Locked nucleic acid: modality, diversity, and drug discovery

Hagedorn, Peter H.; Persson, Robert; Funder, Erik D.; Albæk, Nanna; Diemer, Sanna L.; Hansen, Dennis J.; Møller, Marianne R; Papargyri, Natalia; Christiansen, Helle; Hansen, Bo R.

Published in:
Drug Discovery Today

Link to article, DOI:
10.1016/j.drudis.2017.09.018

Publication date:
2018

Document Version
Peer reviewed version

Link back to DTU Orbit

Citation (APA):
Locked nucleic acid: modality, diversity, and drug discovery

Peter H. Hagedorn¹, Robert Persson², Erik D. Funder¹, Nanna Albæk¹, Sanna L. Diemer¹, Dennis J. Hansen¹, Marianne R. Møller¹, Natalia Papargyri³, Helle Christiansen¹, Bo R. Hansen¹, Henrik F. Hansen¹, Mads A. Jensen¹ and Troels Koch¹

¹Therapeutic Modalities, Roche Pharma Research and Early Development, Roche Innovation Center Copenhagen, 2970 Hørsholm, Denmark
²Pharmaceutical Sciences, Roche Pharma Research and Early Development, Roche Innovation Center Copenhagen, 2970 Hørsholm, Denmark
³Department of Biotechnology and Biomedicine, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark

Over the past 20 years, the field of RNA-targeted therapeutics has advanced based on discoveries of modified oligonucleotide chemistries, and an ever-increasing understanding of how to apply cellular assays to identify oligonucleotides with improved pharmacological properties in vivo. Locked nucleic acid (LNA), which exhibits high binding affinity and potency, is widely used for this purpose. Our understanding of RNA biology has also expanded tremendously, resulting in new approaches to engage RNA as a therapeutic target. Recent observations indicate that each oligonucleotide is a unique entity, and small structural differences between oligonucleotides can often lead to substantial differences in their pharmacological properties. Here, we outline new principles for drug discovery exploiting oligonucleotide diversity to identify rare molecules with unique pharmacological properties.

Introduction
The flow of genetic information is a highly complex process. Of the approximately 43 000 human genes currently annotated by the Ensembl project (release 89), approximately 47% encodes proteins, and the rest are all transcribed into different types of noncoding RNA (ncRNA) [1]. It is clear that these ncRNAs represent a long and growing list of different types of RNA with various functional roles and regulatory interactions [2,3]. This knowledge has provided a stronger basis for the specific annotation of the relationships between structure, type, and function of RNAs and human disease [4]. This is one reason for the increased interest in RNA

Corresponding author: Koch, T. (troels.koch@roche.com)

1339-6446/© 2017 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0).

https://doi.org/10.1016/j.drudis.2017.09.018

Please cite this article in press as: Hagedorn, P.H. et al.: Locked nucleic acid: modality, diversity, and drug discovery, Drug Discov Today (2017), https://doi.org/10.1016/j.drudis.2017.09.018
therapeutics, which focuses not only on classic protein-coding mRNA intervention, but also on how the manipulation of non-coding regulatory RNA may provide novel drugs for previously untreatable diseases.

A rational strategy to target RNA is by synthetic oligonucleotides. Although RNA exists in a complex biological matrix, endogenous enzymatic RNA-processing mechanisms can be exploited and recruited by the heteroduplex formation between RNA and synthetic oligonucleotides. RNA exhibits high turnover rates and is more labile than DNA [5]; thus, inhibitory effects induced by the oligonucleotide binding complement can be executed rapidly and effectively. Zamecnik and Stevenson [6] were the first to show that a single-stranded synthetic DNA oligonucleotide could inhibit the expression of RNA, resulting in the stimulation of substantial research and development activity [7–13].

The past four decades of research and development in oligonucleotide-based RNA therapeutics has produced groundbreaking results. The recent successes are based on improvements in three essential parameters for antisense technologies: (i) medicinal chemistry efforts have devised much-improved nucleic acid modifications for antisense oligonucleotides (AONs), of which the high-affinity LNA [14–17] is among the most widely used; (ii) the knowledge revolution of the functions and biological molecular mechanisms of RNA; and (iii) the translational drug discovery part (i.e., discovery processes and models used in drug discovery). The latter relates to bioinformatics drug design, prediction algorithms, in vitro/vivo assays, better pharmacokinetics/pharmacodynamics (PK/PD) models, absorption, distribution, metabolism, and excretion (ADME) models, and toxicology study designs.

Recent developments have also provided insights into the fundamental properties of the chemical modality. Small chemical modifications of AONs have been shown to produce large property differences even among AONs belonging to the same chemical or structural class. The old concept of the ‘portability of chemistry’ for AONs, whereby nearly all properties are shared for a given chemical class, has turned out not to be true for many important drug properties (e.g., activity and toxicity). In other words, we now know that the structure–activity relationships (SARs) for AONs are more subtle. The fact that AONs belonging to a given structural class only on a few parameters share the same properties and act more as individual unique compounds diversifies and complicates drug discovery [13,18–20]. However, the fact that small chemical modifications, even the configuration of a single bond, can have profound property impacts creates a basis for SAR studies comparable to classic small-molecule drug discovery. Therefore, during the discovery phase, it is now possible to produce more unique compounds for nearly any RNA target with larger therapeutic indexes (TI).

In this review, we focus on three aspects of LNA: (i) state of the art of the modality, including structure, fundamental properties, preferred designs, mechanisms of action, and cellular uptake; (ii) illustrate how subtle structural modifications have profound impacts and, in some cases, produce ‘all or none’ phenotypic effects; and (iii) how the insight or ‘Erkenntnis’ of property diversity can be used to transform and improve the classic oligonucleotide drug discovery process and lead generation.

Chemistry sets boundaries for success
In the first generation of single-stranded therapeutic oligonucleotides, the phosphorothioate (PS) internucleoside linkage was the only chemical modification [8,21–24]. The PS modification replaces one of the nonbridging oxygen atoms with sulfur in the internucleoside phosphate (Fig. 1ai). This modification creates a chiral center at phosphorous producing two isomers: Rp and Sp. Thus, for every PS linkage introduced in an oligonucleotide, two diastereoisomers are formed. Given that conventional solid-phase PS synthesis is not stereoselective, an n-mer PS oligonucleotide contains random mixtures of $2^{n-1}$ diastereoisomers [25–30]. However, recent developments have now made it possible to synthesize individual stereoisomers with high stereoselectivity [31–34], and with good yields [35–37]. Diastereoisomers are different chemical compounds and the diversity of properties for a given PS oligonucleotide will be a gradient spanning from a little to a very large difference in property. Generally, oligonucleotides with a Sp configuration provided better exonuclease resistance compared with oligonucleotides with a Rp configuration, whereas Rp isomers were better substrates for DNA-dependent RNA polymerases, RNase H, and stimulated immune responses [38–40]. Compared with the random mixtures, Rp oligonucleotides exhibited higher melting temperature ($T_m$) against RNA, whereas Sp analogs had lower $T_m$. In the first-generation oligonucleotide drugs, the PS modification was normally not stereocontrolled. Except for rare examples [41], this class of drugs has not provided an adequate TI for AONs [42]. Two reasons for this were that the PS modification decreases the $T_m$ of the hybrid duplex and, although it protects against nucleic degradation to some extent, this protection was still not adequate on its own for most therapeutic purposes [43]. Still, the PS modification, stereocontrolled or not, turned out to be a central modification for later generations. Combined with the increased nuclease stability, PSs also contribute to improved PK and cellular uptake properties of AONs. Unless otherwise stipulated, all antisense data described here come from fully PS-modified AONs.

A preferred strategy to improve the binding affinity of AONs to RNA was to introduce electronegative substituents in the 2'-position of the furanose (Fig. 1aii) [7,8,44–47]. For nucleic acid analogs of this structural class, melting temperatures increased by approximately 0.5–1.5 °C/modification against RNA. The most significant members of this class are the 2’-F, 2’-O-CH$_3$, and 2’-O-CH$_2$CH$_2$O-CH$_3$ (MOE) substituent groups [44,46]. This second generation of drugs shows clear improvements compared with first-generation PS drugs, and two drugs of this type, mipomersen [48] and nusinersen [49], have now been approved by the US Food and Drug Administration (FDA).

A high point for synthetic nucleic acid analogs was reached in 1997 with LNA [15,50] (Fig. 1a(iii)). LNA exhibited unprecedented high RNA-binding affinity [51], providing melting temperature increases of 3–9 °C/modification depending on the oligonucleotide position and design complements [52]. The affinity increase per LNA nucleoside substitution is a reproducible (‘portable’) property [14,51–55] and, combined with the increased nuclease stability, the driver behind the many therapeutic and diagnostic life-science applications that have been demonstrated with, and published on, LNAs [56–63]. Given these significant improvements, LNA is now widely used and acknowledged by many to
belong to a third generation of chemical modifications used in RNA therapeutics.

‘LNA’ refers to the bicyclic structure in which the flexibility of the parent furanose has been locked [15] (Fig. 1aiv). The bicyclic structure of LNA nucleosides is the key to the improved binding properties, and has served as scaffold for many other bicyclic analogs also exhibiting improved binding properties [64–68]. The 2'-O-CH₂-4' linkage is transformational and converts a flexible furanose into a rigid bicycle [69–74]. One thermodynamic consequence of the ‘locking’ compared with native deoxyribose is that the free energy of solvation is 1–2 kcal/mol lower and, thus, more favorable because of the lipophilic nature of the 2'-O-CH₂-4' linkage [75].

The difference between LNA and 2'-OMe RNA, for example, can be observed in the helical structures that they form. Recent structural data indicate that LNA/RNA duplexes perturb the A-type helix, whereas 2'-OMe RNA/RNA attain a more classic A-type structure [76]. LNA/LNA duplexes result in an extreme ‘underwinded’ helix in which the major grove is increased by a dramatic decrease of the slide-and-twist values and a pitch value of 39 Å compared with 29 Å for RNA/RNA. There is a consecutive order of structural change of the classical A-type helix from RNA/RNA, RNA/LNA to LNA/LNA [76,77].

Transforming a flexible furanose moiety into a rigid bicycle reduces the number of conformational degrees of freedom, including rotations around internal bonds [75]. Structurally, this
preorganizes the molecule and reduces the hybridization entropy penalty. Although a net reduction in the entropy change for hybridization (ΔΔS) has been reported as a main driver for the high affinity [78], the situation is more complex. Given that LNA distorts the classic A-form helix, enthalpic contributions, such as hydrophobic interactions and II–II-stacking between the bases, are also increased [77].

LNA designs and RNA intervention mechanisms

LNA designs can be divided in two main categories: mixmers and gapmers (Fig. 1b). In a mixmer, LNA and DNA nucleosides (dark-orange and yellow boxes, respectively in Fig. 1b) are interpersed throughout the sequence of the oligonucleotide, whereas, in a gapmer, two LNA segments at both ends of the oligonucleotide are separated by a central segment or gap of DNA nucleosides. At present, LNA designs are usually made with fully modified non-stereodefined PS internucleoside linkages (red lines in Fig. 1b). Unless otherwise specified, all data mentioned here are obtained from such LNA PS random mixtures.

To inhibit mRNA expression and protein translation (Fig. 1c, mechanism 1), gapmer designs are the most potent. This is because the central DNA/PS segment, which is longer than 7–8 DNA nucleotides (nt), recruits the RNA-cleaving enzyme RNase H when the gapmer is hybridized to the mRNA [79–84]. Gapmers targeting both exons and introns work equally efficiently and, thus, all stages of mRNA and also ncRNA processing can be addressed. Nuclease recruitment by gapmers occurs in both the cytoplasm and the nucleus [85,86], although predominantly in the nucleus [80] (Fig. 1c, mechanism 2). Mixmers can also be used to block translation, but must be designed to bind close to, at, or upstream of the translational start site to do so [83,84] (Fig. 1c, mechanism 3). The mixmer will inhibit translation by blocking the binding of the ribosomal subunits to the mRNA. Alternatively, mixmers binding close to the 5’ end of the pre-mRNA can prevent 5’-cap formation and thereby inhibit translation (Fig. 1c, mechanism 4).

There is sometimes the need to intervene with RNA processing and information flow without cleaving or degrading the target, or inhibiting protein translation. For instance, in the case of splice switching/redirection, or blocking natural antisense or other kinds of ncRNA, mixmers are often the preferred option [87–93] (Fig. 1c, mechanism 5). A mixmer cannot support RNA cleavage but acts as an efficient steric block to mediate a phenotype without destroying the target RNA. This intervention changes the expression profile, creating potentially both ‘loss-of-function’ and ‘gain-of-function’ phenotypes [91–93].

Inhibition, or sequestering, of mRNA is a gain of function that has attracted a special interest [94–97]. miRNAs are an important class of regulatory ncRNAs that can bind to partially complementary sites located in the 3’ untranslated regions (UTRs) of target mRNAs, promoting translational repression or deadenyluation and degradation (Fig. 1c, mechanism 6). Sequestering miRNAs by LNA mixmer hybridization has wide therapeutic potential [98] (Fig. 1c, mechanism 7). The miR-122-antagonizing LNA mixmer ‘miraviren’ was progressed to Phase 2 clinical trials and exhibited safe and potent HCV viral titer reductions [97]. The high binding constant of LNA mixmers produces a slow off-rate, which is the critical parameter for efficient miRNA antagonism. Combined with the long tissue half-life of LNA, the slow off-rate produces a long half-life of the LNA/miRNA heteroduplex, leading to long lasting derepression of the proteins that are under the control of the miRNA.

Affinity, potency, and specificity of LNA oligonucleotides

It is well documented that the high RNA-binding affinity of LNA is linked to its potency [58,57,99]. The term ‘potency’ is used here for the activity per administered dose and is not necessarily related to the concentration in cells or in tissues. Although high target affinity is a driver behind higher potency [14,18], many other factors are also determinants, and higher affinity can be balanced out by other potency-limiting parameters. Examples of such limiting factors are: design, RNase H recruitment [100], tissue/cellular uptake, and/or tissue distribution. For instance, it has been shown that a fully LNA-modified 14-mer targeting the coding region of a transcript did not inhibit protein synthesis [84,101], because the ribosome ‘read through’ the binding site of the LNA, despite the high RNA affinity. When some of the central LNA nt were substituted with DNA-nt so that the 14-mer became a LNA gapmer, the gapmer recruited RNase H and mediated cleavage and degradation of the message, despite the lower affinity [84].

It has also been shown that affinity increase by length expansion of gapmers only improved the potency to a certain compound-specific point/threshold, after which the potency decreased as the length, and affinity, increased [99]. This ‘threshold affinity’ was studied further by Pedersen et al. using a kinetic model to simulate LNA gapmer recruitment by RNase H [100]. The model showed very fast LNA–RNA hybridization on-rate and recruitment of RNase H. IC50 curves were generated as a function of LNA concentrations and, when these were plotted as a function of LNA–RNA affinity, a parabola emerged. This illustrated in affinity terms that potency increased as the free energy of hybridization decreased to approximately ∆ΔG* = 20–25 kcal/mol, but below this point, and for compounds with even higher affinity, the effect reversed and potency decreased. The existence of such ‘optimal’ affinity was explained to be a tradeoff of initial efficient target binding coupled with inefficient deassociation of the cleaved LNA–RNA duplex. Nonspecific sequestration of LNAs through protein binding might also be length dependent because of the negative charge contributed by each PS linkage, and might also contribute to the observed length-dependent potency of LNAs [99].

It is a cornerstone in RNA-based drug discovery to ensure that oligonucleotides interact specifically with their intended RNA targets. For LNA gapmers that bind with high affinity to fully complementary RNA target regions, extra care must be taken to ensure that they do not bind efficiently to partially mismatched, unintended RNA targets [18]. Although the mismatched base pairs in the unintended RNAs reduce the binding affinity compared with the fully matched target, it has been reported that, in some instances, affinity of LNAs can be designed so high that duplexes between mismatched unintended RNAs and gapmers persist long enough, and in sufficient amounts, to allow effective RNase H cleavage [18,102–104]. However, character- and thermodynamics-based bioinformatics algorithms allow for computational gapmer sequence analysis where the likelihood of binding to mismatched unintended RNA can be predicted. Thus, bioinformatics enable the selection of sequence-specific oligonucleotides designed to
have either no or only a few mismatched hybridizations. The sequence specificity of gampers can also be evaluated experimentally using transcriptome-wide profiling approaches, such as microarrays or RNA-seq [18]. Taken together, these computational and experimental methods are important tools to identify compounds with high sequence specificity and low off-target potential [105].

Cellular uptake
Entry of oligonucleotides into cells is a complex process and a comprehensive report is beyond the scope of this review. Here, we focus on central elements that are relevant for LNAs, but the interested reader may consult recent reviews for more information [106–109].

When oligonucleotides are transfected with cationic lipids, the lipids fuse with the plasma membrane and deliver oligonucleotides directly into the cytoplasm. However, the oligonucleotides concentrate quickly in the nucleus, which is advantageous for antisense because RNase H is predominantly found in the nucleus [80]. When oligonucleotides are incubated without added transfection agents, ‘naked’ uptake also occurs [106,110–112]. It was illustrated that ‘naked’ LNAs were taken up and exhibited potent activity in cell cultures [113]. This process was named ‘Gymnosis’ after the Greek ‘gymnos’ or ‘naked’, and activity was found for nearly all cell types, including ‘hard-to-transfect’ cells, such as suspension cell lines and T cells. Gymnastic silencing provides more AON-specific phenotypes. Given that transfection/delivery lipids are not used in gymnastic assays, a cellular response is only related to, or is a consequence of, the added AON [113–115].

Cellular uptake of naked oligonucleotides and LNA is predominantly driven by endocytosis. Many different pathways and cellular proteins are engaged in the uptake, such as clathrin, dynamin, and caveolar-dependent endocytosis [116,117]. However, clathrin-independent endocytosis also exists [118], and it has been shown that antisense activity in hepatocytes can be blocked by adaptor protein AP2M1 but not by clathrin and caveolin [119].

It was recently shown that LNA gampers were taken up in gymnastic assays in human primary T cells via macropinocytosis (MP) [120]. MP is a fluid-phase endocytosis pathway driven by actin–myosin cell protrusions creating larger vesicles. The fate of macropinosomes differs from that of other endosomes and they are directed toward the late-stage endosomes/lysosomes. Once taken up, LNAs were released from MPs/late endosomes and showed potent and specific RNA target knockdown.

The mechanisms by which oligonucleotides are released from endosomes/macropinosomes are still unclear, but maturation of the endosomes from early to late endosomes (multivesicular bodies) is important [85]. It was also shown the LNAs co-localize with various cellular bodies (GW and P-bodies) and that proteins, such as Ago2, TPC-, Hsc-70, and PKC-α proteins, which bind to endosomes, are involved in the endosomal maturation process and probably also in their release [121]. Once oligonucleotides have escaped from the endosomes, activity occurs in the cytoplasm [85] or, after rapid trafficking, also in the nucleus [122,123].

One of the most successful uptake enhancers is the triantennary N-acetylgalactosamine (GalNAc) ligand, which targets the high-density asialoglycoprotein receptor (ASGR) expressed on hepatocytes [108,124]. When oligonucleotides are conjugated to GalNAc, the high turnover of ASGR will internalize receptor-GalNAc-oligonucleotide efficiently via clathrin receptor-mediated endocytosis. GalNAc-conjugated oligonucleotides typically provide 10–20-fold reductions in the dose where half maximum responses were seen, and have revolutionized hepatic RNA targeting [108].

It is possible that the major uptake observed from the endosomal/MP pathways operates together with other uptake mechanisms. In this way, high uptake/low productivity pathways (endocytosis) might operate in parallel with low uptake/high productivity pathways. For example, it has been shown that nucleic acid channels exist in rat kidney and brain cells [125], although it has not been demonstrated that this is a general uptake mechanism. Cellular uptake is a highly diverse and differentiated process that is dependent on many parameters, such as cell type, proliferation/cell cycle state, extracellular composition (both in vitro and in vivo), oligonucleotide design/substitutions pattern, concentration, and phosphorothioate configuration (see below).

Diversity: an inherent attribute for oligonucleotides
Structural diversity
It is normal practice in the field of small-molecule drug discovery to produce collections or libraries of chemical entities. It is generally appreciated that library size is not a goal in itself, but rather the generation of structural, and thereby functional, diversity, and property refinement [126]. For oligonucleotides, structural differences also define functional diversities, but because of the particular compound class, the diversity framework is derived from different property determinants. The diversity framework can be divided in five components or determinants for structural diversity: (i) nucleobase or sequence diversity: the variation in nucleobase composition that constitutes oligonucleotide sequence; (ii) nucleoside diversity: the variation in chemical structure of the nucleoside building blocks used to produce a given oligonucleotide; (iii) backbone diversity: the variation in backbone internucleoside linkage structure, including variations in the chiral orientation of any phosphorothioate linkages; (iv) architectural diversity: the variation in design and combinatorial nucleotide make-up of an oligonucleotide; and (v) macromolecular structure diversity: the variation in structure and topology of the entire molecule.

The first four diversity determinants are classic oligonucleotide parameters. They represent the ‘digital’ Watson–Crick hybridization properties. The antisense literature has covered these in great detail over the past 30 years, but it was only recently appreciated that the nucleobase sequence is not the only determinant, but that the global properties are derived from the position of each structural unit in the context of the entire molecule. The term ‘sequence-related properties’ has been used frequently and some properties can be related to a nucleobase sequence motif alone. However, this description is at best inadequate, and can be meaningless without reference to the specific composition of the molecule.

Macromolecular structure diversity is related to the variation in overall structure and topology of the oligonucleotide. This parameter relates more to the ‘analog’ or aptamer binding properties. When the first four determinants are fixed, a resulting single molecule will be able to attain many different shapes and overall
structures each representing different energy levels. This diversity parameter can be illustrated by modeling, and quantum mechanical (QM) calculations have demonstrated the 3D relations in molecular structure, topology, and electrostatics of LNA PS oligonucleotides [127,128]. Among all possible macromolecular structures, QM or molecular dynamics calculations will be able to identify a specific minimized structure that is most representative for the molecule [75,127]. Below, we first describe functional diversity for the classic PS random mixtures, and then diversities for stereo-defined PS LNAs.

**Functional diversity in vitro**

Human pre-mRNA sequences are typically between 10 000 and 65 000 nt in length with a median length of approximately 25 000 nt [1]. Therefore, tens of thousands of possible sites on a pre-mRNA can be targeted by gapmers. Libraries of gapmers that are positioned at unique and nonoverlapping sites on a RNA target manifest diversity with respect to their nucleobase composition. It is widely recognized that gapmer potency is affected by the target nucleobase sequence. This is usually attributed to the structure of the RNA in the target region, where less-structured RNA leads to more-potent gapmer responses [129]. However, other factors that affect potency can also be attributed to the nucleobase sequence, such as sequence-specific intracellular localization [109].

Nucleobase diversity was illustrated in a study of 57 and 71 fully phosphorothioated gapmers (20mers, mixtures of 219 diastereoisomers). The gapmers were designed with five 2′-O-methoxyethyl (MOE) modifications in each flank (5–10–5 design). A range of activities was seen when screened in mouse primary hepatocytes and human T-24 cells [130]. Some gapmers were not active at all and showed no knockdown of the target. Most compounds showed some activity and a few in both libraries were judged to be potent.

In another study, more than 200-fold differences in potency in vitro were reported for a library of 47 LNA-phosphorothiate gapmers targeted to various regions of the HCV RNA genome [131]. Most of these nucleobase sequence-diverse gapmers were 16mers with four LNAs in each flank (a 4–8–4 design, mixtures of 215 diastereoisomers). Interestingly, when screening for general cytotoxic effects in a cell proliferation assay, the gapmers were observed to also demonstrate more than 200-fold differences in the concentrations at which half-maximal cytotoxicity was observed. These properties were not closely correlated, and the authors were able to select seven gapmers that were both well tolerated and highly potent in vitro. Out of the seven ‘hits’, the gapmer with the highest potency reduced antiviral activity to <5%. This target region was explored in detail by designing 21 additional LNA gapmers with a 4–8–4 design closely tiled across this region. Upon screening of this library, a range of activities was again observed. Some compounds did not affect HCV activity at all, but three gapmers, consecutively spaced only 1 nt apart, were even more potent than the original ‘lead’. The authors also briefly explored architectural diversity by changing the LNA-modification pattern of the gapmer with highest potency and designed an additional 47 gapmers. They observed that 4-8-4 and 3-10-3 designs had similar potency, whereas 5-6-5 and 2-12-2 designs were three and five times less potent, respectively. Notably, although the 4-8-4 and 3-10-3 designs were of similar potency in vitro, the 3-10-3 design had a twofold longer half-life in kidney after intravenous (i.v.) administration to mice [131].

An evaluation of both architectural and nucleoside diversity was reported by Stanton et al. [132]. An initial screening campaign identified three potent PS LNA-gapmers targeted to the Glucocorticoid Receptor mRNA. Each of the three gapmers was of the 3–8–3 design but with completely different nucleobase sequences. The authors designed a library of 109 gapmers where they varied lengths between 14 and 20 nt, and modified flanks with various numbers of LNA, MOE, 2′-OMe, and 2′F nt. Some gapmers contained both LNA and 2′-OMe, or LNA and 2′-F in the flanks. In vitro and in vivo potency were observed to be highly sensitive to small changes in length, nucleoside composition, and number of modifications in the flanks. Thus, the study demonstrated an apparently unpredictable and wide property diversity of oligonucleotides.

Another example of how architectural diversity can impact potency in vitro is shown in Fig. 2. Here, 32 iso-sequential LNA-gapmers targeting HIF1A, were screened for target knockdown at 5 μM in HeLa cells. The LNA design for these gapmers was LxxxDDDDDDxxL, where L stands for LNA, D for DNA, and x for either LNA or DNA, giving 32 different combinations. Measured knockdowns ranged from no effect to 80%, achieved by the compound with the design LLLLDDDDDDDL. The predicted Tm [100] of the designs ranged from 40°C to 60°C, and did not correlate with potency.

**Functional diversity in vivo**

Administered oligonucleotides will interact differentially at each of the steps they encounter on their way to the site of action. These steps include absorption at the site of administration and plasma protein binding in the circulation; the uptake in different tissues and protein binding at the cell surface; the uptake into cells; and their intracellular transport to the site of action. Under in vivo conditions, the five diversity parameters described above will also collectively govern the properties of the specific molecule. Functional diversity is very high for some properties and smaller for others.
Until recently, the least diversity had been observed in the systemic exposure to LNAs after subcutaneous administration. In circulation, LNAs are bound to plasma proteins to a high degree, primarily serum albumin, although the exposure is species dependent. Mice need doses ten times higher than humans per kg to reach the same exposure in plasma [133]. Such comparisons have been discussed for 2’-O-(2-methoxyethyl)-modified oligonucleotides [133] and are important for our understanding of species-specific effects. At therapeutically relevant plasma concentrations (up to 10 μg/ml), 95–99% of LNAs are bound to plasma proteins in the rat. In monkey and humans, a lower fraction is bound to plasma proteins (90–95%) and, in the mouse, even a lower fraction [134]. The free (unbound) fraction of oligonucleotide is secreted via the urine. Within the first 24 h post-dose, the plasma concentration is reduced by two to three orders of magnitude. This reduction results from the rapid redistribution of the oligonucleotide to different tissues, predominantly liver and kidney, followed by skin, bone marrow, spleen, uterus, lymph node, and lung [56,101].

Using a dosing schedule of 15 mg/kg on days 0, 3, 7, 10, and 14 of nine LNA oligonucleotides in mice (dosed, i.v.), a different uptake and accumulation was revealed in the liver (green bars in Fig. 3). The 16mer gapmer (3–10–3) oligonucleotides with different nucleobase sequences were designed against the same RNA target. At sacrifice on day 16, the liver concentration of some of these oligonucleotides was as low as 15 μg/g liver tissue, whereas others showed values of almost 80 μg/g liver tissue. These results are by no means unique. Other studies showed differences in the liver uptake of up to 25-fold. Uptake of LNA oligonucleotides in kidney and other tissues show similar diversities. A kidney: liver disposition ratio spanning over more than two orders of magnitude has been observed in mice (H.F. Hansen, pers. commun.).

Taken together, tissue uptake of LNAs shows a high degree of diversity. The question is whether the differences in tissue uptake among a series of compounds are correlated with target activity and adverse effect of the oligonucleotides. The short answer is no: there is no direct correlation between the tissue uptake across a panel of different compounds and activity. In the study illustrated in Fig. 3, the target knockdown (blue bars) and tissue uptake (green bars) varied fivefold. Again, there was no direct correlation between the liver uptake and target knockdown across the compounds. Thus, an oligonucleotide with a very low liver uptake could be very active.

The higher affinity of LNAs and LNA analogs has been reported to result in hepatotoxicity [135]. By contrast, it has also been reported that hepatotoxicity was significantly reduced by using shorter version of the LNAs (from 20mer to 14mer) [135]. One should be careful to calculate any property correlations in relation to a single determinant, such as length. In other words, does a property difference result from a length change or because different compounds are being compared? For instance, it was shown that a 13mer LNA was toxic, but, by adding just one LNA nt, the corresponding 14mer was nontoxic [19]. The specific composition and the nucleotide positioning (including LNA nt) is also a determinant because LNAs with the same nucleobase sequence and with different positions of the LNA in the flanks can have different hepatotoxic profiles [19,136]. These studies indicate that some sequence motifs influence the hepatotoxic potential. Recent experiments in mice suggest that the number of unintended RNA targets that are effectively reduced in the liver after systemic administration of high-affinity gapmers, such as LNA gapmers, are also correlated with the hepatotoxic potential of the gapmers [96–98]. High-affinity gapmers can also affect unintended partially mismatched RNA targets [137]. These results are consistent with a model relating increased unintended targeting with increased risk of some of these targets being involved in critical cellular functions leading to hepatotoxicity. Alternatively, it could be that the concentration of cleaved RNA in the hepatocytes can reach a level that is in itself toxic to the cells. That hepatotoxicity can be a consequence of off-target effects also helps explain why structural diversity affects the hepatotoxic potential. As discussed above, structural diversity clearly affects activity on the intended target and, therefore, it is reasonable to expect that structural diversity also can affect activity on off-targets.

Using the dosing schedule described above, five administrations over 15 days (every third to fourth day) of 15 mg/kg, and measuring alanine aminotransferase (ALT) in plasma/serum as a measure for hepatotoxicity, high hepatotoxicity was found for both oligonucleotides taken up at high and very low tissue concentrations (red bars in Fig. 3). Likewise, there are oligonucleotides that are taken up in high concentrations without being toxic. Taken together, these results suggest that the liver uptake for different oligonucleotides does not necessarily relate to activity or toxicity even for oligonucleotides that are seen as ‘similar’ (i.e., with same length, design, and chemistry).

It has been reported that some LNAs are toxic [19,102–104,131,132,135]. The toxicity cannot be caused by the LNA nucleosides per se, because there are also many reports of LNAs that are not toxic [19,97,99,102,103,131]. It has also been clearly demonstrated that toxicity can be mitigated by either increasing or
decreasing the number of LNA nucleotides or, for iso-sequential AONs, by changing the nucleoside composition (architectural diversity), or, as observed in Fig. 3, by changing the nucleobase sequence (nucleobase diversity). Irrespective of the underlying mechanism affected by these structural changes, it is apparent that diversity can be exploited to mitigate toxicity as well as optimizing other properties.

### Stereodefined LNA oligonucleotides

Nearly all AONs are synthesized without stereodefinition of the PS internucleoside linkage. Thus, for the LNA PSs described above, each ‘compound’ exists as a random mixture of diastereoisomers. Fundamentally, diastereoisomers are not identical and do have different properties. Moving away from the random mixtures and making single stereodefined PS AONs offers a new dimension to RNA therapeutics [32]. One aspect of this is that macromolecular diversity can now be studied, such as by computer modeling (see above). An interesting outcome of this was QM calculations illustrating that changing a single PS bond configuration can produce a macromolecular structure diversity to the same extent as, for example, nucleobase or nucleotide changes [127,128].

In an in vitro experiment, 236 stereodefined LNAs, randomly selected from 4096 possible isomers of a 13mer targeting human HIF1A, were tested. A difference in knockdown ranging from 0% to 77% was observed (Fig. 4a). The distribution of HIF1A mRNA levels after treatment with gapmer is represented in a histogram in Fig. 4b (gray bars). The solid line in Fig. 4b represents a normal distribution with mean and standard deviation estimated as the median (70%) and the median absolute deviation (22%) of the knockdown data. From this, it can be seen that the distribution of the knockdown data as represented in the histogram is well approximated by the normal distribution (Fig. 4b). The original random-mixture LNA gapmer reduced HIF1A mRNA to approximately 50%. If it can be assumed that the extent of the knockdown achieved by each of the 4096 isomers follows a normal distribution, as suggested from the 236 isomers measured and represented in Fig. 4b, we can infer that approximately 20% of all 4096 isomers will reduce HIF1A mRNA to a larger extent than the original random-mixture LNA gapmer.

In another experiment, six 13mer stereodefined ApoB-targeting LNAs were tested in mice. The LNAs were dosed i.v. at 1.0 mg/kg (Fig. 5). The stereodefined LNAs (A–F) downregulated ApoB mRNA from 87% to 20%, with the random-mixture LNA, RTR3833, reducing mRNA to approximately 50% (blue bars in Fig. 5). The uptake in liver spanned one order of magnitude (1.34–0.135 µg/g, green bars in Fig. 5). It was interesting to note that high uptake did not relate to the highest activity for the six stereodefined compounds (Pearson’s correlation r = 0.35 and P = 0.44; Fig. 5). Only small differences in ALT concentrations...
were seen for the stereodefined LNAs (A–F) and none of them were above twofold of the ALT concentration observed in saline-treated controls (red bars in Fig. 5). Consequently, none of the stereodefined LNAs displayed markedly increased hepatotoxicity. Thus, in line with the QM observations, in vitro knockdown and in vivo uptake and knockdown are highly diverse among diastereoisomers.

The observation that individual stereoisomers can have substantially improved pharmacological properties compared with a random mixture was also recently reported for the 5-10-5 MOE-gapmer mipomersen [33]. Mipomersen comprises 524 288 stereoisomers and targets APOB RNA. The authors demonstrated for a set of six different fully stereodefined versions of mipomersen that three were more stable than the random mixture in both rat liver homogenate and in rat serum, and two were cleaved faster by RNase H in vitro [33]. The stereoisomer that was both more stable and cleaved significantly faster was also tested in mice, and showed significantly longer-lasting lowering of serum ApoB protein levels compared with the random mixture when dosed intraperitoneally eight times at 10 mg/kg [33]. Interestingly, this stereoisomer had a triplet stereochiral motif, SSR, within the gap, which the authors argued could be particularly suited to promote target RNA cleavage by RNase H. The authors tested the code in stereodefined gapmers targeted to FOXO1 RNA and APOC3 RNA, and demonstrated an improvement compared with the random mixtures. However, this finding is in contrast to an earlier study, where, among a set of 31 stereoisomers, no clear differences in in vitro potency was observed between the stereoisomers containing the SSR code and those that did not [34]. More comprehensive analyses with larger sets of stereoisomers are needed to establish whether such a preferred triplet code is a general phenomenon.

The diversity data observed for stereodefined AONs also point the way to a novel understanding of how AON chemistry affects function. In charged AONs, the PS linkage is required for optimal antisense activity [23]. This is also true for AONs in which the PS linkage is not required for making them nuclease resistant. Mechanistically, the importance of the PS linkage is often assigned to the heavy atom effect of sulfur, which alters the sphere of hydration surrounding the molecule relative to the PO linkage. This, plus the increased polarizability of the PS linkage and its ability to form salt bridges, alters protein binding in addition to increasing AON lipophilicity. However, because all possible combinations of Rp and Sp diastereomers are present in the classic nonstereodefined AON random mixtures (see above), the effects of each individual Rp and Sp are so diminutive that the PSs collectively appear ‘pseudosymmetric’, essentially neutralizing the chiral contributions of a PS linkage. However, the nature of a classic random mixture, essentially a library of diastereoisomers, might in itself have a positive role. Given that random mixtures exhibit all possible isomers, including isomers with extreme and/or desired properties, such as high potency, they might outbalance less active, or inactive, isomers [34,138]. It is possible that such a ‘library effect’ and the heavy atom effect work together to promote AON PS activity. By contrast, advantages with AON technology might occur when both the heavy atom effects of sulfur combined with pure phosphorus chirality are used (see below). Naturally, this can only occur for stereodefined PSs.

**Diversity offers new opportunities for RNA therapeutics**

The standard, target-centric, model of RNA drug discovery begins with the identification of a target known to be causally involved in a particular disease (Fig. 6a). From the sequence of the RNA target, libraries of complementary AONs can then be designed and synthesized. The functional properties of the synthesized AONs are evaluated in a screening cascade of cell-based assays and animal models, to identify active and well-tolerated compounds (Fig. 6b).

In this approach, medicinal chemists are confronted with a large chemical space to explore. For gapmers designed to knock down the target RNA, as well as mixmers inhibiting miRNAs or modulating splice switching on pre-mRNA, the number of chemically feasible compounds against any target is on the order of $10^{20}$ (Box 1). Despite great advances in high-throughput synthesis and screening technology, such a large space prohibits exhaustive evaluation.

The current target-centric discovery approach relies on the principle of ‘evaluation in-parallel and local optimization’. That is, libraries of AONs that predominantly explore one parameter (for example potency) are designed, synthesized, and screened in parallel, and, from these, sets of AON hits are selected that are optimal with respect to that parameter. Given that synthesis and screening of AONs takes on the order of weeks for the in vitro part of the screening cascade, and on the order of months for the in vivo part, parallel evaluation of libraries of AONs is necessary to keep the timespan of the discovery project on the order of 1–2 years. Using the analogy of chemical space, the map guiding the search for active and tolerated AONs is as such constructed ‘along the way’, and the path forward determined only by what is known in the local regions explored up until that point. This approach does not necessarily lead to the identification of the globally most-active and best-tolerated AON, but historically have nevertheless performed well and produced numerous drug candidates that have entered clinical testing in human trials (see above).

To make these concepts more concrete, we illustrate in Fig. 6c,d a recent drug discovery project in which local regions of chemical space were explored with respect to nucleobase as well as architectural diversity (Fig. 6c). The gapmers from the first scanning library (green dots in Fig. 6c,d) were screened for activity on the intended target as well as cellular toxicity. As seen in Fig. 6d, some were active and well tolerated, but many were not. The scanning library primarily explores the nucleobase determinant, a classic ‘gene-walk’ (Fig. 6c). Next, the nucleobase sequences for a handful of interesting compounds were identified and used as a basis for designing optimization libraries exploring nucleoside and/or architectural diversity (blue dots in Fig. 6c,d). As seen in Fig. 6d, the optimization libraries identified a higher fraction of leads that were potent and well tolerated compared with the hits in first scanning libraries (upper right quadrant in Fig. 6d).

Furthermore, having identified leads where both nucleobase sequence and architectural design had been locally optimized, additional improvements in functional properties might be possible by synthesizing new libraries of fully stereodefined versions of these gapmers (Fig. 6a). One such example was presented above for fully stereodefined gapmers targeting HIF1A mRNA (Fig. 4). The high property diversity observed for PS diastereoisomers offers a new macromolecular diversity parameter for AON optimization.
Also, in contrast to classic random mixture AONs, PS stereodefined oligonucleotides offer a unique opportunity: it is now possible to study the complex interactions between AONs and, for example, proteins by computational modeling.

The understanding of how structural diversity determinants affect functional properties is essential to guide heuristic approaches in drug discovery. For example, what is the most efficient path to highly potent gapmers? Is it scanning the target RNA comprehensively in 1000 regions with one selected standard design (broad strategy)? Is it scanning the target in only 100 different regions but testing ten different gapmer designs in each (narrow strategy)? Or is it evaluating only ten different target regions but with 100 different gapmer designs in each region (deep strategy)? As calculated in Box 1, thousands of different modified gapmers can be designed against any target region, so each of these strategies is clearly possible. The answer depends on whether nucleobase or architectural diversity affects the therapeutic index the most, and what balance between them explores chemical space most efficiently. Traditionally, nucleobase diversity has been explored the most, with the aim of searching for the regions on the transcript most accessible to the AONs. However, as the examples presented in the sections above show, we are observing considerable impact on potency from other diversity determinants. Consequently, we are moving away from pure broad strategies with only a single fixed design, towards more balanced and deeper strategies where...
**BOX 1**

**The chemical space of RNA therapeutics**

Chemical space can be thought of as an analogy to the cosmological universe, with chemical compounds populating this space instead of stars [139]. How many molecules populate this space? For small organic molecules with up to 17 atoms of C, N, O, S, and halogenes, and taking chemical stability and synthetic feasibility in account, the number has been estimated to be more than $10^{11}$ compounds [140]. Here, we estimate the number of possible gapmers that can be designed against an RNA target. For an RNA that is $n$ nucleotides long, and where the gapmer is $l$ nucleotides long, in total there will then be $n - l + 1$ possible target regions of length $l$. Considering next the backbone chemistry, for a gapmer of length $l$, there will be $l - 1$ linkages between nucleotides, each of which can be in either $R$- or $S$-form. This gives $2^{l-1}$ different stereoisomers. Finally, any type of $m$ possible modified nucleosides can in principle be used in the regions flanking the DNA gap. Therefore, for a gapmer of length $l$, with a gap of length $g$, the number of different modified gapmers that can be synthesized equals $m^{l-g}$. Thus, considering all gapmers of lengths between $l_{\text{min}}$ and $l_{\text{max}}$ taken together the number of different, fully stereodefined, gapmers is equivalent to Eq. (1):

$$
\sum_{l=l_{\text{min}}}^{l_{\text{max}}} (n - l + 1)2^{l-1}m^{l-g}.
$$

For a standard pre-mRNA that is $n = 25,000$ nt in length, and considering gapmer lengths between $l_{\text{min}} = 12$ nt and $l_{\text{max}} = 20$ nt, with a minimal gap of $g = 6$ nt, and where $m = 5$ types of nucleosides (DNA, LNA, MOE, 2’OMe, and 2’F) are considered, there are $8.89 \times 10^{19} \approx 10^{20}$ different possible gapmers. Equation 1 can also be used when estimating the number of mixmers designed to target a narrow region of RNA, such as a miRNA or a splice-site. In this case, with the same parameters as for gapmers but setting $n = 20$ nt and $g = 0$ nt, there are $6.17 \times 10^{19} \approx 10^{20}$ different possible mixmers, about the same number as for gapmers. The smaller region of RNA that can be targeted is balanced out by having the entire oligonucleotide available for modifications, where, for gapmers, there is a gap where only DNA can be used.

multiple designs are tested in highly overlapping regions on the target RNA.

Another point to consider is the multidimensional nature of the local optimization in chemical space, where several properties of gapmers have to be acceptable and drug-like at the same time. Should we then measure all functional properties for the entire library of gapmers? Or are there some properties that are not necessary to consider early on, because diversity can be introduced at a later stage to ensure that the wanted functional properties can be achieved? Say, for example, that exploring nucleobase diversity has identified potent but not well-tolerated gapmer hits. When then nucleoside, architectural, or backbone diversity is explored, will the chance of finding leads that are both potent and well tolerated be higher if we had not begun with hits only optimized for potency?

If AONs shared the same properties simply by membership of the same chemical class (‘portability of chemistry’), only sequence would differentiate properties and broad gene walks would be required to have enough compounds to select from (‘broad strategy’). However, the pronounced functional effects that small changes in chemical structure have (see above) offer new opportunities for AON drug discovery. Diversity has now led us to search much deeper and in a narrower target sequence space (narrow or deep strategy). If we on top of this add the diversity of diastereoisomers, then a large diversity/property spectrum will unfold even from just a single classic AON PS compound mixture (ultra-deep strategy). Deep or ultra-deep strategies offer not only an introduction of a much larger structural space to search in, but more importantly, also a possibility to begin with much ‘tougher’ selection criteria. It is possible to begin with higher sequence restrictions and/or requirements. If, for instance, many and strict requirements are applied, such as species crossreactivity, sequence specificity to only the intended target [15,18], deselection of predicted sequence motifs associated with increased toxic potential [18,19,136], deselection of unwanted sequence motifs from a chemistry, manufacturing, and control perspective, and so on, only a very few sequences will finally comply. In the case where a few 16mer compound mixtures have been identified complying with all selection criteria, it should then be possible to find compounds with even better properties among the 32,768 structurally different diastereoisomers of each 16mer. Deep or ultra-deep strategies will also be advantageous for discovery programs focused on, for example, splice-switching or miRNA inhibition, because here the relevant RNA target regions sequence wise are very narrow.

There is no doubt that harnessing the full potential offered by diversity drives AON drug discovery in the direction of small-molecule concepts. This is with respect to ‘higher’-throughput screenings and working with single molecules (fully stereodefined AONs). However, AON drug discovery exhibits inherent advantages over small-molecule principles. AON drug discovery concepts are more ‘rational’. Watson–Crick base-pairing rules govern target RNA binding, whereas, for small molecules, complex computational docking studies must be undertaken. Also, importantly, oligonucleotide production by solid-phase synthesis follows shared principles for all compounds. Furthermore, as we develop our understanding of these new SARs, we expect to apply an even more ‘rational’ approach to AON drug discovery, such as by developing in silico algorithms that are more predictive for AON screening outcomes taking diversity into account. This has already been done for hepatotoxic potential [19]. In this way, bioinformatics and computational modeling can enable faster discovery processes and, throughout the process, more informative decisions, thereby creating stronger lead candidates.

**Concluding remarks**

The fact that structural differences also define functional diversities is a given for all compound classes. We have shown that not only sequence, but also the design of an AON is a strong property determinant (Fig. 2). A clear SAR exists within a defined target nucleobase sequence simply by varying the sugar and phosphate backbone within that sequence (Fig. 4). This creates AONs with different pharmacological properties (Figs 3 and 5). Exploration of the large property space is the reason why drug discovery libraries in our laboratories have expanded greatly. Initially, we used mainly a broad screening strategy (gene walk), but now we are also exploring compound diversities within selected nucleobase sequences (deep strategy) (Fig. 6).
The diversity we experienced for stereodefined LNAs (with the same sequence and nucleoside design) represents an extra dimension by which even more unique leads might be identified. The basis for their uniqueness might, in part, be created by the positive heavy atom effects of sulfur and the introduction of chirality on phosphorous, which we think act as structural diversity parameters in themselves. A chiral phosphorous scaffold might be ideal for designing drugs interacting with the chiral world of biology.

Given that AONs are based on something as predictable as Watson–Crick base-pairing to RNA, binding to intended and unintended RNA targets can be predicted computationally and measured experimentally [18]. This provides a means for assessing unintended hybridization-induced biological effects of AONs [105]. Still, they are, perhaps unsurprisingly, unique molecules that, in a therapeutic setting, provide unique and yet ‘hard-to-predict’ pharmacological properties. Clearly, more work is needed to build understanding of the underlying mechanism of AON diversity, and to determine how much the TI can be expanded by these new strategies.

However, what is clear at this point is that antisense has developed immensely over the past four decades, and the new drug discovery strategies are producing increasingly better RNA therapeutics. We have realized that the old concept of antisense of ‘portability of chemistry’ only provides a starting point for RNA drug discovery and, as shown here, structural diversity at all levels (nucleobase, nucleoside, backbone, and architectural) can be used to produce AONs with improved drug properties.

Conflict of interest
All authors except N.P. are full time employees at Roche Innovation Centre Copenhagen A/S.

Acknowledgments
We would like to thank Cy A. Stein for helpful comments on the manuscript, and Mikkel Adriansen for in-depth analysis of the mass spectra of stereo-defined oligonucleotides. For their excellent technical assistance, we would like to thank Helle Nymark and Annette Gildov for qPCR high-throughput screening, Charlotte Øverup for mouse liver and serum measurements, Rikke Selberg, Camilla T. Haugeard, and Lisbeth Bang for handling of mice in the animal facilities, and Valesi Mukaka for the synthesis and formulation of oligonucleotides for mouse studies.

References
17 Swaye, E.E. et al. (2007) Antisense oligonucleotides containing locked nucleic acid improve potency but cause significant hepatotoxicity in animals. Nucleic Acids Res. 35, 687–700
19 Hagedorn, P.H. et al. (2013) Hepatotoxic potential of therapeutic oligonucleotides can be predicted from their sequence and modification pattern. Nucleic Acid Ther. 23, 302–310