

Design and Evaluation of Oligonucleotide Analogues

Christian J. Leumann

Abstract: This article contains a short account of the chemical and biophysical properties of the oligonucleotide analogues bicyclo-DNA, tricyclo-DNA, [3.2.1]-DNA, [3.2.1]amide-DNA and the PNA analogue OPA, recently developed in our laboratory. Emphasis is given to various aspects of conformational restriction as a designer tool for enhancing the performance of oligonucleotide analogues and for exploring the supramolecular chemistry of DNA. A short summary of current problems in the field of DNA triple-helix chemistry as well as contributions from our laboratory in this area concludes this communication.

Keywords: Conformational restriction · DNA · Oligonucleotide analogues · RNA · Triple helix

Introduction

More than two decades ago, in two seminal papers, Stephenson and Zamecnik [1][2] disclosed the successful use of oligonucleotides to inhibit the expression of specific genes at the level of translation *in vitro*, thus paving the way for what is now more generally called the antisense approach in human therapy. This therapeutic concept is based on the interaction of an oligonucleotide analogue (antisense oligonucleotide) with a complementary mRNA, to inhibit translation of the message into a protein. The attractiveness of this concept, compared to traditional protein targeting with small molecules, lies in the fact that only minimal structural knowledge about the target, namely its specific gene sequence, is required in order to rationally design a highly selective inhibitor. A major effort in the search of oligonucleotide analogues with favorable chemical and biological properties was the result. A drug acting against cytomegalovirus (CMV) induced retinitis (Vit-ravene™), entirely based on an oligonucleotide analogue, recently came to market and demonstrates the feasibility of the

therapeutic antisense concept. On the other hand, rapid progress in molecular biology and molecular medicine, which has recently culminated in the sequence determination of the human genome, demands the continuous development of new DNA-based analytic and diagnostic methods such as chip technology [3], for high-throughput sequencing of DNA, fast screening of gene function, or real time fluorescence detection of genetic mutations. A further interest in oligonucleotide chemistry has arisen from the almost unique opportunity for constructing well-defined DNA-based molecular architectures including rods, pseudoknots, borromean rings, and two- and three dimensional networks [4][5] by oligonucleotide sequence design, and their eventual use in nanotechnology [6][7]. All of this, however, would not have been possible, were powerful methods for solid phase synthesis of oligonucleotides unavailable [8].

In this realm, our own research activities focus on various supramolecular aspects of DNA/RNA association and folding. With computer-aided design and modeling, the tools of organic synthesis, basic molecular biology and biophysical analysis, we evaluate DNA-analogues with the goal of determining the details of the interrelation between the structure of the monomeric constituents and the properties of the corresponding oligomers. This is pivotal in order to be able to produce oligonucleotide analogues with tailor-made properties for use in the above-mentioned areas. Besides this,

such studies will almost certainly lead to new insight into the basic, non-covalent factors that determine biomolecular organization in general. While we have adequate knowledge to predict conformational preferences of a (bio)molecule, amongst other properties, we are still far from being able to rationalize and thus predict the exact effect of hydrophobic forces, stacking interactions and solvation on structure. In the following, an account of current research projects with special emphasis on conformational and non-covalent aspects of DNA association is given.

Conformationally Restricted Oligonucleotide Analogues

Inspired by the chemical properties of homo-DNA [9] we started more than ten years ago with the relatively simple idea to adapt the concept of structural preorganization [10] to oligonucleotides. By constraining the conformational states of a single-stranded oligonucleotide to match the geometry in its duplexed form, an entropic benefit for complexation is to be expected which should result in a higher thermodynamic stability of the duplex. With this background the oligonucleotide analogues bicyclo-DNA and tricyclo-DNA, showing conformational restriction in their sugar units, were prepared and evaluated (Fig. 1). The concept as such has now been widely applied by a number of research groups and was covered in a recent review [11].

*Correspondence: Prof. Dr. C.J. Leumann
Department of Chemistry & Biochemistry
University of Bern
Freiestrasse 3
CH-3012 Bern
Tel.: +41 31 631 4355
Fax: +41 31 631 3422
E-Mail: leumann@ioc.unibe.ch

From detailed analysis of a series of different bicyclo-DNA and tricyclo-DNA oligonucleotide sequences, 6–12 nucleotide units in length, we found that, in contrast to natural DNA, homopurine/homopyrimidine sequences of these analogues preferentially form duplexes of the Hoogsteen and reversed Hoogsteen type with strongly enhanced thermal stability [12][13]. This unexpected change in the base-pairing preferences was found to be mainly the consequence of a geometric change of one of the six torsion angles (torsion angle γ), defining the backbone geometry of DNA, from the *gauche* to the *antiperiplanar* conformation (Fig. 2). Tricyclo-DNA behaves similarly and leads to a very strong Hoogsteen base-pairing system. Denaturation temperatures (T_m) of homoadenine- and homothymine- nonamer duplexes of tricyclo-DNA can be higher by almost 50 °C compared to the parent, natural (Watson-Crick) base-paired duplex.

The repertoire of secondary structural motifs of RNA is generally much larger than that of genomic DNA because of the lack of a complement. Typically, ribosomal RNA, mRNA, catalytic RNA, and single stranded DNA-aptamer sequences therefore fold intramolecularly to form complex three-dimensional structures with unusual base-base arrangements, hairpin-loops and bulges. To improve the performance of catalytic RNA or DNA-aptamers it is sometimes desirable to stabilize such secondary structural motifs. Fig. 3 shows a self-complementary DNA dodecamer which occurs as an equilibrium mixture between a bimolecular duplex and a monomolecular hairpin in solution. Doping the loop-region of the oligonucleotide with individual tricyclo-DNA units leads to distinct differences in the equilibrium mixture and the hairpin stability. Exchanging the 5'-thymidine residue by a tricyclothymidine unit *e.g.* leads to an increase in T_m of the hairpin by approximately 20 °C (red curve) and a complete shift of the equilibrium towards the hairpin structure, as can be inferred qualitatively from the corresponding UV-melting curves. Exchanging the 3'-deoxyadenosine residue by a tricyclo-adenosine unit (cyan curve) shows the opposite effect, namely a stabilization and equilibrium shift towards the bimolecular duplex structure. Thus besides being a structural probe for local loop conformations, such rigid nucleosides with altered backbone geometry may prove useful for enhancing the binding efficiency of DNA-aptamers or the catalytic efficiency of ribozymes.

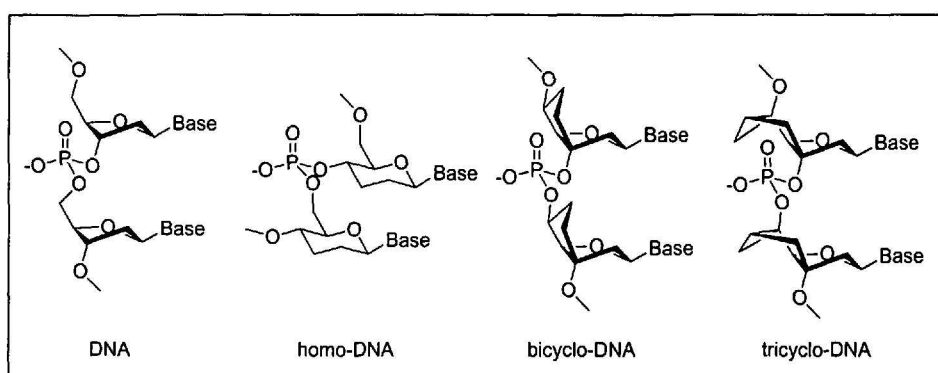


Fig. 1. Chemical structures of a selection of oligonucleotide analogues built from conformationally constrained nucleoside analogues.

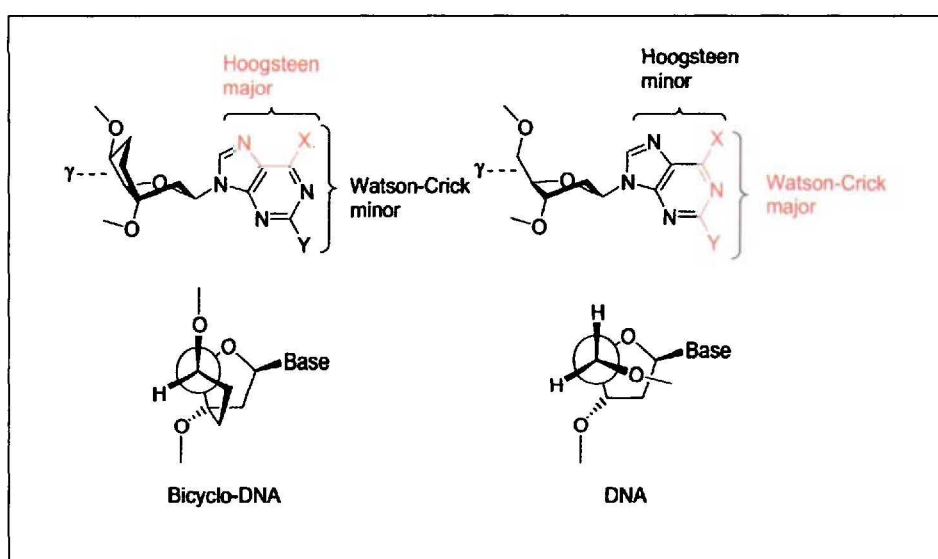


Fig. 2. The principal differences in the preferred recognition mode of complementary strands between oligopurine sequences of bicyclo-DNA and DNA (top) as a consequence of the difference in the preferred conformation around torsion angle γ (bottom).

In a different design project, we have replaced the cyclic furanose unit of a deoxynucleoside by a linear linker unit while maintaining the conformational restriction of the bonds directly involved in the repeating unit of the sugar phosphate backbone. Thus we synthesized and evaluated the two analogues shown in Fig. 4. While the number of bonds between the backbone unit and the base is equal to that of DNA in [3.2.1]DNA, it is longer by one bond in [3.2.1]amide-DNA. With this design we tried to determine whether a cyclic sugar unit is a necessary requirement for obtaining a complementary base-recognition system, and whether the 'three bond rule' for the distance between base and backbone must be strictly obeyed. At the outset this was not at all clear, since there did not exist a functional phosphodiester-based DNA analogue with acyclic connection of the nucleobases to the backbone. While we have studied a number of oligonucleotides in the case of [3.2.1]DNA, we are only begin-

ning to investigate the [3.2.1]amide-DNA.

We found the [3.2.1]DNA system to be a competent base-pairing system forming duplexes with natural DNA and RNA with thermal stability that is inferior by *ca.* 1–3 °C in T_m per base-pair compared to natural DNA [14]. Only anti-parallel and no-parallel duplexes were formed, with both natural DNA and itself as the complexing partner, which points towards a remarkable remote stereocontrol of the sugar-phosphate backbone in determining the polarity of strand association [15]. Furthermore, the Watson-Crick pairing mode is maintained as inferred from CD-spectroscopic analysis. Base-mismatch discrimination is similar compared to natural DNA. These facts corroborate our current view that a cyclic unit between base and backbone is not a necessary requirement for obtaining a competent, selective and complementary base-pairing system, provided that the backbone is structurally preorganized.

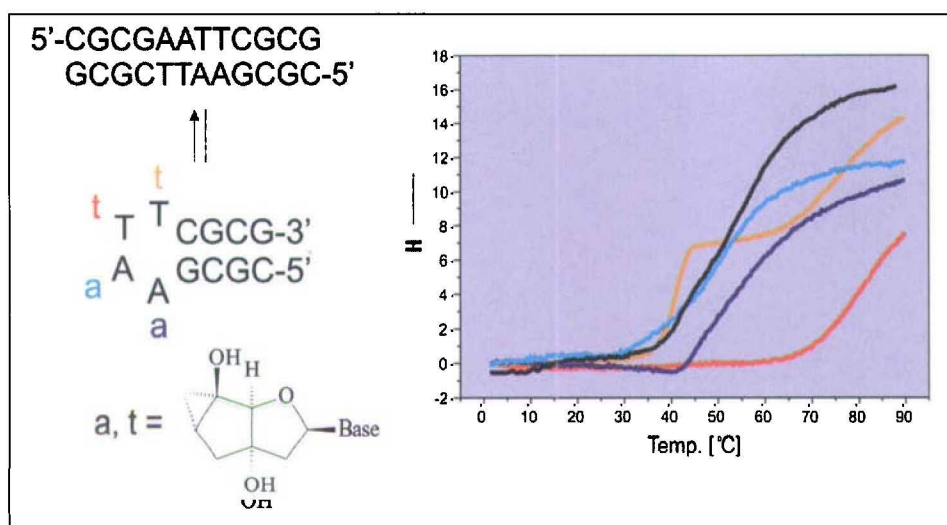


Fig. 3. Schematic view of the equilibrium between the biomolecular duplex and the tetra-loop hairpin structure of the indicated DNA-dodecamer (left) and corresponding UV-melting curves (right). The black curve corresponds to the melting profile of the unmodified DNA dodecamer, while the others correspond to that of the single tricyclo-DNA mutant sequences (indicated with small letters at the appropriate position). Some of the UV-melting curves display biphasic transitions (black, blue, orange) indicating the presence of both hairpin (transition at higher temperature) and bimolecular duplex (transition at lower temperature) in solution. The cyan curve is reminiscent of a single bimolecular duplex to single-strand transition, while the red curve corresponds to a monomolecular hairpin to single-strand transition.

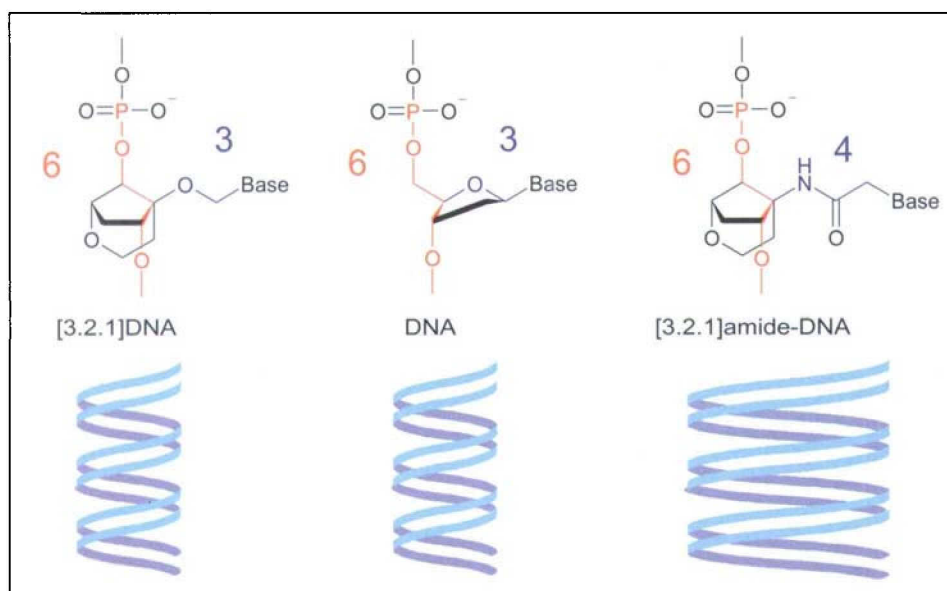


Fig. 4. Chemical structures of DNA, [3.2.1]-DNA and [3.2.1]amide-DNA. Structural elements in red correspond to the backbone elements and those in blue to the 'base-linker' elements. While the number of bonds for the backbone element is the same in all three systems, that of the 'base-linker' element is longer by one bond in [3.2.1]amide-DNA. This elongation is expressed in the cartoon of the double helix by a wider helix.

While we do not yet have a detailed picture of the properties of [3.2.1]amide-DNA, we have been able to show that a corresponding homoadenine decamer forms a duplex with the natural DNA-complement, as well as with a homothymidine decamer of the same backbone type of similar stability as natural DNA [16]. This is interesting in view of the increased distance between base and backbone. Preliminary investigations by CD-spectroscopy revealed an almost normal Watson-Crick duplex structure for the DNA/[3.2.1]amide-DNA hybrid duplex, but almost a perfect mirror image for the pure [3.2.1]amide-DNA decamer duplex, indicating a left-handed Watson-Crick double helical structure. This unexpected property suggests a potential degeneracy in enantioselection of the [3.2.1]amide-DNA system. Indeed we have shown very recently that the [3.2.1]amide-DNA

homoadenine decamer not only binds to a (natural) D-RNA-complement but also to the enantiomeric L-RNA with similar affinity, forming enantiomeric Watson-Crick double helices (Fig. 5). Thus [3.2.1]amide-DNA is the first example of a chiral, non-chiroselective base-pairing system. The structural basis for this degeneracy as well as the eventual consequences with respect to the molecular evolution of homochirality on earth are currently being investigated.

A completely different aspect of conformational restriction was the basis for our design of the PNA (polyamide nucleic acid) analogue OPA (Fig. 6). PNA is an achiral amide-based nucleic acid analogue which only has the nucleobases in common with DNA. Although structurally completely different than DNA, it displays very similar functional properties and forms stable duplex and triplex struc-

tures with RNA and DNA. The properties of PNA have recently been reviewed extensively [17].

Structural investigations of PNA/DNA and PNA/RNA complexes revealed that the tertiary amide function, connecting the base-linker with the backbone, is unidirectionally aligned with the carbonyl oxygen pointing in the direction of the C-terminus of the chain. In the PNA monomers, however, a mixture of both rotameric forms occurs in solution. With the knowledge from the X-ray structure that the carbonyl oxygen is not involved in any direct hydrogen bonding, it was appealing to ask whether this intrinsically conformationally labile amide unit could be replaced by an isostructural, configurationally fixed C=C double bond. Such a replacement would not only allow us to separately address the two (*E* and *Z*) rotameric forms of PNA in

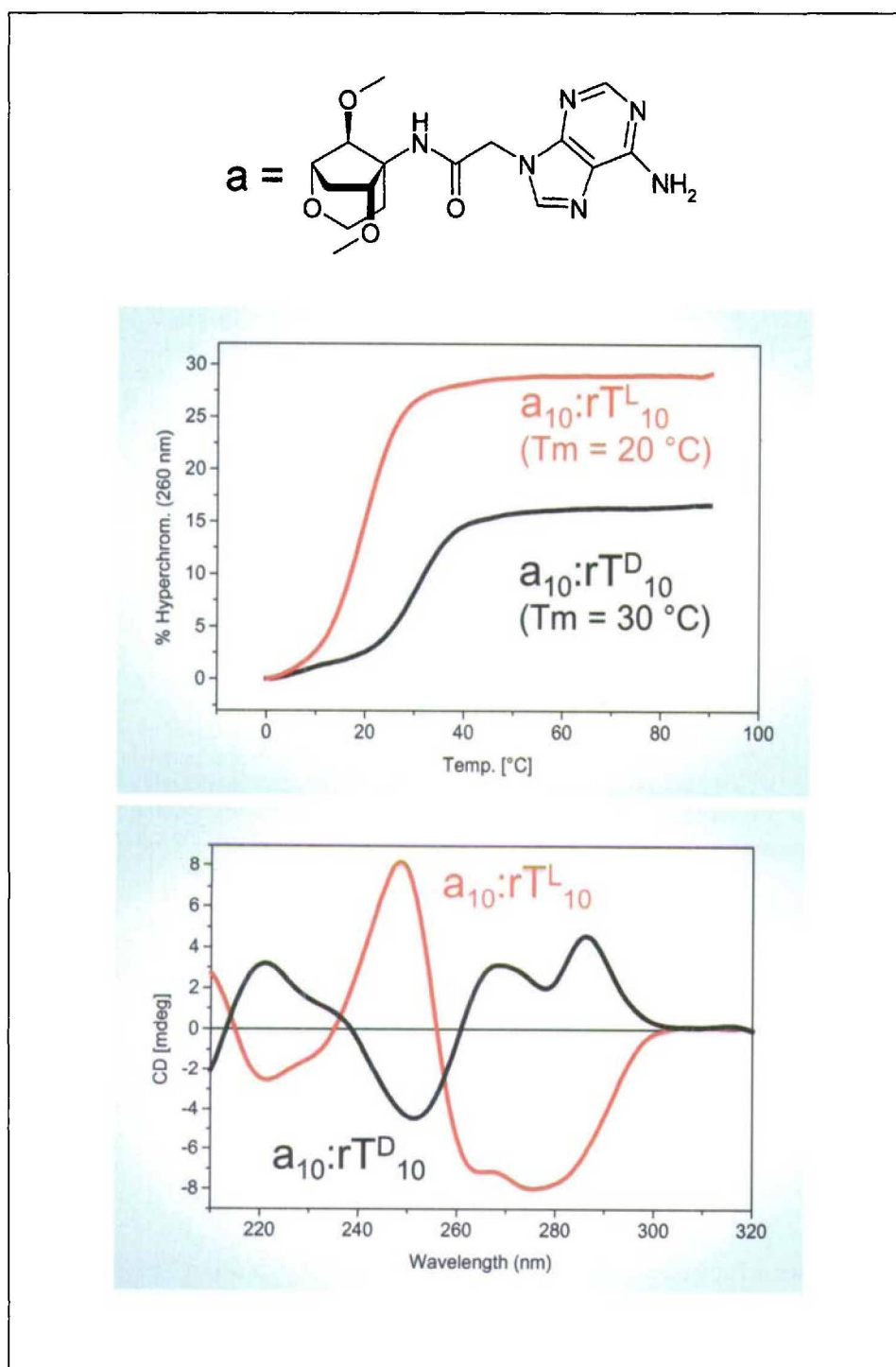


Fig. 5. UV-melting curves (top) and CD-spectra (bottom) of the [3.2.1]amide-DNA homoadenosine decamer duplex with the complementary RNA decamer built from D-ribothymidine (rT^D) or the enantiomeric (unnatural) L-ribothymidine (rT^L).

complex formation, but could also provide fundamental information about the consequences of replacing a hydrophilic by a hydrophobic functional group within an isostructural environment.

Indeed, our preliminary results show that *E*-OPA, which has its bases structurally arranged as in the complexed form of PNA, exhibits subtle differences in pairing to DNA compared to PNA. Homobasic thymine-containing decamers have lost their propensity for triple-helix formation and bind with reduced affinity to DNA. Furthermore, the preferred strand association in non self-complementary adenine- and thymine-containing *E*-OPA

oligomers has changed from antiparallel to parallel [18]. All the information we have at present indicates that non-trivial electrostatic properties of the tertiary amide function, rather than geometry, dominate complementary DNA recognition. In our view this is a prime example of our inability to predict and account for such subtle effects in the molecular design process.

The OPAs represent a convenient structural platform to investigate the role of electrostatic effects in detail. For example, the remaining H-atom of the double bond in *E*-OPA can be replaced by appropriate functional groups to study

hydrophobic interactions, dipole interactions and solvation effects separately. This will lead to a better understanding of the relative contribution of these effects to stability and may eventually yield PNA analogues with tailor-made chemical properties. Work in this direction is currently in progress.

Base-modified Oligonucleotide Analogues for DNA Triple-helix Formation

A related field of interest focuses on recognition of the major groove of dou-

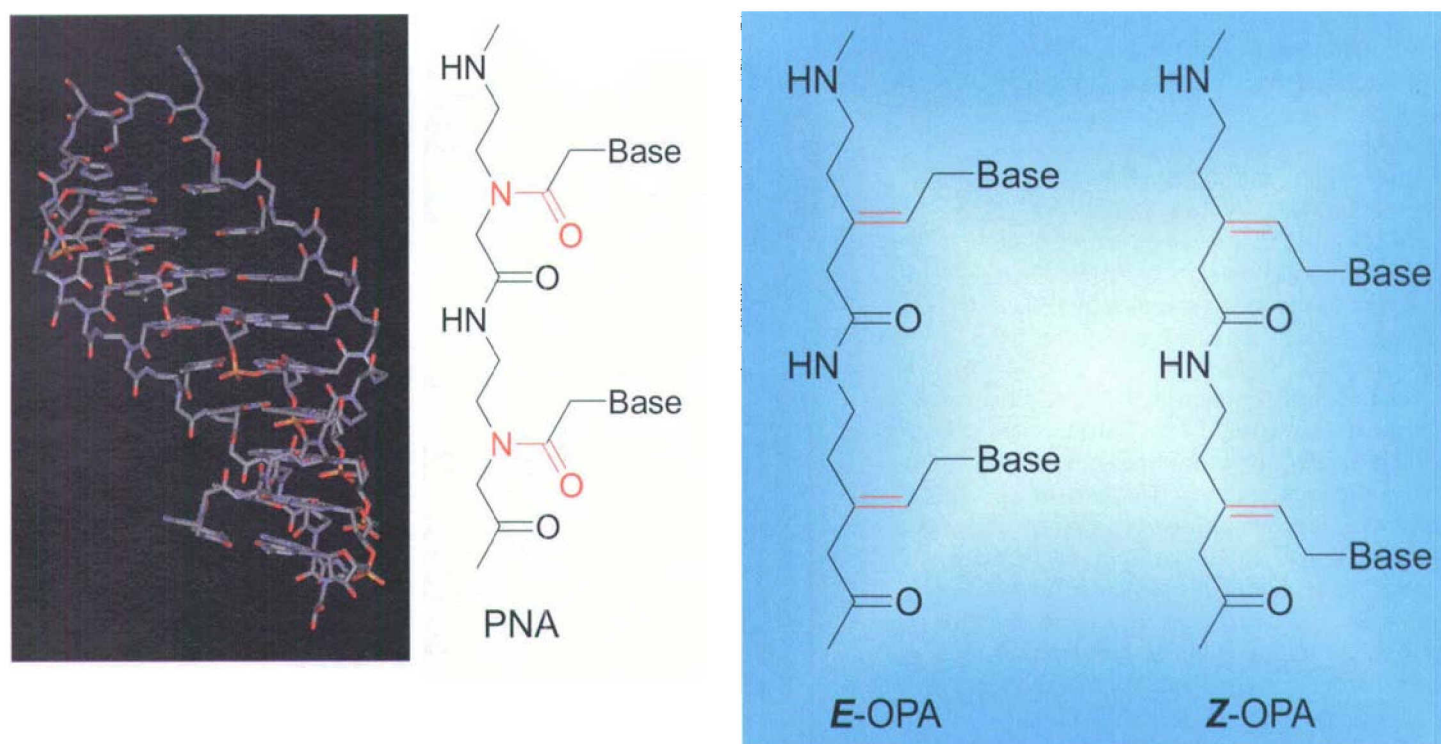


Fig. 6. Part of the X-ray structure of a DNA/PNA₂ triplex, displaying the non-involvement of the carbonyl oxygen of the central amide functionality between the base-linker and the backbone of the PNA-strand in intra- or interstrand hydrogen bonding (left), and molecular structures of the two analogues, *E*- and *Z*-OPA, developed in our laboratory (right).

ble-stranded DNA by oligonucleotides, forming triple-helical DNA structures. Sequence specific blockage of the major groove of DNA with oligonucleotides offers an alternative means of selectively interfering with the genetic expression machinery at the level of transcription. This approach is often referred to as the 'antigene strategy'. There are two stable triple helical binding motifs known today (Fig. 7). In the py:pu-py motif an oligonucleotide third strand consisting exclusively of pyrimidine bases binds in a parallel fashion to the purine strand of the underlying duplex forming T:A-T and C+:G-C base-triples. In the alternative motif, a purine-rich third strand binds in an antiparallel fashion to the purine strand of the duplex to give specific A:A-T, T:A-T, or G:G-C base-triples. The requirement for protonation of the cytosine base in the py:pu-py motif, an unfavorable process at neutral pH, leads to a strong reduction of affinity under physiological conditions. This has been identified as one of the major limitations of the parallel triple-helical binding motif. Furthermore, for both recognition motifs it is absolutely required that the target DNA duplex consists of a homopurine/homopyrimidine sequence tract. This requirement severely limits the range of amenable DNA-targets and has triggered an intense search for sugar and/or base-

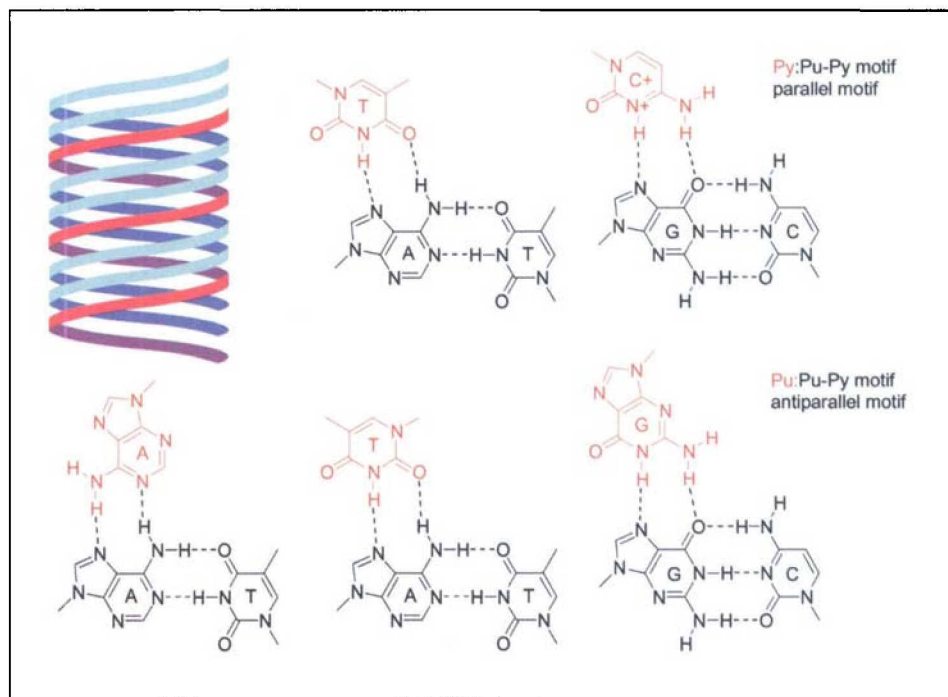


Fig. 7. Schematic view of a third strand oligonucleotide (red) aligned in the major groove of a DNA double helix (light blue) forming a DNA triple-helix (top, left) together with the base-triples occurring in the parallel (top, right) and the antiparallel (bottom) triple-helical binding motif.

modified oligonucleotide third strands able to overcome this sequence restriction.

In order to circumvent the unfavorable properties associated with the protonation of cytosine by base design we focused on the simple concept of remov-

ing unnecessary functional groups and heteroatoms in order to increase the intrinsic basicity of the required aromatic ring-nitrogen. We thus came up with an aminopyridine-containing C-glycoside as a replacement for cytidine (Fig. 8). This heterocycle shows increased basici-

ty by *ca.* 2 pKa units. We were indeed able to show that replacement of 5-methyl cytosine (MeC) by the respective aminopyridine unit (MeP) leads to a notable increase in triplex stability even at a pH as high as 7.5, and in sequence contexts where natural DNA third strands notoriously fail to form triple helices. Furthermore, there was no decrease in selectivity of triplex formation as revealed by DNase I footprint analysis in a competition experiment (Fig. 8) [19].

The central problem in DNA triplex chemistry, however, is its restriction to homopurine/homopyrimidine target sequences. Despite almost 15 years of intense research, a solution to this problem is still elusive. In one of our projects aimed at overcoming the sequence recognition requirements, we proposed the concept of targeting natural pyrimidine bases by one conventional bond and one non-classical (hypothetical) CHO hydrogen bond. This concept arose from earlier observations in experiments with N7-hypoxanthine as a third strand base [20], and from X-ray data on a RNA hexamer duplex [21][22], in which the presence of such non-classical CHO hydrogen bonds was either postulated or observed. We were able to demonstrate that up to three uracil bases as part of 15-mer DNA target cassette can be recognized by a third strand oligonucleotide containing N7-hypoxanthine opposite the uridine residues [23]. This corresponds to a DNA target site containing 20% pyrimidine bases. We have recently expanded on this theme and have shown that the base 4-deoxythymine (4^HT), as part of a third strand, recognizes a cytosine base within the purine target sequence with high selectivity but moderate affinity [24] (Fig. 9).

At this point we have no unambiguous confirmation of the existence of such unconventional hydrogen bonds. Nor do we know yet whether such an arrangement of functional groups exhibits a small positive energetic contribution to triplex stability or whether it is just neutral and non-repulsive. In any case, the integration of such energetically weak but potentially selectivity enhancing elements into the design of third strand bases is an interesting addition to standard approaches involving classical hydrogen bonding schemes.

Outlook and Conclusions

Although antisense research has reached a plateau after a steep initial burst phase, there will definitely be fur-

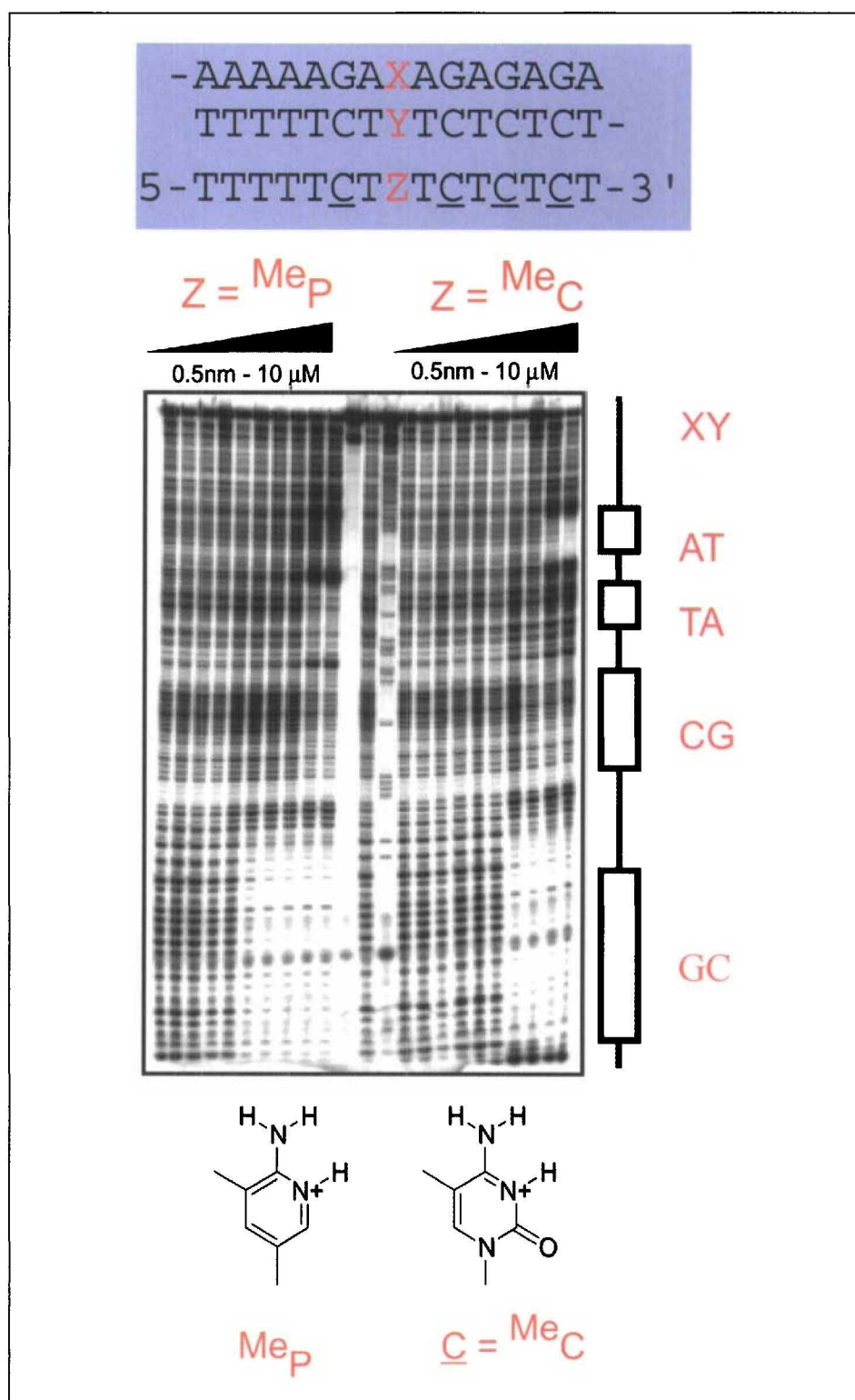


Fig. 8. DNase I footprint assay of third strand binding to duplex DNA containing four target cassettes (top) that are identical with the exception of the base-pair (XY) opposite the base of interest (Z) of the third strand. Protected (white) zones on the gel correspond to areas where third strand binding occurred. The left part of the gel (third strand with $Z = MeP$) clearly shows strong binding to the GC cassette and no binding, even at high mM concentrations to all of the other cassettes.

ther progress in the future, and second or third generation oligonucleotide analogues with improved chemical and pharmacological properties against various targets will be developed. Furthermore, antisense technology will be of invaluable help in functional genomics in order

to elucidate the function of newly identified proteins or receptors. More sophisticated DNA-based analytical tools will continue to replace more traditional technologies in applications ranging from molecular medicine to the forensic sciences. Biotechnology will see new tools

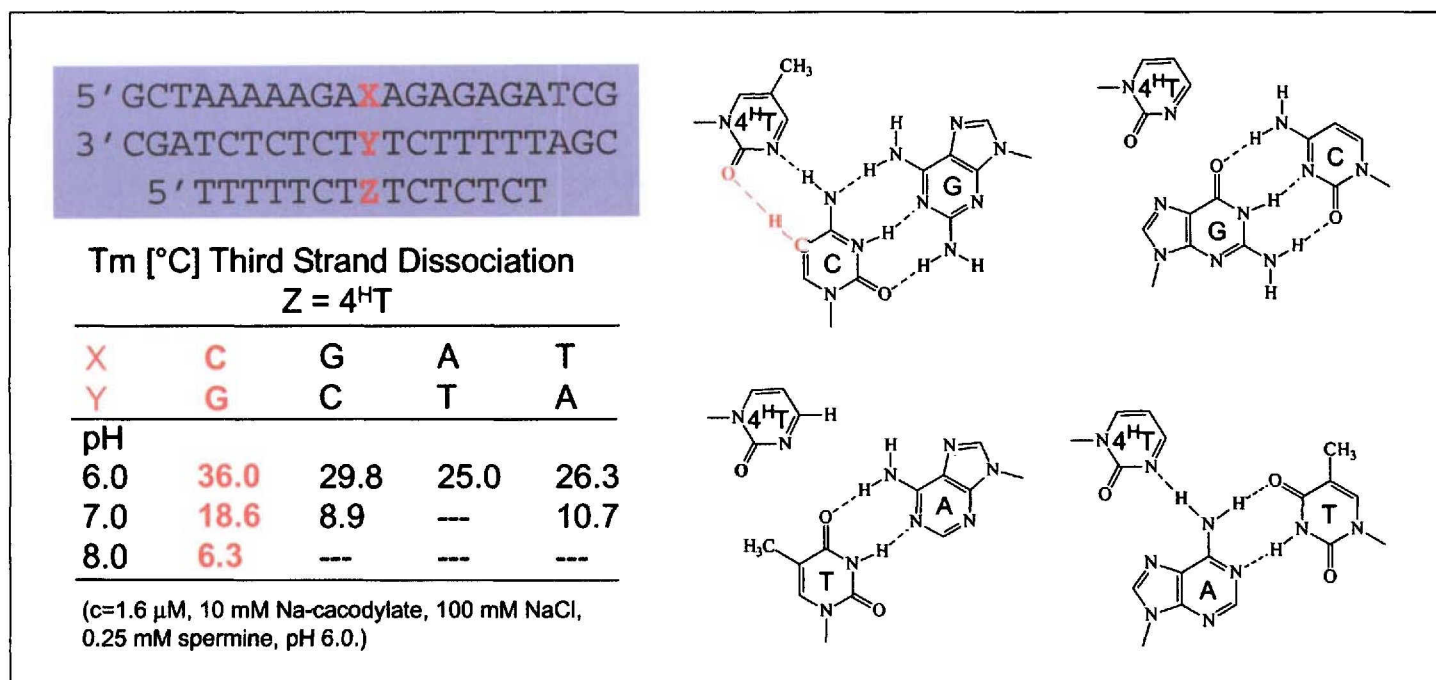


Fig. 9. Sequence context and four base-triple arrangements with 4-deoxythymine (4^{HT}) as third strand base. The table displays T_m data from third strand melting at the conditions indicated. The arrangement X,Y,Z = C,G, 4^{HT} corresponds to the most stable of the four possibilities.

arising from the DNA-based expansion of the genetic alphabet, enhancing the performance of expression systems of proteins containing unnatural amino acids. Last but not least, a whole new area of nanotechnology and biocomputing based on intelligent informational (bio) macromolecules, like DNA, with distinct logical properties is being developed. Taking all this into account there is 'plenty of room down at the bottom' to create new oligonucleotide analogues with distinct chemical, physical and biological properties from simple monomeric units, yet to be designed.

Acknowledgments

I would like to thank all my past and present coworkers for their enthusiastic contributions to this challenging, interdisciplinary field of research. Financial support from the Swiss National Science Foundation, the Roche Foundation, as well as from Novartis AG, Basel and Roche Diagnostics GmbH, Penzberg, is gratefully acknowledged.

Received: February 20, 2001

- [1] M.L. Stephenson, P.C. Zamecnik, *Proc. Natl. Acad. Sci. USA* **1978**, *75*, 285–288.
 [2] P.C. Zamecnik, M.L. Stephenson, *Proc. Natl. Acad. Sci. USA* **1978**, *75*, 280–284.
 [3] M. Chee, R. Yang, E. Hubbell, A. Berno, X.C. Huang, D. Stern, J. Winkler, D.J.

- Lockhart, M.S. Morris, S.P. Fodor, *Science* **1996**, *274*, 610–614.
 [4] N.C. Seeman, *Trends Biotechnol.* **1999**, *17*, 437–443.
 [5] N.C. Seeman, *Curr. Opin. Struct. Biol.* **1996**, *6*, 519–526.
 [6] A.P. Alivisatos, K.P. Johnsson, X. Peng, T.E. Wilson, C.J. Loweth, M.P. Bruchez Jr, P.G. Schultz, *Nature* **1996**, *382*, 609–611.
 [7] C.A. Mirkin, R.L. Letsinger, R.C. Mucic, J.J. Storhoff, *Nature* **1996**, *382*, 607–609.
 [8] M.H. Caruthers, *Acc. Chem. Res.* **1991**, *24*, 278–284.
 [9] J. Hunziker, H.-J. Roth, M. Böhringer, A. Giger, U. Diederichsen, M. Göbel, R. Krishnan, B. Jaun, C.J. Leumann, A. Eschenmoser, *Helv. Chim. Acta* **1993**, *76*, 259–352.
 [10] D.J. Cram, *Angew. Chem.* **1988**, *100*, 1041–1052.
 [11] M. Meldgaard, J. Wengel, *J. Chem. Soc. Perkin 1* **2000**, 3539–3554.
 [12] M. Bolli, C. Litten, R. Schütz, C.J. Leumann, *Chem. Biol.* **1996**, *3*, 197–206.
 [13] R. Steffens, C.J. Leumann, *J. Am. Chem. Soc.* **1999**, *121*, 3249–3255.
 [14] C. Epple, C.J. Leumann, *Chem. Biol.* **1998**, *5*, 209–216.
 [15] B.M. Keller, C.J. Leumann, *Angew. Chem.* **2000**, *112*, 2367–2369.
 [16] A. Egger, C.J. Leumann, *Synlett* **1999**, *S1*, 913–916.
 [17] E. Uhlmann, A. Peyman, G. Breipohl, D.W. Will, *Angew. Chem.* **1998**, *110*, 2954–2983.
 [18] R. Schütz, M. Cantin, C.D. Roberts, B. Greiner, E. Uhlmann, C.J. Leumann, *Angew. Chem.* **2000**, *112*, 1305–1308.
 [19] S. Hildbrand, A. Blaser, S.P. Parel, C.J. Leumann, *J. Am. Chem. Soc.* **1997**, *119*, 5499–5511.
 [20] J. Marfurt, S.P. Parel, C.J. Leumann, *Nucleic Acids Res.* **1997**, *25*, 1875–1882.
 [21] M.C. Wahl, M. Sundaralingam, *TIBS* **1997**, *22*, 97–102.
 [22] M.C. Wahl, S.T. Rao, M. Sundaralingam, *Nature Struct. Biol.* **1996**, *3*, 24–31.
 [23] J. Marfurt, C.J. Leumann, *Angew. Chem.* **1998**, *110*, 184–187.
 [24] I. Prévot-Halter, C.J. Leumann, *Bioorg. Med. Chem. Lett.* **1999**, *9*, 2657–2660.