Design and applications of modified oligonucleotides

M. Gallo1, J.M. Montserrat1,2 and A.M. Iribarren1,3

1Laboratorio de Química de Acidos Nucleicos, Instituto de Investigaciones en Ingeniería Genética y Biología Molecular, CONICET-UBA, Buenos Aires, Argentina
2Instituto de Ciencias, Universidad Nacional de General Sarmiento, Buenos Aires, Argentina
3Centro de Estudios e Investigaciones, Universidad Nacional de Quilmes, Buenos Aires, Argentina

Abstract

Oligonucleotides have a wide range of applications in fields such as biotechnology, molecular biology, diagnosis and therapy. However, the spectrum of uses can be broadened by introducing chemical modifications into their structures. The most prolific field in the search for new oligonucleotide analogs is the antisense strategy, where chemical modifications confer appropriate characteristics such as hybridization, resistance to nucleases, cellular uptake, selectivity and, basically, good pharmacokinetic and pharmacodynamic properties. Combinatorial technology is another research area where oligonucleotides and their analogs are extensively employed. Aptamers, new catalytic ribozymes and deoxyribozymes are RNA or DNA molecules individualized from a randomly synthesized library on the basis of a particular property. They are identified by repeated cycles of selection and amplification, using PCR technologies. Modified nucleotides can be introduced either during the amplification procedure or after selection.

Introduction

Oligonucleotides are single-stranded oligomers of nucleoside units linked together by phosphodiester bridges formed between a 3'-hydroxyl group of one monomer and the 5'-hydroxyl group of another. They can also be defined as short fragments of single-stranded DNA or RNA. Oligonucleotides can be synthesized chemically and are the basis of a research area called nucleic acid chemistry, which has received increasing attention over the last few years.

Khorana was the pioneer in this field since he developed the first chemical synthesis of oligodeoxynucleotides. Although this idea was initially considered to lack practical interest, the discovery of molecular biology tools such as restriction enzymes, DNA sequencing and PCR has provided a wide range of applications for these molecules. The use of oligonucleotides was further facilitated by the development of solid phase synthesis in the early 80’s, which allowed the preparation of longer fragments and simplification of the procedures.

Therefore, synthetic oligonucleotides are today utilized for a wide range of purposes in fields such as biotechnology, molecular biology, diagnosis and therapy. Natural oligonucleotides have been extensively used for gene construction, selection and determina-
tion of DNA sequences (hybridization probes or affinity purification), and site-directed mutagenesis. They constitute the basic reagents for PCR technologies, they are also used to selectively inhibit gene expression in *in vitro* studies and more recently, their ability to interact with proteins has also been explored.

However, the spectrum of oligonucleotide applications can be broadened by chemically modifying their structures. Conjugated oligonucleotides resulting from the coupling with one or more molecules that confer particular properties represent useful analogs. The coupled moieties can provide a particular reactivity, such as cross-linking with other nucleic acids or proteins, or cleave nucleic acids in a selective way. They can also be useful for detection as nonradioactive probes, or for automated sequencing. Other important conjugates are those that provide special intermolecular interactions such as intercalation, enhanced cellular uptake and binding characteristics. All these molecules are widely employed in molecular biology and diagnosis.

Another example of modified oligonucleotides are those that carry modifications in the nucleotide moiety. The main purpose of these alterations is to confer increased stability in biological fluids, since in these media (cell culture media, serum or cytoplasm) natural oligonucleotides are rapidly degraded by enzymes called nucleases.

The modifications can be introduced in the base, sugar or phosphate moieties, but the design of new modified oligonucleotides is a difficult task due to chemical, electronic and steric concerns.

Since these analogs are essential for *in vivo* applications, extensive efforts have been made in this field during the last years. Perhaps the area where research on new analogs has been more prolific is the antisense strategy. Therefore this review will focus mainly on this approach and on one of the latest oligonucleotide applications, i.e., aptamers and catalytic molecules obtained using combinatorial approaches. Ongoing projects of the authors will also be considered.

**Antisense chemistry**

The improvements achieved in molecular biology have allowed a better understanding of diseases at the molecular level, and this has facilitated the development of new drugs based on a rational design.

Antisense therapy is a particular application of rational design based on oligonucleotide chemistry. In 1978 Zamenick and Stephenson (1) first illustrated the idea of this therapy demonstrating that oligonucleotides could inhibit replication of Rous sarcoma virus in a cellular system. The principle involved in the antisense approach is simple: inhibition of expression of a specific gene at the RNA level can be attained by using complementary oligonucleotides, called antisense oligonucleotides, thereby blocking expression of the protein encoded by the target RNA. If the inhibited proteins are disease-related ones, then oligonucleotides become a new class of drugs.

In theory, hybridization of a short antisense oligonucleotide of 15-20 bases to its complementary mRNA by Watson-Crick base pairing should provide high specificity. On the other hand, this strategy offers a general approach since several different diseases could be targeted due to the fact that all the proteins are synthesized by the same mechanism: the gene carrying the genetic information for a particular protein is transcribed to a single-stranded mRNA which is the intermediate carrier of information and serves as template for the protein synthesis by the ribosome (Figure 1).

Numerous reports have demonstrated the ability of antisense oligonucleotides to block the function of specific genes *in vitro* as well as *in vivo*. These molecules are currently being investigated as therapeutic agents for the treatment of viral infections, cancers and
immunological diseases, and several clinical trials are under way (2). Moreover, the first drug based on this strategy has been recently approved by the FDA (Vitravene™ for treating CMV retinitis in AIDS patients).

Although first attempts used unmodified phosphodiester oligonucleotides, it soon became evident that their use as therapeutic agents was limited because of their instability in biological fluids due to the ubiquitous nucleases. These enzymes rapidly degrade the DNA and RNA oligonucleotides by hydrolyzing the phosphodiester backbone. To overcome this problem, considerable efforts have been made to develop more stable analogs that display a reasonable half-life in vivo (3). The design of chemically modified oligonucleotides is a complex process since not only resistance to nucleases but also stable duplexes should be attained. The importance of hybridization is demonstrated by correlation of antisense activity observed in cell culture assays and in vivo experiments, with the hybridization affinity expressed as the Tm (4,5). In addition, the modification should also improve cellular uptake and tissue and cell distribution for a particular target.

Hybridization of antisense oligonucleotides to their target mRNA can lead to processing, transport or translation arrest by physical blockage, or activation of the intracellular RNase H, which then causes cleavage of the target mRNA at the RNA:DNA duplex site. The action of this enzyme degrades the mRNA and the oligonucleotide remains available for binding to a new mRNA target and therefore it can act catalytically and produce an increased biological activity. Unfortunately, this nuclease is very sensitive to structural alterations of the antisense oligonucleotide, and most of the modifications are not compatible with this mechanism. Moreover, unspecific results have also been reported based on the action of this enzyme.

Chemical modifications of the sugar moiety can be prepared in the internucleotide phosphate backbone or in the nucleobase. The scope of nucleobase modifications is rather limited as the requirements for base pairing must be respected in order to maintain the specificity of the antisense strategy. Nevertheless, there are a number of base-modified oligonucleotides that have been shown to be potentially useful. This kind of analogs will not be included in this review; however, there are excellent reports that should be consulted for a comprehensive discussion of this topic (6,7). The present review will focus on the most extensively used modifications thus far, since more extensive descriptions of available analogs have been published elsewhere (8,9).

**Backbone modifications**

The target of nuclease action is the internucleotide phosphate, and therefore the first attempted modifications involved this group. A common problem associated with anionic compounds is the ineffective permeation of cell membranes. Based on this observation, neutral analogs, methylphosphonate oligomers (Figure 2-1a) and phosphoramidates (Figure 2-1b) have been readily prepared on solid supports. Although these analogs have an increased resistance to nucleases, they show a lower RNA binding affinity than natural phosphodiester-based parent sequences. This reduced affinity has been at...
tributed to the asymmetric center generated at each modified phosphorus atom, which under solid phase synthesis leads to the formation of diastereomeric mixtures with variable affinity for the RNA. However, the use of these analogs is limited due to lack of RNase H recruitment and problems related to a still poor cell permeation.

Phosphorothioates are the most widely studied analogs (Figure 2-1c) and have been referred to as “the first generation antisense oligonucleotides”. Phosphorothioates are readily synthesized in high yield, are capable of activating RNase H, and exhibit greatly enhanced nuclease resistance and favorable pharmacokinetic properties. Vitravene carries this modification and other antisense phosphorothioates have shown promising activities in ongoing clinical trials (10). Nevertheless, these analogs, such as methylphosphonates, render the phosphate linkage chiral, and therefore they exhibit a slightly reduced RNA binding affinity. However, the major drawback for their use in antisense therapeutics is their tendency to bind nonspecifically to proteins, which can be the cause of undesirable side effects upon in vivo administration (11). Thus, this backbone modification does not provide a fully satisfactory or general solution to the problem of designing potent and highly specific antisense inhibitors of gene expression.

Other backbone modifications worth mentioning are the (N3’→P5’) phosphoramidates (Figure 2-1d) in which the 3’-O-bridging has been replaced with NH. Although these molecules do not activate RNase H, they have shown encouraging results since they dramatically increase the $T_m$ and provide very good resistance to nucleases (12).

**Sugar modifications**

A wide variety of sugar modifications have been introduced into antisense oligonucleotides to enhance nuclease stability and binding affinity. The most promising ribose modifications involve the 2’ position which defines the conformation of the furanose ring. An electronegative substituent shifts

![Figure 2. Oligonucleotide modifications.](image-url)
the conformational equilibrium of the sugar moiety toward the northern conformation (C3'-endo) consistent with RNA duplexes. This conformation in the antisense strand is expected to preorganize it for proper binding to RNA. The more electronegative the 2' substituent, the greater the shift and the higher the $T_m$. Thus, the 2'-deoxy-2'-fluoro oligonucleotides (Figure 2-2a) form the most stable hybrids but they do not show enough resistance to nucleases for *in vivo* applications. In contrast, the 2'-amino substituted oligonucleotides (Figure 2-2b), due to the lower electronegativity of the nitrogen atom, tend to adopt the south conformation and therefore display a reduced binding affinity.

The 2'-O-alkyl oligonucleotides (Figure 2-2c) (13) also stabilize the duplex, with smaller substituents resulting in greater stability than larger ones, probably due to steric interference by the larger alkyl chains with other parts of the duplex or disruption of water structure in the minor groove. On the other hand, larger groups at the 2'-position improve nuclease resistance. Considering this tendency, it was surprising that 2'-O-alkoxy substituents derived from ethylene glycol, such as 2'-methoxy-ethoxy, 2'-methoxy-triethoxy analogs, showed increased binding affinity (5). This result can be explained as a consequence of the structural preorganization of the modified strand due to the C3'-endo conformation and of the gauche orientation adopted by the ethylene glycol moiety, which allows the accommodation of the side chain in the minor groove and a favorable hydration (14). The 2'-O-alkyl derivatives are usually referred to as the second generation of antisense compounds.

Since none of the 2'-modified analogs supports the RNase H mechanism, a strategy to overcome this limitation consists of the use of chimeric oligonucleotide analogs, which bring together the beneficial properties of the two types of chemistry. In general, they have two segments: one that contains an oligonucleotide derivative capable of activating RNase H (like phosphorothioates) and another that provides increased binding affinity and less side effects (like 2'-O-alkyl oligonucleotides). Successful results have been obtained in cell culture experiments and in experimental animal models (15).

Since the relationship between the sugar conformation and the RNA binding affinity is an important variable in the design of modified oligonucleotides, 2'-deoxy-2'-C-alkyl derivatives present an interesting case. The 2'-deoxy-2'-C-methyl nucleosides that carry the alkyl moiety down in relation to the plane defined by the furanose ring (2'R analogs, Figure 2-2d) show a 2'-endo (south) conformation, while the “2'-methyl up” isomers (2'S, Figure 2-2e) exist predominantly in the north conformation. Oligonucleotides carrying these modifications increase greatly the stability against degradation by nucleases (16). As expected, the binding affinity of the 2'-methyl down oligomers for complementary DNA and RNA was reduced in comparison with unmodified sequences, or even no cooperative melting was observed. On the other hand, 2'-methyl up oligonucleotides exhibited a thermal stability that depended on the sequence and the length of the modified oligonucleotide (17,18). The reason why the (2'S)-methyloligodeoxynucleotides hybridize better than the (2'R) analogs could be attributed to the fact that the first ones have the right preorganization to fit the complementary RNA strand (19), since the sugar moieties are locked in the 3'-endo conformation.

All the concepts discussed above indicate that further improvements in the design of antisense oligonucleotides need to be developed in order to obtain proper therapeutic agents.

**Ribozymes**

Ribozymes are RNA molecules with catalytic activity that cleave internucleotide phosphate bonds of the RNA substrates (20). The hammerhead ribozyme, originally found in
plant viroids, is the simplest one. It can act intermolecularly (trans) and can be tailored to recognize different target sequences. The minimum structure (ca. 40 nucleotides) that maintains the catalytic activity is composed of a conserved catalytic core and two flanking regions complementary to the substrate sequence (21).

As a result of ribozyme attack, the hydrolyzed RNA substrate is biologically inactivated and therefore ribozymes are attractive molecules for the development of new drugs based on the gene therapy principle (22). Like other antisense strategies, ribozymes potentially offer a broad action spectrum since they can be tailored to recognize different mRNAs and by their catalytic action they can produce the desired inhibitory effect at a lower concentration than the corresponding antisense oligonucleotides.

Ribozymes have already proved to be effective gene expression inhibitors in vitro as well as in vivo (23,24).

In order to obtain therapeutic agents based on synthetic ribozymes, it is mandatory to chemically modify theses structures. These modifications, as in the traditional antisense strategy, should confer resistance to nucleases, selectivity and proper hybridization and uptake characteristics. In the case of ribozymes, the design of new modified nucleotides becomes more complex since a proper folding of the nucleic acid is needed in order to maintain the catalytic activity. Several studies (X-ray structure elucidation and mapping with modified nucleotides) have shown that the presence of the 2'-hydroxyl group at specific positions in the catalytic core is essential for hydrolytic activity. Up to now the 2'-modifications lack this group and therefore the synthesis of building blocks carrying this function (2'-C-methyl nucleotides, Figure 2-2f) seems to be an interesting goal. The introduction of 2'-C-methyluridine at those positions of the catalytic core of a hammerhead ribozyme where the 2'-hydroxyls were found to be essential produced \textit{in vitro} active ribozymes with increased resistance to degradation by nucleases (Gallo M, Sczakiel G and Iribarren AM, unpublished results). Further biological and structural studies are being carried out.

**Combinatorial libraries of modified oligonucleotides**

In the last twelve years there has been significant development in the use of libraries of compounds for the discovery of molecules with new binding or catalytic properties (25). The basic idea is to shorten the time and effort associated with the search of drug leads by developing a huge combination of molecules which are further tested against a certain target or for a specific activity. The main problem of this technique is the design of a proper strategy to determine the identity of the active molecules from the informational chaos present in the original pool. To solve this problem (26) either untagged or tagged methodologies have been developed.

In the first case several strategies were applied, for example light-directed spatially addressable combinatorial libraries (27), pin arrangements (28), or recursive deconvolution processes (29).

The tagged methodologies establish a univocal relationship between each member of the library and a coding tag that denotes its identity (Figure 3). In this sense, one of the most powerful tag technologies is the phage display that joins peptide and oligonucleotide properties with tools of molecular biology (30). The principle is that the library is presented on the surface of bacteriophages, each of them displaying a unique peptide that is coded by its genome. This idea has been used by Brenner and Lerner (31) to develop a versatile method named encoded combinatorial chemical library based on alternating parallel combinatorial synthesis where the sequence of a certain peptide is tagged by a natural oligonucleotide. In these examples the library diversity is developed...
by peptide-based chemistry and requires the fulfillment of three requisites: the chemistries between the tag and the diversity branches must be orthogonal, both chemistries should be suitable for automatic synthesis protocols, and the tag should be easily decoded. A related approach is based on a modified oligonucleotide library encoded by a natural oligonucleotide chain, as can be seen in Figure 3 (Sacca P, Fontana A, Montserrat JM and Iribarren AM, unpublished results).

There is a quite different strategy that makes use of modified oligonucleotide libraries called Selective Enhancement of Ligands by Exponential Enrichment (SELEX) (32). The basic principle of this technique is the execution of multiple cycles of selection and amplification resulting in the competition between the most suitable molecules that allows the identification of the best candidates, or set of candidates, from a huge population. In this sense, individual members of a pool are separated from each other on the basis of their ability to perform an arbitrary task, usually a binding or a catalytic activity.

The strategy followed in the selection process differs according to whether DNA or RNA is being used (Figure 4). In the first case, the process usually starts with a pool of natural DNA that is amplified using standard PCR techniques in the presence of a modified nucleoside triphosphate. The oligonucleotide sequences of the library contain a random section of 20-40 nucleotides (indicated by thin lines in Figure 4) flanked at each end that generates $10^{12}$-$10^{18}$ different compounds. This central section is flanked by fixed sequences corresponding to the primers necessary for the amplification process (indicated by bold lines in Figure 4). This huge combination of molecules is then subjected to the selection process. In particular,
the development of catalytic RNA or DNA, ribozymes or deoxyribozymes (33) requires special strategies such as activation of the catalytic activity with organic or metallic cofactors. On the other hand, aptamers, which are selected oligonucleotides with specific binding properties, are usually obtained by the immobilization of the target molecule on a solid support with the configuration of a membrane or a column. In this case, the library is eluted through the immobilized target and the oligonucleotides with better binding properties are retained. Then, the bound sequences are eluted with the help of a solvent with a high salt content or a free small ligand. The selected sequences are amplified by PCR and subjected to a second selection cycle. Usually after 10-15 rounds a discrete number of candidates with good binding properties can be isolated. The sequences are separated and identified using standard molecular biology protocols.

If an RNA pool is used as starting material the selection-amplification cycle (Figure 4) is not so straightforward. Usually, a DNA pool is transcribed with the help of a transcriptase into RNA using modified ribonucleotides (indicated by a star in Figure 4). Then, the modified RNA pool is subjected to a selection process, as previously described for DNA, and a set of sequences with enhanced properties is isolated. The sequences are separated and identified using standard molecular biology protocols.

The use of modified nucleotides in SELEX techniques requires the fulfillment of basic requirements. If DNA is used, then the nucleotide must be the substrate of a Taq polymerase. On the other hand, if a modified RNA library is developed, the modified ribonucleotides must be transcriptase substrates and the modified RNA must be recognized by the reverse transcriptase. In this sense, several research groups have focused on the development of modified nucleotides with added functionality as building blocks for in vitro selection procedures. The majority of these studies deal with the modification and selection of DNA, presumably due to its enhanced chemical stability relative to RNA. But on the other hand, RNA sequences in comparison to DNA sequences may offer a wider variety of folded structural motifs that enrich the diversity space of the library. For example, the presence of 2'-hydroxyls may assist in interdomain binding to provide better tertiary structure stability.

Some modifications in deoxynucleotides were made at the 5'-position of deoxyuridine derivatives. The first report was the application of 5-(1-pentynyl)-2'-deoxyuridine (34) in the development of a novel thrombin aptamer. It is interesting to note that the isolated sequences of this modified library were strikingly different compared to the one selected by using natural bases (35).

Further studies focused attention on the investigation of different functionalizations tolerated at the 5'-position by Taq polymerase and other enzymes. It was not until recently that modified purine deoxyribonucleotides were synthesized and tested as Sequenase T7 and Taq polymerase substrates. Several examples of modified ribonucleotides applied to SELEX techniques can also be found, such as, 2'-amino-2'-deoxypyrimidines used to identify high affinity ligands specific for human neutrophil elastase and 5-pyridylmethylcarboxamid-uridine triphosphate used to obtain an RNA sequence able to catalyze the Diels-Alder reaction (36).

In conclusion, SELEX is a powerful tool that can provide both modified RNA and DNA with binding or catalytic properties. Therefore, the development of functionalized nucleotides that behave as substrates of the involved enzymatic machinery seems to be a fruitful field of research.
References


