by healthy, premutation or mutation. The disease associated repeats occur in promoter, 5'UTR, exon, intron or 3'UTR as shown in Figure 9. The expanded mutation can lead to epigenetic dysregulation and loss of gene expression, toxic gain of RNA function, repeat associated non-AUG (RAN) translated proteotoxicity, toxic gain of protein function, and change in protein function [169, 215-218]. Most known and frequent are trinucleotide repeat disorders like Huntington's disease (HD), Myotonic dystrophy type 1 (DM1), and Friedreich's ataxia (FRDA), although diseases resulting from the expansion of tetra-, penta- and hexa-nucleotide tandem repeats exist as well [219-222].

The expanded number of the repeats a person has does not only mean the penetrance of the disease but influences the disease phenotype as well. Higher number of the repeats leads to earlier onset and increased disease severity [223, 224]. Since changes occur in germ cells these

diseases do not fall under mendelian pattern of inheritance but rather under anticipation, where each progeny carrying the mutation most likely experiences greater expansion, causing worse disease prognosis [224, 225]. Disease onset and instability of the repeat length is also negatively correlated to interruptions in the repeated sequence [226-229]. Additional markers that affect disease progression, so called genetic modifiers, are genes involved in DNA metabolism, such as damage response. These are most likely responsible for the expansion in somatic cells [230].

1.5.1 Huntington's Disease

HD is an autosomal dominant, neurodegenerative disorder caused by the CAG•CTG expansion in exon 1 of the *HTT* gene [231]. Consequent expression of both, toxic gain-of-function mRNA and protein, results in neuronal death mainly in striatum and leads to mobility, psychological and cognitive impairment [232]. Cognitive and behavioural issues usually arise approximately 15 years prior the decrease of motor function [233]. Recently, it has been shown that mutant HTT (mtHTT) protein influences the neurodevelopment of the cortex as well [234]. Additionally, small RNAs, specifically sCAG produced by DICER cleavage of the self-folded mRNA, and exon 1 peptide translated from aberrantly spliced pre-mRNA contribute to neuronal death [235-238]. RAN translated proteins with poly-alanine, -serine, -leucine and -cysteine, have also been discovered in HD patient brains [239]. They appear to be non-toxic in knock-in mouse models and therefore additional studies need to be performed to evaluate their effect in the human brain [240].

HD symptoms manifest if the tandem repeat expands to, or above, 40. In healthy people these repeats span between 6-35. Repeats up to 26 are considered normal. Between 27-35 they are unstable, and offspring have an increased risk of acquiring the disease. 36-39 are characterized with incomplete penetrance, meaning the person might or might not develop the disease [241, 242]. Although the disease can be inherited from both the mother and the father, larger intergenerational changes are mostly observed in paternal transmission [243]. The symptoms appear in young adulthood or middle age, except in juvenile HD.

There are two key moments in the progression of the HD. First, where the DNA repeat number in a patient has exceeded the threshold of repeat stability, hence start to expand in somatic cells and second, where the repeats expand to the extent that they become pathologic [244]. The major contributor to the age at disease onset is the uninterrupted repeat size of the longer allele [226, 245]. Additionally, genome wide association studies (GWAS) helped identify several genetic modifiers involved in DNA repair, such as FAN1, which stabilizes repeat expansion, or MLH1 and MSH3, which contribute to expansion [246-249]. It is becoming more evident that the repeat expansion occurs in germ cells during DNA replication and in somatic cells during RNA transcription. During DNA replication the repeats expand due to slipped strand formation on the nascent strand and misalignment, or fork stalling and reversal [191]. In somatic cells non-perfect hairpins, can form by dsDNA looping out or by self-folding of the ssDNA during RNA transcription. Hairpins are resolved by the DNA mismatch repair enzyme

MutSβ and endonuclease complex MutLα. The cut-out gap is filled by DNA polymerase and DNA ligase LIG1, which can lead to expansion [250].

There is currently no cure for HD and patients have to rely on supportive therapy until death that usually occurs 15 to 20 years post-diagnosis. *In vivo* studies on animal models have shown that some reduction of wtHTT is safe [251, 252]. Nevertheless, complete knock-out affects synaptic connectivity, gene regulation, and survival of striatal projection neurons, which results in similar symptoms as HD and this risk should be considered in new therapies [253]. This could be due to wtHTT's role in intracellular trafficking, regulation of transcription and synaptic connectivity [254-256]. Several different therapeutic methods are being developed at the moment in pre-clinical and clinical studies, such as CRISPR-Cas9, TALEN, Zinc Finger Protein and Nucleases, DNA repair modulators, small molecules and therapeutic ONs [257-261].

Therapeutic ONs target mt*HTT* and wt*HTT* mRNA. Until March 2021 clinical trials were being carried out by Ionis pharmaceuticals/Roche (Tominersen or IONIS-HTTRx or RG6042) and by Wave Life Sciences (WVE-120101 or Precision HD1, WVE-120101 or Precision HD2). Both were halted due to therapeutic ineffectiveness or worsening of the disease progression [262]. UniQure Biopharma (AMT-130) still continues its phase I/II [263]. Tominersen, Precision HD1, and HD2 are gapmer ASOs. Tominersen is targeting both mt and wt pre-mRNA and/or mRNA [264]. Precision HD1 and HD2 are designed for patients carrying a single nucleotide polymorphism, making the ONs mutant specific. AMT-130 is an adeno-associated virus serotype 5 (AAV5) carrying an artificial miRNA that targets mt and wt *HTT* mRNA [265, 266].

Contrary to RNA-targeting ONs, AGOs are still in the early phases of development. However, the advantage of using the anti-gene strategy is especially evident in diseases, where the toxic effect of the transcribed mutated gene is not only caused by a malformed protein but also by pre-mRNA and mRNA [215, 267]. *In vitro* evidence showed that direct targeting of expanded CAG•CTG trinucleotide repeats in DNA significantly reduces both mRNA and protein levels in HD patient fibroblasts. Additionally, lower phosphorylation of *HTT* gene-associated RNA-polymerase II suggests reduced transcription downstream of the ON-targeted repeat [268].

2 RESEARCH AIMS

The overall aim of this thesis was to evaluate the design and the therapeutic effect of AGOs.

More specifically in:

Paper I

The aim was to evaluate the effect of the *HTT* gene targeting PS modified LNA/DNA mixmer strand-invading AGO on the expression of *HTT* mRNA and protein during HD induced Pluripotent Stem Cells (iPSCs) *in vitro* neurogenesis.

Paper II

The aim was to optimize the design of the LNA/DNA mixmer clamp-type AGO using an M3 intercalator and improve its strand-invading efficiency. Furthermore, the goal was to evaluate the AGO's ability to affect the expression of the targeted gene.

Paper III

The aim was to analyze the non-B-DNA forming sequence present in the *MYC* promoter region *in vitro* and evaluate its effect on the strand-invasion efficiency of the LNA/DNA mixmer tail-clamp AGO.

3 MATERIALS AND METHODS

The following materials and methods were used to evaluate and analyze the AGO binding to DNA *in vitro* [269], the RNA transcription alteration and protein levels in disease cell models.

3.1 ON BINDING TO dsDNA

3.1.1 S1 nuclease assay

This assay was used to determine the dsDNA strand-invasion efficiency of the AGO, such as clamps (Paper II) and tail-clamps (Paper III). The hybridization reactions of an AGO in concentration raging from 16 nM to 10 μM and 1 μg of plasmid were carried out in 10 μl total volume at 37 °C in intranuclear pH and salt conditions (50 mM Tris-acetate [pH 7.4], 120 mM KCl, 5 mM NaCl, 0.5 mM $Mg(OAc)_2$) for 24 h to 72 h. Strand-invasion and binding of the AGO to one strand of the DNA leaves an opposite exposed single strand, the displaced loop (D-loop). This is recognized by the S1 nuclease which cleaves the phosphodiester backbone at the exposed region [270]. For the digestion, which is performed on ice, 2.5 μl of the hybridized sample is cut with 24 U of the enzyme. The enzyme is inactivated by the addition of 3 mM EDTA after 6 min. The result is a nicked or a linear plasmid, which can be easily distinguished from the supercoiled plasmid by the migration pattern in the agarose gel electrophoresis. Upon imaging of the agarose gel and signal intensity quantification, the percentage of the strand-invasion is calculated from the ratio of the linear plus nicked plasmid to the supercoiled plasmid. Since some non-specific nicking can occur also in non-invaded plasmid, a mock treated plasmid is used for normalization.

3.1.2 Electrophoretic Mobility Shift Assay (EMSA)

In paper III EMSA was used to determine the triplex formation between the dsDNA target and Pu or Py rich TFO. ^{32}P labelled target was incubated with several concentrations of unlabeled TFOs in ratio from 1:10 to 1:1000 for 19 h at 4 °C. The conditions for Py TFO were adjusted to pH 6.5 and Na^+ to 100 mM. The pH 7.4 and 10 mM Mg^{2+} was used for hybridization of Pu TFO. The samples were separated on a non-denaturing polyacrylamide gel electrophoresis (PAGE), which allows separation of hybridized products. PAGE separation of the duplex and triplex, formed upon the Hoogsteen binding of the TFO, occurs due to slower migration of the triplex. The change of triplex mobility if affected by the change in molecular mass and conformational changes. The ^{32}P signal was detected using phosphorimager.

3.2 BQQ-OP DETECTION OF H-DNA

Benzoquinoquinoxaline (BQQ) is a heterocyclic compound with an aminoalkyl side chain, which intercalates between bases in a DNA triplex and stabilizes its conformation [271]. Coupling of BQQ to 1,10-phenanthroline (ortho-phenanthroline [OP]) results in a triplex-specific cleaving agent, termed BQQ-OP [272]. BQQ-OP binds the triplex and cleaves the dsDNA in the presence of Cu^{2+} ions, which chelate to OP, and a reducing agent leading to *in*

situ production of radicals. These non-diffusible radicals in the immediate vicinity of BQQ-OP intercalating site ultimately cause a break of the DNA backbone.

In Paper III BQQ-OP cleavage was used to determine the presence of the H-DNA structure formed by the sequence of the nuclease hypersensitivity element III_1 ($NHEIII_1$), which is cloned into a supercoiled plasmid. After the BQQ-OP cleavage reaction, when the plasmid is linearized, a restriction enzyme was used to cut the plasmid. This results in two distinct fragments which can be easily separated using agarose gel electrophoresis.

3.3 CELL CULTURE AND ASSAYS

The therapeutic potential of AGOs presented in this thesis was evaluated using disease cell models. This type of evaluation is generally a first step in drug discovery, used for screening a multitude of potential leads. Besides the on-target effect of the tested ON, off-target, cell uptake and toxicity can be approximated as well.

In Paper I, we used iPSC transformed from HD and WT human primary fibroblasts, with 109 or 17/18 CAG•CTG repeats, respectively. The cells were obtained from the National Institute of Neurological Disorders and Stroke (NINDS) Human Cell and Data Repository at the Coriell Institute for Medical Research and the NINDS Human Cell and Data Repository at RUCDR Infinite Biologics. The iPSCs were differentiated towards neuronal lineage by dual SMAD inhibition combined with retinoid signaling [273]. The neuronal stem cells (NSC) were further differentiated to neurons using BrainPhys™ Neuronal Media supplemented with brain-derived neurotrophic factor (BDNF) and human recombinant glial-derived neurotrophic factor (GDNF). The use of iPSCs and directed differentiation towards desired lineage allows investigation of the effect in the cells which would be targeted *in vivo*. This enables more accurate assessment than a study done in immortalized cell cultures. Additionally, these disease models to some extent allow replacement of animal models, which is desirable from an ethical standpoint [274].

HD human primary fibroblasts harboring 68 and 17 CAG•CTG repeats in the *HTT* gene, were used in paper II. The cells were obtained from the Coriell Institute for Medical research. These fibroblasts grow in a monolayer and were maintained in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10 % Fetal Bovine Serum (FBS).

3.3.1 Transfection

The most common method to efficiently deliver nucleic acids in the cells are non-viral transfection reagents. The transfection efficiency depends on the type of the cell and cell density, the cell culture media, and ratio between the nucleic acid and the reagent.

In paper II we used lipofection. In this technique positively charged lipids surround negatively charged AGOs, forming lipoplexes that enter via endocytosis. To form stable complexes, AGOs were incubated together with Lipofectamine RNAiMAX (Thermo Fisher Scientific) in serum free media for 20 min at room temperature. The cells were transfected 24 h after plating

when the confluence reached approximately 80 %. The treatment lasted for 4 to 10 days without changing of the media.

Magnetofection was used in Paper I to deliver the AGO to NSC and early neurons. These biocompatible nanoparticles are generally used for the delivery of the genetic material into hard-to-transfect cells. Similar to lipofection, magnetic nanoparticles and AGOs are first incubated together in serum free media to form stable complexes. These are then added to the cell culture 24 h after the last media change in a drop by drop manner to allow equal distribution. The cells are then put on a magnetic plate for 20 min. This forces the nanoparticles into the cell, which allows that the complexes are simultaneously associated with the cultured cells within minutes. The treatment was terminated after 2 days.

3.3.2 Gymnosis

Gymnosis is a method where the ON is added to the cell culture and taken up by the cell without the aid of a transfection reagent [275]. The naked ON enters the cell via endocytosis, where it then has to escape the endosome to reach the target in the cytoplasm or the nucleus. The efficiency of the cell entry depends on the cell type, density and growth, and on chemical modifications of the ON. Compared to the use of transfection reagents, this delivery mode demands much higher concentration of the ON to elicit any detectable on-target response. However, the cells experience less toxicity and it resembles the *in vivo* conditions more closely, where the treatment does not include any delivery vectors. Since this is true for the treatment through the intrathecal injections, for example for SMA and HD, this method was chosen in Paper I where we evaluated the AGO for its *HTT* downregulating capabilities.

The cell culture protocol for maintaining iPSC and differentiating them first into NSC and then into neurons requires daily cell culture media changes. Since naked ON requires more time to enter the cells, we maintained a 2 μM concentration of the AGO in the media by adding it every day.

3.3.3 Cell viability assay

Cell proliferation reagent WST-1 (Roche) can be used for measurement of cell proliferation and cytotoxicity analysis of tested compounds. Cytotoxic compounds elicit changes in structures and processes needed for cell homeostasis. Therefore, it must be confirmed that the gene of interest (GOI) expression alteration by an AGO comes from interaction with the target and not due to toxicity. WST-1 is a tetrazolium salt that is added to the cell culture media and converted to formazan by mitochondrial dehydrogenases, whose activity correlates with an expansion of viable cells. Formazan has a dark red color with an absorbance between 420 to 480 nm that can be measured by spectrophotometer. To ensure accuracy, the measurement includes a blank, a reference wavelength above 600 nm and three to four technical replicates.

In Paper II, the viability was measured after 4 days long treatment with the clamp AGOs. The absorbance was measured at 450 and 650 nm. The average of three technical measurements was normalized to non-treated cells (NT), for which the viability was set to 100%.

3.4 GENE EXPRESSION AND PROTEIN ANALYSIS

3.4.1 RT-qPCR

Polymerase chain reaction (PCR) is a method where a single copy of either DNA or RNA can be amplified to the power of n, generating billions of copies (n is a number of amplification cycles). This is achieved through design of primers spanning a desired region, a DNA polymerase, and repeating temperature steps allowing DNA denaturation, primer annealing and DNA polymerization. Since DNA polymerase cannot run on RNA, it must be first reverse transcribed (RT) to DNA by reverse transcriptase. In quantitative PCR (qPCR) an intercalating dye or a probe with conjugated quencher and fluorophore is included in the reaction. The increasing fluorescence signal generated from the intercalation or fluorophore release can be measured in real time by the real-time PCR system. In RT-qPCR gene expression analysis, primers and probe are designed for the GOI and a housekeeping control, which allows relative calculation through the ΔΔCt method. Optimization of the reaction conditions, such as amount of starting material, primer and probe concentration, and annealing temperature, allows that the amplification efficiency is between 90 – 110 %. Purity and integrity of the starting material can influence the reaction and should be tested. Salt and protein contamination can be determined by the absorbance ratios at 260/230 and 260/280, respectively. RNA integrity can be verified by using agarose gel electrophoresis.

RT-qPCR was used for expression analysis of *HTT* mRNA and off-target transcripts in Paper I and Paper II. Details of the reaction conditions are described in the Material and Methods section of both papers.

3.4.2 Western blot

Western blotting (WB) is used for protein analysis. Proteins extracted from treated cells are separated by denaturing PAGE. The percentage gradient of polyacrylamide should be selected based on the size of the investigated protein, with higher percentages used for smaller proteins and lower for larger proteins. After the completed run, the proteins are electro-transferred to a nitrocellulose or polyvinylidene difluoride (PVDF) membrane and are detected by the combination of a primary antibody (Ab) directed towards the specific protein and a secondary Ab carrying the fluorophore, which allows imaging.

The WB was used for the assessment of HTT in Paper I and II, and Cleaved caspase 3 in Paper I. To semi quantify the protein levels and evaluate the effect of the AGO treatment, the housekeeping protein was used for normalization.

4 RESULTS AND DISCUSSION

4.1 PAPER I

Oligonucleotides targeting DNA repeats down-regulate *huntingtin* gene expression in Huntington's patient-derived neural model system

The focus of this paper was to evaluate the transcription downregulating effect of a strand-invading AGO targeting the repeat region of the *HTT* gene during *in vitro* neural development. We treated the HD cell model from the iPSC stage through the differentiation to embryoid body (EB), neural rosette structures, NSC and all the way to neurons. Furthermore, we studied off-targeting of the AGO and the effect of the AGO on the process of differentiation, such as formation of the rosette structures and self-renewal of the NSC.

For this study, we used iPSCs and differentiated them first into NSC and then into neurons. We first determined the presence of pluripotent markers Stage-specific embryonic antigen 4 (SSEA4) and octamer-binding transcription factor 4 (OCT4) at the iPSC stage using immunofluorescence analysis and confocal microscopy, and RT-qPCR. To confirm that these cells differentiate into the neuronal lineage, we monitored the mRNA levels of both *OCT4* and *Paired box protein Pax-6* (*PAX6*) mRNA at three distinct differentiation stages: EB formation, rosette structures and presence of NSC. By determining the decreased mRNA levels of *OCT4* and rising of *PAX6* we confirmed that cells are losing pluripotency and are differentiating into the neuronal lineage. Additionally, we used confocal microscopy to confirm the morphology of the rosette structures and the presence of neuron-specific β-III tubulin in the NSC.

For the treatment of these cells we used a 19 nt long LNA/DNA mixmer with a PS backbone, named CAG19, which was previously shown in HD primary patient fibroblasts to elicit strong downregulation of both *HTT* mRNA and protein [127]. CAG19 is designed to target the template strand of the *HTT* gene repeat region. Since this AGO can potentially be used as a therapeutic, we used gymnosis in order to mimic the treatment *in vivo*. During a 22 days long continuous treatment, we evaluated the *HTT* mRNA levels at the end of neural induction of EBs, in prominent neural rosette structures, and in NSCs and early neurons. We confirmed that the CAG19 downregulates the *HTT* mRNA up to 50% during all stages compared to two repeat scrambled controls – SCR14 and SCR19. Additionally, when stopping the treatment for the last 10 days of the differentiation a comparable effect can still be observed. This suggests that the AGO is stable and can be effective for a prolonged period of time, which is not surprising since PS modification inhibits degradation by nucleases [25, 276].

During these 22 days we confirmed that the significant downregulation of the *HTT* mRNA does not interfere with the NSC organization into neural rosette structures. Using confocal microscopy, we detected prominent PAX6 positive structures regardless of the treatment. Additionally, we measured equal diameter of the rosettes and detected the presence of the cell cycle marker Proliferation marker protein Ki-67 (Ki67) in the apical part of these structures. This is important, since these structures serve as a niche for the NSC maintenance and

proliferation, with a potential to further differentiate into neurons or glia cells, and it is suggested that the neurogenic ventricular zone is present in the adult mammalian brain [277].

In the next step we further differentiated NSC into neurons and astrocytes which was confirmed by immunofluorescence analysis and confocal microscopy at day 43, 48 and 53. The cells were positive for PAX6, β-III tubulin, mature neuron-specific microtubule-associated protein 2, isoforms a and b (MAP2ab) and astrocyte-specific S100 calcium-binding protein (S100). Starting at day 38 until day 53, when the cells are not dividing, we treated them with the CAG19 for 15 days using gymnosis and measured the levels of *HTT* mRNA and protein. As in early stages of the differentiation, CAG19 significantly downregulated the transcription of *HTT* mRNA compared to the SCR19 in all three timepoints (5, 7 and 15 days of treatment), most efficiently by 61% after 15 days. By stopping the treatment for the last seven days we showed that the AGO can exert its function for a prolonged period of time also in neurons. Additionally, we evaluated the level of HTT protein at day 53, however despite observing an indication of the reduction, this difference was not statistically significant compared to SCR19.

Besides the *HTT* gene there are several other genes with CAG•CTG repeats. Therefore, the AGO could potentially induce unwanted adverse events by anti-gene off-targeting. We selected four genes expressed in the brain: *POU-homeodomain transcription factor BRN2* (*POU3F2*), *Atrophin-1* (*ATN1*), *Ataxin-2* (*ATXN2*) and *Androgen receptor* (*AR*), and measured the levels of expressed mRNA in neurons using RT-qPCR. In all three timepoints we observed no significant off-target effect compared to the SCR19. Additionally, we evaluated the antisense off-targeting of the CAG19 by measuring the mRNA levels of *DM1 protein kinase* (*DMPK*) and *BRI3 binding protein* (*BRI3BP*), two CTG•CAG repeat harboring genes. We observed a significant downregulation of *DMPK* mRNA at day 5, however this effect disappeared in later days. Despite positive results obtained in this analysis, we investigated the expression of only several genes and a more detailed analysis using RNAseq would be needed to make firm conclusions.

4.2 PAPER II

Anti-gene oligonucleotide clamps strand-invade into dsDNA down-regulating *huntingtin* expression

The aim of Paper II was firstly, to improve the design of an LNA/DNA mixmer clamp AGO using an intercalator, M3, with the goal to achieve high invasion into dsDNA *in vitro*, and secondly, evaluate transcription downregulation capabilities of designed ONs in cell models. The ONs were designed against two specific Pu•Py sequences, present in the 3'UTR of the *HTT* gene or upstream the P2 promoter of the *MYC* gene.

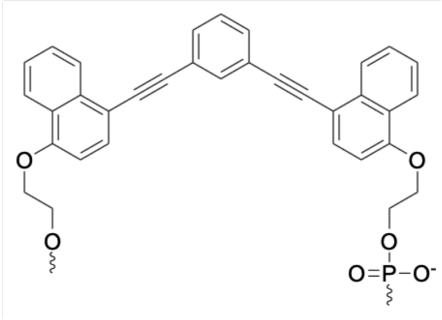

Figure 10: Chemical structure of the M3

The general design of clamp AGO combines a strand-invading ON arm, a TFO arm and a linker connecting them. The clamp strand-invades into dsDNA, hybridizing to one DNA strand by combination of WC and Hoogsteen base pairing, leaving the opposite strand single stranded in the form of a D-loop. In previous studies, clamps invaded less efficiently compared to tail-clamp AGOs, where the WC arm is longer than the TFO [131, 132]. We aimed at improving this by introducing a stacking intercalator, which previously enhanced the invasion of tail-clamps [132]. To evaluate the optimal position of the M3 within the clamp, or the number of M3s in the ON, we positioned the intercalator as the linker and/or at the end of the WC or TFO arm, both being 15 nt long. If the M3 was not used as linker, a non-modified DNA linker with the sequence -tctct- was used instead.

In vitro strand-invasion efficiency at intranuclear pH and salt conditions was determined using the S1 nuclease assay after a 24 h long hybridization of selected MYC-ONs and a supercoiled plasmid containing a single target site. We used five concentrations of the ON in 1:5 dilution series, spanning from the 1:0.64 to 1:400 plasmid:ON ratio (this setup was used for all following experiments). From the initial experiments we determined that the M3 in the linker is crucial for any efficient strand-invasion to take place, since the clamps without it, did essentially not strand-invade at all, even at the highest concentration. Additionally, both arms need to be physically linked considering the corresponding unlinked arms with conjugated M3 did not strand-invade.

Since the clamp AGO combines two modes of binding, the efficiency of the strand-invasion depends both on TFO target recognition/binding and on the strand-displacement capability of the invading arm. Both are influenced by the number of guanines, since the cytosine Hoogsteen base pairing depends on the pH and the G:C WC base pairing in the dsDNA enhances duplex stability [278-280]. While the Pu•Py *MYC* target sequence is mostly composed of adenosines, there is a difference between the 3' and 5' ends, the latter having two consecutive guanosines with additional three in the flanking region. The *HTT* target has three consecutive guanosines one base from the 3' end. To evaluate if these motifs influence the invasion we linked either the 5' end of the WC arm and 3' end of the TFO arm (later termed 3'-oriented clamps) or 3'end of the WC arm and 5' end of the TFO arm (later termed 5'-oriented clamps). As already described, the M3 served as the linker and was also added to the end of the WC or TFO arm. All three HTT-clamps and two MYC-clamps with the M3-linker adjacent to the G-rich motif invade more efficiently. This is in agreement with previously published results on triplex formation [281] and tail-clamp strand-invasion [132] suggesting stable base-pairing at the end of the TFO-arm is necessary to allow invasion of the WC-arm. Additionally, in the optimal clamp orientation the conjugation of a second M3 does not enhance the invasion.

To evaluate if these clamps have a therapeutic potential, we substituted the PO backbone, which is susceptible to nuclease degradation, with the PS modification in all AGOs. *In vitro* PS modified clamps achieved lower strand-invasion after 24 h, which is not surprising since PS backbone reduces the binding affinity of an ON to the DNA duplex [282]. However, after 72 h they reached almost equal invasion percentage as their PO equivalents.

To test the AGO effect on the gene expression we used HD primary patient fibroblasts and transfected the ON with a lipid-based transfection reagent. HTT-clamps target the template strand of the *HTT* gene and should upon hybridization interfere with the RNA polymerase II. After 4 days, all *HTT* targeting clamps successfully downregulated the *HTT* mRNA to clinically relevant level in a dose response manner [251, 283]. This effect was maintained for the next 6 days indicating that the selected modifications allow prolonged anti-gene activity. Furthermore, the ONs did not affect the cell viability which is consistent with previous *in vitro* safety studies of LNA and PS modifications [42].

4.3 PAPER III

Oligonucleotide binding to non-B-DNA in *MYC*

In this project we aimed at examining the non-B-DNA structure formed by the $NHEIII_1$ sequence present in the promoter region of the *MYC Proto-Oncogene* (*MYC*) gene. We then used an *in vitro* model to evaluate structure's effect on the strand-invasion of a tail-clamp AGO.

It was previously shown that this sequence can form either a G-quadruplex or an H-DNA [284]. We focused on the H-DNA and hybridized either the Purine rich TFO (Pu TFO) or the Pyrimidine rich TFO (Py TFO) with the duplex DNA target at pH and salt conditions most suitable for the triplex formation of each TFO. Using EMSA, we determined that this particular sequence rather forms an intermolecular, antiparallel triplex, forming reverse Hoogsteen bonds between the Pu TFO and the dsDNA, which is in accordance to previous publications [285].

To detect an intramolecular H-DNA in the $NHEIII_1$, we used supercoiled bacterial plasmids including or lacking this sequence, termed pMycNHE+ and pMycNHE-, respectively. We incubated them together with BQQ-OP, which specifically recognizes triplex DNA. By adding Cu^{2+} and a reducing agent, free radicals are produced, which results in cleavage of the dsDNA. A second enzymatic cleavage of the plasmid DNA outside the $NHEIII_1$ produces two distinct fragments that can be separated using agarose gel electrophoresis and visualized by fluorescence imaging. Compared to plasmid lacking this sequence, cleavage of the pMycNHE+ resulted in three fragments – linear plasmid and two shorter fragments directly corresponding to the combined cleavage of the BQQ-OP and the restriction enzyme. This confirmed that the H-DNA is formed in the supercoiled plasmid.

The main aim of this project was to analyze the effect of the non-B-DNA structure on the strand-invasion of AGOs. The target site for the AGO was located 11 bases from the $NHEIII_1$ sequence. We used an invading 30 nt long LNA/DNA mixmer WC ON (WC-m44) and two 45 nt long LNA-DNA mixmer tail-clamps AGO, with 5 nt or an M3 intercalating linker, termed Cy3-bis-m44 or Cy3-bis-m44-M3. Previously, the AGOs showed different strand-invading efficiencies, with the Cy3-bis-m44-M3 being superior [132]. All three AGOs were hybridized with either pMycNHE+ or pMycNHE- for 24 or 72 h in intranuclear pH and salt conditions. The strand-invasion was determined using the S1 nuclease assay. The invasion of the Cy3-bis-m44-M3 was equal regardless of the H-DNA forming sequence at both timepoints. Contrary, the generally less efficient Cy3-bis-m44 strand-invasion was improved by the H-DNA structure at 72 h, which was especially evident when the H-DNA was stabilized by the BQQ. Surprisingly, the invasion of the WC-m44 was higher in the absence of the $NHEIII_1$ sequence.

5 CONCLUSIONS AND FUTURE PERSPECTIVES

This thesis aimed at furthering the field of AGOs, using several different types of DNA targeting ON in *in vitro* DNA binding assays and disease cell models.

It describes the use of an *HTT* gene targeting strand-invading ON during *in vitro* neurogenesis from iPSCs through NSC to neurons. It shows that the *HTT* mRNA and protein downregulation achieved by this AGO does not affect the process of neurogenesis, more specifically the formation of neural rosette structures and the pool of NSCs. Additionally, although further studies are needed, it appears that this CAG•CTG repeat targeting ON affects *HTT* specifically and does not influence at least six other genes harboring the repeats in the template or the coding strand. Based on the evidence collected in this project and the previous publication presenting the screening of several variations of this AGO, the LNA/DNA 19 nt mixmer presents itself as a good candidate for further *in vivo* studies.

The thesis further focuses on the optimization and characterization of clamp-type AGOs and their use for the treatment of HD. Based on the results obtained from the S1 nuclease assay, it is evident that the efficient strand-invasion of these ONs depends on the incorporation of the stacking linker between the WC and the TFO arm. Additionally, positioning the linker adjacent to the target sequence motif that could negatively influence TFO binding improves invasion. Furthermore, these AGOs affect the gene expression without affecting the cell viability when directed towards the 3'UTR of the *HTT* template strand. In the future, additional studies are needed to determine the safety profile *in vivo* and potential mutagenic effects of the ON due to the incorporation of the intercalating moiety.

Lastly the thesis pivots towards determination of the antiparallel H-DNA in the $NHEIII_1$ sequence. It aims further at understanding the influence of this non-B-DNA structure on the strand-invasion of tail-clamp AGO, indicating that H-DNA promotes strand-invasion of suboptimal design of this AGO. Although this study suggests that the conformation of the dsDNA influences the strand-invasion efficiency of AGOs, more evidence is needed both in cell models and looking at other non-canonical structures to understand general rules of ON and DNA interaction.

To summarize, this thesis shows that several types of AGOs can be optimized to achieve good binding to the dsDNA and when delivered to the nucleus affect gene transcription of the targeted gene. Combined with the success observed with the RNA targeting ONs, this gives evidence that AGOs have a great potential to be used as therapeutics in the future.

6 ACKNOWLEDGEMENTS

Throughout my doctoral studies I have received ample support and guidance without which this journey would be much more difficult.

First, I would like to thank **Rula Zain** for being my main supervisor during these four years. Thank you for guiding and supporting me through the roller-coaster of what is a PhD education, for being available whenever I needed your help and for allowing me to explore my interests outside of the oligo field.

Edvard Smith, thank you for accepting me first for my master thesis project, offering me a PhD position and giving me the opportunity to work within the Marie Curie network. You have been a great co-supervisor, giving me both space to work and support when I needed it.

Thank you, **Karin Lundin** for being my co-supervisor, for all your input and help, and most importantly for reassuring me that mistakes and failures are always part of science. Thank you for accompanying me to Marie Curie meetings and your constant support.

I would like to thank all the people within the MCG group with whom I shared these four years. **Olof**, for encouraging me and laughing with me throughout my master thesis project and the beginning of my PhD. Thank you for introducing me to the lab work and for teaching me all the techniques. Also, for being the first one who actually told me I have the ability to do the PhD. **Negin**, for laughs in the office and collaboration on the projects. It's been a pleasure learning from you. And thanks for allowing me to steal your cell culture media (and other material) :D. **Osama**, I will never forget the fun we had during our ping-pong games (sorry for smashing the ball so many times :D). It really helped me through these years. **Raul**, for being a great friend and unconditionally helping me with anything I needed. You were my first student, but I think I learned more from you than I ever managed to teach you. I really enjoyed working with you and discussing on common projects. **Salomé**, for nice conversations and positive energy. **Yesid**, for all the good times in the lab, being the only one who understood the quality of sound and being a good friend. **Anna**, for all the interesting conversations about traveling and music and for encouraging me in my hobby. **Vladimir,** for all the fun times in and outside the lab. **Burcu** and **Cristina**, our time together in the lab was short but thanks for helping me in the beginning of my journey in this group. Thanks also to **Eman**, **Qing**, **Yuye** and **Narmeen** for being nice colleagues.

Thank you also to the bachelor and master students who joined our group briefly. **Baranca,** for being the most positive person and always providing the good mood in the office. **Amra**, for trusting me to be your supervisor and allowing me to develop as a teacher. You are still the best mentee I had ;). **Fiona**, for your kindness and good times climbing. **Tina**, for introducing me the EIT Health, the good times in the office and for helping me realize things that hopefully made me a better person and colleague. **Claudia**, grazie for your contagious positive energy which made any dull moment exciting and for giving me insight into my color pallet (my wardrobe thanks you). **Hayley**, for being a great friend and sharing the pictures of chonky pets which definitely lifted my spirits. Thanks also to **Cecilia**, **Haneen**, **Clara** and **Viktoria** for the good company.

I would also like to thank past and present members of the BMM and Evox. **Samir el Andaloussi**, you have a gift in gathering fantastic people and scientists. **Oskar**, thank you for being an awesome friend in and outside the lab from the beginning of my stay in Sweden. Thank you for including me into your circle of friends, which made Stockholm more like a home and not just a place where I spent my PhD. It's been a pleasure climbing, playing DnD and hanging out. **Julia**, I could write a chapter about all the great times I spent with you. Your friendship means the world. Thank you for listening to me even though it wasn't always easy and for being with me at my highs and my lows. **Jeremy**, for the good vibes, great talks and understanding the shoe obsession. **Giulia**, for all the laughs, for the help and encouragement during the last months of my PhD and for letting me procrastinate in your office (I hope I didn't bother you too much ☺). **Rim**, for being just as crazy as I am, thinking that working out in the lab is an excellent idea and for dancing and singing with me when everyone else is trying to work. **Samantha,** my fellow Slovenian, thanks for all your Slovenian meals, giving me a piece of home. I'm happy I can call you a friend. Also, thanks for making me not the clumsiest person in the lab :D. **Kariem**, **Dhanu**, **Antje**, **Daniel** and **Oscar** for scientific advice and great company. **Demir** and **Angus**, for the good times and the laughs. **Manuela**, for taking care of the lab and providing a good environment to work at. Thanks also to other members **Wenyi**, **Tom**, **Ming**, **Noriyasu**, **Beklem**, **André**, **Joel**, **Mattias**, **Safa**, **Doste**, **Taavi**, **Helena** and **Svetlana**, without which, the positive atmosphere in the lab would not be the same.

Thank you also to the CRG group. **Anthony Wright**, for the help with the PhD paperwork. **Laia** for the scientific advice, nice cakes and good company. **Amir** and **Gustav** for fun conversations during lunches.

Thanks also to **Kathrin**, **Emelie** and **Kirsti** for their help in everything related to the administration and paperwork.

The work presented in this thesis would not be possible without established collaborations outside the Karolinska Institutet.

From the University of Southern Denmark, I would like to thank **Jesper Wengel** and his group for providing the oligonucleotides and for scientific input in all my manuscripts, and **Per Trolle Jørgensen** for the input and communication when I had issues with certain oligonucleotides.

From the Sahlgrenska Academy at the University of Gothenburg, I would like to thank **Eric Hanse**, **Henrik Zetterberg**, **Thomas Olsson** and especially **Dzeneta Vizlin-Hodzic**, for collaboration on the work with AGOs and Huntington's Patient-Derived Neural Model System and for introducing me to the world of iPSC.

I would also like to thank **Karin Sollander** and **Helen Bergquist** for their work and help on the project Oligonucleotide Binding to Non-B-DNA in *MYC*.

During my studies I was a part of MMBio, which is a Marie Curie Initial Training Network. I would like to thank all the group leaders and **Liisa van Vliet** for organizing all the international meetings and events, providing me and my fellow early stage researchers, with the best experience we could have. I would also like to thank other ESRs, **Alisa**, **Remkes**, **Joanna**, **Lucia**, **Olivia**, **Madhuri**, **Søren**, **Valerio**, **Enrico**, **Konstantinos**, **Susanna**, **Lise**, **Kristina** and **Rouven**, for all the fun we had and for making the usually overly long meetings enjoyable.

Thank you also to **Roger Strömberg** for accepting me in your group as part of my secondment and introducing me to the world of chemistry, even if it was only for a very short time. **Malgo**, thank you for working with me during this time and trying to teach me the peptide synthesis.

Last but not least I would like to thank my family and friends in Slovenia. It is usually hard to realize and appreciate the positive influence someone has on you when growing up. I do know now that without my parents, who thought me to value experience over material things, who always encouraged me to strive for more and who allowed me to explore whatever I wanted to do, even if it was something crazy, like going to New Zealand for several months, I wouldn't be writing this thesis. Leaving home for Stockholm wasn't the easiest decision but I knew that with your support and encouragement I can make it.

7 REFERENCES

1. Griffith, F., *The Significance of Pneumococcal Types.* Journal of Hygiene, 1928. **27**(2): p. 113-159.

2. Avery, O.T., C.M. Macleod, and M. McCarty, *Studies on the chemical nature of the substance inducing transformation of pneumococcal types: induction of transformation by a deoxyribonucleic acid fraction isolated from pneumococcus type III.* The Journal of experimental medicine, 1944. **79**(2): p. 137.

3. Tatum, E.L. and J. Lederberg, *Gene Recombination in the Bacterium Escherichia coli.* The Journal of Bacteriology, 1947. **53**(6): p. 673.

4. Zinder, N.D. and J. Lederberg, *Genetic exchange in Salmonella.* The Journal of Bacteriology, 1952. **64**(5): p. 679.

5. Watson, J.D. and F.H.C. Crick, *Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid.* Nature, 1953. **171**(4356): p. 737.

6. Gros, F., et al., *Unstable Ribonucleic Acid Revealed by Pulse Labelling of Escherichia Coli.* Nature, 1961. **190**(4776): p. 581-585.

7. Brenner, S., F. Jacob, and M. Meselson, *An Unstable Intermediate Carrying Information from Genes to Ribosomes for Protein Synthesis.* Nature, 1961. **190**(4776): p. 576-581.

8. Jacob, F. and J. Monod, *Genetic regulatory mechanisms in the synthesis of proteins.* Journal of molecular biology, 1961. **3**(3): p. 318-356.

9. Szybalska, E.H. and W. Szybalski, *Genetics of human cess line. IV. DNA- mediated heritable transformation of a biochemical trait.* Proceedings of the National Academy of Sciences of the United States of America, 1962. **48**: p. 2026.

10. Sambrook, J., et al., *The integrated state of viral DNA in SV40- transformed cells.* Proceedings of the National Academy of Sciences of the United States of America, 1968. **60**(4): p. 1288.

11. Stephenson, M.L. and P.C. Zamecnik, *Inhibition of Rous sarcoma viral RNA translation by a specific oligodeoxyribonucleotide.* Proceedings of the National Academy of Sciences of the United States of America, 1978. **75**(1): p. 285.

12. Zamecnik, P.C. and M.L. Stephenson, *Inhibition of Rous sarcoma virus replication and cell transformation by a specific oligodeoxynucleotide.* Proceedings of the National Academy of Sciences of the United States of America, 1978. **75**(1): p. 280-284.

13. Donis-Keller, H., *Site specific enzymatic cleavage of RNA.* Nucleic acids research, 1979. **7**(1): p. 179-192.

14. Fire, A., et al., *Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans.* Nature, 1998. **391**(6669): p. 806-11.

15. Hammond, S.M., et al., *An RNA-directed nuclease mediates post-transcriptional gene silencing in Drosophila cells.* Nature (London), 2000. **404**(6775): p. 293-296.

16. Adams, M.D., et al., *The Sequence of the Human Genome.* Science (American Association for the Advancement of Science), 2001. **291**(5507): p. 1304-1351.

17. Zody, M.C., et al., *Initial sequencing and analysis of the human genome.* Nature (London), 2001. **409**(6822): p. 860-921.

18. Terheggen, H.G., et al., *Unsuccessful trial of gene replacement in arginase deficiency.* Z. Kinder-Heilk., 1975. **119**(1): p. 1-3.

19. Cline, M.J., et al. *Insertion of New Genes into Bone Marrow Cells of Mice*. 1981. Berlin, Heidelberg: Springer Berlin Heidelberg.

20. Wade, N., *UCLA gene therapy racked by friendly fire.* Science, 1980. **210**(4469): p. 509-511.

21. Rosenberg, S.A., et al., *The development of gene therapy for the treatment of cancer.* Annals of surgery, 1993. **218**(4): p. 455.

22. Blaese, R.M., et al., *T lymphocyte- directed gene therapy for ADA- SCID: initial trial results after 4 years.(adenosine deaminase deficiency; severe combined immunodeficiency).* Science, 1995. **270**(5235): p. 475.

23. de Smet, M.D., C.J. Meenken, and G.J. van Den Horn, *Fomivirsen - a phosphorothioate oligonucleotide for the treatment of CMV retinitis.* Ocular immunology and inflammation, 1999. **7**(3-4): p. 189.

24. Wang, F., T. Zuroske, and J.K. Watts, *RNA therapeutics on the rise.* Nat Rev Drug Discov, 2020. **19**(7): p. 441-442.

25. Eder, P.S., et al., *Substrate specificity and kinetics of degradation of antisense oligonucleotides by a 3' exonuclease in plasma.* Antisense Res Dev, 1991. **1**(2): p. 141-51.

26. Agrawal, S., et al., *Pharmacokinetics of antisense oligonucleotides.* Clin Pharmacokinet, 1995. **28**(1): p. 7-16.

27. Mickle, K., et al., *The Effectiveness and Value of Patisiran and Inotersen for Hereditary Transthyretin Amyloidosis.* J Manag Care Spec Pharm, 2019. **25**(1): p. 10-15.

28. Zuluaga-Sanchez, S., et al., *Cost Effectiveness of Nusinersen in the Treatment of Patients with Infantile-Onset and Later-Onset Spinal Muscular Atrophy in Sweden.* Pharmacoeconomics, 2019. **37**(6): p. 845-865.

29. *The cost of getting personal.* Nature Medicine, 2019. **25**(12): p. 1797-1797.

30. Gryaznov, S.M., *Oligonucleotide n3'-->p5' phosphoramidates and thio-phoshoramidates as potential therapeutic agents.* Chem Biodivers, 2010. **7**(3): p. 477-93.

31. Letsinger, R.L., et al., *Cationic oligonucleotides.* Journal of the American Chemical Society, 1988. **110**(13): p. 4470-4471.

32. Eckstein, F. and H. Gindl, *Polyribonucleotides Containing a Phosphorothioate Backbone.* European journal of biochemistry, 1970. **13**(3): p. 558-564.

33. Eckstein, F., *Nucleoside Phosphorothioates.* Annual review of biochemistry, 1985. **54**(1): p. 367-402.

34. Iwamoto, N., et al., *Control of phosphorothioate stereochemistry substantially increases the efficacy of antisense oligonucleotides.* Nature biotechnology, 2017. **35**(9): p. 845-851.

35. Kanehara, H., et al., *Influence of a Thiophosphate Linkage on the Duplex Stability - Does Sp Configuration Always Lead to Higher Stability Than Rp?* Nucleosides and Nucleotides, 1996. **15**(6): p. 1169-1178.

36. Nikiforov, T.T., et al., *The use of phosphorothioate primers and exonuclease hydrolysis for the preparation of single-stranded PCR products and their detection by solid-phase hybridization.* PCR Methods Appl, 1994. **3**(5): p. 285-91.

37. Spitzer, S. and F. Eckstein, *Inhibition of deoxyribonucleases by phosphorothioate groups in oligodeoxyribonucleotides.* Nucleic acids research, 1988. **16**(24): p. 11691-11704.

38. Miller, C.M., et al., *Receptor-Mediated Uptake of Phosphorothioate Antisense Oligonucleotides in Different Cell Types of the Liver.* Nucleic acid therapeutics, 2018. **28**(3): p. 119-127.

39. Juliano, R.L., et al., *Cellular uptake and intracellular trafficking of oligonucleotides: implications for oligonucleotide pharmacology.* Nucleic acid therapeutics, 2014. **24**(2): p. 101-113.

40. Frazier, K.S., *Antisense Oligonucleotide Therapies: The Promise and the Challenges from a Toxicologic Pathologist's Perspective.* Toxicologic pathology, 2015. **43**(1): p. 78-89.

41. Shen, W., et al., *Chemical modification of PS-ASO therapeutics reduces cellular protein-binding and improves the therapeutic index.* Nature biotechnology, 2019. **37**(6): p. 640-650.

42. Funder, E.D., et al., *Refining LNA safety profile by controlling phosphorothioate stereochemistry.* PloS one, 2020. **15**(6): p. e0232603.

43. Henry, S., et al., *Chemically Modified Oligonucleotides Exhibit Decreased Immune Stimulation in Mice.* The Journal of pharmacology and experimental therapeutics, 2000. **292**(2): p. 468-479.

44. Summerton, J. and D. Weller, *Morpholino antisense oligomers: design, preparation, and properties.* Antisense Nucleic Acid Drug Dev, 1997. **7**(3): p. 187-95.

45. Summerton, J., *Morpholino antisense oligomers: the case for an RNase H-independent structural type.* Biochim Biophys Acta, 1999. **1489**(1): p. 141-58.

46. Hudziak, R.M., et al., *Resistance of morpholino phosphorodiamidate oligomers to enzymatic degradation.* Antisense Nucleic Acid Drug Dev, 1996. **6**(4): p. 267-72.

47. Nielsen, P.E., et al., *Sequence-Selective Recognition of DNA by Strand Displacement with a Thymine-Substituted Polyamide.* Science, 1991. **254**(5037): p. 1497-1500.

48. Egholm, M., et al., *PNA hybridizes to complementary oligonucleotides obeying the Watson–Crick hydrogen-bonding rules.* Nature, 1993. **365**(6446): p. 566-568.

49. Nielsen, P.E., M. Egholm, and O. Buchardt, *Evidence for (PNA)2/DNA triplex structure upon binding of PNA to dsDNA by strand displacement.* J Mol Recognit, 1994. **7**(3): p. 165-70.

50. Nielsen, P.E., *Addressing the challenges of cellular delivery and bioavailability of peptide nucleic acids (PNA).* Quarterly Reviews of Biophysics, 2005. **38**(4): p. 345-350.

51. Wancewicz, E.V., et al., *Peptide nucleic acids conjugated to short basic peptides show improved pharmacokinetics and antisense activity in adipose tissue.* Journal of medicinal chemistry, 2010. **53**(10): p. 3919-3926.

52. Zeng, Z., et al., *A Tat-conjugated Peptide Nucleic Acid Tat-PNA-DR Inhibits Hepatitis B Virus Replication In Vitro and In Vivo by Targeting LTR Direct Repeats of HBV RNA.* Molecular Therapy - Nucleic Acids, 2016. **5**(3): p. e295.

53. Inoue, H., et al., *Synthesis and hybridization studies on two complementary nona(2'-O-methyl)ribonucleotides.* Nucleic acids research, 1987. **15**(15): p. 6131.

54. Kawasaki, A.M., et al., *Uniformly modified 2'-deoxy-2'-fluoro phosphorothioate oligonucleotides as nuclease- resistant antisense compounds with high affinity and specificity for RNA targets.* Journal of medicinal chemistry, 1993. **36**(7): p. 831.

55. Lima, W.F. and S.T. Crooke, *Binding affinity and specificity of Escherichia coli RNase H1: impact on the kinetics of catalysis of antisense oligonucleotide-RNA hybrids.* Biochemistry, 1997. **36**(2): p. 390-8.

56. Jirka, S.M., et al., *Evaluation of 2'-Deoxy-2'-fluoro Antisense Oligonucleotides for Exon Skipping in Duchenne Muscular Dystrophy.* Mol Ther Nucleic Acids, 2015. **4**(12): p. e265.

57. Rigo, F., et al., *Synthetic oligonucleotides recruit ILF2/3 to RNA transcripts to modulate splicing.* Nat Chem Biol, 2012. **8**(6): p. 555-61.

58. Cummins, L.L., et al., *Characterization of fully 2- modified oligoribonucleotide hetero- and homoduplex hybridization and nuclease sensitivity.* Nucleic acids research, 1995. **23**(11): p. 2019.

59. Monia, B.P., et al., *Nuclease resistance and antisense activity of modified oligonucleotides targeted to Ha-ras.* J Biol Chem, 1996. **271**(24): p. 14533-40.

60. Monia, B.P., et al., *Evaluation of 2'-modified oligonucleotides containing 2'-deoxy gaps as antisense inhibitors of gene expression.* J Biol Chem, 1993. **268**(19): p. 14514-22.

61. Kraynack, B.A. and B.F. Baker, *Small interfering RNAs containing full 2'-O-methylribonucleotide-modified sense strands display Argonaute2/eIF2C2-dependent activity.* RNA (New York, N.Y.), 2006. **12**(1): p. 163-176.

62. Egli, M., et al., *Probing the Influence of Stereoelectronic Effects on the Biophysical Properties of Oligonucleotides: Comprehensive Analysis of the RNA Affinity, Nuclease Resistance, and Crystal Structure of Ten 2'-O-Ribonucleic Acid Modifications.* Biochemistry, 2005. **44**(25): p. 9045-9057.

63. Teplova, M., et al., *Crystal structure and improved antisense properties of 2'-O-(2-methoxyethyl)-RNA.* Nature Structural Biology, 1999. **6**(6): p. 535-539.

64. K. Singh, S., et al., *LNA (locked nucleic acids): synthesis and high-affinity nucleic acid recognition.* Chemical Communications, 1998(4): p. 455-456.

65. Koshkin, A.A., et al., *LNA (Locked Nucleic Acids): Synthesis of the adenine, cytosine, guanine, 5- methylcytosine, thymine and uracil bicyclonucleoside monomers, oligomerisation, and unprecedented nucleic acid recognition.* Tetrahedron, 1998. **54**(14): p. 3607-3630.

66. Christensen, U., et al., *Stopped-flow kinetics of locked nucleic acid (LNA)-oligonucleotide duplex formation: studies of LNA-DNA and DNA-DNA interactions.* Biochem J, 2001. **354**(Pt 3): p. 481-4.

67. Singh, S.K. and J. Wengel, *Universality of LNA-mediated high-affinity nucleic acid recognition.* Chemical communications (Cambridge, England), 1998(12): p. 1247-1248.

68. Frieden, M., H.F. Hansen, and T. Koch, *Nuclease Stability of LNA Oligonucleotides and LNA-DNA Chimeras.* Nucleosides, nucleotides & nucleic acids, 2003. **22**(5-8): p. 1041-1043.

69. Wahlestedt, C., et al., *Potent and nontoxic antisense oligonucleotides containing locked nucleic acids.* Proceedings of the National Academy of Sciences, 2000. **97**(10): p. 5633-5638.

70. You, Y., et al., *Design of LNA probes that improve mismatch discrimination.* Nucleic Acids Res, 2006. **34**(8): p. e60.

71. Højland, T., et al., *LNA (locked nucleic acid) and analogs as triplex- forming oligonucleotides.* Organic & biomolecular chemistry, 2007. **5**(15): p. 2375.

72. Wengel, C.L., et al., *Oligonucleotides Containing Aminated 2'-Amino-LNA Nucleotides: Synthesis and Strong Binding to Complementary DNA and RNA.* 2017.

73. Hvam, M.L., et al., *Fatty Acid-Modified Gapmer Antisense Oligonucleotide and Serum Albumin Constructs for Pharmacokinetic Modulation.* Molecular therapy, 2017. **25**(7): p. 1710-1717.

74. Seley-Radtke, K.L. and M.K. Yates, *The evolution of nucleoside analogue antivirals: A review for chemists and non-chemists. Part 1: Early structural modifications to the nucleoside scaffold.* Antiviral research, 2018. **154**: p. 66-86.

75. Tsesmetzis, N., et al., *Nucleobase and Nucleoside Analogues: Resistance and Re-Sensitisation at the Level of Pharmacokinetics, Pharmacodynamics and Metabolism.* Cancers, 2018. **10**(7): p. 240.

76. Wang, S. and E.T. Kool, *Origins of the large differences in stability of DNA and RNA helices: C-5 methyl and 2'-hydroxyl effects.* Biochemistry (Easton), 1995. **34**(12): p. 4125-4132.

77. Gutierrez, A.J., et al., *5-Heteroaryl-2'-deoxyuridine Analogs. Synthesis and Incorporation into High-Affinity Oligonucleotides.* Journal of the American Chemical Society, 1994. **116**(13): p. 5540-5544.

78. Wagner, R., et al., *Antisense gene inhibition by oligonucleotides containing C-5 propyne pyrimidines.* Science, 1993. **260**(5113): p. 1510-1513.

79. Luyten, I. and P. Herdewijn, *Hybridization properties of base-modified oligonucleotides within the double and triple helix motif.* European Journal of Medicinal Chemistry, 1998. **33**(7): p. 515-576.

80. Stein, H. and P. Hausen, *Enzyme from Calf Thymus Degrading the RNA Moiety of DNA-RNA Hybrids: Effect on DNA-Dependent RNA Polymerase.* Science (American Association for the Advancement of Science), 1969. **166**(3903): p. 393-395.

81. Wu, H., et al., *Human RNase H1 is associated with protein P32 and is involved in mitochondrial pre-rRNA processing.* PloS one, 2013. **8**(8): p. e71006-e71006.

82. Shen, W., et al., *Dynamic nucleoplasmic and nucleolar localization of mammalian RNase H1 in response to RNAP I transcriptional R-loops.* Nucleic acids research, 2017. **45**(18): p. 10672-10692.

83. Turchi, J.J., et al., *Enzymatic Completion of Mammalian Lagging-Strand DNA Replication.* Proceedings of the National Academy of Sciences - PNAS, 1994. **91**(21): p. 9803-9807.

84. Liang, X.-H., et al., *RNase H1-Dependent Antisense Oligonucleotides Are Robustly Active in Directing RNA Cleavage in Both the Cytoplasm and the Nucleus.* Molecular therapy, 2017. **25**(9): p. 2075-2092.

85. Boiziau, C., et al., *Inhibition of translation initiation by antisense oligonucleotides via an RNase-H independent mechanism.* Nucleic acids research, 1991. **19**(5): p. 1113-1119.

86. Liang, X.-H., W. Shen, and S.T. Crooke, *Specific Increase of Protein Levels by Enhancing Translation Using Antisense Oligonucleotides Targeting Upstream Open Frames.* RNA Activation, 2017. **983**: p. 129-146.

87. Liang, X.-H., et al., *Antisense oligonucleotides targeting translation inhibitory elements in 5′ UTRs can selectively increase protein levels.* Nucleic acids research, 2017. **45**(16): p. 9528-9546.

88. Marsollier, A.-C., et al., *Antisense targeting of 3′ end elements involved in DUX4 mRNA processing is an efficient therapeutic strategy for facioscapulohumeral dystrophy: a new gene-silencing approach.* Human molecular genetics, 2016. **25**(8): p. 1468-1478.

89. Derbis, M., et al., *Short antisense oligonucleotides alleviate the pleiotropic toxicity of RNA harboring expanded CGG repeats.* Nature communications, 2021. **12**(1): p. 1265-17.

90. Rué, L., et al., *Targeting CAG repeat RNAs reduces Huntington's disease phenotype independently of huntingtin levels.* The Journal of Clinical Investigation, 2016. **126**(11): p. 4319-4330.

91. Zhang, T., et al., *Antisense oligonucleotides targeting the RNA binding region of the NP gene inhibit replication of highly pathogenic avian influenza virus H5N1.* International Immunopharmacology, 2011. **11**(12): p. 2057-2061.

92. Kabanov, A.V., et al., *A new class of antivirals: antisense oligonucleotides combined with a hydrophobic substituent effectively inhibit influenza virus reproduction and synthesis of virus-specific proteins in MDCK cells.* FEBS Letters, 1990. **259**(2): p. 327-330.

93. Heald, A.E., et al., *AVI-7288 for Marburg Virus in Nonhuman Primates and Humans.* N Engl J Med, 2015. **373**(4): p. 339-48.

94. Beigel, J.H., et al., *Safety, tolerability, and pharmacokinetics of radavirsen (AVI-7100), an antisense oligonucleotide targeting influenza a M1/M2 translation.* Br J Clin Pharmacol, 2018. **84**(1): p. 25-34.

95. Mann, C.J., et al., *Improved antisense oligonucleotide induced exon skipping in the mdx mouse model of muscular dystrophy.* The Journal of Gene Medicine, 2002. **4**(6): p. 644-654.

96. Aartsma-Rus, A., et al., *Therapeutic Modulation of DMD Splicing by Blocking Exonic Splicing Enhancer Sites with Antisense Oligonucleotides.* Annals of the New York Academy of Sciences, 2006. **1082**(1): p. 74-76.

97. Aartsma-Rus, A. and G.J. van Ommen, *Antisense-mediated exon skipping: a versatile tool with therapeutic and research applications.* Rna, 2007. **13**(10): p. 1609-24.

98. Aartsma-Rus, A., et al., *Development of Exon Skipping Therapies for Duchenne Muscular Dystrophy: A Critical Review and a Perspective on the Outstanding Issues.* Nucleic Acid Ther, 2017. **27**(5): p. 251-259.

99. Singh, R.N. and N.N. Singh, *Mechanism of Splicing Regulation of Spinal Muscular Atrophy Genes.* Advances in neurobiology, 2018. **20**: p. 31-61.

100. Titeux, M., et al., *Antisense-Mediated Splice Modulation to Reframe Transcripts.* Methods Mol Biol, 2018. **1828**: p. 531-552.

101. Bauman, J.A., et al., *Anti-tumor activity of splice-switching oligonucleotides.* Nucleic Acids Res, 2010. **38**(22): p. 8348-56.

102. Disterer, P., et al., *Exon Skipping of Hepatic APOB Pre-mRNA With Splice-switching Oligonucleotides Reduces LDL Cholesterol In Vivo.* Molecular Therapy, 2013. **21**(3): p. 602-609.

103. Tan, F.L. and J.Q. Yin, *RNAi, a new therapeutic strategy against viral infection.* Cell Research, 2004. **14**(6): p. 460-466.

104. Caplen, N.J., *Gene Therapy Progress and Prospects. Downregulating gene expression: the impact of RNA interference.* Gene Therapy, 2004. **11**(16): p. 1241-1248.

105. Patrick Walton, S., et al., *Designing highly active siRNAs for therapeutic applications.* The FEBS Journal, 2010. **277**(23): p. 4806-4813.

106. Wang, Z., *The guideline of the design and validation of MiRNA mimics.* Methods Mol Biol, 2011. **676**: p. 211-23.

107. Nair, J.K., et al., *Multivalent N-acetylgalactosamine-conjugated siRNA localizes in hepatocytes and elicits robust RNAi-mediated gene silencing.* J Am Chem Soc, 2014. **136**(49): p. 16958-61.

108. Osborn, M.F. and A. Khvorova, *Improving siRNA Delivery In Vivo Through Lipid Conjugation.* Nucleic Acid Ther, 2018. **28**(3): p. 128-136.

109. Nikam, R.R. and K.R. Gore, *Journey of siRNA: Clinical Developments and Targeted Delivery.* Nucleic Acid Ther, 2018. **28**(4): p. 209-224.

110. Letai, A.G., et al., *Specificity in formation of triple-stranded nucleic acid helical complexes: studies with agarose-linked polyribonucleotide affinity columns.* Biochemistry, 1988. **27**(26): p. 9108-9112.

111. Moser, H.E. and P.B. Dervan, *Sequence-Specific Cleavage of Double Helical DNA by Triple Helix Formation.* Science (American Association for the Advancement of Science), 1987. **238**(4827): p. 645-650.

112. Le Doan, T., et al., *Sequence-specific recognition, photocrosslinking and cleavage of the DNA double helix by an oligo-[alpha]-thymidylate covalently linked to an azidoproflavine derivative.* Nucleic Acids Res, 1987. **15**(19): p. 7749-60.

113. Carbone, G.M., et al., *DNA binding and antigene activity of a daunomycin-conjugated triplex- forming oligonucleotide targeting the P2 promoter of the human c- myc gene.* Nucleic Acids Research, 2004. **32**(8): p. 2396.

114. Napoli, S., et al., *Growth inhibition and apoptosis induced by daunomycin-conjugated triplex- forming oligonucleotides targeting the c- myc gene in prostate cancer cells.* Nucleic Acids Research, 2006. **34**(2): p. 734.

115. Janowski, B.A., et al., *Inhibiting gene expression at transcription start sites in chromosomal DNA with antigene RNAs.* Nature Chemical Biology, 2005. **1**(4): p. 216-222.

116. Kovacs, A., et al., *Triple helix-forming oligonucleotide corresponding to the polypyrimidine sequence in the rat alpha 1(I) collagen promoter specifically inhibits factor binding and transcription.* The Journal of biological chemistry, 1996. **271**(3): p. 1805.

117. Faria, M., et al., *Targeted inhibition of transcription elongation in cells mediated by triplex-forming oligonucleotides.* Proc Natl Acad Sci U S A, 2000. **97**(8): p. 3862-7.

118. Joy, C., et al., *Design of embedded chimeric peptide nucleic acids that efficiently enter and accurately reactivate gene expression in vivo.* Proceedings of the National Academy of Sciences, 2010. **107**(39): p. 16846.

119. Chin, J.Y., F. Reza, and P.M. Glazer, *Triplex-forming peptide nucleic acids induce heritable elevations in gamma-globin expression in hematopoietic progenitor cells.* Mol Ther, 2013. **21**(3): p. 580-7.

120. Nagatsugi, F., et al., *Site-specific mutagenesis by triple helix- forming oligonucleotides containing a reactive nucleoside analog.* Nucleic acids research, 2003. **31**(6): p. e31.

121. Ricciardi, A.S., et al., *In utero nanoparticle delivery for site-specific genome editing.* Nature communications, 2018. **9**(1): p. 2481-2481.

122. Ricciardi, A.S., et al., *Targeted genome modification via triple helix formation.* Methods in molecular biology (Clifton, N.J.), 2014. **1176**: p. 89.

123. Lohani, N. and M.R. Rajeswari, *Antigene and Antiproliferative Effects of Triplex-Forming Oligonucleotide (TFO) Targeted on hmgb1 Gene in Human Hepatoma Cells.* Anticancer Agents Med Chem, 2020.

124. Taniguchi, Y., et al., *Development of novel C-nucleoside analogues for the formation of antiparallel-type triplex DNA with duplex DNA that includes TA and dUA base pairs.* Org Biomol Chem, 2020. **18**(15): p. 2845-2851.

125. Ohkubo, A., et al., *Chemical synthesis and properties of modified oligonucleotides containing 5'-amino-5'-deoxy-5'-hydroxymethylthymidine residues.* Bioorg Med Chem, 2020. **28**(8): p. 115407.

126. Okamura, H., Y. Taniguchi, and S. Sasaki, *Aminopyridinyl-Pseudodeoxycytidine Derivatives Selectively Stabilize Antiparallel Triplex DNA with Multiple CG Inversion Sites.* Angew Chem Int Ed Engl, 2016. **55**(40): p. 12445-9.

127. Zaghloul, E.M., et al., *CTG repeat-targeting oligonucleotides for down-regulating Huntingtin expression.* Nucleic Acids Research, 2017. **45**(9): p. 5153-5169.

128. Gissberg, O.I., et al., *Fast and efficient synthesis of Zorro- LNA type 3-5-5-3 oligonucleotide conjugates via parallel in situ stepwise conjugation.* Organic & Biomolecular Chemistry, 2016. **14**(14): p. 3584-3590.

129. Zaghloul, E.M., et al., *Optimizing anti- gene oligonucleotide ' Zorro- LNA' for improved strand invasion into duplex DNA.* Nucleic Acids Research, 2011. **39**(3): p. 1142.

130. Ge, R., et al., *Zorro locked nucleic acid induces sequence- specific gene silencing.* FASEB journal : official publication of the Federation of American Societies for Experimental Biology, 2007. **21**(8): p. 1902.

131. Moreno, P.M.D., et al., *Development of bis- locked nucleic acid (bisLNA) oligonucleotides for efficient invasion of supercoiled duplex DNA.* Nucleic Acids Research, 2013. **41**(5): p. 3257-3273.

132. Geny, S., et al., *Next-generation bis-locked nucleic acids with stacking linker and 2'-glycylamino-LNA show enhanced DNA invasion into supercoiled duplexes.* Nucleic Acids Research, 2016. **44**(5): p. 2007.

133. Peffer, N.J., et al., *Strand- invasion of duplex DNA by peptide nucleic acid oligomers.* Proceedings of the National Academy of Sciences of the United States of America, 1993. **90**(22): p. 10648.

134. Sergey, V.S., et al., *Enhancement of strand invasion by oligonucleotides through manipulation of backbone charge.* Nature Biotechnology, 1996. **14**(13): p. 1700.

135. Bergquist, H., et al., *Disruption of Higher Order DNA Structures in Friedreich's Ataxia (GAA)n Repeats by PNA or LNA Targeting.* PLoS One, 2016. **11**(11): p. e0165788.

136. Egholm, M., et al., *Efficient pH-independent sequence-specific DNA binding by pseudoisocytosine-containing bis-PNA.* Nucleic Acids Res, 1995. **23**(2): p. 217-22.

137. Bentin, T., H.J. Larsen, and P.E. Nielsen, *Combined triplex/duplex invasion of double-stranded DNA by "tail-clamp" peptide nucleic acid.* Biochemistry, 2003. **42**(47): p. 13987-95.

138. Kuhn, H., et al., *An experimental study of mechanism and specificity of peptide nucleic acid (PNA) binding to duplex DNA.* J Mol Biol, 1999. **286**(5): p. 1337-45.

139. Economos, N.G., et al., *Peptide Nucleic Acids and Gene Editing: Perspectives on Structure and Repair.* Molecules (Basel, Switzerland), 2020. **25**(3): p. 735.

140. Franklin, R.E. and R.G. Gosling, *Molecular configuration in sodium thymonucleate.* Nature, 1953. **171**(4356): p. 740-1.

141. Shortle, B.E., et al., *Physical and Enzymatic Studies on Poly d(I-C).Poly d(I-C), an Unusual Double-helical DNA.* Nature (London), 1970. **228**(5277): p. 1166-1169.

142. Zhao, J., et al., *Non-B DNA structure-induced genetic instability and evolution.* Cellular and molecular life sciences : CMLS, 2010. **67**(1): p. 43-62.

143. Bedrat, A., L. Lacroix, and J.-L. Mergny, *Re-evaluation of G-quadruplex propensity with G4Hunter.* Nucleic acids research, 2016. **44**(4): p. 1746-1759.

144. Kouzine, F., et al., *Permanganate/S1 Nuclease Footprinting Reveals Non-B DNA Structures with Regulatory Potential across a Mammalian Genome.* Cell systems, 2017. **4**(3): p. 344-356.e7.

145. Kouzine, F. and D. Levens, *Supercoil-driven DNA structures regulate genetic transactions.* Frontiers in bioscience, 2007. **12**(8-12): p. 4409-4423.

146. Zhabinskaya, D. and C.J. Benham, *Competitive superhelical transitions involving cruciform extrusion.* Nucleic acids research, 2013. **41**(21): p. 9610-9621.

147. Delic, J., R. Onclercq, and M. Moisan-Coppey, *Inhibition and enhancement of eukaryotic gene expression by potential non-B DNA sequences.* Biochemical and biophysical research communications, 1991. **181**(2): p. 818-826.

148. Wölfl, S., B. Wittig, and A. Rich, *Identification of transcriptionally induced Z-DNA segments in the human c- myc gene.* Biochimica et biophysica acta. Gene structure and expression, 1995. **1264**(3): p. 294-302.

149. Tang, C. and T. MacCarthy, *Characterization of DNA G-Quadruplex Structures in Human Immunoglobulin Heavy Variable (IGHV) Genes.* Frontiers in immunology, 2021. **12**: p. 671944-671944.

150. Jurikova, K., et al., *Role of folding kinetics of secondary structures in telomeric G-overhangs in the regulation of telomere maintenance in Saccharomyces cerevisiae.* The Journal of biological chemistry, 2020. **295**(27): p. 8958-8971.

151. Jara-Espejo, M. and S.R. Peres Line, *DNA G-quadruplex stability, position and chromatin accessibility are associated with CpG island methylation.* The FEBS journal, 2020. **287**(3): p. 483-495.

152. Bacolla, A., et al., *Translocation and deletion breakpoints in cancer genomes are associated with potential non-B DNA-forming sequences.* Nucleic acids research, 2016. **44**(12): p. 5673-5688.

153. Wojciechowska, M., et al., *Non-B DNA Conformations Formed by Long Repeating Tracts of Myotonic Dystrophy Type 1, Myotonic Dystrophy Type 2, and Friedreich's Ataxia Genes, Not the Sequences per se, Promote Mutagenesis in Flanking Regions.* The Journal of biological chemistry, 2006. **281**(34): p. 24531-24543.

154. Bacolla, A., et al., *The involvement of non-B DNA structures in gross chromosomal rearrangements.* DNA Repair, 2006. **5**(9): p. 1161-1170.

155. Gonitel, R., et al., *DNA Instability in Postmitotic Neurons.* Proceedings of the National Academy of Sciences - PNAS, 2008. **105**(9): p. 3467-3472.

156. Rindler, P.M. and S.I. Bidichandani, *Role of transcript and interplay between transcription and replication in triplet-repeat instability in mammalian cells.* Nucleic acids research, 2011. **39**(2): p. 526-535.

157. Hanvey, J.C., M. Shimizu, and R.D. Wells, *Intramolecular DNA triplexes in supercoiled plasmids.* Proceedings of the National Academy of Sciences - PNAS, 1988. **85**(17): p. 6292-6296.

158. Lyamichev, V.I., et al., *DNA H form requires a homopurine-homopyrimidine mirror repeat.* Nature (London), 1987. **330**(6147): p. 495-497.

159. Zain, R. and J.S. Sun, *Do natural DNA triple-helical structures occur and function in vivo?* Cellular and Molecular Life Sciences CMLS, 2003. **60**(5): p. 862-870.

160. Palecek, E., *Local supercoil-stabilized DNA structures.* Critical reviews in biochemistry and molecular biology, 1991. **26**(2): p. 151.

161. Yoshinori, K. and K.-S. Terumi, *Magnesium Ion-Dependent Triple-Helix Structure Formed by Homopurine-Homopyrimidine Sequences in Supercoiled Plasmid DNA.* Proceedings of the National Academy of Sciences - PNAS, 1988. **85**(11): p. 3781-3785.

162. Campos, J.L. and J.A. Subirana, *The influence of Mg++ and Zn++ on polypurine-polypyrimidine multistranded helices.* J Biomol Struct Dyn, 1991. **8**(4): p. 793-800.

163. Kang, S., F. Wohlrab, and R.D. Wells, *GC-rich flanking tracts decrease the kinetics of intramolecular DNA triplex formation.* The Journal of biological chemistry, 1992. **267**(27): p. 19435-19442.

164. Behe, M.J., *An overabundance of long oligopurine tracts occurs in the genome of simple and complex eukaryotes.* Nucleic Acids Res, 1995. **23**(4): p. 689-95.

165. Brahmachari, S.K., et al., *Polypurine/polypyrimidine sequences as cis-acting transcriptional regulators.* Gene, 1997. **190**(1): p. 17-26.

166. Schroth, G.P. and P.S. Ho, *Occurrence of potential cruciform and H-DNA forming sequences in genomic DNA.* Nucleic Acids Res, 1995. **23**(11): p. 1977-83.

167. Guliang, W. and K.M. Vasquez, *Naturally occurring H-DNA-forming sequences are mutagenic in mammalian cells.* Proceedings of the National Academy of Sciences of the United States of America, 2004. **101**(37): p. 13448.

168. Gacy, A.M., et al., *GAA Instability in Friedreich's Ataxia Shares a Common, DNA-Directed and Intraallelic Mechanism with Other Trinucleotide Diseases.* Molecular cell, 1998. **1**(4): p. 583-593.

169. Bidichandani, S.I., T. Ashizawa, and P.I. Patel, *The GAA Triplet-Repeat Expansion in Friedreich Ataxia Interferes with Transcription and May Be Associated with an Unusual DNA Structure.* American journal of human genetics, 1998. **62**(1): p. 111-121.

170. Henderson, E., et al., *Telomeric DNA oligonucleotides form novel intramolecular structures containing guanine·guanine base pairs.* Cell, 1987. **51**(6): p. 899-908.

171. Wang, Y. and D.J. Patel, *Guanine residues in d(T2AG3) and d(T2G4) form parallel-stranded potassium cation stabilized G-quadruplexes with anti glycosidic torsion angles in solution.* Biochemistry (Easton), 1992. **31**(35): p. 8112-8119.

172. Dai, J., et al., *An Intramolecular G-Quadruplex Structure with Mixed Parallel/Antiparallel G-Strands Formed in the Human BCL-2 Promoter Region in Solution.* Journal of the American Chemical Society, 2006. **128**(4): p. 1096-1098.

173. Wang, Y. and D.J. Patel, *Solution structure of the human telomeric repeat d[AG3(T2AG3)3] G-tetraplex.* Structure (London), 1993. **1**(4): p. 263-282.

174. Huppert, J.L. and S. Balasubramanian, *Prevalence of quadruplexes in the human genome.* Nucleic acids research, 2005. **33**(9): p. 2908-2916.

175. Todd, A.K., M. Johnston, and S. Neidle, *Highly prevalent putative quadruplex sequence motifs in human DNA.* Nucleic acids research, 2005. **33**(9): p. 2901-2907.

176. Chambers, V.S., et al., *High-throughput sequencing of DNA G-quadruplex structures in the human genome.* Nature biotechnology, 2015. **33**(8): p. 877-881.

177. Hänsel-Hertsch, R., et al., *G-quadruplex structures mark human regulatory chromatin.* Nature genetics, 2016. **48**(10): p. 1267-1272.

178. Huppert, J.L. and S. Balasubramanian, *G-quadruplexes in promoters throughout the human genome.* Nucleic acids research, 2006. **35**(2): p. 406-413.

179. Besnard, E., et al., *Unraveling cell type-specific and reprogrammable human replication origin signatures associated with G-quadruplex consensus motifs.* Nature structural & molecular biology, 2012. **19**(8): p. 837-844.

180. Guilbaud, G., et al., *Local epigenetic reprogramming induced by G-quadruplex ligands.* Nature chemistry, 2017. **9**(11): p. 1110-1117.

181. Nayun, K., *The Interplay between G-quadruplex and Transcription.* Current Medicinal Chemistry, 2019. **26**(16): p. 2898-2917.

182. Johnson, F.B., *Fundamentals of G-quadruplex biology.* Annual reports in medicinal chemistry, 2020. **54**: p. 3-44.

183. Biffi, G., et al., *Quantitative visualization of DNA G-quadruplex structures in human cells.* Nature chemistry, 2013. **5**(3): p. 182-186.

184. De, S. and F. Michor, *DNA secondary structures and epigenetic determinants of cancer genome evolution.* Nature structural & molecular biology, 2011. **18**(8): p. 950-955.

185. Biffi, G., et al., *Elevated levels of G-quadruplex formation in human stomach and liver cancer tissues.* PloS one, 2014. **9**(7): p. e102711-e102711.

186. Dott, P.J., C.R. Chuang, and G.F. Saunders, *Inverted repetitive sequences in the human genome.* Biochemistry (Easton), 1976. **15**(18): p. 4120-4125.

187. Smith, G.R., *Meeting DNA palindromes head-to-head.* Genes & development, 2008. **22**(19): p. 2612-2620.

188. Lilley, D.M., *The inverted repeat as a recognizable structural feature in supercoiled DNA molecules.* Proceedings of the National Academy of Sciences, 1980. **77**(11): p. 6468-6472.

189. Panayotatos, N. and R.D. Wells, *Cruciform structures in supercoiled DNA.* Nature (London), 1981. **289**(5797): p. 466-470.

190. Oussatcheva, E.A., et al., *Influence of Global DNA Topology on Cruciform Formation in Supercoiled DNA.* Journal of molecular biology, 2004. **338**(4): p. 735-743.

191. Liu, G., et al., *Replication-dependent instability at (CTG) x (CAG) repeat hairpins in human cells.* Nat Chem Biol, 2010. **6**(9): p. 652-9.

192. Nag, D.K. and T.D. Petes, *Seven-base-pair inverted repeats in DNA form stable hairpins in vivo in Saccharomyces cerevisiae.* Genetics, 1991. **129**(3): p. 669-73.

193. Pearson, C.E., et al., *Inverted repeats, stem-loops, and cruciforms: Significance for initiation of DNA replication.* Journal of cellular biochemistry, 1996. **63**(1): p. 1-22.

194. Yamamoto, Y., O. Miura, and T. Ohyama, *Cruciform Formable Sequences within Pou5f1 Enhancer Are Indispensable for Mouse ES Cell Integrity.* International journal of molecular sciences, 2021. **22**(7): p. 3399.

195. Zannis-Hadjopoulos, M., W. Yahyaoui, and M. Callejo, *14-3-3 Cruciform-binding proteins as regulators of eukaryotic DNA replication.* Trends in biochemical sciences (Amsterdam. Regular ed.), 2008. **33**(1): p. 44-50.

196. Saintamand, A., et al., *Deciphering the importance of the palindromic architecture of the immunoglobulin heavy-chain 3' regulatory region.* Nature communications, 2016. **7**(1): p. 10730-10730.

197. Bakhtiari, M., et al., *Variable number tandem repeats mediate the expression of proximal genes.* Nature communications, 2021. **12**(1): p. 2075-2075.

198. Brázda, V., et al., *Cruciform structures are a common DNA feature important for regulating biological processes.* BMC Molecular Biology, 2011. **12**(1): p. 33.

199. Warburton, P.E., et al., *Inverted repeat structure of the human genome: the X-chromosome contains a preponderance of large, highly homologous inverted repeats that contain testes genes.* Genome research, 2004. **14**(10A): p. 1861-1869.

200. Brázda, V., et al., *Global analysis of inverted repeat sequences in human gene promoters reveals their non-random distribution and association with specific biological pathways.* Genomics (San Diego, Calif.), 2020. **112**(4): p. 2772-2777.

201. Ganapathiraju, M.K., et al., *A reference catalog of DNA palindromes in the human genome and their variations in 1000 Genomes.* Human genome variation, 2020. **7**(1): p. 40-40.

202. Lu, L., et al., *The human genome-wide distribution of DNA palindromes.* Functional & integrative genomics, 2007. **7**(3): p. 221-227.

203. Catasti, P., et al., *Unusual Structures of the Tandem Repetitive DNA Sequences Located at Human Centromeres.* Biochemistry (Easton), 1994. **33**(13): p. 3819-3830.

204. Gordenin, D.A., et al., *Inverted DNA repeats: a source of eukaryotic genomic instability.* Molecular and Cellular Biology, 1993. **13**(9): p. 5315-5322.

205. Lu, S., et al., *Short Inverted Repeats Are Hotspots for Genetic Instability: Relevance to Cancer Genomes.* Cell reports (Cambridge), 2015. **10**(10): p. 1674-1680.

206. Tanaka, H., et al., *Short Inverted Repeats Initiate Gene Amplification through the Formation of a Large DNA Palindrome in Mammalian Cells.* Proceedings of the National Academy of Sciences - PNAS, 2002. **99**(13): p. 8772-8777.

207. Voineagu, I., et al., *Replication stalling at unstable inverted repeats: Interplay between DNA hairpins and fork stabilizing proteins.* Proceedings of the National Academy of Sciences, 2008. **105**(29): p. 9936-9941.

208. Lobachev, K.S., et al., *Factors affecting inverted repeat stimulation of recombination and deletion in Saccharomyces cerevisiae.* Genetics (Austin), 1998. **148**(4): p. 1507-1524.

209. Mitsuhashi, S., M.C. Frith, and N. Matsumoto, *Genome-wide survey of tandem repeats by nanopore sequencing shows that disease-associated repeats are more polymorphic in the general population.* BMC medical genomics, 2021. **14**(1): p. 17-17.

210. Trost, B., et al., *Genome-wide detection of tandem DNA repeats that are expanded in autism.* Nature, 2020. **586**(7827): p. 80-86.

211. Willems, T., et al., *Population-Scale Sequencing Data Enable Precise Estimates of Y-STR Mutation Rates.* American journal of human genetics, 2016. **98**(5): p. 919-933.

212. Fondon, r.J.W. and H.R. Garner, *Molecular Origins of Rapid and Continuous Morphological Evolution.* Proceedings of the National Academy of Sciences - PNAS, 2004. **101**(52): p. 18058-18063.

213. Jakubosky, D., et al., *Properties of structural variants and short tandem repeats associated with gene expression and complex traits.* Nature communications, 2020. **11**(1): p. 2927-2927.

214. Jansen, A., et al., *Intragenic tandem repeats generate functional variability.* Nature genetics, 2005. **37**(9): p. 986-990.

215. Fardaei, M., et al., *Three proteins, MBNL, MBLL and MBXL, co-localize in vivo with nuclear foci of expanded-repeat transcripts in DM1 and DM2 cells.* Human Molecular Genetics, 2002. **11**(7): p. 805-814.

216. Cleary, J.D., A. Pattamatta, and L.P.W. Ranum, *Repeat-associated non-ATG (RAN) translation.* The Journal of biological chemistry, 2018. **293**(42): p. 16127-16141.

217. Jafar-Nejad, P., et al., *Opposing effects of polyglutamine expansion on native protein complexes contribute to SCA1.* Nature (London), 2008. **452**(7188): p. 713-718.

218. Nasrallah, M.P., et al., *Differential effects of a polyalanine tract expansion in Arx on neural development and gene expression.* Human molecular genetics, 2012. **21**(5): p. 1090-1098.

219. Liquori, C.L., et al., *Myotonic dystrophy type 2 caused by a CCTG expansion in intron 1 of ZNF9.* Science, 2001. **293**(5531): p. 864-7.

220. Sato, N., et al., *Spinocerebellar Ataxia Type 31 Is Associated with "Inserted" Penta-Nucleotide Repeats Containing (TGGAA)n.* The American Journal of Human Genetics, 2009. **85**(5): p. 544-557.

221. DeJesus-Hernandez, M., et al., *Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS.* Neuron, 2011. **72**(2): p. 245-256.

222. Zain, R. and C.I.E. Smith, *Targeted Oligonucleotides for Treating Neurodegenerative Tandem Repeat Diseases.* Neurotherapeutics, 2019. **16**(2): p. 248-262.

223. Starr, E., et al., *The relationship between trinucleotide (CAG) repeat length and clinical features of Huntington's disease.* Nature genetics, 1993. **4**(4): p. 398-403.

224. Monrós, E., et al., *Phenotype Correlation and Intergenerational Dynamics of the Friedreich Ataxia GAA Trinucleotide Repeat.* American journal of human genetics, 1997. **61**(1): p. 101-110.

225. Pratte, A., et al., *Anticipation in myotonic dystrophy type 1 parents with small CTG expansions.* American journal of medical genetics. Part A, 2015. **167A**(4): p. 708-714.

226. Genetic Modifiers of Huntington's Disease Consortium. Electronic address, g.h.m.h.e. and C. Genetic Modifiers of Huntington's Disease, *CAG Repeat Not Polyglutamine Length Determines Timing of Huntington's Disease Onset.* Cell, 2019. **178**(4): p. 887-900 e14.

227. Santoro, M., et al., *Myotonic dystrophy type 1: role of CCG, CTC and CGG interruptions within DMPK alleles in the pathogenesis and molecular diagnosis.* Clin Genet, 2017. **92**(4): p. 355-364.

228. Reiss, A.L., et al., *Length of uninterrupted CGG repeats determines instability in the FMR1 gene.* Nature genetics, 1994. **8**(1): p. 88-94.

229. Du, J., et al., *Length-dependent CTG.CAG triplet-repeat expansion in myotonic dystrophy patient-derived induced pluripotent stem cells.* Hum Mol Genet, 2013. **22**(25): p. 5276-87.

230. Holmans, P.A., T.H. Massey, and L. Jones, *Genetic modifiers of Mendelian disease: Huntington's disease and the trinucleotide repeat disorders.* Human molecular genetics, 2017. **26**(R2): p. R83-R90.

231. *A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. The Huntington's Disease Collaborative Research Group.* Cell, 1993. **72**(6): p. 971-83.

232. Snowden, J.S., *The Neuropsychology of Huntington's Disease.* Archives of Clinical Neuropsychology, 2017. **32**(7): p. 876-887.

233. Marder, K., et al., *Rate of functional decline in Huntington's disease.* Neurology, 2000. **54**(2): p. 452-458.

234. Barnat, M., et al., *Huntington's disease alters human neurodevelopment.* Science, 2020: p. eaax3338.

235. Rué, L., et al., *Targeting CAG repeat RNAs reduces Huntington's disease phenotype independently of huntingtin levels.* The Journal of clinical investigation, 2016. **126**(11): p. 4319-4330.

236. Creus-Muncunill, J., et al., *Huntington's disease brain-derived small RNAs recapitulate associated neuropathology in mice.* Acta Neuropathol, 2021.

237. Sathasivam, K., et al., *Aberrant splicing of HTT generates the pathogenic exon 1 protein in Huntington disease.* Proceedings of the National Academy of Sciences - PNAS, 2013. **110**(6): p. 2366-2370.

238. Neueder, A., et al., *The pathogenic exon 1 HTT protein is produced by incomplete splicing in Huntington's disease patients.* Sci Rep, 2017. **7**(1): p. 1307.

239. Banez-Coronel, M., et al., *RAN Translation in Huntington Disease.* Neuron, 2015. **88**(4): p. 667-77.

240. Yang, S., et al., *Lack of RAN-mediated toxicity in Huntington's disease knock-in mice.* Proceedings of the National Academy of Sciences - PNAS, 2020. **117**(8): p. 4411-4417.

241. Losekoot, M., et al., *EMQN/CMGS best practice guidelines for the molecular genetic testing of Huntington disease.* European journal of human genetics : EJHG, 2013. **21**(5): p. 480-486.

242. Jarem, D.A. and S. Delaney, *Premutation huntingtin allele adopts a non-B conformation and contains a hot spot for DNA damage.* Biochemical and biophysical research communications, 2011. **416**(1): p. 146-152.

243. Kremer, H.P.H., et al., *Sex-dependent mechanisms for expansions and contractions of the CAG repeat on affected Huntington disease chromosomes.* American journal of human genetics, 1995. **57**(2): p. 343-350.

244. Swami, M., et al., *Somatic expansion of the Huntington's disease CAG repeat in the brain is associated with an earlier age of disease onset.* Human molecular genetics, 2009. **18**(16): p. 3039-3047.

245. Lee, J.M., et al., *CAG repeat expansion in Huntington disease determines age at onset in a fully dominant fashion.* Neurology, 2012. **78**(10): p. 690-695.

246. Lee, J.-M., et al., *Identification of Genetic Factors that Modify Clinical Onset of Huntington's Disease.* Cell, 2015. **162**(3): p. 516-526.

247. Goold, R., et al., *FAN1 modifies Huntington's disease progression by stabilizing the expanded HTT CAG repeat.* Hum Mol Genet, 2019. **28**(4): p. 650-661.

248. Tomé, S., et al., *MSH3 polymorphisms and protein levels affect CAG repeat instability in Huntington's disease mice.* PLoS genetics, 2013. **9**(2): p. e1003280.

249. Pinto, R.M., et al., *Mismatch repair genes Mlh1 and Mlh3 modify CAG instability in Huntington's disease mice: genome-wide and candidate approaches.* PLoS genetics, 2013. **9**(10): p. e1003930.

250. Schmidt, M.H.M. and C.E. Pearson, *Disease-associated repeat instability and mismatch repair.* DNA repair, 2016. **38**: p. 117-126.

251. Boudreau, R.L., et al., *Nonallele-specific silencing of mutant and wild-type huntingtin demonstrates therapeutic efficacy in Huntington's disease mice.* Mol Ther, 2009. **17**(6): p. 1053-63.

252. McBride, J.L., et al., *Preclinical safety of RNAi-mediated HTT suppression in the rhesus macaque as a potential therapy for Huntington's disease.* Mol Ther, 2011. **19**(12): p. 2152-62.

253. Burrus, C.J., et al., *Striatal Projection Neurons Require Huntingtin for Synaptic Connectivity and Survival.* Cell Rep, 2020. **30**(3): p. 642-657 e6.

254. Colin, E., et al., *Huntingtin phosphorylation acts as a molecular switch for anterograde/retrograde transport in neurons.* The EMBO journal, 2008. **27**(15): p. 2124-2134.

255. Timmusk, T., et al., *Huntingtin interacts with REST/NRSF to modulate the transcription of NRSE-controlled neuronal genes.* Nature genetics, 2003. **35**(1): p. 76-83.

256. Marcora, E. and M.B. Kennedy, *The Huntington's disease mutation impairs Huntingtin's role in the transport of NF-κB from the synapse to the nucleus.* Human molecular genetics, 2010. **19**(22): p. 4373-4384.

257. Yang, S., et al., *CRISPR/Cas9-mediated gene editing ameliorates neurotoxicity in mouse model of Huntington's disease.* J Clin Invest, 2017. **127**(7): p. 2719-2724.

258. Zeitler, B., et al., *Allele-selective transcriptional repression of mutant HTT for the treatment of Huntington's disease.* Nat Med, 2019. **25**(7): p. 1131-1142.

259. Fink, K.D., et al., *Allele-Specific Reduction of the Mutant Huntingtin Allele Using Transcription Activator-Like Effectors in Human Huntington's Disease Fibroblasts.* Cell transplantation, 2016. **25**(4): p. 677-686.

260. Nakamori, M., et al., *A slipped-CAG DNA-binding small molecule induces trinucleotide-repeat contractions in vivo.* Nat Genet, 2020. **52**(2): p. 146-159.

261. PTC Therapeutics, I., *PTC Therapeutics Announces that PTC518 Has Entered into a Phase 1 Clinical Trial for the Huntington's Disease Program*

. 2020: CISION PR Newswire.

262. Kwon, D., *Failure of genetic therapies for Huntington's devastates community.* Nature, 2021.

263. Rodrigues, F.B. and E.J. Wild, *Huntington's Disease Clinical Trials Corner: April 2020.* J Huntingtons Dis, 2020. **9**(2): p. 185-197.

264. Kordasiewicz, H.B., et al., *Sustained therapeutic reversal of Huntington's disease by transient repression of huntingtin synthesis.* Neuron, 2012. **74**(6): p. 1031-44.

265. Evers, M.M., et al., *AAV5-miHTT Gene Therapy Demonstrates Broad Distribution and Strong Human Mutant Huntingtin Lowering in a Huntington's Disease Minipig Model.* Mol Ther, 2018. **26**(9): p. 2163-2177.

266. Spronck, E.A., et al., *AAV5-miHTT Gene Therapy Demonstrates Sustained Huntingtin Lowering and Functional Improvement in Huntington Disease Mouse Models.* Molecular Therapy - Methods & Clinical Development, 2019. **13**: p. 334-343.

267. Martí, E., *RNA toxicity induced by expanded CAG repeats in Huntington's disease.* Brain Pathol, 2016. **26**(6): p. 779-786.

268. Zaghloul, E.M., et al., *CTG repeat-targeting oligonucleotides for down-regulating Huntingtin expression.* Nucleic Acids Research, 2017.

269. Mozafari, N. and T. Umek, *Assessing Oligonucleotide Binding to Double-Stranded DNA.* Methods in molecular biology (Clifton, N.J.), 2019. **2036**: p. 91.

270. Vogt, V.M., *Purification and Further Properties of Single-Strand-Specific Nuclease from Aspergillus oryzae.* European journal of biochemistry, 1973. **33**(1): p. 192-200.

271. Christophe, E., et al., *Rational design of a triple helix- specific intercalating ligand.* Proceedings of the National Academy of Sciences of the United States of America, 1998. **95**(7): p. 3591.

272. Zaid, A., et al., *Triple-Helix Directed Cleavage of Double-Stranded DNA by Benzoquinoquinoxaline-1,10-phenanthroline Conjugates.* ChemBioChem, 2004. **5**(11): p. 1550-1557.

273. Shi, Y., P. Kirwan, and F.J. Livesey, *Directed differentiation of human pluripotent stem cells to cerebral cortex neurons and neural networks.* Nat Protoc, 2012. **7**(10): p. 1836-46.

274. Avior, Y., I. Sagi, and N. Benvenisty, *Pluripotent stem cells in disease modelling and drug discovery.* Nature reviews. Molecular cell biology, 2016. **17**(3): p. 170-182.

275. Stein, C.A., et al., *Efficient gene silencing by delivery of locked nucleic acid antisense oligonucleotides, unassisted by transfection reagents.* Nucleic Acids Research, 2010. **38**(1): p. e3.

276. Fisher, T.L., et al., *Intracellular disposition and metabolism of fluorescently-labeled unmodified and modified oligonucleotides microinjected into mammalian cells.* Nucleic Acids Res, 1993. **21**(16): p. 3857-3865.

277. Mirzadeh, Z., et al., *Neural stem cells confer unique pinwheel architecture to the ventricular surface in neurogenic regions of the adult brain.* Cell Stem Cell, 2008. **3**(3): p. 265-78.

278. Povsic, T.J. and P.B. Dervan, *Triple helix formation by oligonucleotides on DNA extended to the physiological pH range.* Journal of the American Chemical Society, 1989. **111**(8): p. 3059-3061.

279. James, P.L., T. Brown, and K.R. Fox, *Thermodynamic and kinetic stability of intermolecular triple helices containing different proportions of C+·GC and T·AT triplets.* Nucleic acids research, 2003. **31**(19): p. 5598-5606.

280. Yakovchuk, P., E. Protozanova, and M.D. Frank-Kamenetskii, *Base-stacking and base-pairing contributions into thermal stability of the DNA double helix.* Nucleic acids research, 2006. **34**(2): p. 564-574.

281. Pabon-Martinez, Y.V., et al., *LNA effects on DNA binding and conformation: from single strand to duplex and triplex structures.* Sci Rep, 2017. **7**(1): p. 11043.

282. Boczkowska, M., P. Guga, and W.J. Stec, *Stereodefined Phosphorothioate Analogues of DNA: Relative Thermodynamic Stability of the Model PS-DNA/DNA and PS-DNA/RNA Complexes.* Biochemistry (Easton), 2002. **41**(41): p. 12483-12487.

283. Grondin, R., et al., *Six-month partial suppression of Huntingtin is well tolerated in the adult rhesus striatum.* Brain, 2012. **135**(Pt 4): p. 1197-209.

284. Kinniburgh, A.J., *A cis- acting transcription element of the c- myc gene can assume an H- DNA conformation.* Nucleic Acids Research, 1989. **17**(19): p. 7771.

285. Cooney, M., et al., *Site- specific oligonucleotide binding represses transcription of the human c- myc gene in vitro.* Science, 1988. **241**(4864): p. 456.