BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

# Nucleic acid chemistry

Edited by Hans-Achim Wagenknecht

# Imprint

Beilstein Journal of Organic Chemistry
www.bjoc.org
ISSN 1860-5397
Email: journals-support@beilstein-institut.de

The *Beilstein Journal of Organic Chemistry* is published by the Beilstein-Institut zur Förderung der Chemischen Wissenschaften.

Beilstein-Institut zur Förderung der Chemischen Wissenschaften
Trakehner Straße 7–9
60487 Frankfurt am Main
Germany
www.beilstein-institut.de

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

# Nucleic acid chemistry

Hans-Achim Wagenknecht

Editorial

Open Access

Address:
Institute of Organic Chemistry, Karlsruhe Institute of Technology (KIT), Fritz-Haber-Weg 6, 76131 Karlsruhe, Germany

Email:
Hans-Achim Wagenknecht - wagenknecht@kit.edu

Keywords:
nucleic acids

*Beilstein J. Org. Chem.* **2014,** *10,* 2928–2929.
doi:10.3762/bjoc.10.311

Received: 14 November 2014
Accepted: 20 November 2014
Published: 10 December 2014

This article is part of the Thematic Series "Nucleic acid chemistry".

Guest Editor: H.-A. Wagenknecht



Since the discovery of the structure of the DNA double helix in 1953 by Watson and Crick [1], we know that DNA is of critical importance, carrying the genetic information for all living organisms. Only a few years later appeared the first reports on the chemical synthesis of oligonucleotides with a natural 3'-5' phosphodiester linker by Michelsen and Todd [2]. It took another two decades until the first step towards building block chemistry (the so-called phosphoramidte chemistry) was published by Letsinger and Lunsford, enabling efficient and fully synthetic access to oligonucleotides [3]. These authors discovered that DNA building blocks based on phosphorous(III) were significantly more reactive than phosphordiesters or -triesters. Finally, this approach using phosphoramidites as nucleoside building blocks was significantly further developed in 1981 by Beaucage and Caruthers [4]. Since then, oligonucleotides of up to 50-mers in length have become available by an extremely efficient solid-phase methodology that runs automatically on machines, so-called DNA synthesizers. With special care during each synthesis step, even longer oligonucleotides can be prepared in a similar, almost routine fashion. The field of nucleic acid chemistry has exploded: natural and artificial functionalities, as well as probes, markers or other biologically active molecules, can be synthetically introduced into DNA (as well as RNA) by preparing the corresponding artificial DNA and RNA building blocks. More recently, the chemistry was further developed for building blocks that are otherwise synthetically not obtainable. In such cases, postsynthetic strategies can allow for the desired oligonucleotides modification. This becomes an even more important issue for functionalities, probes or biologically relevant molecules that are incompatible with routine phosphoramidite chemistry. Although nucleic acid chemistry appears to be a mature part of organic and bioorganic chemistry, the many questions that are still being raised by research in biology and chemical biology give sufficient motivation to continue to synthesize new nucleic acid probes and thereby further develop nucleic acid chemistry. Moreover, in addition to their biological functionality, DNA and RNA are considered as increasingly important architectures and scaffolds for two- and three-dimensional objects, networks and materials for nanosciences. In conclusion, nucleic acid chemistry continues to maintain its appeal and affords good reason to focus on in this Thematic Series of the Beilstein Journal of Organic Chemistry.

Hans-Achim Wagenknecht

Karlsruhe, November 2014

## References

1. Watson, J. D.; Crick, F. H. C. *Nature* **1953,** *171,* 737–738. doi:10.1038/171737a0
2. Michelson, A. M.; Todd, A. R. *J. Chem. Soc.* **1955,** 2632–2638. doi:10.1039/jr9550002632
3. Letsinger, R. L.; Lunsford, W. B. *J. Am. Chem. Soc.* **1976,** *98,* 3655–3661. doi:10.1021/ja00428a045
4. Beaucage, S. L.; Caruthers, M. H. *Tetrahedron Lett.* **1981,** *22,* 1859–1862. doi:10.1016/S0040-4039(01)90461-7

## License and Terms

Beilstein Journal of Organic Chemistry

# A new building block for DNA network formation by self-assembly and polymerase chain reaction

Holger Bußkamp‡, Sascha Keller‡, Marta Robotta, Malte Drescher and Andreas Marx*§

Full Research Paper

Open Access

Address:
Department of Chemistry and Konstanz Research School Chemical Biology, University of Konstanz, Universitätsstraße 10, 78457 Konstanz, Germany

Email:
Andreas Marx* - Andreas.Marx@uni-konstanz.de

* Corresponding author ‡ Equal contributors
§ Tel. +49-7531-88 5139; Fax. +49- 7531-88 5140

Keywords:
AFM; branched DNA; DNA; DNA polymerase; nanotechnology; nucleic acids; PCR; self-assembly

*Beilstein J. Org. Chem.* **2014,** *10,* 1037–1046.
doi:10.3762/bjoc.10.104

Received: 02 February 2014
Accepted: 10 April 2014
Published: 07 May 2014

This article is part of the Thematic Series "Nucleic acid chemistry".

Guest Editor: H.-A. Wagenknecht



## Abstract

The predictability of DNA self-assembly is exploited in many nanotechnological approaches. Inspired by naturally existing self-assembled DNA architectures, branched DNA has been developed that allows self-assembly to predesigned architectures with dimensions on the nanometer scale. DNA is an attractive material for generation of nanostructures due to a plethora of enzymes which modify DNA with high accuracy, providing a toolbox for many different manipulations to construct nanometer scaled objects. We present a straightforward synthesis of a rigid DNA branching building block successfully used for the generation of DNA networks by self-assembly and network formation by enzymatic DNA synthesis. The Y-shaped 3-armed DNA construct, bearing 3 primer strands is accepted by *Taq* DNA polymerase. The enzyme uses each arm as primer strand and incorporates the branched construct into large assemblies during PCR. The networks were investigated by agarose gel electrophoresis, atomic force microscopy, dynamic light scattering, and electron paramagnetic resonance spectroscopy. The findings indicate that rather rigid DNA networks were formed. This presents a new bottom-up approach for DNA material formation and might find applications like in the generation of functional hydrogels.

## Introduction

DNA has found applications in the field of nanotechnology due to its inherent properties. The simplicity and predictability of DNA secondary structure are of outstanding potential for the design of self-assembled architectures [1-3]. Inspired by naturally existing self-assembled DNA architectures known as Holliday junctions, Seeman envisioned the approach to organize DNA with branched DNA (bDNA) and thereby initiated the field of structural DNA nanotechnology [4]. Since then several

reports have described the generation of bDNAs self-assembling to predesigned architectures with dimensions on the nanometer scale [5-9]. Based on this, numerous examples of 2D arrays [10-14], DNA origami [15] and complex 3D DNA nanostructures [16-20] were generated by intelligent algorithmic assembly design strategies [5-9]. Besides, the directed assembly of cells was achieved by using duplex DNA to drive the connections of cells and thereby providing access to microtissues [21,22] or extracellular matrices [23,24] by DNA-based 2D-arrays.

DNA bears the inherent potential, that nature evolved a large toolbox of different enzymes for manipulation of DNA. These enzymes can be used to manipulate DNA for the construction of DNA nanometer scaled objects. For example, DNA ligases were applied to covalently attach DNA strands to each other to form covalently linked objects [16,18,25,26]. Furthermore, by using branched DNA constructs and ligases, a DNA hydrogel was generated [27-29]. Luo et al. used a DNA based network, manufactured in that fashion, that can act as a protein producing gel and can be used as an efficient cell-free translation system [30]. Recently, ordered 2D DNA scaffolds were reported in which a nanometer precise arrangement of enzymes on these scaffolds leads to efficient enzymatic communication [31]. DNA polymerases have been applied for the assembly of DNA nanostructures. Joyce et al. employed a DNA polymerase to synthesize long single-stranded DNA that folds into an octahedron by assistance of scaffolding DNA oligomers [17]. In another approach rolling circle amplification was used for enzymatic amplification of DNA nanostructures [32,33]. Recently, this approach was extended to operate in cells [34,35].

Furthermore, it has been shown that branched DNA constructs can form materials by self-assembly [36]. Richert et al. were able to generate DNA based materials based on branched DNA molecules which are non-covalently bound to each other by only hybridization of 2 nucleotides [37]. Interestingly, the formed DNA networks are remarkably stabilized (up to 95 °C) compared to the non-branched counterparts.

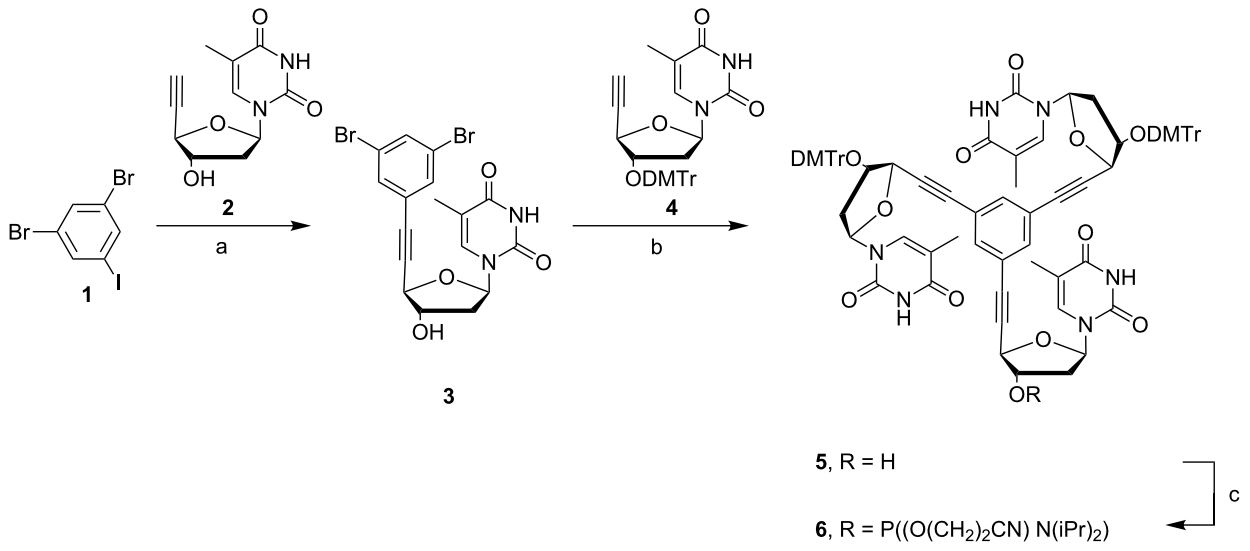
We previously reported an approach to construct three dimensional DNA networks that were generated and amplified by DNA polymerase chain reactions (PCR). In order to construct the network we developed covalently connected, 3-armed bDNA constructs (Y-motif) that act as primer and reverse primer strands in PCR [38-40]. The branching of the DNA was realized via a flexible alkyl chain that was connected to the nucleobase. Although the primer strands were covalently connected, they were accepted by a DNA polymerase and DNA networks formed by the enzyme. Based on this observation, we aim at investigating the impact of geometric constrains within the covalently branching unit on the network forming behavior of the branched DNA and the ability of DNA polymerases to form DNA networks by PCR. Thus, we developed a synthetic strategy for branched DNA by using a rigid branching point (Bp), based on the 1,3,5-triethynylbenzene scaffold. After the synthesis, the branched DNA was investigated towards its properties in network formation by self-assembly and PCR. We found that, despite the geometric restriction of the branching unit, the enzymatic generation of complex DNA networks by PCR was feasible. The novel generated DNA networks were investigated by agarose gel electrophoresis, atomic force microscopy, dynamics light scattering, and electron paramagnetic resonance spectroscopy on surfaces and in solution.

## Results and Discussion

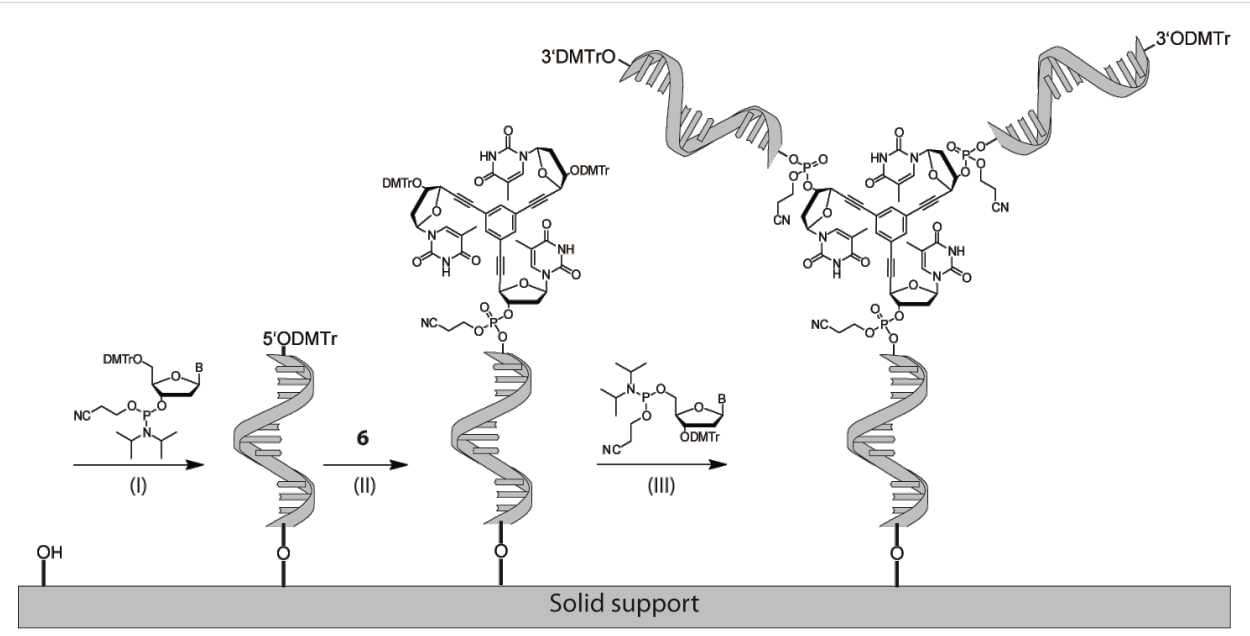
**Design and synthesis of the branching molecule**. In order to investigate the impact of rigidity of the branching core on DNA hybridization and usage of the constructs for network formation by PCR, we aimed at synthesizing a branching molecule based on the 1,3,5-triethynylbenzene scaffold. The oligonucleotides should be directly fused to the benzene ring via the three acetylene modifications resulting in a Y-shape topology. Thereby a rigid core with reduced degrees of rotation will be generated in contrast to other approaches that used more flexible branching molecules. The synthesis strategy was designed to meet the requirements of standard DNA solid support synthesis. Stepwise Sonogashira reaction was employed using the higher reactivity of iodide in the presence of bromide within 1,3-dibromo-5-iodobenzene (**1**) employing the known compounds **2** [41] and **4** [42] yielding **5** in acceptable yields (Scheme 1). Finally, compound **5** was transformed into **6** by phosphitylation resulting in a building block that bears protection groups and reactive groups that are standard in solid phase DNA oligonucleotide synthesis.

**Synthesis of branched oligonucleotides.** DNA oligonucleotide synthesis was performed at 0.2 μmol scale (trityl-on mode) employing the standard phosphoramidites and **6** which was diluted in a mixture containing 10% $CH_2Cl_2$ in $CH_3CN$ to a final concentration of 0.12 M. 3000 Å LCAA-CPG support was used, derivatized with the respective 3'-nucleotide of the respective DNA oligomers. Since we later intended to investigate whether the oligonucleotide branches are used as primers in PCR (vide infra), the oligonucleotides have to terminate with a free 3'-hydroxy group. This requires a particular synthesis strategy (Scheme 2).

The synthesis strategy was adapted in a way that all branches have the same sequence and terminate with a free 3'-OH group required for processing by DNA polymerases. This was




**Scheme 1:** Reagents and conditions: (a) $PdCl_2(PPh_3)_2$, DMF, CuI, $NEt_3$, 55 °C, microwave, 82%; (b) $PdCl_2(PPh_3)_2$, DMF, CuI, $NEt_3$, 80 °C, microwave, 48%; (c) $CH_2Cl_2$, 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite, *N*,*N*-diisopropylethylamine, 0 °C, 84%.




**Scheme 2:** Stepwise solid-phase synthesis for branched oligonucleotides. (I): The first oligonucleotide branch is synthesized in 3'–5' direction using standard phosphoramidites. (II): Incorporation of the branching point by usage of **6**. (III): Simultaneous synthesis of the two remaining oligonucleotide branches in 5'–3' direction using inverse protected phosphoramidites.

achieved by the employment of standard phosphoramidites until the incorporation of the branching unit **6**. Afterwards, the inverse-phosphoramidite strategy was used for the synthesis of both DNA strands. Following this approach a series of branched oligonucleotides were synthesized (Figure 1). The average coupling yield was always higher than 95% requiring a coupling time of 5 min only for the reaction of **6**. DNA-oligomers were purified twice by HPLC and characterized by ESI–IT–MS (see Supporting Information File 1).

**Characterization of bDNA by thermal denaturation studies and CD spectroscopy.** Formation of stable duplexes with complementary DNA strands is a prerequisite for the employment of the bDNA **ODN I** and **ODN II** in PCR experiments.


**Figure 1:** (A) Depiction of synthesized branched oligonucleotides; (B) Sequences of all synthesized branched oligonucleotides. Bp: branchpoint.

Thus, we investigated duplex formation properties. In order to investigate whether the thymidine residues that are directly linked to the benzene core are amenable to participate in duplex formation, complementary oligonucleotides with varied lengths were investigated and compared to linear, non-branched reference duplexes. The thermal denaturation studies (see Supporting Information File 1) indicate that dependent on length, Y-shaped bDNA **ODN I** behave comparably to the linear non-branched counterpart **7** (Table 1). Increases in $T_m$ were observed with increasing duplex length. Noteworthy, the obtained results show that the terminal thymidine that is covalently connected to the benzene core via a rigid ethylene bridge is amenable to contribute to duplex stability. This was evidenced by an increase in $T_{m,}$ when **ODN I** was hybridized to **10** in comparison to the one nucleotide shorter **9**.

Next, bDNA constructs were characterized by circular dichroism (CD) spectroscopy. The CD spectra of Y-shaped bDNA **ODN I** bound to oligonucleotides **8**–**11**, respectively, show dichroic peaks similar to those of unmodified DNA duplexes, indicating that bDNA maintain the B-DNA form (see Supporting Information File 1).

As the Y-shaped bDNA is able to hybridize with complementary linear strands, we next investigated, if these constructs can undergo self-assembly to form DNA networks. For this purpose, self-complementary (sc) bDNA constructs were synthesized (**ODN-sc-I** to **ODN-sc-IV**, cf. Figure 1) and the melting characteristics were addressed by thermal denaturation studies (cf. Figure 2).

The self-complementary oligonucleotides (5 μM) were annealed in 10 mM triethylammonium acetate buffer at pH 7 in the presence of sodium chloride (150 mM), sodium chloride and magnesium chloride (150 mM and 100 mM, respectively) or in the absence of salts. The solutions were heated and the UV absorbance was recorded in dependence on the temperature.

**Table 1:** Thermal denaturation studies comparing linear and branched oligonucleotides hybridization. Incorporated phosphoramidite **6** is depicted as Bp for branched **ODN I**.[a]

| Linear DNA | | | bDNA | | |
|---|---|---|---|---|---|
| duplex | | $T_m$ [°C] | duplex | | $T_m$ [°C] |
| 5'-TGGTGATGGTGATGGT<br>3'-ACCACTACCACTACCACT | **8 7** | 61.5 | 5'-TGGTGATGGTGATGGT<br>(3'-ACCACTACCACTACCACT)$_3$ Bp | **8 ODN I** | 60.4 |
| 5'-TGGTGATGGTGATGGTG<br>3'-ACCACTACCACTACCACT | **9 7** | 63.0 | 5'-TGGTGATGGTGATGGTG<br>(3'-ACCACTACCACTACCACT)$_3$ Bp | **9 ODN I** | 62.4 |
| 5'-TGGTGATGGTGATGGTGA<br>3'-ACCACTACCACTACCACT | **10 7** | 65.2 | 5'-TGGTGATGGTGATGGTGA<br>(3'-ACCACTACCACTACCACT)$_3$ Bp | **10 ODN I** | 64.0 |
| 5'-TGGTGATGGTGATGGTGAC<br>3'-ACCACTACCACTACCACT | **11 7** | 63.6 | 5'-TGGTGATGGTGATGGTGAC<br>(3'-ACCACTACCACTACCACT)$_3$ Bp | **11 ODN I** | 62.3 |

[a]Temperatures were determined with ±0.5 °C accuracy.

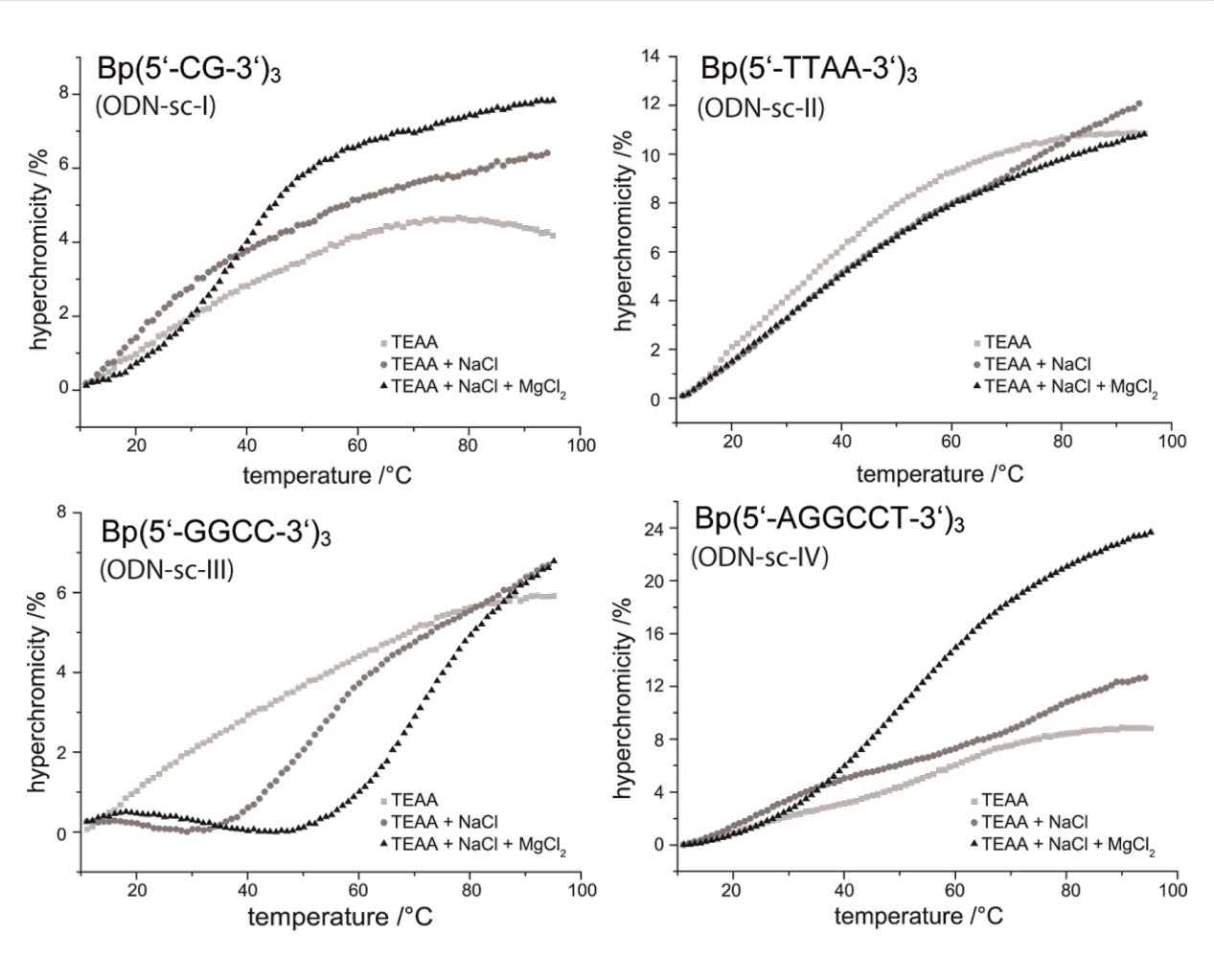



**Figure 2:** Thermal denaturating studies of self-complementary branched oligonucleotides. Different conditions are indicated in the figure. TEAA: triethylammonium acetate.

Interestingly, self-complementary constructs with high GC-content show higher melting temperatures and a strong increase of the melting temperature on the addition of $Mg^{2+}$, whereas the linear control oligonucleotides showed no melting in all buffers tested (data not shown). In magnesium ion-containing buffer the CG 2-mer (**ODN-sc-I**) shows a higher melting temperature (37.6 °C) as the **ODN-sc-II** 4-mer (24.3 °C). The GC-rich 4-mer **ODN-sc-III** shows high melting temperatures even in the presence of only sodium chloride (58.0 °C) and an even higher melting temperature in the presence of additional magnesium (73.0 °C). The melting temperature in presence of sodium chloride and magnesium is even higher than the respective melting temperature of the 6-mer **ODN-sc-IV** (50.7 °C). All in all, one can conclude that the self-assembly of GC-rich self-complementary constructs is stronger than the self-assembly of AT containing constructs. Furthermore, the self-assembly is much stronger in the presence of magnesium ions for the GC-rich constructs.

**Enzyme catalyzed network growth.** Next we investigated whether the synthesized **ODN I** and **II** are suitable primers for DNA network formation by PCR. We used a 1062 nt open reading frame of human DNA polymerase β as template and *Thermus aquaticus* (*Taq*) DNA polymerase for amplification. In earlier studies [17] of flexible Y-motifs strong dependence of the network formation on the annealing temperature was found. Therefore, we varied this parameter of PCR and investigated product formation by agarose gel electrophoresis (Figure 3).

Depending on the cycle number the formation of slower migrating products was observed. At the highest cycle number (28 cycles) at all investigated annealing temperatures ranging from 52–66 °C the formation of DNA networks with hardly any mobility was observed. Noteworthy, using linear primer strands of the same sequence resulted in the formation of the expected linear reaction products that migrated as expected (see Supporting Information File 1). No amplification products were

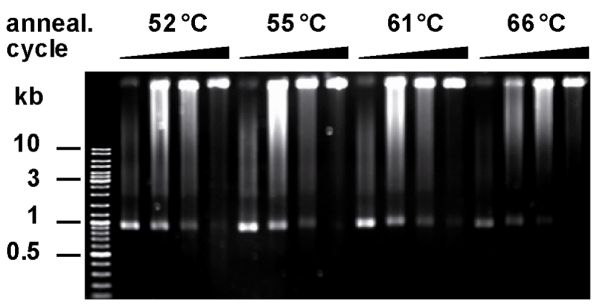



**Figure 3:** Generation of DNA networks with *Taq* DNA polymerase from 1062 nt template using primer **ODN I** and **ODN II**. Monitoring of DNA network growth influenced by varying of annealing temperature (as indicated) and increasing cycles (10, 18, 22, 28, respectively).

observed in the absence of template DNA under identical conditions (see Supporting Information File 1).

**Characterization of DNA networks.** The formed DNA networks were studied by atomic force microscopy (AFM) using the tapping mode [43,44]. The original sample solutions were diluted to a total amount of 10 ng/µL DNA for AFM measurements with buffer (10 mM Tris, pH 7.4, 1 mM $NiCl_2$). Freshly cleaved mica was incubated with the sample following a multistep protocol (see Supporting Information File 1).

As control, the PCR products of standard linear primer strands were investigated first (Figure 4D). Long linear double stranded DNA (dsDNA) is flexible. During AFM measurement linear or coiled structures with a height ranging from 0.5 to 1.1 nm (theoretic diameter 2 nm) and a width of about 15 nm were observed. Due to the force, which is applied by the scanning tip, the DNA is flattened, which might explain the smaller height of the observed DNA [45]. The increased width of the observed objects is a result of the finite size of the scanning tip, leading to a shape broadening of objects [46]. The PCR products derived from branched primer strands **ODN I** and **II** were found to be extended DNA networks (Figure 4A–C). An overview scan of 10 × 10 µm (Figure 4A) showed the diversity of shapes. Further AFM scans with higher resolution (Figure 4B,C) depicted DNA networks with dimensions from 0.6 to 2 µm in the surface dimension and a height up to 7 nm (see Supporting Information File 1). These findings correlate well with the decreased mobility of the structures in the agarose gel electrophoresis (Figure 3) demonstrating one covalently connected migrating DNA molecule. The irregular shapes of the DNA networks might result from the collapse of three-dimensional structures forced by ionic interactions and induced by nickel ions on the mica surface.

More information about the DNA network character was observed taking the first derivative of the height output channel resulting in Figure 4E. The principle of this mathematic operation is to assume a local maximum at an edge resulting from cross section measurements of objects (see Supporting Information File 1). Using this operation the data depicted in Figure 4C could be transferred into the data depicted in Figure 4E. The

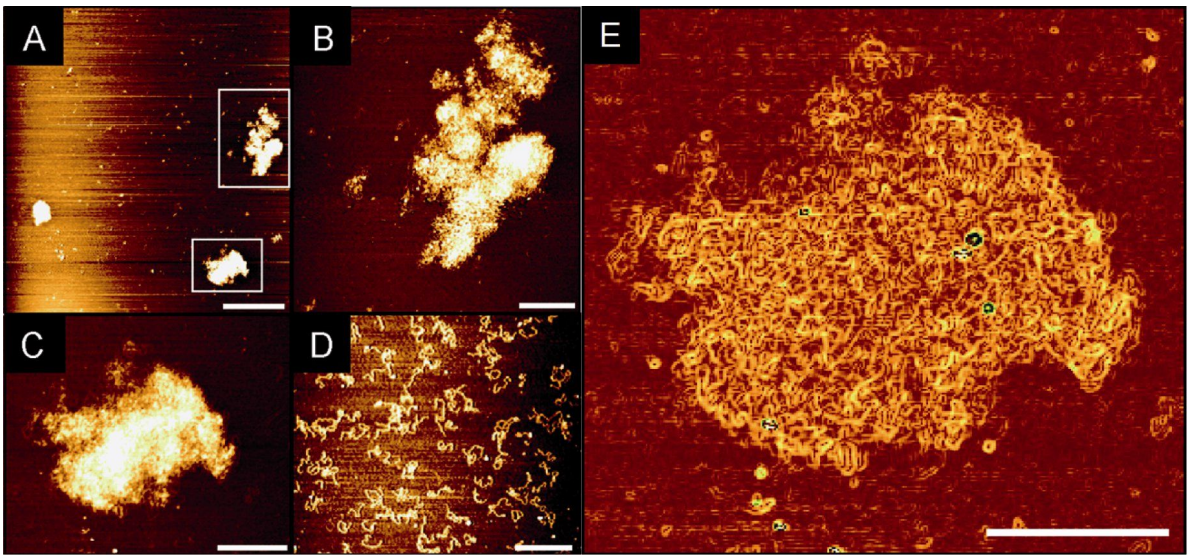



**Figure 4:** AFM analysis of DNA structures: (A) Overview-scan 10 × 10 µm showing DNA networks generated after 28 cycles with 66 °C annealing temperature (bar is 2 µm). (B, C) Zoom in picture of DNA network (all bars are 0.5 µm, respectively). (D) AFM when non-branched primers were employed (bar is 0.5 µm). (E) Software analysis of AFM measurement taking first derivative of height channel. DNA network generated after 28 cycles with 66 °C annealing temperature (bar is 0.5 µm).

AFM picture showed a better contrast for the visualization of DNA networks. Proceeding in this way, the DNA strands were better resolved and gave an impression of the shape of DNA networks. The observed DNA networks showed a less ordered shape which can be related to DNA flexibility. This flexibility was also observed with linear DNA (Figure 4D).

In order to review whether DNA networks form in solution the network generated by PCR was investigated by dynamic light scattering (DLS). The obtained data indicate that DNA networks are also present in solution. The measured average hydrodynamic diameter ($D_H$) at 90° after 10 measurements were partially 67 nm and 593 nm while the linear DNA showed a $D_H$ of 11 nm (see Supporting Information File 1). This $D_H$ is variable at different angles because DNA networks were not expected to have spherical morphology in solution [47]. Furthermore, different $D_H$ values in the case of branched DNA samples were owed to the dynamics in solution. Comparing scattering intensity of linear DNA and branched DNA networks, gave the possibility to conclude that DNA networks existed in solution in variable shapes.

Next we employed electron paramagnetic resonance (EPR) spectroscopy for investigation of the DNA networks. EPR is a widespread technique for the studies of structural and dynamic properties of biological macromolecules, e.g., DNA [48-53]. Since the systems studied in the current investigation are diamagnetic, nitroxide labels had to be inserted enzymatically [54,55]. EPR spectra of nitroxide labels are sensitive to dynamics on the picosecond to microsecond timescales and these dynamics are greatly altered upon attachment to a macromolecule. The nitroxide labels report not only the dynamics of the macromolecules as a whole, but additionally, they undergo rotations around the chemical bonds of the linker, and furthermore, the site of attachment can undergo conformational changes compared to the rest of the macromolecule. As the EPR-signal arises from these three processes, the interpretation of EPR spectra is rendered difficult. Measurements at a single frequency do not allow complete description of the spin label motion. However, a spectrum often can be approximated by a simple motional model to provide information on the properties of the macromolecule [56].

Since DNA polymerases are known to tolerate several dNTP modifications [57,58], we next investigated whether dTTP can be partially replaced by the nitroxide labeled dT*TP to generate DNA networks containing spin labels covalently bound to DNA.

All EPR spectra were measured in X-band (9.5 GHz) at room temperature in solution and are shown as first derivatives. For later comparison, spin-labeled TTP analogue dT*TP (Figure 5D) was measured and results to EPR spectra with three narrow lines indicating a high rotational mobility averaging the anisotropy of the hyperfine interaction to the $^{14}N$ nucleus (Figure 5A) [54].

Using the same approach as described above, we generated a linear DNA construct as well as a DNA network using a 1062 bp long template in the presence of a 1:1 ratio of dT*TP to TTP and natural dNTPs. We obtained spectra showing that the spin-labeled nucleotide was indeed incorporated into the DNA (Figure 5B,C). All spectra were quantitatively analyzed by spectral line shape simulations. Thereby, the dynamics of the nitroxide spin-labels are reflected in rotational correlation times $\tau_c$ assuming isotropic rotation of the label. The $^{13}C$-satellites (Figure 5A) were not taken into account for simulations. While spectra of Figure 5A and B are satisfactorily described by this approach, two components featuring two different rotational correlation times $\tau_c$ were required in case of Figure 5C. The two component fit was more consistent than assuming anisotropic rotational diffusion or a log-Gaussian distribution of correlation times [59]. The main component of the spectrum features a drastically reduced rotational mobility of the spin-label. The rotational correlation times as well as the fraction of both spectral components for spin labeled DNA networks as derived by spectral simulations are summarized in Table S1 (see Supporting Information File 1). The drastic decrease in mobility of the probe shown in the EPR spectrum upon enzyme-catalyzed DNA network growth (Figure 5C) clearly indicated network formation in solution and suggests rather rigid DNA networks. A more detailed analysis of the spectrum results in two components; the first component (contributing 8%) features a rotational correlation time ($\tau_c = 0.17 \pm 0.02$ ns) almost identical to $\tau_c$ of spin labels in linear DNA ($\tau_c = 0.15 \pm 0.02$ ns). The second component (92%) gives rise to a correlation time of $\tau_c = 11.2 \pm 0.3$ ns and is allocated to spin-labels incorporated into the DNA network. Since the presence of linear, non-branched DNA in the sample was excluded by gel filtration, we concluded, that the first component originated from “dangling DNA” at the edges of DNA networks. In a simple model, two-dimensional, approximately circularly shaped networks consisting of some 600 hexagons would contain about 8% dangling DNA. The diameter of such a structure could be estimated to several micrometers and is thus consistent with the AFM results. Concentrations of spin labels were obtained from the simulations fitted to the experimental data and used to derive the degree of labeling. We found that 1% of incorporated thymidines were replaced by the spin labeled analogue in PCR when linear strands as well as networks were formed. Thus, one dsDNA strand contains approximately 4 incorporated spin labels.

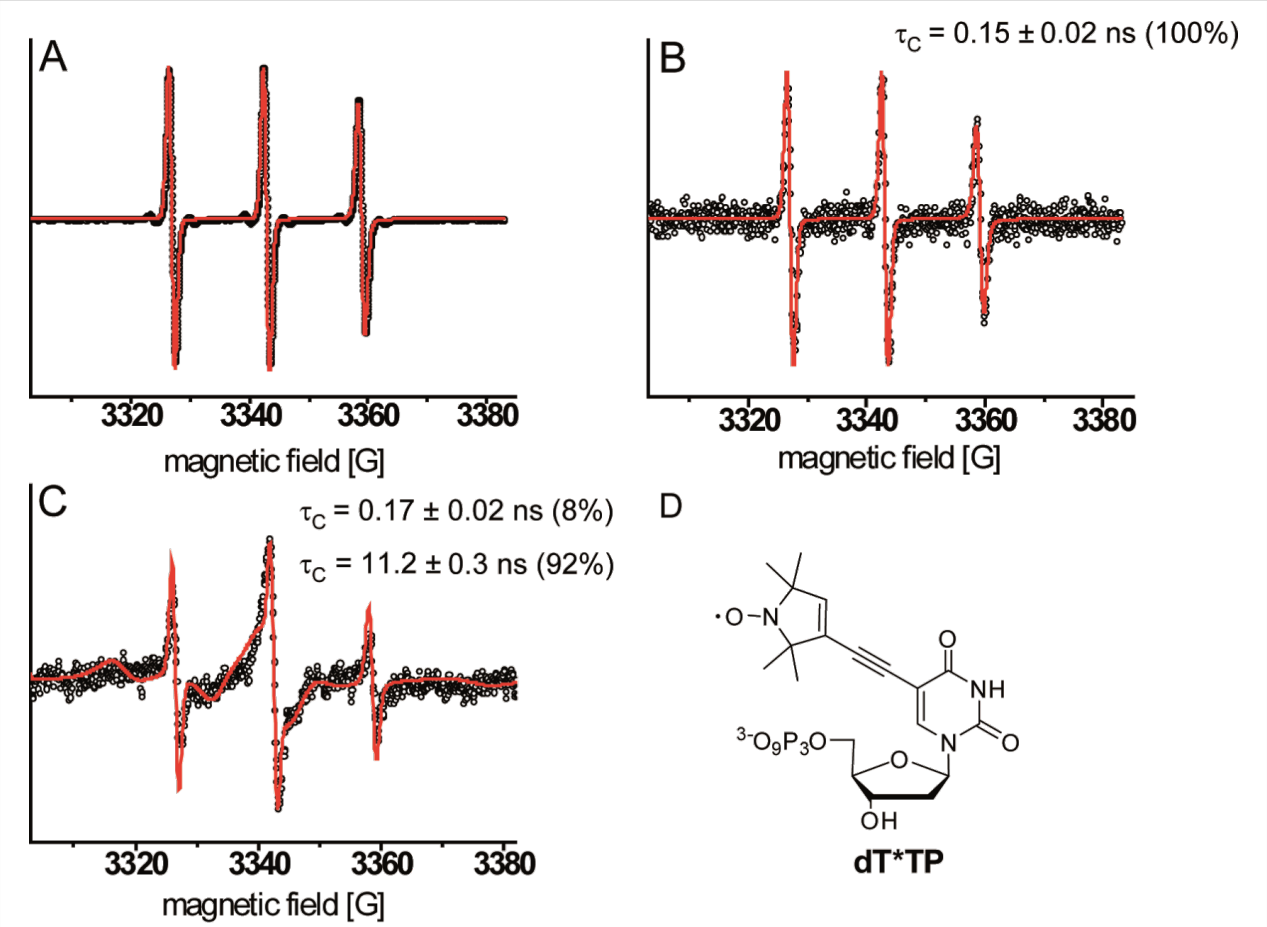



**Figure 5:** EPR spectra with corresponding spectral simulations (red line) of (A) dT*TP, (B) PCR reaction product resulting from non-branched primers in the presence of 50% dT*TP; The simulated spectrum and the corresponding $\tau_C$ are derived from a one component fit, (C) DNA networks generated by PCR in the presence of 50% dTTP* using branched primer; the simulated spectrum and corresponding $\tau_C$ are derived from a two component fit, (D) structure of modified dT*TP employed in PCR.

## Conclusion

To summarize, branched DNA primers were applied for the generation of DNA networks via enzymatic elongation of DNA by PCR. A straightforward synthesis was developed resulting in a rigid DNA branching building block **6** that was successfully used in solid phase DNA synthesis. In thermal denaturation studies Y-shaped bDNA indicates comparable behavior to the linear non-branched counterpart, thus forming a B-DNA conformation. Furthermore, the branched DNA self assembles into stable networks when short self-complimentary DNA sequences are used. We found that despite its rigidity, *Taq* DNA polymerase accepts the branched DNA construct as primers and builds DNA networks that grow cycle by cycle in PCR. Further, the generated DNA networks can be visualized using AFM. Generated images present surface covering structures in different shapes which are further characterized in solution by DLS. EPR measurements further corroborate network formation and suggest rather rigid DNA networks. As demonstrated, the approach allows to additionally modify the networks by using chemically modified nucleotides during PCR. The depicted approach might find applications like in the generation of functional hydrogels or tissue engineering.

## Supporting Information

Experimental procedures and full characterization data for all new compounds, materials and general procedures are given. Oligonucleotide synthesis and characterization of oligonucleotides, original melting curves of oligonucleotides and CD spectra of double stranded Y-motif are also shown. Further AFM studies, DLS spectra and EPR related experiments.

### Supporting Information File 1

Experimental part.
[http://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-10-104-S1.pdf]

## Acknowledgements

We acknowledge funding by the Konstanz Research School Chemical Biology and within the programme Molekulare Bionik of the Ministerium für Wissenschaft, Forschung und Kunst, Baden-Württemberg. This project was supported by the DFG (DR 743/2-1), we gratefully acknowledge Sebastian Höfel´s contribution to EPR data analysis.

## References

1. Seeman, N. C. *J. Biomol. Struct. Dyn.* **1990,** *8,* 573–581. doi:10.1080/07391102.1990.10507829
2. Wei, B.; Wang, Z.; Mi, Y. *J. Comput. Theor. Nanosci.* **2007,** *4,* 133–141.
3. Zuker, M. *Nucleic Acids Res.* **2003,** *31,* 3406–3415. doi:10.1093/nar/gkg595
4. Seeman, N. C. *J. Theor. Biol.* **1982,** *99,* 237–247. doi:10.1016/0022-5193(82)90002-9
5. Seeman, N. C. *Nature* **2003,** *421,* 427–431. doi:10.1038/nature01406
6. Deng, Z.; Lee, S.-H.; Mao, C. *J. Nanosci. Nanotechnol.* **2005,** *5,* 1954–1963. doi:10.1166/jnn.2005.504
7. Feldkamp, U.; Niemeyer, C. M. *Angew. Chem., Int. Ed.* **2006,** *45,* 1856–1876. doi:10.1002/anie.200502358
8. Lin, C.; Liu, Y.; Rinker, S.; Yan, H. *ChemPhysChem* **2006,** *7,* 1641–1647. doi:10.1002/cphc.200600260
9. Aldaye, F. A.; Palmer, A. L.; Sleiman, H. F. *Science* **2008,** *321,* 1795–1799. doi:10.1126/science.1154533
10. Liu, D.; Wang, M.; Deng, Z.; Walulu, R.; Mao, C. *J. Am. Chem. Soc.* **2004,** *126,* 2324–2325. doi:10.1021/ja031754r
11. LaBean, T. H.; Yan, H.; Kopatsch, J.; Liu, F.; Winfree, E.; Reif, J. H.; Seeman, N. C. *J. Am. Chem. Soc.* **2000,** *122,* 1848–1860. doi:10.1021/ja993393e
12. Mao, C.; Sun, W.; Seeman, N. C. *J. Am. Chem. Soc.* **1999,** *121,* 5437–5443. doi:10.1021/ja9900398
13. Winfree, E.; Liu, F.; Wenzler, L. A.; Seeman, N. C. *Nature* **1998,** *394,* 539–544. doi:10.1038/28998
14. Fu, T. J.; Seeman, N. C. *Biochemistry* **1993,** *32,* 3211–3220. doi:10.1021/bi00064a003
15. Rothemund, P. W. K. *Nature* **2006,** *440,* 297–302. doi:10.1038/nature04586
16. Chen, J.; Seeman, N. C. *Nature* **1991,** *350,* 631–633. doi:10.1038/350631a0
17. Shih, W. M.; Quispe, J. D.; Joyce, G. F. *Nature* **2004,** *427,* 618–621. doi:10.1038/nature02307
18. He, Y.; Ye, T.; Su, M.; Zhang, C.; Ribbe, A. E.; Jiang, W.; Mao, C. *Nature* **2008,** *452,* 198–201. doi:10.1038/nature06597
19. Liedl, T.; Högberg, B.; Tytell, J.; Ingber, D. E.; Shih, W. M. *Nat. Nanotechnol.* **2010,** *5,* 520–524. doi:10.1038/nnano.2010.107
20. Langecker, M.; Arnaut, V.; Martin, T. G.; List, J.; Renner, S.; Mayer, M.; Dietz, H.; Simmel, F. C. *Science* **2012,** *338,* 932–936. doi:10.1126/science.1225624
21. Gartner, Z. J.; Bertozzi, C. R. *Proc. Natl. Acad. Sci. U. S. A.* **2009,** *106,* 4606–4610. doi:10.1073/pnas.0900717106
22. Selden, N. S.; Todhunter, M. E.; Jee, N. Y.; Liu, J. S.; Broaders, K. E.; Gartner, Z. J. *J. Am. Chem. Soc.* **2012,** *134,* 765–768. doi:10.1021/ja2080949
23. Aldaye, F. A.; Senapedis, W. T.; Silver, P. A.; Way, J. C. *J. Am. Chem. Soc.* **2010,** *132,* 14727–14729. doi:10.1021/ja105431h
24. Vermesh, U.; Vermesh, O.; Wang, J.; Kwong, G. A.; Ma, C.; Hwang, K.; Heath, J. R. *Angew. Chem., Int. Ed.* **2011,** *50,* 7378–7380. doi:10.1002/anie.201102249
25. Goodman, R. P.; Schaap, I. A. T.; Tardin, C. F.; Erben, C. M.; Berry, R. M.; Schmidt, C. F.; Turberfield, A. J. *Science* **2005,** *310,* 1661–1665. doi:10.1126/science.1120367
26. Zhang, Y.; Seeman, N. C. *J. Am. Chem. Soc.* **1994,** *116,* 1661–1669. doi:10.1021/ja00084a006
27. Li, Y.; Cu, Y. T. H.; Luo, D. *Nat. Biotechnol.* **2005,** *23,* 885–889. doi:10.1038/nbt1106
28. Li, Y.; Tseng, Y. D.; Kwon, S. Y.; d'Espaux, L.; Bunch, J. S.; McEuen, P. L.; Luo, D. *Nat. Mater.* **2004,** *3,* 38–42. doi:10.1038/nmat1045
29. Um, S. H.; Lee, J. B.; Park, N.; Kwon, S. Y.; Umbach, C. C.; Luo, D. *Nat. Mater.* **2006,** *5,* 797–801. doi:10.1038/nmat1741
30. Park, N.; Um, S. H.; Funabashi, H.; Xu, J.; Luo, D. *Nat. Mater.* **2009,** *8,* 432–437. doi:10.1038/nmat2419
31. Wilner, O. I.; Weizmann, Y.; Gill, R.; Lioubashevski, O.; Freeman, R.; Willner, I. *Nat. Nanotechnol.* **2009,** *4,* 249–254. doi:10.1038/nnano.2009.50
32. Lin, C.; Xie, M.; Chen, J. J. L.; Liu, Y.; Yan, H. *Angew. Chem., Int. Ed.* **2006,** *45,* 7537–7539. doi:10.1002/anie.200602113
33. Lin, C.; Wang, X.; Liu, Y.; Seeman, N. C.; Yan, H. *J. Am. Chem. Soc.* **2007,** *129,* 14475–14481. doi:10.1021/ja0760980
34. Lin, C.; Rinker, S.; Wang, X.; Liu, Y.; Seeman, N. C.; Yan, H. *Proc. Natl. Acad. Sci. U. S. A.* **2008,** *105,* 17626–17631. doi:10.1073/pnas.0805416105
35. Paukstelis, P. J.; Ellington, A. D. *Proc. Natl. Acad. Sci. U. S. A.* **2008,** *105,* 17593–17594. doi:10.1073/pnas.0810029105
36. Meng, M.; Ahlborn, C.; Bauer, M.; Plietzsch, O.; Soomro, S. A.; Singh, A.; Muller, T.; Wenzel, W.; Bräse, S.; Richert, C. *ChemBioChem* **2009,** *10,* 1335–1339. doi:10.1002/cbic.200900162
37. Singh, A.; Tolev, M.; Meng, M.; Klenin, K.; Plietzsch, O.; Schilling, C. I.; Muller, T.; Nieger, M.; Bräse, S.; Wenzel, W.; Richert, C. *Angew. Chem., Int. Ed.* **2011,** *50,* 3227–3231. doi:10.1002/anie.201006992
38. Keller, S.; Wang, J.; Chandra, M.; Berger, R.; Marx, A. *J. Am. Chem. Soc.* **2008,** *130,* 13188–13189. doi:10.1021/ja8045348
39. Chandra, M.; Keller, S.; Gloeckner, C.; Bornemann, B.; Marx, A. *Chem.–Eur. J.* **2007,** *13,* 3558–3564. doi:10.1002/chem.200601473
40. Keller, S.; Marx, A. *Chem. Soc. Rev.* **2011,** *40,* 5690–5697. doi:10.1039/c1cs15040e
41. Sharma, R. A.; Bobek, M. *J. Org. Chem.* **1978,** *43,* 367–369. doi:10.1021/jo00396a051
42. Nuzzi, A.; Massi, A.; Dondoni, A. *QSAR Comb. Sci.* **2007,** *26,* 1191–1199. doi:10.1002/qsar.200740079
43. Ding, K.; Alemdaroglu, F. E.; Börsch, M.; Berger, R.; Herrmann, A. *Angew. Chem., Int. Ed.* **2007,** *46,* 1172–1175. doi:10.1002/anie.200603064
44. Hansma, H. G. *Annu. Rev. Phys. Chem.* **2001,** *52,* 71–92. doi:10.1146/annurev.physchem.52.1.71
45. Magonov, S. N.; Elings, V.; Whangbo, M.-H. *Surf. Sci.* **1997,** *375,* L385–L391. doi:10.1016/S0039-6028(96)01591-9
46. Bustamante, C.; Keller, D.; Yang, G. *Curr. Opin. Struct. Biol.* **1993,** *3,* 363–372. doi:10.1016/S0959-440X(05)80107-1
47. Carneiro, K. M. M.; Aldaye, F. A.; Sleiman, H. F. *J. Am. Chem. Soc.* **2010,** *132,* 679–685. doi:10.1021/ja907735m
48. Barhate, N.; Cekan, P.; Massey, A. P.; Sigurdsson, S. T. *Angew. Chem., Int. Ed.* **2007,** *46,* 2655–2658. doi:10.1002/anie.200603993

49. Qin, P. Z.; Haworth, I. S.; Cai, Q.; Kusnetzow, A. K.; Grant, G. P. G.; Price, E. A.; Sowa, G. Z.; Popova, A.; Herreros, B.; He, H. *Nat. Protoc.* **2007,** *2,* 2354–2365. doi:10.1038/nprot.2007.308
50. Schiemann, O.; Piton, N.; Mu, Y.; Stock, G.; Engels, J. W.; Prisner, T. F. *J. Am. Chem. Soc.* **2004,** *126,* 5722–5729. doi:10.1021/ja0393877
51. Schiemann, O.; Piton, N.; Plackmeyer, J.; Bode, B. E.; Prisner, T. F.; Engels, J. W. *Nat. Protoc.* **2007,** *2,* 904–923. doi:10.1038/nprot.2007.97
52. Ward, R.; Keeble, D. J.; El-Mkami, H.; Norman, D. G. *ChemBioChem* **2007,** *8,* 1957–1964. doi:10.1002/cbic.200700245
53. Singh, V.; Azarkh, M.; Exner, T. E.; Hartig, J. S.; Drescher, M. *Angew. Chem., Int. Ed.* **2009,** *48,* 9728–9730. doi:10.1002/anie.200902146
54. Obeid, S.; Yulikov, M.; Jeschke, G.; Marx, A. *Angew. Chem., Int. Ed.* **2008,** *47,* 6782–6785. doi:10.1002/anie.200802314
55. Obeid, S.; Baccaro, A.; Welte, W.; Diedrichs, K.; Marx, A. *Proc. Natl. Acad. Sci. U. S. A.* **2010,** *107,* 21327–21331. doi:10.1073/pnas.1013804107
56. Brustolon, M.; Giamello, E., Eds. *Electron Paramagnetic Resonance;* John Wiley & Sons: Hoboken, 2009. doi:10.1002/9780470432235
57. Weisbrod, S.; Marx, A. *Chem. Commun.* **2008,** 5675–5685. doi:10.1039/b809528k
58. Hocek, M.; Fojta, M. *Org. Biomol. Chem.* **2008,** *6,* 2233–2241. doi:10.1039/b803664k
59. Faetti, M.; Giordano, M.; Leporini, D.; Pardi, L. *Macromolecules* **1999,** *32,* 1876–1882. doi:10.1021/ma981178x

## License and Terms

Beilstein Journal of Organic Chemistry

# Pyrene-modified PNAs: Stacking interactions and selective excimer emission in PNA$_2$DNA triplexes

Alex Manicardi*[1], Lucia Guidi[1], Alice Ghidini[1,2] and Roberto Corradini*[1]

Full Research Paper

Open Access

Address:
[1]Department of Chemistry, University of Parma, Parco Area delle Scienze 17/A, 43124, Parma, Italy. Fax: +39 0521 905472; Tel: +39 0521 905410 and [2]Present Address: Department of Biosciences and Nutrition, Karolinska Institutet, Novum, Hälsovägen 7, 14183, Huddinge, Sweden

Email:
Alex Manicardi* - alex.manicardi@unipr.it; Roberto Corradini* - roberto.corradini@unipr.it

* Corresponding author

Keywords:
modified nucleobase; nucleic acids; PNA; pyrene excimer; SNP recognition; triplex stabilization

*Beilstein J. Org. Chem.* **2014,** *10,* 1495–1503.
doi:10.3762/bjoc.10.154

Received: 11 March 2014
Accepted: 03 June 2014
Published: 02 July 2014

This article is part of the Thematic Series "Nucleic acid chemistry".

Guest Editor: H.-A. Wagenknecht



## Abstract

Pyrene derivatives can be incorporated into nucleic acid analogs in order to obtain switchable probes or supramolecular architectures. In this paper, peptide nucleic acids (PNAs) containing 1 to 3 1-pyreneacetic acid units (**PNA1**–**6**) with a sequence with prevalence of pyrimidine bases, complementary to cystic fibrosis W1282X point mutation were synthesized. These compounds showed sequence-selective switch-on of pyrene excimer emission in the presence of target DNA, due to PNA$_2$DNA triplex formation, with stability depending on the number and positioning of the pyrene units along the chain. An increase in triplex stability and a very high mismatch-selectivity, derived from combined stacking and base-pairing interactions, were found for **PNA2**, bearing two distant pyrene units.

## Introduction

Peptide nucleic acid (PNA) probes are very selective in the recognition of DNA and have been used in a large variety of diagnostic methods, easily allowing the detection of point mutations at very low concentrations [1-3]. Poly-pyrimidine PNA can form very stable triplexes of the type PNA/DNA/PNA with poly-purine DNA, via both Watson–Crick and Hoogsteen base pairing (Figure 1). These structures are so stable that dsDNA undergoes displacement of the non-complementary strand [4-7]. However, the formation of triplex structures is limited to homopyrimidine sequences since the presence of one or more purine residues destabilizes these complexes and favour the formation of less stable duplexes [8]. Therefore it would be of great value to adopt strategies for the stabilization of triplex structures even in the presence of non-pyrimidine bases. From the available structural data on these complexes [7], it is possible to envisage that any pair of groups protruding from both thymines methyl groups of a TAT triplet and able to give rise to attractive interactions (Figure 1a) would stabilize the

triplex. If these groups are aromatic fluorophores, changes in the fluorescence properties can be observed upon interaction with DNA, thus enabling to study the occurring interactions and to produce switching PNA probes.

Fluorescent switching probes for DNA detection are very useful tools in diagnostics applications such as real-time PCR and in situ hybridisation [9,10]. Among the possible reporter groups, pyrene has been proposed in several in vitro detection systems, due to the sensitivity of its fluorescence properties to microenvironment and due to its ability to produce stabilizing stacking interactions and to show excimer fluorescence [11-20]. Furthermore, pyrene has been shown to favour self-assembly processes of supramolecular structures [21-28] and interact with carbon nanostructures such as nanotubes [29] or graphene [30], thus allowing to create composite material with special properties. PNA fluorescent probes bearing pyrene units as "universal base" were described [31,32], and recently, pyrrolidinyl-PNA bearing a uracil-linker pyrene unit showed good fluorescence response and mismatch recognition [33]; though terminal pyrene units were shown to stabilize triplexes formed by oligonucleotide probes [34], the effect of single- or multiple pyrene units on PNA in the formation of triplex structures has still to be addressed.

We have recently reported the modification of uracil at C5 by hydroxymethylation, followed by substitution with chloride and then with azide, which can be used for click chemistry or as a masked amino group both in a PNA monomer and in PNA oligomers, allowing to produce a variety of modified PNAs from a single precursor [35]. This chemistry introduces a moderate degree of flexibility which can be useful for allowing interactions with other groups to occur within the major groove.

In this work we applied this strategy to the synthesis of new mono-, di- and tri-functionalised PNA containing a 1-pyreneacetic acid residue linked to this C5-aminomethyl group (Figure 1b). As a model sequence, we chose a 9-mer (Figure 1c) complementary to a purine-rich tract of DNA which is present in the mutated form of the human cystic fibrosis (CFTR) gene, and which was previously studied in our lab using PNA and modified PNA probes [36,37]; this mutation is characterised by the presence of an adenine instead of guanine, and corresponds to one of the most frequent point mutations connected with cystic fibrosis (M-W1282X).

## Results and Discussion

### Synthesis of the PNA strands

Two different approaches were followed for the introduction of the pyrene units in the PNA strands. The probe containing only 1 pyrene unit (**PNA1**, Figure 1c) was synthesized by on-resin modification of 5-azidomethyluracil precursor, as described previously [35], whereas a pyrene-containing modified monomer **1** (Scheme 1), more suitable for automated synthesis, was

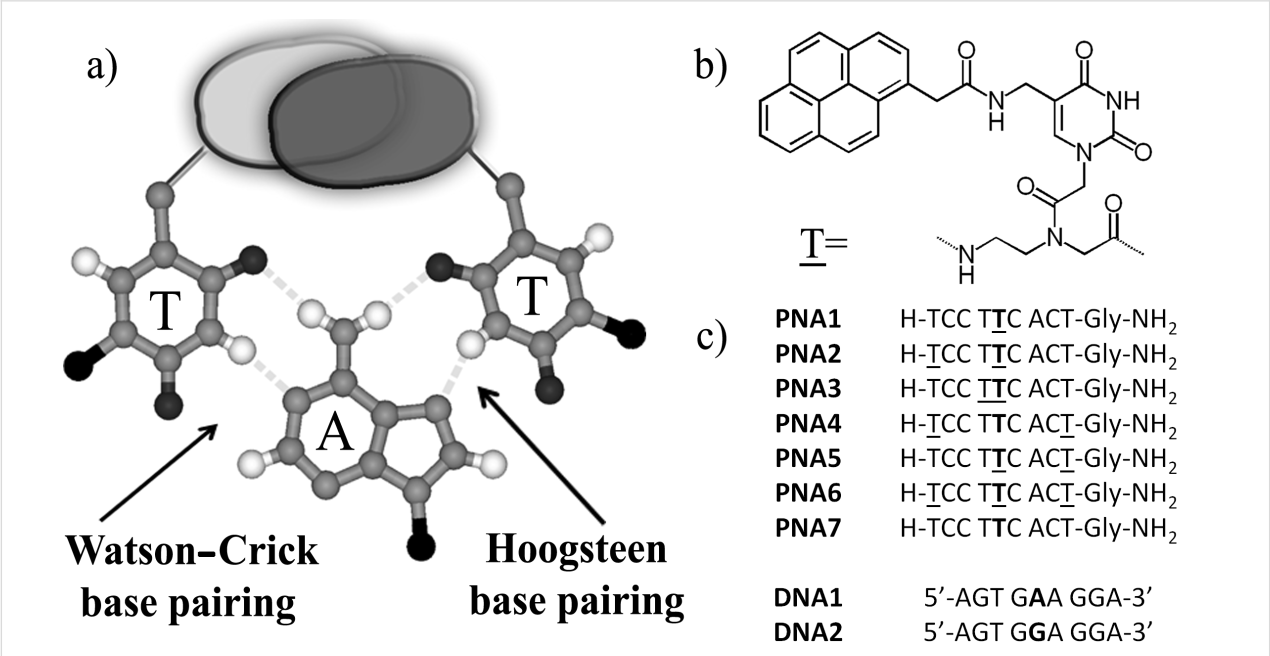



**Figure 1:** (a) TAT triplet structure showing Watson–Crick and Hoogsteen base pairing; the binding can be reinforced by the concurrent interaction between two groups protruding from C5 position of thymine into the major groove; (b) pyrene-modified uracil derivative used in PNA monomer in the present study; (c) sequences of PNA and DNA used. T indicates pyrene modified nucleobases; bold letters indicate the position of W1282X point mutation.

designed for the realization of all the other oligomers (**PNA2–6,** Figure 1c).

For the synthesis of the modified monomer bearing the pyrene moiety, we started from the 5-azidomethyluracil building block **2** previously synthesized by our lab [35]. The azide function was first reduced under Staudinger conditions to the corresponding amine **3**, on which 1-pyreneacetic acid was linked using HBTU/DIPEA as condensing agent. The ester **4** was then hydrolyzed to the acid **5**, and linked to the Fmoc-protected PNA backbone using EDC/DhBtOH as activating mixture; the PNA monomer **1**, was then obtained by ester hydrolysis of **6** under acidic conditions.

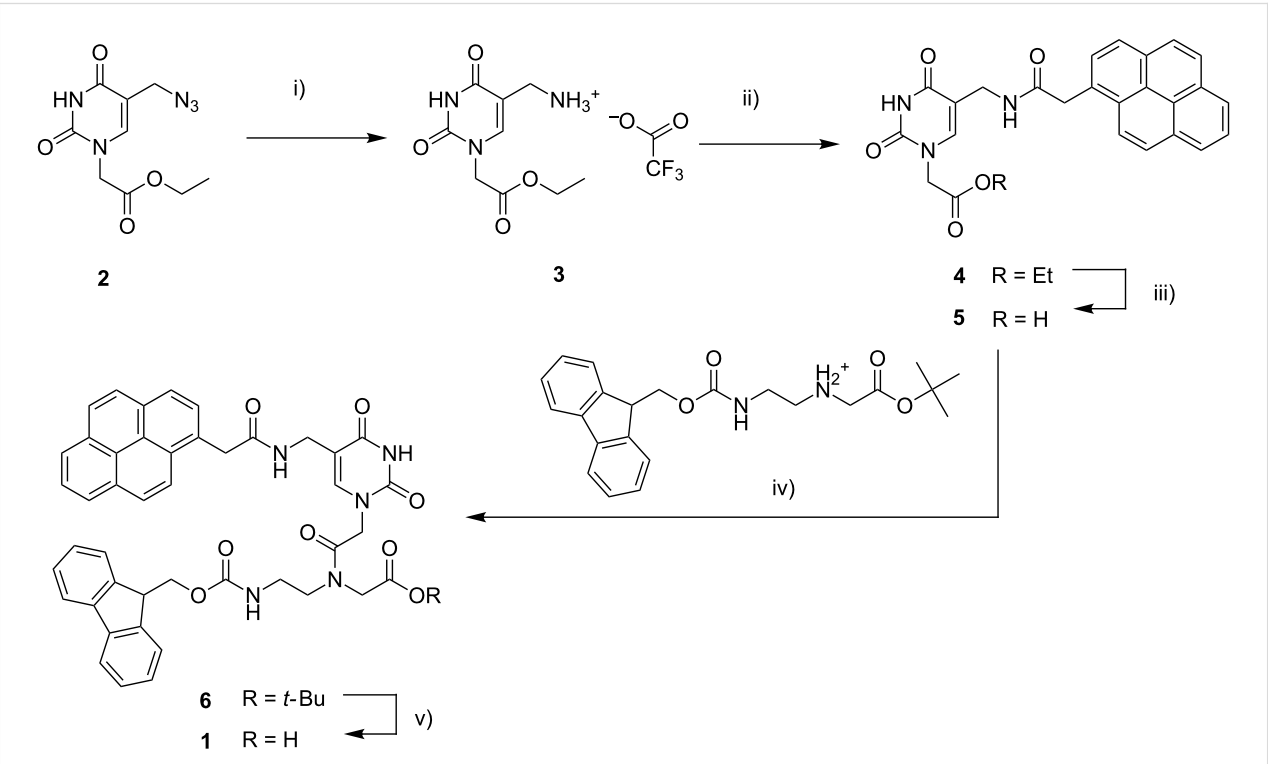
The PNAs sequence was designed to be complementary to the W1282X mutated form of CFTR gene, and all PNAs were synthesized using standard Fmoc-based manual solid phase protocol. The crude products were purified by RP-HPLC and characterized by HPLC–UV–MS (Supporting Information File 1, Figures S6–S11).

## Thermal stability of PNA:DNA complexes

The introduction of a modification in a PNA stand can lead to different effects, electronic or steric, which affect both self-aggregation of the PNA and their interactions with complementary DNA strand. Substitution at the C-5 position of the uracil ring allows positioning of the substituent in the direction of the major groove of the double helix, thus reducing the destabilization induced by steric factors; moreover the large aromatic portion introduced with the pyrene ring can interact with the flanking bases of the strand through π–π stacking interactions, thus stabilizing the complex formed.

For the evaluation of the sum of all this effects we measured the melting temperatures of the complexes formed between the PNAs and the full matched **DNA1** or the single mismatched **DNA2** (corresponding to the wild type CFTR gene), using both UV (Table 1, and Supporting Information File 1, Figures S12–S14) and fluorescence (Supporting Information File 1, Figure S15). The stability of these complexes was indeed found to be strongly dependent on the presence, the positioning and the number of pyrene units within the PNA strand.

The presence of a single pyrene unit (**PNA1**) destabilizes the PNA:DNA complex. The introduction of a second pyrene unit adjacent to the first one (**PNA3**) results in a further destabilization, whereas distal positioning of pyrene units (**PNA2**, **4**, **5**) leads to stabilization if compared to **PNA1**, but to an extent depending on the position of the second pyrene unit. For **PNA2** the additional interactions lead to the highest stability and very




**Scheme 1:** Synthesis of the PNA monomer 1: i) 1. $PPh_3$, $H_2O$, THF; 2. TFA, 71%; ii) 1-pyreneacetic acid, HBTU, DIPEA, DMF, 66%; iii) NaOH, $H_2O$/MeOH (2:1), 91%; iv) EDC·HCl, DIPEA, DhBtOH, DMF, 68%; v) TFA, DCM, 86%.

**Table 1:** UV melting temperature of PNA:DNA complexes. All measurements were done in PBS at pH 7 with 1 μM strand concentration except for unmodified PNA measurements (5 μM strand concentration).

| PNA | $T_m$ PNA:**DNA1** (°C) | $T_m$ PNA:**DNA2** (°C) | $\Delta T_m$ (°C) |
|---|---|---|---|
| **PNA1** | 26 | 20 | 6 |
| **PNA2** | 39 | 19 | 20 |
| **PNA3** | 24 | < 18 | n.d. |
| **PNA4** | 33[a] | 22 | 11 |
| **PNA5** | 28 | 22 | 6 |
| **PNA6** | n.d.[b] | n.d.[b] | n.d. |
| **PNA7** | 34 | 24 | 10 |

[a]Broad transition observed. [b]Continuous drift, no net transition observed.

high selectivity, with $\Delta T_m$ strongly increased compared to the unmodified **PNA7** (20 °C vs 10 °C). The presence of a second pyrene unit at N-terminal position is more stabilizing than that at C-term (compare **PNA5** and **PNA2, 4**). **PNA4** is characterized by a broad melting curve, whereas for **PNA6** a continuous drift was observed already for the PNA alone, and in the presence of DNA no clear-cut transition was detected, suggesting a pyrene-mediated strong aggregation of the probe itself.

As described below, all the probes showed excimer emission in the 460–480 nm range upon hybridization (Figure 2). The temperature dependence of the excimer band in the presence of DNA was found to be in accordance with the UV melting measurements (Supporting Information File 1, Figure S15).

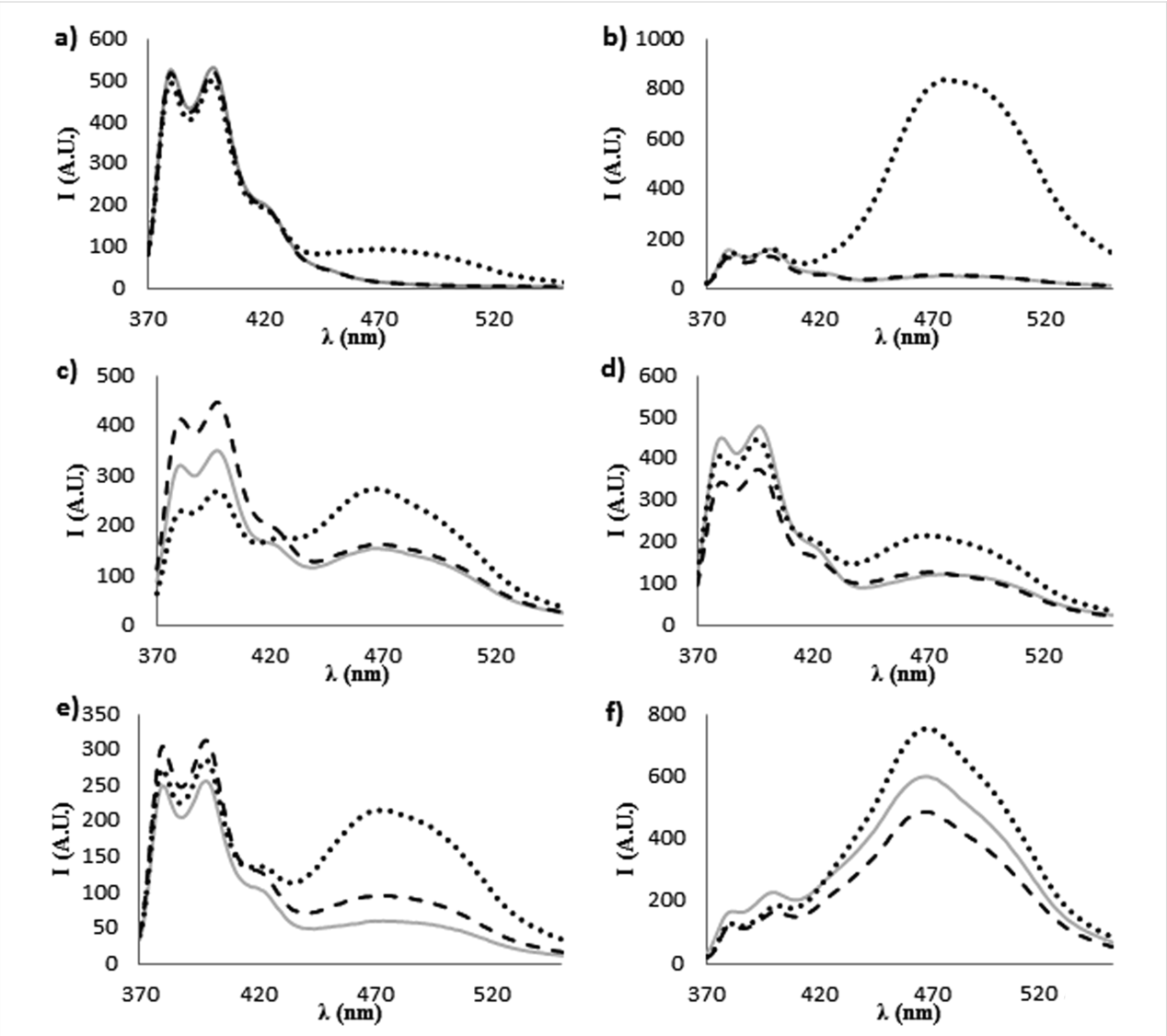



**Figure 2:** Fluorescence spectra at 347 nm excitation, recorded at 20 °C of: (a) **PNA1**, (b) **PNA2**, (c) **PNA3**, (d) **PNA4**, (e) **PNA5**, (f) **PNA6**. All measurements were done in PBS buffer, pH 7; concentration of each strand was 1 μM. Full lines are for ssPNA solutions, dotted lines are for PNA:**DNA1** solutions and broken lines are for PNA:**DNA2** solutions.

## Fluorescence studies

Beside the modification of thermal stability and selectivity induced by the incorporation of pyrene moieties described above, we evaluated the fluorescence properties of these PNA in the absence and in the presence of DNA. The evaluation of the pyrene quantum yields showed that these probes are much less fluorescent than the 1-pyreneacetic acid precursor in water (23 times lower quantum yield for **PNA2**, see Supporting Information File 1, Figure S16), probably due to the quenching effect of nucleobase units; however, the most important data are related to changes in the fluorescence spectrum upon hybridization with DNA, since this property is strongly related to the environment around the fluorophores [38] and can reveal interactions between pyrene units in the PNA:DNA complexes.

In Figure 2 the fluorescence emission spectra of the PNA probes in the absence or in the presence of complementary **DNA1** or mismatched **DNA2** are reported.

For **PNA1**, having only one pyrene unit, a typical pyrene excimer band was observed in the presence of **DNA1** (Figure 2a); this band cannot evidently derive from an intramolecular excimer and thus it must be due to a DNA-templated association of two PNA units. Furthermore, the same band was not observed in the presence of **DNA2**, thus indicating that the excimer formation is sequence-specific. **PNA2**, which has two distant pyrene units, showed a weak excimer emission, due to weak self-association (this band tend to disappear with dilution, see Supporting Information File 1, Figure S16), which underwent a dramatic enhancement when **PNA2** was bound to **DNA1**, whereas it remained unchanged by addition of **DNA2** (Figure 2b and Figure 3). This resulted in a very high increase in the excimer to monomer emission ratio (Figure 4), which can be exploited for analytical purposes in the case of the biologically relevant **DNA1** (mutated form) and **DNA2** (wilt type). The fluorescent responses for the other two mismatched DNA (**DNA3**: 5'-AGTG**C**AGGA-3' and **DNA4**, 5'-AGTG**T**AGGA-3'), were also measured (Supporting Information File 1, Figure S18) and were shown to give rise to results comparable to that of **DNA2**. Accordingly, no melting transitions were observed for **PNA2** with **DNA3** and **DNA4** above 18 °C (data not

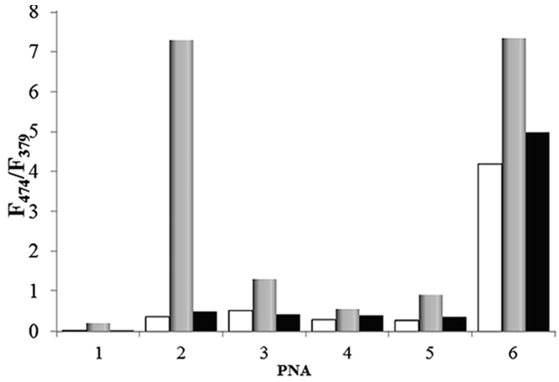



**Figure 4:** Ratio of the intensities of the pyrene excimer ($F_{474}$) and monomer emission ($F_{379}$) for the PNA probes in the absence of DNA (white bars), in the presence of **DNA1** (full match, grey bars), and **DNA2** (mismatch, black bars). Experiments were done at 20 °C in PBS at 1 μM PNA concentration (0.5 μM DNA concentration).

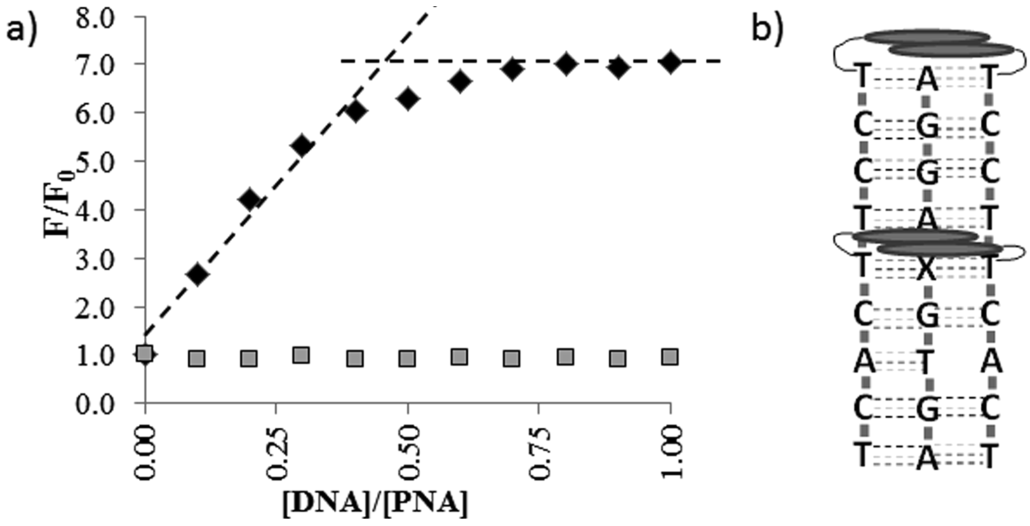



**Figure 3:** (a) Increase in fluorescence intensity of the excimer band for **PNA2** upon addition of complementary **DNA1** (black diamonds) or mismatched **DNA2** (grey squares) at 25 °C); (b) model of interaction showing both base recognition through hydrogen bonding and stacking interactions. X = A for full match, G for mismatch.

shown). Thus the intensity of the excimer band was found to follow the expected sequence selectivity of the hydrogen-bonding scheme.

**PNA3** has two proximal pyrene units on adjacent bases, and therefore the free probe already shows excimer emission; this band was enhanced in the presence of the templating **DNA1**, whereas in the presence of the mutated **DNA2** the excimer band remained as in single strand and only a slight enhancement in the monomer emission was observed; thus, the excimer to monomer emission ratio (Figure 4) was slightly reduced. **PNA4** and **PNA5** showed an increase in the excimer fluorescence intensity signal upon hybridization with the templating DNA, though lower than for **PNA2**. **PNA6** has already a strong excimer emission as single-strand, but this band was slightly enhanced upon interaction with the full-match **DNA1**, whereas it was slightly reduced in the presence of **DNA2**.

For all PNAs, very similar results were obtained in fluorescence response induced by **DNA3** and **DNA4**, except that for **PNA6** the difference observed with **DNA4** was less pronounced than with other mismatches (Supporting Information File 1, Figure S18).

The DNA-induced formation or enhancement of these excimer bands can be explained if a $PNA_2DNA$ triplex is formed, favoured by the prevalence of pyrimidines in the PNA [22,23]. The $PNA_2DNA$ triplex, in this sequence, is destabilized by the presence of a pyrimidine base (T) in the 5'-end of the DNA; thus this sequence represents a good model for evaluating the stabilization/destabilization effects due to the presence of pyrene units. The nature of these PNA:DNA complexes was confirmed by titration experiments; for **PNA7** CD titrations revealed a 2:1 stoichiometry (Supporting Information File 1, Figure S19); the same stoichiometry was found for **PNA2** by following the increase in the excimer emission as a function of DNA concentration (Figure 3a).

The hysteresis observed between melting and annealing curves (Supporting Information File 1, Table S1) is also indicative of the formation of triplex structures between the PNAs and the DNA.

## Evaluation of pyrene-modified **PNA2** as fluorescent probe

In Figure 4 the ratios between the excimer and monomer emission of each probe alone and in the presence of 0.5 equivalents of the full match **DNA1** and mismatched **DNA2** are reported. Under these conditions, **PNA1** showed an increased, though weak, excimer emission, whereas **PNA2** showed a dramatic increase in the presence of **DNA1** and very low one in the presence of **DNA2**. Thus, **PNA2** showed best performances in terms of excimer signal intensity $\Delta F_{FM}$ (difference between the fluorescence in the presence of full match **DNA1** and that of single strand PNA) and of selectivity compared to the wild-type mismatch (MM) reaching $\Delta F_{FM}/\Delta F_{MM} = 180$.

Using a 1 μM **PNA2** solution we calculated the limit of detection (LOD) of the fluorescence detection of **DNA1** using this probe. A linear regression was obtained in the low nanomolar range, and a LOD of 18.7 nM for **DNA1** was calculated (see Supporting Information File 1).

## Effect of pyrene units on stability and sensing

According to the occurrence of excimer bands in the fluorescence spectra of the PNA probes (Figure 2 and Figure 4) the presence of pyrene favours self-association of two PNA strands; strong interactions should be observed for **PNA2**–**5** and even stronger for **PNA6**, i.e., with the increase of the number of the pyrene units, as indeed experimentally observed. The following model can be used to rationalise the observed data for DNA interaction. The stacking interactions (Figure 1a) occurring between pyrene units of different PNA strands (schematically depicted in Figure 3b for **PNA2**) affect triplex formation, which is also biased by both steric and conformational effects; the base pairings of the adenines in the target DNA with the modified uracil units allow the two pyrene residues to be kept close enough to interact (generating an excimer band), but this process can result in destabilization of the overall structure (see $T_m$ of **PNA1** in Table 1). However, for **PNA2**, the combined effect of two pyrene pairs properly positioned allows to increase both stability and selectivity of PNA compared to unmodified one. The N-terminal pyrene unit, in addition to the central one (which has the same position as in **PNA1**), stabilizes the triplex structure through the occurrence of combined stacking interactions (Figure 3b). Thus, the presence of a single mismatch facing the central modified monomer results in destabilization not only of the excimer corresponding to this nucleobase, but of the entire triplex, leading to high mismatch recognition. This induces the very high selectivity in the switch-on of the excimer fluorescence emission (Figures 2, 3 and 4). All the other tested dispositions are not so effective in terms of stabilization, fluorescence response and selectivity; for **PNA3** this is attributable to steric hindrance between the adjacent pyrene units; **PNA4** and **PNA5** containing one pyrene unit in the C-term, at the end of a segment in which the triplex structure is destabilized by the presence on the PNA of one adenine unit, show less selectivity; furthermore a gradual transition was observed for **PNA4**, suggesting weak cooperativity in the stacking interaction. The presence of three pyrene residues (**PNA6**) instead, induce a strong self-aggregation of the PNA alone; this assembling process is favoured by the presence of **DNA1** and to a lesser

extent **DNA2** (Figure 2 and Figure 4a); however, the strong excimer signal of the PNA alone prevents its use as an efficient probe for DNA.

## Conclusion

In conclusion, we have demonstrated that introduction of two pyrene units protruding into the major groove and properly positioned along the PNA strand (as in **PNA2**) can stabilize $PNA_2DNA$ triplex structures by additional stacking interactions which combine with Watson–Crick and Hoogsteen base pairing; these interactions are clearly detectable by the formation of the excimer band of pyrene in the fluorescence spectra. Thus this work makes a significant step toward the objective of stabilizing triplexes even in the presence of pyrimidines on the target sequence, while still maintaining and even increasing sequence selectivity. Moreover, for diagnostics, it is important that a very high and sequence-selective excimer to monomer ratio can be obtained, as with **PNA2**, upon hybridization, a property which is very important also in the case of more elaborated methods such as gated detection. These characteristics make **PNA2** a very good fluorescent probe, with very high single-base selectivity in both thermal stability and excimer formation upon binding to target DNA. Thus, the present results can be very useful in the design of new probes for single point mutations and single nucleotide polymorphisms (SNPs), highly relevant in the genomic as well as in the clinical fields.

## Experimental

### General information

Reagents were purchased from Sigma-Aldrich, Fluka, Merck, Carlo Erba, TCI Europe, Link, ASM and used without further purification. All reactions were carried out under a nitrogen atmosphere with dry solvents under anhydrous conditions, unless otherwise noted. Anhydrous solvents were obtained by distillation or anhydrification with molecular sieves. Reactions were monitored by TLC carried out on 0.25 mm E. Merck silica-gel plates (60F-254) by using UV light as visualizing agent and ninhydrin solution and heat as developing agents. E. Merck silica gel (60, particle size 0.040–0.063 mm) was used for flash-column chromatography. NMR spectra were recorded on Bruker Avance 400 or 300 instruments and calibrated by using residual undeuterated solvent as an internal reference. The following abbreviations were used to explain the multiplicities: s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet and br = broad. IR spectra were measured using a FTIR Thermo Nicolet 5700, in transmission mode using KBr or NaCl. HPLC–UV–MS were recorded by using a Waters Alliance 2695 HPLC with Micromass Quattro microAPI spectrometer, a Waters 996 PDA and equipped with a Phenomenex Jupiter column (250 × 4.6 mm, 5 μm, C18, 300 Å) (method A, 5 minutes in $H_2O$ 0.2% formic acid (FA), then linear gradient to 50% MeCN 0.2% FA in 30 minutes at a flow rate of 1 mL/min). PNA oligomers were purified with RP-HPLC using a XTerra Prep $RP_{18}$ column (7.8 × 300 mm, 10 μm) (method B, linear gradient from $H_2O$ 0.1% TFA to 50% MeCN 0.1% TFA in 30 minutes at a flow rate of 4.0 mL/min). HRMS were recorded using a Thermo LTQ-Orbitrap XL.

**Synthesis and characterization of compounds 1, 3–6** are reported in Supporting Information File 1.

**Synthesis and characterization of PNAs.** The synthesis of **PNA1** was already described in a previous work [35]. The syntheses of all the other PNAs, bearing multiple pyrene units (**PNA2**, **PNA3**, **PNA4**, **PNA5** and **PNA6**), were performed with standard Fmoc-based manual synthesis protocol using **1** in addition to standard monomers, on a Rink amide resin loaded with Fmoc-Gly-OH as first monomer (0.2 mmol/g). The unmodified **PNA7** was synthesized using a standard Boc-based manual protocol using commercial monomers on a MBHA resin loaded with Fmoc-Gly-OH as first monomer (0.2 mmol/g). PNA purifications were performed by RP-HPLC with UV detection at 260 nm (gradient B). The purity and identity of the purified PNAs were determined by HPLC–UV–MS (gradient A). **PNA2**: 9%; $t_r$: 24.7 min; ESI-MS (*m*/*z*): calcd for [M]: 2932.1426; found: 1466.9 $[MH_2]^{2+}$, 978.2 $[MH_3]^{3+}$, 733.9 $[MH_4]^{4+}$, 587.2 $[MH_5]^{5+}$; **PNA3**: 10%; $t_r$: 24.6 min; ESI-MS: (*m*/*z*): calcd for [M]: 2932.1426; found: 1466.9 $[MH_2]^{2+}$, 978.2 $[MH_3]^{3+}$, 733.9 $[MH_4]^{4+}$, 587.3 $[MH_5]^{5+}$; **PNA4**: 11%; $t_R$: 23.7 min; ESI-MS (*m*/*z*): calcd for [M]: 2932.1426; found: 1466.8 $[MH_2]^{2+}$, 978.2 $[MH_3]^{3+}$, 733.9 $[MH_4]^{4+}$, 587.3 $[MH_5]^{5+}$; **PNA5**: 15%; $t_R$: 24.0 min; ESI-MS (*m*/*z*): calcd for [M]: 2932.1426; found: 1467.1 $[MH_2]^{2+}$, 978.2 $[MH_3]^{3+}$, 733.8 $[MH_4]^{4+}$, 587.3 $[MH_5]^{5+}$; **PNA6**: 11%; $t_R$: 27.4 min; ESI-MS (*m*/*z*): calcd for [M]: 3189.2266; found: 1064.1 $[MH_3]^{3+}$, 798.2 $[MH_4]^{4+}$, 638.7 $[MH_5]^{5+}$; **PNA7**: 25%; $t_R$: 18.5 min; ESI-MS (*m*/*z*): calcd for [M]: 2419.4159; found: 1210.4 $[MH_2]^{2+}$, 807.3 $[MH_3]^{3+}$, 605.7 $[MH_4]^{4+}$, 484.8 $[MH_5]^{5+}$. Yields reported in % for each PNA are those of purified products, calculated by UV–vis analysis.

**UV measurements**. Stock solutions of PNA and DNA synthetic oligonucleotides (Thermo-Fisher Scientific, HPLC-grade) were prepared in double-distilled water, and the PNA concentration was calculated by UV absorbance with the following extinction coefficients ($\varepsilon_{260}$ $[M^{-1}cm^{-1}]$) for the nucleobases: T 8600, T* 14938 (pyrene-modified monomer, see Supporting Information File 1, Figure S12 for the calculation of this value), C 6600, A 13700, and G 11700. For DNA the data provided by the producer were used. From these, solutions containing single stranded PNA and DNA or PNA:DNA duplexes were prepared. Measurement conditions: [PNA] = [DNA] = 1 μM in PBS (100 mM NaCl, 10 mM $NaH_2PO_4{\cdot}H_2O$, 0.1 mM EDTA, pH

7.0, 1.25% DMF). All the samples were first incubated at 90 °C for 5 min, then slowly cooled to room temperature. Thermal denaturation profiles ($A_{260}$ versus $T$) of the hybrids were measured with a UV–vis Lambda Bio 20 spectrophotometer equipped with a Peltier temperature programmer PTP6 interfaced to a personal computer. For the temperature range 18–50 °C, $A_{260}$ values were recorded at 0.1 °C increments, with a temperature ramp of 1 °C/min. Both melting and annealing curves were recorded for each solution. The melting temperature ($T_m$) was determined from the maximum of the first derivative of the melting curves.

**Fluorescence measurements.** Fluorescence spectra were recorded on a Perkin Elmer LS55 luminescence spectrometer equipped with a LAUDA ECOline RE104 temperature control system, exciting at 347 nm (slit: 5.0 nm), scanning from 370 nm to 550 nm, a scan speed of 200 nm/min was used and 3 accumulation for each spectrum. Samples were prepared as reported for UV measurements. From the stock solutions, described above, solutions of PNA alone (1 μM in PBS) and of PNA/DNA 2:1 (1 μM PNA and 0.5 μM DNA in PBS) were prepared. All the samples were first incubated at 90 °C for 5 min, and then slowly cooled to the temperature of analysis. Fluorescence emission spectra were recorded with an excitation wavelength of 347 nm (slit excitation: 5.0 nm), scanning from 370 nm to 550 nm (slit emission: 10.0 nm), a scan speed of 200 nm/min was used with 3 spectra accumulation for each solution. All measurements were compensated for lamp fluctuations by normalization using as reference a 20 nM 1-pyreneacetic acid solution in PBS. Equilibration of the solution and complete formation of the complexes were checked by repeating the analysis after 10 minutes, to ensure that no significant variation of the fluorescence profiles was present. Variable temperature fluorescence measurements are reported in Supporting Information File 1.

**Fluorescence titration of PNA2 and PNA3.** From the stock solution described above single stranded PNA solutions (1 μM in PBS) and single stranded DNA solutions (10 μM in PBS) were prepared. PNA solutions were first incubated at the experimental temperature, then spectra were recorded after addition of portions of DNA (10% of the PNA amount each), allowing an equilibration time of 8 min. Fluorescence emission spectra were recorded with an excitation wavelength of 347 nm (slit: 5.0 nm), scanning from 370 nm to 550 nm, with scan speed of 200 nm/min, and 3 spectra accumulation for each solution. All measurements were corrected for dilution, and compensated for lamp fluctuations by normalization using as reference a 20 nM 1-pyreneacetic acid solution in PBS. Fluorescence titration of **PNA3** at 10 °C is reported in Supporting Information File 1, Figure S17.

## Supporting Information

### Supporting Information File 1

Synthesis, characterization, and spectral data of compounds **1, 3–6**, HPLC–MS analyses of **PNA1–7**, additional UV, fluorescence and CD data.
[http://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-10-154-S1.pdf]

## Acknowledgements

This work was partially supported by MIUR (PRIN20093N774P)

## References

1. Nielsen, P. E. *Curr. Opin. Biotechnol.* **2001,** *12,* 16–20. doi:10.1016/S0958-1669(00)00170-1
2. Nielsen, P. E. *Peptide Nucleic Acids: Protocols and Applications;* Horizon Bioscience: Norwich, UK, 2004.
3. Bertucci, A.; Manicardi, A.; Corradini, R. Advanced Molecular Probes for Sequence-Specific DNA Recognition. In *Detection of non-amplified Genomic DNA;* Spoto, G.; Corradini, R., Eds.; Springer: Dordrecht, The Netherlands, 2012; pp 89–124. doi:10.1007/978-94-007-1226-3_4
4. Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. *Science* **1991,** *254,* 1497–1500. doi:10.1126/science.1962210
5. Wittung, P.; Nielsen, P. E.; Nordén, B. *J. Am. Chem. Soc.* **1996,** *118,* 7049–7054. doi:10.1021/ja960521f
6. Demidov, V. V.; Yavnilovich, M. V.; Belotserkovski, B. P.; Frank-Kamenetskii, M. D.; Nielsen, P. E. *Proc. Natl. Acad. Sci. U. S. A.* **1995,** *92,* 2637–2641. doi:10.1073/pnas.92.7.2637
7. Betts, L.; Josey, J. A.; Veal, J. M.; Jordan, S. R. *Science* **1995,** *270,* 1838–1841. doi:10.1126/science.270.5243.1838
8. Egholm, M.; Behrens, C.; Christensen, L.; Berg, R. H.; Nielsen, P. E.; Buchardt, O. *J. Chem. Soc., Chem. Commun.* **1993,** 800–801. doi:10.1039/c39930000800
9. Ranasinghe, R. T.; Brown, T. *Chem. Commun.* **2011,** *47,* 3717–3735. doi:10.1039/c0cc04215c
10. Huang, K.; Martí, A. A. *Anal. Bioanal. Chem.* **2012,** *402,* 3091–3102. doi:10.1007/s00216-011-5570-6
11. Østergaard, M. E.; Hrdlicka, P. J. *Chem. Soc. Rev.* **2011,** *40,* 5771–5788. doi:10.1039/c1cs15014f
12. Ebata, K.; Masuko, M.; Ohtani, H.; Kashiwasake-Jibu, M. *Photochem. Photobiol.* **1995,** *62,* 836–839. doi:10.1111/j.1751-1097.1995.tb09144.x
13. Paris, P. L.; Langenhan, J. M.; Kool, E. T. *Nucleic Acids Res.* **1998,** *26,* 3789–3793. doi:10.1093/nar/26.16.3789
14. Seo, Y. J.; Ryu, J. H.; Kim, B. H. *Org. Lett.* **2005,** *7,* 4931–4933. doi:10.1021/ol0518582
15. Seo, Y. J.; Lee, I. J.; Yi, J. W.; Kim, B. H. *Chem. Commun.* **2007,** 2817–2819. doi:10.1039/b707278c
16. Yamana, K.; Ohshita, Y.; Fukunaga, Y.; Nakamura, M.; Maruyama, A. *Bioorg. Med. Chem.* **2008,** *16,* 78–83. doi:10.1016/j.bmc.2007.04.053
17. Conlon, P.; Yang, C. J.; Wu, Y.; Chen, Y.; Martinez, K.; Kim, Y.; Stevens, N.; Marti, A. A.; Jockusch, S.; Turro, N. J.; Tan, W. *J. Am. Chem. Soc.* **2008,** *130,* 336–342. doi:10.1021/ja076411y
18. Sau, S. P.; Kumar, T. S.; Hrdlicka, P. J. *Org. Biomol. Chem.* **2010,** *8,* 2028–2036. doi:10.1039/b923465a

19. Häner, R.; Biner, S. M.; Langenegger, S. M.; Meng, T.; Malinovskii, V. L. *Angew. Chem., Int. Ed.* **2010,** *49,* 1227–1230. doi:10.1002/anie.200905829
20. Bains, G.; Patel, A. B.; Narayanaswami, V. *Molecules* **2011,** *16,* 7909–7935. doi:10.3390/molecules16097909
21. Nakamura, M.; Murakami, Y.; Sasa, K.; Hayashi, H.; Yamana, K. *J. Am. Chem. Soc.* **2008,** *130,* 6904–6905. doi:10.1021/ja801054t
22. Kashida, H.; Asanuma, H. *Phys. Chem. Chem. Phys.* **2012,** *14,* 7196–7204. doi:10.1039/c2cp40520b
23. Teo, Y. N.; Kool, E. T. *Chem. Rev.* **2012,** *112,* 4221–4245. doi:10.1021/cr100351g
24. Cho, Y.; Kool, E. T. *ChemBioChem* **2006,** *7,* 669–672. doi:10.1002/cbic.200500515
25. Varghese, R.; Wagenknecht, H.-A. *Chem. Commun.* **2009,** 2615–2624. doi:10.1039/b821728a
26. Mayer-Enthart, E.; Wagenknecht, H.-A. *Angew. Chem., Int. Ed.* **2006,** *45,* 3372–3375. doi:10.1002/anie.200504210
27. Malinovskii, V. L.; Samain, F.; Häner, R. *Angew. Chem., Int. Ed.* **2007,** *46,* 4464–4467. doi:10.1002/anie.200700891
28. Adeyemi, O. O.; Malinovskii, V. L.; Biner, S. M.; Calzaferri, G.; Häner, R. *Chem. Commun.* **2012,** *48,* 9589–9591. doi:10.1039/c2cc34183b
29. Schmucker, W.; Klumpp, S.; Hennrich, F.; Kappes, M.; Wagenknecht, H.-A. *RSC Adv.* **2013,** *3,* 6331–6333. doi:10.1039/c3ra00163f
30. Ghosh, S.; An, X.; Shah, R.; Rawat, D.; Dave, B.; Kar, S.; Talapatra, S. *J. Phys. Chem. C* **2012,** *116,* 20688–20693. doi:10.1021/jp303339f
31. MacKinnon, K. F.; Qualley, D. F.; Wosky, S. A. *Tetrahedron Lett.* **2007,** *48,* 8074–8077. doi:10.1016/j.tetlet.2007.09.019
32. Tedeschi, T.; Tonelli, A.; Sforza, S.; Corradini, R.; Marchelli, R. *Artif. DNA: PNA XNA* **2010,** *1,* 83–89. doi:10.4161/adna.1.2.13899
33. Boonlua, C.; Vilaivan, C.; Wagenknecht, H.-A.; Vilaivan, T. *Chem.–Asian J.* **2011,** *6,* 3251–3259. doi:10.1002/asia.201100490
34. Trkulja, I.; Biner, S. M.; Langenegger, S. M.; Häner, R. *ChemBioChem* **2007,** *8,* 25–27. doi:10.1002/cbic.200600378
35. Manicardi, A.; Accetta, A.; Tedeschi, T.; Sforza, S.; Marchelli, R.; Corradini, R. *Artif. DNA: PNA XNA* **2012,** *3,* 53–62. doi:10.4161/adna.20158
36. Feriotto, G.; Corradini, R.; Sforza, S.; Bianchi, N.; Mischiati, C.; Marchelli, R.; Gambari, R. *Lab. Invest.* **2001,** *81,* 1415–1427. doi:10.1038/labinvest.3780355
37. Corradini, R.; Feriotto, G.; Sforza, S.; Marchelli, R.; Gambari, R. *J. Mol. Recognit.* **2004,** *17,* 76–84. doi:10.1002/jmr.646
38. Kalyanasundaram, K.; Thomas, J. K. *J. Phys. Chem.* **1977,** *81,* 2176–2180. doi:10.1021/j100538a008

## License and Terms

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

# Influence of perylenediimide–pyrene supramolecular interactions on the stability of DNA-based hybrids: Importance of electrostatic complementarity

Christian B. Winiger, Simon M. Langenegger, Oleg Khorev and Robert Häner*

Full Research Paper

Open Access

Address:
Department of Chemistry and Biochemistry, University of Bern, Freiestrasse 3, CH-3012 Bern, Switzerland

Email:
Robert Häner* - robert.haener@dcb.unibe.ch

* Corresponding author

Keywords:
DNA; hybridization; nucleic acids; perylenediimide; pyrene

*Beilstein J. Org. Chem.* **2014,** *10,* 1589–1595.
doi:10.3762/bjoc.10.164

Received: 10 April 2014
Accepted: 17 June 2014
Published: 11 July 2014

This article is part of the Thematic Series "Nucleic acid chemistry".

Guest Editor: H.-A. Wagenknecht



## Abstract

Aromatic π–π stacking interactions are ubiquitous in nature, medicinal chemistry and materials sciences. They play a crucial role in the stacking of nucleobases, thus stabilising the DNA double helix. The following paper describes a series of chimeric DNA–polycyclic aromatic hydrocarbon (PAH) hybrids. The PAH building blocks are electron-rich pyrene and electron-poor perylenediimide (PDI), and were incorporated into complementary DNA strands. The hybrids contain different numbers of pyrene–PDI interactions that were found to directly influence duplex stability. As the pyrene–PDI ratio approaches 1:1, the stability of the duplexes increases with an average value of 7.5 °C per pyrene–PDI supramolecular interaction indicating the importance of electrostatic complementarity for aromatic π–π stacking interactions.

## Introduction

When two aromatic molecules are in close proximity they often have a tendency to interact non-covalently in a face-to-face stacking arrangement. Face-centered, parallel aromatic π–π stacking interactions have been studied and reviewed in great detail [1-5]. These interactions are especially important for polycyclic aromatic hydrocarbons (PAHs) [6,7]. The interaction is the result of solvophobicity, as well as van der Waals, electrostatic and charge transfer interactions that can lead to a thermodynamically favourable association [8]. It is an important interaction in biological systems, drug receptor interactions, materials sciences, and supramolecular chemistry [8-12]. Such interactions are strongly dependent on the electron density and distribution of the partners [2,9,13-16]. In particular, the interaction between electron-rich (donor) and electron-deficient (acceptor) aromatic rings results in stable aggregates [17-22].

In the DNA duplex, the interaction of the two complementary strands is governed mainly by aromatic π–π stacking interactions, hydrogen bonds, and electrostatic repulsion from the

negatively-charged sugar phosphate backbone [10,23-28]. DNA can be regarded as an amphiphilic polymer in which aromatic residues are linked by negatively charged phosphodiester groups [29]. The importance of aromatic and hydrophobic factors for duplex stability was demonstrated by replacing the natural nucleobases by size expanded analogs [30-35].

A classic example of polymeric donor–acceptor complexes are the aedamers (*a*romatic *e*lectron *d*onor *a*cceptor oligo*mers*) pioneered by Iverson and coworkers [18,36,37]. They consist of face-to-face stacked electron-rich naphthalene and electron-poor naphthalenediimide (NDI) chromophores and belong to the broader area of foldamers [38].

DNA has been described as a molecular scaffold for arranging various types of chromophores [39-44]. Recently, we reported that oligoarenotides (oligomers with an alternating phosphodiester-aromatic hydrocarbon motif) exhibit similar structural properties as nucleic acids, and although the aromatic hydrocarbons cannot engage in any sort of Watson–Crick related hydrogen bonding, the individual strands interact via an interstrand stacking motif [45-48]. Herein we describe a series of DNA-based hybrids (Figure 1 and Table 1) containing electron-rich 1,8-dialkynylpyrenes (**Y**) and electron-poor perylenediimides (PDI, **E**).

PDIs (Figure 1A) are some of the most widely studied organic chromophores [49-52]. Moreover, we have reported on the aggregation and stacking properties of 1,8- and 1,6-dialkynylpyrene [53,54]. Figure 1B shows the electrostatic potential surface of 1,8-diprop-1-ynylpyrene and *N*,*N*'-dimethyl-PDI. The former is considerably more electron-rich/higher electron density (red) than the latter, which is expected to favour an alternating aromatic π–π stacking arrangement of **E** and **Y** due to electrostatic complementarity.

We show herein that duplex formation by our chimeric DNA-oligoarenotide strands proceeds in a selective manner, the chromophores on opposite strands interdigitate and stack face-to-face in an organised controlled assembly.

## Results and Discussion

The principle of the system is illustrated in Figure 1. All oligomers are composed of a DNA part and a modified section containing a total of four PDIs (blue) and/or pyrenes (green). Oligomers **1**–**7** consisting of varying numbers of pyrene or PDI moieties covalently linked to complementary DNA strands were prepared by automated oligonucleotide synthesis using the previously described phosphoramidite pyrene [53] and PDI [55] building blocks.

**Table 1:** $T_m$ values of the hybrids determined by thermal denaturation experiments.[a]

| | Sequence | $T_m$ (°C) | Number of pyrene–PDI interactions |
|---|---|---|---|
| Ref | 5' GCGTTA<br>3' CGCAAT | 13.0 | |
| **1**<br>**2** | 5' GCGTTA **YYYY**<br>3' CGCAAT **YYYY** | 50.5 | 0 |
| **1**<br>**3** | 5' GCGTTA **YYYY**<br>3' CGCAAT **YYEY** | 54.5 | 2 |
| **1**<br>**4** | 5' GCGTTA **YYYY**<br>3' CGCAAT **YEYE** | 58.5 | 4 |
| **1**<br>**5** | 5' GCGTTA **YYYY**<br>3' CGCAAT **YEEE** | 61.0 | 6 |
| **1**<br>**6** | 5' GCGTTA **YYYY**<br>3' CGCAAT **EEEE** | 64.5 | 7 |
| **7**<br>**2** | 5' GCGTTA **EEEE**<br>3' CGCAAT **YYYY** | 66.5 | 7 |
| **7**<br>**6** | 5' GCGTTA **EEEE**<br>3' CGCAAT **EEEE** | 52.0 | 0 |
| **1**<br>**7** | 5' GCGTTA **YYYY**<br>5' GCGTTA **EEEE** | –[b] | n/a |
| **2**<br>**6** | 3' CGCAAT **YYYY**<br>3' CGCAAT **EEEE** | –[b] | n/a |

[a]Conditions: oligomer conc. 2.5 μM single strand, 10 mM sodium phosphate buffer, pH 7.2, 0.1 M NaCl, absorption monitored at 260 nm; error ±0.5 °C; [b]no transition observed (see Supporting Information File 1).

The DNA stem acts as a supramolecular scaffold, and together with the flexible, negatively-charged phosphate linker allows the chromphores to adopt optimal conformations in aqueous solution and increases the solubility. The strands were hybridised in various combinations (Table 1), and their stability and photophysical properties were investigated. Since the DNA duplex is identical in all hybrids, differences in stability must originate from the modified section. The sequence of the modified part is changed in such a way that annealing of different strands leads to a varying number of pyrene–PDI stacking interactions. Strand **1** is common to all hybrids. The complementary strands **2**–**6** possess an increasing number of PDIs. Thus, in the resultant hybrids, the number of pyrene–PDI face-to-face stacking interactions also increases steadily from left to right, e.g., duplex **1*2** contains only pyrene–pyrene interactions, whereas duplex **1*6** has the maximum number pyrene–PDI interactions.

### Thermal denaturation experiments

Thermal denaturation experiments revealed a clear trend in duplex stability (Figure 2). The thermal stability correlates with the number of pyrene–PDI interactions [56] and increases linearly in the series. The melting temperature ($T_m$) values are summarized in Table 1.

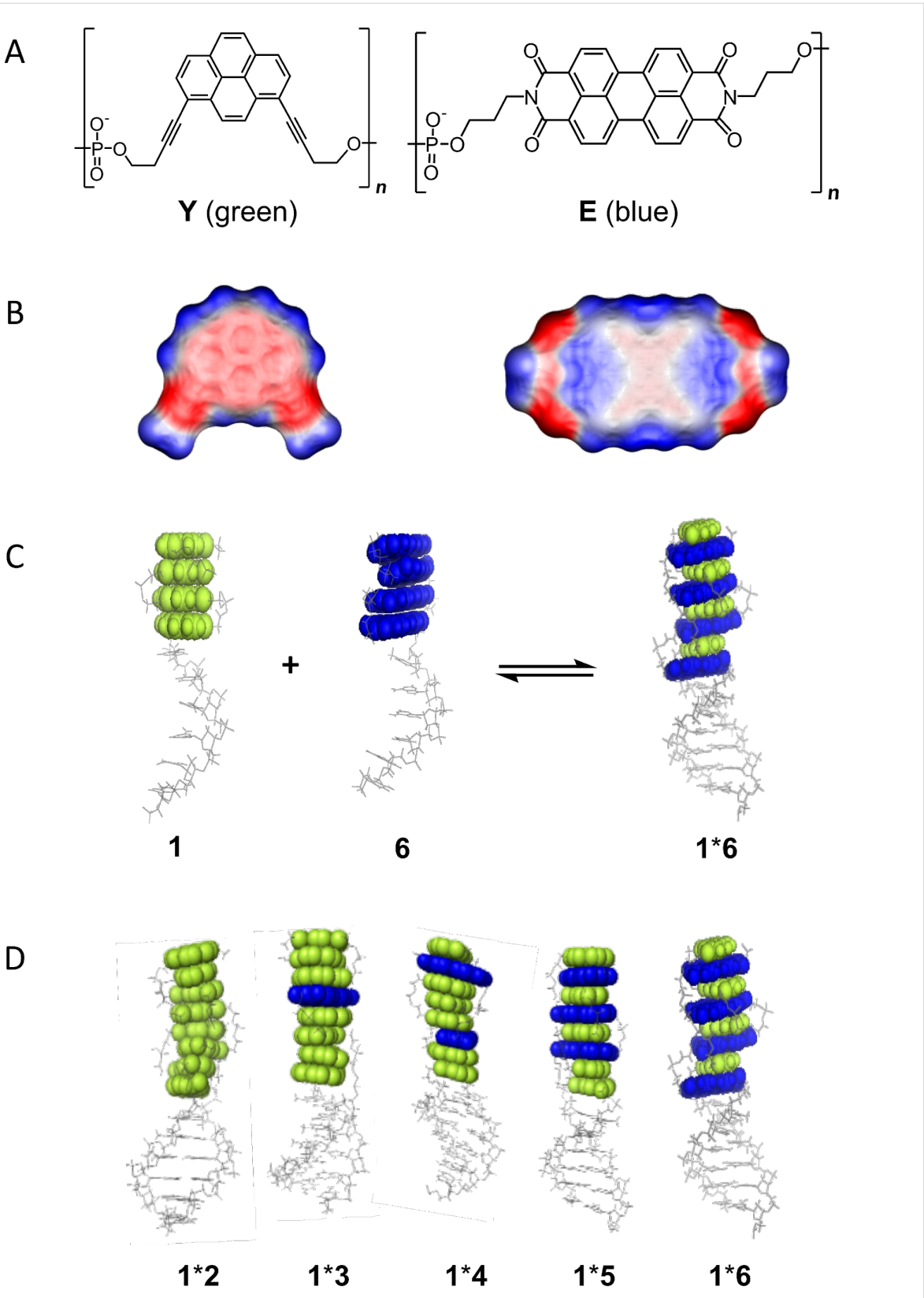



**Figure 1:** (A) Structures of 1,8-dialkynylpyrene (**Y**) and PDI (**E**); (B) illustration of the electrostatic potential surface of 1,8-diprop-1-ynylpyrene (left) and *N*,*N*′-dimethyl-PDI (right); (C) illustration of duplex formation with chimeric oligomers; (D) hybrids **1*****2** to **1*****6**. The number of pyrene–PDI interactions increases from left to right.

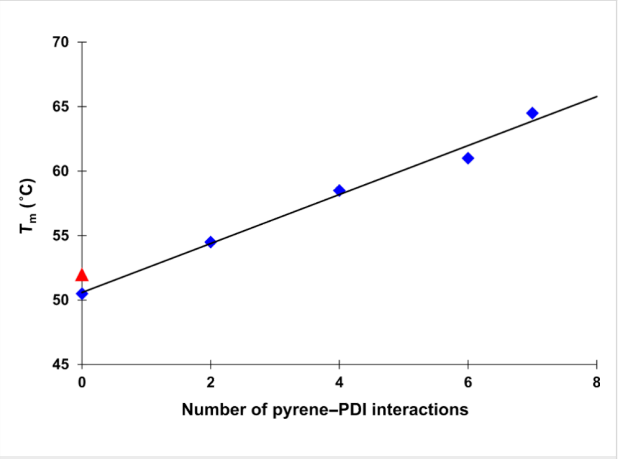



**Figure 2:** A plot of melting temperature ($T_m$) versus the number of pyrene–PDI interactions for duplexes **1*2** to **1*6** (from left to right) presented in Table 1. The $T_m$ was recorded at 260 nm; $R^2$ = 0.987. The red triangle represents the $T_m$ of the control hybrid **7*6**.

Hybrid **1*2** has a $T_m$ of 50.5 °C which is 37.5 °C higher than the reference DNA duplex ($T_m$ = 13 °C). Since hybrid **1*2** has seven pyrene–pyrene interactions, one of these interactions ($\Delta T_{m/(Y-Y)}$) contributes ≈ 5.4 °C to hybrid stability. Likewise, a value for $\Delta T_{m/(E-E)}$ = 5.6 °C is calculated for hybrid **7*6**. The average contribution of a pyrene–PDI interaction can be calculated from the $T_m$ difference ($T_m$ = 51.5 °C) between **1*6** and the reference duplex. A value of $\Delta T_{m/(Y-E)}$ = 7.4 °C is obtained in this way. Hybrid **7*2** serves as a further control. In this duplex, the DNA and the modified parts of the two strands have been interchanged relative to **1*6**. The $T_m$ value of **7*2** is in the same range as **1*6** (66.5 and 64.5 °C, respectively), which translates into $\Delta T_{m/(Y-E)}$ = 7.7 °C. Thus, the average contributions to the hybrid stabilities are as follows: $\Delta T_{m/(Y-E)}$ ≈ 7.5 °C, whereas $\Delta T_{m/(Y-Y)}$ and $\Delta T_{m/(E-E)}$ ≈ 5.5 °C.

The results of electrostatic complementarity between an electron-rich pyrene and an electron-poor PDI is highlighted by the fact that duplexes with only pyrene or PDI are considerably less stable (Table 1, hybrids **1*2** and **7*6**) than hybrids containing both types of aromatic compounds. Although the actual stability of such duplexes strongly depends on several parameters like, e.g., the geometry of the building blocks and the flexibility of the linkers, a general trend can be deduced from the thermal denaturation results that accounts for the above mentioned design of building blocks and sequences.

## UV–vis absorption spectroscopy

The stacking interactions of **Y** and **E** in the hybrids could be followed by UV–vis absorption spectroscopy. A significant change in the vibronic band ratio supports the model of an alternating interstrand interdigitation interaction between pyrene and PDI chromophores as proposed in Figure 1. In general, the vibronic band ratio of PAHs gives valuable information on the aggregation state of the molecules [57]. In a stack of only pyrenes (**1*2**) the vibronic band at 370 nm is higher than that at 390 nm (Figure 3), indicating that the pyrenes are stacked parallel and face-to-face. In contrast, in duplex **1*6** the intensity of the vibronic band 390 nm is higher indicating that the pyrenes are separated from each other by intercalating PDIs [58]. The same absorption behaviour is seen for the PDI vibronic band intensities.

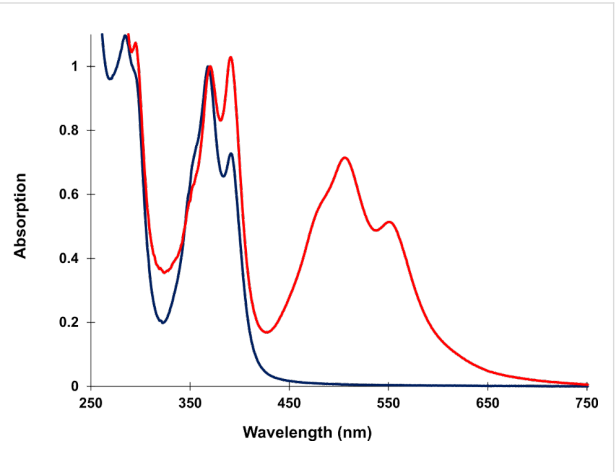



**Figure 3:** UV–vis absorption spectra (scaled) of duplexes **1*2** (blue) and **1*6** (red) at 20 °C. Conditions: see Table 1.

Figure 4 focuses on the vibronic bands of pyrene's $S_0 \rightarrow S_1$ absorption band in duplexes **1*2** to **1*6**. An increasing number of PDIs in a stack leads to a stronger vibronic band at 390 nm.

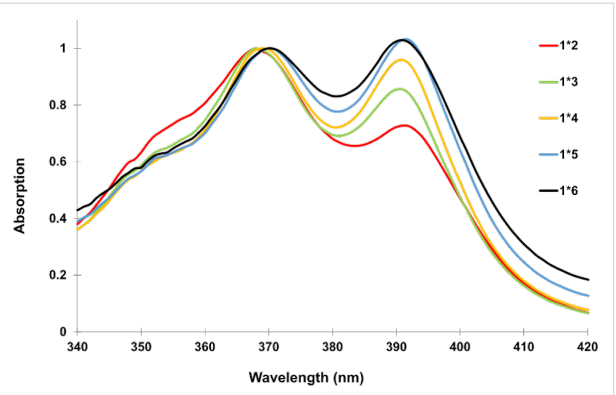



**Figure 4:** UV–vis absorption spectra (scaled) of duplexes **1*2** to **1*6** at 20 °C. Conditions: see Table 1.

This is in stark contrast to the effect of thermally denaturing duplex **1*2** into two single strands (Figure 5). There, the vibronic band at 370 nm is always the highest indicating that the pyrenes are stacked even at 90 °C in the single strands. Such behaviour was also observed in chrysene-modified DNA [59].

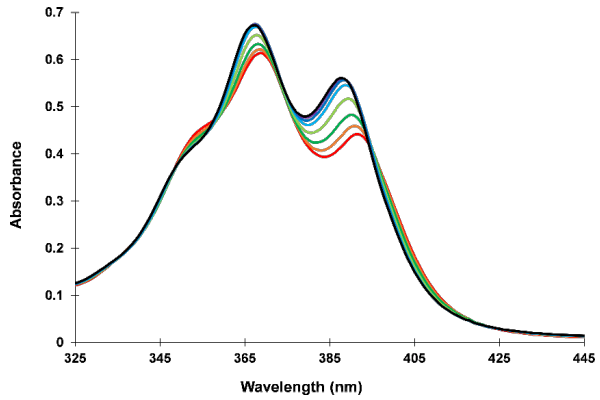



**Figure 5:** Temperature-dependent UV–vis absorption spectrum of **1*2**. Conditions: see Table 1.

## Fluorescence spectroscopy

The interaction of two or more dialkynylpyrenes (**Y**) results in a pronounced excimer fluorescence [53]. Hybridization of single strands **1** and **2** increases the intensity of the excimer (Figure 6), whereas hybridization of single strands **1** and **6** results in an extinction of excimer fluorescence. Such behaviour was also observed in previous work and was explained by an alternating interdigitation interaction of the pyrene with the PDI building blocks [60].

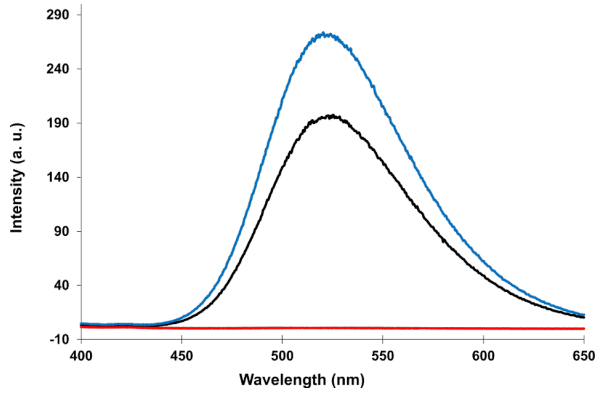



**Figure 6:** Fluorescence spectra of oligomer **1** (black), duplex **1*2** (blue) and duplex **1*6** (red) at 20 °C. Excitation: 370 nm. Conditions: see Table 1.

## Gel migration experiments

The electrophoretic mobility of relatively small (<1000 kbp), linear DNA strands is inversely proportional to their molecular weight [61]. It serves as a reliable method to demonstrate the formation of double versus single stranded DNA structure. Therefore, the formation of defined short duplexes has been further investigated using polyacrylamide gel electrophoresis (PAGE) experiments. Oligomer single strands **1**, **6** and **7** migrate with the same velocity as the 13 bp reference (Figure 7). Strands **1** and **6**, however, form a duplex and thus have lower electrophoretic mobility, similar to an 18–20 bp reference. Oligomer **7** has the same DNA sequence as **6**, but with 4 PDI molecules instead of 4 pyrenes (see Table 1). Thus when combined, single strands **1** and **7** do not form a duplex due to having non-complementary DNA parts, and migrate on the gel like single strands **1**, **6**, and **7**. These results underline the importance of the complementary DNA segments in aligning the pyrene and PDI chromophores for optimal interaction.

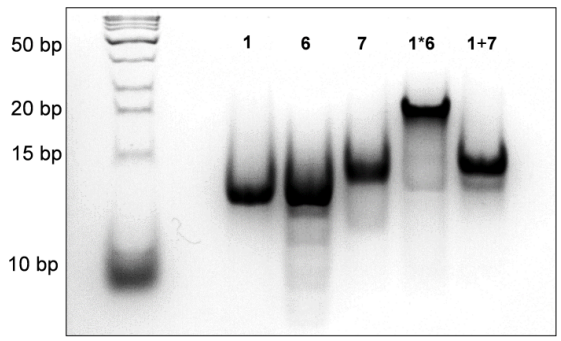



**Figure 7:** PAGE experiment. All oligomers were used in a total amount of 150 pmol in 10 mM sodium phosphate buffer, 100 mM NaCl and 10% loading buffer, 20% polyacrylamide gel with a 10% loading gel, 1 h 40 min, 4 °C, 170 V, 6 mA, 2 W. Left lane: DNA ladder.

## Conclusion

A series of DNA oligonucleotides functionalised with electron-poor perylenediimide (PDI, **E**) and electron-rich 1,8-dialkynylpyrene (**Y**) chromophores has been synthesized and their photophysical and thermal melting properties were investigated. UV–vis absorption and fluorescence spectra indicate an alternate, face-to-face, stacking of PDI and pyrene moieties. The DNA portion serves as an ideal scaffold to align the pyrene and PDI chromophores and to study their interaction in solution. The stability of the duplexes shows a clear dependence on the number of pyrene–PDI interactions within the duplex. As the pyrene–PDI ratio approaches 1:1, the stability of the duplexes increases with up to 7.5 °C per pyrene–PDI supramolecular interaction underlining the importance of electrostatic complementarity for aromatic π–π stacking interactions.

## Supporting Information

### Supporting Information File 1

Detailed experimental procedures and supplementary spectroscopic data.
[http://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-10-164-S1.pdf]

## Acknowledgements

We would like to express our sincere gratitude to Prof. Dr. Gion Calzaferri for invaluable suggestions and discussions regarding this work. This work was supported by the Swiss National Foundation (Grant 200020-149148).

## References

1. Hunter, C. A.; Sanders, J. K. M. *J. Am. Chem. Soc.* **1990,** *112,* 5525–5534. doi:10.1021/ja00170a016
2. Hunter, C. A.; Lawson, K. R.; Perkins, J.; Urch, C. J. *J. Chem. Soc., Perkin Trans. 2* **2001,** 651–669. doi:10.1039/b008495f
3. Salonen, L. M.; Ellermann, M.; Diederich, F. *Angew. Chem., Int. Ed.* **2011,** *50,* 4808–4842. doi:10.1002/anie.201007560
4. Grimme, S. *Angew. Chem., Int. Ed.* **2008,** *47,* 3430–3434. doi:10.1002/anie.200705157
5. Martinez, C. R.; Iverson, B. L. *Chem. Sci.* **2012,** *3,* 2191–2201. doi:10.1039/c2sc20045g
6. Watson, M. D.; Fechtenkötter, A.; Müllen, K. *Chem. Rev.* **2001,** *101,* 1267–1300. doi:10.1021/cr990322p
7. Feng, C.; Lin, C. S.; Fan, W.; Zhang, R. Q.; Van Hove, M. A. *J. Chem. Phys.* **2009,** *131,* 194702. doi:10.1063/1.3251785
8. Hunter, C. A. *Chem. Soc. Rev.* **1994,** *23,* 101–109. doi:10.1039/cs9942300101
9. Meyer, E. A.; Castellano, R. K.; Diederich, F. *Angew. Chem., Int. Ed.* **2003,** *42,* 1210–1250. doi:10.1002/anie.200390319
10. Hunter, C. A. *J. Mol. Biol.* **1993,** *230,* 1025–1054. doi:10.1006/jmbi.1993.1217
11. Philp, D.; Stoddart, J. F. *Angew. Chem., Int. Ed. Engl.* **1996,** *35,* 1154–1196. doi:10.1002/anie.199611541
12. Sakai, N.; Matile, S. *Beilstein J. Org. Chem.* **2012,** *8,* 897–904. doi:10.3762/bjoc.8.102
13. Cockroft, S. L.; Perkins, J.; Zonta, C.; Adams, H.; Spey, S. E.; Low, C. M. R.; Vinter, J. G.; Lawson, K. R.; Urch, C. J.; Hunter, C. A. *Org. Biomol. Chem.* **2007,** *5,* 1062–1080. doi:10.1039/b617576g
14. Wheeler, S. E. *Acc. Chem. Res.* **2013,** *46,* 1029–1038. doi:10.1021/ar300109n
15. Collings, J. C.; Roscoe, K. P.; Robins, E. G.; Batsanov, A. S.; Stimson, L. M.; Howard, J. A. K.; Clark, S. J.; Marder, T. B. *New J. Chem.* **2002,** *26,* 1740–1746. doi:10.1039/b207102a
16. Ponzini, F.; Zagha, R.; Hardcastle, K.; Siegel, J. S. *Angew. Chem., Int. Ed.* **2000,** *39,* 2323–2325. doi:10.1002/1521-3773(20000703)39:13<2323::AID-ANIE2323>3.0.CO;2-X
17. Lokey, R. S.; Iverson, B. L. *Nature* **1995,** *375,* 303–305. doi:10.1038/375303a0
18. Gabriel, G. J.; Sorey, S.; Iverson, B. L. *J. Am. Chem. Soc.* **2005,** *127,* 2637–2640. doi:10.1021/ja046722y
19. Mathis, G.; Hunziker, J. *Angew. Chem., Int. Ed.* **2002,** *41,* 3203–3205. doi:10.1002/1521-3773(20020902)41:17<3203::AID-ANIE3203>3.0.CO;2-K
20. Tanaka, H.; Bollot, G.; Mareda, J.; Litvinchuk, S.; Tran, D.-H.; Sakai, N.; Matile, S. *Org. Biomol. Chem.* **2007,** *5,* 1369–1380. doi:10.1039/b702255g
21. Das, A.; Molla, M. R.; Maity, B.; Koley, D.; Ghosh, S. *Chem.–Eur. J.* **2012,** *18,* 9849–9859. doi:10.1002/chem.201201140
22. Ghosh, S.; Ramakrishnan, S. *Macromolecules* **2005,** *38,* 676–686. doi:10.1021/ma0478759
23. Hunter, C. A. *BioEssays* **1996,** *18,* 157–162. doi:10.1002/bies.950180212
24. Eschenmoser, A. *Science* **1999,** *284,* 2118–2124. doi:10.1126/science.284.5423.2118
25. Kool, E. T. *Chem. Rev.* **1997,** *97,* 1473–1488. doi:10.1021/cr9603791
26. Herdewijn, P. *Biochim. Biophys. Acta, Gene Struct. Expression* **1999,** *1489,* 167–179. doi:10.1016/S0167-4781(99)00152-9
27. Benner, S. A. *Acc. Chem. Res.* **2004,** *37,* 784–797. doi:10.1021/ar040004z
28. Wengel, J. *Acc. Chem. Res.* **1999,** *32,* 301–310. doi:10.1021/ar980051p
29. Benner, S. A.; Hutter, D. *Bioorg. Chem.* **2002,** *30,* 62–80. doi:10.1006/bioo.2001.1232
30. Yang, Z. Y.; Hutter, D.; Sheng, P.; Sismour, A. M.; Benner, S. A. *Nucleic Acids Res.* **2006,** *34,* 6095–6101. doi:10.1093/nar/gkl633
31. Hirao, I.; Kimoto, M.; Yamashige, R. *Acc. Chem. Res.* **2012,** *45,* 2055–2065. doi:10.1021/ar200257x
32. Wu, Y. Q.; Ogawa, A. K.; Berger, M.; McMinn, D. L.; Schultz, P. G.; Romesberg, F. E. *J. Am. Chem. Soc.* **2000,** *122,* 7621–7632. doi:10.1021/ja0009931
33. Berger, M.; Ogawa, A. K.; McMinn, D. L.; Wu, Y.; Schultz, P. G.; Romesberg, F. E. *Angew. Chem., Int. Ed.* **2000,** *39,* 2940–2942. doi:10.1002/1521-3773(20000818)39:16<2940::AID-ANIE2940>3.0.CO;2-#
34. Liu, H.; Gao, J.; Maynard, L.; Saito, Y. D.; Kool, E. T. *J. Am. Chem. Soc.* **2004,** *126,* 1102–1109. doi:10.1021/ja038384r
35. Brotschi, C.; Häberli, A.; Leumann, C. J. *Angew. Chem., Int. Ed.* **2001,** *40,* 3012–3014. doi:10.1002/1521-3773(20010817)40:16<3012::AID-ANIE3012>3.0.CO;2-Y
36. Lokey, R. S. *Curr. Opin. Chem. Biol.* **2003,** *7,* 91–96. doi:10.1016/S1367-5931(02)00002-9
37. Ikkanda, B. A.; Samuel, S. A.; Iverson, B. L. *J. Org. Chem.* **2014,** *79,* 2029–2037. doi:10.1021/jo402704z
38. Hecht, S.; Huc, I., Eds. *Foldamers - Structure, Properties, and Applications;* Wiley-VCH: Weinheim, 2007.
39. Filichev, V. V.; Pedersen, E. B. DNA-Conjugated Organic Chromophores in DNA Stacking Interactions. In *Wiley Encyclopedia of Chemical Biology;* Begley, T. P., Ed.; Wiley: Hoboken, 2009; Vol. 1.
40. Varghese, R.; Wagenknecht, H.-A. *Chem. Commun.* **2009,** 2615–2624. doi:10.1039/b821728a
41. Malinovskii, V. L.; Wenger, D.; Häner, R. *Chem. Soc. Rev.* **2010,** *39,* 410–422. doi:10.1039/b910030j
42. Østergaard, M. E.; Hrdlicka, P. J. *Chem. Soc. Rev.* **2011,** *40,* 5771–5788. doi:10.1039/c1cs15014f
43. Stulz, E. *Chem.–Eur. J.* **2012,** *18,* 4456–4469. doi:10.1002/chem.201102908
44. Asanuma, H.; Fujii, T.; Kato, T.; Kashida, H. *J. Photochem. Photobiol., C* **2012,** *13,* 124–135. doi:10.1016/j.jphotochemrev.2012.04.002
45. Häner, R.; Garo, F.; Wenger, D.; Malinovskii, V. L. *J. Am. Chem. Soc.* **2010,** *132,* 7466–7471. doi:10.1021/ja102042p
46. Rudnev, A. V.; Malinovskii, V. L.; Nussbaumer, A. L.; Mishchenko, A.; Häner, R.; Wandlowski, T. *Macromolecules* **2012,** *45,* 5986–5992. doi:10.1021/ma3007619
47. Simona, F.; Nussbaumer, A. L.; Häner, R.; Cascella, M. *J. Phys. Chem. B* **2013,** *117,* 2576–2585. doi:10.1021/jp310320f
48. Häner, R.; Samain, F.; Malinovskii, V. L. *Chem.–Eur. J.* **2009,** *15,* 5701–5708. doi:10.1002/chem.200900369

49. Langhals, H. *Helv. Chim. Acta* **2005,** *88,* 1309–1343. doi:10.1002/hlca.200590107
50. Weil, T.; Vosch, T.; Hofkens, J.; Peneva, K.; Müllen, K. *Angew. Chem., Int. Ed.* **2010,** *49,* 9068–9093. doi:10.1002/anie.200902532
51. Würthner, F. *Chem. Commun.* **2004,** 1564–1579. doi:10.1039/b401630k
52. Neelakandan, P. P.; Zeidan, T. A.; McCullagh, M.; Schatz, G. C.; Vura-Weis, J.; Kim, C. H.; Wasielewski, M. R.; Lewis, F. D. *Chem. Sci.* **2014,** *5,* 973–981. doi:10.1039/c3sc52908h
53. Bittermann, H.; Siegemund, D.; Malinovskii, V. L.; Häner, R. *J. Am. Chem. Soc.* **2008,** *130,* 15285–15287. doi:10.1021/ja806747h
54. Vybornyi, M.; Rudnev, A. V.; Langenegger, S. M.; Wandlowski, T.; Calzaferri, G.; Häner, R. *Angew. Chem., Int. Ed.* **2013,** *52,* 11488–11493. doi:10.1002/anie.201307029
55. Rahe, N.; Rinn, C.; Carell, T. *Chem. Commun.* **2003,** 2120–2121. doi:10.1039/B307395E
56. It must be noted at this point, that the exact sequence of pyrene/PDI alternation can only be guessed, since the first residue stacking on the DNA stem could belong to either of the two strands. To simplify all subsequent description and discussion we have chosen that the first modified residue is the one attached at the nucleotide 5'-position.
57. Clark, A. E.; Qin, C.; Li, A. D. Q. *J. Am. Chem. Soc.* **2007,** *129,* 7586–7595. doi:10.1021/ja0687724
58. Biner, S. M.; Kummer, D.; Malinovskii, V. L.; Häner, R. *Org. Biomol. Chem.* **2011,** *9,* 2628–2633. doi:10.1039/c0ob01132k
59. Khorev, O.; Bösch, C. D.; Probst, M.; Häner, R. *Chem. Sci.* **2014,** *5,* 1506–1512. doi:10.1039/c3sc53316f
60. Bouquin, N.; Malinovskii, V. L.; Häner, R. *Chem. Commun.* **2008,** 1974–1976. doi:10.1039/b802193g
61. Bloomfield, V. A.; Crothers, D. M.; Tinoco, I., Jr. *Nucleic Acids - Structures, Properties, and Functions;* University Science Books: Sausalito, CA, 2000.


## License and Terms

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

# Structure/affinity studies in the bicyclo-DNA series: Synthesis and properties of oligonucleotides containing bc$^{en}$-T and iso-tricyclo-T nucleosides

Branislav Dugovic, Michael Wagner and Christian J. Leumann*

Full Research Paper

Open Access

Address:
Department of Chemistry and Biochemistry, University of Bern, Freiestrasse 3, CH-3012 Bern, Switzerland

Email:
Christian J. Leumann* - christian.leumann@dcb.unibe.ch

* Corresponding author

Keywords:
DNA/RNA affinity; nucleic acids; nucleosides; oligonucleotides; oligonucleotide therapy; X-ray structures

*Beilstein J. Org. Chem.* **2014,** *10,* 1840–1847.
doi:10.3762/bjoc.10.194

Received: 21 May 2014
Accepted: 22 July 2014
Published: 12 August 2014

This article is part of the Thematic Series "Nucleic acid chemistry".

Guest Editor: H.-A. Wagenknecht



## Abstract

We present the synthesis of the two novel nucleosides iso-tc-T and bc$^{en}$-T, belonging to the bicyclo-/tricyclo-DNA molecular platform. In both modifications the torsion around C6'–C7' within the carbocyclic ring is planarized by either the presence of a C6'–C7' double bond or a cyclopropane ring. Structural analysis of these two nucleosides by X-ray analysis reveals a clear preference of torsion angle γ for the gauche orientation with the furanose ring in a near perfect 2'-endo conformation. Both modifications were incorporated into oligodeoxynucleotides and their thermal melting behavior with DNA and RNA as complements was assessed. We found that the iso-tc-T modification was significantly more destabilizing in duplex formation compared to the bc$^{en}$-T modification. In addition, duplexes with complementary RNA were less stable as compared to duplexes with DNA as complement. A structure/affinity analysis, including the already known bc-T and tc-T modifications, does not lead to a clear correlation of the orientation of torsion angle γ with DNA or RNA affinity. There is, however, some correlation between furanose conformation (N- or S-type) and affinity in the sense that a preference for a 3'-endo like conformation is associated with a preference for RNA as complement. As a general rule it appears that $T_m$ data of single modifications with nucleosides of the bicyclo-/tricyclo-DNA platform within deoxyoligonucleotides are not predictive for the stability of fully modified oligonucleotides.

## Introduction

Antisense oligonucleotides (ASOs) can interfere with gene expression via various biological mechanisms, depending on the nature of the cellular RNA target [1]. First and foremost they can inhibit translation by targeting a mature mRNA in either its coding or non-coding part of the sequence by a steric block or an RNase H dependent degradation mechanism. Furthermore, it has recently been shown that ASOs can alter RNA splicing when targeting exon/intron junctions or splice enhancer or silencer binding sites on pre-mRNAs, thus leading to alternative splicing [2,3], to exon skipping [4,5] or to exon inclu-

sion [6]. In addition they can restore the function of mRNAs containing extended aberrant repeat sequences in their non-coding region by either restoring correct cellular localization or inhibiting vital protein sequestration by the aberrant repeats [7]. Last but not least, there is an ever growing number of micro RNAs (miRNAs) that are involved in genetic and epigenetic regulation of gene expression. Their misregulation stays at the onset of various forms of cancer and other metabolic diseases, and targeting of such miRNAs with ASOs (antimirs or anatagomirs) has been shown in the recent past to be a promising therapeutic principle [8].

There exists a multitude of chemical modifications in ASOs. Historically, the first modification was the replacement of the phosphodiester linking units in DNA by phosphorothioate groups, thus conferring higher metabolic stability to ASOs in plasma and tissue [9,10]. Another site of modification is the 2'-OH group of RNA that can be equipped or replaced with various chemical entities typically aiming at higher affinities to the corresponding RNA targets [11-14]. More diverse analogues include structures in which the sugar phosphate backbone has been replaced by a charge neutral peptide backbone, such as the peptide nucleic acids (PNAs) [15] or by a nucleotide derived phosphorodiamidate backbone, such as the morpholino oligonucleotides (PMOs) [16]. Of particular interest is the class of conformationally constrained oligonucleotides. Members of this class are amongst others the locked nucleic acids (LNA) [17,18], the hexose nucleic acids (HNAs) [19] and the family of bi- and tricyclo-DNA (Figure 1) [20-23]. These analogues aim at increasing RNA affinity by structurally preorganizing single strands for duplex formation.

Bc- and tc-DNA have been conceived to reduce the entropy upon duplex formation with a nucleic acid target by reducing the conformational flexibility around the C3'–C4' and C4'–C5' bonds, while achieving as much as possible of a geometric match with the backbone conformation of DNA in duplexed form. From this a gain in the free energy of duplex formation and, hence, more stable duplexes are expected [24]. Over the years we became interested in determining the structure/RNA affinity relationship of the underlying sugar scaffold and to develop them into a molecular platform for oligonucleotide therapeutics. Given the exclusiveness of the ethylene bridge between the centers 3' and 5' with respect to DNA or RNA we have identified this structural element to be the primary goal for chemical modification [25-30]. In continuation of this work we decided to investigate on two novel thymine nucleosides with restricted conformation of the C6'–C7' bond, namely $bc^{en}$-T and iso-tc-T (Figure 1). Here we present the synthesis and X-ray structural characterization of the respective nucleosides, their incorporation into oligodeoxynucleotides by phosphoramidite chemistry as well as the DNA and RNA affinity profiles of the modified oligonucleotides.

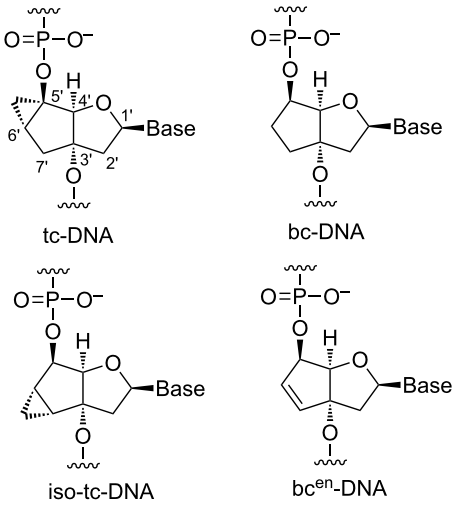
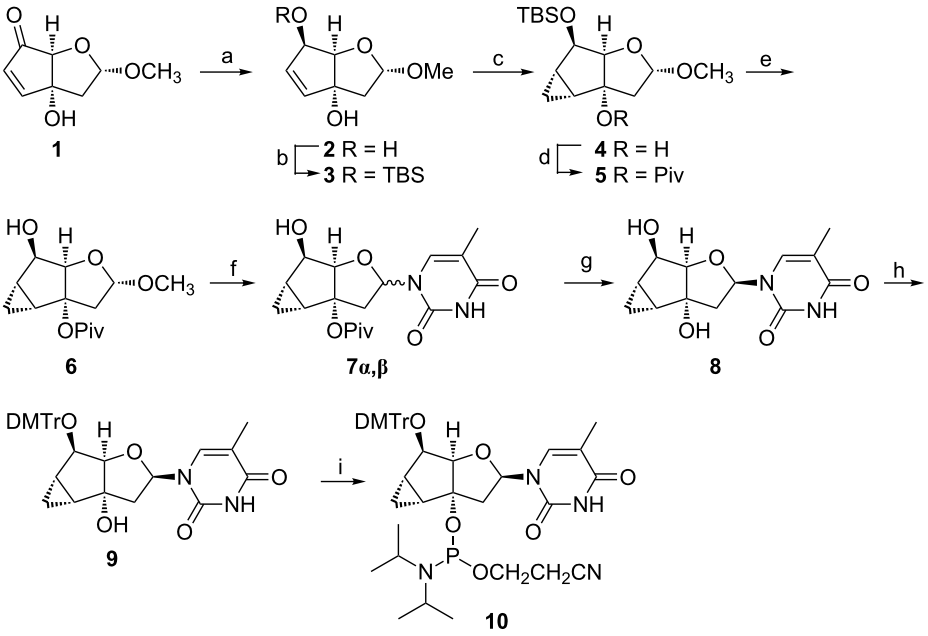



**Figure 1:** Chemical structures and carbon numbering scheme of tricyclo(tc)-DNA (top, left), bicyclo(bc)-DNA (top, right) and the newly synthesized iso-tricyclo(iso-tc)-DNA (bottom, left) and bicyclo-en($bc^{en}$)-DNA (bottom, right).

# Results

## Synthesis of building blocks

The synthesis of the phosphoramidite **10** started with the bicyclic intermediate **1** that had previously been described on the way to related bicyclo-DNA derivatives (Scheme 1) [30]. Following an obvious synthetic strategy, compound **1** was subjected to carbonyl reduction which occurred with high stereoselectivity from the less hindered, convex side of the bicyclic system, resulting in alcohol **2** along with traces of its epimer. To increase the stereoselectivity of the upcoming cyclopropanation reaction it seemed appropriate to protect the secondary hydroxy group as TBS ether (→ **3**). Indeed cyclopropanation of **3** with diethylzinc and $CH_2I_2$ proceeded stereospecifically, again from the convex side of the bicyclic system, to give **4**. Subsequent nucleosidation of **4** via the Vorbrüggen procedure [31,32] with transient protection of the tertiary hydroxy group in **4** as TMS ether, however, was unsuccessful and yielded only the corresponding α-nucleoside in yields below 25%. We reasoned that the exclusive formation of α-nucleosides is due to the steric bulk of the TBS group, further suppressing the intrinsically disfavored β-(endo)-face attack of the base. To counterbalance these effects we chose to protect the tertiary hydroxy group as a pivaloyl ester thus increasing the steric bulk on the α-face, and relieving that on the β-face by replacing the TBS by a transient TMS group. The conversion of **4** → **6** proceeded smoothly and indeed, the use of compound **6** as nucleobase acceptor improved the yield of nucleosides **7α**,**β**



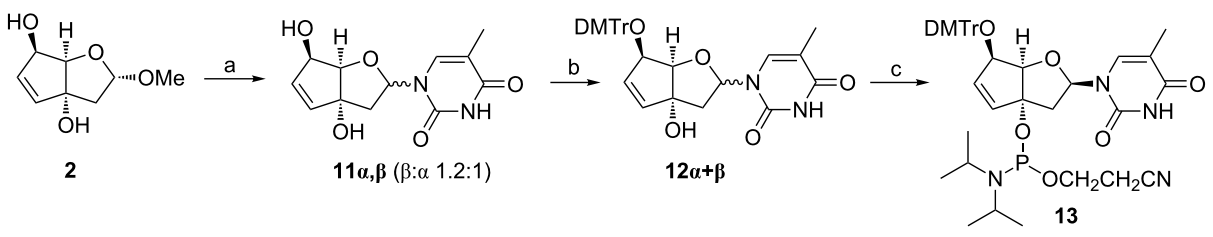
**Scheme 1:** Conditions: (a) $NaBH_4$, $CeCl_3{\cdot}7H_2O$, MeOH, −78 °C → rt, 1.5 h, 73% (+9% of C6-epimer); (b) TBS-Cl, imidazole, $CH_2Cl_2$, rt, 16 h, 79%; (c) $Et_2Zn$ in hexane (1 M), $CH_2I_2$, $CH_2Cl_2$, 0 °C → rt, 16 h, 86%; (d) PivCl, DMAP, pyridine, $ClH_2C–CH_2Cl$, 70 °C, quant.; (e) TBAF, THF, rt, 19 h, 95%; (f) thymine, BSA, $SnCl_4$, $CH_3CN$, 0 °C → rt, 17 h, 56% **7β** + 22% (**7α/β** 2.5:1); (g) $Bu_4NOH$, $H_2O$/dioxane, rt, 16 h, 94%; (h) DMTrCl, pyridine, $CH_2Cl_2$, rt, 24 h, 98%; (i) CEP-Cl, DIPEA, THF, rt, 4 h, 89%.

in general and led to an acceptable β:α = 2.5:1 ratio of anomers. Subsequent saponification of **7α**,**β** (unseparable by flash chromatography) proved to be tricky and after testing a series of standard techniques, only treatment with $Bu_4NOH$ in a mixed organic/aqueous solvent gave nucleosides **8α**,**β** in good yield. It was at this step where the two anomers could be readily separated by flash chromatography. Continuing with **8β** the synthesis of **10** was concluded by standard tritylation (→ **9**) and phosphitylation.

To extend on the structure/nucleic acid affinity profile of this modification we also became interested in nucleoside **11β**, containing a double bond instead of the cyclopropane ring. In the context of oligonucleotides this derivative seemed appropriate to investigate the direct steric influence of the cyclopropyl/methylene group in a bicyclic sugar scaffold that is otherwise very similar in flexibility and geometry. The corresponding building block **13** (Scheme 2) was easily available via nucleosidation of sugar intermediate **2** (in situ TMS protection of both hydroxy groups) leading to the mixture of anomeric nucleosides **11α**,**β** in a ratio of β:α = 1.2:1. After standard tritylation of **11α**,**β** the anomeric mixture **12α**,**β** became separable by flash chromatography and the corresponding β-nucleoside **12β** could be smoothly converted into the phosphoramidite **13** by standard methods.

## Structural properties of nucleosides

To get an independent proof on the relative configuration around the cyclopropane ring and the glycosidic bond and to obtain insight into the conformational properties of the central bicyclic sugar scaffold, crystals of **8β** and **11β** were grown and subjected to X-ray analysis (Figure 2, Table 1). It clearly



**Scheme 2:** Conditions: (a) thymine, BSA, TMSOTf, TMSCl, $CH_3CN$, rt, 2.5 h; (b) DMTrCl, pyridine, rt, 16 h, 29% of **12α** and 34% of **12β** (over two steps); (c) CEP-Cl, DIPEA, THF, rt, 1 h, 94%.

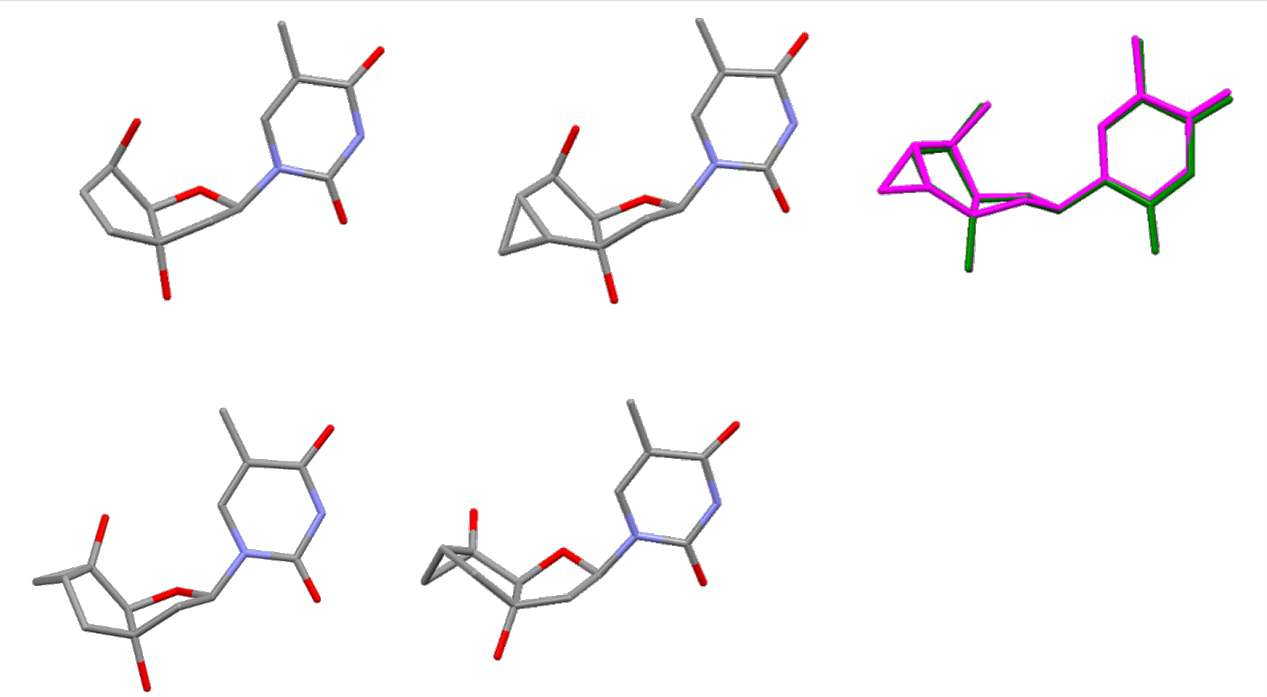


**Figure 2:** X-ray structure of top row: nucleosides **8β** (left), **11β** (center) and overlay of both structures (right); bottom row: tc-T (Mol A, left, Mol B, right).

emerges that in both nucleosides the furanose unit appears in an almost perfect 2'-endo conformation giving rise to a *trans* arrangement of torsion angle δ (O3'–C3'–C4'–C5'). In both structures the cyclopentane ring exists in a shallow envelope conformation with C5' being slightly out of plane. This leads to a *gauche* orientation of torsion angle γ (C3'–C4'–C5'–O5'). In both structures the base thymine is, as expected, in the *anti*-orientation. The overlay of both structures clearly highlights the similarity of both structures, indicating that the extra methylene group in **8β** plays no direct role in controlling the conformation of the bicyclic ring system.

**Table 1:** Selected backbone torsion angles and sugar pucker data for **8β** and **11β** and related bi/tricyclo-nucleosides from X-ray structures.

| | γ | δ | χ | *P* | $\nu_{max}$ |
|---|---|---|---|---|---|
| **8β** | 86.8° | 150.1° | −106.4° | 167.8° | 36.1 |
| **11β** | 86.9° | 146.0° | −115.7° | 160.7° | 36.3 |
| bc-T[a] | 149.3° | 126.5° | −112.7° | 128.4° | 42.4 |
| tc-T[b] Mol A | 125.0° | 152.5° | −130.4° | 172.2° | 36.7 |
| tc-T[b] Mol B | 154.8° | 98.7° | −120.3° | 94.4° | 36.0 |

[a]Ref [33]; [b]two structurally independent molecules per asymmetric unit.

A comparison of **8β** and **11β** with bc-T, having a saturated cyclopentane unit clearly reveals structural differences. The largest deviation is associated with the position of the 5'-OH group which is in a pseudoequatorial orientation in bc-T, giving rise to a torsion angle γ in the anticlinal range. The saturation of the carbocyclic ring translates to a lesser extent also into the furanose ring where a 1'-exo instead of a 2'-endo conformation is observed in bc-T. Both furanose conformations, however, belong to the *S*-type and are thus structurally related in the context of nucleic acid duplex conformation. Quite interestingly, the original tc-T nucleoside [34], for which we solved the X-ray structure here for the first time (Figure 2), shows considerable conformational variability in the furanose part. The asymmetric unit contains two independent molecules (Mol A and Mol B) of which the furanose part in Mol A adopts a 2'-endo (S-type) conformation, while in Mol B a 4'-endo (N-type) conformation is observed. In summary it appears that rigidifying and planarizing the C5'–C6'-bond as in tc-nucleosides leads to an anticlinal orientation of torsion angle γ and variability between S- and N-type in the furanose conformation, whereas rigidifying and planarizing the C6'–C7'-bond, as in **8β** and **11β**, leads to a synclinal torsion angle γ and a consistent 2'-endo furanose conformation. Saturation of the carbocyclic ring, as in bc-T, leads to an anticlinal arrangement of γ and a somewhat attenuated but clear preference for an S-type furanose conformation.

## Oligonucleotide synthesis

The dodecamers **ON1–4**, shown in Table 2, containing one to two modifications, were synthesized in order to test the conse-

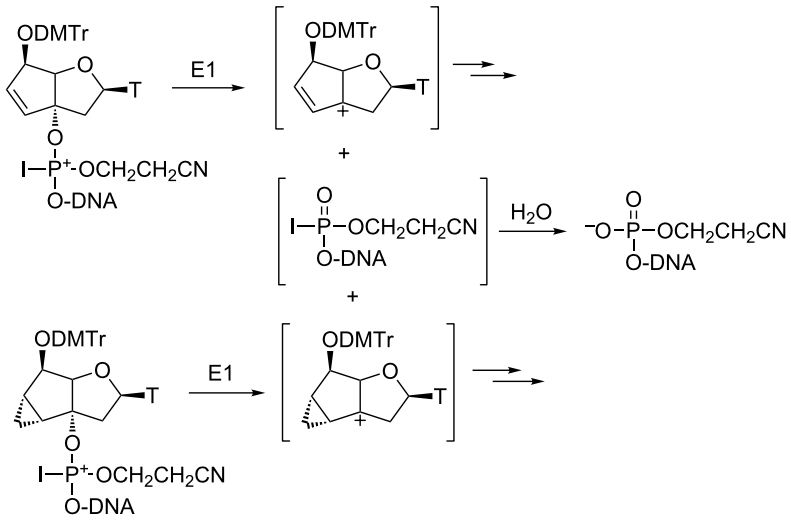
quences of the two modified bicyclic nucleotides on RNA and DNA affinity. **ON5–7**, containing the known tc-T residues in the respective positions, were synthesized for comparison. **ON1–3** were assembled on the 1.3 µmol scale on a DNA synthesizer utilizing standard phosphoramidite chemistry protocols first. The trityl assay after incorporation of **13** and the subsequent building block revealed typically a drop of synthesis yield by roughly 20%. This was also reflected in the HPLC traces after final cleavage from the solid support (33% aq $NH_3$, 55 °C, 16 h) which revealed besides the expected oligonucleotides **ON1–3** also truncated sequences corresponding to 5'-phosphorylated fragments arising from cleavage 3' to the modification as determined by mass spectrometry. Re-subjection of the isolated full length oligonucleotides **ON1–3** to deprotection conditions did not lead to any further degradation, suggesting that E1 elimination of the 3'-P-unit occurs during the oxidation step of the modified residues, most likely on the level of the iodinated phosphite intermediate [35], leading to the formation of an allylic carbocation in the $bc^{en}$-T unit and 5'-phosphorylated DNA fragment (Scheme 3).

The synthesis of oligonucleotides using building block **10** proved to be even more difficult. Using the standard phosphoramidite protocol, the oligonucleotide synthesis failed completely at the site of modification and not even traces of a full length oligonucleotide could be observed after chain assembly and deprotection. Only 5'-phosphorylated, truncated oligonucleotide fragments could be isolated. We reasoned that oxidation with iodine followed by E1 elimination of the 3'-P-unit happend also in this case, leaving behind an alpha-cyclopropyl cation that undergoes subsequent rearrangement (Scheme 3). The fact that elimination is quantitative in the iso-tc-T case may be explained by the release of ring strain during cyclopropyl rearrangement which contributes to the stabilization of the E1 transition state. Based on these assumptions we changed the oxidant from iodine to *t*-BuOOH, which has successfully been used in the past in the allyloxycarbonyl base- and phosphate protecting scheme for oligonucleotide synthesis [36]. Under these conditions, full length oligonucleotide **ON4** could be isolated in 80% as determined by trityl assay. Due to limited availability of the phosphoramidite building block **10**, only this particular oligonucleotide could be obtained in sufficient quantities for biophysical experimentation.

## $T_m$ data

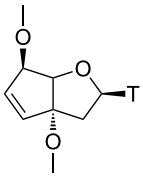
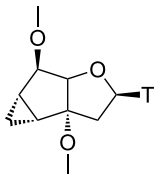
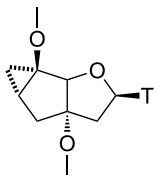
To assess DNA and RNA affinity of the two novel modifications we measured UV-melting curves at 260 nm. With a gradient of 0.5 °C/min the heating and cooling curves are superimposable, indicating equilibrium conditions and excluding degradation of the modified oligonucleotides under the conditions of measurement. The corresponding $T_m$-data are summarized in Table 2.

The $bc^{en}$-T modification destabilizes duplexes with complementary DNA by −1.4 to −2.0 °C per modification relative to dT in a somewhat sequence dependent context. If flanked by two pyrimidine nucleotides (**ON1**) the destabilization is higher as compared to purines as nearest neighbors (**ON2**). Two consecutive residues lead to less destabilization which is in line with earlier observations on tc-DNA where it was found that the highest $T_m$/modification were observed in fully modified oligonucleotides [37]. Duplexes with RNA as complement are also destabilized albeit to a lesser extent ($\Delta T_m$/mod −0.8 to −2.2 °C). The same sequence dependence as for DNA as




**Scheme 3:** Pathways for elimination of the modified nucleotides during the oxidation step in oligonucleotide assembly.

**Table 2:** Sequence information and analytical data of **ON1–7** as well as $T_m$ data from UV-melting curves (260 nm) in 10 mM $NaH_2PO_4/Na_2HPO_4$, 150 mM NaCl, pH 7.0. Duplex concentration: 1.2 µM.

| | Sequence | Modification t | $T_m$ vs DNA[a] | $T_m$ vs RNA[a] |
|---|---|---|---|---|
| **ON1** | d(GGATGTTCtCGA) |  | 45.3 (−2.0) | 45.6 (−2.2) |
| **ON2** | d(GGAtGTTCtCGA) | | 43.9 (−1.7) | 45.0 (−1.4) |
| **ON3** | d(GGATGttCTCGA) | | 44.4 (−1.4) | 46.2 (−0.8) |
| **ON4** | d(GGATGTTCtCGA) |  | 44.8 (−2.5) | 43.0 (−4.8) |
| **ON5** | d(GGATGTTCtCGA) |  | 45.8 (−1.6) | 46.6 (−1.2) |
| **ON6** | d(GGAtGTTCtCGA) | | 46.8 (−0.3) | 46.8 (−0.5) |
| **ON7** | d(GGATGttCTCGA) | | 47.2 (−0.1) | 49.6 (+0.9) |

[a] $T_m$ of unmodified duplex d(GGATGTTCTCGA): 47.3 °C vs DNA; 47.8 °C vs RNA; $\Delta T_m$ per modification in parenthesis.

complement appears and again, two consecutive modifications are associated with the least depression in $T_m$/modification. Thus, it turns out that bc$^{en}$-DNA prefers RNA over DNA as a complement which is remarkable given that the parent nucleoside adopts a 2'-endo (S-type) sugar conformation and not a 3'endo (N-type) as do modifications that typically prefer RNA as complement (e.g., LNA). This is somewhat similar to observations with the α-L-LNA analogue which also prefers RNA over DNA as complement despite being a DNA mimic [38].

Also the iso-tc-T modification (**ON4**) turns out to destabilize duplexes with complementary DNA and RNA. However, in contrast to the bc$^{en}$-modification, where there is essentially no difference in binding to DNA and RNA, destabilization of DNA as complement is lower ($\Delta T_m$ −2.5 °C) while that of RNA is higher ($\Delta T_m$ −4.8 °C), this despite the fact that the sugar conformations of the monomers (see Figure 2) are virtually identical. The differential behavior therefore has to be attributed to steric effects of the cyclopropyl methylene group on the adjacent 3'-phosphodiester function solely.

## Discussion

The two novel bc-/tc-modifications presented in this work are part of our endeavor to understand the structure/affinity relationship of this particular oligonucleotide molecular platform in more detail. More precisely we aimed with these modifications to learn how subtle structural changes influence not only the backbone torsion angle γ but also control the conformation of the furanose ring which is central for duplex structure and stability. From X-ray analysis of monomers (Table 2) we find that planarizing the C5'–C6'-bond (tc-nucleosides) leads to a *trans* orientation of torsion angle γ and some variability between S- and N-type furanose conformation, compared to the ring-saturated bc-nucleosides which have a stronger preference for S-type furanose conformation and also maintain the preference for the trans orientation of torsion angle γ. On the other hand, planarizing the C6'–C7'-bond, as in **8β** and **11β**, leads to a synclinal torsion angle γ and a consistent 2'-endo furanose conformation.

In order to correlate structural features of the four monomers under discussions with thermal affinity of correspondingly modified oligonucleotides we have summarized the $\Delta T_m$/modification data for RNA and DNA binding of the four modifications within the same sequence context for which data was available (Table 3). From the data it becomes evident that there is no clear correlation between torsion angle γ and affinity. For example bc$^{en}$-T, having γ in the for duplexes natural *gauche* orientation, is more destabilizing than bc-T in which it is clearly in the unnatural *trans* orientation. However, there seems to be

**Table 3:** $\Delta T_m$/modification data for four different bi/tricyclo modifications in one sequence context.

| d(GGATGTTCtCGA) t = | $\Delta T_m$/modification vs DNA [°C] | $\Delta T_m$/modification vs RNA [°C] | furanose pucker | torsion angle γ |
|---|---|---|---|---|
| bc-T[a] | +1.5 | −0.5 | 1'-exo | trans |
| $bc^{en}$-T | −2.0 | −2.2 | 2'-endo | gauche |
| tc-T | −1.6 | −1.2 | 2'-endo/<br>4'-endo | trans<br>trans |
| iso-tc-T | −2.5 | −4.8 | 2'-endo | gauche |

[a]Ref. [27].

some correlation between the furanose pucker and affinity. There is a trend that nucleosides preferring an S-type sugar conformation (bc-T, $bc^{en}$-T, iso-tc-T) prefer a DNA over an RNA complement. In the only nucleoside that shows some 3'-endo (N-type) character (tc-T) this is inverted. Probably the clearest correlation can be made regarding the effect of the cyclopropane ring in iso-tc-T. Compared to $bc^{en}$-T, it becomes clear that the additional $CH_2$ group destabilizes duplexes in an otherwise isostructural scaffold. This is most likely due to unfavorable steric interactions with the 3'-phosphate group. This negative effect is not unexpectedly most pronounced with RNA as a complement (A-type helical structure).

It has to be clearly noted here that an analysis based on single incorporations of bc- or tc-modifications does not necessarily reflect the effect of the same residues in fully modified oligonucleotides. For example, a bc-T residue stabilizes a duplex with complementary DNA in the above sequence context. However, a fully modified bc-oligonucleotide has no stabilizing effect upon binding to a DNA or RNA complement [20]. Along the same lines, a tc-T residue in the above sequence context destabilizes duplexes with both a DNA and an RNA complement. On the other hand fully modified tc-oligonucleotides stabilize duplexes with DNA and RNA by 1–3 °C per modification [37]. It thus appears that every modification of the DNA or RNA backbone with a bc- or tc-residue is associated with an energetic penalty which most likely arises from the local structural perturbation of the backbone at the site of modification. The more homogeneous the backbone becomes, the more dominant is the energetic benefit (or penalty) of the modification.

## Conclusion

We have synthesized the two novel, thymine containing bc-/tc-nucleosides **8** and **11β** and incorporated them into oligodeoxynucleotides. Analysis of the monomers by X-ray spectroscopy clearly show a high degree of similarity in the conformation of the underlying bicyclic scaffold of these two nucleosides. Thermal melting analysis of duplexes shows a destabilization with both DNA and RNA as complements. The destabilization is more expressed with the iso-tc-T unit and is due to steric interactions of the extra-$CH_2$ group of the cyclopropane ring with the adjacent 3'-phosphate unit. A structure/affinity analysis including the known bc-T and tc-T nucleosides suggests that it is less the structural variety of torsion angle γ but more the furanose pucker (2'-endo vs 3'-endo) that governs affinity. Furthermore, from the accumulated set of $T_m$ data available it becomes clear that $\Delta T_m$/modification data from oligonucleotides with single incorporations of members of the bc/tc-DNA family in general do not reflect the affinity profile of the corresponding fully modified oligonucleotides.

## Supporting Information

Experimental procedures and analytical data, including copies of $^1H$, $^{13}C$ and $^{31}P$ NMR spectra (where appropriate) for all new compounds as well as details for oligonucleotide synthesis and thermal melting experiments.

### Supporting Information File 1

Experimental part.
[http://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-10-194-S1.pdf]

## Acknowledgements

A postdoctoral grant to B.D by ISIS Pharmaceuticals, 2855 Gazelle Court, Carlsbad, CA-92010 (USA) is gratefully acknowledged. We thank the group of Chemical Crystallography of the University of Bern (PD Dr. P. Macchi and Dr. J. Hauser) for the X-ray structures and the Swiss National Science Foundation (Requip project 206021_128724) for co-funding of the single crystal X-ray diffractometer at the department of Chemistry and Biochemistry of the University of Bern.

## References

1. Bennett, C. F.; Swayze, E. E. *Annu. Rev. Pharmacol. Toxicol.* **2010,** *50,* 259–293. doi:10.1146/annurev.pharmtox.010909.105654
2. Vacek, M.; Sazani, P.; Kole, R. *Cell. Mol. Life Sci.* **2003,** *60,* 825–833.

3. Rigo, F.; Hua, Y.; Chun, S. J.; Prakash, T. P.; Krainer, A. R.; Bennett, C. F. *Nat. Chem. Biol.* **2012,** *8,* 555–561. doi:10.1038/nchembio.939
4. Lu, Q.-L.; Yokota, T.; Takeda, S.; Garcia, L.; Muntoni, F.; Partridge, T. *Mol. Ther.* **2011,** *19,* 9–15. doi:10.1038/mt.2010.219
5. Wood, M. J. A.; Gait, M. J.; Yin, H. *Brain* **2010,** *133,* 957–972. doi:10.1093/brain/awq002
6. Passini, M. A.; Bu, J.; Richards, A. M.; Kinnecom, C.; Sardi, S. P.; Stanek, L. M.; Hua, Y.; Rigo, F.; Matson, J.; Hung, G.; Kaye, E. M.; Shihabuddin, L. S.; Krainer, A. R.; Bennett, C. F.; Cheng, S. H. *Sci. Transl. Med.* **2011,** *3,* 72ra18. doi:10.1126/scitranslmed.3001777
7. Wheeler, T. M.; Leger, A. J.; Pandey, S. K.; MacLeod, A. R.; Nakamori, M.; Cheng, S. H.; Wentworth, B. M.; Bennett, C. F.; Thornton, C. A. *Nature* **2012,** *488,* 111–115. doi:10.1038/nature11362
8. Stenvang, J.; Petri, A.; Lindow, M.; Obad, S.; Kauppinen, S. *Silence* **2012,** *3,* No. 1. doi:10.1186/1758-907X-3-1
9. Spitzer, S.; Eckstein, F. *Nucleic Acids Res.* **1988,** *16,* 11691–11704. doi:10.1093/nar/16.24.11691
10. Stein, C. A.; Tonkinson, J. L.; Yakubov, L. *Pharmacol. Ther.* **1991,** *52,* 365–384. doi:10.1016/0163-7258(91)90032-H
11. Manoharan, M. *Biochim. Biophys. Acta* **1999,** *1489,* 117–130. doi:10.1016/S0167-4781(99)00138-4
12. Pallan, P. S.; Greene, E. M.; Jicman, P. A.; Pandey, R. K.; Manoharan, M.; Rozners, E.; Egli, M. *Nucleic Acids Res.* **2011,** *39,* 3482–3495. doi:10.1093/nar/gkq1270
13. Kalota, A.; Karabon, L.; Swider, C. R.; Viazovkina, E.; Elzagheid, M.; Damha, M. J.; Gewirtz, A. M. *Nucleic Acids Res.* **2006,** *34,* 451–461. doi:10.1093/nar/gkj455
14. Prakash, T. P.; Kawasaki, A. M.; Wancewicz, E. V.; Shen, L.; Monia, B. P.; Ross, B. S.; Bhat, B.; Manoharan, M. *J. Med. Chem.* **2008,** *51,* 2766–2776. doi:10.1021/jm701537z
15. Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. *Science* **1991,** *254,* 1497–1500. doi:10.1126/science.1962210
16. Summerton, J. E. *Biochim. Biophys. Acta* **1999,** *1489,* 141–158. doi:10.1016/S0167-4781(99)00150-5
17. Veedu, R. N.; Wengel, J. *Chem. Biodiversity* **2010,** *7,* 536–542. doi:10.1002/cbdv.200900343
18. Imanishi, T.; Obika, S. *Chem. Commun.* **2002,** 1653–1659. doi:10.1039/b201557a
19. Hendrix, C.; Rosemeyer, H.; Verheggen, I.; Van Aerschot, A.; Seela, F.; Herdewijn, P. *Chem.–Eur. J.* **1997,** *3,* 110–120. doi:10.1002/chem.19970030118
20. Bolli, M.; Trafelet, H. U.; Leumann, C. *Nucleic Acids Res.* **1996,** *24,* 4660–4667. doi:10.1093/nar/24.23.4660
21. Renneberg, D.; Leumann, C. J. *J. Am. Chem. Soc.* **2002,** *124,* 5993–6002. doi:10.1021/ja025569+
22. Renneberg, D.; Bouliong, E.; Reber, U.; Schümperli, D.; Leumann, C. J. *Nucleic Acids Res.* **2002,** *30,* 2751–2757. doi:10.1093/nar/gkf412
23. Murray, S.; Ittig, D.; Koller, E.; Berdeja, A.; Chappell, A.; Prakash, T. P.; Norrbom, M.; Swayze, E. E.; Leumann, C. J.; Seth, P. P. *Nucleic Acids Res.* **2012,** *40,* 6135–6143. doi:10.1093/nar/gks273
24. Tarköy, M.; Leumann, C. *Angew. Chem., Int. Ed. Engl.* **1993,** *32,* 1432–1434. doi:10.1002/anie.199314321
25. Šilhár, P.; Leumann, C. J. *Bioorg. Med. Chem.* **2010,** *18,* 7786–7793. doi:10.1016/j.bmc.2010.09.064
26. Luisier, S.; Leumann, C. J. *Heterocycles* **2010,** *82,* 775–790. doi:10.3987/COM-10-S(E)65
27. Luisier, S.; Leumann, C. J. *ChemBioChem* **2008,** *9,* 2244–2253. doi:10.1002/cbic.200800322
28. Lietard, J.; Leumann, C. J. *J. Org. Chem.* **2012,** *77,* 4566–4577. doi:10.1021/jo300648u
29. Lietard, J.; Ittig, D.; Leumann, C. J. *Bioorg. Med. Chem.* **2011,** *19,* 5869–5875. doi:10.1016/j.bmc.2011.08.022
30. Dugovic, B.; Leumann, C. J. *J. Org. Chem.* **2014,** *79,* 1271–1279. doi:10.1021/jo402690j
31. Vorbrüggen, H.; Krolikiewicz, K.; Bennua, B. *Chem. Ber.* **1981,** *114,* 1234–1255. doi:10.1002/cber.19811140404
32. Vorbrüggen, H.; Bennua, B. *Chem. Ber.* **1981,** *114,* 1279–1286. doi:10.1002/cber.19811140407
33. Tarköy, M.; Bolli, M.; Schweizer, B.; Leumann, C. *Helv. Chim. Acta* **1993,** *76,* 481–510. doi:10.1002/hlca.19930760132
34. Steffens, R.; Leumann, C. *Helv. Chim. Acta* **1997,** *80,* 2426–2439. doi:10.1002/hlca.19970800812
35. Scheuer-Larsen, C.; Dahl, B. M.; Wengel, J.; Dahl, O. *Tetrahedron Lett.* **1998,** *39,* 8361–8364. doi:10.1016/S0040-4039(98)01844-9
36. Hayakawa, Y.; Wakabayashi, S.; Kato, H.; Noyori, R. *J. Am. Chem. Soc.* **1990,** *112,* 1691–1696. doi:10.1021/ja00161a006
37. Ittig, D.; Gerber, A.-B.; Leumann, C. J. *Nucleic Acids Res.* **2011,** *39,* 373–380. doi:10.1093/nar/gkq733
38. Ellemann Nielsen, K. M.; Petersen, M.; Håkansson, A. E.; Wengel, J.; Jacobsen, J. P. *Chem.–Eur. J.* **2002,** *8,* 3001–3009. doi:10.1002/1521-3765(20020703)8:13<3001::AID-CHEM3001>3.0.CO;2-1


## License and Terms

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

# Synthesis of a bifunctional cytidine derivative and its conjugation to RNA for in vitro selection of a cytidine deaminase ribozyme

Nico Rublack and Sabine Müller*

Full Research Paper

Open Access

Address:
Institut für Biochemie, Ernst-Moritz-Arndt-Universität Greifswald, Felix-Hausdorff-Str. 4, D-17487 Greifswald, Germany

Email:
Sabine Müller* - smueller@uni-greifswald.de

* Corresponding author

Keywords:
cytidine deaminase; modified nucleoside; nucleic acids; ribozyme; RNA world; SELEX

*Beilstein J. Org. Chem.* **2014,** *10,* 1906–1913.
doi:10.3762/bjoc.10.198

Received: 28 May 2014
Accepted: 30 July 2014
Published: 15 August 2014

This article is part of the Thematic Series "Nucleic acid chemistry".

Guest Editor: H.-A. Wagenknecht



## Abstract

Over the past 20 years, the generation of functional RNAs by in vitro selection has become a standard technique. Apart from aptamers for simple binding of defined ligands, also RNAs for catalysis of chemical reactions have been selected. In the latter case, a key step often is the conjugation of one of the two reactants to the library, requiring suitable strategies for terminal or internal RNA functionalization. With the aim of selecting a ribozyme for deamination of cytidine, we have set up a selection scheme involving the attachment of the cytidine acting as deamination substrate to the 3'-terminus of the RNAs in the library, and library immobilization. Here, we report the synthesis of a bifunctional cytidine derivative suitable for conjugation to RNA and linkage of the conjugated library to a streptavidine-coated surface. Successful conjugation of the cytidine derivative to the 3'-terminus of a model RNA is demonstrated.

## Introduction

Since the discovery of the first catalytic RNA in the ciliate *Tetrahymena thermophilia* in 1982 [1], a number of naturally occurring ribozymes have been described [2]. Whereas all of these natural ribozymes accelerate transesterifications or, as in the case of the ribosome, the peptide bond formation, artificial ribonucleic acids developed by in vitro selection have been shown to catalyze a wide variety of organic chemical reactions [3-5]. Moreover, several of these developed ribozymes promote reactions of metabolic relevance in modern life organisms. Impressive examples are RNA catalysts that support an aldol reaction between an aldehyde and a ketone, relevant to the synthesis of sugars [6], or the linkage of ribose to nucleobases to generate nucleotides [7]. Such ribozymes are seen as important functional entities underscoring the RNA world theory, where ribonucleic acids are suggested acting as the carrier of genetic information as well as functional players enabling RNA amplification and processing [8]. Even though a number of ribozymes supporting the RNA world theory are known, there is a wide

field of protein-catalyzed reactions in present-day organisms, for which no ribozyme analogue has yet been found. This applies for example to the transformation of cytidine to uridine, which is a well-known RNA editing event in modern cellular chemistry [9]. This process is catalyzed by cytidine deaminases (CDA, EC 3.5.4) belonging to a family of enzymes found in pro- and eukaryotes [10]. In the active center of all CDAs, a zinc ion is responsible for the activation of a water molecule that acts as nucleophile attacking the C4 carbon center of the cytosine residue and thus facilitating deamination [11]. The development of a ribozyme supporting the same kind of reaction would be a valuable addition to the repertoire of RNA activities with relevance to the RNA world theory.

A very useful and often applied technique for the generation of catalytic RNA structures is a variation of the classical SELEX approach (Systematic Evolution of Ligands by EXponential Enrichment) [12,13]. A typical procedure for the selection of a ribozyme, which enhances the reaction between two substrates, involves the conjunction of either of them with the members of an RNA-library. This is mostly achieved by transcription priming [14], which however, allows attachment of a specific reactant merely to the 5'-terminus of library RNAs [15-17]. Alternatively, post-transcriptional protocols can address both the 5'- and the 3'-terminus. In general, post-transcriptional 5'-modification may be achieved by thiophosphorylation with T4-polynucleotide kinase followed by derivatization of the introduced terminal thiophosphate [18,19], or by chemical conversion of the 5'-terminal primary OH group into an amine or azide to be used for further conjugation with NHS-esters [20] or with alkynes [21]. Alternatively, natural and modified nucleosides can be attached to the 3'-terminus by the use of enzymes like poly(U)-polymerase (PUP), poly(A)-polymerase (PAP) or the terminal deoxynucleotidyl transferase (TdT) [22-25]. However, since most enzymatic techniques achieve only moderate yields, chemical strategies may be advantageous. Chemical 3'-end modification uses the unique properties of the RNA's 3'-terminal *cis*-diol, which can be specifically oxidized with metaperiodate [26-28], followed by reaction of the produced dialdehyde with amines or hydrazines linked to a desired functional entity [29].

We set out to select from a random library catalytically active RNAs that support deamination of cytidine to uridine. For this purpose, a bifunctionalized cytidine derivative (Figure 1) was synthesized. Via its 5'-OH group, the cytidine derivative is linked to a hexaethylene glycol tether bearing a primary amino group. At the C4-position of the base, a short linker connected

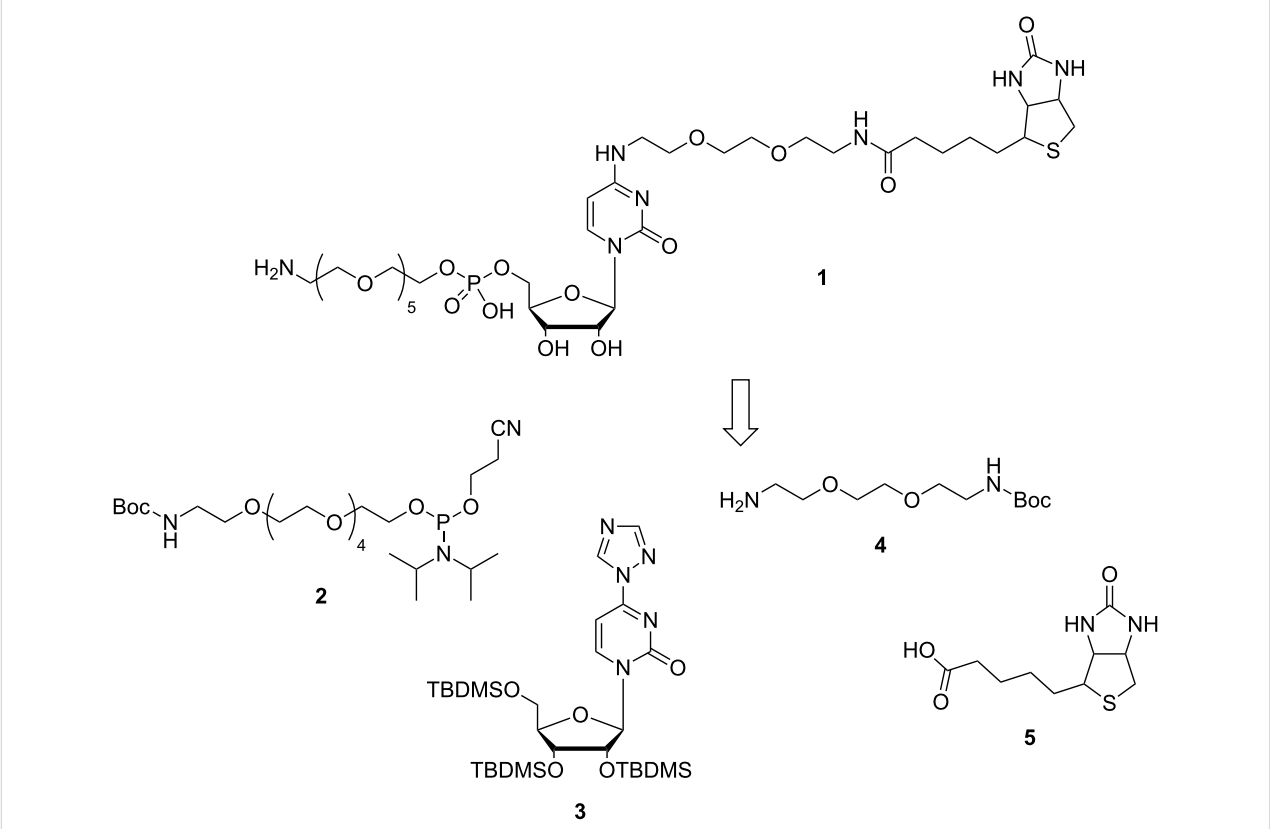



**Figure 1:** Retrosynthetic analysis of the bifunctional cytidine derivative **1** for functionalization of a periodate-oxidized RNA library.

to a biotin was introduced. Upon periodate oxidation of the 3'-*cis* diol of the RNA library molecules, the produced dialdehyde is thought to react with the primary amino group of the cytidine derivative allowing immobilization of the resulting modified RNAs onto a solid surface through biotin–streptavidine interaction. All RNA sequences that can catalyze the desired deamination reaction of cytidine to uridine will be cleaved off and thus released into solution. Upon recovery, beneficial variants can be subjected to the next round of selection. Here we describe the synthetic route to the bifunctionalized cytidine derivative and its successful conjugation to RNA.

## Results and Discussion

Synthesis of the bifunctional cytidine derivative **1** started from uridine making use of four synthons: a protected hexaethylene glycol linker phosphoramidite **2** bearing a primary amine to be used later for RNA functionalization, protected and suitably activated uridine **3**, and a short mono protected diamino linker **4** for the attachment of a biotin moiety **5** in order to immobilize the conjugated RNA molecules onto a streptavidine-coated surface (Figure 1).

First, the sugar hydroxy functions of uridine (**6**) were fully protected with *tert*-butyldimethylsilyl (TBDMS) groups and the tris-silylated uridine **7** [30] was further reacted with $POCl_3$, triazole and triethylamine (Figure 2).

Among the procedures described for conversion of uridine into cytidine derivatives [31-34] we have chosen the path via a triazolyl activated uridine [31], because the triazolyl derivatives are known being stable enough to allow convenient handling and isolation, and they have shown very good reactivity with aliphatic amines. Synthon **3** could be isolated by recrystallization in 76% yield. Next, we introduced a short linker to the $N^4$-position of **3** in order to attach a biotin unit to the modified nucleoside. Linker **4** was synthesized from 2,2'-(ethylenedioxy)diethylamine (**8**) with di-*tert*-butyl dicarbonate ($Boc_2O$) in dioxane [35] (Figure 3).

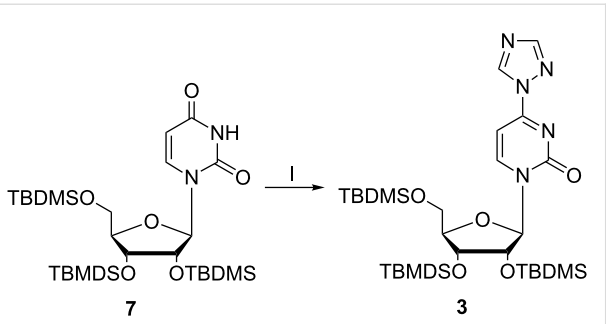



**Figure 2:** Introduction of the triazolyl moiety into the uridine derivative **7** generating synthon **3**. I: 4 equiv $POCl_3$, 16 equiv triazole, 20 equiv $NEt_3$, MeCN, 60 min at 0 °C, 4.5 h at rt, 76%.

The mono protected diamino linker **4** was purified by chromatography on silica gel and subsequently used for reaction with the triazolyl protected uridine derivative **3** to produce compound **9** in 85% yield. Upon removal of the Boc group and sugar deprotection, the remaining aliphatic amine should be used for conjugation to biotin, followed by coupling of the 5'-hydroxy group to a linker phosphoramidite.

In terms of synthesis efficiency, we first tried to simultaneously deprotect the amino functionality at the nucleobase and the sugar 5'-hydroxy group to generate derivative **12**. Thereafter, the greater nucleophilicity of the primary amine over the alcohol should be used for selective amide bond formation with biotin, provided for reaction as *N*-hydroxysuccinimidyl (NHS-) ester, and subsequently the 5'-*O*-phosphoramidite should be prepared. Unfortunately, preparation of **12** in the suggested one step procedure (Figure 4, upper path) was not possible.

According to a previous report by Zhu et al. [36], the 5'-OH group of tris-silylated nucleosides can be selectively removed by treatment with a THF/TFA/$H_2O$ mix (4:1:1, v/v/v) at 0 °C. Since cleavage of Boc-groups requires strong Brønsted acids [37-39], we varied the protocol of Zhu et al. with respect to the acid concentration and to reaction time. However, we could not

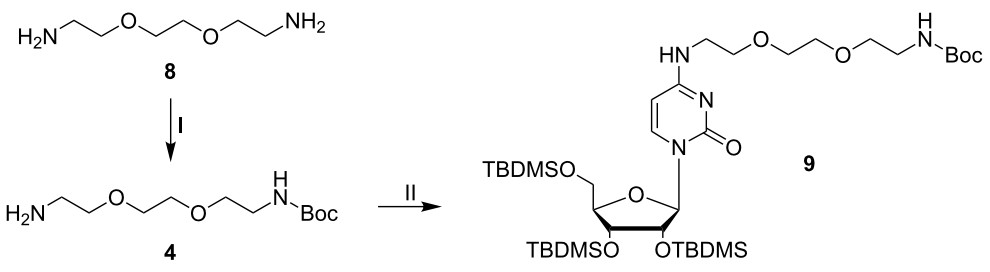



**Figure 3:** Preparation of synthon **4** and substitution of the triazolyl moiety of **3** to form the fully protected cytidine derivative **9**. I: 0.16 equiv $Boc_2O$, dioxane, 3 h, rt, 68% (based on the amount of $Boc_2O$); II: 0.77 equiv triazolyl protected uridine **3**, MeCN, rt, overnight, 85%.

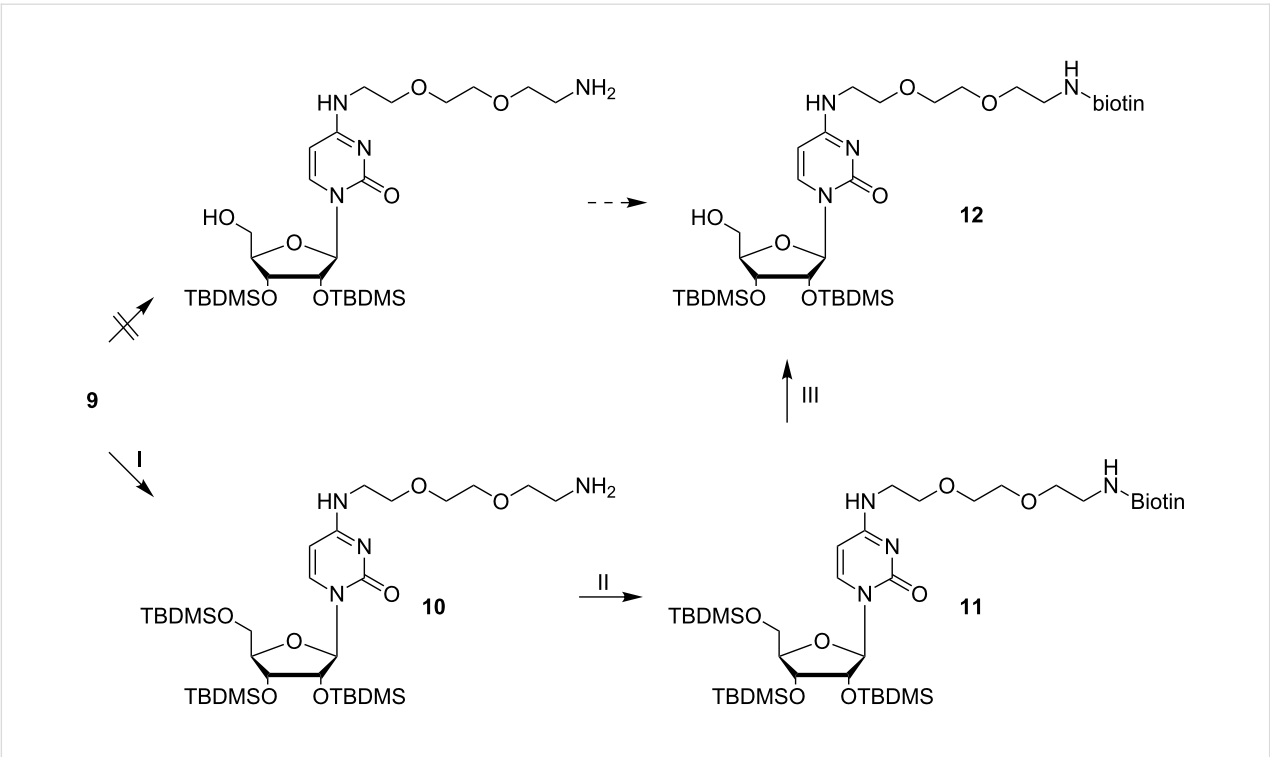



**Figure 4:** Synthesis of 2',3'-bis-*O*-(*tert*-butyldimethylsilyl)-1-[4-(*N*'-biotinyl-3,6-dioxaoctane-1,8-diamine)pyrimidine-2(1*H*)-onyl]-β-D-riboside (**12**). I: 2.6 equiv $ZnBr_2$, DCM, 1 d, rt, Ar, 82%; II: 1.1 equiv EDAC·HCl, 1.1 equiv biotin, DMF, 0 °C → rt, overnight, 65%; III: THF/TFA/$H_2O$ (4:1:1, v/v/v), 0 °C, 5 h, 94%.

identify conditions that led to removal of the Boc group, and leaving the TBDMS groups intact; multiple desilylated products were detected in all cases. These findings forced us to synthesize **12** in an alternative way. Apart from protic acids, the *tert*-butoxycarbonyl moiety of Boc-protected amines can be selectively cleaved with Lewis acids [40]. Thus, stirring of **9** in anhydrous DCM with $ZnBr_2$ delivered the free amine **10** with a yield of 82% (Figure 4). To this, biotin was coupled by in situ activation of the carboxylic acid as an NHS-ester using 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride (EDAC·HCl). Next, the 5'-*O*-TBDMS group was removed by stirring **11** in a mixture of THF/TFA/$H_2O$ (4:1:1, v/v/v) at 0 °C, and solely the 5'-*O*-deprotected cytidine derivative **12** was obtained in 94% yield (Figure 4). The long amino linker **13** was synthesized from hexaethylene glycol with an overall yield of 52% over three steps following a standard literature protocol [41]. After introduction of a Boc group for protection of the linker amino function, the resulting compound **14** was reacted with (2-cyanoethyl-*N*,*N*-diisopropyl)chlorophosphoramidite to yield the phosphoramidite **2**, which was coupled with synthon **12** in anhydrous THF in the presence of 5-benzylmercapto-1*H*-tetrazole (BMT) as activator (Figure 5).

The coupling product **16** was isolated with a yield of 28%. Finally, all remaining protecting groups were cleaved off. The base-labile β-cyanoethyl (CE) group was removed by stirring **16** in a 1:1 mixture of concentrated aqueous ammonia and methylamine. Subsequent treatment of **17** with triethylamine trihydrofluoride (TEA·3HF) at 55 °C for 2 h delivered compound **18**, which was reacted with neat TFA. Each deprotection step was monitored by MALDI mass spectrometry. The bifunctionalized cytidine derivative **1** as the final product was isolated by reversed-phase chromatography with a yield of 83% over the last three steps (Figure 6).

To set-up conditions for conjugation of the bifunctional cytidine derivative **1** to RNA, we synthesized a short model oligonucleotide (GUC AGC CGU CAG GAU CCG UG) corresponding to the 3'-terminal sequence of the envisaged RNA-library. The RNA was oxidized by treatment with sodium periodate [42], and the resulting dialdehyde was reacted with cytidine derivative **1**. The obtained intermediate was reduced with sodium cyanoborohydride to form the more stable amine as final product (Figure 7).

The reaction analysis was carried out by reversed-phase HPLC (Figure 8).

The RP-chromatogram of the crude product (A) shows two peaks (denoted as I) corresponding to the oxidized and

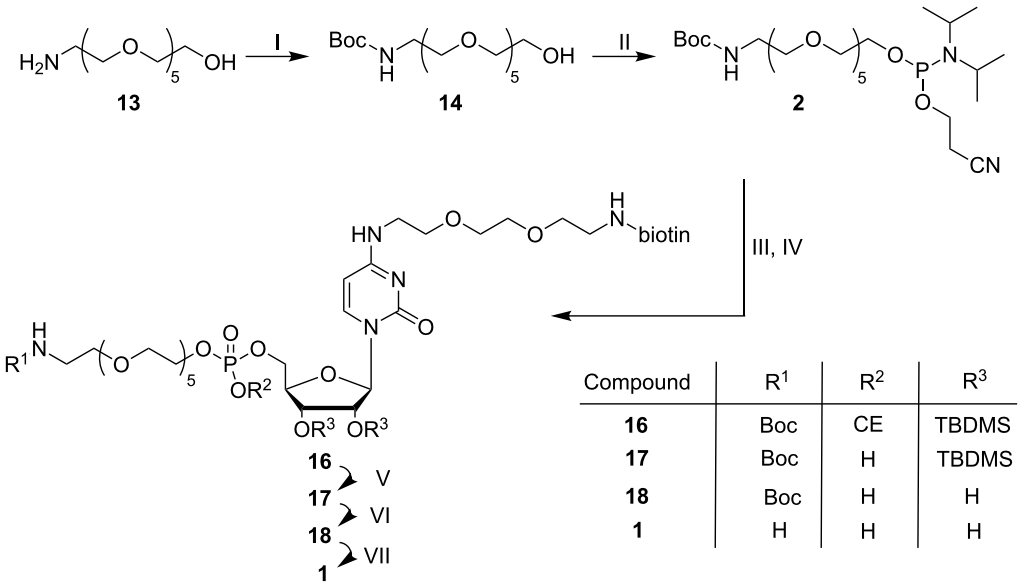



**Figure 5:** Formation of the phosphoramidite **2** from amino alcohol **13**, and subsequent coupling to the 5'-*O*-deprotected nucleoside **12**. I: 1.2 equiv di-*tert*-butyl dicarbonate, EtOAc, 60 °C, 1.5 h, 98%; II: 5.3 equiv DIPEA, 1.3 equiv (2-cyanoethyl-*N*,*N*-diisopropyl)chlorophosphoramidite, DCM, rt, 1.5 h; III: 0.9 equiv **12**, 4.5 equiv BMT, THF, IV: 0.2 M $I_2$ in THF/pyridine/$H_2O$, 10 min, rt, 28% over two steps; V: aq $NH_3$ (30%)/methylamine (8 M), rt, 30 min; VI: TEA·3HF, 55 °C, 2 h; VII: TFA, rt, 2 min; 83 % over three steps.

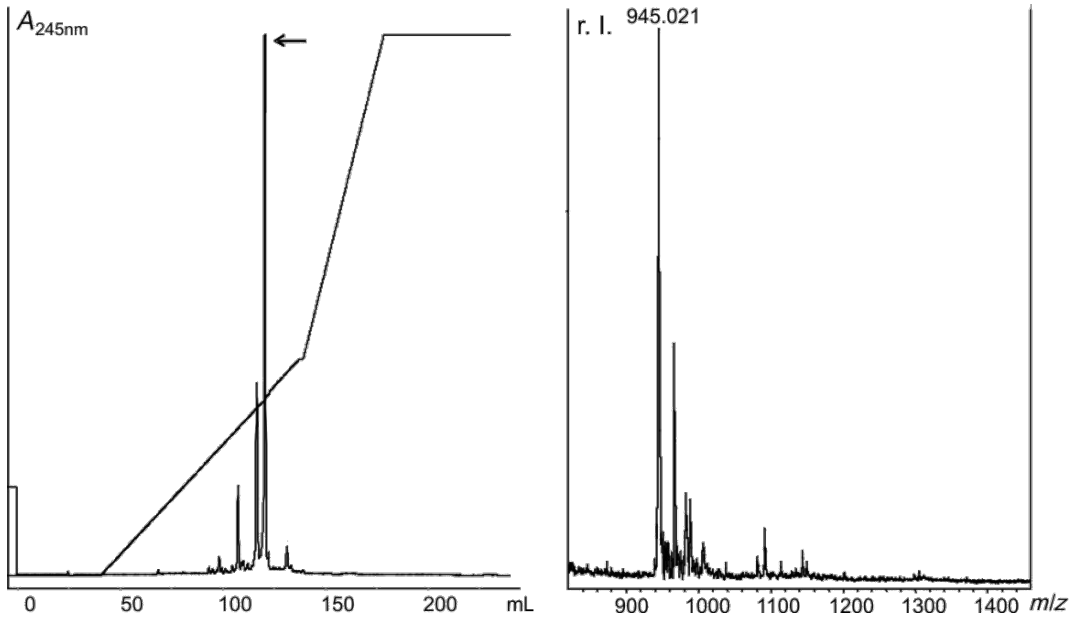



**Figure 6:** A) Reversed-phase HPLC purification of **1** after complete deprotection of **16.** *A* represents the absorption at 254 nm. The fraction indicated by an arrow was collected and analyzed by MALDI mass spectrometry. B) MALDI-spectrum of the pooled collected fractions. The peak at 945.021 (calcd. mass: 944.40 g/mol $[M + H]^+$) corresponds to the cytidine derivative **1**. r.I. represents the relative intensity.

cyanoborohydride reduced RNA (C), peak II corresponding to the cytidine derivative **1** (B), and a new peak (III) indicating the desired RNA–cytidine conjugate, as confirmed by MALDI mass spectrometry (see Supporting Information File 1).

## Conclusion

We successfully synthesized a cytidine derivative for the functionalization of an RNA library following a synthetic route over 16 steps. Two linker moieties of different lengths and function-

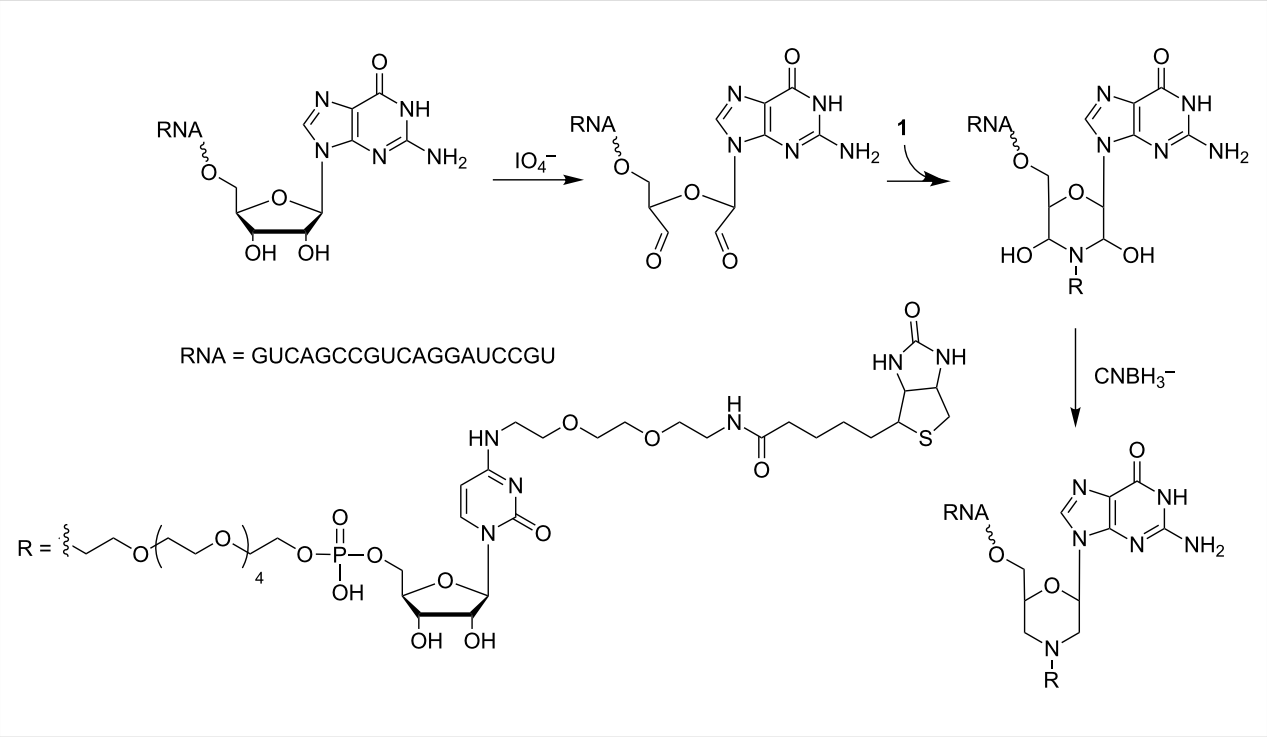



**Figure 7:** Reaction scheme of periodate oxidation of a 20mer model RNA followed by coupling of cytidine derivative **1**.

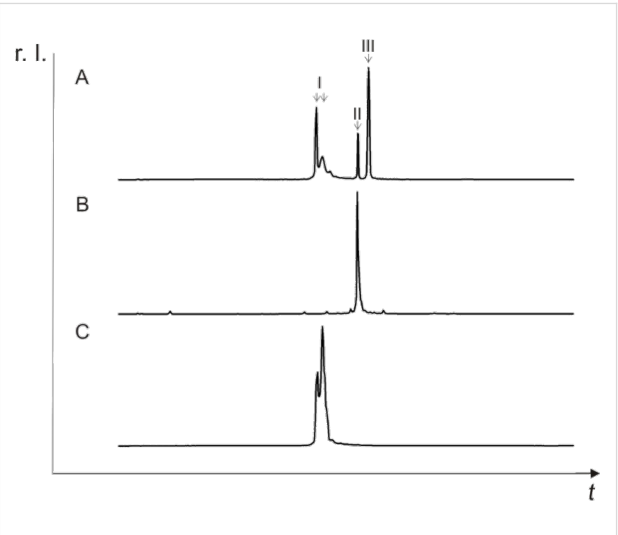



**Figure 8:** Reversed-phase HPLC analysis. A: Crude product of the coupling reaction between the 20mer model RNA and **1.** The peaks I correspond to the periodate oxidized and subsequently reduced RNA. Peak II matches the cytidine derivative **1**. Peak III indicates the RNA–cytidine **1** conjugate. B: Cytidine derivative **1**. C: 20mer model RNA after treatment with sodium periodate and reduction with sodium cyanoborohydride in the absence of cytidine derivative **1**. r. I.: relative intensity.

ality were attached to the nucleobase and to the sugar residue of cytidine, thus enabling for one coupling of the nucleoside to periodate-oxidized RNAs, and for second linking the functionalized library to a surface. The 5'-*O*-coupled linker carrying an aliphatic amino group allows 3'-terminal conjugation of cytidine to the molecules of an RNA library, whereas the biotinylated linker attached to C4 of the nucleobase allows linking the modified library to a streptavidine coated surface. The suitability of the synthesized cytidine derivative was confirmed by successful conjugation to the 3'-terminus of a model RNA, as analyzed by HPLC and MALDI–MS. This set-up is going to be used for selection of a cytidine deaminase ribozyme supporting the conversion of uridine to cytidine. Active molecules will be cleaved from the solid phase and released into solution, such that those can be collected, reverse-transcribed and amplified to enter the next round of selection.

## Experimental

The bifunctional cytidine derivative **1** was synthesized starting from uridine. All reaction steps and the characterization of the obtained products are described in detail in Supporting Information File 1. The RNA GUC AGC CGU CAG GAU CCG UG used as model for studying 3'-terminal conjugation of **1** was synthesized on an Applied Biosystems 394 DNA/RNA Synthesizer following the standard protocol for oligoribonucleotide chain assembly. The synthesized RNA was deprotected using aqueous ammonia (32%)/aq methylamine (40%) (1:1, v/v) at 65 °C for 30 min for removal of base and phosphate protecting groups and cleavage from the support, and TEA·3HF for removal of 2'-*O*-protecting groups, essentially as described previously [43].

## Periodate oxidation

For the oxidation of the RNA's 3'-terminal *cis*-diol, 5 nmol of the 20mer RNA substrate were incubated in a total volume of 1 mL 30 mM sodium metaperiodate, 100 mM NaOAc (pH 5.4) for 1 hour at room temperature in the dark. The oxidized RNA was recovered by ethanol precipitation.

## Conjugation of the cytidine derivative **1** to RNA

The oxidized RNA was dissolved in 250 µL water containing 20 mM imidazole (pH 8), 5 mM $NaCNBH_3$, 1 mM EDTA and 1 mM cytidine derivative **1**. The reaction was carried out at 37 °C. After 2 hours, 25 µL of 50 mM $NaBH_4$ were added, and the reaction mixture was incubated for additional 15 min. Ethanol precipitation yielded the crude coupling product, which was analyzed by reversed-phase HPLC on an Äkta Purifier (Amersham Bioscience). Column: Macherey Nagel EC 250/4 Nucleodur 100-5 C18 ec; Buffers: (A) 0.1 M triethylammonium acetate (pH 7.0), 5% acetonitrile and (B) 0.1 M triethylammonium acetate (pH 7.0), 30% acetonitrile; flow rate 0.5 mL $min^{-1}$; gradient: 0% → 85% (B) in 14 CV, 85% 4 CV, 85% → 100% (B), 100% (B) 4 CV, 100% → 0% (B) 2 CV. The product containing fraction was collected, RNA was lyophilized and analyzed by MALDI–MS: (calcd mass: 7339.9 g/mol, found: 7339.04 $[M + 1]^+$; 3668.515 $[M + 2]^+/2$, Supporting Information File 1).

# Supporting Information

### Supporting Information File 1

Experimental part.
[http://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-10-198-S1.pdf]

# Acknowledgements

The authors wish to thank Dr. Bettina Appel and Julia Drenckhan for RNA synthesis service.

# References

1. Kruger, K.; Grabowski, P. J.; Zaug, A. J.; Sands, J.; Gottschling, D. E.; Cech, T. R. *Cell* **1982,** *31,* 147–157. doi:10.1016/0092-8674(82)90414-7
2. Doudna, J. A.; Cech, T. R. *Nature* **2002,** *418,* 222–228. doi:10.1038/418222a
3. Serganov, A.; Keiper, S.; Malinina, L.; Tereshko, V.; Skripkin, E.; Höbartner, C.; Polonskaia, A.; Phan, A. T.; Wombacher, R.; Micura, R.; Dauter, Z.; Jäschke, A.; Patel, D. J. *Nat. Struct. Mol. Biol.* **2005,** *12,* 218–224. doi:10.1038/nsmb906
4. Sengle, G.; Eisenführ, A.; Arora, P. S.; Nowick, J. S.; Famulok, M. *Chem. Biol.* **2001,** *8,* 459–473. doi:10.1016/S1074-5521(01)00026-6
5. Tsukiji, S.; Ramaswamy, K.; Suga, H. *Pure Appl. Chem.* **2004,** *76,* 1525–1536. doi:10.1351/pac200476071525
6. Fusz, S.; Eisenführ, A.; Srivatsan, S. G.; Heckel, A.; Famulok, M. *Chem. Biol.* **2005,** *12,* 941–950. doi:10.1016/j.chembiol.2005.06.008
7. Unrau, P. J.; Bartel, D. P. *Nature* **1998,** *395,* 260–263. doi:10.1038/26193
8. Gilbert, W. *Nature* **1986,** *319,* 618. doi:10.1038/319618a0
9. Powell, L. M.; Wallis, S. C.; Pease, R. J.; Edwards, Y. H.; Knott, T. J.; Scott, J. *Cell* **1987,** *50,* 831–840. doi:10.1016/0092-8674(87)90510-1
10. Navaratnam, N.; Sarwar, R. *Int. J. Hematol.* **2006,** *83,* 195–200. doi:10.1532/IJH97.06032
11. Carter, C. W., Jr. *Biochimie* **1995,** *77,* 92–98. doi:10.1016/0300-9084(96)88110-7
12. Ellington, A. D.; Szostak, J. W. *Nature* **1990,** *346,* 818–822. doi:10.1038/346818a0
13. Tuerk, C.; Gold, L. *Science* **1990,** *249,* 505–510. doi:10.1126/science.2200121
14. Sampson, J. R.; Uhlenbeck, O. C. *Proc. Natl. Acad. Sci. U. S. A.* **1988,** *85,* 1033–1037. doi:10.1073/pnas.85.4.1033
15. Wolf, J.; Dombos, V.; Appel, B.; Müller, S. *Org. Biomol. Chem.* **2008,** *6,* 899–907. doi:10.1039/b716151d
16. Skipsey, M.; Hack, G.; Hooper, T. A.; Shankey, M. C.; Conway, L. P.; Schröder, M.; Hodgson, D. R. W. *Nucleosides, Nucleotides Nucleic Acids* **2013,** *32,* 670–681. doi:10.1080/15257770.2013.851393
17. Samanta, A.; Krause, A.; Jäschke, A. *Chem. Commun.* **2014,** *50,* 1313–1316. doi:10.1039/c3cc46132g
18. Harris, M. E.; Christian, E. L. *Methods* **1999,** *18,* 51–59. doi:10.1006/meth.1999.0756
19. Burgin, A. B.; Pace, N. R. *EMBO J.* **1990,** *9,* 4111–4118.
20. Davies, M. J.; Shah, A.; Bruce, I. J. *Chem. Soc. Rev.* **2000,** *29,* 97–107. doi:10.1039/a905230e
21. Paredes, E.; Das, S. R. *ChemBioChem* **2011,** *12,* 125–131. doi:10.1002/cbic.201000466
22. Rosemeyer, V.; Laubrock, A.; Seibl, R. *Anal. Biochem.* **1995,** *224,* 446–449. doi:10.1006/abio.1995.1068
23. Igloi, G. L. *Anal. Biochem.* **1996,** *233,* 124–129. doi:10.1006/abio.1996.0016
24. Martin, G.; Keller, W. *RNA* **1998,** *4,* 226–230.
25. Winz, M.-L.; Samanta, A.; Benzinger, D.; Jäschke, A. *Nucleic Acids Res.* **2012,** *40,* e78. doi:10.1093/nar/gks062
26. Malaprade, L. *Bull. Soc. Chim. Fr.* **1928,** *43,* 683–696.
27. House, H. O. *Modern synthetic reactions,* 2nd ed.; W. A. Benjamin Inc.: Menlo Park, California, 1972; pp 353–359.
28. Marinetti, G. V.; Rouser, G. *J. Am. Chem. Soc.* **1955,** *77,* 5345–5349. doi:10.1021/ja01625a042
29. Winston, S. E.; Fuller, S. A.; Evelegh, M. J.; Hurrell, J. G. R. Conjugation of Enzymes to Antibodies. *Current protocols in molecular biology;* John Wiley & Sons, 2001; Vol. 11, 11.1. doi:10.1002/0471142727.mb1101s50
30. Hisamatsu, Y.; Hasada, K.; Amano, F.; Tsubota, Y.; Wasada-Tsutsui, Y.; Shirai, N.; Ikeda, S.-i.; Odashima, K. *Chem. – Eur. J.* **2006,** *12,* 7733–7741. doi:10.1002/chem.200600099
31. Amantea, A.; Henz, M.; Strazewski, P. *Helv. Chim. Acta* **1996,** *79,* 244–254. doi:10.1002/hlca.19960790125
32. Sung, W. L. *Nucleic Acids Res.* **1981,** *9,* 6139–6152. doi:10.1093/nar/9.22.6139
33. Reese, C. B.; Ubazawa, A. *Tetrahedron Lett.* **1980,** *21,* 2265–2268. doi:10.1016/0040-4039(80)80020-7

34. Höbartner, C.; Kreutz, C.; Flecker, E.; Ottenschläger, E.; Pils, W.; Grubmayr, K.; Micura, R. *Monatsh. Chem.* **2003,** *134,* 851–873. doi:10.1007/s00706-003-0592-1
35. Boeijen, A.; Liskamp, R. M. J. *Eur. J. Org. Chem.* **1999,** 2127–2135. doi:10.1002/(SICI)1099-0690(199909)1999:9<2127::AID-EJOC2127>3.0.CO;2-T
36. Zhu, X.-F.; Williams, H. J.; Scott, A. I. *J. Chem. Soc., Perkin Trans. 1* **2000,** 2305–2306. doi:10.1039/b003562i
37. Stahl, G. L.; Walter, R.; Smith, C. W. *J. Org. Chem.* **1978,** *43,* 2285–2286. doi:10.1021/jo00405a045
38. Sakai, N.; Ohfune, Y. *J. Am. Chem. Soc.* **1992,** *114,* 998–1010. doi:10.1021/ja00029a031
39. Shendage, D. M.; Fröhlich, R.; Haufe, G. *Org. Lett.* **2004,** *6,* 3675–3678. doi:10.1021/ol048771l
40. Kaul, R.; Brouillette, Y.; Sajjadi, Z.; Hansford, K. A.; Lubell, W. D. *J. Org. Chem.* **2004,** *69,* 6131–6133. doi:10.1021/jo0491206
41. Miller, J. S.; Béthencourt, M. I.; Hahn, M.; Lee, T. R.; West, J. L. *Biotechnol. Bioeng.* **2006,** *93,* 1060–1068. doi:10.1002/bit.20809
42. Zahler, N. H.; Harris, M. E. In *Handbook of RNA biochemistry;* Hartmann, R. K.; Bindereif, A.; Schön, A.; Westhof, E., Eds.; Wiley-VCH: Weinheim, Germany, 2005; Vol. 1, pp 75–85.
43. Drude, I.; Strahl, A.; Galla, D.; Müller, O.; Müller, S. *FEBS J.* **2011,** *278,* 622–633. doi:10.1111/j.1742-4658.2010.07983.x

## License and Terms

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

# Syntheses of $^{15}$N-labeled pre-queuosine nucleobase derivatives

Jasmin Levic and Ronald Micura*

Full Research Paper

Open Access

Address:
Institute of Organic Chemistry, University of Innsbruck and Center for Molecular Biosciences Innsbruck, Innrain 80–82, 6020 Innsbruck, Austria

Email:
Ronald Micura* - ronald.micura@uibk.ac.at

* Corresponding author

Keywords:
heterocycles; ligands; nucleic acids; nucleobases; nucleosides; pyrrolopyrimidinones

*Beilstein J. Org. Chem.* **2014,** *10,* 1914–1918.
doi:10.3762/bjoc.10.199

Received: 13 April 2014
Accepted: 14 July 2014
Published: 18 August 2014

This article is part of the Thematic Series "Nucleic acid chemistry".

Guest Editor: H.-A. Wagenknecht



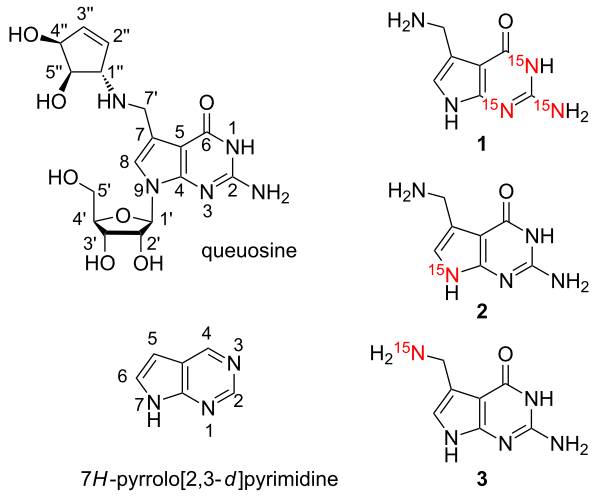
## Abstract

Pre-queuosine or queuine (preQ$_1$) is a guanine derivative that is involved in the biosynthetic pathway of the hypermodified tRNA nucleoside queuosine (Que). The core structure of preQ$_1$ is represented by 7-(aminomethyl)-7-deazaguanine (preQ$_1$ base). Here, we report the synthesis of three preQ$_1$ base derivatives with complementary $^{15}$N-labeling patterns, utilizing [$^{15}$N]-KCN, [$^{15}$N]-phthalimide, and [$^{15}$N$_3$]-guanidine as cost-affordable $^{15}$N sources. Such derivatives are required to explore the binding process of the preQ$_1$ base to RNA targets using advanced NMR spectroscopic methods. PreQ$_1$ base specifically binds to bacterial mRNA domains and thereby regulates genes that are required for queuosine biosynthesis.

## Introduction

The small pyrrolo[2,3-*d*]pyrimidine 7-(aminomethyl)-7-deazaguanine is a natural product, also termed prequeuosine base (preQ$_1$ base) [1,2]. This guanine derivative is involved in the complex biosynthetic pathway of the hypermodified tRNA nucleoside queuosine [3]. Recently, preQ$_1$ base has attracted considerable attention because this nucleobase specifically binds to bacterial mRNA domains and regulates genes that are required for queuosine biosynthesis, by a so-called riboswitch mechanism [4-8]. To explore the binding process of preQ$_1$ base to the RNA and to shed light on the dynamics underpinning this process advanced NMR spectroscopic methods exist for which $^{15}$N-labeled preQ$_1$ base derivatives would be highly beneficial. Here, we report efficient routes for the synthesis of three derivatives with complementary $^{15}$N-labeling patterns (Scheme 1).

The synthesis of preQ$_1$ base has been described first in 1979 by Goto and coworkers from 2-methylthio-6-methoxy-7-methyl-7-deazapurine in 13 steps [9]. Another early, but more efficient procedure was reported by Nishimura in 1988 based on the Mannich reaction using dibenzylamine–formaldehyde and 2-acylaminopyrrolo[2,3-*d*]pyrimidin-4(3*H*)-one, which resulted in the selective introduction of the dibenzylaminomethyl group



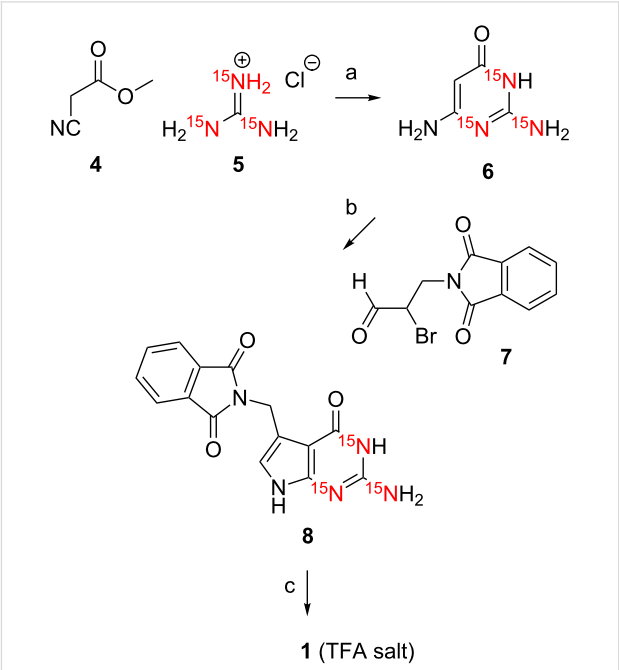

**Scheme 1:** The hypermodified nucleoside queuosine (Q) and the synthetic targets of preQ$_1$ bases **1** to **3** with complementary $^{15}$N labeling patterns for potential NMR spectroscopic applications. Purine and systematic numbering as indicated.

[10]. The following amine exchange reaction of the dibenzylamine function in the Mannich base with ammonia resulted in the preQ$_1$ base. More recently, Carell and coworkers developed a straightforward pathway based on the key reaction of in situ α-brominated 3-phthalimidopropanal with 2,6-diaminopyrimidin-4-one [11], inspired by Grubb's synthesis of the Q base (queuine) [12]. Alternatively, Klebe and coworkers employed a Michael addition of the same pyrimidinone to the nitroolefin 2-[(2*E*)-3-nitroprop-2-en-1-yl]-1*H*-isoindole-1,3(2*H*)-dione [13], however, this route seemed inconvenient for our purposes because access to the nitroolefin requires several additional steps.

## Results and Discussion

Our aim was to develop a robust synthetic pathway to the preQ$_1$ derivatives with the three complementary $^{15}$N labeling patterns depicted in Scheme 1. For this undertaking we considered Carell's synthesis [11] of preQ$_1$ base as a solid foundation that we intended to adapt and modify accordingly, under the premises of efficacy and cost-minimization for $^{15}$N incorporation.

For [$^{15}$N1,$^{15}$N3,H$_2$$^{15}$N(C2)]-7-(aminomethyl)-7-deazaguanine (**1**), we started with the reaction of methyl cyanoacetate (**4**) and [$^{15}$N$_3$]-guanidine hydrochloride (**5**) under basic conditions to give the corresponding [$^{15}$N1,$^{15}$N3,H$_2$$^{15}$N(C2)]-2,6-diaminopyrimidin-4-one (**6**) in high purity after work-up and reversed-phase column chromatography (C18) (Scheme 2) [14]. Then, the α-bromo aldehyde **7** was obtained in two steps from commercially available 3-phthalimidopropan-1-ol that was oxidized using Dess–Martin periodinane. Subsequent in situ bromination of the 3-phthalimidopropan-1-al with CH$_3$SiBr, described previously by others [11,12], did not work reliable in our hands.

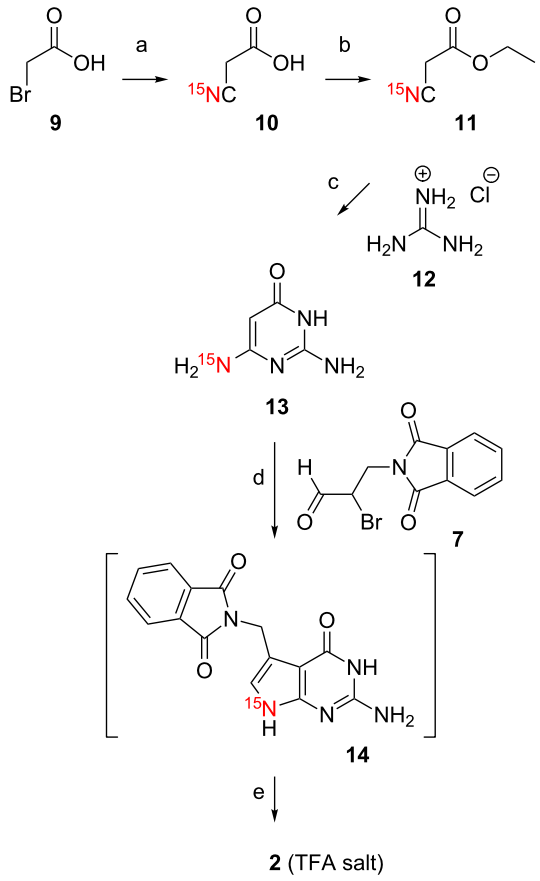



**Scheme 2:** Synthesis of [$^{15}$N1,$^{15}$N3,H$_2$$^{15}$N(C2)]-preQ$_1$ base (**1**). a) CH$_3$ONa (10 equiv) CH$_3$OH, reflux, 10 h, RP C18 chromatography, 35%. b) NaOAc·3H$_2$O (2 equiv), acetonitrile/water, 40 °C, 4 h, 63%. c) H$_2$N-NH$_2$.H$_2$O (10 equiv), ethanol, reflux, 14 h, RP C18 chromatography, 75%. Compound **1** was isolated as salt of trifluoroacetic acid (TFA).

Therefore, in anlogy to Grubb [15] and a more detailed protocol by Yamaguchi [16], we applied 5,5-dibromobarbituric acid [17] to obtain the α-bromo aldehyde **7** which was well stable during purification by column chromatography on SiO$_2$ and isolated in good yields. The pyrrolo[2,3-*d*]pyrimidine ring system of preQ$_1$ base was built in good yields via the cyclocondensation reaction between [$^{15}$N1,$^{15}$N3,H$_2$$^{15}$N(C2)]-2,6-diaminopyrimidin-4-one (**6**) and the 2-bromo-3-phthalimidopropan-1-al (**7**). Finally, deprotection was performed with hydrazine hydrate. The previously published route [11] recommended *N*-Boc functionalization of the preQ$_1$ base in the crude reaction mixture to enable flash chromatography on SiO$_2$ followed by cleavage of the auxiliary function, however, although robust in handling, the yields were rather modest. We therefore decided to directly purify the crude product by reversed-phase column chromatography (HPLC) and obtained compound **1** in excellent yield and purity. We mention that compound **1** was isolated as salt of trifluoroacetic acid (TFA) using 1% TFA in the eluent and it was assumed to exist in 1:1 stoichiometry (preQ$_1$:TFA) based on $^1$H NMR spectra and consideration of p$K_a$ values. However, it is noteworthy that a crystal structure of the preQ$_1$·TFA salt

that was crystallized from saturated aqueous solution showed the co-existence of mono- (N(C'7)) and dications (N(C'7), N3) in the crystal [11].

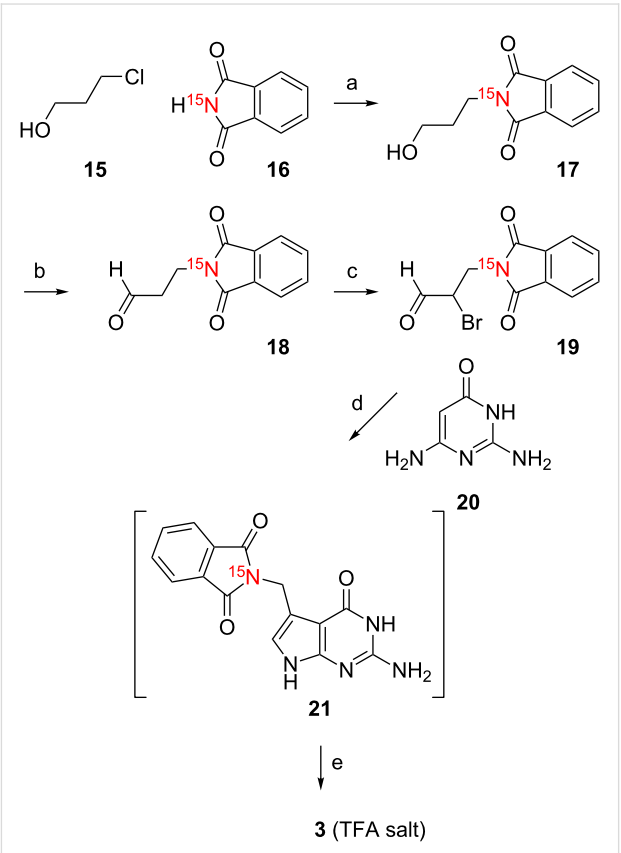
The synthetic track for the [$^{15}N_3$]-preQ$_1$ base (**1**) was designed with the concept in mind to access the complementary $^{15}N$ patterns of [$^{15}$N9]-preQ$_1$ base (**2**) and [H$_2{}^{15}$N(C7')]-preQ$_1$ base (**3**) by employing the same key steps. In this sense, the key intermediate [H$_2{}^{15}$N(C6)]-2,6-diaminopyrimidin-4-one (**13**) for target **2** was accessible by first synthesizing ethyl [$^{15}$N]-2-cyanoacetate (**11**) from 2-bromoacetic acid (**9**) and potassium cyanide [$^{15}$N]-KCN, followed by esterification (Scheme 3).




**Scheme 3:** Synthesis of [$^{15}$N9]-preQ$_1$ base (**2**). a) [$^{15}$N]-KCN (1 equiv), $Na_2CO_3$, $H_2O$, pH 9, 80 °C, 3 h, then room temperature, 20 h, 90%. b) Ethanol (5 equiv), $H_2SO_4$ cat., reflux, 18 h, 92%. c) $CH_3ONa$ (10 equiv) $CH_3OH$, reflux, 10 h, RP C18 chromatography, 40%. d) NaOAc·3$H_2O$ (2 equiv), acetonitrile/water, 40 °C, 4 h, 52%. e) $H_2N$-$NH_2$.$H_2O$ (10 equiv), ethanol, reflux, 14 h, RP C18 chromatography, 92%. Compound **2** was isolated as salt of trifluoroacetic acid (TFA).

All further steps were conducted in direct analogy as described for target **1**, namely reaction with guanidine hydrochloride to furnish compound **13**, followed by cyclocondensation with 2-bromo-3-phthalimidopropan-1-al (**7**) to give the protected [$^{15}$N9]-preQ$_1$ base **14** for subsequent deprotection yielding the desired [$^{15}$N9]-preQ$_1$ base (**2**) (Scheme 3).

Also for the third target, [H$_2{}^{15}$N(C7')]-preQ$_1$ base (**3**), our strategy turn out to be highly convenient. First, we prepared the $^{15}$N-labeled aldehyde **18** as the key intermediate (Scheme 4). This was achieved by reaction of 3-chloropropanol (**15**) with [$^{15}$N]-phthalimide **16** to give [$^{15}$N]-3-phthalimidopropan-1-ol (**17**). All further steps were in direct analogy as described for targets **1** and **2**, namely reaction with 5,5-dibromobarbituric acid [17] to obtain [$^{15}$N]-2-bromo-3-phthalimidopropan-1-al (**19**), followed by cyclocondensation with commercially available 2,6-diaminopyrimidin-4-one (**20**) to give the protected [$^{15}$N(C7')]-preQ$_1$ base **21** for subsequent deprotection yielding the desired [$^{15}$N(C7')]-preQ$_1$ base (**3**) (Scheme 4).




**Scheme 4:** Synthesis of [H$_2{}^{15}$N(C7')] preQ$_1$ base (**3**). a) $K_2CO_3$ (1.5 equiv), DMF, 70 °C, 14 h, 47%. b) Dess–Martin periodinane (1.5 equiv), $CH_2Cl_2$, 3 h, room temperature. c) 5,5-dibromobarbituric acid (0.6 equiv), acetonitrile, reflux, 2 h, 45%. d) NaOAc·3$H_2O$ (2 equiv), acetonitrile/water, 40 °C, 4 h, 58%. e) $H_2N$-$NH_2$.$H_2O$ (10 equiv), ethanol, reflux, 14 h, RP C18 chromatography, 82%. Compound **3** was isolated as salt of trifluoroacetic acid (TFA).

Finally, a direct comparison of $^1$H NMR spectra of the three $^{15}$N labeled preQ$_1$ bases synthesized here is provided in Figure 1.

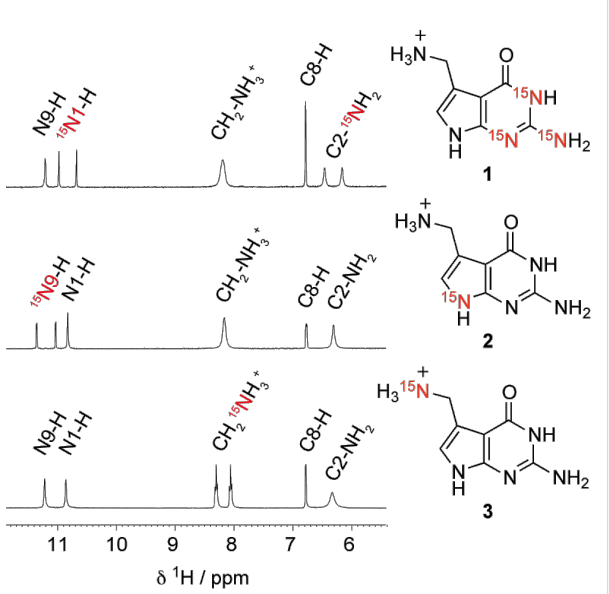



**Figure 1:** Comparison of $^{1}$H NMR spectra of the preQ$_1$ bases **1**, **2** and **3** with complementary $^{15}$N labeling patterns. Conditions: $c_{preQ1\ base}$ = 1 mM; DMSO-$d_6$, 298 K.

## Conclusion

In this short note, an efficient and cost-minimizing route for $^{15}$N labeled preQ$_1$ base derivatives has been described. The synthesis of the pyrrolo[2,3-*d*]pyrimidine ring system is based on the cyclocondensation reaction between α-bromoaldehydes and 2,6-diaminopyrimidin-4-ones and utilizes [$^{15}$N]-KCN, [$^{15}$N]-phthalimide, and [$^{15}$N$_3$]-guanidine for $^{15}$N sources to achieve three complementary labeling patterns that cover all five nitrogen atoms of preQ$_1$ base. The new derivatives carry the potential for modern NMR spectroscopic applications to study the recognition process of these small molecules with RNA aptamer domains from the three preQ$_1$ riboswitch classes known to this date [4,5,18].

## Experimental

**General.**Chemical reagents and solvents were purchased from commercial suppliers (Sigma-Aldrich) and used without further purification. Organic solvents for reactions were dried overnight over freshly activated molecular sieves (4 Å). The reactions were carried out under argon atmosphere. Analytical thin-layer chromatography (TLC) was carried out on Marchery-Nagel Polygram SIL G/UV254 plates. Column chromatography was carried out on silica gel 60 (70–230 mesh). Reversed-phase column chromatography was performed on a GE Healthcare Äktaprime system using a commercial Götec-Labortechnik GmbH 310-25 LiChroprep RP-18 (40–63 µm) column (Merck Lobar compatible). The LC separation was monitored by ultraviolet (UV) detection at 280 nm. Solvent systems were as described below for the individual compounds. $^{1}$H and $^{13}$C NMR spectra were recorded on Bruker DRX 300 MHz and Bruker Avance II+ 600 MHz instruments. The chemical shifts (δ) are reported relative to tetramethylsilane (TMS) and referenced to the residual proton or carbon signal of the deuterated solvent: $CDCl_3$ (7.26 ppm), DMSO-$d_6$ (2.49 ppm), for $^{1}$H NMR spectra; $CDCl_3$ (77.0 ppm) or DMSO-$d_6$ (39.5 ppm) for $^{13}$C NMR spectra. $^{1}$H and $^{13}$C assignments are based on COSY and HSQC experiments. MS experiments were performed on a Bruker 7T FT-ICR instrument with an electrospray ion source. Samples were analyzed in the positive-ion mode.

**[$^{15}$N1,$^{15}$N3,H$_2$$^{15}$N(C2)]-2,6-Diaminopyrimidin-4(3*H*)-one (6).** To a solution of methyl cyanoacetate (**4**, 90 µL, 1.02 mmol) and [$^{15}$N$_3$]-guanidine hydrochloride (**5**, 100 mg, 1.02 mmol) in methanol (6.6 mL) was added dropwise $NaOCH_3$ (0.53 g, 9.85 mmol) in methanol (4.1 mL). After the addition was complete, the mixture was refluxed for 10 hours and allowed to cool to room temperature. The mixture was filtrated, and the filtrate evaporated to dryness. The residue was redissolved in water (1 mL) at 90 °C. The yellow solution was then acidified to pH 6 by acetic acid. The crude product was purified by reversed-phase (C18) column chromatography (eluent A: water, eluent B: acetonitrile; 0–15% B in 40 min, 5 mL/min) as eluent to yield 46 mg of compound **6** (35%) as a yellow solid. $^{1}$H NMR (300 MHz, DMSO-$d_6$) δ 4.41 (s, 1H, CH), 5.82 (s, 2H, $NH_2$), 6.70 (d, $J$ = 89.03 Hz, 2H, $^{15}NH_2$), 7.91 (d, $J$ = 89.51 Hz, 1H, $^{15}$NH) ppm.

**[$^{15}$N1,$^{15}$N3,H$_2$$^{15}$N(C2)]-2-[(2-Amino-4,7-dihydro-4-oxo-1*H*-pyrrolo[2,3-*d*]pyrimidin-5-yl)methyl]-1,3-dihydro-2*H*-isoindole-1,3-dione (8).** Compound **6** (46 mg, 0.37 mmol) and α-bromoaldehyde **7** [15] (100 mg, 0.37 mmol) were suspended in a mixture of acetonitrile and water (1.6 mL; 1:1). Sodium acetate trihydrate (97 mg, 0.72 mmol) was added, and the suspension stirred at 40 °C. After 10 minutes, all solids were disolved and a yellow solution was obtained which rapidly turned into a suspension again, indicating that product **8** started to precipitate. The mixture was stirred for 4 hours, then cooled to room temperature and filtered. The residue was dried under reduced pressure to yield 70 mg of compound **8** (63%) as yellow solid. $^{1}$H NMR (300 MHz, DMSO-$d_6$) δ 4.84 (s, 2H, $CH_2$-N), 6.16 (d, $J$ = 88.96 Hz, 2H, $^{15}NH_2$-C2), 6.36 (s, 1H, H-C8), 7.85–7.87 (m, 4H, arom H), 10.42 (d, $J$ = 89.21 Hz, 1H, $^{15}$N1-H), 10.81 (s, 1H, NH) ppm.

**Trifluoroacetate salt of [$^{15}$N1,$^{15}$N3,H$_2$$^{15}$N(C2)]-7-(aminomethyl)-7-deazaguanine (1).** Compound **8** (70 mg, 0.22 mmol) was added to a solution containing hydrazine hydrate (111 µL, 2.2 mmol) and ethanol (3 mL). The mixture was refluxed overnight and evaporated to dryness. The residue was purified by reversed-phase (C18) column chromatograpy (eluent A: 1%

trifluoroacetic acid in $H_2O$; eluent B: acetonitrile; 0–20% B in 50 min, 4 mL/min) to yield 49 mg of compound **1** (75%; calculated as mono TFA salt) as a light yellow solid. $^1$H NMR (300 MHz, DMSO-$d_6$) δ 4.01 (s, 2H,$CH_2$), 6.30 (d, $J$ = 89.1 Hz, 2H, $^{15}NH_2$-C2), 6.78 (s, 1H, H-C8), 8.18 (s, 3H, $NH_3^+$), 10.82 (d, $J$ = 89.6 Hz, 1H, $^{15}$N1-H), 11.20 (s, 1H, NH) ppm; $^{13}$C NMR (75 MHz, $D_2O$) δ 34.57 ($CH_2NH_3^+$), 98.32 (C7), 110.29 (C5), 113.88 (C2), 118.55 (C8), 141.05 (C4), 150.60 (C6=O), 159.09 ($COO^-$), 162.18 ($CF_3$) ppm; HRMS–ESI $m/z$: $[M + H]^+$ calcd for $C_7H_9N_2{}^{15}N_3O$, 183.07909; found, 183.07865.

**UV spectroscopic analysis of unlabeled 7-(aminomethyl)-7-deazaguanine (preQ$_1$ base):** UV ($H_2O$) $\lambda_{max}$ (ε) = 218 (13630), 258 (7940) nm. For comparison, see UV spectroscopic data and extinction coefficients of 7-deazaguanine, *N*9-methyl-7-deazaguanine and preQ$_1$ nucleoside in references [19-21].

## Supporting Information

$^1$H and $^{13}$C NMR spectra are provided in Supporting Information File 1. Synthetic procedures for the syntheses of compounds **2**, **3**, **10**, **11**, **13**, **14**, **17**–**19**, and **21**.

### Supporting Information File 1

Synthetic procedures and NMR spectra of the most typical compounds.
[http://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-10-199-S1.pdf]

## Acknowledgements

The authors thank Marina Frener (Micura group) for synthetic contributions to 2-bromo-3-phthalimidopropan-1-al (**7**), Kathrin Breuker for FT ICR mass spectrometric analysis, Christoph Kreutz and Tobias Santner for discussions. Funding by the Austrian Science Fund FWF (P21641, I1040) is gratefully acknowledged.

## References

1. Iwata-Reuyl, D. *Bioorg. Chem.* **2003,** *31,* 24–43. doi:10.1016/S0045-2068(02)00513-8
2. Vinayak, M.; Pathak, C. *Biosci. Rep.* **2009,** *30,* 135–148. doi:10.1042/BSR20090057
3. McCarty, R. M.; Bandarian, V. *Bioorg. Chem.* **2012,** *43,* 15–25. doi:10.1016/j.bioorg.2012.01.001
4. Roth, A.; Winkler, W. C.; Regulski, E. E.; Lee, B. W. K.; Lim, J.; Jona, I.; Barrick, J. E.; Ritwik, A.; Kim, J. N.; Welz, R.; Iwata-Reuyl, D.; Breaker, R. R. *Nat. Struct. Mol. Biol.* **2007,** *14,* 308–317. doi:10.1038/nsmb1224
5. Meyer, M. M.; Roth, A.; Chervin, S. M.; Garcia, G. A.; Breaker, R. R. *RNA* **2008,** *14,* 685–695. doi:10.1261/rna.937308
6. Santner, T.; Rieder, U.; Kreutz, C.; Micura, R. *J. Am. Chem. Soc.* **2012,** *134,* 11928–11931. doi:10.1021/ja3049964
7. Soulière, M. F.; Altman, R. B.; Schwarz, V.; Haller, A.; Blanchard, S. C.; Micura, R. *Proc. Natl. Acad. Sci. U. S. A.* **2013,** *110,* E3256–E3264. doi:10.1073/pnas.1304585110
8. Kang, M.; Eichhorn, C. D.; Feigon, J. *Proc. Natl. Acad. Sci. U. S. A.* **2014,** *111,* E663–E671. doi:10.1073/pnas.1400126111
9. Ohgi, T.; Kondo, T.; Goto, T. *Chem. Lett.* **1979,** *8,* 1283–1286. doi:10.1246/cl.1979.1283
10. Akimoto, H.; Imamiya, E.; Hitaka, T.; Nomura, H.; Nishimura, S. *J. Chem. Soc., Perkin Trans. 1* **1988,** 1637–1644. doi:10.1039/p19880001637
11. Klepper, F.; Polborn, K.; Carell, T. *Helv. Chim. Acta* **2005,** *88,* 2610–2616. doi:10.1002/hlca.200590201
12. Barnett, C. J.; Grubb, L. M. *Tetrahedron* **2000,** *56,* 9221–9225. doi:10.1016/S0040-4020(00)00895-4
13. Gerber, H.-D.; Klebe, G. *Org. Biomol. Chem.* **2012,** *10,* 8660–8668. doi:10.1039/c2ob26387d
14. Lolli, M.; Medana, C.; Romagnano, S.; Castoldi, F.; Pozzoli, S.; Vago, F.; Fanelli, R.; Airoldi, L. *J. Labelled Compd. Radiopharm.* **1998,** *41,* 243–252. doi:10.1002/(SICI)1099-1344(199803)41:3<243::AID-JLCR73>3.0.CO;2-H
15. Barnett, C. J.; Grubb, L. M. *Tetrahedron Lett.* **2000,** *41,* 9741–9745. doi:10.1016/S0040-4039(00)01768-8
16. Yamaguchi, H. Thiazolimine compound and oxazolimine compound. Eur. Pat. Appl. EP1640369 A1, March 29, 2006; pp 26–27.
17. Grundke, G.; Keese, W.; Rimpler, M. *Chem. Ber.* **1985,** *118,* 4288–4291. doi:10.1002/cber.19851181037
18. McCown, P. J.; Liang, J. J.; Weinberg, Z.; Breaker, R. R. *Chem. Biol.* **2014,** *21,* 880–889. doi:10.1016/j.chembiol.2014.05.015
19. Seela, F.; Lüpke, U. *Chem. Ber.* **1977,** *110,* 1462–1469. doi:10.1002/cber.19771100428
20. Seela, F.; Kehne, A.; Winkeler, H.-D. *Liebigs Ann. Chem.* **1983,** 137–146. doi:10.1002/jlac.198319830113
21. Kondo, T.; Okamoto, K.; Ohgi, T.; Goto, T. *Tetrahedron* **1986,** *42,* 207–213. doi:10.1016/S0040-4020(01)87419-6

## License and Terms

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

# Pyrrolidine nucleotide analogs with a tunable conformation

Lenka Poštová Slavětínská, Dominik Rejman and Radek Pohl*

Full Research Paper

Open Access

Address:
Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, Flemingovo nám. 2, 166 10 Prague 6, Czech Republic

Email:
Radek Pohl* - pohl@uochb.cas.cz

* Corresponding author

Keywords:
conformation; NMR; nucleic acids; nucleotide analog; phosphonic acid; pseudorotation; pyrrolidine

*Beilstein J. Org. Chem.* **2014,** *10,* 1967–1980.
doi:10.3762/bjoc.10.205

Received: 28 May 2014
Accepted: 05 August 2014
Published: 22 August 2014

This article is part of the Thematic Series "Nucleic acid chemistry".

Guest Editor: H.-A. Wagenknecht



## Abstract

Conformational preferences of the pyrrolidine ring in nucleotide analogs **7–14** were investigated by means of NMR and molecular modeling. The effect of the relative configuration of hydroxy and nucleobase substituents as well as the effect of the alkylation or acylation of the pyrrolidine nitrogen atom on the conformation of the pyrrolidine ring were studied. The results of a conformational analysis show that the alkylation/acylation can be effectively used for tuning the pyrrolidine conformation over the whole pseudorotation cycle.

## Introduction

Nucleotides, nucleosides and nucleobases play an important role in all biological systems. Therefore, it is not surprising that many of their analogs possess interesting biological properties. Potent antiviral drugs based on phosphonate nucleotides **1a–c** [1,2], **2a–d** and **3a–d** (Figure 1) have been reported. Prodrugs of acyclic compounds **1a–c** are currently in clinical use for the treatment of diseases caused by DNA viruses and retroviruses. Cyclic analogs **2** and **3** were reported to exhibit antiviral activity against HIV strains [3]. These examples show that the modification of the sugar-phosphate moiety in nucleotides is a successful approach in developing antiviral therapeutics.

Our long-term interest in the synthesis and evaluation of biological properties of phosphonate azanucleotides has yielded several potent inhibitors of nucleoside/nucleotide metabolizing enzymes: thymine derivatives **4** and **11** (Figure 1 and Figure 2) – inhibitors of thymidine phosphorylase isolated from spontaneous lymphoma of SD rats ($IC_{50}$ = 15 and 11 nM, respectively) [4], guanine derivative **5** – a potent inhibitor of human purine nucleoside phosphorylase PNP ($K_i$ = 10 nM) [5], and finally guanine derivative **6** – exhibiting inhibitory activity against 6-oxopurine phosphoribosyltransferases from *Escherichia coli* [6]. These biologically active analogs contain a five-membered pyrrolidine ring, whose conformation has not been explored so far.

It is known that the conformation of a five-membered ribose or deoxyribose ring plays an important role in the nucleoside,

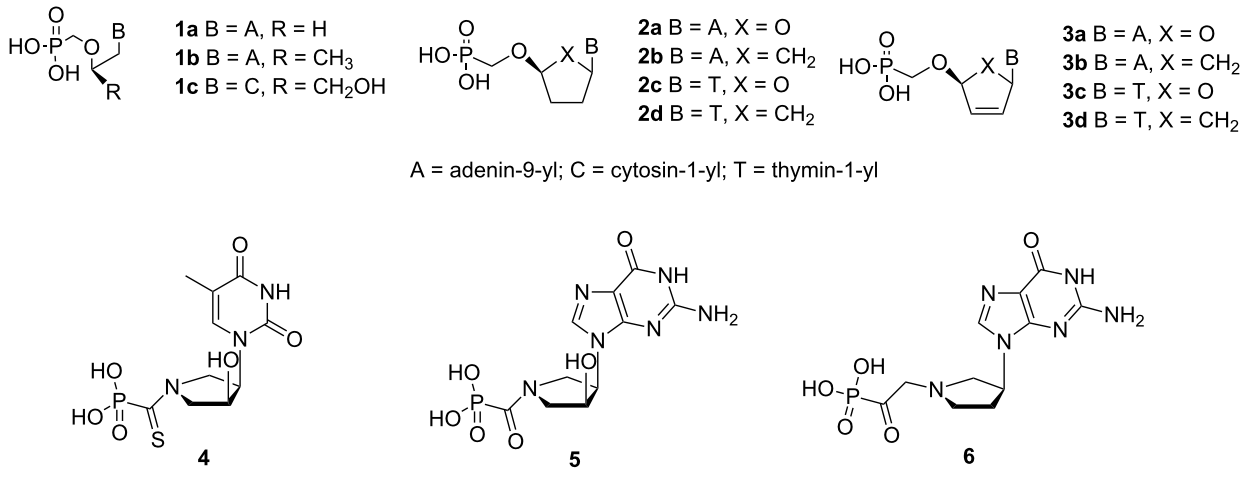



**Figure 1:** Examples of biologically active acyclic and cyclic nucleotide analogs.

nucleotide and oligonucleotide spatial structure. The conformation predefines, for example, the stability of DNA:RNA duplexes – complexes between oligodeoxyribonucleotide and a complementary RNA strain [7] or an overall shape of oligonucleotide. It is also well known that the structure of DNA may alternate between A, B and Z forms depending on the hydration and the type and concentration of metal ions, which is also reflected in the conformation of the sugar part. Finally, the conformation of ribose or its mimics also significantly predetermines the binding of a nucleoside, nucleotide or an analog thereof to an active site of a particular enzyme. The knowledge of the conformation of nucleosides, nucleotides and their analogs is therefore essential for the understanding or even prediction of their biological properties.

There are several approaches providing information on the conformation of a five-membered ring. Conformational analysis using vicinal proton–proton scalar couplings ($^3J_{HH}$) of ring protons obtained from $^1$H NMR spectra is a traditional and well-established experimental method. $^3J_{HH}$ encodes information on the exocyclic dihedral angle between coupled protons based on the Karplus relationship [8]. Exocyclic dihedral angles are in turn in direct relation with endocyclic dihedral angles between ring atoms defining the conformation of the five-membered ring based on the pseudorotation concept [9]. This straightforward approach is complicated by the presence of more than one rapidly interconverting conformer in solution, resulting in the observation of averaged experimental $^3J_{HH}$. Nevertheless, even these averaged $^3J_{HH}$ can be used for conformational analysis by means of the program PSEUROT [10], which minimizes differences between experimental and calculated $^3J_{HH}$, assuming a two-state equilibrium of conformers. The output of the PSEUROT provides pseudorotation parameters: phase angles ($P$), the maximum puckering amplitudes ($\Phi_{max}$) and the relative amounts of individual conformers. The original PSEUROT program was later complemented with the program MULDER [11]. Recently, the Matlab Pseudorotation GUI version of the program has been developed, enabling the creation of conformational maps [12].

The conformational analysis using PSEUROT-based programs was originally designed for ribose or deoxyribose, but in general it can be used for any saturated five-membered ring. The program PSEUROT was, for example, used in a conformational analysis of the prolyl ring in aminoethylprolyl peptide nucleic acid monomers [13], or in a conformational analysis of 2'-fluoro-4'-thioarabinothymidine [14]. The performance of the Matlab Pseudorotation GUI program was tested on two 4'-thio-2'deoxynucleoside analogs [12].

In this publication, we present a conformational analysis of pyrrolidine azanucleotide analogs **7–14** containing thymine and adenine as examples of pyrimidine and purine nucleobases, respectively (Figure 2), and show how the conformation is affected by the mode of the phosphonate moiety attachment to the pyrrolidine ring. Particularly, we show that the alkylation or acylation (an amine or an amide bond formation) of the pyrrolidine nitrogen atom can be used for a simple but effective tuning of pyrrolidine-ring conformation in compounds **7–14**.

# Results and Discussion

## Chemistry

Phosphonomethyl derivatives **7–10** were prepared from pyrrolidine azanucleosides **15a–d** [5,15] via a Mannich-type reaction with diisopropyl phosphite and aqueous formaldehyde at elevated temperature in good yields (Scheme 1). The obtained

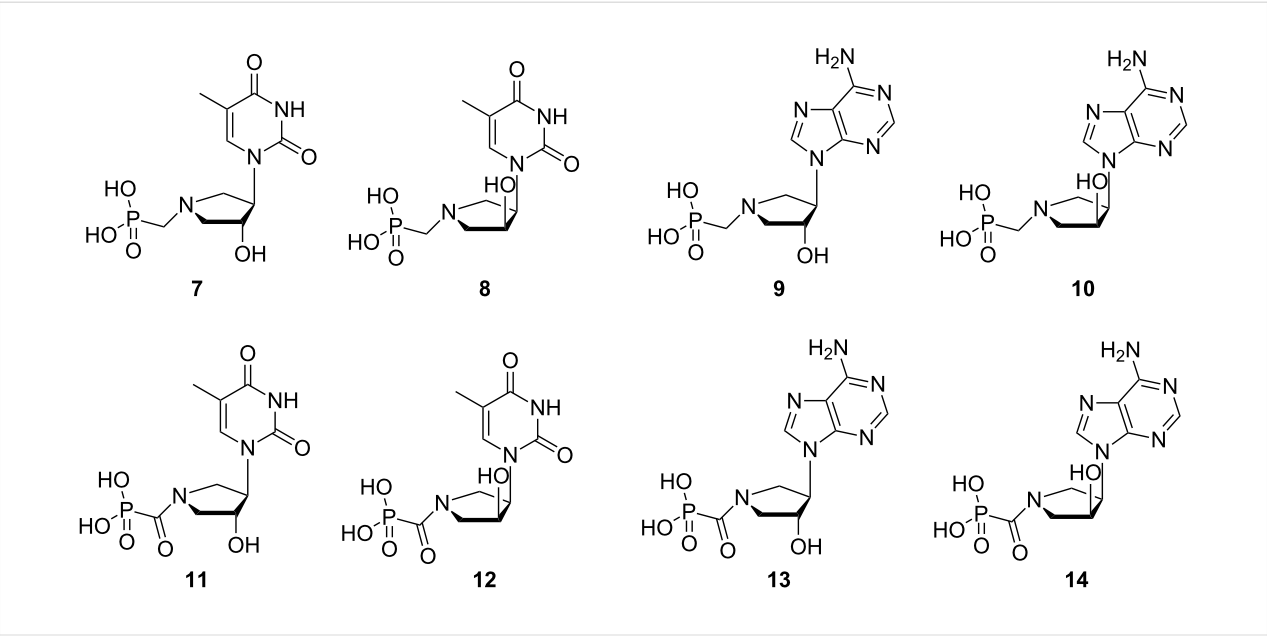
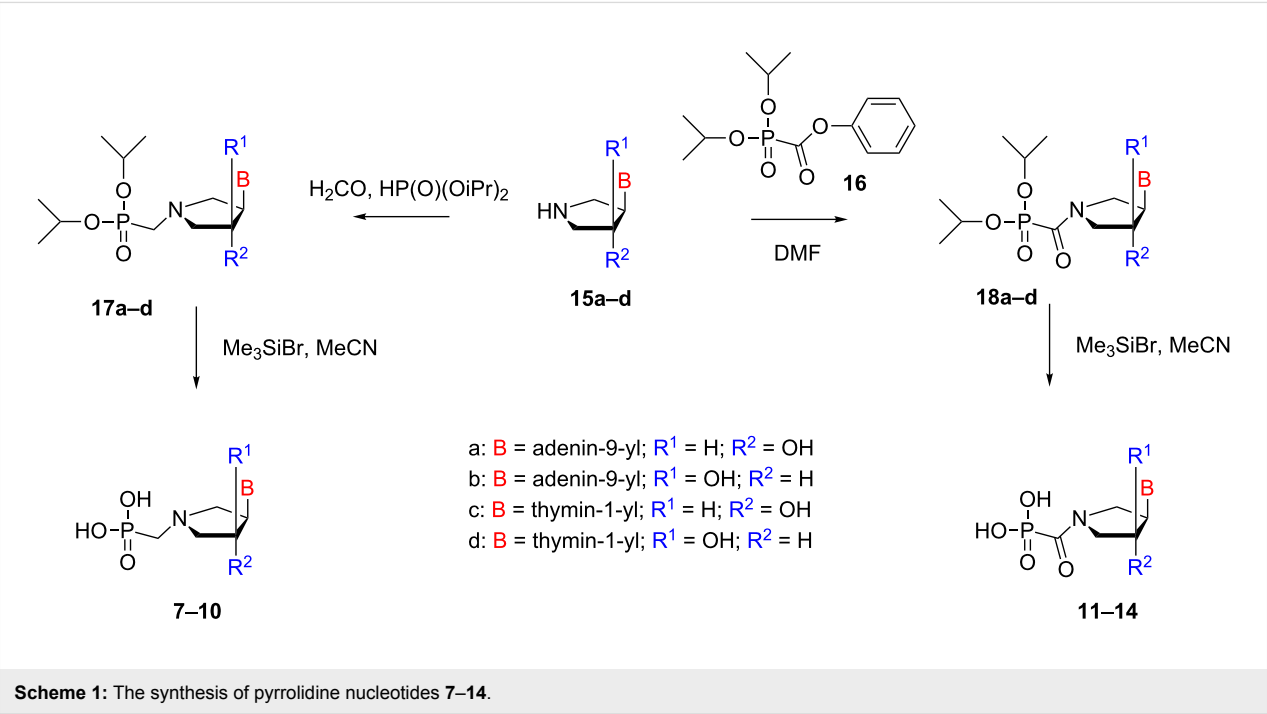



**Figure 2:** The pyrrolidine nucleotide analogs investigated in this study.




**Scheme 1:** The synthesis of pyrrolidine nucleotides **7–14**.

diisopropyl esters **17a–d** were deprotected by treatment with trimethylsilyl bromide in acetonitrile at room temperature.

Phosphonoformyl derivatives **11–14** were prepared from **15a–d** by the reaction with diisopropyl phenylphosphonoformate (**16**) at elevated temperature in good yields. The obtained diisopropyl esters **18a–d** were deprotected by treatment with trimethylsilyl bromide in acetonitrile at room temperature.

## NMR and conformational analysis

Phosphonomethyl **7–10** and phosphonoformyl **11–14** analogs were characterized by $^{1}$H, $^{13}$C and $^{31}$P NMR. The NMR spectra were acquired using 50 mM solutions of the studied compounds in $D_2O$ and all resonances were assigned based on H,H-COSY, H,H-ROESY, H,C-HSQC and H,C-HMBC experiments (for a complete signal assignment and copies of NMR spectra, see Supporting Information File 1). The numbering of

the pyrrolidine ring, the nucleobase and the endocyclic phase angles $\Phi_0$–$\Phi_4$ for the purposes of the conformational analysis is shown in Figure 3. There are fundamental differences in the NMR spectra of phosphonomethyl and phosphonoformyl derivatives, which will be commented on in detail in the following paragraphs.

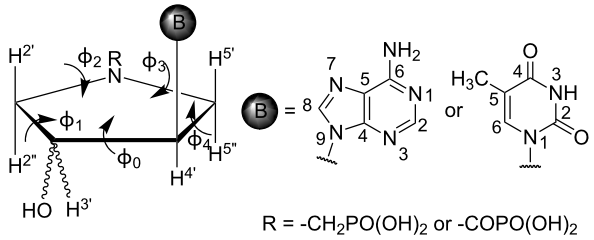



**Figure 3:** The numbering of the pyrrolidine ring, the nucleobase and the endocyclic phase angles for the purposes of the conformational analysis.

## NMR study of phosphonomethyl derivatives

Phosphonomethyl derivatives **7–10** contain both basic (pyrrolidine) and acidic (phosphonic acid) functionalities. Therefore, several protonation/deprotonation transitions are expected to take place when scanning a pH scale. In order to determine these transitions, we measured NMR spectra of **7–10** at different pD values using $D_2O$ solutions of DCl or NaOD for decreasing or increasing pD, respectively. We found that the NMR spectra of **7–10** are strongly pD-dependent as demonstrated for example on the $^1$H NMR spectrum of adenine derivative **9** (Figure 4).

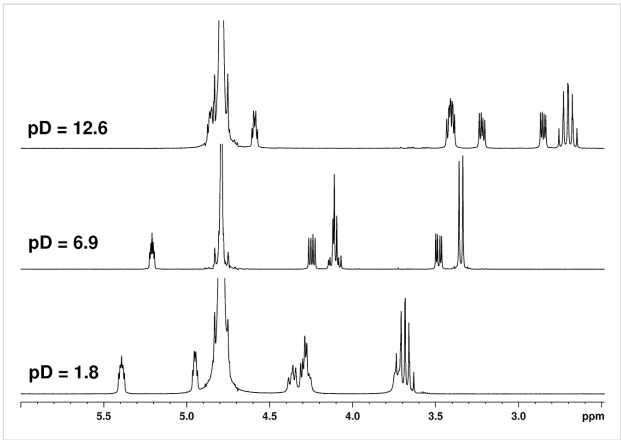



**Figure 4:** The aliphatic part (pyrrolidine protons) of the $^1$H NMR spectra of **9** measured in $D_2O$ at different pD values.

As follows from pD titrations (Figure 5; for more details see Supporting Information File 1), phosphonomethyl analogs **7–10**, both *cis* (**8**, **10**) and *trans* (**7**, **9**), exist at a very low pD value (<3) as free phosphonic acids with a deuterated pyrrolidine nitrogen atom (Figure 6). The phosphonic acid moiety goes through a two-stage deuteration/dedeuteration transition at pD ~ 3 and 5 and the pyrrolidine nitrogen atom stays deuterated until pD ~ 9. This suggests that the compounds exist as zwitterions with a hydrogen bond between the negatively charged phosphonate moiety and the positively charged pyrrolidine nitrogen atom in the range of pD 3–9. The pD titration experiments have also revealed the deuteration/dedeuteration of adenine and thymine nucleobases. In agreement with the data reported [16,17], adenine becomes deuterated predominantly on

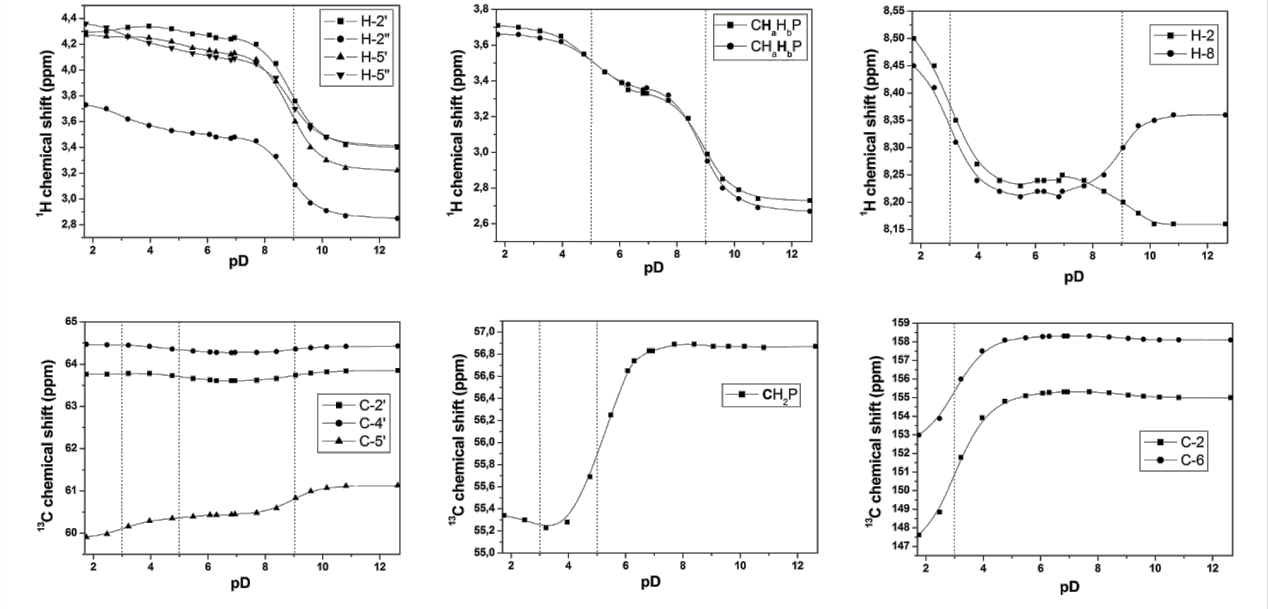



**Figure 5:** Changes of selected $^1$H and $^{13}$C chemical shifts of **9** upon pD change.

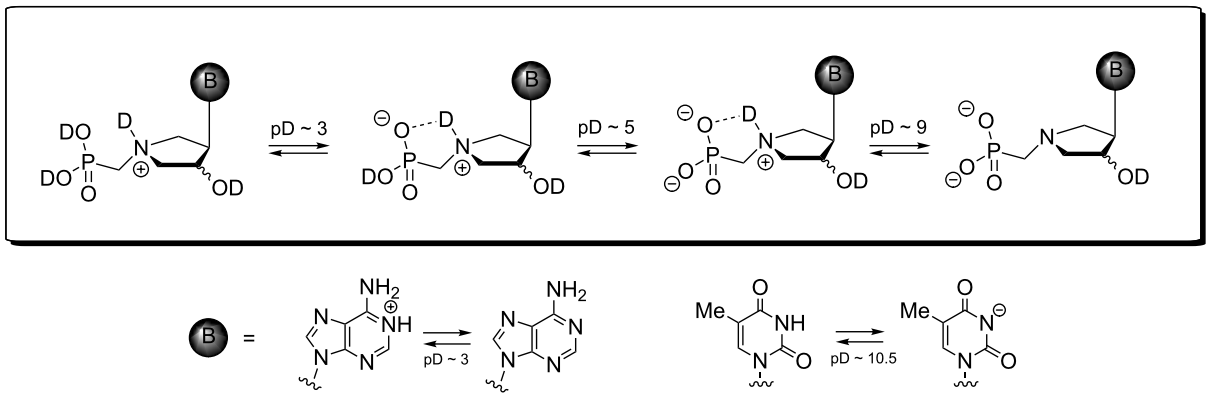



**Figure 6:** The deuteration equilibria of phosphonomethyl derivatives **7–10**.

N-1 at pD ~ 3 while thymine releases its deuterium at N-3 at pD ~ 10.5 and becomes negatively charged. Typically, a pD value of 5.7–6.9 was reached when 50 mM solutions of **7–10** in $D_2O$ were prepared for the NMR measurement, which indicates that isolated compounds are in the form of the monosodium monozwitterionic salt.

The orientation of the phosphonomethyl moiety relative to the nucleobase in the zwitterionic form could be theoretically deduced from the H,H-ROESY spectra. In real spectra, however, protons from $CH_2PO(OH)_2$ provide NOE contacts to protons on both sides of the pyrrolidine ring, suggesting an exchange of deuterium connected to the pyrrolidine nitrogen.

## NMR study of phosphonoformyl derivatives

In contrast to **7–10**, an amidic pyrrolidine nitrogen atom in phosphonoformyl derivatives **11–14** is not involved in protonation/deprotonation transitions. Only negligible changes of $^1$H chemical shifts were observed when the pD value of the solution was changed as demonstrated in the example of compound **13** (Figure 7).

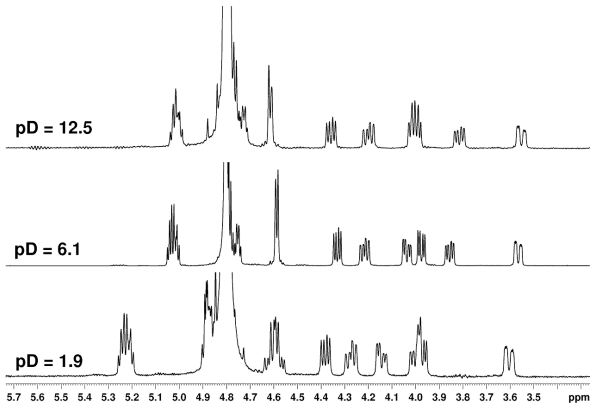



**Figure 7:** The aliphatic part (pyrrolidine protons) of the $^1$H NMR spectra of **13** measured in $D_2O$ at different pD values.

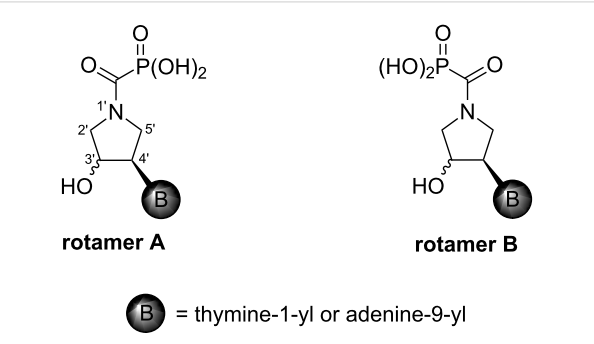



**Figure 8:** Amide rotamers of phosphonoformyl derivatives **11–14**.

Indeed, the reason is the delocalization of the lone electron pair on the pyrrolidine nitrogen atom by resonance. This brings, in addition to the lack of basicity, the existence of two amide rotamers A and B (Figure 8), which can be observed in NMR spectra as two sets of signals. The ratio of A:B is about 1:1 with a slight excess of rotamer B (see the Experimental and Supporting Information File 1).

The existence of the two rotamers prompted us to estimate the energy barrier of their interconversion. We therefore performed variable $^{31}$P NMR measurements of **14** at different temperatures (starting at 25 °C with a 10 °C step, with the last measurement taken at 100 °C) and a line-shape analysis of an uncoupled, exchanging two-site $^{31}$P system (Figure 9). This enabled us to obtain the rate constants of the exchange at different temperatures and estimate the activation parameters of the interconversion such as the Gibbs free energy of activation $\Delta G^{\ddagger}_{298}$ = 80.7 kJ/mol; the enthalpy of activation $\Delta H^{\ddagger}$ = 23.8 kJ/mol; and the entropy of activation $\Delta S^{\ddagger}$ = −190.6 J/(K·mol) (for details,

see Supporting Information File 1). The relatively low value of $\Delta G^{\ddagger}_{298}$ = 80.7 kJ/mol implies that the preparative isolation of the individual rotamers A and B at room temperature would not be possible.

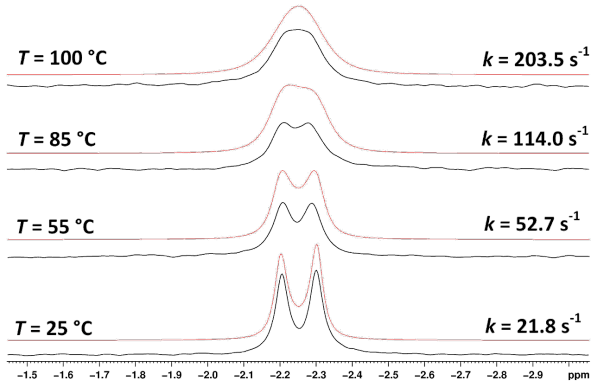



**Figure 9:** The $^{31}$P NMR spectra (202.3 MHz) of **14** measured (the black curve) and simulated (the red curve) at various temperatures.

For the purposes of conformational analysis, it was necessary to assign correctly the NMR signals of the individual conformers. We noticed that only one of the C-2' or C-5' signals in the particular rotamer is split in the $^{13}$C NMR spectrum due to $^{3}J$(C,P) spin–spin interaction. In harmony with the general dependence of the vicinal coupling constant on the dihedral angle according to the well-known Karplus relationship, one would expect the $^{3}J$(C,P) of the carbon in the *trans* arrangement to the phosphorus atom to evince larger splitting than the other arranged in the *cis* arrangement. Thus, $^{3}J$(C2',P) > $^{3}J$(C5',P) in rotamer A and $^{3}J$(C5',P) > $^{3}J$(C2',P) in rotamer B should be observed. In order to support this assumption, we calculated $^{3}J$(C,P) of the most stable conformers (see the chapter Conformational analysis) of the adenine derivatives **13** and **14** for both A and B rotamers using the DFT B3LYP/6-31++G* method. The results summarized in Table 1 show a good agreement of the calculated and experimental values, confirming correct assignment of the rotamers (Figure 10).

**Table 1:** The comparison of the observed and calculated (in parentheses) $^{3}J$(C2',P) and $^{3}J$(C5',P) for rotamers **13A,B** and **14A,B**.

| Rotamer | $^{3}J$(C2',P) | $^{3}J$(C5',P) |
|---|---|---|
| **13A** | 4.6 (3.8) | 0 (−0.8) |
| **13B** | 0 (−0.3) | 4.9 (4.8) |
| **14A** | 4.5 (4.2) | 0 (−0.8) |
| **14B** | 0 (−0.3) | 4.9 (4.8) |

## Conformational analysis

The geometry of the pyrrolidine ring in the compounds studied can occupy various envelope (E) and twisted (T) conformations as depicted in Figure 11. The particular conformation is

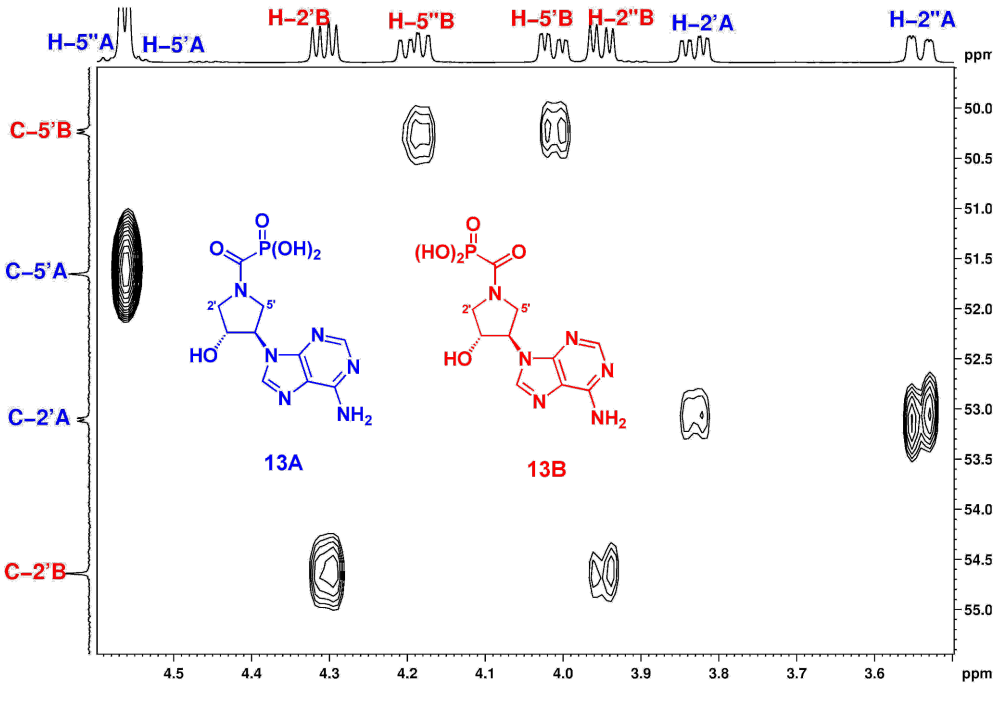



**Figure 10:** A part of the H,C-HSQC spectrum of derivative **13**, showing the assignment of rotamers A and B.

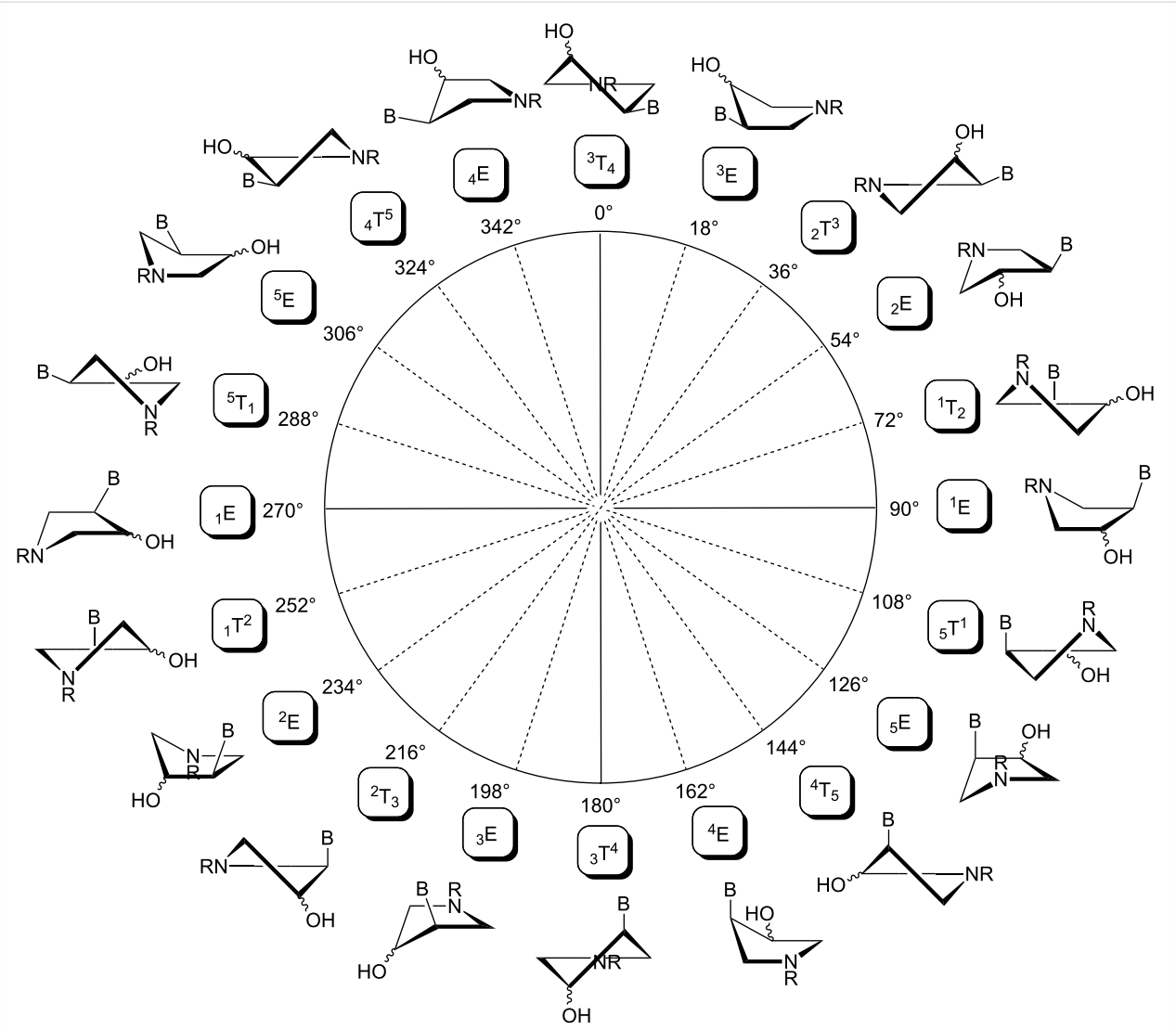



**Figure 11:** The pseudorotation pathway of the pyrrolidine ring in the compounds studied. The sign B stands for a nucleobase (thymin-1-yl or adenin-9-yl) and R means a phosphonomethyl or phosphonoformyl moiety.

described by two pseudorotation parameters: by the phase angle ($P$) and by the maximum puckering amplitude ($\Phi_{max}$) [9]. The phase angle is a periodic variable indicating which ring atoms are situated outside the ring plane and can reach 0°–360°. The maximum puckering amplitude describes the degree of distortion of the five-membered ring out of the plane and its value is usually in the range of 35°–45°.

Conformational analysis using experimental vicinal $^{1}H–^{1}H$ scalar coupling constants ($^{3}J_{HH}$) together with PSEUROT-based programs is aimed at fitting the calculated $^{3}J_{HH}$ to experimental values assuming a two-state conformation equilibrium. In this study, we used the Matlab Pseudorotation GUI program [12] to perform this task. The program requires several input parameters that include experimental $^{3}J_{HH}$ and constants A and B, describing the relation between exocyclic $\Phi_{exo}$ (related to $^{3}J_{HH}$) and endocyclic $\Phi_{endo}$ (related to $P$ and $\Phi_{max}$) dihedral angles ($\Phi_{exo} = A\Phi_{endo} + B$).

The values of $^{3}J_{HH}$ (Table 2) were extracted from the $^{1}H$ NMR spectra, in which the signals were stereospecifically assigned based on ROESY cross-peaks as demonstrated on the example of compound **14** in Figure 12.

The parameters A and B for the five-membered ribose and deoxyribose ring had been traditionally obtained from X-ray data [18]. In the case of a five-membered ring for which X-ray data are not available, DFT-optimized geometries of a number of conformations are used [19]. This molecular modeling approach was also employed in the case of the studied com-

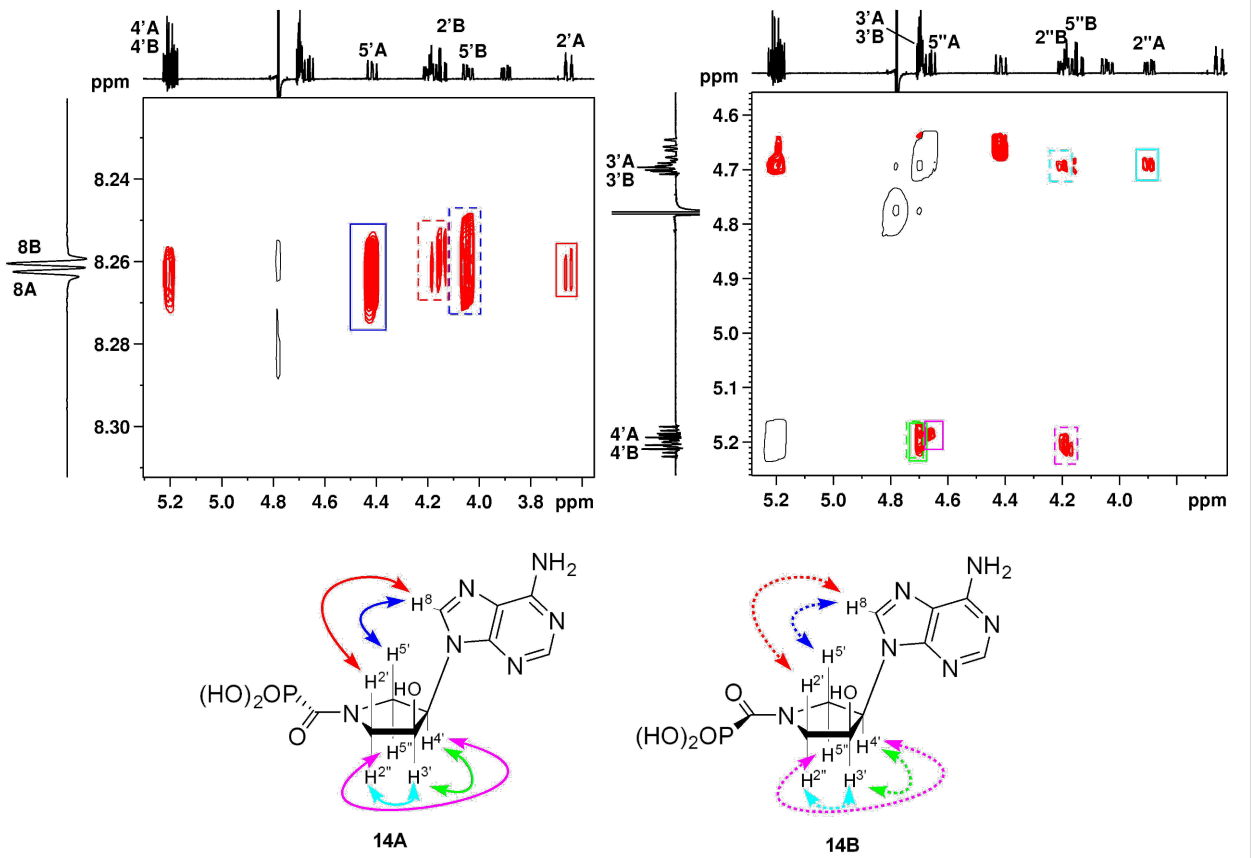



**Figure 12:** An example of the stereospecific assignment of pyrrolidine-ring protons of **14** in the H,H-ROESY spectrum.

**Table 2:** The $^3J_{HH}$ of the pyrrolidine-ring protons used in the conformational analysis.

| | 2',3' | 2",3' | 3',4' | 4',5' | 4',5" |
|---|---|---|---|---|---|
| **7** | 6.9 | 6.4 | 4.1 | 5.6 | 9.1 |
| **8** | 2.0 | 5.2 | 6.0 | 6.9 | 9.4 |
| **9** | 6.5 | 5.1 | 3.0 | 4.4 | 7.3 |
| **10** | 3.5 | 6.1 | 6.2 | 5.0 | 7.9 |
| **11A** | 6.8 | 4.8 | 5.4 | 6.2 | 7.7 |
| **11B** | 6.1 | 5.7 | 5.5 | 6.1 | 8.5 |
| **12A** | 2.1 | 5.2 | 4.5 | 9.5 | 8.1 |
| **12B** | 3.3 | 4.8 | 4.6 | 8.5 | 8.5 |
| **13A** | 6.1 | 4.1 | 4.5 | 5.6 | 6.3 |
| **13B** | 5.8 | 5.2 | 5.3 | 5.5 | 7.8 |
| **14A** | 2.1 | 5.1 | 4.2 | 9.2 | 7.6 |
| **14B** | 3.5 | 4.6 | 4.3 | 8.3 | 8.2 |

pounds **7**–**14**. We performed the ring parameterization for adenine derivatives **9**, **10**, **13** and **14** (see Supporting Information File 1 for details) and assumed that replacing adenine for thymine as a nucleobase would not dramatically affect the constants A and B.

The results of the conformational analysis obtained using the Matlab Pseudorotation GUI program are summarized in Table 3. Instead of presenting the pseudorotation parameters *P* and $\Phi_{max}$ in numbers, we rather present the pseudorotation maps generated by the program in which the contour plots indicate the root-mean-square-deviation between the fitted and experimental $^3J_{HH}$. These conformation maps provide a more realistic view of the conformational behavior of compounds **7**–**14** than a single numeric value for *P* and $\Phi_{max}$.

The pseudorotation maps reveal the effect of the substitution of the pyrrolidine ring on its conformation. The conformation of phosphonomethyl derivatives **7** and **9** with *trans* configuration on 3' and 4' carbon atoms tends to occupy mainly (~60%) the $^1T_2$ or $^1E$ conformation in an equilibrium with the less populated (~40%) $_1E$ conformer. This equilibrium is significantly shifted in the case of *cis* derivatives **8** and **10** towards the $_2T^3$ conformation (>80%) due to the steric hindrance of 3' and 4' substituents, where the nucleobase occupies predominantly the pseudo-equatorial position. There is an obvious contrast between minor conformations of **8** and **10**, which might be caused by hydrogen bonding between the 3'-OH group and the 4'-nucleobase substituent.

A similar effect in the conformer populations was observed for phosphonoformyl derivatives. While *trans* derivatives **11** and **13** exist as an about 1:1 equilibrium mixture of $^{3}T_{4}$ and $_{3}T^{4}$ conformations, *cis* derivatives **12** and **14** prefer (>80%) the $^{3}T_{4}$ conformation. The orientation of the phosphonoformyl moiety (rotamers A and B) has a negligible effect on the conformation of the pyrrolidine ring; therefore, only conformation maps for rotamer A are presented in Table 3. The effect of the alkylation

**Table 3:** The results of the conformational analysis.

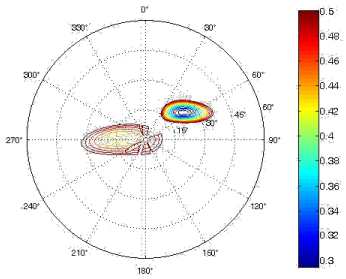
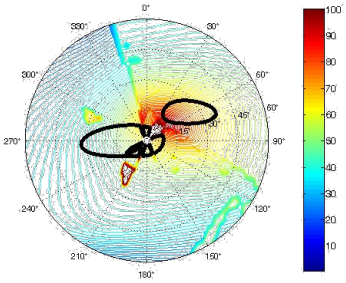
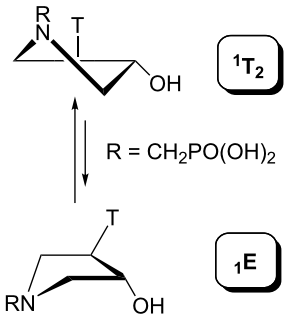
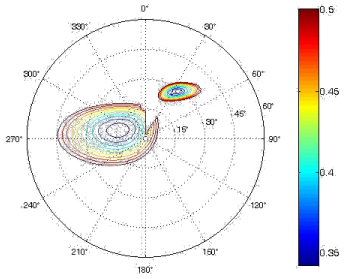
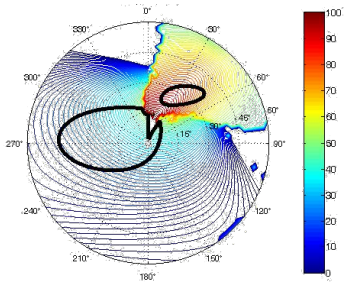
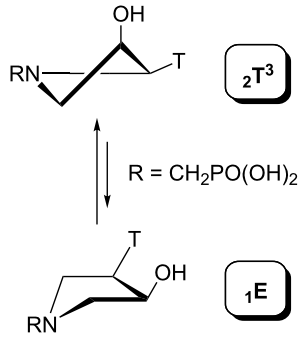
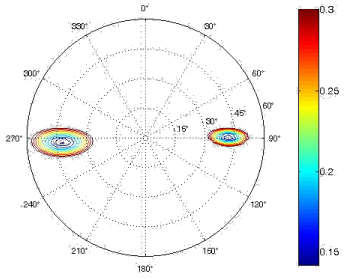
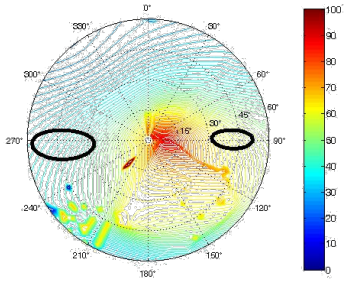
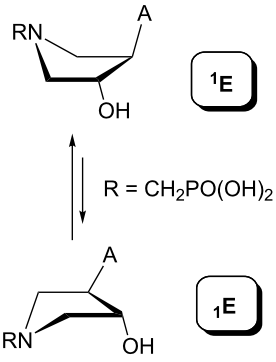
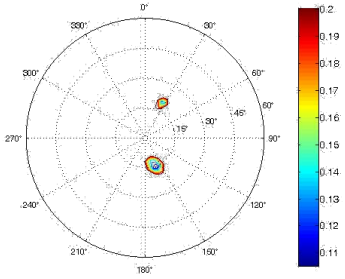
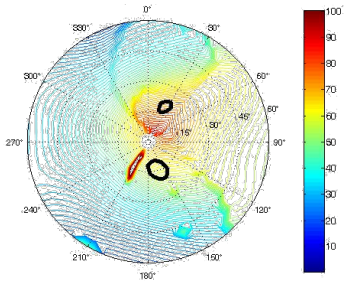
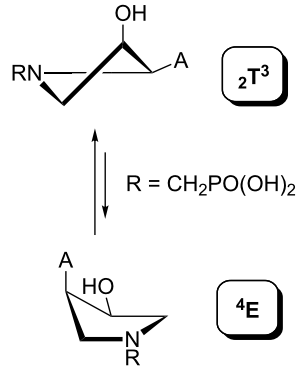
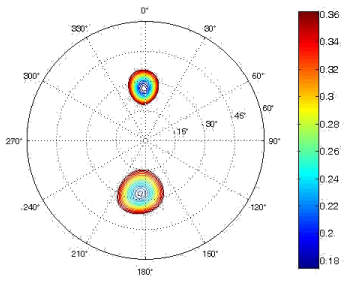
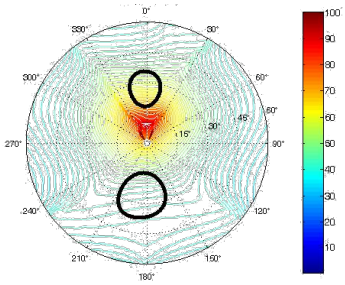
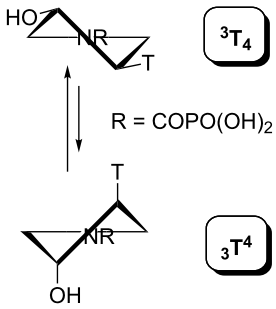
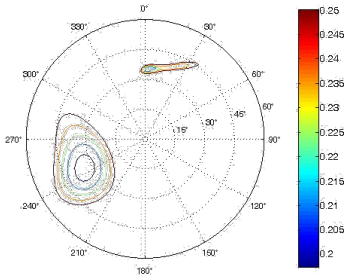
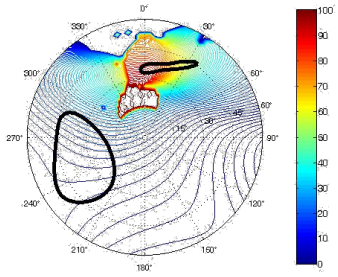
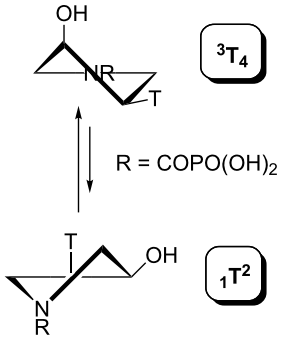
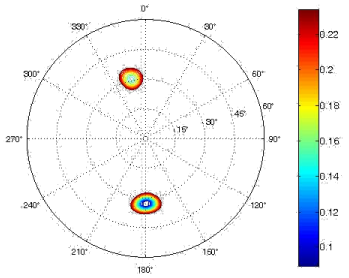
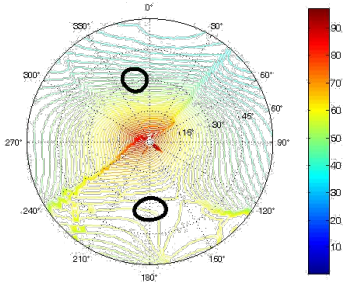
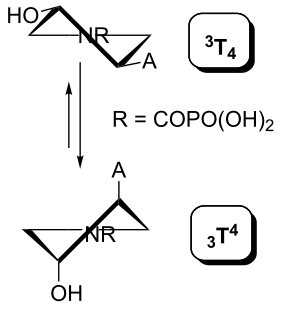
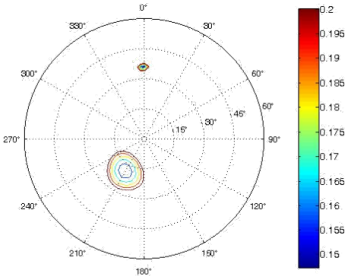
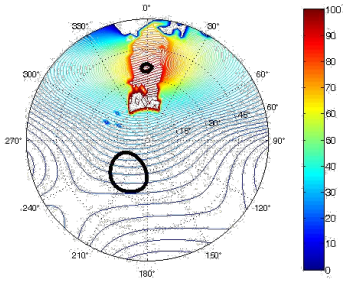
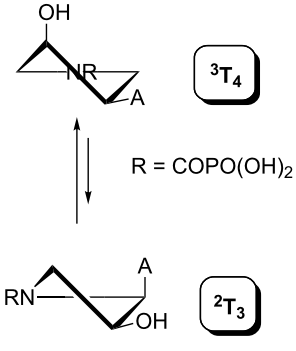
| Analog | Pseudorotation map[a] | Conformer population[b] | Conformation preferences[c] |
|---|---|---|---|
| **7** |  |  |  |
| **8** |  |  |  |
| **9** |  |  |  |
| **10** |  |  |  |

**Table 3:** The results of the conformational analysis. (continued)

| | | | |
|---|---|---|---|
| **11A** |  |  |  |
| **12A** |  |  |  |
| **13A** |  |  |  |
| **14A** |  |  |  |

[a]The contour lines indicate the total RMSD of the fit. [b]The contour lines indicate the percentage of the conformation present in the best fit. The thicker black line is the outer contour line of the two dominant conformations from the second column of the table. [c]A graphical illustration of the conformations obtained from the pseudorotation maps.

or acylation of the pyrrolidine nitrogen is also fairly visible in the pseudorotation maps. Phosphonoformyl derivatives **11–14** keep the conformation of the pyrrolidine ring in narrow northern ($P \sim 0°$) and southern (P ~ 180°) regions, which arises from the plane arrangement of the C2'–N1'–C5' fragment of the pyrrolidine ring (minor conformers of *cis* derivatives **12A** and **14A** do not follow this trend). In contrast to that, the conformation of the pyrrolidine ring in phosphonomethyl derivatives **7–10** is more flexible and rather occupies the eastern and western segments of the pseudorotation wheel, reacting sensi-

tively to the hydrogen bonding between the 3'-OH group and the 4'-nucleobase substituent.

In order to supplement our experimental observation, we calculated the energy profile of the five-membered pyrrolidine ring pseudorotation for adenine derivatives **9** and **13**. We generated 20 conformers covering the whole pseudorotation pathway in 18-degree steps with constant maximum puckering amplitude of 40 degrees. For each conformer, the endocyclic torsion angles $\Phi_0$ and $\Phi_3$ were kept fixed and the geometry of the molecule was optimized using the DFT B3LYP/6-31G* method in vacuo. Predominant conformations can be found as the energy minima by plotting the calculated energy against phase angle *P* (Figure 13). The curves in Figure 13 proved a two-state equilibrium used as a prerequisite in PSEUROT-based programs and known for furanose conformation in natural nucleosides and nucleotides.

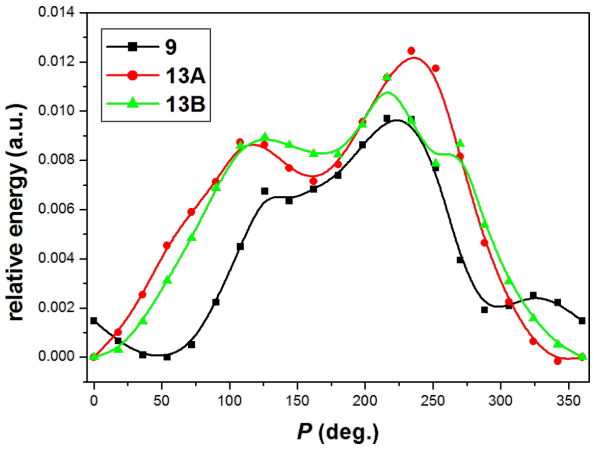



**Figure 13:** The energy profile of the five-membered pyrrolidine ring pseudorotation for adenine derivatives **9** and **13**.

The two energy minima were then fully optimized by B3LYP/6-31++G* in water (the CPCM model). The optimized conformations were compared with those obtained by conformational analysis using $^3J_{HH}$ and the results showed a good agreement of the molecular modeling and the experiment (cf. Table 3 and Figure 14).

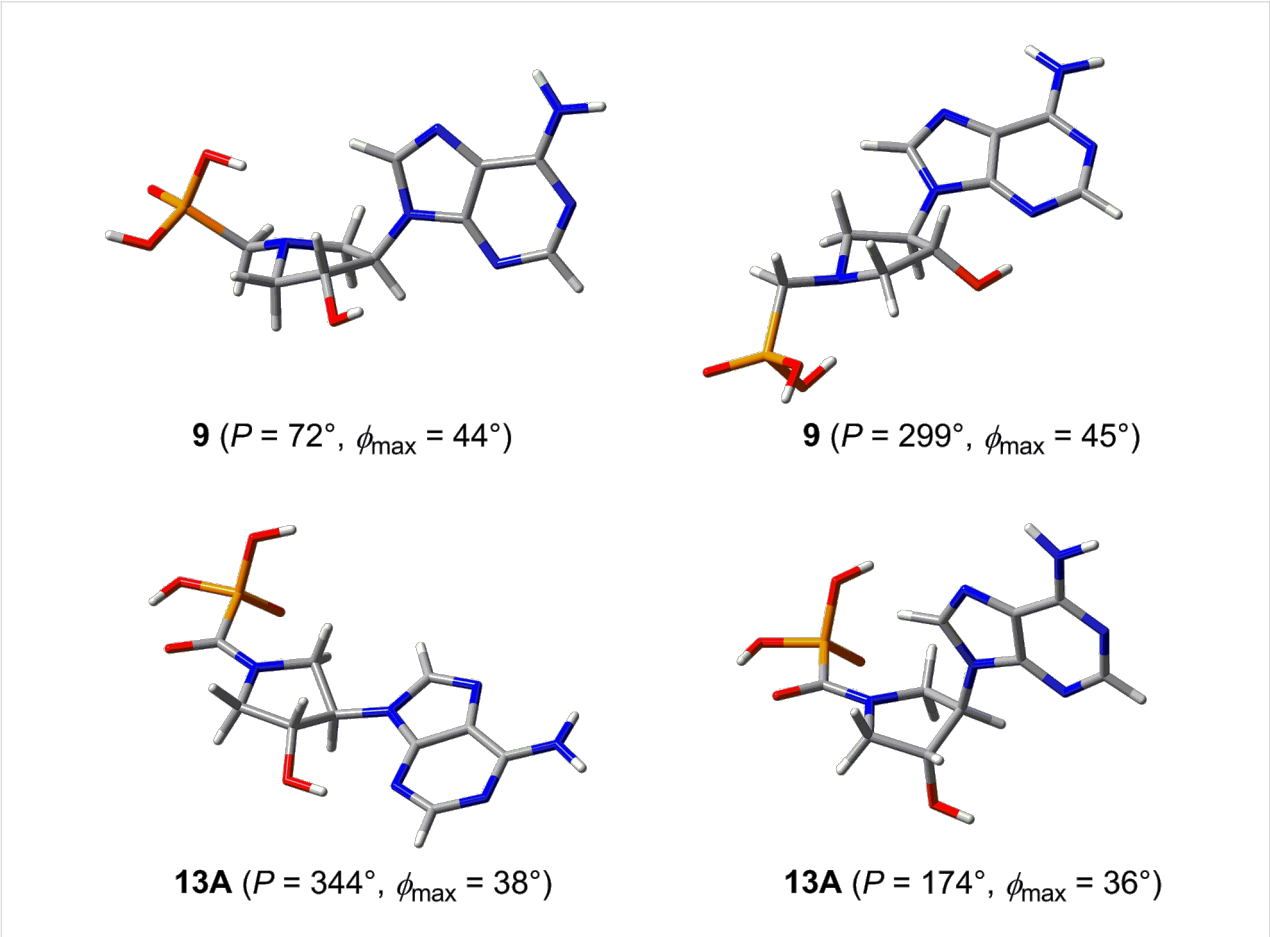



**Figure 14:** The most stable conformations of adenine derivatives **9** and **13A** calculated by the B3LYP/6-31++G* method in water (CPCM).

## Conclusion

In summary, we presented an NMR and DFT conformational analysis of pyrrolidine nucleotide analogs bearing phosphonomethyl or phosphonoformyl substituents attached to the pyrrolidine nitrogen atom. The mode of the phosphonate moiety attachment to the pyrrolidine ring results in tuning the conformation of the five-membered pyrrolidine ring over the whole pseudorotation wheel. While phosphonoformyl derivatives **11–14** keep the conformation of the pyrrolidine ring in narrow northern ($P \sim 0°$) and southern ($P \sim 180°$) regions, which arises from the plane arrangement of the C2'–N1'–C5' fragment of the pyrrolidine ring, the conformation of phosphonomethyl derivatives **7–10** is more flexible and rather occupies the eastern and western segment of the pseudorotation wheel. In general, this simple yet effective tuning of the pyrrolidine-ring conformation can be used for a number of other pyrrolidine derivatives.

## Experimental

### Chemistry

Phosphonoalkyl derivatives **7**–**10** were prepared according to the procedures described in [20] while thymine carbonylphosphonic acids **11** and **12** were prepared according to the procedures described in [4]. Unless stated otherwise, all solvents used were anhydrous. TLC was performed on TLC plates precoated with silica gel (silica gel/TLC-cards, UV 254, Merck). The compounds were detected using UV light (254 nm), by spraying with 1% ethanolic solution of ninhydrine to visualize amines, or by spraying with a 1% solution of 4-(4-nitrobenzyl)pyridine in ethanol, followed by heating and treatment with gaseous ammonia (for the detection of alkylating agents, such as dialkyl phosphonates; blue color). The course of the reactions and the purity of the prepared compounds were determined by LC–MS using a Waters AutoPurification System with a 2545 Quaternary Gradient Module and a 3100 Single Quadrupole Mass Detector using a Luna C18 column (100 × 4.6 mm, 3 μm, Phenomenex) at a flow rate of 1 mL/min. The following mobile phase was used, where A, B, and C represent 50 mM $NH_4HCO_3$, 50 mM $NH_4HCO_3$ in 50% aq $CH_3CN$, and $CH_3CN$, respectively: A→B over 10 min, B→C over 10 min and C for 5 min. Preparative RP HPLC was performed on an LC5000 Liquid Chromatograph (INGOS-PIKRON, CR) using a Luna C18 (2) column (250 × 21.2 mm, 5 μm) at a flow rate of 10 mL/min. A gradient elution of methanol in pH 7.5 0.1 M TEAB (A, 0.1 M TEAB; B, 0.1 M TEAB in 50% aq methanol; C, methanol) was used. All final compounds were lyophilized. Mass spectra were collected on an LTQ Orbitrap XL (Thermo Fisher Scientific) using ESI ionization. The pD titrations were performed on pH meter IQ150 Scientific Instruments equipped with a pH probe PH47-SS. NMR spectra were acquired in $D_2O$ on a Bruker AVANCE 600 ($^1H$ at 600.1 MHz and $^{13}C$ at 150.9 MHz), Bruker AVANCE 500 ($^1H$ at 499.8 MHz, $^{13}C$ at 125.7 MHz and $^{31}P$ at 202.3 MHz) and/or Bruker AVANCE 400 ($^1H$ at 400.0 MHz, $^{13}C$ at 100.6 MHz and $^{31}P$ at 162.0 MHz) NMR spectrometers. Chemical shifts (in ppm, δ scale) were referenced to the 1,4-dioxane signal (3.75 ppm for $^1H$ and 69.3 ppm for $^{13}C$) as an internal or external standard. $^{31}P$ NMR spectra were referenced to $H_3PO_4$ (0 ppm) as an external standard. Coupling constants ($J$) are given in Hz. The complete assignment of $^1H$ and $^{13}C$ signals was performed by an analysis of the correlated homonuclear H,H-COSY, H,H-ROESY and heteronuclear H,C-HSQC and H,C-HMBC spectra.

### Calculations

All calculations were carried out using the Gaussian 09 software package [21]. DFT calculations were performed using Becke3-LYP [22,23] with 6-31G* basis set for a fixed conformation geometry in the calculation of the energy profile (Figure 13) or 6-31++G* basis set for full geometry optimizations. The full geometry optimizations were carried out using conductor-like polarizable continuum model (CPCM) [24].

### General method 1: Synthesis of phosphonoacyl derivatives (**13** and **14**)

General method 1 followed the similar procedure described previously [6]. The mixture of pyrrolidine nucleosides **15a,b** [15] (1 mmol) and diisopropyl phenylphosphonoformiate (1.2 mmol) in DMF (10 mL) was stirred at 80 °C for 1 h (the course of the reaction was followed by TLC using the solvent system H1 (EtOAc/acetone/EtOH/$H_2O$ 4:1:1:1) or LC–MS). The reaction mixture was concentrated in vacuo and the intermediate **18a,b** was obtained in a pure form by chromatography on silica gel using a linear gradient of H1 in ethyl acetate. The intermediate **18a,b** (1 mmol) was dissolved in acetonitrile (10 mL) and bromotrimethylsilane (5 mmol) was added under argon atmosphere. The reaction mixture was stirred overnight at room temperature under argon atmosphere. The reaction mixture was concentrated in vacuo, dissolved in 0.5 M aqueous TEAB (5 mL) and evaporated. The final product was obtained by preparative HPLC, converted to its sodium salt by passing through a column of Dowex 50 in a $Na^+$ form (10 mL/mmol) and lyophilized from water.

### (3*R*,4*R*)-4-(Adenin-9-yl)-3-hydroxypyrrolidin-1-*N*-ylcarbonylphosphonic acid (**13**)

The title compound was prepared according to general method 1 from pyrrolidine nucleoside **15a** (0.08 g, 0.36 mmol) in 56% overall yield (75.2 mg, 0.2 mmol) in the form of a white amorphous solid. HRMS–ESI: $[M + H]^+$ calcd for $C_{10}H_{13}O_5N_6NaP$, 351.05736; found, 351.05728. The NMR spectra showed a 10:11 mixture of rotamers A:B. Rotamer A: $^1H$ NMR

(600.1 MHz, $D_2O$, $T$ = 25 °C, pD = 6.1) 3.54 (dddd, $J_{gem}$ = 13.6, $J_{2",3'}$ = 4.1, $J_{H,P}$ = 1.9, $J_{2",5"}$ = 0.9, 1H, H-2"), 3.83 (ddd, $J_{gem}$ = 13.6, $J_{2',3'}$ = 6.1, $J_{H,P}$ = 1.9, 1H, H-2'), 4.55 (ddd, $J_{gem}$ = 13.0, $J_{5',4'}$ = 5.6, $J_{5',3'}$ = 0.9, 1H, H-5'), 4.57 (ddd, $J_{gem}$ = 13.0, $J_{5",4'}$ = 6.3, $J_{5",2"}$ = 0.9, 1H, H-5"), 4.73 (dddd, $J_{3',2'}$ = 6.1, $J_{3',4'}$ = 4.5, $J_{3',2"}$ = 4.1, $J_{3',5'}$ = 0.9, 1H, H-3'), 5.01 (ddd, $J_{4',5"}$ = 6.3, $J_{4',5'}$ = 5.6, $J_{4',3'}$ = 4.5, 1H, H-4'), 8.134 (s, 1H, H-8), 8.151 (s, 1H, H-2); $^{13}$C NMR (150.9 MHz, $D_2O$, $T$ = 25 °C, pD = 6.1) 51.65 (C-5'), 53.11 (d, $J_{C,P}$ = 4.6, C-2'), 62.87 (C-4'), 73.75 (C-3'), 121.29 (C-5), 142.56 (C-8), 151.58 (C-4), 154.92 (C-2), 157.98 (C-6), 179.15 (d, $J_{C,P}$ = 197.5, P-CO); $^{31}$P{1H} NMR (202.3 MHz, $D_2O$, $T$ = 25 °C, pD = 6.1) −2.25. Rotamer B: $^1$H NMR (600.1 MHz, $D_2O$, $T$ = 25 °C, pD = 6.1) 3.95 (dd, $J_{gem}$ = 12.6, $J_{2",3'}$ = 5.2, 1H, H-2"), 4.01 (ddd, $J_{gem}$ = 13.6, $J_{5',4'}$ = 5.5, $J_{H,P}$ = 1.9, 1H, H-5'), 4.19 (ddd, $J_{gem}$ = 13.6, $J_{5",4'}$ = 7.8, $J_{H,P}$ = 2.0, 1H, H-5"), 4.31 (ddd, $J_{gem}$ = 12.6, $J_{2',3'}$ = 5.8, $J_{2',5'}$ = 0.8, 1H, H-2'), 4.76 (ddd, $J_{3',2'}$ = 5.8, $J_{3',4'}$ = 5.3, $J_{3',2"}$ = 5.2, 1H, H-3'), 4.99 (ddd, $J_{4',5"}$ = 7.8, $J_{4',5'}$ = 5.5, $J_{4',3'}$ = 5.3, 1H, H-4'), 8.130 (s, 1H, H-8), 8.149 (s, 1H, H-2); $^{13}$C NMR (150.9 MHz, $D_2O$, $T$ = 25 °C, pD = 6.1) 50.23 (d, $J_{C,P}$ = 4.9, C-5'), 54.64 (C-2'), 60.88 (C-4'), 75.44 (C-3'), 121.29 (C-5), 142.73 (C-8), 151.62 (C-4), 154.95 (C-2), 158.00 (C-6), 179.24 (d, $J_{C,P}$ = 197.0, P-CO); $^{31}$P{1H} NMR (202.3 MHz, $D_2O$, $T$ = 25 °C, pD = 6.1) −2.23.

### (3*S*,4*R*)-4-(Adenin-9-yl)-3-hydroxypyrrolidin-1-*N*-ylcarbonylphosphonic acid (**14**)

The title compound was prepared according to general method 1 from pyrrolidine nucleoside **15b** (0.19 g, 0.86 mmol) in 28% overall yield (91 mg, 0.25 mmol) in the form of a white amorphous solid. HRMS–ESI: $[M + H]^+$ calcd for $C_{10}H_{13}O_5N_6NaP$, 351.05736; found, 351.05773. The NMR spectra showed a 10:11 mixture of rotamers A:B. Rotamer A: $^1$H NMR (600.1 MHz, $D_2O$, $T$ = 25 °C, pD = 5.9) 3.65 (dt, $J_{gem}$ = 13.7, $J_{2',3'} = J_{H,P}$ = 2.1, 1H, H-2'), 3.90 (ddd, $J_{gem}$ = 13.7, $J_{2",3'}$ = 5.1, $J_{H,P}$ = 1.9, 1H, H-2"), 4.41 (dd, $J_{gem}$ = 11.6, $J_{5',4'}$ = 9.2, 1H, H-5'), 4.66 (dd, $J_{gem}$ = 11.6, $J_{5",4'}$ = 7.6, 1H, H-5"), 4.693 (ddd, $J_{3',2"}$ = 5.1, $J_{3',4'}$ = 4.2, $J_{3',2'}$ = 2.1, 1H, H-3'), 5.19 (ddd, $J_{4',5'}$ = 9.2, $J_{4',5"}$ = 7.6, $J_{4',3'}$ = 4.2, 1H, H-4'), 8.158 (s, 1H, H-2), 8.262 (s, 1H, H-8); $^{13}$C NMR (150.9 MHz, $D_2O$, $T$ = 25 °C, pD = 5.9) 50.18 (C-5'), 54.30 (d, $J_{C,P}$ = 4.5, C-2'), 58.95 (C-4'), 70.35 (C-3'), 120.80 (C-5), 143.65 (C-8), 151.93 (C-4), 154.89 (C-2), 157.95 (C-6), 178.88 (d, $J_{C,P}$ = 198.3, P-CO); $^{31}$P{$^1$H} NMR (202.3 MHz, $D_2O$, $T$ = 25 °C, pD = 5.9) −2.21. Rotamer B: $^1$H NMR (600.1 MHz, $D_2O$, $T$ = 25 °C, pD = 5.9) 4.04 (ddd, $J_{gem}$ = 12.8, $J_{5',4'}$ = 8.3, $J_{H,P}$ = 1.9, 1H, H-5'), 4.14 (dd, $J_{gem}$ = 13.0, $J_{2',3'}$ = 3.5, 1H, H-2'), 4.18 (ddd, $J_{gem}$ = 12.8, $J_{5",4'}$ = 8.2, $J_{H,P}$ = 1.8, 1H, H-5"), 4.20 (dd, $J_{gem}$ = 12.6, $J_{2",3'}$ = 4.6, 1H, H-2"), 4.698 (ddd, $J_{3',2"}$ = 4.6, $J_{3',4'}$ = 4.3, $J_{3',2'}$ = 3.5, 1H, H-3'), 5.21 (ddd, $J_{4',5'}$ = 8.3, $J_{4',5"}$ = 8.2, $J_{4',3'}$ = 4.3, 1H, H-4'), 8.163 (s, 1H, H-2), 8.261 (s, 1H, H-8); $^{13}$C NMR (150.9 MHz, $D_2O$, $T$ = 25 °C, pD = 5.9) 49.06 (d, $J_{C,P}$ = 5.2, C-5'), 55.63 (C-2'), 57.39 (C-4'), 72.34 (C-3'), 120.82 (C-5), 143.66 (C-8), 152.04 (C-4), 154.90 (C-2), 157.95 (C-6), 178.91 (d, $J_{C,P}$ = 197.9, P-CO); $^{31}$P{$^1$H} NMR (202.3 MHz, $D_2O$, $T$ = 25 °C, pD = 5.9) −2.30.

## Supporting Information

The pD titration curves of **9** and **10**, the variable temperature NMR of **14**, the pyrrolidine-ring parameterization for **9**, **10**, **13** and **14**, the NMR signal assignments and copies of the $^1$H, $^{13}$C and $^{31}$P NMR spectra are given.

### Supporting Information File 1

Additional experimental data.
[http://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-10-205-S1.pdf]

## Acknowledgements

This work was supported by the Czech Science Foundation (Project No. 13-24880S).

## References

1. De Clercq, E.; Sakuma, T.; Baba, M.; Pauwels, R.; Balzarini, J.; Rosenberg, I.; Holý, A. *Antiviral Res.* **1987,** *8,* 261–272. doi:10.1016/S0166-3542(87)80004-9
2. Balzarini, J.; Naesens, L.; Herdewijn, P.; Rosenberg, I.; Holý, A.; Pauwels, R.; Baba, M.; Johns, D. G.; De Clercq, E. *Proc. Natl. Acad. Sci. U. S. A.* **1989,** *86,* 332–336. doi:10.1073/pnas.86.1.332
3. Mackman, R. L.; Boojamra, C. G.; Prasad, V.; Zhang, L.; Lin, K.-Y.; Petrakovsky, O.; Babusis, D.; Chen, J.; Douglas, J.; Grant, D.; Hui, H. C.; Kim, C. U.; Markevitch, D. Y.; Vela, J.; Ray, A.; Cihlar, T. *Bioorg. Med. Chem. Lett.* **2007,** *17,* 6785–6789. doi:10.1016/j.bmcl.2007.10.038
4. Kočalka, P.; Rejman, D.; Vaněk, V.; Rinnová, M.; Tomečková, I.; Králiková, Š.; Petrová, M.; Páv, O.; Pohl, R.; Buděšínský, M.; Liboska, R.; Točík, Z.; Panová, N.; Votruba, I.; Rosenberg, I. *Bioorg. Med. Chem. Lett.* **2010,** *20,* 862–865. doi:10.1016/j.bmcl.2009.12.081
5. Rejman, D.; Panová, N.; Klener, P.; Maswabi, B.; Pohl, R.; Rosenberg, I. *J. Med. Chem.* **2012,** *55,* 1612–1621. doi:10.1021/jm201409u
6. Keough, D. T.; Hocková, D.; Rejman, D.; Špaček, P.; Vrbková, S.; Krečmerová, M.; Eng, W. S.; Jans, H.; West, N. P.; Naesens, L. M. J.; de Jersey, J.; Guddat, L. W. *J. Med. Chem.* **2013,** *56,* 6967–6984. doi:10.1021/jm400779n
7. Freier, S. M.; Altmann, K.-H. *Nucleic Acids Res.* **1997,** *25,* 4429–4443. doi:10.1093/nar/25.22.4429
8. Karplus, M. *J. Am. Chem. Soc.* **1963,** *85,* 2870–2871. doi:10.1021/ja00901a059
9. Altona, C.; Sundaralingam, M. *J. Am. Chem. Soc.* **1972,** *94,* 8205–8212. doi:10.1021/ja00778a043

10. De Leeuw, F. A. A. M.; Altona, C. *J. Comput. Chem.* **1983,** *4,* 428–437. doi:10.1002/jcc.540040319
11. Padrta, P.; Sklenář, V. *J. Biomol. NMR* **2002,** *24,* 339–349. doi:10.1023/A:1021656808607
12. Hendrickx, P. M. S.; Martins, J. C. *Chem. Cent. J.* **2008,** *2,* No. 20. doi:10.1186/1752-153X-2-20
13. Sharma, N. K.; Ganesh, K. N. *Tetrahedron* **2010,** *66,* 9165–9170. doi:10.1016/j.tet.2010.09.082
14. Watts, J. K.; Sadalapure, K.; Choubdar, N.; Pinto, B. M.; Damha, M. J. *J. Org. Chem.* **2006,** *71,* 921–925. doi:10.1021/jo051844+
15. Rejman, D.; Kočalka, P.; Buděšínský, M.; Pohl, R.; Rosenberg, I. *Tetrahedron* **2007,** *63,* 1243–1253. doi:10.1016/j.tet.2006.11.047
16. Christensen, J. J.; Rytting, J. H.; Izatt, R. M. *J. Phys. Chem.* **1967,** *71,* 2700–2705. doi:10.1021/j100867a047
17. Christensen, J. J.; Rytting, J. H.; Izatt, R. M. *Biochemistry* **1970,** *9,* 4907–4913. doi:10.1021/bi00827a012
18. De Leeuw, H. P. M.; Haasnoot, C. A. G.; Altona, C. *Isr. J. Chem.* **1980,** *20,* 108–126. doi:10.1002/ijch.198000059
19. Houseknecht, J.; Altona, C.; Hadad, C. M.; Lowary, T. L. *J. Org. Chem.* **2002,** *67,* 4647–4651. doi:10.1021/jo025635q
20. Rejman, D.; Pohl, R.; Kočalka, P.; Masojídková, M.; Rosenberg, I. *Tetrahedron* **2009,** *65,* 3673–3681. doi:10.1016/j.tet.2009.02.071
21. Gaussian 09, Revision D.01, Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J., J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A.; Gaussian, Inc.: Wallingford CT, 2009.
22. Becke, A. D. *J. Chem. Phys.* **1993,** *98,* 5648–5652. doi:10.1063/1.464913
23. Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988,** *37,* 785–789. doi:10.1103/PhysRevB.37.785
24. Cossi, M.; Rega, N.; Scalmani, G.; Barone, V. *J. Comput. Chem.* **2003,** *24,* 669–681. doi:10.1002/jcc.10189

## License and Terms

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

# Synthesis of phosphoramidites of isoGNA, an isomer of glycerol nucleic acid

Keunsoo Kim, Venkateshwarlu Punna, Phaneendrasai Karri and Ramanarayanan Krishnamurthy*

Full Research Paper

Open Access

Address:
Department of Chemistry, The Scripps Research Institute, 10550 North Torrey Pines Rd, La Jolla, CA 92037, USA

Email:
Ramanarayanan Krishnamurthy* - rkrishna@scripps.edu

* Corresponding author

Keywords:
acyclic nucleic acids; glycerol nucleic acids; isoGNA; oligonucleotides; phosphoramidites

*Beilstein J. Org. Chem.* **2014,** *10,* 2131–2138.
doi:10.3762/bjoc.10.220

Received: 01 May 2014
Accepted: 15 August 2014
Published: 08 September 2014

This article is part of the Thematic Series "Nucleic acid chemistry".

Guest Editor: H.-A. Wagenknecht



## Abstract

IsoGNA, an isomer of glycerol nucleic acid GNA, is a flexible (acyclic) nucleic acid with bases directly attached to its linear backbone. IsoGNA exhibits (limited) base-pairing properties which are unique compared to other known flexible nucleic acids. Herein, we report on the details of the preparation of isoGNA phosphoramidites and an alternative route for the synthesis of the adenine derivative. The synthetic improvements described here enable an easy access to isoGNA and allows for the further exploration of this structural unit in oligonucleotide chemistry thereby spurring investigations of its usefulness and applicability.

## Introduction

Acyclic nucleic acids have garnered a lot of attention and are becoming an important component in nucleic acid chemistry [1-3]. Lately, there has been an explosion in the number of investigated candidates [4-7]. We have recently reported on the base-pairing properties of isoGNA, an isomer of glycerol derived nucleic acid (Figure 1) [8]. Our motivation was driven by the question “how structurally simple and minimal can an oligonucleotide be and still exhibit base-pairing?” When we studied isoGNA we were surprised to find that the base-pairing properties were unpredictably different from other closely related acyclic nucleic acids [8]. This indicated that the nature of the backbone exerts a strong influence on the disposition of the nucleobases, and there seems to be no straightforward correlation between the nature of the backbone and base-pairing properties [9].

We are continuing our investigation of isoGNA oligonucleotides within the context of chimeric sequences and collaborative intercalation studies and thus have a continuing need for the synthesis of isoGNA building blocks. While iso-glycerol nucleosides are known [10-15], the published synthetic routes and efficiencies of these vary widely and primarily depend on

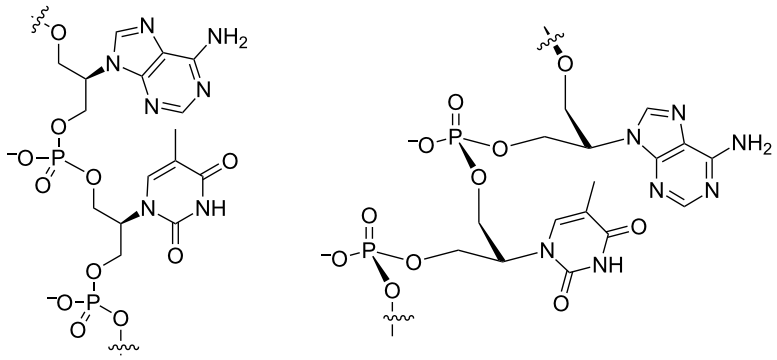


**Figure 1:** A constitutional and conformational (idealized) representation of isoGNA, an isomer of glycerol nucleic acid.

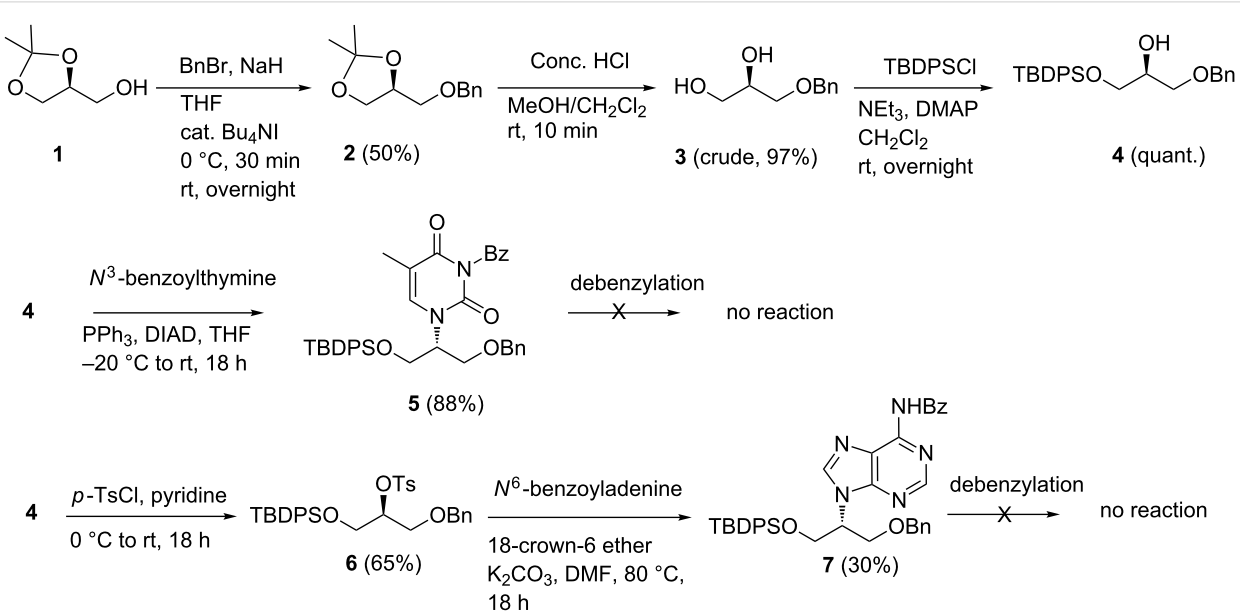
the nature of the nucleobase. Herein, we detail our synthetic efforts and the current improvements for the synthesis of all four isoGNA nucleosides and their phosphoramidites.

# Results and Discussion

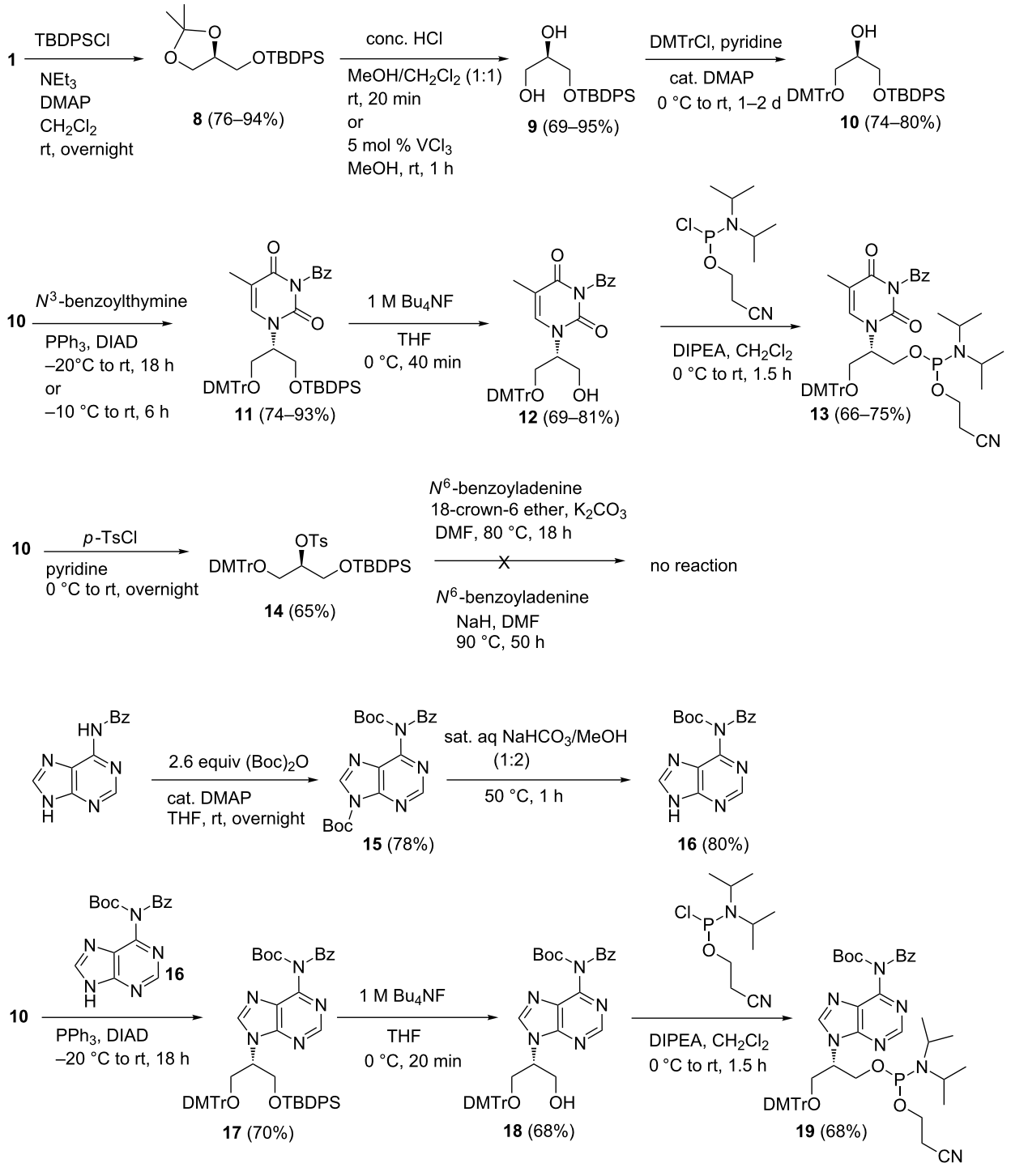
## Initial synthetic routes

We started with commercially available (*S*)-solketal (**1**), which was protected as the benzyl ether **2**. Ketal deprotection followed by the silylation of the primary hydroxy group with *tert*-butyldiphenylsilyl chloride provided the substrate **4**, which is suitable for the introduction of the canonical nucleobases by a $S_N2$ displacement reaction. At this point, we chose the Mitsunobu reaction (Scheme 1), which has been widely employed [16]. The reaction with the $N^3$-benzoyl protected thymine delivered **5** in good yield. However, the reaction with $N^6$-benzoyladenine proved problematic and no product was detected. Therefore, in the adenine case, we tried a direct $S_N2$ reaction of $N^6$-benzoyladenine with tosyl derivative **6**, which was only moderately successful (30% yield of **7**). Unexpectedly, in both the thymine and adenine cases, the subsequent debenzylation reaction failed under a variety of conditions ($H_2$, 10% Pd/C, MeOH or EtOAc; Pd/C, DMF, cyclohexadiene; 10% Pd/C, $NH_4HCOO$, acetone, reflux; 10% Pd-black, DMF, cyclohexadiene; $(CH_3)_3SiI$, $CHCl_3$, 25 °C), for reasons that were not obvious to us. Consequently, we abandoned this approach and pursued a different protecting group approach.

We silylated **1** to afford **8** whose ketal was deprotected followed by tritylation to give the derivative **10**, which is expected to be suited for the introduction of nucleobases by an $S_N2$ reaction (Scheme 2). Once again, the Mitsunobu reaction with thymine proceeded smoothly. However, the reaction with




**Scheme 1:** Initial routes towards the synthesis of iso-glycerol nucleosides.



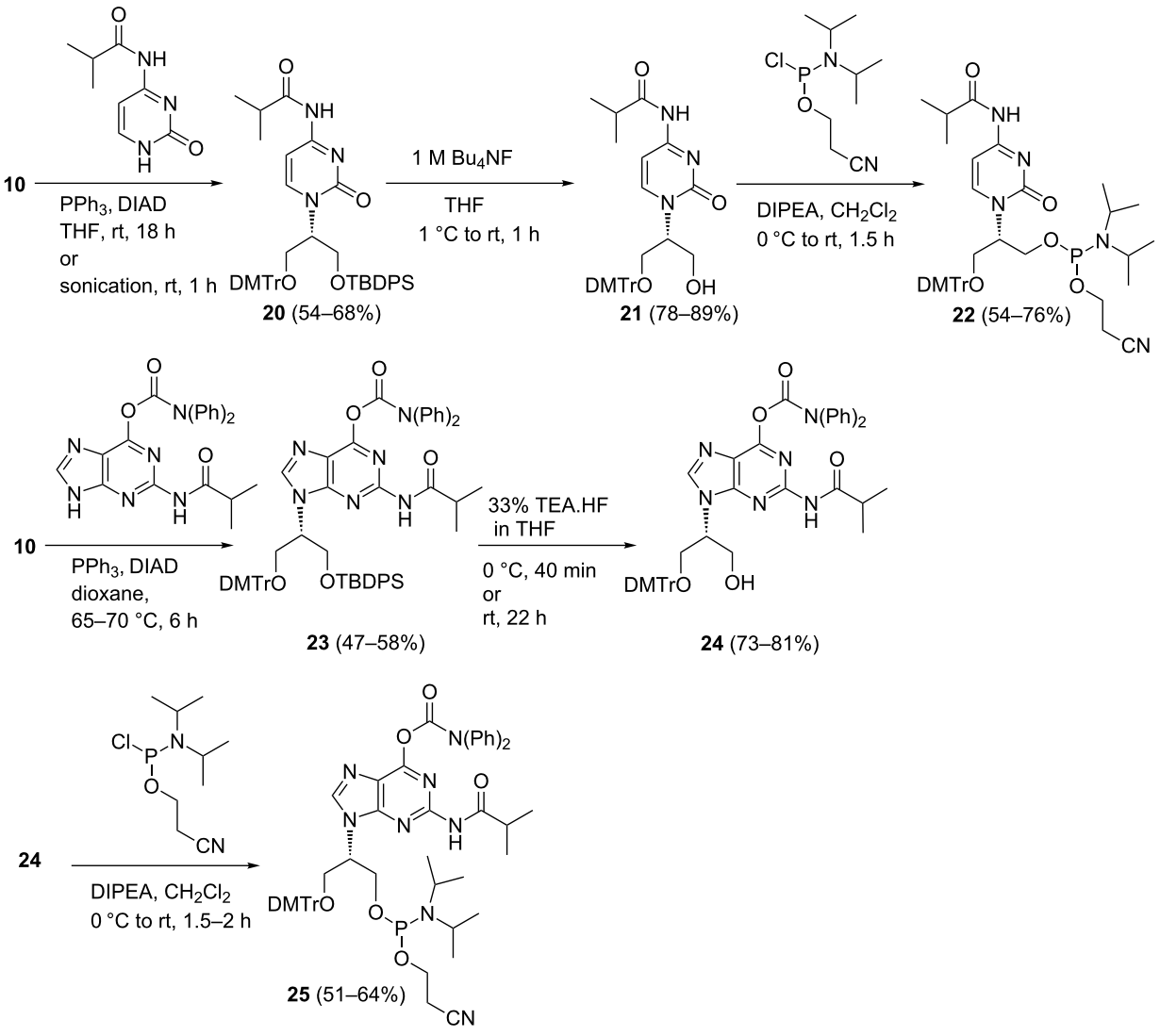

**Scheme 2:** Preparation of thymine and adenine containing phosphoramidites of isoGNA.

$N^6$-benzoyladenine proved problematic and did not proceed even with the tosyl derivative **14**. The solubility issues with benzoylated adenine and the bulk of the O-TBDPS group may be the underlying reasons behind this difficulty. Therefore, we proposed to use a doubly $N^6$-protected adenine, the $N^6$-benzoyl-Boc derivative **16**, which was synthesized starting from $N^6$-benzoyladenine in two steps. $N^6$-Benzoyladenine was treated with 2.6 equiv $(Boc)_2O$ to afford $N^6$-benzoyl-$N^6$,$N^9$-di-Boc-adenine **15**; sat. aq $NaHCO_3$/MeOH (1:2) conditions cleaved the $N^9$-Boc group of **15** specifically [17,18] to yield **16**.

The use of **16** in the subsequent Mitsunobu reaction afforded 68% of the desired product **17**. With derivatives **11** and **17** in hand we proceeded to prepare the corresponding phosphoramidite derivatives **13** and **19**, respectively, under standard conditions as outlined in Scheme 2.

We proceeded to prepare the cytosine- and guanine-containing phosphoramidite by using the tritylated intermediate **10** as the starting point (Scheme 3). In the cytosine series we found that the normal $N^4$-benzoyl or $N^4$-acetyl protected cytosine did not give satisfactory yields under the Mitsunobu conditions. The best yields were achieved with $N^4$-isobutyrylcytosine as the nucleophile [19-21]. In the guanine series, the $N^2$-isobutyryl, $O^6$-diphenylcarbamoyl protected guanine worked well (Scheme 3) [22].

With all the four phosphoramidites in hand we proceeded to the automated synthesis of the oligomers, which proceeded smoothly with ≥95% coupling efficiency as determined by trityl assays. It was at the stage of the removal of the protecting groups to afford the free oligos that we encountered a severe problem with the adenine derivative. While the aqueous ammonia/methylamine base deprotection removed all base labile protecting groups (acetate, isobutyrate, diphenylcarbamoyl groups), the $N^6$-Boc protecting group on adenine was impervious. We tried different acidic conditions (trifluoroacetic acid, acetic acid with heat, $AlCl_3$ [23-25]), but all of these conditions started to degrade the oligomer without completely removing the *tert*-Boc group (as seen by MALDI–TOF mass spectral data in combination with ion-exchange HPLC monitoring) [8].




**Scheme 3:** Preparation of guanine- and cytosine-containing phosphoramidites of isoGNA.

## Improved synthesis of adenine-isoGNA building block

We went back in our synthesis in order to address the adenine protecting group problem and to re-work the inefficient Mitsunobu reaction of $N^6$-benzoyladenine. We explored various conditions (different solvents, temperature etc.) and found that the sonication [26] of the reaction mixture by using $N^6$-benzoyladenine in anhydrous dioxane led to respectable yields (34–36%) of **26** accompanied by a (unidentified) byproduct (Scheme 4). Without sonication no product **26** was formed (by TLC).

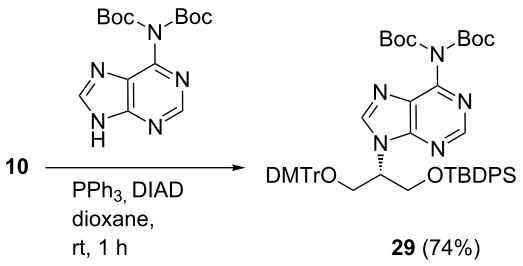
However, the yield of the sonication reaction varied capriciously and the isolation of the compound was complicated by interference from the DEAD side-products during column chromatographic purification. We did manage to get enough amounts of **26** to prepare the phosphoramidite **28**, and proceeded with the oligomer synthesis for our studies. However, this was far from satisfactory, and the continued demand for isoGNA oligomers within our group and our collaborative work prompted us to look for an improvement of the preparation of the adenine building block. Considering the possible reasons for the inefficient reaction of the protected adenine derivatives in the Mitsunobu reaction, we focused on the modification of adenine in order to improve its solubility.

We considered three different derivatives of adenine based on increased solubility: 6-chloroadenine has been utilized to overcome the solubility and regioselectivity ($N^7$ versus $N^9$ nucleosidation) issues. Yet, the conversion of the chloro to the $NH_2$ group is not efficient enough to consider this option an attractive one [27]. The second option was the $N^6$,$N^6$-dibenzoyladenine derivative, but this compound was found to be highly unstable for isolation and handling (one of the $N^6$-benzoyl group falls off very easily). The third option was the use of the $N^6$,$N^6$-di-Boc-adenine derivative, even though this entails the difficulty in removing these Boc groups, which are incompatible with the oligonucleotide backbone. Therefore, we considered the option of removing these Boc groups after the Mitsunobu reaction is performed, and to introduce the $N^6$-benzoyl group before the oligo stage.

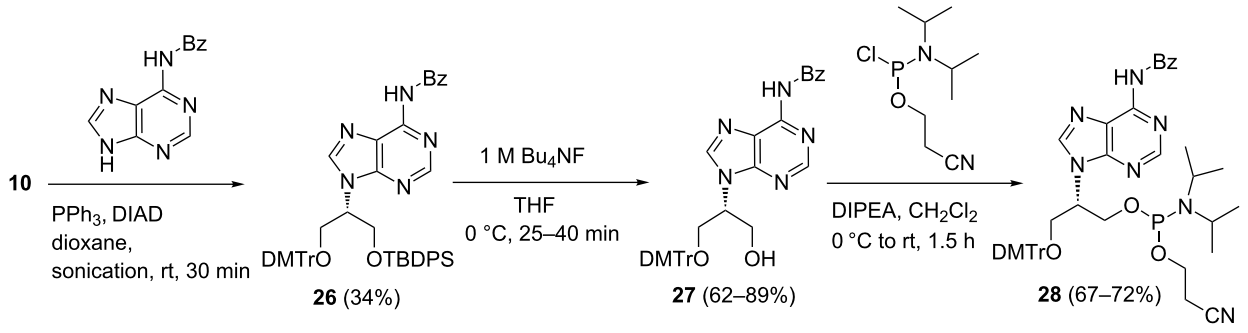
The Schneller group has reported that $N^6$-amino-di-Boc-protected adenine undergoes an efficient reaction under classic Mitsunobu conditions as a result of its increased solubility [28]. The regioselective attack at the $N^9$-position is a consequence of the steric bulk of the di-Boc protection at the $N^6$-position, which renders the $N^7$-position inaccessible. As expected, a high yield was observed under mild conditions with substrates containing a secondary alcohol [28].

Di-Boc protected adenine was readily prepared via protection and followed by selective deprotection according to a literature procedure [17,18]. The Mitsunobu reaction between the di-Boc-adenine and **10** was investigated, and it was found that dioxane as a solvent almost doubled the yield (74%) compared to THF (36%; Scheme 5). Interestingly, sonication afforded no tangible advantage. Doubling the amount of DIAD or di-Boc-adenine or $PPh_3$ did not give increased yields. The desired product **29** was easily purified on silica gel due to its lower polarity (when compared with the $N^6$-benzoyladenine reaction).

Initially, we tried to simultaneously deprotect both Boc groups and the TBDPS group in the presence of TBAF in a one-pot




**Scheme 5:** Mitsunobu reaction with the di-Boc-adenine.




**Scheme 4:** Preparation of adenine containing phosphoramidite of isoGNA.

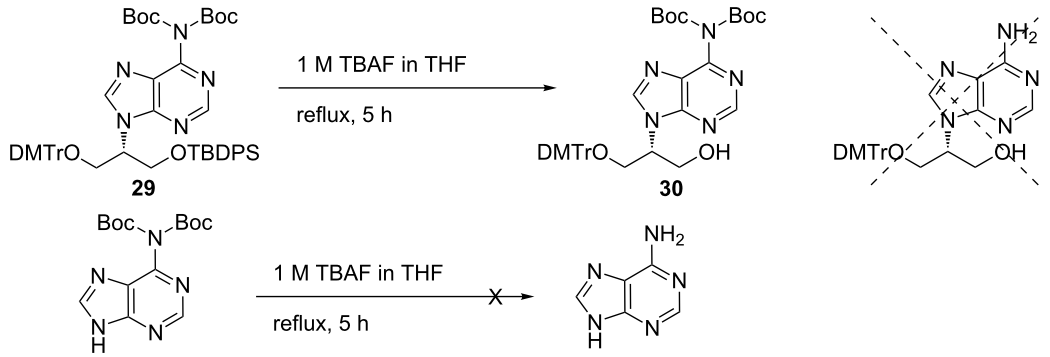
reaction [29]. However, only the TBDPS group was removed and a quantitative amount of mono-protected compound **30** was produced (Scheme 6). The fluoride ion was not effective enough to remove at least one of the Boc groups in **29**. We checked the deprotection conditions for the removal of the Boc moiety by fluoride ions on the model substrate $N^6$,$N^6$-di-Boc-adenine, and found that it was very challenging because free adenine was not afforded even under reflux with 1 M TBAF for 4 days.

Finally, we decided to remove the Boc groups on adenine with trifluoroacetic acid (TFA), even though we also expected the loss of the DMTr group on **29**. When **29** was treated with 20% of TFA in $CH_2Cl_2$, a mixture of **32** and **33** formed after 4 h. The TBDPS moiety was found to be stable under these conditions (Table 1, entry 1). 50% of TFA in $CH_2Cl_2$ completely removed both Boc groups on adenine at either 4 °C or rt to afford the desired **33** in 87% yield.

The trityl group was reintroduced on **33** under standard conditions to give **34**, and subsequent benzoylation afforded **26** in overall good yields (Scheme 7). The NMR spectra of **26** were in good agreement with those data obtained previously [8] via Mitsunobu reaction with $N^6$-benzoyladenine.

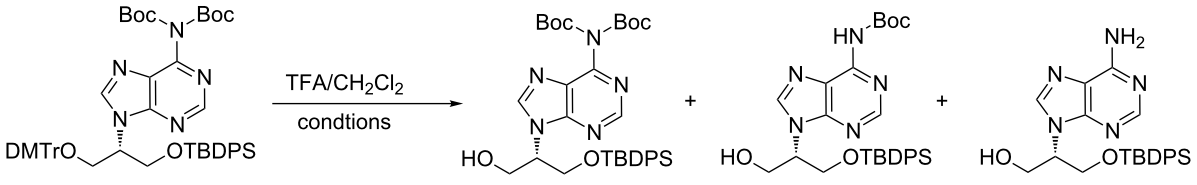
Overall, **26** was afforded in 43% yield from the Mitsunobu reaction of **10** with $N^6$,$N^6$-di-Boc-adenine. In addition to the slight improvement compared to the direct Mitsunobu reaction of **10** with $N^6$-benzoyladenine, this protocol allows for the synthesis of large amounts of **26** with an exact control of the regioselectivity at the $N^9$-position of adenine.

Having solved the "adenine problem", we focused on each of the remaining steps and optimized the synthesis of other isoGNA phosphoramidites. For example, the deprotection of **8** to **9** was initially achieved by the treatment with conc. HCl. However, in scaling up the reaction, the loss of the silyl group




**Scheme 6:** An attempt to remove both Boc and silyl groups simultaneously.

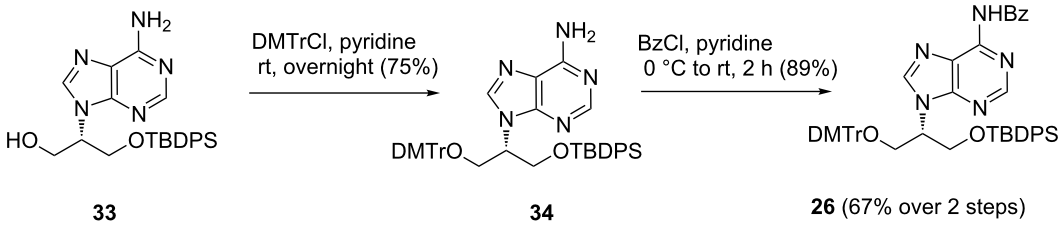
**Table 1:** Result of Boc deprotection of **29** with TFA.




**29** **31** **32** **33**

| Entry | Ratio of TFA (%) | Temp.[a] | Time (h) | Ratio (%)[b] (**31**:**32**:**33**) | Yield |
|---|---|---|---|---|---|
| 1 | 20 | 0 °C to rt | 4 | 0:33:66 | – |
| 2 | 50 | 0 °C to 4 °C | 2 | 0:20:80 | – |
| | | | 4 | 0:0:100 | – |
| 3[c] | 50 | 0 °C to rt | 2 | 0:0:100 | 87% |

[a]TFA (99%) was added under ice-bath (−2 to 2 °C) cooling. [b]The ratio is roughly calculated by the intensity of visualized spots on TLC by UV lamp. [c]This reaction was performed on a large scale.




**Scheme 7:** The synthesis of **26** via tritylation and benzoylation.

was found to occur. Therefore, we sought an alternative route and found that the deprotection mediated by vanadium trichloride gave the desired product **9** cleanly in large-scale reactions [30]. The desilylation of intermediates **11**, **20** and **26** was effected by the treatment of 1 M TBAF in THF at 0 °C to room temperature. In the case of guanine derivative **23**, we observed that the use of the triethylamine·HF complex instead of TBAF was preferable at room temperature. The free alcohols **12**, **18**, **21** and **24** were phosphitylated under standard conditions to obtain the phosphoramidites suitable for the automated solid phase synthesis.

## Conclusion

We have described herein improvements to and the optimization of the synthesis of isoGNA phosphoramidite building blocks. We observed that the Boc protecting group on adenine while not being compatible with oligonucleotide chemistry, can nevertheless be used as a temporary protecting group, which is extremely useful for increasing the solubility and regioselectivity of adenine in the Mitsunobu reaction. The optimized procedures described here allowed us to access a stable supply of isoGNA phosphoramidites for further investigation in our laboratory.

## Supporting Information

### Supporting Information File 1

Experimental procedures.
[http://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-10-220-S1.pdf]

### Supporting Information File 2

$^{1}H$, $^{13}C$ and $^{31}P$ NMR spectra.
[http://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-10-220-S2.pdf]

## Acknowledgements

This work was jointly supported by the NSF and NASA Astrobiology Program under the NSF Center for Chemical Evolution, Grant CHE-1004570 and by NASA Astrobiology: Exobiology and Evolutionary Biology Program (Grant NNX09AM96G).

## References

1. Vandendriessche, F.; Augustyns, K.; Van Aerschot, A.; Busson, R.; Hoogmartens, J.; Herdewijn, P. *Tetrahedron* **1993,** *49,* 7223–7238. doi:10.1016/S0040-4020(01)87200-8
2. Nielsen, P.; Dreiøe, L. H.; Wengel, J. *Bioorg. Med. Chem.* **1995,** *3,* 19–28. doi:10.1016/0968-0896(94)00143-Q
3. Zhang, S.; Switzer, C.; Chaput, J. C. *Chem. Biodiversity* **2010,** *7,* 245–258. doi:10.1002/cbdv.200900281
4. Kumar, V.; Gore, K. R.; Pradeepkumar, P. I.; Kesavan, V. *Org. Biomol. Chem.* **2013,** *11,* 5853–5865. doi:10.1039/c3ob41244j
5. Murayama, K.; Tanaka, Y.; Toda, T.; Kashida, H.; Asanuma, H. *Chem. – Eur. J.* **2013,** *19,* 14151–14158. doi:10.1002/chem.201301578
6. Perlíková, P.; Karlsen, K. K.; Pedersen, E. B.; Wengel, J. *ChemBioChem* **2014,** *15,* 146–156. doi:10.1002/cbic.201300567
7. Li, P.; Sun, J.; Su, M.; Yang, X.; Tang, X. *Org. Biomol. Chem.* **2014,** *12,* 2263–2272. doi:10.1039/c3ob42291g
8. Karri, P.; Punna, V.; Kim, K.; Krishnamurthy, R. *Angew. Chem., Int. Ed.* **2013,** *52,* 5840–5844. doi:10.1002/anie.201300795
9. Krishnamurthy, R. *Synlett* **2014,** *25,* 1511–1517. doi:10.1055/s-0033-1340183
10. Mikhailov, S. N.; Smrt, J. *Collect. Czech. Chem. Commun.* **1975,** *40,* 3080–3085. doi:10.1135/cccc19753080
11. Hiller, S. A.; Zarin, D. E.; Zhuk, R. A. *Nucleic Acids Res.* **1976,** *3,* 721–728. doi:10.1093/nar/3.3.721
12. Capaldi, D. C.; Eleuteri, A.; Chen, Q.; Schinazi, R. F. *Nucleosides Nucleotides* **1997,** *16,* 403–416. doi:10.1080/07328319708001358
13. Cadet, G.; Chan, C.-S.; Daniel, R. Y.; Davis, C. P.; Guiadeen, D.; Rodriguez, G.; Thomas, T.; Walcott, S.; Scheiner, P. *J. Org. Chem.* **1998,** *63,* 4574–4580. doi:10.1021/jo9715231
14. Rosenberg, I.; Holý, A.; Masojídoková, M. *Collect. Czech. Chem. Commun.* **1988,** *53,* 2753–2777. doi:10.1135/cccc19882753
15. Vrbovská, S.; Holý, A.; Pohl, R.; Masojídoková, M. *Collect. Czech. Chem. Commun.* **2006,** *71,* 543–566. doi:10.1135/cccc20060543
16. Lowe, G.; Vilaivan, T. *J. Chem. Soc., Perkin Trans. 1* **1997,** 539–546. doi:10.1039/a603696a
17. Dey, S.; Garner, P. *J. Org. Chem.* **2000,** *65,* 7697–7699. doi:10.1021/jo000983i
18. Porcheddu, A.; Giacomelli, G.; Píredda, I.; Carta, M.; Nieddu, G. *Eur. J. Org. Chem.* **2008,** 5786–5797. doi:10.1002/ejoc.200800891

19. Antle, V. D.; Caperelli, C. A. *Nucleosides Nucleotides* **1999,** *18,* 1911–1928. doi:10.1080/07328319908044853
20. Choo, H.; Chong, Y.; Chu, C. K. *Org. Lett.* **2001,** *3,* 1471–1473. doi:10.1021/ol015783n
21. Shatila, R. S.; Bouhadir, K. H. *Tetrahedron Lett.* **2006,** *47,* 1767–1770. doi:10.1016/j.tetlet.2006.01.035
22. Lu, W.; Sengupta, S.; Petersen, J. L.; Akhmedov, N. G.; Shi, X. *J. Org. Chem.* **2007,** *72,* 5012–5015. doi:10.1021/jo070515+
23. James, K. D.; Ellington, A. D. *Tetrahedron Lett.* **1998,** *39,* 175–178. doi:10.1016/S0040-4039(97)10509-3
24. James, K. D.; von Kiedrowski, G.; Ellington, A. D. *Nucleosides Nucleotides* **1997,** *16,* 1821–1836. doi:10.1080/07328319708002536
25. Pothukanuri, S.; Pianowski, Z.; Winssinger, N. *Eur. J. Org. Chem.* **2008,** 3141–3148. doi:10.1002/ejoc.200800141
26. Lepore, S. D.; He, Y. *J. Org. Chem.* **2003,** *68,* 8261–8263. doi:10.1021/jo0345751
27. Perez-Perez, M.-J.; Rozenski, J.; Busson, R.; Herdewijn, P. *J. Org. Chem.* **1995,** *60,* 1531–1537. doi:10.1021/jo00111a010
28. Yin, X.-q.; Li, W.-k.; Schneller, S. W. *Tetrahedron Lett.* **2006,** *47,* 9187–9189. doi:10.1016/j.tetlet.2006.10.126
29. Jacquemard, U.; Bénéteau, V.; Lefoix, M.; Routier, S.; Mérour, J.-Y.; Coudert, G. *Tetrahedron* **2004,** *60,* 10039–10047. doi:10.1016/j.tet.2004.07.071
30. Sabitha, G.; Reddy, G. S. K. K.; Reddy, K. B.; Reddy, N. M.; Yadav, J. S. *J. Mol. Catal. A: Chem.* **2005,** *238,* 229–232. doi:10.1016/j.molcata.2005.05.028

## License and Terms

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

# Second generation silver(I)-mediated imidazole base pairs

Susanne Hensel, Nicole Megger, Kristina Schweizer and Jens Müller*

Full Research Paper

Open Access

Address:
Institut für Anorganische und Analytische Chemie, Westfälische Wilhelms-Universität Münster, Corrensstr. 28/30, 48149 Münster, Germany

Email:
Jens Müller* - mueller.j@uni-muenster.de

* Corresponding author

Keywords:
bioinorganic chemistry; DNA; imidazole; metal-mediated base pairs; nucleic acids; nucleosides

*Beilstein J. Org. Chem.* **2014,** *10,* 2139–2144.
doi:10.3762/bjoc.10.221

Received: 23 June 2014
Accepted: 27 August 2014
Published: 09 September 2014

This article is part of the Thematic Series "Nucleic acid chemistry".

Guest Editor: H.-A. Wagenknecht



## Abstract

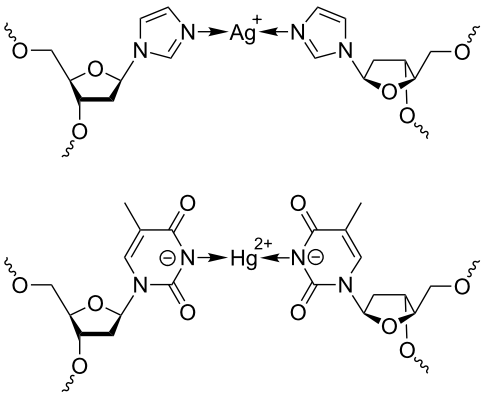
The imidazole–Ag(I)–imidazole base pair is one of the best-investigated artificial metal-mediated base pairs. We show here that its stability can be further improved by formally replacing the imidazole moiety by a 2-methylimidazole or 4-methylimidazole moiety. A comparison of the thermal stability of several double helices shows that the addition of one equivalent of Ag(I) leads to a 50% larger increase in the melting temperature when a DNA duplex with methylated imidazole nucleosides is applied. This significant effect can likely be attributed to a better steric shielding of the metal ion within the metal-mediated base pair.

## Introduction

Nucleic acids with their evolutionary optimized self-assembling properties represent an ideal basis for the generation of artificial, site-specifically functionalized supramolecular aggregates [1-4]. In this context, numerous functional moieties have been introduced into various nucleic acid building blocks, including, for example, modified nucleobases and modified sugar entities. A recent addition is the application of ligand systems as nucleobases surrogates [5-9]. In the presence of suitable transition metal ions, so-called metal-mediated base pairs between two complementary ligand-derived artificial nucleobases can be formed. Accordingly, metal-based functionality can be introduced site-specifically into nucleic acids. Various applications have been reported for such metal-modified nucleic acids [5], including increased charge transfer capability [10,11] and DNA-based logic gates [12,13]. Applying the concepts of coordination chemistry, a wide range of metal-specific ligands have been incorporated successfully into nucleic acids and nucleic acid derivatives. Most recently, even coordination patterns of the [2 + 1] type could be obtained [14,15]. While most examples focus on the use of DNA, other systems such as RNA [16], GNA [17,18] (glycol nucleic acid), PNA [19-21], and other nucleic acid derivatives [22] have been modified as well. It is interesting to note that metal-mediated base pairs do not necessarily require the presence of artificial ligands. In contrast, natural nucleobases have also been successfully applied in the generation of metal-mediated base pairs such as thyminate–Hg(II)–thyminate or cytosine–Ag(I)–cytosine [23-25]. It has even been shown that polymerases are capable of processing metal-mediated base pairs [26,27], resulting, e.g., in the formation of cytosine–Ag(I)–adenine base pairs [28]. Struc-

tural analyses of short oligonucleotide duplexes comprising one or more metal-mediated base pairs indicate that the large conformational space of nucleic acids exists also for metal-modified nucleic acids. In particular, the existence of B-DNA- and Z-DNA-type conformations have been reported in this context [29-33].

The imidazole–Ag(I)–imidazole base pair represents one of the best investigated metal-mediated base pairs to date [29,30,34,35]. It comprises a linearly coordinated silver(I) ion inserted in-between two complementary artificial imidazole nucleosides (Scheme 1, top). The NMR-based solution structure of a DNA duplex with three contiguous imidazole–Ag(I)–imidazole represented the first experimental proof that a B-DNA conformation is compatible with the presence of metal-mediated base pairs, as only minor distortions at the base pair steps between natural and artificial base pairs were observed [29,30]. Additional experiments have shown that neighbouring imidazole–Ag(I)–imidazole base pairs are formed cooperatively, suggesting the possibility of creating long contiguous metal-containing sections within oligonucleotide duplexes [35]. A closer inspection of the experimental NMR structure reveals that the silver(I) ions, despite being lined up inside the duplex along the helical axis, still appear to be accessible from the outside of the duplex (Figure 1). This contrasts the situation in a duplex with thyminate–Hg(II)–thyminate base pairs (Scheme 1, bottom) [31,32]. As a result of the keto groups on the thymine residues, the mercury(II) ions included in this type of metal-mediated base pair are clearly shielded from outside access [31]. In an attempt to increase the stability of the imidazole–Ag(I)–imidazole base pairs, additional shielding of the central silver(I) ions could be accomplished by introducing methyl groups on the imidazole moieties. A comparison of the p$K_a$ values [36] of 1-methylimidazole (p$K_a$ = 7.21), 1,2-dimethylimidazole (p$K_a$ = 8.22) and 1,4-dimethylimidazole (p$K_a$ = 7.75) indicates that this modification is also expected to slightly increase the basicity of the ligand, thereby also contributing to the formation of a more stable metal complex.




**Scheme 1:** Schematic representation of an imidazole–Ag(I)–imidazole base pair (top) and a thyminate–Hg(II)–thyminate base pair (bottom).

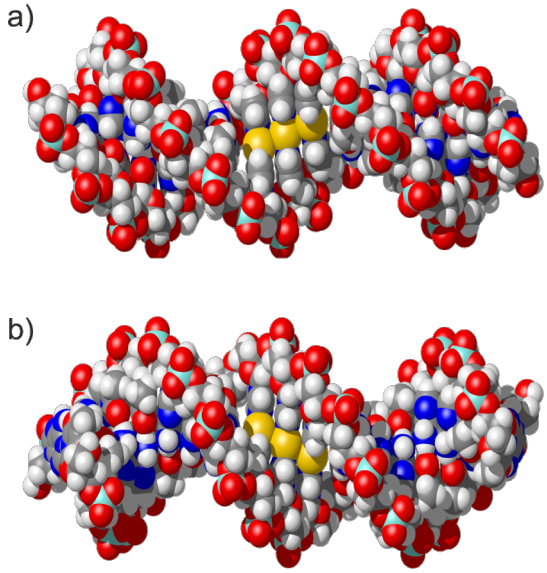
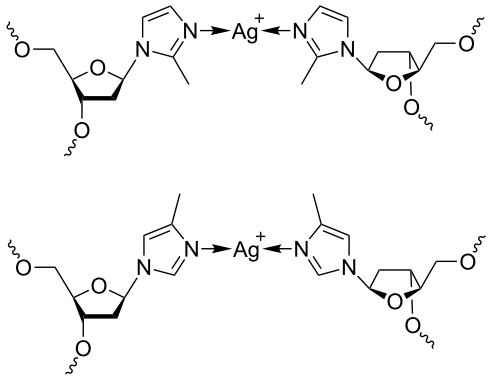



**Figure 1:** Different views of a DNA duplex comprising three neighbouring imidazole–Ag(I)–imidazole base pairs. a) View from the major groove; b) view from the minor groove. The silver(I) ions shown in yellow are clearly visible and appear to be solvent accessible. Coordinates have been taken from PDB 2M54. This figure was prepared using MolMol [37].

# Results and Discussion

## Synthesis and characterization of the nucleosides

Scheme 2 shows silver(I)-mediated methylimidazole homo base pairs involving either 2-methylimidazole (top) or 4-methylimidazole (bottom), resulting in a shielded access to the silver(I) ions from the minor and major groove, respectively. The phosphoramidites required for automated DNA solid-phase synthesis were obtained using synthetic procedures established previously for the non-methylated imidazole analogue (Scheme 3) [34,38]. Methylimidazole, deprotonated in situ via the addition of sodium hydride, has been reacted with Hoffer's chloro sugar to give the *p*-toluoyl-protected nucleoside (**1a**, **1b**). In the case of 2-methylimidazole, only one product (2-methylimidazole nucleoside, **1a**) was formed due to the symmetry of the deprotonated starting compound. When performing this reaction with 4-methylimidazole, two product isomers were obtained as a result of the asymmetric substitution of the deprotonated starting compound, namely 4-methylimidazole nucleoside (**1b**) and 5-methylimidazole nucleoside. The latter isomer was obtained in very low yield only (isomer ratio of ~15:1). It was not further investigated, because its methyl group is oriented in a direction that is not expected to lead to an



**Scheme 2:** Schematic representation of silver(I)-mediated imidazole homo base pairs involving 2-methylimidazole (top) and 4-methylimidazole (bottom), designed to shield access to the silver(I) ions from minor and major groove, respectively.

increased shielding of the silver(I) ions in the corresponding metal-mediated base pairs. As a result, all further experiments were performed using either **1a** or **1b** only. Deprotection of **1a**/**1b** by means of aqueous ammonia in methanol led to the free nucleosides **2a**/**2b**. The next two steps involved the orthogonal protection of the two hydroxy functions with the dimethoxytrityl and the 2-cyanoethyl-*N*,*N*-diisopropylphosphoramidite moieties, finally resulting in the formation of **4a**/**4b** suitable for automated solid-phase oligonucleotide synthesis.

The p$K_a$ values of the free nucleosides **2a**/**2b** were determined by pD-dependent $^1$H NMR spectroscopy. They amount to 6.61 and 6.50, respectively. As expected, the methylimidazole nucleosides are slightly more basic than the corresponding non-methylated parent nucleoside (p$K_a$ = 6.01) [34].

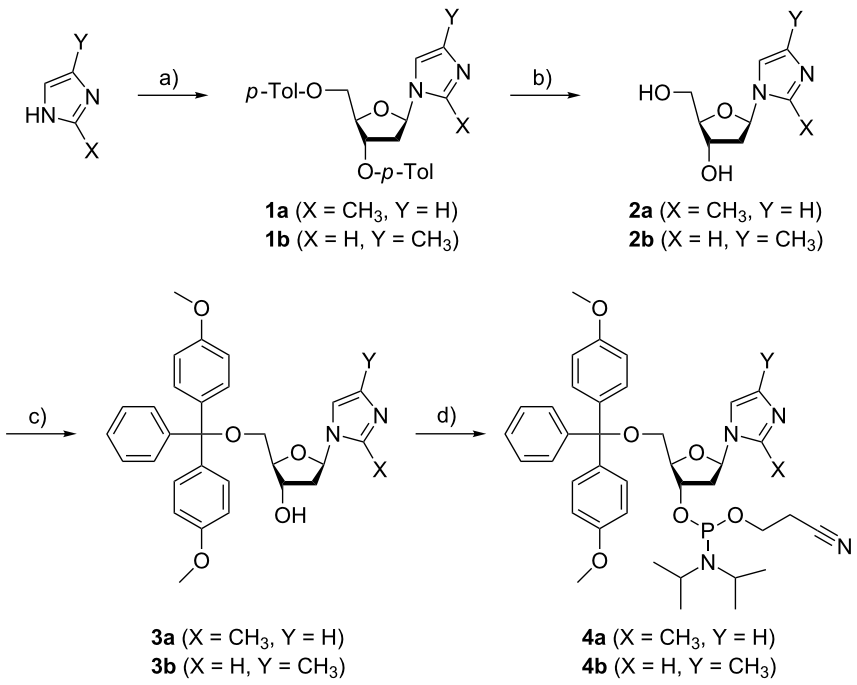
## Investigation of the oligonucleotides

For each of the artificial nucleosides, two DNA double helices were synthesized comprising either one or two central methylimidazole:methylimidazole base pairs (Scheme 4). The

I 5'-d(TTT GTT TGT TTG XTT GTT TTT TTT TT)-3'
3'-d(AAA CAA ACA AAC XAA CAA AAA AAA AA)-5'

II 5'-d(TTT GTT TGT TTG XXT GTT TTT TTT TT)-3'
3'-d(AAA CAA ACA AAC XXA CAA AAA AAA AA)-5'

**Scheme 4:** Oligonucleotide sequences under investigation (X = 2-methylimidazole **2a** or 4-methylimidazole **2b**).




**Scheme 3:** Synthesis of methylimidazole-based nucleosides and their corresponding phosphoramidites required for automated solid-phase DNA synthesis. a) 1. NaH, 2. Hoffer's chloro sugar, $CH_3CN$, 0 °C to ambient temperature, 3 h; b) aqueous $NH_3$ (25%), $CH_3OH$, ambient temperature, 16 h; c) dimethoxytrityl chloride, 4-dimethylaminopyridine, pyridine, ambient temperature, 3 h; d) 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite, diisopropylethylamine, $CH_2Cl_2$, ambient temperature, 30 min.

sequences were chosen because they match those that had thoroughly been investigated previously with non-methylated imidazole nucleosides [35], enabling a direct observation of the expected shielding effect. Figure 2 shows the melting curves of two of the double helices, determined by temperature-dependent UV spectroscopy (for the other two duplexes, see Supporting Information File 1). As can clearly be seen, the addition of one equivalent [39] of Ag(I) leads to a significant increase in the melting temperature $T_m$. Table 1 lists the experimentally determined melting temperatures. As expected, the double helix comprising two neighbouring methylimidazole:methylimidazole mispairs displays a lower $T_m$ compared to the duplex comprising one mispair only. It is interesting to note that even in the absence of Ag(I) the duplexes containing the methylated imidazole nucleosides melt at a significantly higher $T_m$ (by about 3–4 °C), possibly due to an additional stabilising hydrophobic interaction in the groove. The position of the methyl group (2-methylimidazole vs 4-methylimidazole) does not influence the absolute melting temperature. In the presence of one equivalent of Ag(I), the thermal stability of the duplexes increases significantly. Notably, the increase in $T_m$ is larger by about 50% for the methylated imidazole nucleosides compared with the non-methylated one, amounting to about 8.5 °C per silver(I)-mediated base pair. The significantly increased thermal stability of the duplexes comprising either 2-methylimidazole or 4-methylimidazole therefore cannot be attributed to the above-mentioned hydrophobic interaction only but must also be a direct effect of the formation of the metal-mediated base pairs. This indicates that the concept of increasing the stability of silver(I)-mediated imidazole base pairs by restricting the access to the central metal ion via introduction of additional methyl groups on the ligand is successful.

The CD spectra shown in Figure 3 (and Supporting Information File 1) are in excellent agreement with those recorded

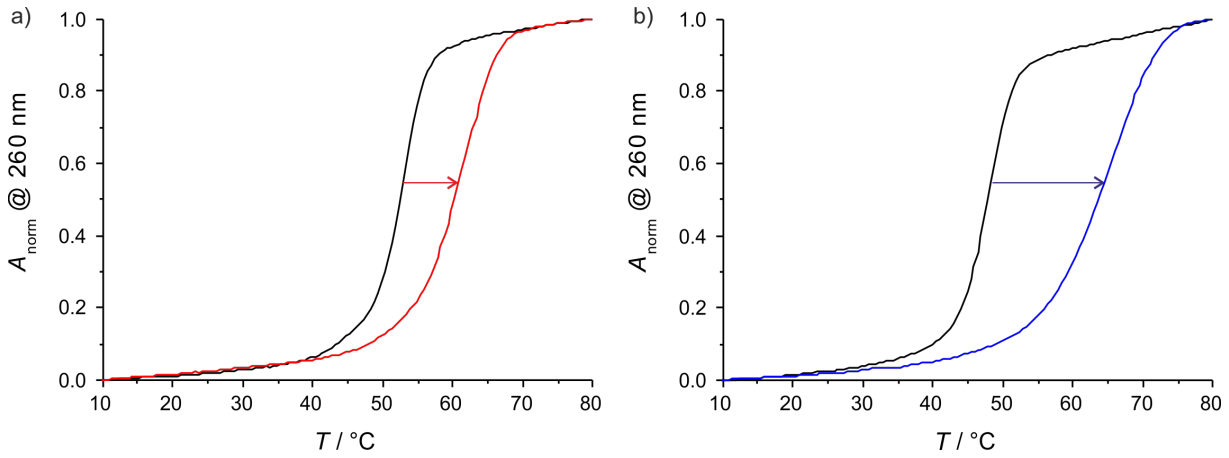



**Figure 2:** Melting curves based on normalized UV absorbance at 260 nm of a) duplex **I** with X = 4-methylimidazole and b) duplex **II** with X = 2-methylimidazole in the absence (black) and presence of one equivalent of Ag(I) (coloured). For the sequences, see Scheme 4. Experimental conditions: 1 μM duplex, 150 mM $NaClO_4$, 5 mM MOPS (pH 6.8).

**Table 1:** Absolute and relative melting temperatures $T_m$ (°C) of DNA duplexes containing one or two (methyl)imidazole:(methyl)imidazole base pairs in the absence and presence of silver(I) (**Im** = imidazole; **2** = 2-methylimidazole; **4** = 4-methylimidazole; for sequences, see Scheme 4). The estimated standard deviation of $T_m$ amounts to 1 °C. Experimental conditions: 1 μM duplex, 150 mM $NaClO_4$, 5 mM MOPS (pH 6.8).

| artificial nucleoside X | $T_m$ (no Ag(I)) | $T_m$ (+ Ag(I)) | $\Delta T_m$ (0 → 1 equiv Ag(I)) |
|---|---|---|---|
| **Im**:**Im**[a] | 49.5 | 55.5 | + 6.0 |
| **2**:**2** | 53.0 | 62.0 | + 9.0 |
| **4**:**4** | 53.3 | 61.3 | + 8.0 |
| **ImIm**:**ImIm**[a] | 44.5 | 55.5 | + 11.0 |
| **22**:**22** | 48.4 | 64.9 | + 16.5 |
| **44**:**44** | 48.2 | 65.7 | + 17.5 |

[a]Data taken from [35].

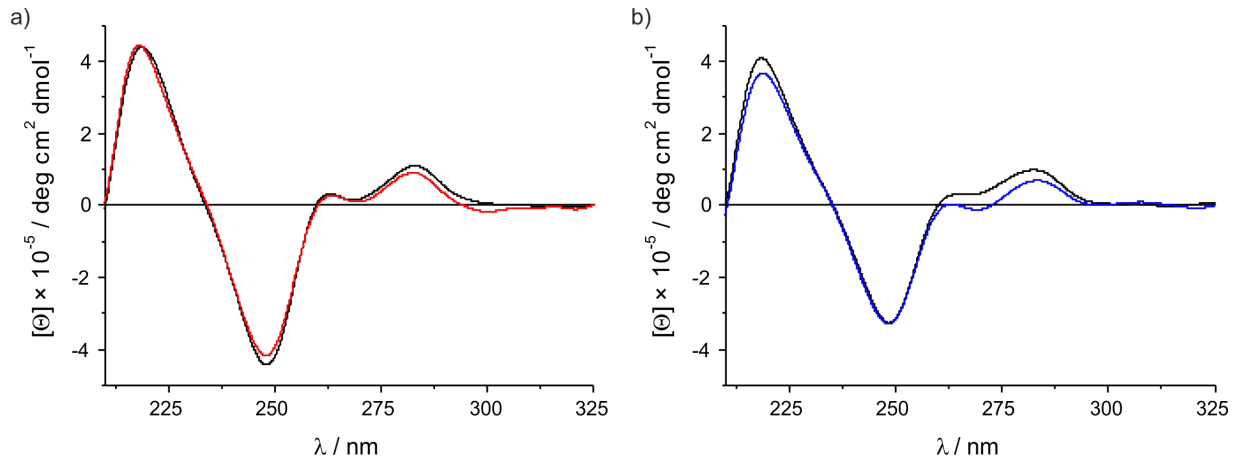



**Figure 3:** CD spectra of a) duplex **I** with X = 4-methylimidazole and b) duplex **II** with X = 2-methylimidazole in the absence (black) and presence of one equivalent of Ag(I) (coloured). For the sequences, see Scheme 4. Experimental conditions: 1 μM duplex, 150 mM $NaClO_4$, 5 mM MOPS (pH 6.8).

previously for the analogous duplexes containing the non-methylated imidazole nucleoside. They indicate that the duplexes belong to the B-DNA family. The low intensity of the positive Cotton effects at about 262 nm and 283 nm is typical for duplexes with this type of sequence and has, for example, also been reported for poly[d(A)]·poly[d(T)] [40]. The presence of Ag(I) does not alter the wavelengths of the Cotton effects and only marginally affects the ellipticity at a given wavelength. Taken together, it can be concluded that neither the introduction of a methyl group nor the formation of the silver(I)-mediated base pair does influence the overall DNA duplex conformation. As the synthesis of 2-methylimidazole nucleoside from 2-methylimidazole yields one isomer only whereas the synthesis of 4-methylimidazole nucleoside from 4-methylimidazole also gives 5-methylimidazole as an undesired side product, it is recommended to replace imidazole by 2-methylimidazole in future research on imidazole–Ag(I)–imidazole base pairs to obtain nucleic acid systems of superior stability with otherwise identical properties.

## Conclusion

The formal substitution of an imidazole moiety in DNA duplexes comprising silver(I)-mediated imidazole–Ag(I)–imidazole base pairs by a 2-methylimidazole or 4-methylimidazole moiety leads to a significant increase in thermal stability of the nucleic acid. The increase in $T_m$ upon the addition of one equivalent of Ag(I) rises by about 50% for the methylated nucleosides compared to the non-methylated one. The molecular structure of a DNA duplex with imidazole–Ag(I)–imidazole base pairs supports the notion that the increased stability may be due to a better shielding of the Ag(I) from the surrounding solvent.

## Supporting Information

UV melting curves and CD spectra of duplex **I** with X = 2-methylimidazole and of duplex **II** with X = 4-methylimidazole. Full experimental details and characterization data of the compounds shown in Scheme 3 and Scheme 4.

### Supporting Information File 1

Experimental data.
[http://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-10-221-S1.pdf]

## Acknowledgements

Financial support by the Deutsche Forschungsgemeinschaft (SFB 858) is gratefully acknowledged.

## References

1. Bandy, T. J.; Brewer, A.; Burns, J. R.; Marth, G.; Nguyen, T.; Stulz, E. *Chem. Soc. Rev.* **2011,** *40,* 138–148. doi:10.1039/b820255a
2. Varghese, R.; Wagenknecht, H.-A. *Chem. Commun.* **2009,** 2615–2624. doi:10.1039/b821728a
3. Schreiber, R.; Do, J.; Roller, E.-M.; Zhang, T.; Schüller, V. J.; Nickels, P. C.; Feldmann, J.; Liedl, T. *Nat. Nanotechnol.* **2014,** *9,* 74–78. doi:10.1038/nnano.2013.253
4. Wilner, O. I.; Willner, I. *Chem. Rev.* **2012,** *112,* 2528–2556. doi:10.1021/cr200104q
5. Scharf, P.; Müller, J. *ChemPlusChem* **2013,** *78,* 20–34. doi:10.1002/cplu.201200256
6. Müller, J. *Eur. J. Inorg. Chem.* **2008,** 3749–3763. doi:10.1002/ejic.200800301
7. Clever, G. H.; Shionoya, M. *Met. Ions Life Sci.* **2012,** *10,* 269–294. doi:10.1007/978-94-007-2172-2_10

8. Takezawa, Y.; Shionoya, M. *Acc. Chem. Res.* **2012,** *45,* 2066–2076. doi:10.1021/ar200313h
9. Clever, G. H.; Kaul, C.; Carell, T. *Angew. Chem., Int. Ed.* **2007,** *46,* 6226–6236. doi:10.1002/anie.200701185
10. Liu, S.; Clever, G. H.; Takezawa, Y.; Kaneko, M.; Tanaka, K.; Guo, X.; Shionoya, M. *Angew. Chem., Int. Ed.* **2011,** *50,* 8886–8890. doi:10.1002/anie.201102980
11. Ehrenschwender, T.; Schmucker, W.; Wellner, C.; Augenstein, T.; Carl, P.; Harmer, J.; Breher, F.; Wagenknecht, H.-A. *Chem. – Eur. J.* **2013,** *19,* 12547–12552. doi:10.1002/chem.201300593
12. Carell, T. *Nature* **2011,** *469,* 45–46. doi:10.1038/469045a
13. Park, K. S.; Jung, C.; Park, H. G. *Angew. Chem., Int. Ed.* **2010,** *49,* 9757–9760. doi:10.1002/anie.201004406
14. Richters, T.; Müller, J. *Eur. J. Inorg. Chem.* **2014,** 437–441. doi:10.1002/ejic.201301491
15. Richters, T.; Krug, O.; Kösters, J.; Hepp, A.; Müller, J. *Chem. – Eur. J.* **2014,** *20,* 7811–7818. doi:10.1002/chem.201402221
16. Johannsen, S.; Paulus, S.; Düpre, N.; Müller, J.; Sigel, R. K. O. *J. Inorg. Biochem.* **2008,** *102,* 1141–1151. doi:10.1016/j.jinorgbio.2007.12.023
17. Schlegel, M. K.; Zhang, L.; Pagano, N.; Meggers, E. *Org. Biomol. Chem.* **2009,** *7,* 476–482. doi:10.1039/b816142a
18. Zhang, L.; Meggers, E. *J. Am. Chem. Soc.* **2005,** *127,* 74–75. doi:10.1021/ja043904j
19. Ma, Z.; Olechnowicz, F.; Skorik, A. Y.; Achim, C. *Inorg. Chem.* **2011,** *50,* 6083–6092. doi:10.1021/ic200138b
20. Franzini, R. M.; Watson, R. M.; Patra, G. K.; Breece, R. M.; Tierney, D. L.; Hendrich, M. P.; Achim, C. *Inorg. Chem.* **2006,** *45,* 9798–9811. doi:10.1021/ic0609610
21. Küsel, A.; Zhang, J.; Alvariño Gil, M.; Stückl, A. C.; Meyer-Klaucke, W.; Meyer, F.; Diederichsen, U. *Eur. J. Inorg. Chem.* **2005,** 4317–4324. doi:10.1002/ejic.200500464
22. Taherpour, S.; Lönnberg, H.; Lönnberg, T. *Org. Biomol. Chem.* **2013,** *11,* 991–1000. doi:10.1039/c2ob26885j
23. Ono, A.; Torigoe, H.; Tanaka, Y.; Okamoto, I. *Chem. Soc. Rev.* **2011,** *40,* 5855–5866. doi:10.1039/c1cs15149e
24. Megger, D. A.; Megger, N.; Müller, J. *Met. Ions Life Sci.* **2012,** *10,* 295–317. doi:10.1007/978-94-007-2172-2_11
25. Megger, D. A.; Fonseca Guerra, C.; Bickelhaupt, F. M.; Müller, J. *J. Inorg. Biochem.* **2011,** *105,* 1398–1404. doi:10.1016/j.jinorgbio.2011.07.005
26. Urata, H.; Yamaguchi, E.; Funai, T.; Matsumura, Y.; Wada, S.-i. *Angew. Chem., Int. Ed.* **2010,** *49,* 6516–6519. doi:10.1002/anie.201002142
27. Kaul, C.; Müller, M.; Wagner, M.; Schneider, S.; Carell, T. *Nat. Chem.* **2011,** *3,* 794–800. doi:10.1038/nchem.1117
28. Funai, T.; Miyazaki, Y.; Aotani, M.; Yamaguchi, E.; Nakagawa, O.; Wada, S.-i.; Torigoe, H.; Ono, A.; Urata, H. *Angew. Chem., Int. Ed.* **2012,** *51,* 6464–6466. doi:10.1002/anie.201109191
29. Johannsen, S.; Megger, N.; Böhme, D.; Sigel, R. K. O.; Müller, J. *Nat. Chem.* **2010,** *2,* 229–234. doi:10.1038/nchem.512
30. Kumbhar, S.; Johannsen, S.; Sigel, R. K. O.; Waller, M. P.; Müller, J. *J. Inorg. Biochem.* **2013,** *127,* 203–210. doi:10.1016/j.jinorgbio.2013.03.009
31. Yamaguchi, H.; Šebera, J.; Kondo, J.; Oda, S.; Komuro, T.; Kawamura, T.; Dairaku, T.; Kondo, Y.; Okamoto, O.; Ono, A.; Burda, J. V.; Kojima, C.; Sychrovský, V.; Tanaka, Y. *Nucleic Acids Res.* **2014,** *42,* 4094–4099. doi:10.1093/nar/gkt1344
32. Kondo, Y.; Yamada, T.; Hirose, C.; Okamoto, I.; Tanaka, Y.; Ono, A. *Angew. Chem., Int. Ed.* **2014,** *53,* 2385–2388. doi:10.1002/anie.201309066
33. Atwell, S.; Meggers, E.; Spraggon, G.; Schultz, P. G. *J. Am. Chem. Soc.* **2001,** *123,* 12364–12367. doi:10.1021/ja011822e
34. Müller, J.; Böhme, D.; Lax, P.; Morell Cerdà, M.; Roitzsch, M. *Chem. – Eur. J.* **2005,** *11,* 6246–6253. doi:10.1002/chem.200500511
35. Petrovec, K.; Ravoo, B. J.; Müller, J. *Chem. Commun.* **2012,** *48,* 11844–11846. doi:10.1039/c2cc36169h
36. Lenarcik, B.; Ojczenasz, P. *J. Heterocycl. Chem.* **2002,** *39,* 287–290. doi:10.1002/jhet.5570390206
37. Koradi, R.; Billeter, M.; Wüthrich, K. *J. Mol. Graphics* **1996,** *14,* 51–55. doi:10.1016/0263-7855(96)00009-4
38. Durland, R. H.; Rao, T. S.; Bodepudi, V.; Seth, D. M.; Jayaraman, K.; Revankar, G. R. *Nucleic Acids Res.* **1995,** *23,* 647–653. doi:10.1093/nar/23.4.647
39. Throughout this work, one equivalent of Ag(I) is defined as the number of Ag(I) ions that are required to saturate all imidazole:imidazole mispairs. Hence, one equivalent of Ag(I) amounts to 1 Ag(I) for duplex I and to 2 Ag(I) for duplex II.
40. Vorlíčková, M.; Kejnovská, I.; Bednářová, K.; Renčiuk, D.; Kypr, J. *Chirality* **2012,** *24,* 691–698. doi:10.1002/chir.22064


## License and Terms

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

# Synthesis and optical properties of pyrrolidinyl peptide nucleic acid carrying a clicked Nile red label

Nattawut Yotapan[1], Chayan Charoenpakdee[1], Pawinee Wathanathavorn[1], Boonsong Ditmangklo[1], Hans-Achim Wagenknecht[*2] and Tirayut Vilaivan[*1]

Full Research Paper

Open Access

Address:
[1]Organic Synthesis Research Unit, Department of Chemistry, Faculty of Science, Chulalongkorn University, Phayathai Road, Patumwan, Bangkok 10330, Thailand and [2]Institute of Organic Chemistry, Karlsruhe Institute of Technology (KIT), Fritz-Haber-Weg 6, 76131 Karlsruhe, Germany

Email:
Hans-Achim Wagenknecht[*] - wagenknecht@kit.edu;
Tirayut Vilaivan[*] - vtirayut@chula.ac.th

* Corresponding author

Keywords:
click chemistry; deoxyribonucleic acid; DNA bulge; fluorescence; nucleic acids; solvatochromism

*Beilstein J. Org. Chem.* **2014,** *10,* 2166–2174.
doi:10.3762/bjoc.10.224

Received: 08 May 2014
Accepted: 19 August 2014
Published: 11 September 2014

This article is part of the Thematic Series "Nucleic acid chemistry".

Associate Editor: S. Flitsch



## Abstract

DNA or its analogues with an environment-sensitive fluorescent label are potentially useful as a probe for studying the structure and dynamics of nucleic acids. In this work, pyrrolidinyl peptide nucleic acid (acpcPNA) was labeled at its backbone with Nile red, a solvatochromic benzophenoxazine dye, by means of click chemistry. The optical properties of the Nile red-labeled acpcPNA were investigated by UV–vis and fluorescence spectroscopy in the absence and in the presence of DNA. In contrast to the usual quenching observed in Nile red-labeled DNA, the hybridization with DNA resulted in blue shifting and an enhanced fluorescence regardless of the neighboring bases. More pronounced blue shifts and fluorescence enhancements were observed when the DNA target carried a base insertion in close proximity to the Nile red label. The results indicate that the Nile red label is located in a more hydrophobic environment in acpcPNA–DNA duplexes than in the single-stranded acpcPNA. The different fluorescence properties of the acpcPNA hybrids of complementary DNA and DNA carrying a base insertion are suggestive of different interactions between the Nile red label and the duplexes.

## Introduction

Fluorescent labels are important tools for investigating the structure and dynamics of biomolecular interactions [1-3]. Traditionally, the biological macromolecules are labeled with two or more dyes which can interact in a conformation/distant-dependent manner via Förster resonance energy transfer (FRET) [4-6]. Alternatively, the FRET pairs can be replaced by an environmentally sensitive label that can change its fluorescence in response to its altered micro-environment [7-9]. Nile

red is a member of the benzophenoxazine dye family which exhibits several interesting features including a high photostability, high fluorescence quantum yield, broad working pH range, long excitation and emission wavelengths, and solvatochromic properties [10]. Applications of Nile red as a staining dye in histology [11,12] and as a probe for the sensing of polarity and hydrophobicity [13] are well-known. Nile red has been used as a DNA label, either as a base modifier [14-16], a base replacement [17] or a backbone-tethered label [18,19]. However, in most cases the formation of DNA duplexes does not yield significant fluorescence changes in the Nile red, unless it is used in combination with another dye such as pyrene to form an energy transfer pair [20,21]. A related phenoxazine dye – Nile blue – has also been incorporated into DNA as a base replacement, again without showing significant structure-induced fluorescence change [22].

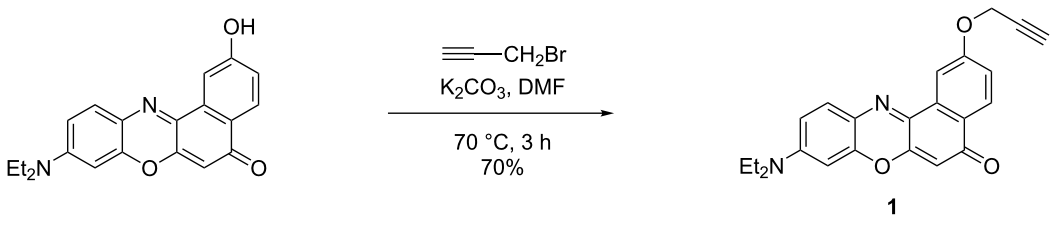
Peptide nucleic acid or PNA is an electrostatically neutral analogue of DNA which can form very stable duplexes with DNA and RNA in a highly sequence specific fashion. PNA–DNA duplexes have different structural morphology and electrostatic potential surfaces from DNA–DNA duplexes and therefore they are interacting differently with DNA-binding dyes [23-25]. We had recently introduced a new conformationally constrained pyrrolidinyl PNA known as acpcPNA that shows several unique properties [26-28]. Its potential applications as a probe for DNA sensing are well-established. For example, we have developed singly labeled acpcPNA probes that can give a fluorescence change in response to hybridization to DNA [29-31]. Due to its solvatochromic properties, Nile red is a potential candidate to be used in combination with PNA to develop a hybridization probe which not only differentiates between complementary and non-complementary DNA, but can also report on local structural variations. Except for an example from our group on the conformation control of Nile red-labeled DNA by acpcPNA [32], no combination of the Nile red label and PNA has yet been reported in the literature. It is therefore our purpose to develop a facile method for synthesizing Nile red-labeled acpcPNA based on the click strategy [33-36] and to investigate its optical properties.

## Results and Discussion

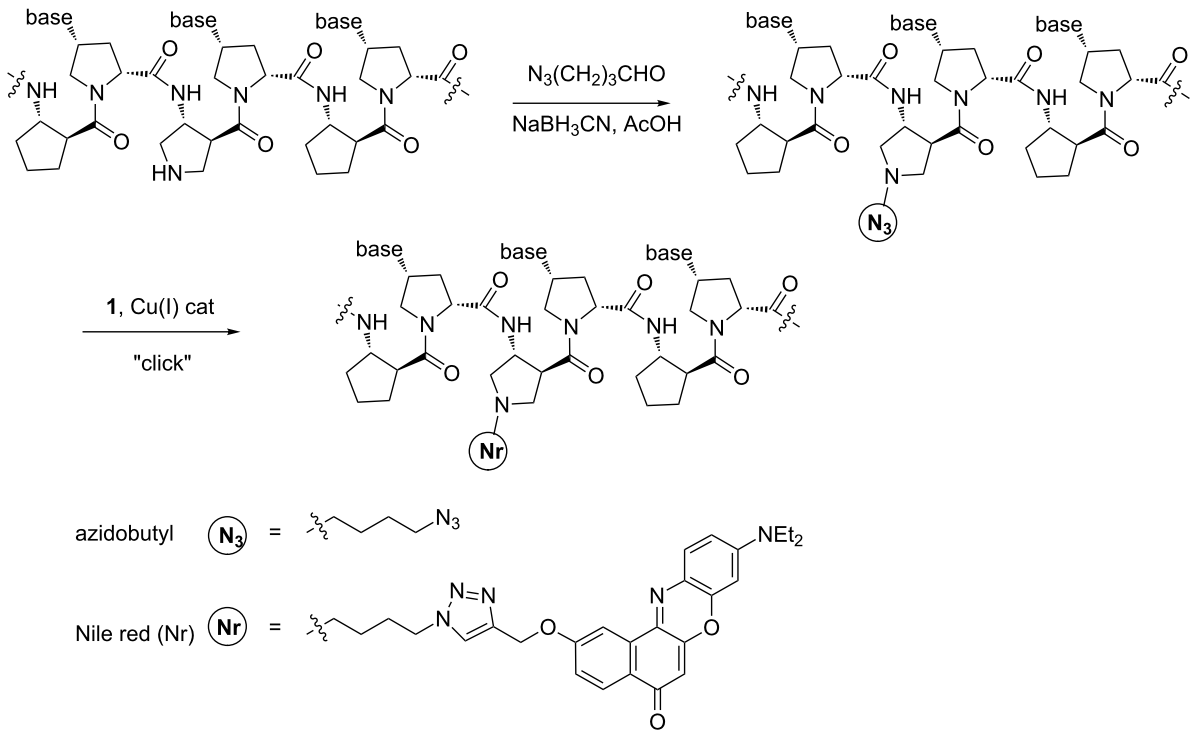
At least two different means to introduce the Nile red label onto DNA by using click chemistry have been reported in the literature. One involves the clicking of in situ generated 5-azido-deoxyuridine-containing DNA with propargyl Nile red [16], the other employs DNA bearing 2'-propargylated nucleotides and azide-modified Nile red [19]. An approach related to the latter was used in this work, although the azide function was placed on the PNA instead of the Nile red, and the clicking was carried out on the solid support rather than in solution phase. The propargylated Nile red label **1** was synthesized in 70% yield by alkylation of the known 2-hydroxy Nile red [37,38] with propargyl bromide in the presence of $K_2CO_3$ in DMF (Scheme 1).

The alkyne-containing compound **1** was clicked onto the backbone of acpcPNA that had been pre-functionalized with an azidobutyl group by a reductive alkylation strategy previously reported by our group (Scheme 2) [31]. Two lysine residues were incorporated at the N- and C-termini of the acpcPNA to ensure a sufficient solubility in aqueous solution. Five acpcPNA sequences, each of which singly labeled at the backbone with Nile red under different sequence context, were successfully synthesized and characterized by MALDI–TOF MS (Figure 1 and Table 1). Isolated yields in the range of 6–18% were obtained (0.5 μmol scale), which are typical for solid phase synthesis, whereby the majority of material loss occurred during HPLC purification. All Nile red-labeled acpcPNAs are freely soluble in water (>1 mM), providing a bright blue solution.

The UV–vis spectrum of the free propargyl Nile red label **1** in acetonitrile (Figure 2a) showed absorption and fluorescence emission maxima at 538 nm and 620 nm, with a fluorescence quantum yield ($\Phi_F$) of 0.65, which is quite comparable to other Nile red derivatives reported in the literature [10]. As the polarity of the solvent increases with the addition of aqueous phosphate buffer, both the absorption and the fluorescence maxima shifted to longer wavelengths with a concomitant decrease in the fluorescence quantum yields (50% MeCN: $\lambda_{abs}$ 564 nm, $\lambda_{em}$ 641 nm, $\Phi_F$ 0.29; 20% MeCN: $\lambda_{abs}$ 587 nm, $\lambda_{em}$




**Scheme 1:** Synthesis of propargylated Nile red **1**.




**Scheme 2:** Synthesis of azidobutyl- and Nile red-modified acpcPNA.

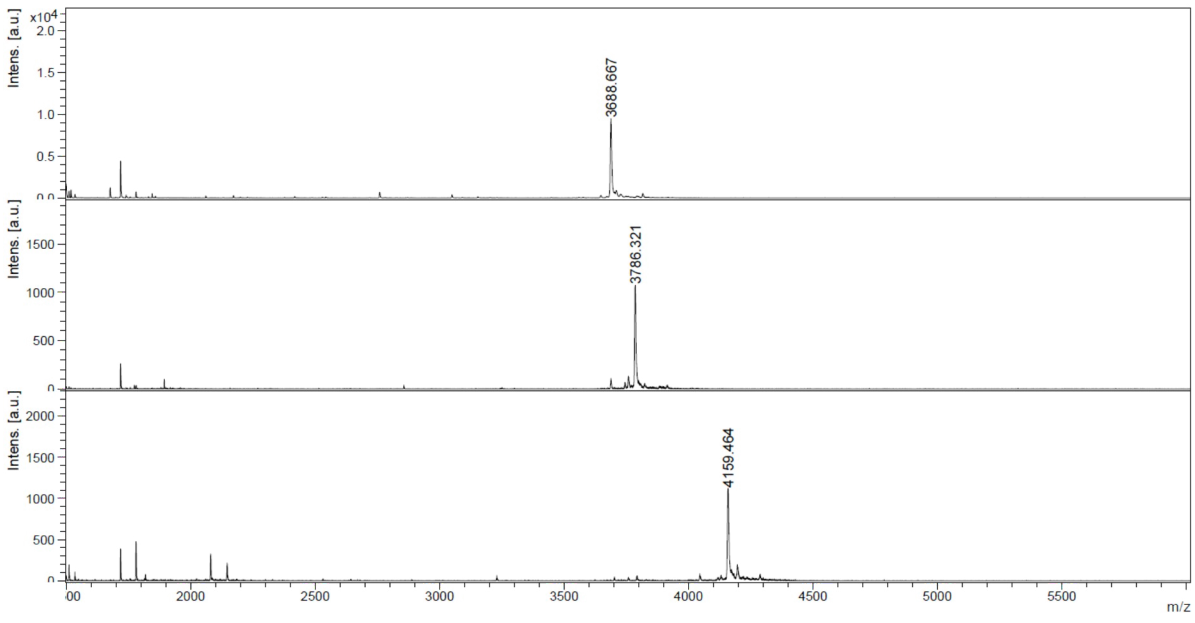



**Figure 1:** MALDI–TOF mass spectra of the crude 10mer acpcPNA before (top) (calcd *m*/*z* 3688.0), and after functionalizing with the azidobutyl group (middle) (calcd *m*/*z* 3785.1), and followed by clicking with Nile red (bottom) (calcd *m*/*z* 4157.5).

651 nm; $\Phi_F$ 0.14). In aqueous phosphate buffer (10 mM, pH 7.0), the Nile red-labeled acpcPNA **10mer-Nr** exhibited a broad absorption peak centered at 575 nm and a fluorescence emission at 656 nm, respectively (Figure 2b). The solvatochromic property of the Nile red-labeled acpcPNA is demonstrated as shown by the progressively blue-shifted absorption

**Table 1:** Sequence, isolated yield and characterization data of Nile red-labeled acpcPNA.

| PNA | sequence (N to C)[a] | $t_R$ (min)[b] | % yield[c] | *m*/*z*[d] (calcd) | *m*/*z*[e] (found) |
|---|---|---|---|---|---|
| **10mer-Nr** | GTAGA(Nr)TCACT | 33.6 | 6.3 | 4157.5 | 4155.1 |
| **11merAA-Nr** | CATAA(Nr)AATACG | 34.2 | 18.5 | 4491.9 | 4491.1 |
| **11merCC-Nr** | CATAC(Nr)CATACG | 34.6 | 11.4 | 4443.9 | 4441.8 |
| **11merGG-Nr** | CATAG(Nr)GATACG | 32.8 | 15.2 | 4523.9 | 4523.2 |
| **11merTT-Nr** | CATAT(Nr)TATACG | 34.3 | 8.4 | 4473.9 | 4473.0 |

[a]All sequences were end-capped at N- and C-termini with *N*-acetyl-L-lysine and L-lysinamide, respectively. [b]HPLC conditions: C18 column 4.6 × 50 mm, 3 µ, gradient 0.1% TFA in $H_2O$:MeOH 90:10 for 5 min then linear gradient to 10:90 over 30 min, flow rate 0.5 mL/min, 260 nm. [c]Isolated yield (determined spectrophotometrically) after HPLC purification. [d]Average mass of $[M + H]^+$. [e]MALDI–TOF.

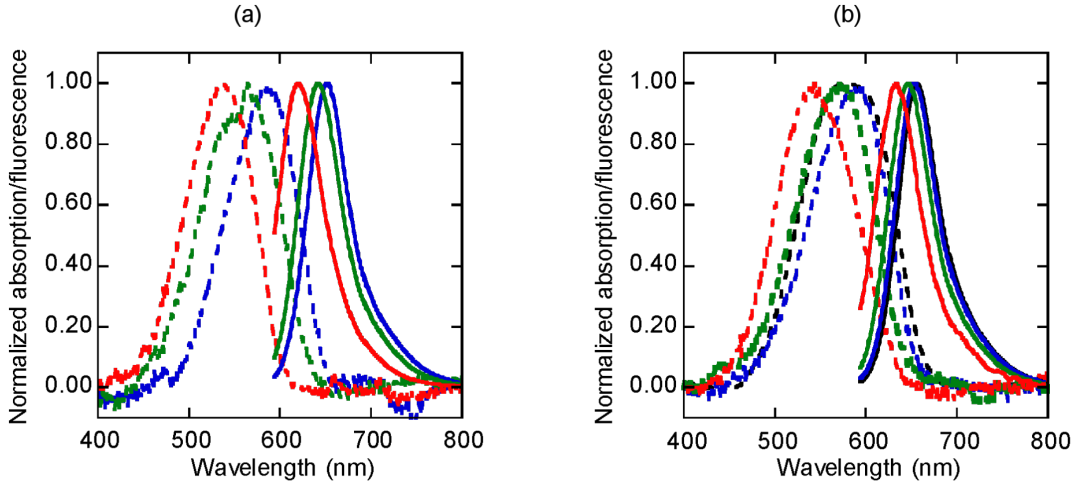



**Figure 2:** Normalized absorption (---) and fluorescence (—) spectra of (a) propargyl Nile red **1** and (b) Nile red-labeled acpcPNA **10mer-Nr** in 10 mM sodium phosphate buffer (pH 7.0):acetonitrile: black = 100:0; blue = 80:20; green = 50:50; red = 0:100. All measurements were carried out in 10 mM sodium phosphate buffer (pH 7.0), [PNA] and [**1**] = 1.0 µM at 20 °C with $\lambda_{ex}$ = 580 nm.

and fluorescence emission maxima in the presence of acetonitrile (100% MeCN: $\lambda_{abs}$ 544 nm, $\lambda_{em}$ 634 nm; 50% MeCN: $\lambda_{abs}$ 569 nm, $\lambda_{em}$ 647 nm; 20% MeCN: $\lambda_{abs}$ 592 nm, $\lambda_{em}$ 654 nm). The fluorescence spectrum of the Nile red-labeled acpcPNA is notably red-shifted compared to the fluorescence spectrum of the free Nile red label in the same solvent. This suggests possible interactions between the Nile red label and the PNA base or backbone. Nevertheless, the Nile red in the acpcPNA **10mer-Nr** was not quenched as indicated by the quantum yield value being similar to that of the free label (**10mer-Nr**: $\Phi_F$ = 0.16; **1**: $\Phi_F$ = 0.14 in 20% MeCN). Unfortunately, the value can only be compared in an acetonitrile:water mixture as compound **1** is essentially insoluble in water.

Melting temperature data and optical properties of the Nile red-labeled acpcPNAs and their hybrids with various DNA are summarized in Table 2. Thermal denaturation experiments suggest that the Nile red-labeled acpcPNA **10mer-Nr** can form a stable hybrid with complementary DNA. In contrast to some other labels such as pyrene which usually destabilize acpcPNA–DNA duplexes [30,31], the $T_m$ of the complementary DNA hybrid of Nile red-labeled PNA (58.8 °C by UV or 56.9 °C by fluorescence) was comparable to the $T_m$ of unlabeled acpcPNA with an identical sequence (57.6 °C) [26]. Absorption spectra of **10mer-Nr** and its complementary DNA duplex show differences in the Nile red region as shown by the sharpening and red shift of the absorption maxima ($\lambda_{max}$ = 575 and 598 nm in single stranded and duplex, respectively). Importantly, fluorescence spectra of the duplex showed a pronounced blue shift of the emission maxima (647 nm) compared to the single-stranded PNA (656 nm) as well as a small fluorescence increase (1.36 fold at 643 nm) (Figure 3a,b). The blue-shifted fluorescence maxima and increased quantum yields of Nile red suggest that the Nile red chromophore in complementary PNA–DNA duplexes was placed in a less polar environment compared to its placement in single-stranded PNA. Although

**Table 2:** $T_m$ and optical properties of Nile red-labeled acpcPNA[a].

| PNA | DNA (5′ to 3′)[b] | $T_m$ (°C)[c] | $\lambda_{abs}$ | $\lambda_{em}$[d] | $\Phi_F$[e] | $F/F_0$[f] |
|---|---|---|---|---|---|---|
| **10mer-Nr** | none | – | 575 | 656 | 0.11 | – |
| | AGTGATCTAC | 58.8 (56.9) | 598 | 647 | 0.15 | 1.36 |
| | AGTG**C**TCTAC | 37.5 (36.7) | 585 | 648 | 0.17 | 1.72 |
| | AGT**C**ATCTAC | 37.7 (37.5) | 589 | 649 | 0.14 | 1.43 |
| | AGTGA**C**CTAC | 40.8 (41.3) | 592 | 652 | 0.15 | 1.70 |
| | AGTGA**C**TCTAC | 46.9 (52.0) | 589 | 643 | 0.29 | 2.91 |
| | AGTGA**A**TCTAC | 40.1 (46.7) | 588 | 648 | 0.20 | 2.60 |
| | AGTGA**T**TCTAC | 54.7 (52.7) | 591 | 649 | 0.22 | 2.65 |
| | AGTGA**G**TCTAC | 42.2 (46.0) | 593 | 645 | 0.20 | 3.31 |
| | AGT**C**GATCTAC | N.D.[g] | 591 | 648 | 0.14 | 1.16 |
| | AGTGATC**C**TAC | N.D.[g] | 597 | 653 | 0.10 | 0.87 |
| | AGTG**C**TCTAC | N.D.[g] | 585 | 648 | 0.17 | 1.45 |
| | AGTGA**C**CCTAC | N.D.[g] | 593 | 652 | 0.13 | 1.23 |
| | AGTGA**C**TC**C**AC | N.D.[g] | 585 | 647 | 0.13 | 1.24 |
| **11merAA-Nr** | none | – | 598 | 657 | 0.15 | – |
| | CGTATTTTATG | 76.0 | 600 | 651 | 0.23 | 1.55 |
| | CGTATT**C**TTATG | 74.7 | 593 | 645 | 0.33 | 2.34 |
| **11merCC-Nr** | none | – | 594 | 656 | 0.08 | – |
| | CGTATGGTATG | 67.4 | 594 | 646 | 0.19 | 2.49 |
| | CGTATG**C**GTATG | 71.2 | 598 | 644 | 0.13 | 1.79 |
| **11merGG-Nr** | none | – | 599 | 660 | 0.04 | – |
| | CGTATCCTATG | (62.3) | 599 | 655 | 0.15 | 3.27 |
| | CGTATC**C**CTATG | (62.3) | 599 | 652 | 0.10 | 2.46 |
| **11merTT-Nr** | none | – | 588 | 655 | 0.09 | – |
| | CGTATAATATG | 76.6 | 602 | 654 | 0.19 | 2.05 |
| | CGTATA**C**ATATG | 76.0 | 590 | 646 | 0.29 | 3.44 |

[a]All measurements were carried out in 10 mM sodium phosphate buffer (pH 7.0), [PNA] = 1.0 μM; [DNA] = 1.2 μM at 20 °C. [b]Underlined and boldface letters in DNA sequences indicate the position of mismatch and base insertion, respectively. [c]Determined by UV–vis (260 nm) and/or fluorescence spectrophotometry (643 nm, shown in parentheses). [d]$\lambda_{ex}$ = 580 nm. [e]Cresyl violet was used as a standard (Φ = 0.54 in MeOH) [41]. [f]$F/F_0$ was calculated from the ratio of fluorescence of duplex divided by the single-stranded PNA at 643 nm. [g]Not determined.

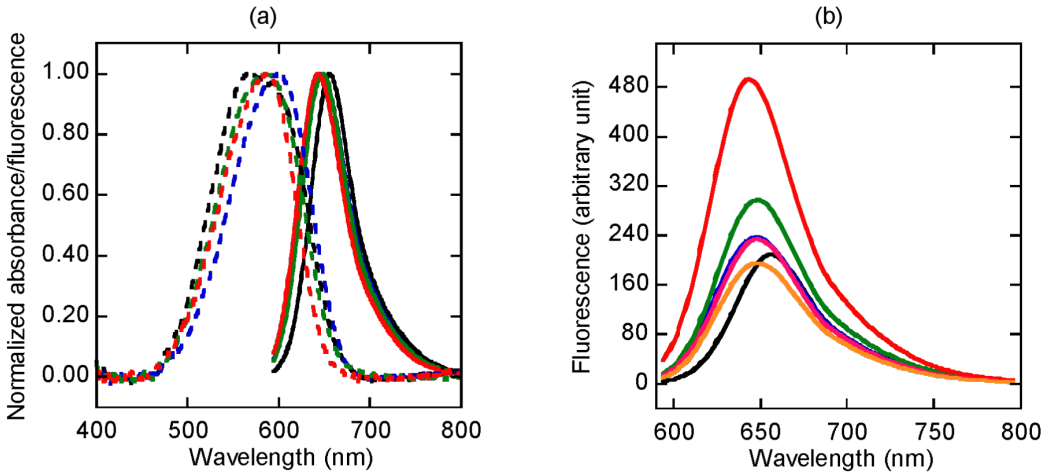



**Figure 3:** (a) Normalized absorption (---) and fluorescence (—) spectra and (b) fluorescence spectra of Nile red-labeled acpcPNA **10mer-Nr** in the absence (black) and presence of various DNA sequences (5'→3'): blue = AGTGATCTAC (complementary); green = AGTGCTCTAC (mismatched); red = AGTGA**C**TCTAC (bulged); pink = AGTGC**C**TCTAC (bulged with mismatch); orange = AGT**C**GATCTAC (misplaced bulge). All measurements were carried out in 10 mM sodium phosphate buffer (pH 7.0), [PNA] = 1.0 μM; [DNA] = 1.2 μM at 20 °C with $\lambda_{ex}$ = 580 nm.

the related benzophenoxazine dye Nile blue binds to DNA by intercalation, the binding results in a bathochromic shifting and a quenched fluorescence [39,40]. Based on this, together with the fact that PNA–DNA duplexes are not good substrates for intercalative binding [23], we propose that the Nile red is more likely to bind to the groove of the PNA–DNA duplex rather than intercalate into the base stacks.

In the presence of a mismatched base in the DNA strand, the $T_m$ of the duplex was decreased as expected. For example, the mismatched hybrid of **10mer-Nr** with the DNA carrying a mismatched base in close proximity to the position of the Nile red label showed $T_m$ values in the range of 37–41 °C. Fluorescence spectra of these mismatched DNA duplexes showed blue-shifted and enhanced fluorescence emissions similar to the complementary duplex, which ranged from 1.4 to 1.7 fold compared to the one of single-stranded Nile red-labeled acpcPNA. This suggests that the Nile red label can still noticeably interact with the mismatched duplexes.

The most remarkable feature of the Nile red-labeled acpcPNA **10mer-Nr** is the relatively large fluorescence increase after the hybridization with DNA targets which carry a base insertion in close proximity to the Nile red label. As shown in Table 2, these duplexes showed more pronounced fluorescence increases (2.6–3.3 fold) than the complementary or mismatched duplexes (1.4–1.7 fold relative to the single-stranded **10mer-Nr**). The fluorescence maxima were also further blue-shifted relative to the single-stranded and complementary duplexes of Nile red-labeled acpcPNA (Figure 3a,b). These results suggest that the Nile red label in these duplexes adopts a different configuration to the complementary and the single-mismatched duplexes. When the DNA strand carries an extra inserted base, the only way it can form a stable hybrid is to form a bulge on the DNA strand at the insertion site. This is supported by the unusually high $T_m$ values (46–53 °C) for these duplexes. In addition, consistent fluorescence increases were observed regardless of the nature of the inserted base. Based on these findings we propose that the extra DNA base is looped out to form a bulge which can accommodate the Nile red label. A similar binding mode has been proposed earlier for pyrene-labeled DNA [42]. When the inserted base was misplaced, i.e., away from the Nile red label, or when a mismatch base was introduced elsewhere in the DNA strand, the fluorescence change was small (Figure 3b, Figure S11 and Figure S12, Supporting Information File 1), suggesting that the fluorescence increase was due to a specific interaction between the Nile red label and the bulge site. The addition of β-cyclodextrin to the bulged duplexes caused no change in the fluorescence of the Nile red label, whereas a marked blue shift was observed with single-stranded PNA (8 nm) (Figure S17, Supporting Information File 1). This experiment confirmed that the Nile red label is buried well within the hydrophobic pocket of the bulged duplexes and therefore not available to form an inclusion complex with the cyclodextrins [17]. Less pronounced shifts were observed with complementary and mismatched duplexes (2 and 4 nm, respectively) upon the addition of cyclodextrin, which indicates possible interactions between the groove-bound Nile red and cyclodextrin.

To better understand the effect of neighboring bases on the optical properties of the Nile red-labeled acpcPNA and its duplexes with DNA, four Nile red-labeled acpcPNA sequences (**11merAA-Nr**, **11merCC-Nr**, **11merGG-Nr**, **11merTT-Nr**) were synthesized. These four Nile red-labeled acpcPNA sequence exhibited differences only at the bases flanking the Nile red label, For single-stranded PNA, the order of fluorescence quantum yields was AA > TT ~ CC > GG. This suggests a more efficient quenching of the Nile red label by neighboring G than by other nucleobases. A is almost non-quenching as shown by the high quantum yield of **11merAA-Nr** ($\Phi_F$ = 0.15), which is comparable to the high quantum yield of the free Nile red label ($\Phi_F$ = 0.14). On the other hand, the presence of two flanking G in **11merGG-Nr** resulted in a decrease of the quantum yield of more than 70% ($\Phi_F$ = 0.04). Upon hybridization with complementary DNA targets, the fluorescence quantum yields of all four hybrids were increased to a similar range ($\Phi_F$ = 0.15–0.23). However, since the initial fluorescence of **11merGG-Nr** was low, the fluorescence change was more pronounced (3.27 fold) than other sequences (1.55–2.49 fold). When a bulge was introduced, a larger fluorescence increase was observed in duplexes with neighboring A-T than G-C pairs. Accordingly, the bulged duplexes of **11merAA-Nr** and **11merTT-Nr** became more fluorescent than the corresponding complementary duplexes. Opposite results were observed with **11merCC-Nr** and **11merGG-Nr**, that is, the complementary duplexes exhibited higher fluorescence than the bulged duplexes. In all cases the fluorescence emission maxima of the bulged duplex were at shorter wavelengths than fluorescence emission maxima of the complementary duplexes and single-stranded PNA. These results clearly support our hypothesis of the different placement of the Nile red label in the complementary and bulged duplexes.

Although Nile red has been previously incorporated in DNA either through a base substitution or modification, these Nile red-labeled DNAs do not show appreciable fluorescence changes in response to the hybridization with DNA regardless of the mode of Nile red attachment [14,17,18]. In most cases, the pairing of Nile red-labeled DNA with another DNA strand resulted in unchanged or decreased fluorescence quantum yields. This, together with the red-shifted absorption and fluo-

rescence spectra in comparison to single-stranded DNA, suggests that Nile red may intercalate into the base stack of DNA–DNA duplexes. The behavior of the Nile red label in acpcPNA is completely different, as shown by the consistent increase in the fluorescence quantum yield upon hybridization with complementary DNA irrespective of the nature of the flanking bases. In such duplexes, the Nile red is expected to interact with the PNA–DNA duplexes by means of groove binding, which results in a lower localized polarity around the Nile red chromophore and gives rise to the observed blue shifts and the fluorescence increase. In addition, a larger increase in fluorescence was also observed with DNA targets that can form a bulge in the vicinity of the Nile red label. We propose that the Nile red label is buried within the looped out structure of the bulged duplex, which gives rise to an even more pronounced blue shift and fluorescence increase, except when there are nearby G-C base pairs which may quench the fluorescence.

## Conclusion

We successfully synthesized the propargylated Nile red **1** and clicked it onto acpcPNA with an azide-modified backbone. The solvatochromic properties of the Nile red label is retained in the labeled acpcPNA. Quenching of the Nile red label by neighboring bases in acpcPNA increases in the order of A > T ~ C > G. The hybridization with fully complementary DNA and DNA with an inserted base consistently resulted in blue-shifted and enhanced fluorescence. This indicates that the Nile red label in acpcPNA–DNA duplexes is in a more hydrophobic environment compared to when the Nile red label is in single-stranded acpcPNA. Based on spectroscopic evidence, we propose that either the Nile red label resides within the groove (complementary duplexes) or the hydrophobic pocket formed by the looped-out base (bulged duplexes) rather than intercalating into the base stacks.

## Experimental

### General remarks

All chemicals were obtained from standard suppliers and used as received. Anhydrous DMF for peptide synthesis was purchased from RCI Labscan (Thailand). All other solvents were AR or HPLC grade and were used without further purification. Oligonucleotides were obtained from Pacific Science (Thailand) or BioDesign (Thailand). The water used in all experiments was obtained from an ultrapure water system fitted with a Millipak® 40 filter unit.

### Synthesis of propargyl Nile red **1**

2-Hydroxy Nile red [37] (140.1 mg, 0.41 mmol) was dissolved in anhydrous DMF (5 mL). The solution was cooled in an ice bath followed by the addition of potassium carbonate (210.7 mg, 1.5 mmol) and propargyl bromide (100 µL, 1.5 mmol). The mixture was heated at 70 °C and stirred for 3 h. After the removal of the solvent, the residue was purified by column chromatography (EtOAc/hexanes 1:4) to obtain **1** (108.1 mg, 70% yield) as a dark purple solid. $^1$H NMR (400 MHz, DMSO-$d_6$) δ 1.17 (t, $J$ = 6.9 Hz, 6H, C$H_3$CH$_2$), 3.51 (q, $J$ = 6.9 Hz, 4H, CH$_3$C$H_2$), 3.67 (t, $J$ = 2.2 Hz, 1H, ≡CH), 5.03 (d, $J$ = 2.2 Hz, 2H, OCH$_2$), 6.20 (s, 1H, ArH), 6.65 (d, $J$ = 2.6 Hz, 1H, ArH), 6.83 (dd, $J$ = 9.1 and 2.6 Hz, 1H, ArH), 7.32 (dd, $J$ = 8.7 and 2.5 Hz, 1H, ArH), 7.64 (d, $J$ = 9.1 Hz, 1H, ArH), 8.04 (d, $J$ = 2.5 Hz, 1H, ArH), 8.08 (d, $J$ = 8.7 Hz, 2H, ArH); $^{13}$C NMR (100 MHz, DMSO-$d_6$) δ 181.3, 159.8, 151.8, 150.9, 146.5, 138.1, 133.5, 131.0, 127.2, 125.5, 124.0, 118.2, 110.1, 107.0, 104.1, 96.0, 78.8, 78.7, 55.9, 44.4, 12.4; IR (ATR) $\nu_{max}$: 3281.7, 3190.3, 2968.3, 2920.4, 2846.4, 1675.4, 1584.0, 1405.6 cm$^{-1}$; HRMS (ESI–TOF): [M + H]$^+$ calcd for $C_{23}H_{21}N_2O_3$, 373.1552; found, 373.1581; UV (MeOH) $\lambda_{max}$ (ε): 553 (3.2 × 10$^4$).

### Synthesis of Nile red-modified acpcPNA

The 3-aminopyrrolidine-4-carboxylic acid (apc) modified acpcPNA was manually synthesized at a 1.5 µmol scale on Tentagel S-RAM resin (Fluka, 0.24 mmol/g) from the four Fmoc-protected pyrrolidinyl PNA monomers (A$^{Bz}$, T, C$^{Bz}$, G$^{Ibu}$) and spacers [Fmoc-(1*S*,2*S*)-2-aminocycolpentanecarboxylic acid or (3*R*,4*S*)-3-(Fmoc-amino)-1-trifluoroacetylpyrrolidine-4-carboxylic acid] [43] according to the previously published protocol [26,27,31]. Lysine was included at both C- and N-termini to improve the water solubility. After the completion of the synthesis, the N-terminal Fmoc group was removed and the free amino group was capped by acetylation. The acpcPNA on the solid support was spilt to 0.5 µmol portions for a further labeling experiment and treated with 1:1 dioxane/aqueous $NH_3$ at 60 °C overnight to remove the nucleobase- and apc-protecting groups. Fully deprotected acpcPNA (0.5 µmol) was treated with 4-azidobutanal (15 µmol, 30 equiv) in the presence of $NaBH_3CN$ (30 µmol, 60 equiv) and HOAc (30 µmol, 60 equiv) in MeOH (200 µL) at room temperature overnight. After an exhaustive washing with MeOH, the azide-modified acpcPNA was reacted with **1** (7.5 µmol, 15 equiv) in the presence of tris[(benzyl-1*H*-1,2,3-triazol-4-yl)methyl]amine [44] (TBTA, 30 µmol, 60 equiv), tetrakis(acetonitrile) copper(I) hexafluorophosphate (15 µmol, 30 equiv) and (+)-sodium-L-ascorbate (60 µmol, 120 equiv) in 3:1 (v/v) DMSO/*t*-BuOH at room temperature overnight. After the reaction was completed, the labeled acpcPNA was cleaved from the solid support with trifluoroacetic acid (500 µL × 30 min × 3). After drying and washing with diethyl ether, the residue was purified by reversed-phase HPLC and characterized by MALDI–TOF mass spectrometry (Microflex, Bruker Daltonics; α-cyano-4-hydroxycinnamic acid matrix, positive linear ion mode).

### Spectroscopic studies

Samples for melting temperature, UV–vis, and fluorescence studies were prepared in 10 mM sodium phosphate buffer (pH 7.0) at concentrations of PNA = 1.0 μM and DNA = 1.2 μM. UV–vis and thermal denaturation experiments were performed on a CARY 100 UV–vis spectrophotometer (Varian, Australia) equipped with a thermal melting system. Fluorescence spectra were measured on a Cary Eclipse Fluorescence Spectrophotometer (Varian, Australia) at an excitation wavelength of 580 nm with 5 nm excitation and emission slits. Fluorescence melting experiments were performed under the same conditions by measuring the sample that was pre-heated to the specified temperature at 5 °C intervals and left for equilibration at that temperature for at least 5 min. The emission at 643 nm was divided by the emission at 643 nm of the single-stranded acpcPNA at the same temperature and plotted against the temperature to obtain the fluorescence melting curves. Melting temperatures were obtained from the UV or fluorescence melting curves by first derivative plots.

## Supporting Information

### Supporting Information File 1

NMR spectra, HPLC chromatogram, mass spectra and additional spectroscopic data.
[http://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-10-224-S1.pdf]

## Acknowledgements

Financial support of this work from the Thailand Research Fund (DPG5780002), the Ratchadaphiseksomphot Endowment Fund of the Chulalongkorn University (RES560530126-AM), and the Thai government stimulus package 2 (TKK2555, SP2) under the Project for the Establishment of Comprehensive Center for Innovative Food, Health Products and Agriculture and the Royal Golden Jubilee Ph.D. Program (Grant No. PHD/0316/2550, to NY and TV) is gratefully acknowledged.

## References

1. Murphy, M. C.; Rasnik, I.; Cheng, W.; Lohman, T. M.; Ha, T. *Biophys. J.* **2004,** *86,* 2530–2537. doi:10.1016/S0006-3495(04)74308-8
2. Sindbert, S.; Kalinin, S.; Nguyen, H.; Kienzler, A.; Clima, L.; Bannwarth, W.; Appel, B.; Müller, S.; Seidel, C. A. M. *J. Am. Chem. Soc.* **2011,** *133,* 2463–2480. doi:10.1021/ja105725e
3. Börjesson, K.; Preus, S.; El-Sagheer, A.-H.; Brown, T.; Albinsson, B.; Wilhelmsson, L. M. *J. Am. Chem. Soc.* **2009,** *131,* 4288–4293. doi:10.1021/ja806944w
4. Marras, S. A. E. *Mol. Biotechnol.* **2008,** *38,* 247–255. doi:10.1007/s12033-007-9012-9
5. Sapsford, K. E.; Berti, L.; Medintz, I. L. *Angew. Chem., Int. Ed.* **2006,** *45,* 4562–4588. doi:10.1002/anie.200503873
6. Sahoo, H. *J. Photochem. Photobiol., C* **2011,** *12,* 20–30. doi:10.1016/j.jphotochemrev.2011.05.001
7. Lavis, L. D.; Raines, R. T. *ACS Chem. Biol.* **2008,** *3,* 142–155. doi:10.1021/cb700248m
8. Waegele, M. M.; Culik, R. M.; Gai, F. *J. Phys. Chem. Lett.* **2011,** *2,* 2598–2609. doi:10.1021/jz201161b
9. Dziuba, D.; Postupalenko, V. Y.; Spadafora, M.; Klymchenko, A. S.; Guérineau, V.; Mély, Y.; Benhida, R.; Burger, A. *J. Am. Chem. Soc.* **2012,** *134,* 10209–10213. doi:10.1021/ja3030388
10. Jose, J.; Burgess, K. *Tetrahedron* **2006,** *62,* 11021–11037. doi:10.1016/j.tet.2006.08.056
11. Fowler, S. D.; Greenspan, P. *J. Histochem. Cytochem.* **1985,** *33,* 833–836. doi:10.1177/33.8.4020099
12. Greenspan, P.; Mayer, E. P.; Fowler, S. D. *J. Cell Biol.* **1985,** *100,* 965–973. doi:10.1083/jcb.100.3.965
13. Sackett, D. L.; Wolff, J. *Anal. Biochem.* **1987,** *167,* 228–234. doi:10.1016/0003-2697(87)90157-6
14. Varghese, R.; Gajula, P. K.; Chakraborty, T. K.; Wagenknecht, H.-A. *Synlett* **2009,** 3252–3257. doi:10.1055/s-0029-1218384
15. Varghese, R.; Wagenknecht, H.-A. *Chem. – Eur. J.* **2010,** *16,* 9040–9046. doi:10.1002/chem.201001136
16. Beyer, C.; Wagenknecht, H.-A. *Chem. Commun.* **2010,** *46,* 2230–2231. doi:10.1039/b924471a
17. Okamoto, A.; Tainaka, K.; Fujiwara, Y. *J. Org. Chem.* **2006,** *71,* 3592–3598. doi:10.1021/jo060168o
18. Prokhorenko, I. A.; Dioubankova, N. N.; Korshun, V. A. *Nucleosides, Nucleotides Nucleic Acids* **2004,** *23,* 509–520. doi:10.1081/NCN-120028344
19. Wenge, U.; Ehrenschwender, T.; Wagenknecht, H.-A. *Bioconjugate Chem.* **2013,** *24,* 301–304. doi:10.1021/bc300624m
20. Varghese, R.; Wagenknecht, H.-A. *Org. Biomol. Chem.* **2010,** *8,* 526–528. doi:10.1039/b918381g
21. Varghese, R.; Wagenknecht, H.-A. *Chem. – Eur. J.* **2009,** *15,* 9307–9310. doi:10.1002/chem.200901147
22. Lachmann, D.; Berndl, S.; Wolfbeis, O. S.; Wagenknecht, H.-A. *Beilstein J. Org. Chem.* **2010,** *6,* No. 13. doi:10.3762/bjoc.6.13
23. Wittung, P.; Kim, S. K.; Buchardt, O.; Nielsen, P.; Nordén, B. *Nucleic Acids Res.* **1994,** *22,* 5371–5377. doi:10.1093/nar/22.24.5371
24. Dilek, I.; Madrid, M.; Singh, R.; Urrea, C. P.; Armitage, B. A. *J. Am. Chem. Soc.* **2005,** *127,* 3339–3345. doi:10.1021/ja045145a
25. Armitage, B. A. *Top. Curr. Chem.* **2005,** *253,* 55–76. doi:10.1007/b100442
26. Vilaivan, T.; Srisuwannaket, C. *Org. Lett.* **2006,** *8,* 1897–1900. doi:10.1021/ol060448q
27. Vilaivan, C.; Srisuwannaket, C.; Ananthanawat, C.; Suparpprom, C.; Kawakami, J.; Yamaguchi, Y.; Tanaka, Y.; Vilaivan, T. *Artificial DNA: PNA XNA* **2011,** *2,* 50–59. doi:10.4161/adna.2.2.16340
28. Mansawat, W.; Vilaivan, C.; Balázs, Á.; Aitken, D. J.; Vilaivan, T. *Org. Lett.* **2012,** *14,* 1440–1443. doi:10.1021/ol300190u
29. Boonlua, C.; Vilaivan, C.; Wagenknecht, H.-A.; Vilaivan, T. *Chem. – Asian J.* **2011,** *6,* 3251–3259. doi:10.1002/asia.201100490
30. Boonlua, C.; Ditmangklo, B.; Reenabthue, N.; Suparpprom, C.; Poomsuk, N.; Siriwong, K.; Vilaivan, T. *RSC Adv.* **2014,** *4,* 8817–8827. doi:10.1039/c3ra47997h
31. Ditmangklo, B.; Boonlua, C.; Suparpprom, C.; Vilaivan, T. *Bioconjugate Chem.* **2013,** *24,* 614–625. doi:10.1021/bc3005914
32. Sezi, S.; Varghese, R.; Vilaivan, T.; Wagenknecht, H.-A. *ChemistryOpen* **2012,** *1,* 173–176. doi:10.1002/open.201200016
33. Meldal, M.; Tornøe, C. W. *Chem. Rev.* **2008,** *108,* 2952–3015. doi:10.1021/cr0783479

34. Gramlich, P. M. E.; Wirges, C. T.; Manetto, A.; Carell, T. *Angew. Chem., Int. Ed.* **2008,** *47,* 8350–8358. doi:10.1002/anie.200802077
35. Amblard, F.; Cho, J. H.; Schinazi, R. F. *Chem. Rev.* **2009,** *109,* 4207–4220. doi:10.1021/cr9001462
36. El-Sagheer, A. H.; Brown, T. *Chem. Soc. Rev.* **2010,** *39,* 1388–1405. doi:10.1039/b901971p
37. Briggs, M. S. J.; Bruce, I.; Miller, J. N.; Moody, C. J.; Simmonds, A. C.; Swann, E. *J. Chem. Soc., Perkin Trans. 1* **1997,** 1051–1058. doi:10.1039/a605012c
38. Jose, J.; Burgess, K. *J. Org. Chem.* **2006,** *71,* 7835–7839. doi:10.1021/jo061369v
39. Chen, Q.-y.; Li, D.-h.; Yang, H.-h.; Zhu, Q.-z.; Xu, J.-g.; Zhao, Y. *Analyst* **1999,** *124,* 901–906. doi:10.1039/a901174i
40. Zhao, G.-C.; Zhu, J.-J.; Chen, H.-Y.; Wang, X.-M.; Lu, Z.-H. *Chin. J. Chem.* **2002,** *20,* 57–62. doi:10.1002/cjoc.20020200112
41. Magde, D.; Brannon, J. H.; Cremers, T. L.; Olmsted, J., III. *J. Phys. Chem.* **1979,** *83,* 696–699. doi:10.1021/j100469a012
42. Okamoto, A.; Ichiba, T.; Saito, I. *J. Am. Chem. Soc.* **2004,** *126,* 8364–8365. doi:10.1021/ja049061d
43. Reenabthue, N.; Boonlua, C.; Vilaivan, C.; Vilaivan, T.; Suparpprom, C. *Bioorg. Med. Chem. Lett.* **2011,** *21,* 6465–6469. doi:10.1016/j.bmcl.2011.08.079
44. Chan, T. R.; Hilgraf, R.; Sharpless, K. B.; Fokin, V. V. *Org. Lett.* **2004,** *6,* 2853–2855. doi:10.1021/ol0493094

## License and Terms

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

# Molecular recognition of AT-DNA sequences by the induced CD pattern of dibenzotetraaza[14]annulene (DBTAA)–adenine derivatives

Marijana Radić Stojković[1], Marko Škugor[1], Łukasz Dudek[2], Jarosław Grolik[2], Julita Eilmes[2] and Ivo Piantanida[*1]

Full Research Paper

Open Access

Address:
[1]Laboratory for Study of Interactions of Biomacromolecules, Division of Organic Chemistry and Biochemistry, Ruđer Bošković Institute, Bijenička cesta 54, PO Box 180, HR-10002 Zagreb, Croatia and [2]Department of Chemistry, Jagiellonian University, Ingardena 3, 30-060 Kraków, Poland

Email:
Ivo Piantanida[*] - pianta@irb.hr

* Corresponding author

Keywords:
AT-DNA recognition; circular dichroism; DBTAA-adenine conjugate; ds-DNA/RNA; minor groove binding; nucleic acids

*Beilstein J. Org. Chem.* **2014,** *10,* 2175–2185.
doi:10.3762/bjoc.10.225

Received: 12 May 2014
Accepted: 27 August 2014
Published: 12 September 2014

This article is part of the Thematic Series "Nucleic acid chemistry".

Guest Editor: H.-A. Wagenknecht



## Abstract

An investigation of the interactions of two novel and several known DBTAA–adenine conjugates with double-stranded DNA and RNA has revealed the DNA/RNA groove as the dominant binding site, which is in contrast to the majority of previously studied DBTAA analogues (DNA/RNA intercalators). Only DBTAA–propyladenine conjugates revealed the molecular recognition of AT-DNA by an ICD band pattern > 300 nm, whereas significant ICD bands did not appear for other ds-DNA/RNA. A structure–activity relation for the studied series of compounds showed that the essential structural features for the ICD recognition are a) the presence of DNA-binding appendages (adenine side chain and positively charged side chain) on both DBTAA side chains, and b) the presence of a short propyl linker, which does not support intramolecular aromatic stacking between DBTAA and adenine. The observed AT-DNA-ICD pattern differs from previously reported ss-DNA (poly dT) ICD recognition by a strong negative ICD band at 350 nm, which allows for the dynamic differentiation between ss-DNA (poly dT) and coupled ds-AT-DNA.

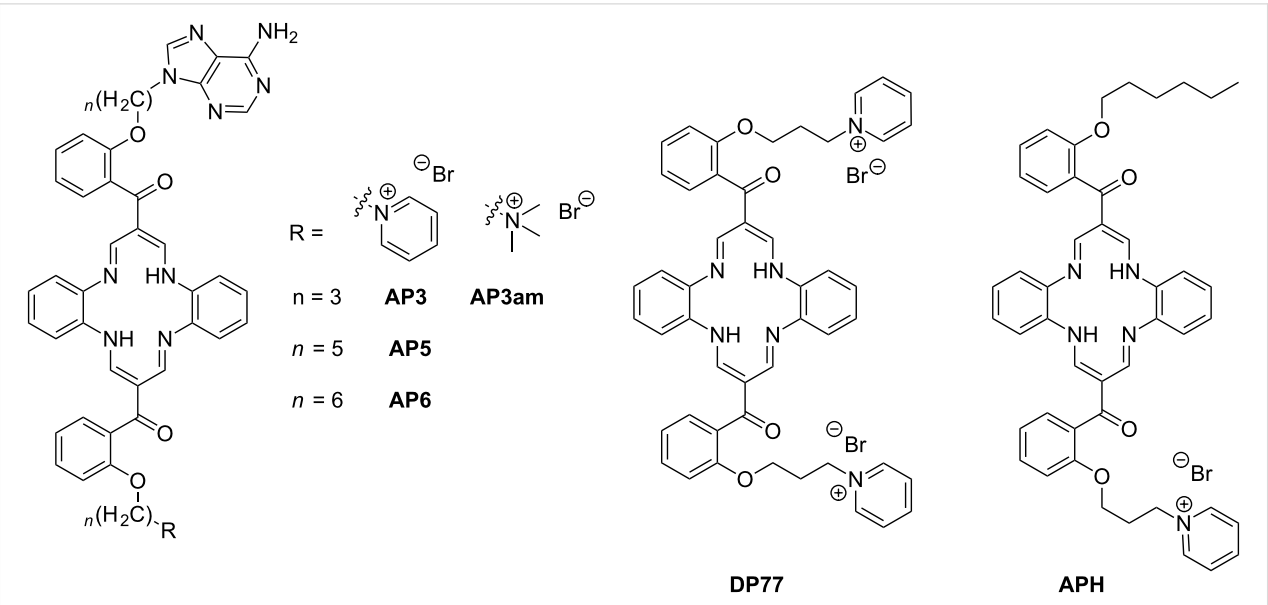
## Introduction

The majority of natural and artificial applications involving small molecule-DNA/RNA recognition depend on several non-covalent binding modes. Typical examples are double-stranded (ds) DNA/RNA intercalation, minor or major groove binding, and external electrostatic binding [1]. However, non-covalent interactions involving small molecules ($M_w$ < 600) can only rely on a small number of interacting groups, while the steric parameters of DNA/RNA binding sites are also quite limited.

Nevertheless, the dynamic nature of nucleic acids, combined with structural differences between various DNA/RNA sequences, offer numerous highly interesting targets. Within the last two decades, the design of small binding molecules has mostly relied on the three-dimensional recognition of various targeted DNA/RNA sites [2,3], frequently relying upon new knowledge gained from supramolecular chemistry [4].

Among the large number of DNA/RNA sequences of biochemical interest, long homogeneous AT tracts have attracted significant attention as small molecule targets. An illustrative natural example can be found in proteins which use selective binding interactions of an arginine-rich side arm inside the AT sequence minor groove to broaden the related DNA major groove, within which the protein biological action takes place [5]. However, the majority of known small molecules are not able to distinguish AT tracts by length and composition. This has been demonstrated, for instance, by the inability to distinguish between longer tracts (more than 20–30 base pairs) of homo and alternating distribution of AT base pairs.

Our recent studies revealed a novel class of cyanine dyes characterized by bulky phosphonium substituents as intriguing AT binders, which showed very rare kinetic differentiation between alternating- and homo-AT-DNA sequences [6]. Moreover, these dyes efficiently entered cells and were shown to be non-toxic, mitochondria-specific fluorescent markers [7]. Previous studies also revealed the advantage of nucleobase incorporation in small molecule structures for the recognition of complementary nucleotides/polynucleotides [8-10]. Recent studies of the novel DBTAA-adenine conjugates **AP3** and **AP6** (Scheme 1) showed a highly selective binding of only **AP3** to poly dT among all other ss-DNA/RNA, as characterized by the induction of a specific CD band response [11]. Observed oligo dT specificity provoked the intriguing question of whether such a specificity would be observed in double-stranded AT-DNA sequences. The experimental design to investigate this question was supported by the aforementioned importance of AT tracts and the selectivity of cyanine dyes. On this basis, we were inspired to prepare the novel binders **AP5** with a DBTAA–adenine linker length between **AP3** (oligo dT specific) and **AP6** (oligo-dT inert) and **AP3am**, in which pyridinium is exchanged by a permethylated amine, with the aim of determining the importance of aromatic stacking interactions. The results were compared with the reference **APH** (lacking adenine) and previously studied **DP77** [12] (Scheme 1), which, having pyridinium instead of adenine, can also be regarded as a reference structure. Double-stranded DNA/RNA targets chosen for this study are long (<100 base pairs) synthetic polynucleotides poly dG–poly dC, poly dA–poly dT, poly dAdT–poly dAdT and poly rA–poly rU, each associated with specific structural properties of the minor/major groove as the anticipated binding site (**APH**, **AP3**, **AP6** didn't intercalate into ct-DNA [11]). Namely, parameters such as the groove width and depth, steric obstructions like the amino groups of guanine, H-bonding patterns, as well as polynucleotide charge density and the curvature of the ds-helix backbone all differ significantly across the double stranded examples mentioned above (Table S1, Supporting Information File 1). The choice of long polynucleotides ensures that significant binding of the DBTAA moiety at the ends of the double




**Scheme 1:** Studied DBTAA–adenine conjugates (**AP3**, **AP3am**, **AP5**, **AP6**) [11], and the reference compounds lacking adenine (**APH** [11], **DP77** [12,13]).

strands ("capping") can be neglected and that our investigations indeed sense the differences of the secondary structure (from minor/major groove) of the studied DNA and RNA sequences.

# Results and Discussion

## Synthesis

The synthetic routes to the new adenine–DBTAA conjugates **AP3am** and **AP5** are summarized in Scheme 2, the details of which are given in the Experimental section.

## Spectroscopic characterisation of DBTAA derivatives in aqueous medium

Novel compounds **AP3am** and **AP5** were moderately soluble in aqueous solutions, that is, up to $c_{AP3am} = 1 \times 10^{-3}$ mol dm$^{-3}$ and $c_{AP5} = 1 \times 10^{-4}$ mol dm$^{-3}$, respectively. Buffered aqueous solutions of the studied compounds were stable for several months and their absorbances were proportional to their concentrations up to $c = 3 \times 10^{-5}$ mol dm$^{-3}$. Changes of the UV–vis spectra in response to temperature increases of up to 95 °C were negligible, and the reproducibility of the UV–vis spectra upon cooling back to 25 °C was excellent. The UV–vis spectra of new compounds (**AP3am**, **AP5**) were similar to previously studied analogues [11] (Supporting Information File 1).

## Study of interactions of DBTAA derivatives with ds-DNA/RNA in aqueous medium

Previous studies [11] revealed intriguing recognition of single stranded DNA dT sequences by **AP3** (but not **AP6**, **APH**), whereby intercalation of **AP3**, **AP6**, **APH** in double stranded ct-DNA was excluded. Here, we studied in more detail the interactions of **AP3**, **AP6**, **APH** as well as the novel **AP3am** and **AP5** with a series of synthetic double-stranded polynucleotides.

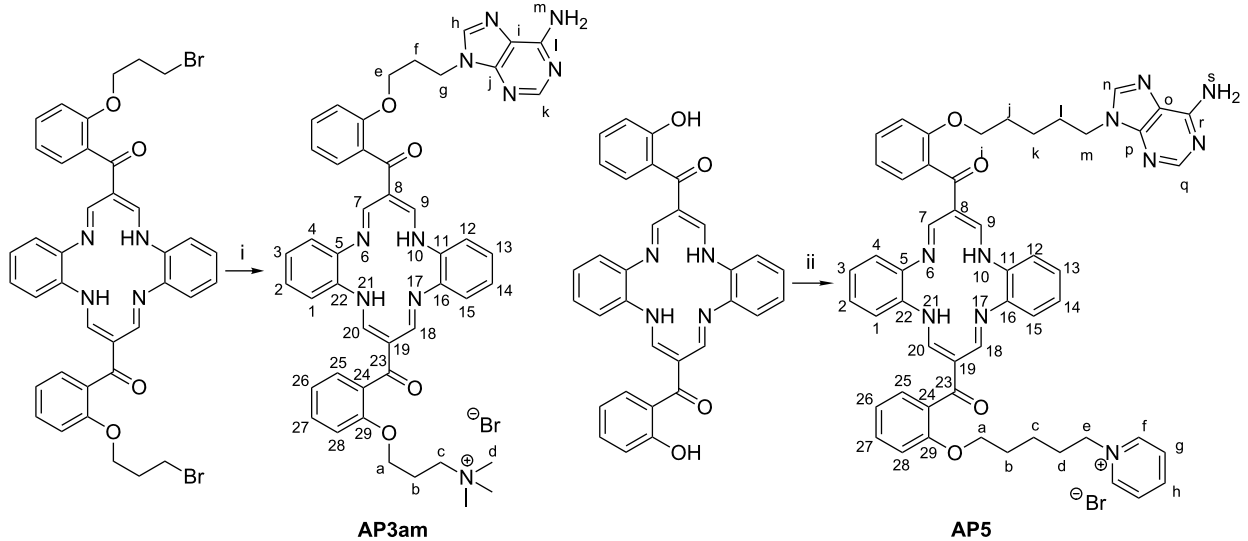
## Thermal denaturation experiments

Non-covalent binding of ligands to ds-DNA/RNA usually induces stabilization of the ds-helix against thermal denaturation resulting in the increase of DNA/RNA-$T_m$. In particular, thermal stabilization is characteristic for the intercalative binding mode due to the strong aromatic stacking interactions between studied condensed aromatic molecule and adjacent base pairs [1,16].

None of the **AP** compounds showed any stabilization effect on any of studied ds-polynucleotides, whereas the previously studied **DP77** analogue [12] showed a significant stabilization effect. The **DP77** analogue essentially differing from the **AP** series by the presence of two positive charges, one each on side arm (Table 1).

## UV–vis titrations and affinity determination

Titration of the **AP** series with any studied polynucleotide resulted in a pronounced decrease of the UV–vis absorbance of the DBTAA chromophore at >300 nm (Figure 1, Table 2). However, no measurable shifts of the UV–vis absorption maxima for any of the **AP** series were observed. In contrast, the previously studied analogue **DP77** exhibited strong bathochromic shifts of ($\Delta\lambda_{346\ nm}$ = 4–13 nm; pyridinium instead of adenine) [12]. The observation of a hypochromic and not a




**Scheme 2:** Preparations of the products **AP3am** and **AP5**. Starting compounds were synthesized according to previously reported procedures [13,14]. i: (1) adenine, NaH, DMF, (2) trimethylamine, DMF; ii: (1) 9-(5-bromopentyl)adenine [15], potassium carbonate, DMF, (2) 1,5-dibromopentane, potassium carbonate, DMF, (3) pyridine.

**Table 1:** The $\Delta T_m$[a] values (°C) of studied ds-polynucleotides upon addition of **AP3**, **AP3am AP5, AP6** and **APH** at pH 7.0 (sodium cacodylate buffer, $I$ = 0.05 mol dm$^{-3}$), ratio $r$ = 0.3[b].

| | AP3 | AP3am | AP5 | AP6 | APH | DP77 [12] |
|---|---|---|---|---|---|---|
| ct-DNA ($T_m$ = 79.5 °C) | 0 | 0 | 0 | 0 | 0 | 12.8 |
| poly A - poly U ($T_m$ = 52.7 °C) | 0.5 | 0 | 0 | 0 | 0 | 28.1 |
| poly dA–poly dT ($T_m$ = 62.7 °C) | 0.7 | 0 | 0 | 0 | 0 | 12.2 |
| poly dAdT–poly dAdT ($T_m$ = 56.0 °C) | <3 | <2 | – | 0 | 0 | – |

[a]Error in $\Delta T_m$ : ±0.5 °C; [b]$r$ = [compound]/[polynucleotide].

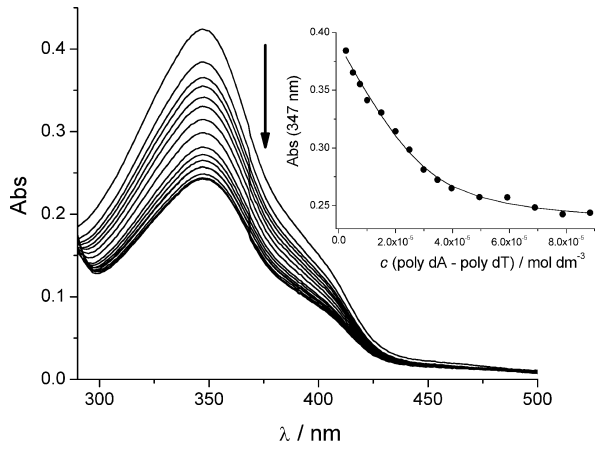



**Figure 1:** Changes in the UV–vis spectrum of **AP3am** ($c$ = 1.0 × 10$^{-5}$ mol dm$^{-3}$) upon titration with poly dA–poly dT; Inset: dependence of the **AP3am** absorbance at $\lambda_{max}$ = 347 nm on $c$ (poly dA–poly dT), pH 7.0, sodium cacodylate buffer, $I$ = 0.05 mol dm$^{-3}$.

bathochromic effect for the **AP** series suggested a different type of aromatic stacking interaction compared to **DP77** (DNA intercalator). There are several possible explanations for the observed hypochromic effect, including the intramolecular stacking of DBTAA with adenine or the intermolecular stacking of two DBTAA chromophores within DNA/RNA grooves, a solvatochromic effect in the DNA/RNA binding site, and even a weak partial intercalation of DBTAA. However, for the accurate elucidation of the precise mechanism, knowledge of the DNA/RNA binding mode is a prerequisite.

Processing of the titration data by means of the Scatchard equation [17,18] gave the binding constants and the density of the binding sites (Table 2, log $K_s$ and ratio $n$, respectively). Data for the **APH** are not presented in the Table 2 due to the formation of the agglomerates, which cause baseline increases, thus preventing the collection of a sufficient number of data (>10 points necessary) for an accurate non-linear fitting.

Examination of the data in Table 2 indicates that all studied compounds showed similar affinity toward ds-DNA and ds-RNA. The density of the binding sites (Table 2, ratio $n$) is mostly too high for the intercalative mode of binding ($n_{intercal.\ teor.}$ < 0.25). This hints at binding within the grooves of DNA/RNA or agglomeration along the DNA/RNA double helix. The possibility of several simultaneous binding modes is excluded by the isoelliptic points observed in CD experiments, which strongly support one dominant binding mode.

**Table 2:** Binding constants (log $K_s$)[a] and ratios $n$[b] ([bound compound]/[polynucleotide phosphate]) calculated from the UV–vis titrations of **AP3**, **AP3am, AP5** and **AP6** with ds-polynucleotides at pH 7.0 (buffer sodium cacodylate, $I$ = 0.05 mol dm$^{-3}$).

| | AP3 | | | AP3am | | | AP5 | | | AP6 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | H[c] | log $K_s$ | $n$ | H[c] | log $K_s$ | $n$ | H[c] | log $K_s$ | $n$ | H[c] | log $K_s$ | $n$ |
| poly A–poly U | 53 | 6.0 | 0.6 | 35 | 5.8 | 0.5 | 44 | 6.6 | 0.7 | 64 | 5.7 | 0.7 |
| poly dA–poly dT | 36 | 6.2 | 0.4 | 40 | 5.9 | 0.3 | 40 | 5.4 | 0.5 | 37 | 5.0 | 0.7 |
| poly dAdT–poly dAdT | 35 | 6.1 | 0.5 | 30 | 5.5 | 0.6 | –[d] | –[d] | –[d] | –[d] | –[d] | –[d] |
| poly dG–poly dC | 64 | 5.5 | 0.5 | 35 | 5.7 | 0.5 | 61 | 5.8 | 0.6 | 74 | 5.6 | 0.5 |

[a]Titration data were processed according to the Scatchard equation [17,18]; [b]Accuracy of $n$ ± 10–30%, consequently log $K_s$ values vary in the same order of magnitude; [c]H/% = (Abs(**APH**, **AP3**, **AP3am**, **AP5**, **AP6**) − Abs(complex))/Abs(**AP3**, **AP3am**, **AP5**, **AP6**)) × 100; [d]Accurate calculation of log $K_s$ values was hampered by precipitation during titration.

The thermal stabilization of ds-DNA/RNA is an essential feature of the intercalative binding mode, as well as for many groove binding molecules. The inability of the entire **AP** series to stabilize ds-DNA/RNA (Table 1), in combination with a considerable binding affinity (Table 2) gives rise to the question of the **AP** series binding mode. To shed more light on the unusual binding process of the **AP** dyes, it was necessary to assess the structure of the **AP** dye/polynucleotide complex. Application of the most informative methods like NMR and X-ray crystallography was hampered by the tendency of the **AP** dye/DNA complexes to form colloidal systems at $c > 0.1$ mM, which neither crystalized nor were suitable for NMR studies due to the extensive broadening and decreasing of proton signals, which eventually merged with the baseline (details about unsuccessful NMR experiments see [11]).

## Circular dichroism (CD) experiments

CD spectropolarimetry offers unique possibilities for the investigation of small molecule–DNA/RNA interactions. The polynucleotide secondary structure is chiral itself, and any binding-induced conformational changes are reflected in the CD spectrum [19]. Moreover, the uniform binding of achiral small molecules within chiral DNA/RNA helix results in an induced CD spectrum (ICD) of the small molecule chromophore, whereby the ICD spectrum range >300 nm (at which DNA/RNA do not absorb) is highly informative about the orientation of the chromophore with respect to DNA/RNA chiral axis [19,20]. For instance, the intercalation of an aromatic moiety inside DNA/RNA should result in a weak negative ICD, groove binding would give a strong positive ICD, while the eventual aromatic-dimer formation within polynucleotide grooves should yield the bisignate coupled exciton electronic coupling (EC) ICD bands [20].

The studied DBTAA derivatives are achiral and are not associated with intrinsic CD spectra. Thus, the appearance of ICD bands >300 nm upon DNA/RNA binding could be used to estimate the orientation of the DBTAA chromophore in the DNA/RNA binding site.

The addition of **APH**, **AP3**, **AP3am**, **AP5** and **AP6** resulted in a decrease of the ds-DNA/RNA CD bands (range from 220 to 300 nm, Supporting Information File 1), whereas the isoelliptic points supported the formation of only one type of the compound/DNA or RNA complex. The intensity decrease of ds-DNA/RNA CD bands is usually associated with the partial disruption of the polynucleotide helical chirality caused by the binding of a small molecule.

The reference compound **APH**, as well as **AP5** and **AP6** (Figure 2C) with all studied ds-DNA/RNA give negligible (if any) ICD bands in the range of >300–450 nm, which along with the moderate affinity (Table 2) and the absence of thermal stabilization (Table 1), suggested a non-specific aggregation of molecules along polynucleotides. Such an unspecific binding mode (possible locations are minor/major groove, not excluding weak partial intercalation of the DBTAA moiety) results in DBTAA chromophores characterized by a variety of orienta-

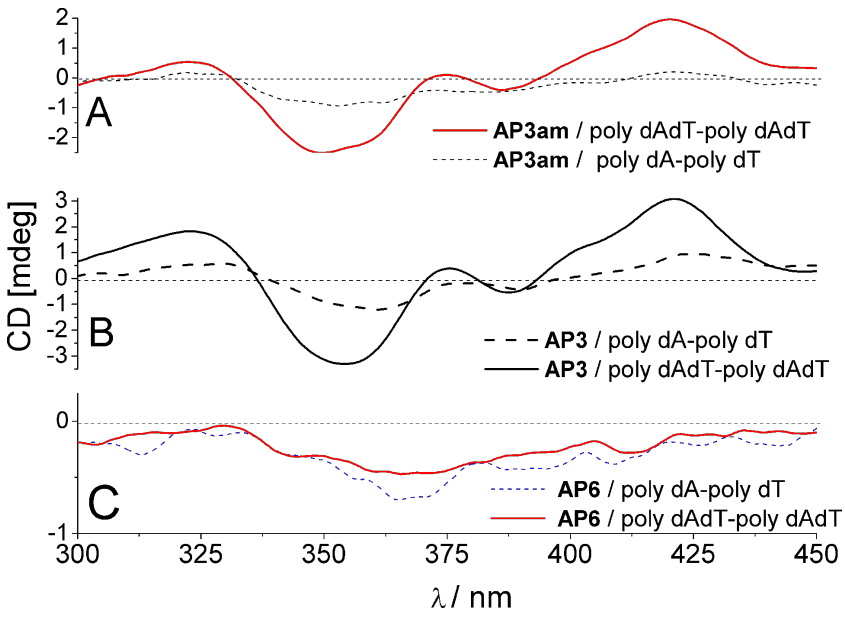



**Figure 2:** Induced CD bands observed for **AP3**, **AP3am**, **AP6** for poly dA–poly dT and poly $(dAdT)_2$ ($c = 3.0 \times 10^{-5}$ mol dm$^{-3}$). The dG–dC and rA–rU polynucleotides did not give any induced (I)CD bands >300 nm.

tions oriented variously in respect to the DNA/RNA chiral axis, which can yield the observed weak non-descriptive ICD pattern [19,20].

By contrast with **APH**/**AP5**/**AP6**, the "propyl" analogues **AP3** and **AP3am** revealed significantly different ICD patterns >300 nm. Strongly induced ICD bands were observed only for AT-containing DNAs (Figure 2A,B), but not for GC-DNA and ds-RNA (Supporting Information File 1). Intriguingly, the ICD pattern observed for **AP3** and **AP3am** (Figure 2A,B; positive–negative–positive sign of ICD bands) were distinctively different from the very weak negative ICD bands of **AP5** (Supporting Information File 1) and **AP6** (Figure 2C). Moreover, save for **AP3**, **AP3am** gave ICD bands (Supporting Information File 1) with mixed sequence ct-DNA (58% AT base pairs).

Although **AP3** and **AP3am** ICD spectra with AT-DNA closely resemble the ICD spectrum of the previously studied **DP77** (no adenine attached to DBTAA), there are several differences between **AP3**/**AP3am** and **DP77**. Only **DP77** [12] gives an ICD spectrum for ds-RNA (poly rA–poly rU) and ICD at low $ratios_{[compd]/[DNA]}$ (0.1; 0.2), which is attributed to **DP77–DBTAA** intercalation. Under the same conditions the ICD bands of **AP3**/**AP3am** exhibit a significant intensity only at r > 0.3, which strongly corroborates the hypothesis of the dimerization of the DBTAA chromophore within the DNA groove [20]. A more detailed analysis of ICD data revealed that ICD bands of **AP3**/**AP3am** were highly sensitive to the secondary structure of DNA/RNA and, in particular, the minor groove properties (Table S1, Supporting Information File 1). For instance, the ICD bands of **AP3**/**AP3am** show the strongest intensity for an alternating AT–AT-polynucleotide characterized by a minor groove size of 6.3 Å, which is ideal for the accommodation of an aromatic dimer (VdW sum of stacked aromatics ~7 Å, Figure 4). In contrast to alternating-AT, the homo-polynucleotide poly dA–poly dT minor groove is much narrower (3.3 Å), and the lower intensity of bisignate ICD bands points to a lower percentage of dimer formation (due to the necessary unwinding and adaptation of homo-polynucleotide to accommodate a DBTAA-dimer). Homo- and alternating-AT-DNA also significantly differ in terms of the charge density and the curvature of the DNA backbone, which can influence interactions of the positively charged arm of **AP3**/**AP3am** with the phosphate backbone and **AP3**/**AP3am**-adenine with thymines of the DNA.

The GC-DNA and AU-RNA minor grooves are not suitable for an aromatic dimer binding. In the case of GC-DNA this is due to the steric hindrance of the guanine amino groups inside the groove. In the case of the AU-RNA, it is because of the very wide and shallow shape, which does not support the binding of a small molecule (Table S1, Supporting Information File 1). This is consistent with the fact that no **AP3**/**AP3am** ICD bands were observed for GC-DNA and ds-RNA.

In a previous study **AP3** revealed an unprecedented recognition of a dT sequence by the appearance of a specific ICD band, while it did not exhibit an ICD signal for other ss-DNA/RNA [11]. A comparison of ICD bands for various ss- and ds-DNA (Figure 3) revealed that the strong negative band at 350 nm is characteristic for ds-DNA, whereas the positive bands at 323 and 421 nm are common for both, ss-dT and ds-AT-DNA. Thus, the negative ICD band at 350 nm can be assigned to a part of the bisignate coupled exciton electronic coupling (EC) ICD band of the **AP3** (also **AP3am**) dimer inside the ds-DNA minor groove. On the other hand, **AP3** binds to poly dT as a single molecule (not stacked with other DBTAA) and, therefore, it is lacking the characteristic bisignate ICD band at 350 nm. Interestingly, the negative ICD band at 350 nm is also visible for ct-DNA (58% AT base pairs), which confirms that any AT-DNA sequence even in mixed polynucleotides will give a characteristic ICD pattern.

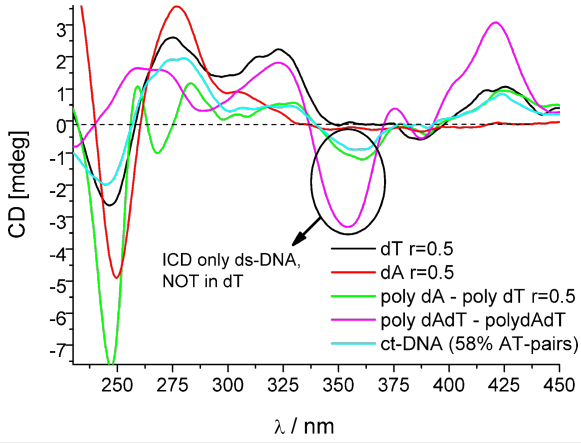



**Figure 3:** Comparison of **AP3** ICD bands for ss-DNA (dA and dT) with alternating- and homo-AT ds-polynucleotides, as well as mixed sequence DNA (ct-DNA, 58% of AT).

## Structure–activity relations of DBTAA–adenine conjugates in AT-DNA binding

Despite the results described above, it is still not clear whether the origin of the DBTAA-ICD pattern proposed for the AT-DNA recognition is an intermolecularly stacked system formed by two DBTAA moieties or, alternatively, a single intramolecularly stacked DBTAA–adenine molecule. The solubility of the system prevented NMR experiments. So we relied on structural differences (linker length connecting DBTAA and adenine) between **AP3**/**AP3am** and **AP5**/**AP6** as free molecules for our proposal of the most likely structure explanation.

The chosen derivatives (**AP5** and **AP3**) were constructed by starting from the crystal structure of their close analogue [13], and then manually self-folded to bring adenine on top of DBTAA. Finally, their chemical bond lengths and angles were corrected by the MM2-optimization in the software Marvin-Suite to assure a realistic presentation (Figure 4). These models serve as a schematic presentation for an easier discussion of the experimental results.

The **AP3**/**AP3am** cannot form efficient intramolecular stacks because the **AP3**/**AP3am** propyl linker is too short to allow adenine to stack above the DBTAA moiety (Figure 4 lower left). Thus, it can be expected that, given an excess of **AP3** over DNA binding sites ($r > 0.3$), two **AP3**-DBTAA moieties will form π–π stacked dimers within the DNA minor groove similar to the crystal structure of close analogue (Figure 4, right). Conversely, the **AP5** and **AP6** aliphatic linkers are long enough to allow self-folding and easy intramolecular π–π stacking of adenine with DBTAA (Figure 4, upper right). During the binding event to ds-DNA (AT minor groove) **AP5** and **AP6** have to combine intramolecular (self-folded) and intermolecular (with DNA) interactions, whereby the best compromise could be the binding of self-folded molecules within a DNA minor groove (see schematic presentation in Figure 5).

The reference compound **APH** has only one side arm interacting with DNA so that a number of orientation options upon DNA/RNA binding is increased, which results in a negligible/non-descriptive ICD pattern. It should be stressed that the absence of **APH**-ICD bands emphasizes the importance of adenine in **AP3**/**AP3am** for the induction of AT-DNA ICD pattern.

## Conclusion

The **AP** derivatives presented herein are found to non-covalently bind to ds-DNA/RNA. The absence of bathochromic shifts in UV–vis titrations, the lack of a thermal stabilization effect on any ds-DNA/RNA, as well as the previously reported [11] viscometry and gel electrophoresis results do not support intercalation as the dominant binding mode (not excluding weak partial intercalation of the DBTAA moiety in some cases). Intrinsically achiral, **AP**-derivatives reveal significantly different induced CD bands upon binding to ds-DNA/RNA. Only DBTAA–propyladenine conjugates (**AP3**/**AP3am**) revealed molecular recognition of AT-DNA by the appearance of an ICD band pattern >300 nm. Significant ICD bands did not appear for other ds-DNA/RNA. A structure–activity relation for the studied series of compounds revealed that the essential structural features for the ICD recognition of AT-DNA are a) the presence of DNA-binding appendages on both DBTAA side chains, i.e., the adenine side chain and the positively charged side chain and b) the presence of a short propyl linker, which does not support intramolecular aromatic stacking between DBTAA and adenine.

The essential difference between **AP3**/**AP3am** and **DP77** is that the former give ICD bands exclusively for AT-DNA, while the latter gives various ICD patterns for all ds-DNA and even ds-RNA. Thus **DP77** is not suitable for the recognition of AT-DNA.

Furthermore, the ICD patterns of **AP3** with ss-DNA (oligo dT) and ds-AT-DNA essentially differ in the appearance of a strong negative band at 350 nm, characteristic for ds-DNA. This band could be used to probe the efficiency of pairing oligo dT

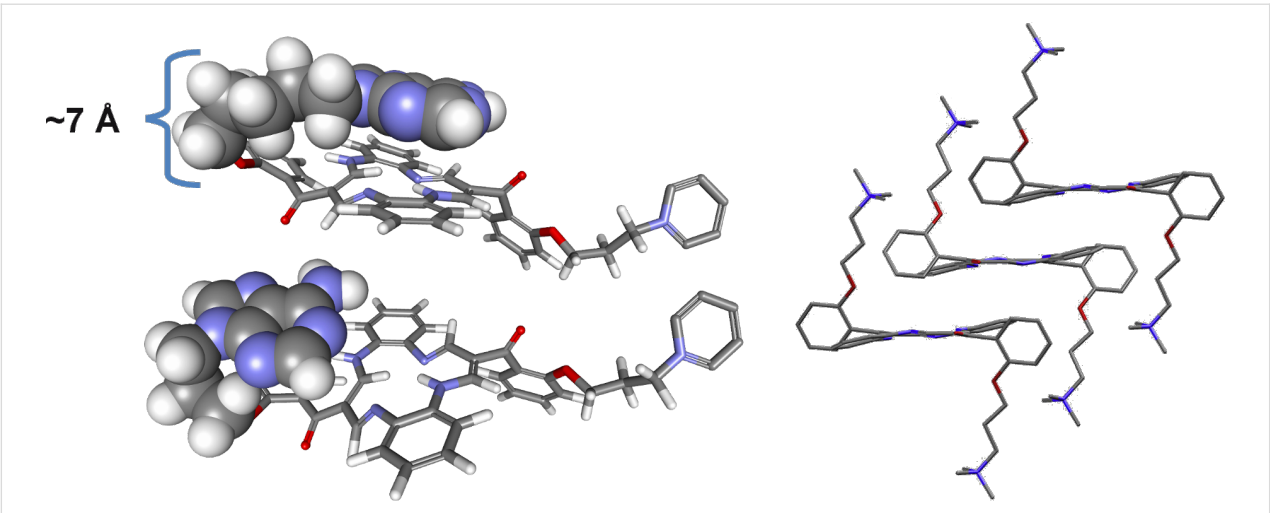



**Figure 4:** Self-folding of **AP5** (upper left – note the stacking of DBTAA and adenine) and **AP3** (lower left – note the short linker which does not allow for the stacking of DBTAA and adenine); (right) detail of three molecules packing in crystal by π–π stacking of parallel DBTAA rings with a distance of ca. 3 Å (see [13] and Figure 8 therein).

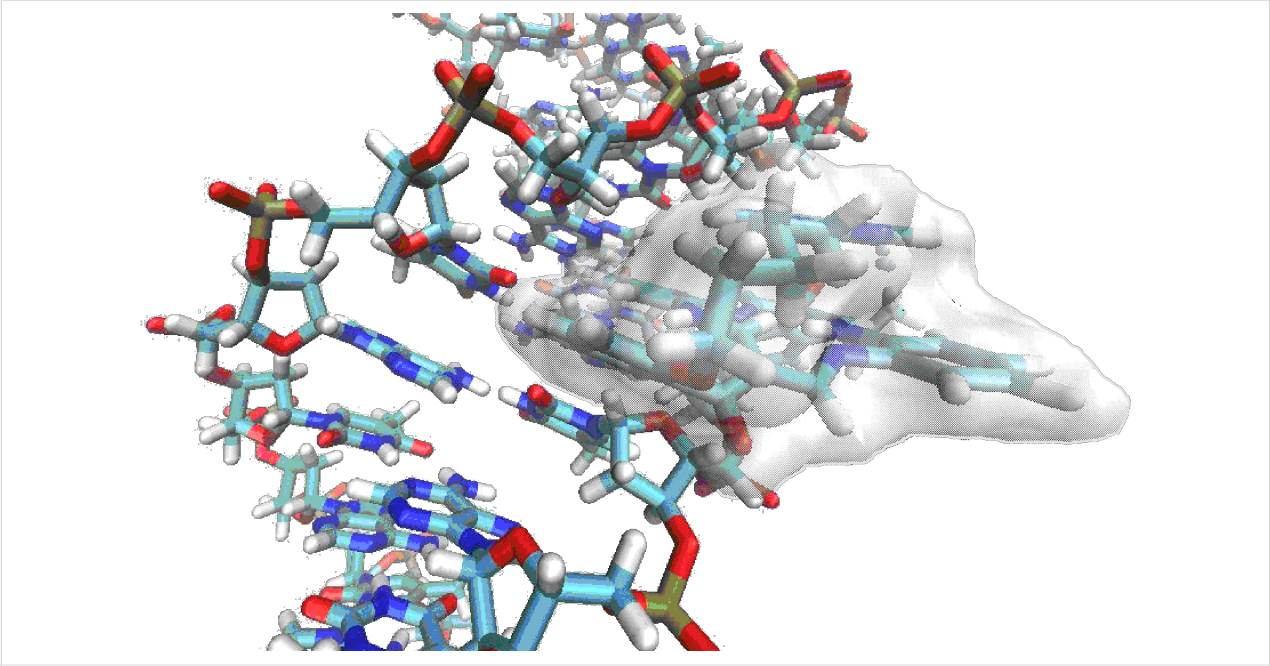


**Figure 5:** Schematic presentation of self-folded **AP5** (free **AP5** see Figure 4; upper left) in the minor groove of poly dAdT–poly dAdT. Double stranded DNA (previously used in [6]) was created with the program NUCGEN, a part of the AMBER11 program suite [21]. The self-folded ligand **AP5** was docked manually into the poly $(dA–dT)_2$ polynucleotide minor groove by using the program VMD [22], taking into account that VdW radii of DNA and ligand do not overlap.

sequences with complementary dA or rA structures, for instance, in any (antisense strategy) developed oligonucleotide, which relies on efficient and precise AT sequence pairing. Namely, the addition of **AP3**/**AP3am** to ss-DNA containing an oligo dT sequence should yield the specific ICD pattern of dT among any other ss-DNA/RNA sequence present in solution [11]. The ICD pattern upon the addition of the exclusively matching oligo A sequence should change by the appearance of a strong negative band at 350 nm, which is characteristic for ds-AT-DNA. Another application could be related to oligo (dT)-cellulose commonly used for tRNA purification [23]. In particular, the difference in **AP3**/**AP3am** ICD pattern between cellulose-tagged dT and coupled double-stranded oligo dT–rA can be reported and its stability and structure studied in detail by monitoring the changes in ICD shape and intensity.

Furthermore, DBTAA–derivatives are azamacrocyclic ligands and thus have additional potential to bind metal cations [24], such metal complexes offering a variety of different interactions with DNA/RNA. Future prospects also include the synthesis of fluorescent DBTAA derivatives, which are expected to transfer observed intriguing DBTAA-DNA/RNA binding properties in a lower concentration range and broaden the set of DBTAA applications by cell-bioimaging. Such applications are also supported by the preliminary bioactivity screening of the AP-series, which revealed low cytotoxicity toward human cell lines.

# Experimental

## General

7,16-Bis[2-(3-bromopropoxy)benzoyl]-5,14-dihydrodibenzo[*b*,*i*][1,4,8,11]tetraazacyclotetradecine, 7,16-Bis[2-hydroxybenzoyl]-5,14-dihydrodibenzo[*b*,*i*][1,4,8,11]tetraazacyclotetradecine and 9-(6-bromopentyl)adenine were prepared by the procedure described earlier [13-15]. All reagents were purchased from commercial sources (Sigma-Aldrich) and were used as received. Solvents were dried by using standard methods and were freshly distilled before use.

$^{1}H$ and $^{13}C$ NMR were run on Bruker AVANCE II 300 and Bruker AVANCE III 600 spectrometers. Chemical shifts (δ) are expressed in parts per million and J values in hertz. Signal multiplicities are denoted as s (singlet), d (doublet), t (triplet), q (quartet), and m (multiplet). The IR-ATR spectra were recorded with a Thermo Fisher Scientific Nicolet IR200. ESI mass spectra were taken on a Bruker Daltonics microTOF-II spectrometer.

## Syntheses

**(7-{2-[3-(Adenin-9-yl)propoxy]benzoyl}-16-{2-[3-(*N*,*N*,*N*-trimethylammonium)propoxy]benzoyl}-5,14-dihydrodibenzo[*b*,*i*][1,4,8,11]tetraazacyclotetradecine bromide) (AP3am):** Similar as described in [11] a mixture consisting of adenine (0.175 g, 1.298 mmol) and 60% NaH (0.039 g, 0.931 mmol) in anhydrous DMF (3 mL) was stirred for 1 h at

room temperature. 7,16-Bis[2-(3-bromopropoxy)benzoyl]-5,14-dihydrodibenzo[*b*,*i*][1,4,8,11]tetraazacyclotetradecine [13] (0.5 g, 0.649 mmol) dissolved in hot anhydrous DMF (40 mL) was added, and the reaction mixture was stirred at 40 °C for 5 min, then for 3 h at room temperature. The reaction mixture was then cooled in a freezer and partitioned between dichloromethane (70 mL) and water (100 mL). The organic layer was separated, washed with water (2 × 100 mL), dried over anhydrous magnesium sulfate, concentrated to a small volume, and chromatographed on a column of silica gel with dichloromethane/methanol (20:0.5 to 20:1 v/v) as an eluent. The main orange fraction was collected and evaporated to dryness. A residue was dissolved in DMF (3 mL) and placed in an ice bath. 0.6 mL of cooled trimethylamine was added, and the mixture was stirred for 5 h at 50 °C. The excess of trimethylamine was removed under diminished pressure, and 30 mL of *tert*-butyl methyl ether was added. Crystallized solid was collected, washed with *tert*-butyl methyl ether, dried in vacuum, and chromatographed on a column with basic aluminum oxide with dichloromethane/methanol (20:0.5 to 20:2 v/v) as an eluent. The main orange fraction was collected, evaporated to dryness, and dissolved in methanol (2 mL). An orange-red microcrystalline product was obtained by the slow diffusion of diethyl ether and the drying isolated solid under vacuum. Yield: 0.046 g (8%). $^1$H NMR (600 MHz, DMSO-$d_6$) δ (ppm) 14.29 (m, 1H, $H^{10}$/$H^{21}$), 14.20 (m, 1H, $H^{10}$/$H^{21}$), 8.55 (d, $J$ = 6.7 Hz, 2H, {$H^7$, $H^9$}/{$H^{18}$, $H^{20}$}), 8.43 (d, $J$ = 6.5 Hz, 2H, {$H^7$, $H^9$}/{$H^{18}$, $H^{20}$}), 7.91 (s, 1H, $H^h$), 7.86 (s, 1H, $H^k$), 7.52 (m, 2H, $H^{27}$, $H^{27'}$), 7.39 (m, 2H, $H^{25}$, $H^{25'}$), 7.09-7.31 (m, 12H, $H^1$-$H^4$, $H^{12}$-$H^{15}$, $H^{26}$, $H^{26'}$, $H^{28}$, $H^{28'}$), 7.07 (s, 2H, $H^m$), 4.09–4.14 (m, 4H, $H^g$, $H^a$), 4.03 (t, $J$ = 5.9 Hz, 2H, $H^e$), 3.28 (m, 2H, $H^c$), 2.92 (s, 9H, $H^d$), 2.11 (m, 2H, $H^f$), 2.03 (m, 2H, $H^b$); $^{13}$C NMR (75 MHz, DMSO-$d_6$) δ (ppm) 191.2, 191.4 ($C^{23}$, $C^{23'}$), 155.7 ($C^l$), 154.8, 155.0 ($C^{29}$, $C^{29'}$), 152.3, 152.7 ($C^7$, $C^9$, $C^{18}$, $C^{22}$), 152.1 ($C^k$), 149.2 ($C^j$), 140.4 ($C^h$), 136.1, 136.3 ($C^5$, $C^{11}$, $C^{16}$, $C^{22}$), 131.4 ($C^{27}$, $C^{27'}$), 129.1, 129.2 ($C^{25}$, $C^{25'}$), 128.7, 128.8 ($C^{24}$, $C^{24'}$), 126.7 ($C^2$, $C^3$, $C^{13}$, $C^{14}$), 120.9, 121.2 ($C^{26}$, $C^{26'}$), 118.6 ($C^i$), 115.3, 115.5 ($C^1$, $C^4$, $C^{12}$, $C^{15}$), 112.7, 113.0 ($C^{28}$, $C^{28'}$), 110.0, 110.2 ($C^{18}$, $C^{19}$), 65.1 ($C^a$, $C^e$), 62.8 ($C^c$), 52.0 ($C^d$), 29.0, 22.5 ($C^b$, $C^f$); ESI-HRMS (*m/z*): $M^+$ calcd for $C_{46}H_{47}N_{10}O_4$, 803.378; found, 803.376; IR-ATR (cm$^{-1}$): 3525, 3374, 3314, 3177, 3083, 2959, 2919, 2868, 1651, 1600, 1562, 1484, 1449, 1417, 1396, 1286, 1263, 1239, 1219.

**(7-{2-[5-(Adenin-9-yl)pentoxy]benzoyl}-16-{2-[5-(*N*-pyridinium-1-yl)pentoxy]benzoyl}-5,14-dihydrodibenzo[*b*,*i*][1,4,8,11]tetraazacyclotetradecine bromide) (AP5):** Similar as described in [11] a reaction mixture consisting of 7,16-Bis[2-hydroxybenzoyl]-5,14-dihydrodibenzo[*b*,*i*][1,4,8,11]tetraazacyclotetradecine [14] (0.2 g, 0.378 mmol), anhydrous potassium carbonate (0.104 g, 0.757 mmol), and 9-(6-bromopentyl)adenine [15] (0.054 g, 0.189 mmol) in anhydrous DMF (40 mL) was stirred for 72 h at room temperature. 1,5-Dibromopentane (0.618 mL, 4.54 mmol) and anhydrous potassium carbonate (0.026 g, 0.189 mmol) were then added, and the stirring was continued for 24 h at room temperature. The reaction mixture was partitioned between dichloromethane (20 mL) and water (100 mL). A small amount of solid KBr was added to improve the separation of the phases. The organic layer was separated and washed thoroughly with water (5 × 30 mL), dried over anhydrous magnesium sulfate, concentrated to a small volume, and chromatographed on a column of silica gel with dichloromethane/methanol (10:0.6 v/v) as an eluent. The second fraction was collected from the two orange ones which displayed the highest intensity. It was evaporated to dryness, dissolved in 3 mL of chloroform, and, once again, chromatographed on a column of silica gel with chloroform/methanol (20:0.4 to 20:0.6 v/v) as an eluent. The main fraction was collected, evaporated to dryness, and a solid residue was dissolved in pyridine (5 mL). The mixture was stirred for 7 h at 45 °C, then pyridine was removed under diminished pressure, and the solid residue was chromatographed on a column with basic aluminum oxide with dichloromethane/methanol (20:0.5 to 20:3 v/v) as an eluent. The main orange fraction was collected and evaporated to dryness. An orange-red microcrystalline product was obtained by drying under vacuum. Yield: 0.035 g (19%). $^1$H NMR (300 MHz, DMSO-$d_6$) δ (ppm) 14.26 (m, 2H, $H^{10}$, $H^{21}$), 8.95 (m, 2H, $H^f$), 8.54 (tt, $J$ = 7.8 Hz, $J$ = 1.3 Hz, 1H, $H^h$), 8.45 (m, 4H, $H^7$, $H^9$, $H^{18}$, $H^{20}$), 8.05 (m, 2H, $H^g$), 7.95 (s, 1H, $H^n$), 7.86 (s, 1H, $H^q$), 7.50 (m, 2H, $H^{27}$, $H^{27'}$), 7.33 (dd, $J$ = 3.7 Hz, $J$ = 1.8 Hz, 1H, $H^{25}$/$H^{25'}$), 7.04–7.26 (m, 15H, $H^1$-$H^4$, $H^{12}$-$H^{15}$, $H^{25}$/$H^{25'}$, $H^{26}$, $H^{26'}$, $H^{28}$, $H^{28'}$, $H^s$), 4.43 (t, $J$ = 7.6 Hz, 2H, $H^e$), 4.00 (m, 4H, $H^a$, $H^m$), 3.87 (t, $J$ = 7.3 Hz, 2H, $H^i$), 1.78 (m, 2H, aliphatic chain), 1.59 (m, 6H, aliphatic chain), 1.24 (m, 4H, aliphatic chain); $^{13}$C NMR (75 MHz, DMSO-$d_6$) δ (ppm) 191.3, 191.4 ($C^{23}$, $C^{23'}$), 155.7 ($C^r$), 155.1, 155.1 ($C^{29}$, $C^{29'}$), 152.4, 152.5 ($C^7$, $C^9$, $C^{18}$, $C^{20}$), 152.1 ($C^q$), 149.3 ($C^p$), 145.3 ($C^h$), 144.4 ($C^f$), 140.2 ($C^n$), 136.1 ($C^5$, $C^{11}$, $C^{16}$, $C^{22}$), 131.3, 131.5 ($C^{27}$, $C^{27'}$), 129.0, 129.2 ($C^{25}$, $C^{25'}$), 128.6 ($C^{24}$, $C^{24'}$), 127.9 ($C^g$), 126.7 ($C^2$, $C^3$, $C^{13}$, $C^{14}$), 120.7 ($C^{26}$, $C^{26'}$), 118.5 ($C^o$), 115.2 ($C^1$, $C^4$, $C^{12}$, $C^{15}$), 112.5, 112.5 ($C^{28}$, $C^{28'}$), 109.9, 110.0 ($C^8$, $C^{19}$), 67.5, 67.5 ($C^a$, $C^i$), 60.3 ($C^e$), 42.4 ($C^m$), 21.8, 22.4, 27.6, 27.8, 28.9, 30.1 ($C^b$–$C^d$, $C^j$–$C^l$); ESI-HRMS (*m/z*): $M^+$ calcd for $C_{52}H_{51}N_{10}O_4$, 879.409; found, 879.407; IR-ATR (cm$^{-1}$): 1249, 1286, 1412, 1446, 1483, 1560, 1588, 1643, 2863, 2935, 3057, 3317.

## Spectrophotometric studies

The UV–vis spectra were recorded on a Varian Cary 100 Bio spectrophotometer and the CD spectra on a JASCO J815 spectrophotometer at 25 °C with appropriate 1 cm path quartz

cuvettes. The study of interactions with DNA and RNA was carried out with aqueous solutions of compounds buffered to pH 7.0 (buffer sodium cacodylate, $I$ = 0.05 mol dm$^{-3}$).

Polynucleotides were purchased as noted: poly dAdT–poly dAdT, poly dG–poly dC, poly dA–poly dT, poly A–poly U (Sigma-Aldrich, St. Louis. USA), calf thymus (ct)-DNA (Aldrich). Polynucleotides were dissolved in sodium cacodylate buffer, $I$ = 0.05 mol dm$^{-3}$, pH 7.0. The calf thymus ct-DNA was additionally sonicated and filtered through a 0.45 mm filter [25]. The polynucleotide concentration was spectroscopically determined as the concentration of nucleobases. Spectrophotometric titrations were performed at pH 7.0 ($I$ = 0.05 mol dm$^{-3}$, buffer sodium cacodylate) by adding portions of polynucleotide solution into the solution of the studied compound for UV–vis. CD experiments were carried out by adding portions of compound stock solution into the solution of polynucleotide. Titration data were processed by the Scatchard equation [17,18]. The values for $K_s$ and $n$ given in Table 2 all have satisfactory correlation coefficients (>0.999). The thermal melting curves for DNA, RNA and their complexes with the studied compounds (Table 1) were determined as previously described [16,26] by following the absorption change at 260 nm as a function of the temperature. The absorbance of the ligands was subtracted from every curve, and the absorbance scale was normalized. $T_m$ values are the midpoints of the transition curves determined from the maximum of the first derivative and checked graphically by the tangent method [26]. The $\Delta T_m$ values were calculated by subtracting $T_m$ of the free nucleic acid from $T_m$ of the complex. Every reported $\Delta T_m$ value was the average of at least two measurements. The error of $\Delta T_m$ is ±0.5 °C.

## Supporting Information

### Supporting Information File 1

Additional NMR spectra for new compounds, additional UV–vis and CD spectra.
[http://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-10-225-S1.pdf]

## Acknowledgements

Financial support by the Ministry of Science, Education, Sport of Croatia (098-0982914-2918), the Grant (WCh-BW) Jagiellonian University, Krakow, Poland, the European Regional Development Fund, and the Polish Innovation Economy Operational Program (contract POIG.02.01.00-12-023/08) is acknowledged. This work was also supported by FP7-REGPOT-2012-2013-1, Grant Agreement Number 316289 – InnoMol. We are grateful to Sanja Tomić for providing us with the constructed and optimized structure of poly $(dAdT)_2$, previously used in [6].

## References

1. Demeunynck, M.; Bailly, C.; Wilson, W. D., Eds. *DNA and RNA Binders;* Wiley-VCH: Weinheim, Germany, 2002. doi:10.1002/3527601783
2. Trinquet, E.; Mathis, G. *Mol. BioSyst.* **2006,** *2,* 380–387. doi:10.1039/b607152j
3. Tumir, L.-M.; Piantanida, I. *Mini-Rev. Med. Chem.* **2010,** *10,* 299–308. doi:10.2174/138955710791330990
4. Garcia-España, E.; Piantanida, I.; Schneider, H.-J. Nucleic Acids as Supramolecular Targets. In *Supramolecular Systems in Biomedical Fields;* Schneider, H.-J., Ed.; Royal Society of Chemistry: Cambridge, 2013; pp 213–259. doi:10.1039/9781849737821-00213
5. Rohs, R.; West, S. M.; Sosinsky, A.; Liu, P.; Mann, R. S.; Honig, B. *Nature* **2009,** *461,* 1248–1253. doi:10.1038/nature08473
6. Tumir, L.-M.; Crnolatac, I.; Deligeorgiev, T.; Vasilev, A.; Kaloyanova, S.; Branilović, M. G.; Tomić, S.; Piantanida, I. *Chem. – Eur. J.* **2012,** *18,* 3859–3864. doi:10.1002/chem.201102968
7. Crnolatac, I.; Tumir, L.-M.; Lesev, N. Y.; Vasilev, A. A.; Deligeorgiev, T. G.; Mišković, K.; Glavaš-Obrovac, L.; Vugrek, O.; Piantanida, I. *ChemMedChem* **2013,** *8,* 1093–1103. doi:10.1002/cmdc.201300085
8. Grabar Branilović, M.; Tomić, S.; Tumir, L.-M.; Piantanida, I. *Mol. BioSyst.* **2013,** *9,* 2051–2062. doi:10.1039/c3mb25578f
9. Tumir, L.-M.; Grabar, M.; Tomić, S.; Piantanida, I. *Tetrahedron* **2010,** *66,* 2501–2513. doi:10.1016/j.tet.2010.01.063
10. Tumir, L.-M.; Piantanida, I.; Juranović, I.; Meić, Z.; Tomić, S.; Žinić, M. *Chem. Commun.* **2005,** 2561–2563. doi:10.1039/b500617a
11. Radić Stojković, M.; Škugor, M.; Tomić, S.; Grabar, M.; Smrečki, V.; Dudek, Ł.; Grolik, J.; Eilmes, J.; Piantanida, I. *Org. Biomol. Chem.* **2013,** *11,* 4077–4085. doi:10.1039/c3ob40519b
12. Radić Stojković, M.; Marjanović, M.; Pawlica, D.; Dudek, L.; Eilmes, J.; Kralj, M.; Piantanida, I. *New J. Chem.* **2010,** *34,* 500–507. doi:10.1039/b9nj00490d
13. Pawlica, D.; Radić Stojković, M.; Dudek, Ł.; Piantanida, I.; Sieroń, L.; Eilmes, J. *Tetrahedron* **2009,** *65,* 3980–3989. doi:10.1016/j.tet.2009.03.030
14. Sigg, I.; Haas, G.; Winkler, T. *Helv. Chim. Acta* **1982,** *65,* 275–279. doi:10.1002/hlca.19820650128
15. Itahara, T. *J. Chem. Soc., Perkin Trans. 2* **1996,** 2695–2700. doi:10.1039/p29960002695
16. Mergny, J.-L.; Lacroix, L. *Oligonucleotides* **2003,** *13,* 515–537. doi:10.1089/154545703322860825
17. Scatchard, G. *Ann. N. Y. Acad. Sci.* **1949,** *51,* 660–672.
18. McGhee, J. D.; von Hippel, P. H. *J. Mol. Biol.* **1976,** *103,* 679–684.
19. Rodger, A.; Norden, B., Eds. *Circular Dichroism and Linear Dichroism;* Oxford University Press: New York, 1997.
20. Eriksson, M.; Nordén, B. Linear and Circular Dichroism of Drug-Nucleic Acid Complexes. In *Methods in Enzymology;* Chaires, J. B.; Waring, M. J., Eds.; Academic Press: San Diego, 2001; Vol. 340, pp 68–98.
21. Structure given by courtesy of Tomić, S.. http://ambermd.org/.
22. Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996,** *14,* 33–38. doi:10.1016/0263-7855(96)00018-5
23. Sambrook, J.; Fritsch, E. F.; Maniatis, T., Eds. *Molecular Cloning, a laboratory manual;* Cold Spring Harbor Laboratory Press: New York, 1989.
24. Lewiński, K.; Eilmes, J. *J. Inclusion Phenom. Macrocyclic Chem.* **2005,** *52,* 261–266. doi:10.1007/s10847-004-8025-2
25. Chaires, J. B.; Dattagupta, N.; Crothers, D. M. *Biochemistry* **1982,** *21,* 3933–3940. doi:10.1021/bi00260a005

26. Piantanida, I.; Palm, B. S.; Čudić, P.; Žinić, M.; Schneider, H.-J. *Tetrahedron* **2004,** *60,* 6225–6231. doi:10.1016/j.tet.2004.05.009

## License and Terms

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

# Solution phase synthesis of short oligoribonucleotides on a precipitative tetrapodal support

Alejandro Gimenez Molina, Amit M. Jabgunde, Pasi Virta and Harri Lönnberg*

Full Research Paper

Open Access

Address:
Department of Chemistry, Faculty of Mathematics and Natural Sciences, University of Turku, FI-20014, Turku, Finland

Email:
Harri Lönnberg* - harlon@utu.fi

* Corresponding author

Keywords:
nucleic acids; oligoribonucleotides; phosphoramidite; soluble support; synthesis

*Beilstein J. Org. Chem.* **2014,** *10,* 2279–2285.
doi:10.3762/bjoc.10.237

Received: 13 June 2014
Accepted: 18 September 2014
Published: 29 September 2014

This article is part of the Thematic Series "Nucleic acid chemistry".

Guest Editor: H.-A. Wagenknecht



## Abstract

An effective method for the synthesis of short oligoribonucleotides in solution has been elaborated. Novel 2'-*O*-(2-cyanoethyl)-5'-*O*-(1-methoxy-1-methylethyl) protected ribonucleoside 3'-phosphoramidites have been prepared and their usefulness as building blocks in RNA synthesis on a soluble support has been demonstrated. As a proof of concept, a pentameric oligoribonucleotide, 3'-UUGCA-5', has been prepared on a precipitative tetrapodal tetrakis(4-azidomethylphenyl)pentaerythritol support. The 3'-terminal nucleoside was coupled to the support as a 3'-*O*-(4-pentynoyl) derivative by Cu(I) promoted 1,3-dipolar cycloaddition. Couplings were carried out with 1.5 equiv of the building block. In each coupling cycle, the small molecular reagents and byproducts were removed by two quantitative precipitations from MeOH, one after oxidation and the second after the 5'-deprotection. After completion of the chain assembly, treatment with triethylamine, ammonia and TBAF released the pentamer in high yields.

## Introduction

Recognition of short noncoding RNAs as regulatory elements of gene expression [1-5] has attracted interest in their physicochemical properties, including structure, dynamics of chain invasion and recognition by small molecular entities [6-10]. The quantities of short RNA sequences required for such studies are often larger than what can easily be obtained by lab-scale solid phase synthesis. In other words, there seems to be a need for a straightforward solution phase approach, allowing assembly of short RNAs in a hundreds of milligrams scale. While several such methods for the synthesis of DNA, based either the phosphoramidite [11-16], H-phosphonate [17-19] or phosphotriester chemistry [20-22], have been introduced, none of them has so far been applied to the synthesis of RNA.

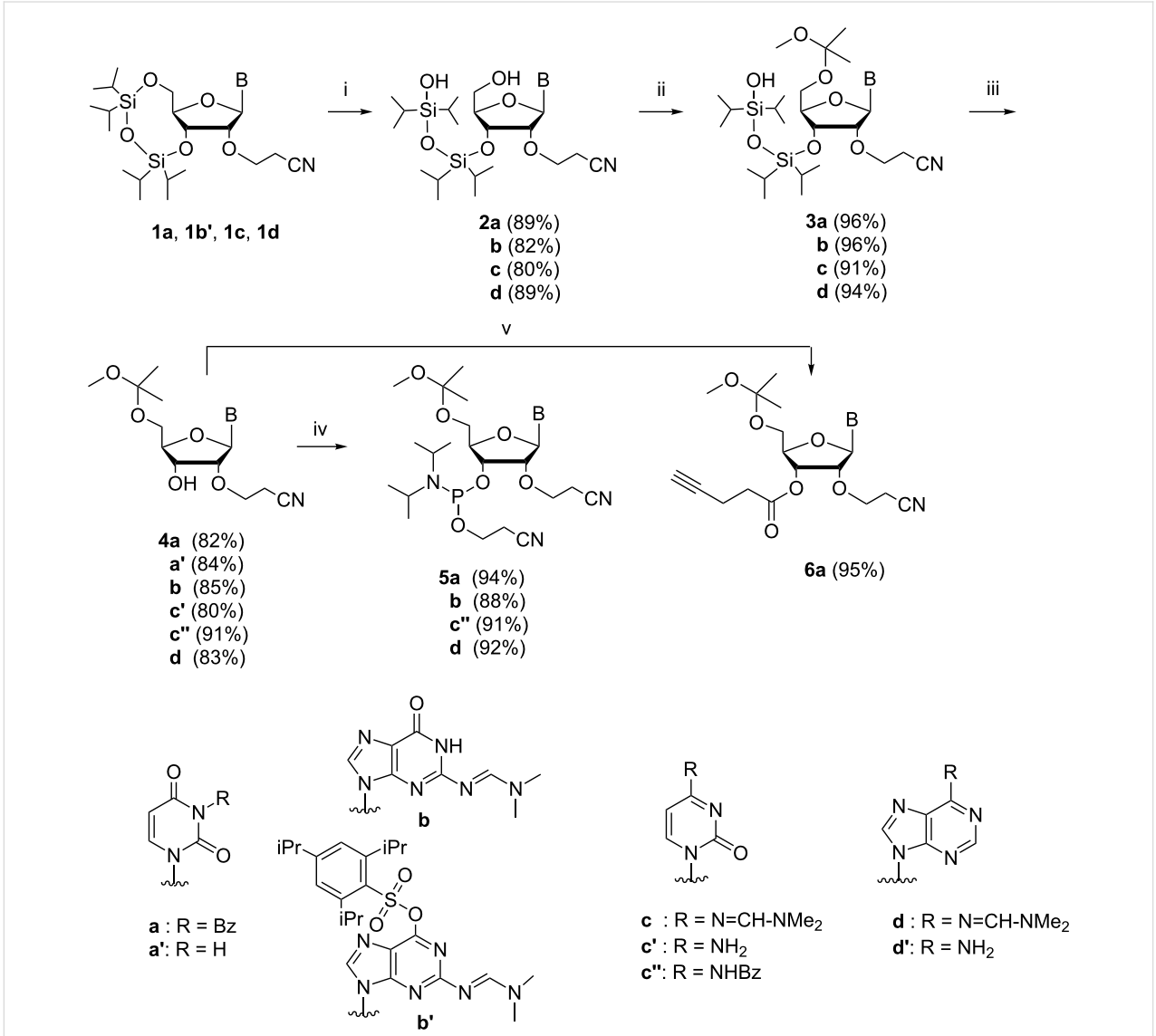
Tetrakis(4-azidomethylphenyl)pentaerythritol has recently been introduced as a practical soluble support for the synthesis of short oligodeoxyribonucleotides [16]. The 3'-terminal nucleoside is attached as a 3'-*O*-(4-pentynoyl) derivative to the support by a Cu(I) catalyzed click reaction and conventional phosphoramidite chemistry is then applied. The advantage of this support

is that the small molecular reagents and byproducts are removed in each coupling cycle by two fully quantitative precipitations from MeOH, one after oxidation and the second after the 5'-deprotection. Owing to the symmetrical tetrapodal structure of the support, the completeness of couplings may be verified by $^1$H NMR spectroscopy. We now report on the synthesis of short oligoribonucleotides on this support. Unfortunately, commercially available building blocks that bear a large hydrophobic protecting group at 2'-O, such as the *tert*-butyldimethylsilyl group, cannot be used, since the increased hydrophobicity of the growing chain prevents precipitation from MeOH. For this reason, 3'-phosphoramidite building blocks bearing the less hydrophobic 2-cyanoethyl group at 2'-O [23] and the 1-methoxy-1-methylethyl group at 5'-O [15] have been prepared and used to assemble a pentamer, 5'-UUGCA-3', on the precipitative pentaerythritol derived support. Precipitation has previously been exploited in conversion of fully protected guanosine 3'-phosphoramidite to cyclic 3',5'-GMP dime by an essentially one-pot synthesis followed by a two-step, one-flask deprotection [24].

# Results and Discussion

## Preparation of nucleosidic building blocks

The synthesis of the nucleosidic building blocks is outlined in Scheme 1. The base moiety protected 2'-*O*-(2-cyanoethyl)-3',5'-*O*-(1,1,3,3-tetraisopropyldisiloxane-1,3-diyl) ribonucleosides (**1a**,**b'**,**c**,**d**), used as starting materials, were prepared as described previously in the literature [23]. The 5'-O–Si bond



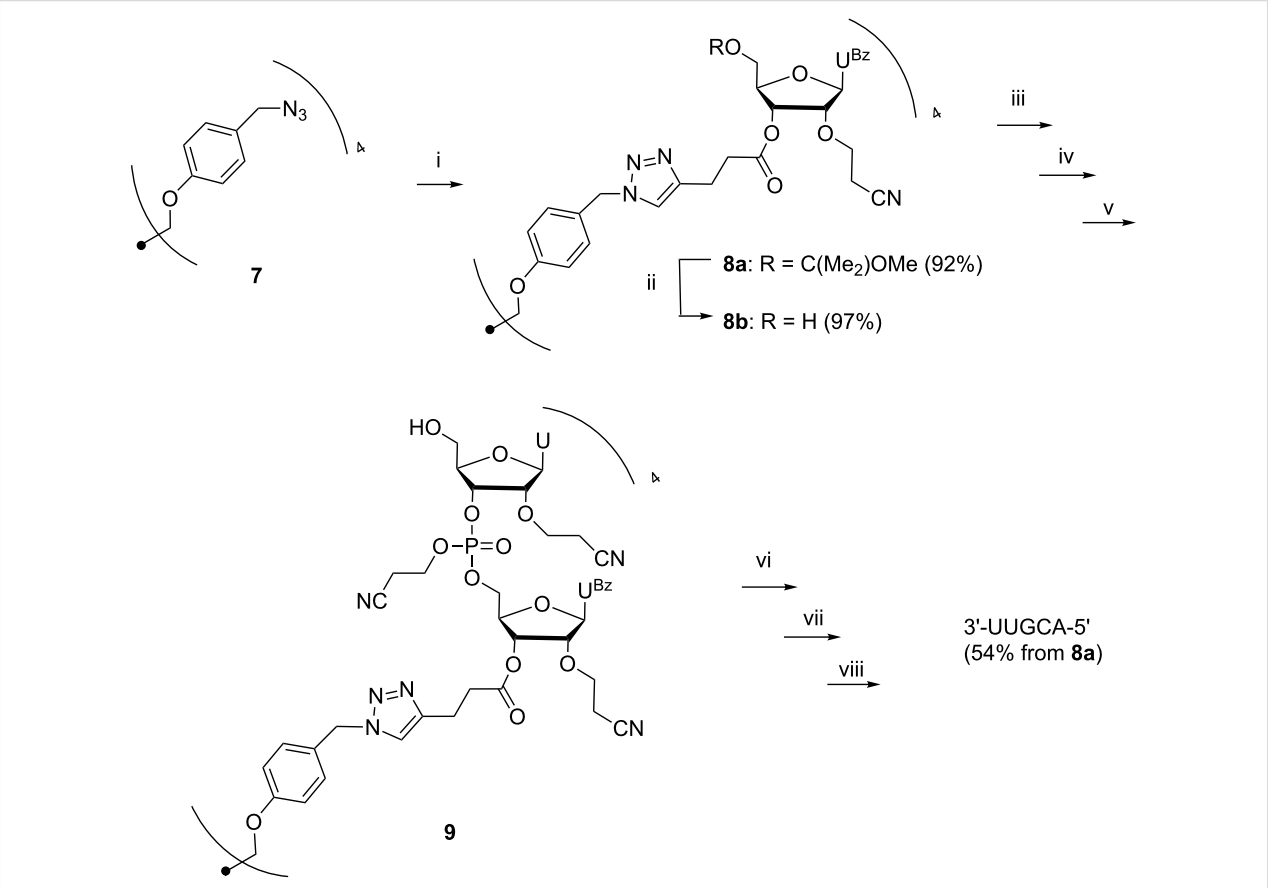

**Scheme 1:** Synthesis of building blocks of the oligoribonucleotide synthesis. (i) TFA, aq THF, 0 °C; (ii) 2-methoxypropene, TsOH, THF; (iii) $NH_4F$, MeOH; (iv) 1-chloro-1-(2-cyanoethoxy)-*N*,*N*-diisopropylphosphinamine, DIPEA, DCM; (v) 1. 4-pentynoic acid, DCC, dioxane, 2. Py, DMAP (cat).

was selectively hydrolyzed with TFA in aqueous THF, leaving the 3'-O–Si linkage intact, and the exposed 5'-OH of compounds **2a**–**d** was subjected to acid-catalyzed acetalization with 2-methoxypropene in THF. The subsequent 3'-*O* desilylation of the fully protected nucleosides (**3a**–**d**) then required careful adjustment of conditions. Desilylation with $NH_4F$ in MeOH turned out to be successful in the sense that the 2'-O and 5'-O protecting groups remained intact, whereas TEA·3HF in THF partly removed the 5'-*O*-(1-methoxy-1-methylethyl) group and TBAF in THF the 2'-*O*-(2-cyanoethyl) group. Unfortunately, even the reaction with $NH_4F$ was in some cases accompanied by removal of the base moiety protecting groups. Cytidine derivative **3c** lost entirely and adenosine derivative **3d** partially the dimethylaminomethylene group, yielding **4c'** and **4d'**, respectively. Reintroduction of the same protecting group to **4d'** gave the desired **4d** in high yield, whereas **4c'** was benzoylated to obtain **4c''**. In addition, the uridine derivative **3a** underwent partial removal of the 3-benzoyl group from **3a** giving **4a'** in addition to **4a**. The phosphoramidite building blocks (**5a'**,**b**,**c''**,**d**) were finally obtained by phosphitylation of compounds **4a'**, **4b**, **4c''** and **4d** with 1-chloro-1-(2-cyanoethoxy)-*N*,*N*-diisopropylphosphinamine in MeCN in the presence of *N*,*N*-diisopropylethylamine.

To enable immobilization of the 3'-terminal nucleoside to the azido functionalized support, 3-benzoyl-2-*O*-cyanoethyl-5-(1-methoxy-1-methylethyl)uridine (**4a**) was transformed to its 3'-*O*-(4-pentynoyl) derivative **6a**. To accomplish this, 4-pentynoic acid was first converted to its anhydride by DCC activation in dioxane and this was then used for the 3'-*O*-acylation of **4a** in pyridine in the presence of 4-dimethylaminopyridine. Experimental details for the preparation of all the building blocks and NMR and MS data for their characterization are given in Supporting Information File 1.

## Oligonucleotide synthesis

Previously reported [16] pentaerythritol-derived tetraazido support **7** was used to assemble a pentameric oligoribonucleotide, 5'-UUGCA-3', from the building blocks described above. The 3'-terminal block **6a** was first attached to the support by Cu(I) promoted 1,3-dipolar cycloaddition [25,26] (Scheme 2). The procedure employed was essentially the same




**Scheme 2:** (i) **6a**, CuI, sodium ascorbate, DMAc; (ii) HCl in dioxane/MeOH 2:1, (iii) **5a'**, 4,5-dicyanoimidazole, MeCN/DMF 1:1; (iv) 1. $I_2$, THF, 2,6-lutidine, $H_2O$, 2. $P(OMe)_3$, 3. Precipitation from MeOH; (v) 1. HCl in DCM/MeOH 5:2, 2. Precipitation from MeOH; (vi) steps iii–v repeated to coupled **5b**, **5c''** and **5d** in this order; (vii) 1. TEA, 2. $NH_4OAc$, aq $NH_3$ (25%), 3. TBAF, $iPrNH_2$, THF; (viii) Precipitation from EtOH with NaOAc.

as described for the attachment of similar 3'-*O*-(4-pentynoyl) acylated 2'-deoxyribonucleosides [16,22]. The tetrapodal nucleoside cluster (**8a**) obtained was purified by silica gel chromatography and the 5'-*O*-(1-methoxy-1-methylethyl) protecting groups were then removed by acid-catalyzed methanolysis with HCl in a mixture of MeOH and dioxane. Removal of volatiles under reduced pressure and treatment of the residue with diethyl ether gave the deprotected support (**8b**) as a white powder. The homogeneity of **8b** was verified by RP-HPLC (see Supporting Information File 1).

The synthetic cycle was composed of three steps, as usual: coupling, oxidation and deprotection. Couplings were carried out in a 1:1 mixture of MeCN and DMF under $N_2$ using 1.5 equiv of the respective phosphoramidite building block per a support-bound 5'-hydroxy group. After coupling, the phosphite triester obtained was oxidized to phosphate triester with iodine in aq THF containing 2,6-lutidine, and the reaction was quenched with $P(OMe)_3$ in DMF. The mixture was concentrated to oil and the support was precipitated from cold MeOH. HPLC analyses of both the filtrate and the precipitate indicated that the support-bound oligonucleotides precipitated quantitatively from MeOH, while the unreacted building block and small molecular reagents remained in the filtrate (Figure 1). No sign of the HPLC signal referring to the support bound nucleotides ($t_R$ = 20.14 min) could be detected in the liquid phase. In addition to the signal at $t_R$ = 20.14 min, a minor signal at $t_R$ = 19.51 min appears in the HPLC trace of the precipitate. This refers to a compound having lost one benzoyl protecting group.

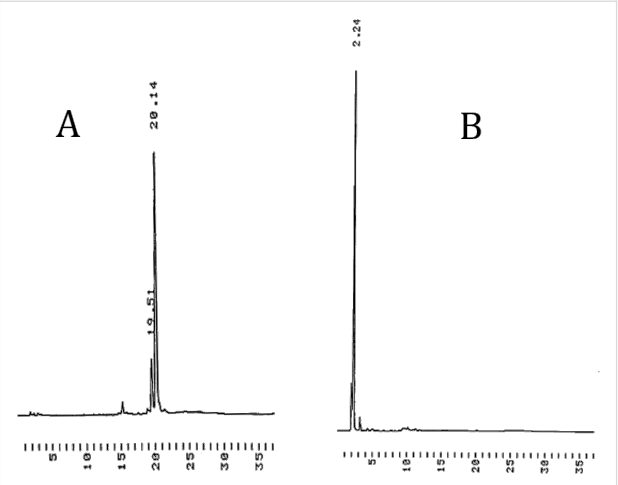



**Figure 1:** HPLC traces referring to precipitation of the tetrapodal support bearing fully protected UU-dimers (**9**), i.e., after the oxidation step: (A) stands for the precipitate and (B) for the filtrate after the precipitation. Chromatographic conditions: a Thermo ODS Hypersil C18 (250 × 4.6 mm, 5 µm) column eluted with a mixture of MeCN and aq $Et_3N$ (0.1 mol $L^{-1}$) at flow rate 1 mL $min^{-1}$. A linear gradient from MeCN 25% at *t* = 0 min to MeCN 100% at *t* = 25 min.

The white precipitate was then disolved in a 5:2 mixture (v/v) of DCM and MeOH and subjected to HCl-catalyzed methanolysis ([HCl] = 0.015 mol $L^{-1}$). After 15 min, the mixture was neutralized by the addition of pyridine, concentrated to oil and the support was again precipitated from cold MeOH (Figure 2). The minor signal at 13 min (a) on the HPLC trace of the precipitate (A) refers to a support having one 2-cyanoethyl group removed.

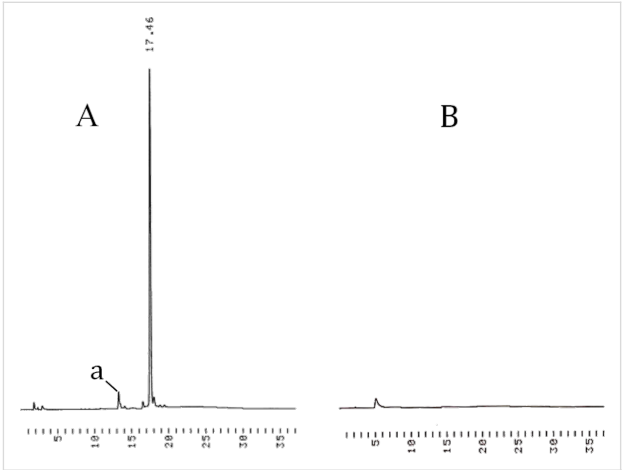



**Figure 2:** HPLC traces referring to precipitation of the tetrapodal support bearing 5'-*O*-deprotected UU-dimers, i.e., after the deprotection step: (A) stands for the precipitate and (B) for the filtrate after the precipitation. For the chromatographic conditions, see the legend of Figure 1.

The identity of the product was after each coupling cycle, i.e., after removal of the 5'-O-(methoxy-1-methylethyl) protecting group, verified by ESIMS. The data obtained are given in the Table 1. The completeness of the coupling was verified by HPLC. To ensure quantitative coupling, an exceptionally long coupling time (12 h) was employed. The coupling time was not really optimized, but preliminary studies revealed that after 4 h a minor part of the support still contained one unreacted branch. Since 4,5-dicyanoimidazole afforded almost quantitative coupling, no other activators were tested.

Upon completion of the assembly of 3'-UUGCA-5', the oligonucleotide was released from the support and deprotected by treatment with triethylamine, ammonia and finally with TBAF. The deprotected pentamer was precipitated from EtOH with NaOAc, The overall yield of the pentamer was 54% (according to the UV absorption), corresponding to 86% average coupling yield. Figure 3 shows the homogeinety and identity of the pentamer.

## Conclusion

An effective method for the synthesis of short oligoribonucleotides in solution has been elaborated. The main characteristics of this method are the use of a precipitative pentaerythritol-

**Table 1:** Negative ion ESIMS of the tetrapodal pentaerythritol (PE) soluble support bearing 5'-*O*-deprotected oligonucleotides.

| Compound | *m*/*z* obsd. | *m*/*z* calcd |
|---|---|---|
| PE-$(U^{Bz})_4$ | 1327.8 $[M + 2Cl]^{2-}$ | 1327.7 $[M + 2Cl]^{2-}$ |
| PE-$[U^{Bz}U]_4$ | 2116.4 $[M - 2H]^{2-}$ | 2116.8 $[M - 2H]^{2-}$ |
| PE$[U^{Bz}UG^{dmf}]_4$ | 3130.8 $[M - 2H]^{2-}$ | 3128.6 $[M - 2H]^{2-}$ |
| PE$[U^{Bz}UG^{dmf}C^{Bz}]_4$ | 2079.5 $[M - 4H]^{4-}$ | 2079.7 $[M - 4H]^{4-}$ |
| PE$[U^{Bz}UG^{dmf}C^{Bz}A^{dmf}]_4$ | 2570.0 $[M - 4H]^{4-}$ | 2570.1 $[M - 4H]^{4-}$ |

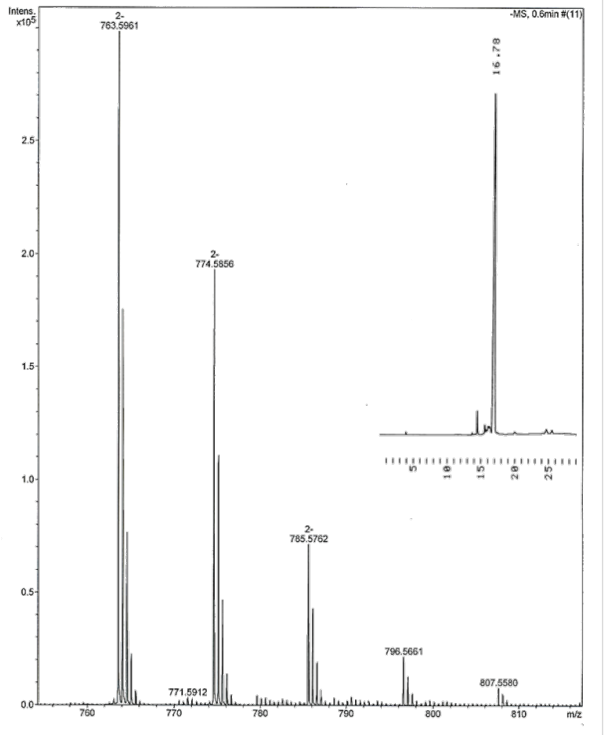



**Figure 3:** Negative ion ESIMS and HPLC traces of the isolated 3'-UUGCA-5'. For the MS over a wider mass range, see Supporting Information File 1. Chromatographic conditions: a Thermo ODS Hypersil C18 (250 × 4.6 mm, 5 µm) column eluted with a mixture of MeCN and $H_2O$ at flow rate 1 mL $min^{-1}$. A linear gradient from MeCN 0% at *t* = 0 min to MeCN 100% at *t* = 20 min. The minor signal at 14.5 min refers to benzamide.

derived soluble support and small reasonably polar sugar moiety protecting groups, viz. the 2'-*O*-(2-cyanoethyl) and the 5'-*O*-(1-methoxy-1-methylethyl) group. These together allow facile purification after each coupling/oxidation and 5'-*O*-deprotection step by quantitative precipitation from MeOH. As a proof of concept, a 3'-UUGCA-5' pentamer was prepared in 54% yield by using 1.5 equiv of the monomeric phosphoramidite building block per 5'-hydroxy function. Release from the support and removal of the cyanoethyl group from the phosphate and the base moiety protecting groups was achieved by consecutive treatments with triethylamine and aqueous $NH_3$/$NH_4OAc$. Treatment of the released oligomer with 1 mol $L^{-1}$ TBAF in THF removed the 2'-*O*-(2-cyanoethyl) protecting group and the deprotected oligomer was isolated by precipitation from EtOH with NaOAc.

## Experimental

**General**. RP-HPLC separations were carried out on a Thermo ODS Hypersil C18 (250 × 4.6 mm, 5 µm) analytical column using UV detection (λ = 260 nm). NMR spectra were recorded on a Bruker Avance spectrometer (500 or 400 MHz) at 25 °C. Chemical shifts are given in ppm from TMS. HRMS analysis were recorded on a Bruker Daltonics spectrometer.

For the preparation and characterization of building blocks **1a**,**b'**,**c**,**d**, **2a**–**d**, **3a**–**d**, **4a**,**a'**,**b**,**c'**,**d**, **5a'**,**b**,**c''**,**d** and **6a**, see the Supporting Information File 1.

**Tetrakis-*O*-(4-azidomethylphenyl)pentaerythritol** (**7**). Preparation of compound **7** has been described previously [16].

**Tetrakis-*O*-{4-{3-[3-benzoyl-2'-*O*-(2-cyanoethyl)-5'-*O*-(1-methoxy-1-methylethyl)uridin-3'-*O*-yl]-3-oxoprop-1-yl}-1*H*-1,2,3-triazol-1-ylmethylphenyl}pentaerythritol** (**8a**). Compounds **6a** (0.544 g; 0.98 mmol) and **7** (0.130 g; 0.20 mmol) were dissolved in dry DMAc (5 mL) in a pyrex tube. The slightly yellow-brownish solution was subjected to degassing according to a previous procedure [27]. A catalytic amount of CuI and sodium ascorbate were added under $N_2$ and the mixture was stirred for 20 h at 45 °C. The solvent was removed under reduced pressure and the residue was subjected to column chromatography on silica gel using a gradient of 1–5% MeOH in DCM containing 1% TEA. Support **8a** was obtained as a white foam in 92% yield (0.520 g; 0.18 mmol).

**Tetrakis-*O*-{4-{3-[3-benzoyl-2'-*O*-(2-cyanoethyl)uridin-3'-*O*-yl]-3-oxoprop-1-yl}-1*H*-1,2,3-triazol-1-ylmethylphenyl}pentaerythritol** (**8b**). Support **8a** (0.520 g; 0.18 mmol) was dissolved in 2:1 mixture (v/v) of dioxane and MeOH (20 mL) and 0.5 mL of 0.1 mol $L^{-1}$ HCl in dioxane was added. The yellowish solution was stirred for 1 h at room temperature and the solvent removed under reduced pressure. A small amount of pyridine was added to neutralize any traces of acid. $Et_2O$

(50 mL) was added affording **8b** instantaneously as white precipitate (0.453 g; 0.17 mmol) in 97% yield. The identity of the product was verified by ESIMS (Table 1) and the homogeinity by HPLC on a Thermo ODS Hypersil C18 (250 × 4.6 mm, 5 μm) column eluted with a mixture of MeCN and $H_2O$ at flow rate 1 mL $min^{-1}$, using a linear gradient from MeCN 25% at $t$ = 0 min to MeCN 100% at $t$ = 25min (for the HPLC traces, see Supporting Information File 1).

**Oligonucleotide synthesis**. Support **8b** (0.151 g; 0.056 mmol) and 6 equiv of the phosphoramidite building block **5a'** (200 mg; 0.35 mmol; 1.5 equiv per branch) were dissolved in an 1:1 mixture (v/v) of MeCN and DMF (5 mL), and 4,5-dicyanoimidazole in MeCN (1.39 mL of 0.25 mol $L^{-1}$ solution, 0.32 mmol) was added. The mixture was stirred for 12 h at room temperature and then $I_2$ (0.13 g; 0.5 mmol) in 4:2:1 mixture (v/v) of THF, water and 2,6-lutidine (4 mL) was added and the mixture was stirred for 20 min. Unreacted $I_2$ was quenched by the addition of $P(OMe)_3$ (0.1 mmol), the mixture was concentrated to oil and the oligonucleotide was precipitated in quantitative yields from cold MeOH (75 mL). The white precipitate was filtered over celite (AW) and isolated.

The white precipitate obtained was dissolved in a 5:2 mixture of DCM and MeOH (35 mL) and 5 mL of a 0.125 mol $L^{-1}$ HCl in MeOH solution were added. After 15 min, pyridine was added to neutralize the acid catalyst. The solution was concentrated to oil and cold MeOH (50 mL) was added. The white precipitate obtained instantaneously was collected by filtration through celite (AW) and dried under vacuum. The identitiy of the clustered dimer PE[3'-$U^{Bz}$U-5'-OH]$_4$ (**9**) was verified by ESIMS (Table 1) and the homogeneity by HPLC (Figure 2). The gravimetrically determined overall yield was 93% (230 mg).

The coupling cycle (coupling, oxidation, precipitation, deprotection and precipitation) was then repeated to couple building blocks **5b**, **5c''** and **5d**, in this order. The identity and homogeneity of the product was verified after each cycle by ESIMS (Table 1) and HPLC.

**Release and deprotection**. A sample of the support bound cluster, 3'-$U^{Bz}UG^{dmf}C^{Bz}A^{dmf}$-5' (20 mg), was dissolved in an 1:1 mixture of DMF and MeCN (4 mL) and treated with TEA (2 mL) to release the protected oligonucleotide from the support [23]. To the mixture, aq $NH_3$ (25%) containing 10% $NH_4OAc$ (w/v) was added and the solution was stirred for 12 h at room temperature [28]. The solvent was removed under reduced pressure and 1 mol $L^{-1}$ solution of TBAF in dry THF (4 mL) and isopropylamine (200 μL) were added. The reaction was monitored by HPLC and was observed to be completed in 12 h at room temperature. The homogeneity and identity of the deprotected pentamer, 3'-UUGCA-5', was verified by HPLC and ESIMS (Figure 3). According to the UV-spectrophotometric assay, the overall yield of the pentamer from **8b** was 54%, corresponding to 86% average coupling yield. The remaining support bound oligonucleotide was then released and deprotected in the same manner, and the deprotected oligonucleotide was precipitated as a sodium salt by adding EtOH (20 mL) to the solution of the oligonucleotide in 3 mol $L^{-1}$ aq NaOAc (1.0 mL) and keeping the mixture at −20 °C for 1.5 h. The precipitate was carefully dried, and the concentration of a weight sample dissolved in water was determined on the basis of UV absorbance at 260 nm. According to this analysis, the oligonucleotide content of the precipitate was 95%.

## Supporting Information

### Supporting Information File 1

Further experimental data.
[http://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-10-237-S1.pdf]

## References


1. He, L.; Hannon, G. J. *Nat. Rev. Genet.* **2004,** *5,* 522–531. doi:10.1038/nrg1379
2. Bumcrot, D.; Manoharan, M.; Koteliansky, V.; Sah, D. W. Y. *Nat. Chem. Biol.* **2006,** *2,* 711–719. doi:10.1038/nchembio839
3. Davidson, B. L.; McCray, P. B. *Nat. Rev. Genet.* **2011,** *12,* 329–340. doi:10.1038/nrg2968
4. Berezikov, E. *Nat. Rev. Genet.* **2011,** *12,* 846–860. doi:10.1038/nrg3079
5. Dong, H.; Lei, J.; Ding, L.; Wen, Y.; Ju, H.; Zhang, X. *Chem. Rev.* **2013,** *113,* 6207–6233. doi:10.1021/cr300362f
6. Graber, D.; Moroder, H.; Micura, R. *J. Am. Chem. Soc.* **2008,** *130,* 17230–17231. doi:10.1021/ja806716s
7. Fauster, K.; Kreutz, C.; Micura, R. *Angew. Chem., Int. Ed.* **2012,** *51,* 13080–13084. doi:10.1002/anie.201207128
8. Granqvist, L.; Virta, P. *J. Org. Chem.* **2014,** *79,* 3529–3536. doi:10.1021/jo500326j
9. Thomas, J. R.; Hergenrother, P. J. *Chem. Rev.* **2008,** *108,* 1171–1224. doi:10.1021/cr0681546
10. Kiviniemi, A.; Virta, P. *Bioconjugate Chem.* **2011,** *22,* 1559–1566. doi:10.1021/bc200101r
11. Bonora, G. M.; Biancotto, G.; Maffini, M.; Scremin, C. L. *Nucleic Acids Res.* **1993,** *21,* 1213–1217. doi:10.1093/nar/21.5.1213
12. De Koning, M. C.; Ghisaidoobe, A. B. T.; Duynstee, H. I.; Ten Kortenaar, P. B. W.; Filippov, D. V.; van der Marel, G. A. *Org. Process Res. Dev.* **2006,** *10,* 1238–1245. doi:10.1021/op060133q
13. Donga, R. A.; Khaliq-Uz-Zaman, S. M.; Chan, T.-H.; Damha, M. J. *J. Org. Chem.* **2006,** *71,* 7907–7910. doi:10.1021/jo061279q
14. Dueymes, C.; Schönberger, A.; Adamo, I.; Navarro, A.-E.; Meyer, A.; Lange, M.; Imbach, J.-L.; Link, F.; Morvan, F.; Vasseur, J.-J. *Org. Lett.* **2005,** *7,* 3485–3488. doi:10.1021/ol0511777
15. Gimenez Molina, A.; Kungurtsev, V.; Virta, P.; Lönnberg, H. *Molecules* **2012,** *17,* 12102–12120. doi:10.3390/molecules171012102

16. Kungurtsev, V.; Laakkonen, J.; Gimenez Molina, A.; Virta, P. *Eur. J. Org. Chem.* **2013,** 6687–6693. doi:10.1002/ejoc.201300864
17. Padiya, K. J.; Salunkhe, M. M. *Bioorg. Med. Chem.* **2000,** *8,* 337–342. doi:10.1016/S0968-0896(99)00287-4
18. Reese, C. B.; Yan, H. *J. Chem. Soc., Perkin Trans. 1* **2002,** 2619–2633. doi:10.1039/b208802a
19. Adamo, I.; Dueymes, C.; Schönberger, A.; Navarro, A.-E.; Meyer, A.; Lange, M.; Imbach, J.-L.; Link, F.; Morvan, F.; Vasseur, J.-J. *Eur. J. Org. Chem.* **2006,** 436–448. doi:10.1002/ejoc.200500547
20. Bonora, G. M.; Scremin, C. L.; Colonna, F. P.; Garbesi, A. *Nucleic Acids Res.* **1990,** *18,* 3155–3159. doi:10.1093/nar/18.11.3155
21. Colonna, F. P.; Scremin, C. L.; Bonora, G. M. *Tetrahedron Lett.* **1991,** *32,* 3251–3254. doi:10.1016/S0040-4039(00)79736-X
22. Kungurtsev, V.; Virta, P.; Lönnberg, H. *Eur. J. Org. Chem.* **2013,** 7886–7890. doi:10.1002/ejoc.201301352
23. Saneyoshi, H.; Seio, K.; Sekine, M. *J. Org. Chem.* **2005,** *70,* 10453–10460. doi:10.1021/jo051741r
24. Gaffney, B. L.; Veliath, E.; Zhao, J.; Jones, R. A. *Org. Lett.* **2010,** *12,* 3269–3271. doi:10.1021/ol101236b
25. Rostovtsev, V. V.; Green, L. G.; Fokin, V. V.; Sharoless, K. B. *Angew. Chem., Int. Ed.* **2002,** *41,* 2596–2599. doi:10.1002/1521-3773(20020715)41:14<2596::AID-ANIE2596>3.0.CO;2-4
26. Tornøe, C. W.; Christensen, C.; Meldal, M. *J. Org. Chem.* **2002,** *67,* 3057–3064. doi:10.1021/jo011148j
27. Nagai, A.; Guo, Z.; Feng, X.; Jin, S.; Chen, X.; Ding, X.; Jiang, D. *Nat. Commun.* **2011,** *2,* No. 536. doi:10.1038/ncomms1542
28. Keller, A. C.; Serva, S.; Knapp, D. C.; Kwiatkowski, M.; Engels, J. W. *Collect. Czech. Chem. Commun.* **2009,** *74,* 515–534. doi:10.1135/cccc2008183

## License and Terms

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

# Reversibly locked thionucleobase pairs in DNA to study base flipping enzymes

Christine Beuck[1] and Elmar Weinhold[*2]

Full Research Paper

Open Access

Address:
[1]Department of Structural & Medicinal Biochemistry, University of Duisburg-Essen, Universitätsstr. 2-5, D-45141 Essen, Germany and [2]Institute of Organic Chemistry, RWTH Aachen University, Landoltweg 1, D-52056 Aachen, Germany

Email:
Elmar Weinhold[*] - elmar.weinhold@oc.rwth-aachen.de

* Corresponding author

Keywords:
DNA base flipping; DNA cross-link; DNA methyltransferase; nucleic acids; oligodeoxynucleotide; thionucleoside

*Beilstein J. Org. Chem.* **2014,** *10,* 2293–2306.
doi:10.3762/bjoc.10.239

Received: 09 June 2014
Accepted: 28 August 2014
Published: 01 October 2014

This article is part of the Thematic Series "Nucleic acid chemistry".

Guest Editor: H.-A. Wagenknecht



## Abstract

Covalently interstrand cross-linked DNA is an interesting tool to study DNA binding proteins that locally open up the DNA duplex by flipping single bases out of the DNA helix or melting whole stretches of base pairs to perform their function. The ideal DNA cross-link to study protein–DNA interactions should be specific and easy to synthesize, be stable during protein binding experiments, have a short covalent linker to avoid steric hindrance of protein binding, and should be available as a mimic for both A/T and G/C base pairs to cover all possible binding specificities. Several covalent interstrand cross-links have been described in the literature, but most of them fall short of at least one of the above criteria. We developed an efficient method to site-specifically and reversibly cross-link thionucleoside base pairs in synthetic duplex oligodeoxynucleotides by bisalkylation with 1,2-diiodoethane resulting in an ethylene-bridged base pair. Both linked A/T and G/C base pair analogs can conveniently be prepared which allows studying any base pair-opening enzyme regardless of its sequence specificity. The cross-link is stable in the absence of reducing agents but the linker can be quickly and tracelessly removed by the addition of thiol reagents like dithiothreitol. This property makes the cross-linking reaction fully reversible and allows for a switching of the linked base pair from locked to unlocked during biochemical experiments. Using the DNA methyltransferase from *Thermus aquaticus* (M.TaqI) as example, we demonstrate that the presented cross-linked DNA with an ethylene-linked A/T base pair analog at the target position is a useful tool to determine the base-flipping equilibrium constant of a base-flipping enzyme which lies mostly on the extrahelical side for M.TaqI.

## Introduction

Covalent interstrand DNA cross-links have long sparked clinical and biochemical interest. The cytotoxicity of bisfunctional alkylating agents like nitrogen mustards or chloroethyl nitrosourea (CENU) derivatives has been attributed to their ability to form interstrand DNA cross-links. The resulting covalently linked bases block any machinery that relies on separating the strands of the DNA duplex, e.g., DNA damage repair, replication and transcription [1-5], which is exploited in using

CENU derivatives and other interstrand cross-linking reagents as antitumor agents. Chemically synthesized DNA cross-links are of great interest to mimic the reaction products of cross-linking agents acting *in vivo* [6-8] and to mechanistically study enzymes that rely on opening Watson–Crick base pairs by base flipping, as observed with DNA methyltransferases (DNA MTases), DNA glycosylases and some restriction endonucleases, or unwinding the DNA helix with DNA helicases [9-17].

In order to study the binding of these proteins to the cross-linked DNA, the linkage has to be specific and stable under the conditions of the binding experiment, the linker should be short to avoid steric interference with protein binding, and, preferentially, the oligodeoxynucleotides (ODN) should be easily obtained on a DNA synthesizer without the need for synthesizing special building blocks. In fact, a large number of modified building blocks, also suitable for further post-synthetic derivatization, is commercially available and allows a convenient access to modified ODN. Several protocols to generate selective DNA interstrand cross-links in short duplex ODN have been described, however, most of them do not meet all these criteria. Most bisalkylating agents that react with native DNA show selectivity for one type of base pair, e.g., CENU reacts with G/C base pairs, however, regioselective introduction of only one cross-linked base pair at a specified position is not possible. Several linked nucleotides mimicking these alkylation products have been synthesized and then converted into the phosphoamidite building block for chemical DNA synthesis [6,7,18-28]. Not only do these protocols require extensive chemical synthesis, but the sequence of the synthesized duplex ODN cannot be freely chosen since after incorporation of the linked dinucleotide both DNA strands are elongated in parallel and therefore are identical in sequence. If the linked base pair is not incorporated at the 5’ end but in the middle of a duplex, the nucleotides 3' to the linkage in the 2$^{nd}$ strand will need to be filled in with orthogonal 5' to 3' chemistry.

A dT nucleoside analog with an exocyclic aziridine moiety has been chemically incorporated into poly-T ODN. Upon hybridization with a complementary poly-A strand, the aziridine group is attacked by the exocyclic amino group of the opposing adenine, resulting in ring opening and formation of an ethylene cross-linked base pair [29,30]. Other aziridine-substituted nucleobases have been incorporated enzymatically by a DNA polymerase, but elongation past the modified nucleoside has not been reported [31].

Another approach uses nucleobases in which an exocyclic functionality has been substituted by a vinyl group. Upon hybridization with a complimentary, unmodified strand, the vinyl group is attacked by the nucleophilic exo- or endocyclic nitrogen of the partnering base, forming a cross-linked base pair [32-34]. While the cross-linking reaction itself is straight-forward, the vinyl-substituted nucleosides need to be chemically synthesized and subsequently converted into their phosphoamidites for chemical DNA synthesis. The same holds true for cross-links based on aldehyde [8,35-37] or click chemistry [38]. In addition, long linkers that might interfere with protein binding are typically employed.

The Verdine group has developed a cross-linking approach based on the easy postsynthetic introduction of cystamine or longer amino-alkylthiol linkers into ODN using a convertible nucleoside approach [39-48]. The amino-alkylthiol linker is introduced as its disulfide dimer to protect the sulfur and ensure that the reaction with the convertible nucleobase occurs via the amino group. Two of these residues are placed in adjacent base pairs and reduction of the linker disulfides frees the thiol groups, which, upon removal of reducing agent, react under oxidative conditions to form a disulfide cross-link. Adjacent A/A, C/C, G/G and T/T cross-links have been synthesized with this method. An interstrand G/G cross-link of adjacent G/C base pairs was used to study binding of the DNA cytosine-C5 MTase M.HaeIII [39]. The two guanine bases were linked in the minor groove via their exocyclic N2 atoms and the linker does not interfere with M.HaeIII binding. However, for other enzymes with a different specificity or binding mode, especially when attempting to cross-link the target base, the length and steric demand of the linker can be critical and hinder protein binding.

A thionucleobase opposing a guanine base has been cross-linked using a bulky bis-bromoacetamide linker [40,41]. Selective alkylation occurs at the sulfur atom of 4-thiouracil and at the N7 ring nitrogen of the opposite guanine base, introducing a positive charge. This cross-link is not ideally suited to study protein–DNA interactions because the linker is sterically very demanding and the positive charge introduced in the guanine makes it susceptible to depurination.

Direct zero-length cross-linking of two opposing thionucleosides like 2’-deoxy-6-thioinosine (dI$^{6S}$) with 2’-deoxy-4-thiothymidine (dT$^{4S}$) or 2’-deoxy-4-thiouridine (dU$^{4S}$) via a disulfide linkage, appears to be very attractive for studying protein–DNA interactions. These thionucleosides can easily be incorporated in synthetic ODN and the linkage is formed without additional linker atoms avoiding the risk of steric interference with protein binding [42-46]. A disulfide cross-link between 2'-deoxy-6-thioguanosine (dG$^{6S}$) and dU$^{4S}$ was used to study the specificity of the human flap endonuclease FEN1 and confirmed that unpairing the two terminal nucleotides of the DNA duplex is crucial for the selection of the target phosphodiester bond [47]. In our hands though, although direct cross-

linking between internal $dI^{6S}$ or $dG^{6S}$ with $dU^{4S}$ residues in duplex ODN could be achieved in yields of up to 80%, the cross-link was not stable during purification. The successful formation of a stable disulfide cross-link might depend on the DNA sequence and position of the thionucleobase pair within the duplex.

We developed a novel method to efficiently cross-link opposite thionucleobases within duplex ODN by exploiting the selective nucleophilic properties of the thiobases. Bisalkylation of a thiobase pair with 1,2-diiodoethane in aqueous solution yields a short linkage with low steric demand. Both A/T and G/C base pair analogs can be site-specifically cross-linked to study enzymes with various sequence specificities. We applied this method to probe the contribution of base flipping to the overall binding affinity of the DNA adenine-N6 MTase from *Thermus aquaticus* (M.TaqI) as an example.

# Results

## Selective cross-linking of thionucleobase pairs in DNA by bis-alkylation with 1,2-diiodoethane

Thionucleobases are excellent soft nucleophiles, and chemoselective alkylation of the sulfur atoms readily occur within the context of duplex DNA without modifying other functionalities [41,48-50]. Reacting a thiobase pair in DNA with a bis-electrophilic linker should thus result in selective cross-link formation. Thionucleosides can easily be incorporated into ODN by using a commercially available cyanoethyl protected phosphoamidite for $dG^{6S}$ and $dU^{4S}$ or by postsynthetic modification of convertible nucleoside precursors for $dI^{6S}$ and $dU^{4S}$ [51].

The ideal geometry of a Watson–Crick base pair within B-DNA is a coplanar orientation of the two nucleobases, stacked between its neighboring base pairs. In contrast, the ideal geometry of saturated linkers will demand non-planar torsion angles. Coleman et al. [44] investigated this topic by performing molecular dynamic simulations for a disulfide cross-linked $I^{6S}/U^{4S}$ base pair in B-DNA and showed that the compromise between the dihedrals preferred by the linker and the planar base pair results in local propeller and buckle motions. Since the ethylene linker in our cross-linked base pairs will likely also deviate slightly from a planar arrangement to evade eclipsed conformations, we chose $U^{4S}$ over $T^{4S}$ to avoid a potential steric conflict of the C5 methyl group with the neighboring base pair and keep the local distortion in the DNA helix to a minimum.

The 14mer duplex ODN $\mathbf{1}^{I6S}\cdot\mathbf{2}^{U4S}$ containing a thio-analog of the A/T base pair and $\mathbf{1}^{G6S}\cdot\mathbf{2}^{U4S}$ containing a thio-analog of the G/C base pair were reacted with 1,2-diiodoethane at room temperature and pH 9.0 (Figure 1a and 1b). The reaction was

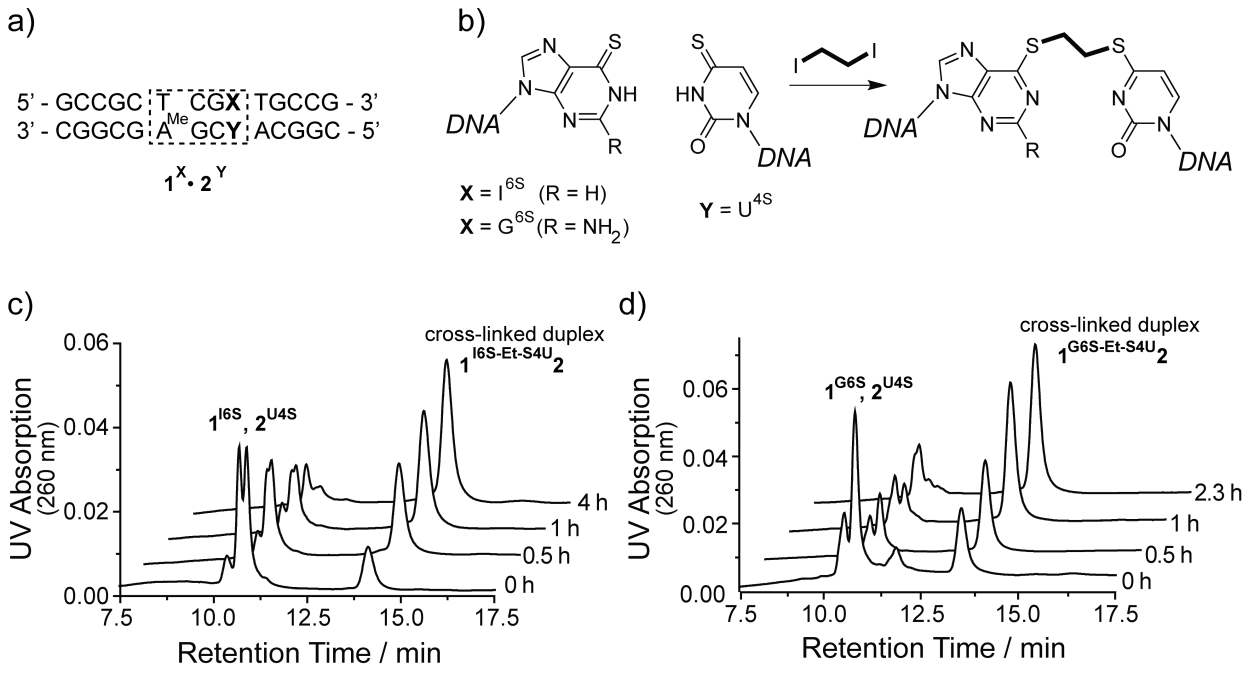



**Figure 1:** Covalent cross-linking of a $I^{6S}/U^{4S}$ or $G^{6S}/U^{4S}$ base pair within duplex ODN **1·2** (a) by bis-alkylation with 1,2-diiodoethane at pH 9.0 (b). The time course for the reaction of duplexes $\mathbf{1}^{I6S}\cdot\mathbf{2}^{U4S}$ (c) and $\mathbf{1}^{G6S}\cdot\mathbf{2}^{U4S}$ (d) is monitored by denaturing anion exchange HPLC. The non-cross-linked parent duplexes dissociate and elute as their respective single strands, while the covalently linked duplexes elute significantly later.

monitored by denaturing anion exchange chromatography. Addition of urea to the elution buffers and heating of the column to 70 °C causes non-linked duplex ODN to dissociate and elute as their respective single strands while the covalently linked duplexes elute significantly later, comparable to a single-stranded ODN of twice the length (Figure 1c and 1d). In contrast, no species eluting at higher retention times are observed when each of the single strands alone is reacted with 1,2-diiodoethane. These observations are consistent with the formation of the interstrand cross-linked duplexes $\mathbf{1}^{\mathbf{I6S\text{-}Et\text{-}S4U}}\mathbf{2}$ (80% yield after 4 h) and $\mathbf{1}^{\mathbf{G6S\text{-}Et\text{-}S4U}}\mathbf{2}$ (72% after 2.3 h). The cross-linked duplexes exhibit a significantly diminished UV absorption at 332 nm compared to the duplex ODN with unmodified thionucleobases. The long wavelength absorption band of thionucleobases is characteristic for their thio–keto tautomer [52]. This band is shifted towards shorter wavelengths upon alkylation at the sulfur atom and loss of the carbon–sulfur double bond. A small fraction of single-stranded ODN with low 332 nm absorption remains and can be assigned to alkylation products where two 1,2-diiodoethane molecules reacted with both thionucleobases in the duplex or where the reaction with 1,2-diiodoethane took place on only one of the thionucleobases but failed to react with the opposing base.

Both cross-linked duplexes $\mathbf{1}^{\mathbf{I6S\text{-}Et\text{-}S4U}}\mathbf{2}$ and $\mathbf{1}^{\mathbf{G6S\text{-}Et\text{-}S4U}}\mathbf{2}$ were isolated by preparative denaturing anion exchange HPLC. An aliquot of each cross-linked duplex was resuspended in buffer without reducing agent. The integrity of the purified cross-linked duplexes was assessed by denaturing HPLC (Figure 2b and Figure 3b). Both cross-linked duplexes elute at 13.4 min and show the characteristic low 332 nm absorption, with only trace amounts of dissociated single strands in the 10.5–10.7 min range. The elution profile did not change over the course of 3 days showing that the cross-link is stable at room temperature in the absence of reducing agents. The UV spectrum of both cross-linked duplexes $\mathbf{1}^{\mathbf{I6S\text{-}Et\text{-}S4U}}\mathbf{2}$ and $\mathbf{1}^{\mathbf{G6S\text{-}Et\text{-}S4U}}\mathbf{2}$ clearly

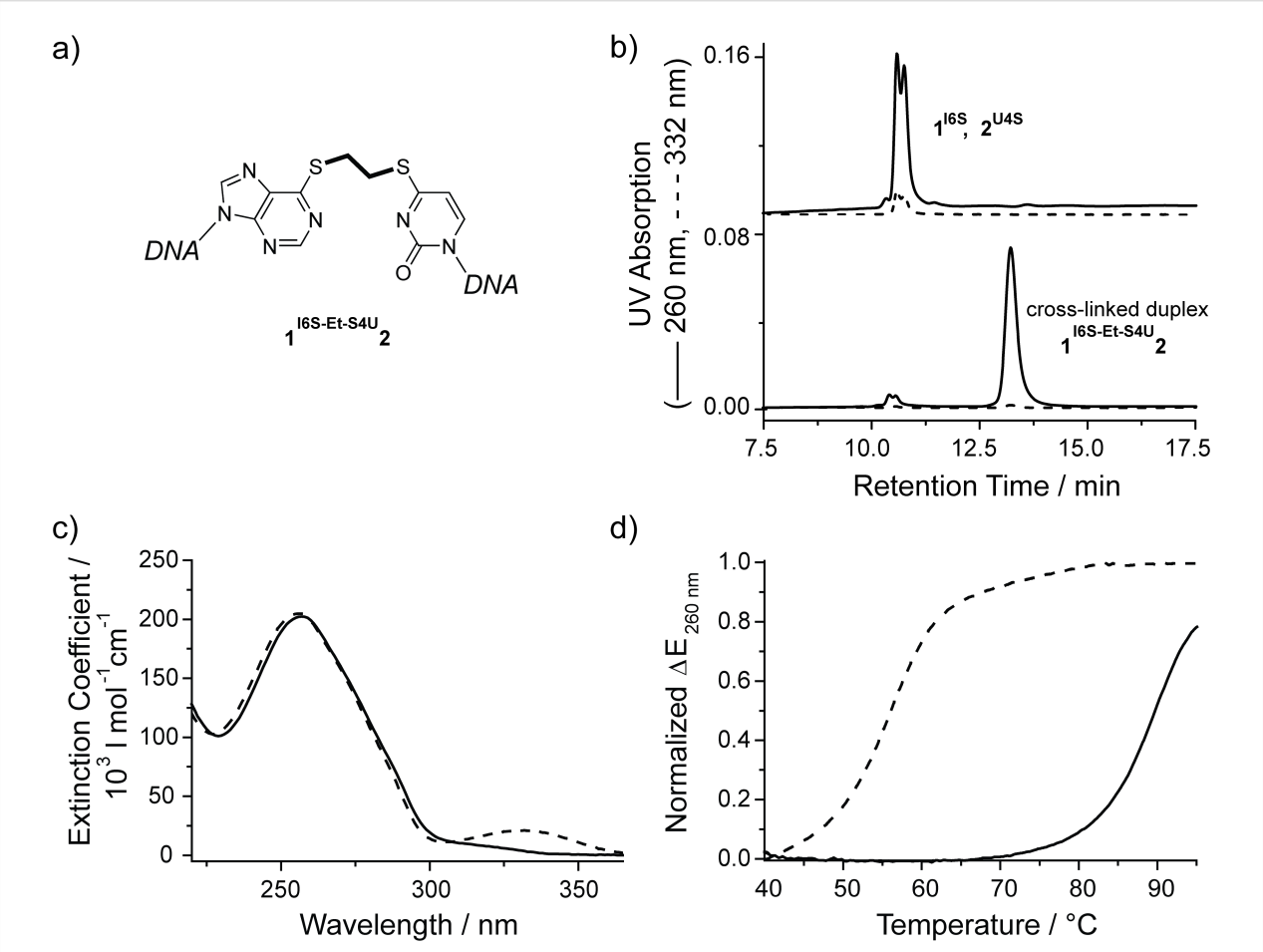



**Figure 2:** Characterization of HPLC-purified cross-linked duplex $\mathbf{1}^{\mathbf{I6S\text{-}Et\text{-}S4U}}\mathbf{2}$ (a) by denaturing anion exchange HPLC (b, bottom) in comparison to the non-linked parent duplex (b, top). Alkylation at the sulfur atoms within the thio-nucleobases is confirmed by the wavelength shift and strong intensity reduction of the absorption band above 300 nm in the UV spectrum (c, solid) compared to the spectrum of $\mathbf{1}^{\mathbf{I6S}}\cdot\mathbf{2}^{\mathbf{U4S}}$ (dashed). The thermal stability of the cross-linked duplex $\mathbf{1}^{\mathbf{I6S\text{-}Et\text{-}S4U}}\mathbf{2}$ (d, solid) is drastically increased compared to $\mathbf{1}^{\mathbf{I6S}}\cdot\mathbf{2}^{\mathbf{U4S}}$ (dashed).

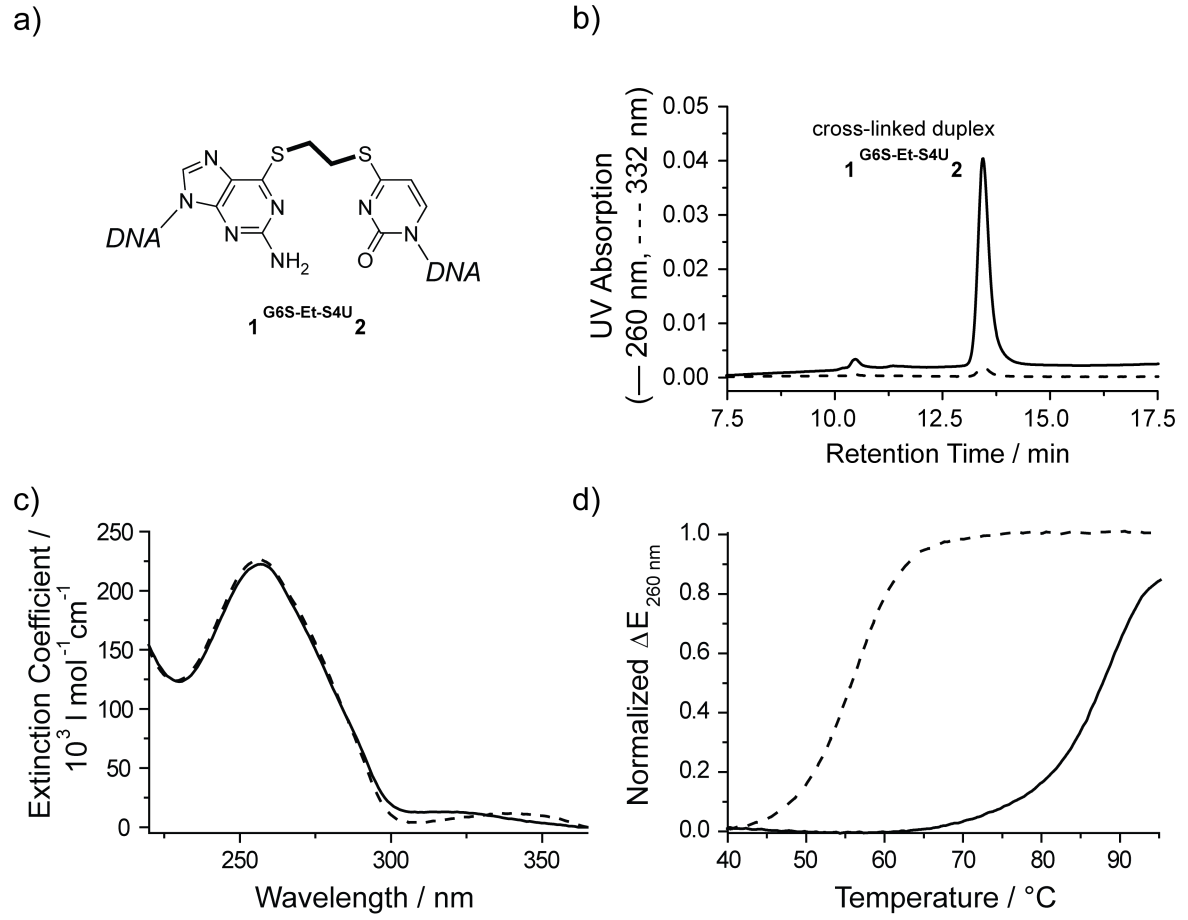



**Figure 3:** Characterization of HPLC-purified cross-linked duplex $\mathbf{1}^{\mathbf{G6S\text{-}Et\text{-}S4U}}\mathbf{2}$ (a) by denaturing anion exchange HPLC (b). Alkylation at the sulfur atoms within the thionucleobases is confirmed by the strong reduction of the absorption band above 300 nm in the UV spectrum of $\mathbf{1}^{\mathbf{G6S\text{-}Et\text{-}S4U}}\mathbf{2}$ (c, solid) compared to the spectrum of $\mathbf{1}^{\mathbf{G6S}}\cdot\mathbf{2}^{\mathbf{U4S}}$ (dashed). The thermal stability of the cross-linked duplex $\mathbf{1}^{\mathbf{G6S\text{-}Et\text{-}S4U}}\mathbf{2}$ (d, solid) is drastically increased compared to $\mathbf{1}^{\mathbf{G6S}}\cdot\mathbf{2}^{\mathbf{U4S}}$ (dashed).

confirms the shift of the thio-nucleobase absorption band over 300 nm towards lower wavelengths as it is typical for S-alkylated thio-nucleobases (Figure 2c and Figure 3c). Circular dichroism [53] spectroscopy is sensitive to the overall global conformation of nucleic acids. The CD spectrum of duplex ODN $\mathbf{1}^{\mathbf{I6S\text{-}Et\text{-}S4U}}\mathbf{2}$ retains the characteristic shape of B-DNA, showing that the ethyl cross-link does not significantly distort the DNA helix (data not shown). The cross-linked duplexes were further characterized by thermal denaturation monitoring the UV absorption at 260 nm (Figure 2d and Figure 3d). A covalent link between the two strands within a DNA duplex is expected to locally stabilize the DNA duplex by preventing the complete dissociation of its strands. Ultimately, at high temperatures all hydrogen-bonded base pairs will be disrupted even in the presence of a covalent link. Therefore, a shift of the melting transition towards higher temperatures is expected for the cross-linked duplexes. The non-cross-linked duplex ODN $\mathbf{1}^{\mathbf{I6S}}\cdot\mathbf{2}^{\mathbf{U4S}}$ and $\mathbf{1}^{\mathbf{G6S}}\cdot\mathbf{2}^{\mathbf{U4S}}$ cooperatively melt at 56 °C, which is comparable to the melting temperature of the same duplex with a natural mismatched base pair, while the 14mer duplex ODN with a native hydrogen-bonded A/T base pair melts at 66 °C. The thionucleobases behave like a mismatch because both thio-nucleobases preferably exist in their thio-keto form and thus no Watson–Crick hydrogen bonds are formed in the $I^{6S}/U^{4S}$ pair and only one is possible in the $G^{6S}/U^{4S}$ pair. Both cross-linked duplexes have melting temperatures of 88–89 °C demonstrating a drastic stabilization compared to the parent duplex ODN with the unmodified thionucleobases and the matched duplex ODN. This 33 °C increase in the melting temperature upon cross-linking is comparable to the stabilization previously observed for other interstrand cross-links [44,54].

We have shown that duplex ODN containing a thionucleobase pair can easily be covalently cross-linked by bis-alkylation with 1,2-diiodoethane. Both duplexes with a linked A/T and G/C base pair analog can be obtained by choosing the corres-

ponding $I^{6S}/U^{4S}$ or $G^{6S}/U^{4S}$ thionucleobase pair. The cross-linked duplexes can be isolated and were stable in the absence of reducing agent.

## Traceless linker removal with thiol nucleophiles

The covalent ethylene cross-links in $\mathbf{1}^{\mathbf{I6S\text{-}Et\text{-}S4U}}\mathbf{2}$ and $\mathbf{1}^{\mathbf{G6S\text{-}Et\text{-}S4U}}\mathbf{2}$ are stable at room temperature in the absence of thiol reagents. Interestingly, when an aliquot of each cross-linked duplex is resuspended in buffer with 1 mM dithiothreitol (DTT) and analyzed by denaturing anion exchange HPLC, no more cross-linked duplex (retention time 13.4 min) is observed. Instead, the duplex dissociates and the single strands elute at retention times between 10.5–10.7 min. These single strands exhibit a fully restored absorption at 332 nm, suggesting that the cross-linking reaction was reversed and the duplex ODN $\mathbf{1}^{\mathbf{I6S}}\mathbf{.2}^{\mathbf{U4S}}$ and $\mathbf{1}^{\mathbf{G6S}}\mathbf{.2}^{\mathbf{U4S}}$ with the thionucleosides in their thio-keto form have been restored (Figure 4). The UV spectra of the cross-linked duplex $\mathbf{1}^{\mathbf{I6S\text{-}Et\text{-}S4U}}\mathbf{2}$ before and after addition of 1 mM DTT show the full reappearance of the band at long wavelength around 330 nm identical to the unlinked duplex $\mathbf{1}^{\mathbf{I6S}}\mathbf{.2}^{\mathbf{U4S}}$, corroborating that the linker was fully removed from both thionucleobases.

In order to demonstrate that the opened duplex in fact contains the unmodified thionucleosides, the duplex ODN resulting from the reaction of $\mathbf{1}^{\mathbf{I6S\text{-}Et\text{-}S4U}}\mathbf{2}$ with DTT was transferred into DTT-free buffer using a desalting column and the cross-linking reagent 1,2-diiodoethane was added again. After 2.5 h the reaction was analyzed by denaturing anion exchange chromatography, showing the reappearance of the cross-linked duplex (data not shown).

Opening of the cross-linked base pair was further examined with other thiol nucleophiles like β-mercaptoethanol (BME) and ethanethiol (EtSH) in addition to DTT. These thiols were added to a solution of the cross-linked duplex $\mathbf{1}^{\mathbf{I6S\text{-}Et\text{-}S4U}}\mathbf{2}$ and the reaction was monitored by recording the UV spectrum as a function of time (Figure 5). DTT and BME are able to remove the alkyl linker and restore the thionucleobase specific UV absorption above 300 nm almost instantaneously (Figure 5a and 5b) while the reaction with EtSH proceeds slower with a half life of less than 5 min (Figure 5c).

We conclude that reversion of the cross-linking reaction occurs by attack of a thiol nucleophile onto the carbon atoms of the linker with the thionucleobases acting as the leaving groups.

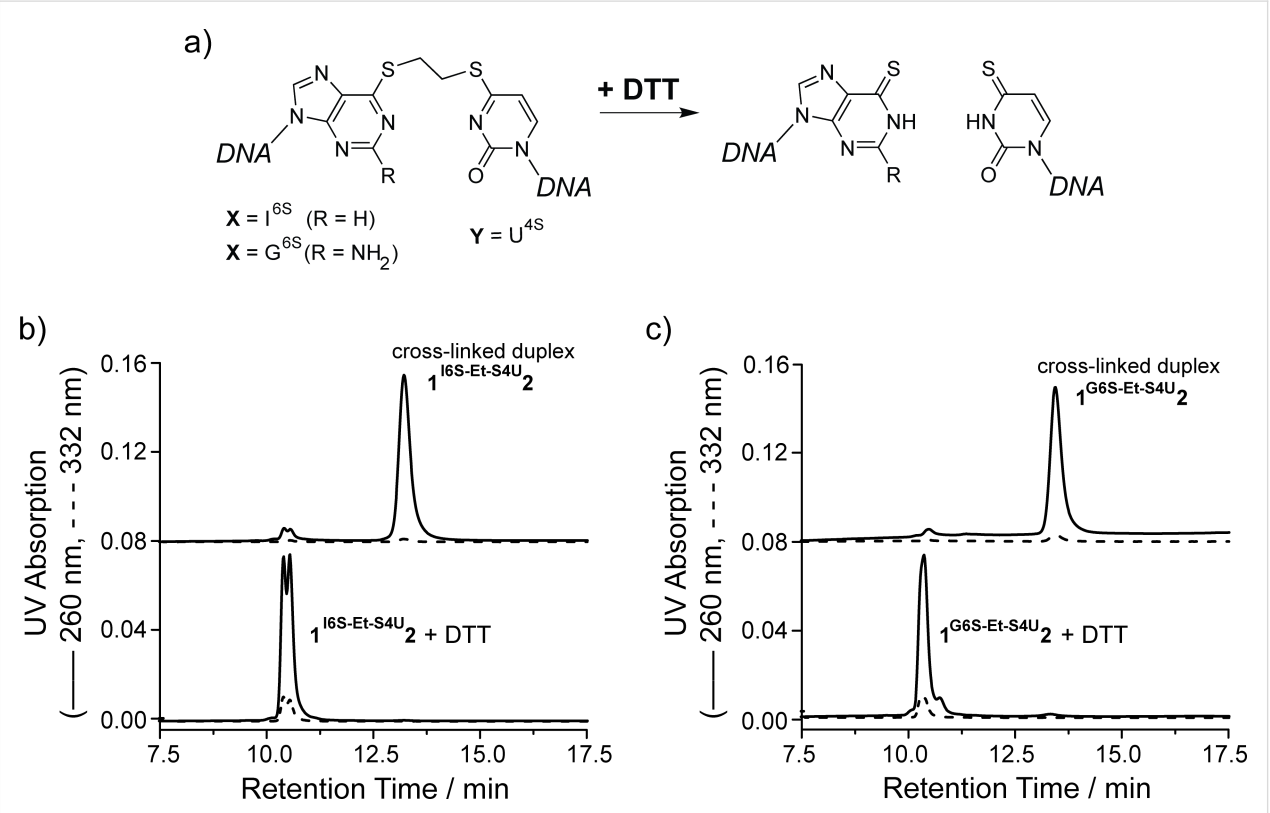



**Figure 4:** Opening and traceless linker removal of the cross-linked duplexes $\mathbf{1}^{\mathbf{I6S\text{-}Et\text{-}S4U}}\mathbf{2}$ and $\mathbf{1}^{\mathbf{G6S\text{-}Et\text{-}S4U}}\mathbf{2}$ by adding DTT (a). Denaturing anion exchange HPLC chromatograms (b, c) demonstrate that the reaction products dissociate and elute at the same retention times as their parent single strands. Furthermore, the UV absorption at 332 nm is fully recovered, indicating that the thio-keto function in the nucleobases has been restored.

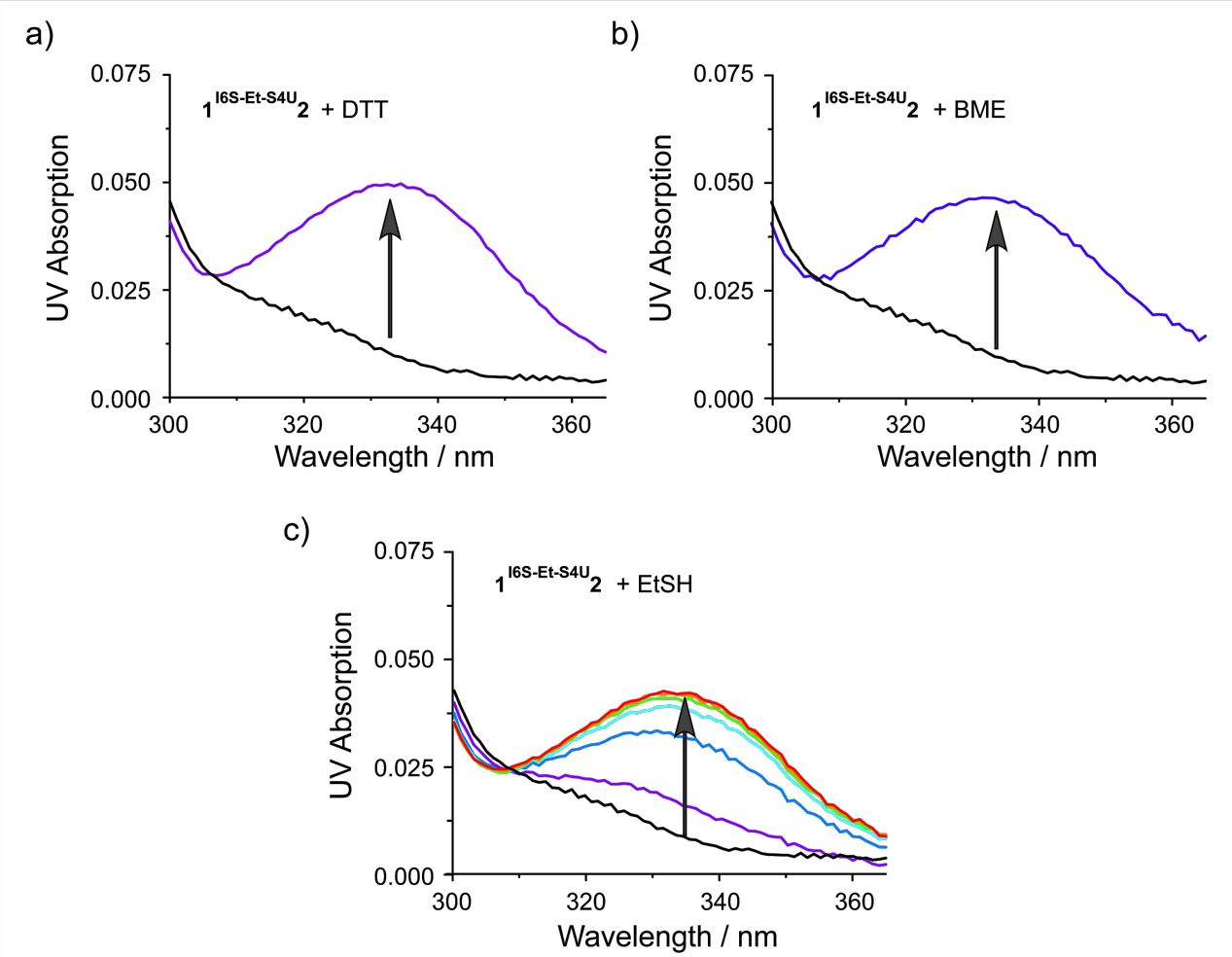



**Figure 5:** UV time course of opening of the cross-link in duplex $\mathbf{1}^{\mathbf{I6S\text{-}Et\text{-}S4U}}\mathbf{2}$ with the thiol nucleophiles DTT (a), β-mercaptoethanol (BME, b) and ethanethiol (EtSH) (c). Upon removal of the linker, the thionucleobase absorption bands above 300 nm are restored. Black, no nucleophile; purple, 0 min; blue, 5 min; cyan, 10 min; green, 15 min; orange, 20 min; red, 25 min.

The re-opened reaction product can be re-cross-linked by renewed addition of 1,2-diiodoethane after removing any thiol reagent (not shown).

While the duplex ODN with an ethylene-linked thionucleobase pair are stable in the absence of reducing agents, the addition of thiol-containing nucleophiles such as DTT quickly and completely removes the ethylene bridge from both thionucleobase pairs. This interesting feature of quick traceless linker removal allows for a switching of the thionucleobase pair from the locked to the unlocked state.

## Cross-linked DNA as a tool to determine the base flipping equilibrium of DNA-modifying enzymes

DNA-modifying enzymes often use a base flipping mechanism to gain access to their target bases [9,11-17,55-68]. Experimental evidence supports a two-step binding model comprised of an initial association equilibrium with the dissociation constant $K_{D,init}$, followed by flipping of the target base with the equilibrium constant $K_{flip}$ (Figure 6a) [55,56,69]. The binding affinity of base-flipping enzymes is often determined using DNA with the fluorescent base analog 2-aminopurine (2AP) at the target site [70-76]. The fluorescence of 2AP is quenched within the base stack of duplex DNA and increases once the DNA is bound and the 2AP base is flipped out of the DNA helix by the base-flipping enzyme. A competition binding assay is used to determine the dissociation constant $K_D$ for a non-fluorescent DNA substrate. Titrating enzyme into a fixed ratio of the fluorescent DNA with known $K_D$ and the non-fluorescent DNA leads to a fluorescence intensity increase which in turn depends on the concentrations and ratio of binding affinities. The contribution of the base flipping step ($K_{flip}$) to the overall observed binding affinity cannot be extracted from these data because the non-flipped and flipped enzyme complexes A and B (Figure 6) cannot be distinguished in a single binding experi-

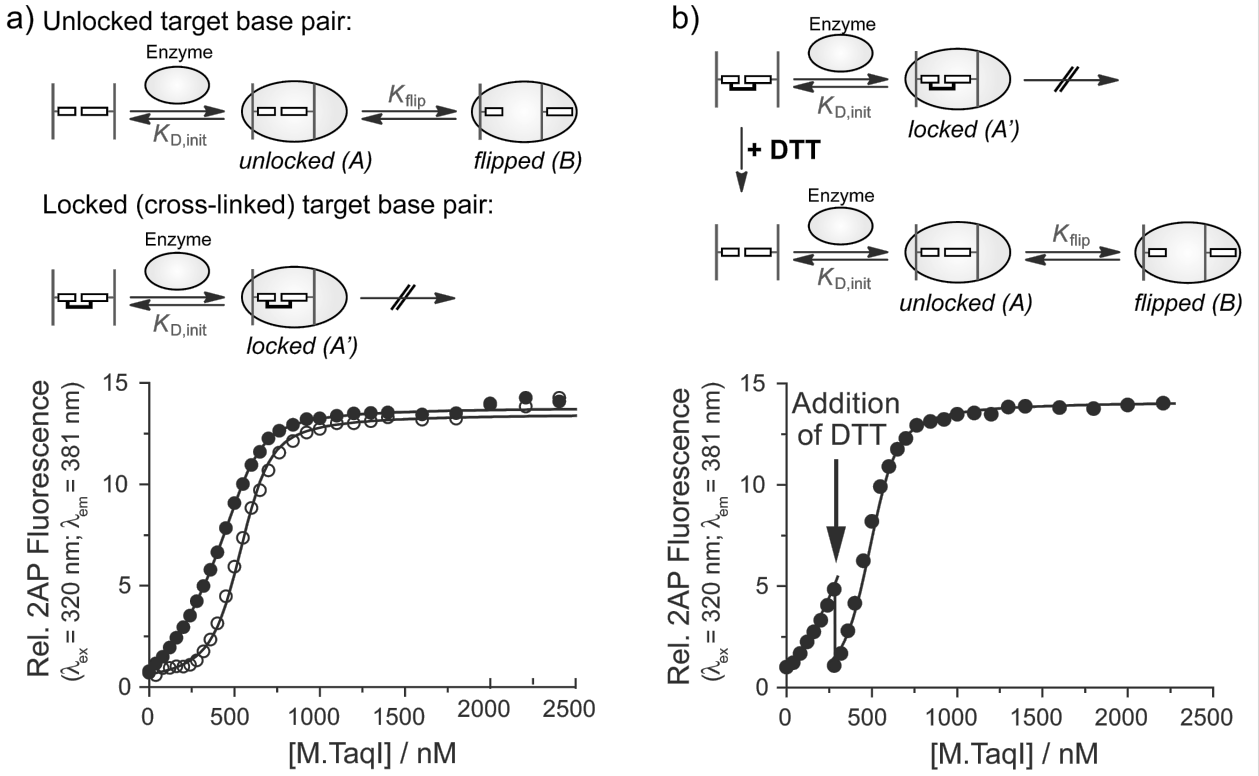



**Figure 6:** Competition binding of M.TaqI to DNA with unlocked and locked target base pairs. Increasing amounts of M.TaqI were added to a mixture of the DNA of interest and a fluorescent DNA with known $K_D$ that contains the target base analog 2-aminopurine (2AP). A more delayed increase in the 2AP fluorescence indicates tighter binding of the non-fluorescent DNA. (a) Binding of M.TaqI to duplex $\mathbf{1}^{\mathbf{I6S}}\cdot\mathbf{2}^{\mathbf{U4S}}$ with unlocked target base pair (top reaction scheme and open circles) allows base flipping of the target base into the active site. The observed dissociation constant $K_D^{unlocked}$ comprises both the initial binding constant $K_{D,init}$ and the base flipping equilibrium constant $K_{flip}$. Binding of M.TaqI to the cross-linked duplex $\mathbf{1}^{\mathbf{I6S\text{-}Et\text{-}S4U}}\mathbf{2}$ with locked target base pair (bottom reaction scheme and closed circles) prevents base flipping, and the observed dissociation constant $K_D^{locked}$ equals $K_{D,init}$. (b) Switching from locked to unlocked target base pair during the M.TaqI binding experiment. The same competition titration was started with the cross-linked duplex $\mathbf{1}^{\mathbf{I6S\text{-}Et\text{-}S4U}}\mathbf{2}$ in buffer without reducing agent preventing base flipping. Upon addition of DTT (arrow) the cross-link is unlocked enabling base flipping and thus tighter binding. The titration is continued and now follows the binding curve for the duplex with unlocked target base pair.

ment. Therefore the overall dissociation constant $K_D^{unlocked}$ is given by (see Supporting Information File 1):

$$K_D^{unlocked} = K_{D,init} / (1 + K_{flip}) \quad (1)$$

In order to calculate $K_{flip}$, the dissociation constant $K_{D,init}$ of the initial encounter complex A must be determined independently. Utilizing DNA with a covalently locked target base pair (Figure 6) the base flipping step is blocked and the dissociation constant $K_D^{locked}$ is equal to the dissociation constant of the initial complex A':

$$K_D^{locked} = K_{D,init} \quad (2)$$

It is of great importance that the presence of the linker does not impair initial binding so that $K_{D,init}$ is identical for the formation of both complexes A and A'. When both binding experiments with locked and unlocked target base pair are performed, $K_{flip}$ is obtained by combination of equations (1) and (2):

$$K_{flip} = \left( K_D^{locked} / K_D^{unlocked} \right) - 1 \quad (3)$$

We used the duplex $\mathbf{1}^{\mathbf{I6S\text{-}Et\text{-}S4U}}\mathbf{2}$ containing a locked A/T base pair analog at the target position within the double-stranded 5'-TCGA-3' recognition site of the DNA adenine-N6 MTase M.TaqI to determine the equilibrium constant $K_{flip}$ for the base flipping step. The dissociation constants for the duplexes with an unlocked ($\mathbf{1}^{\mathbf{I6S}}\cdot\mathbf{2}^{\mathbf{U4S}}$) and with a locked ($\mathbf{1}^{\mathbf{I6S\text{-}Et\text{-}S4U}}\mathbf{2}$) target base pair analog were determined by a competitive 2AP fluorescence binding assay in separate experiments (Figure 6a). With $K_D^{unlocked}$ = 1.0 ± 0.5 nM and $K_D^{locked}$ = 6.0 ± 3 nM equation (3) provides $K_{flip}$ = 5, corresponding to about 80% of the target base being flipped within the M.TaqI-DNA complex. Therefore,

the base flipping equilibrium of M.TaqI strongly lies on the extrahelical side.

To demonstrate that the locked target base pair can be re-opened during the binding experiment, the M.TaqI titration was started with cross-linked duplex ODN $\mathbf{1}^{\mathbf{I6S\text{-}Et\text{-}S4U}}\mathbf{2}$ in buffer without reducing agent and M.TaqI was added up to the point where the binding curves for the duplexes with unlocked and locked target base pair (Figure 6a) differed the most. DTT was then added to unlock the target base pair, allowing base flipping to take place (Figure 6b). The binding affinity is thus improved resulting in more potent competition and a significant drop in 2-AP fluorescence intensity. Continuing the titration with increasing M.TaqI concentrations, the binding curve now follows the one observed with duplex $\mathbf{1}^{\mathbf{I6S}}\mathbf{\cdot 2}^{\mathbf{U4S}}$ with an unlocked target base pair.

Cross-linked duplex DNA with an ethylene-linked thio-base pair has proven to be a useful tool for analyzing the base flipping equilibrium of the DNA MTase M.TaqI. The target base within the M.TaqI-DNA complex is mostly in the extrahelical conformation which places it into the enzyme's active site. The quick and traceless removal of the linker by addition of DTT enables switching of the target base pair from a locked to an unlocked state during the course of the binding experiment.

## Discussion

Here we presented a new and convenient method to engineer a reversibly locked thionucleobase pair in duplex DNA by bisalkylation of either an $I^{6S}/U^{4S}$ or $G^{6S}/U^{4S}$ pair with 1,2-diiodoethane. There are several advantages of the presented cross-linking protocol over most others known in the literature: The cross-link is introduced postsynthetically on the level of duplex DNA in aqueous solution, without the need of extensive chemical synthesis and with no restrictions with respect to the DNA sequence. The cross-linked analogs of both an A/T and a G/C base pair can be obtained by starting with either an $I^{6S}/U^{4S}$ or $G^{6S}/U^{4S}$ base pair, and therefore, this chemistry can be employed to study any DNA-modifying enzyme no matter what target base specificity it possesses. While different linker geometries might work for different DNA-binding proteins, the ethyene-linked thiobase pairs offer the advantage of a very short linker with little steric demand for studying DNA-opening enzymes. The cross-linked duplex ODN are stable at room temperature in buffers without thiol reagents. Addition of thiol nucleophiles results in a quick and traceless removal of the ethylene linker which makes them interesting tools for chemical switching DNA from locked to unlocked base pairs during biophysical experiments. However, this interesting property requires that all proteins studied are stable in buffers without reducing thiol reagent. Even small amounts of these nucleophiles from the protein stock solutions are sufficient to open the linkage.

As a proof of principle, we applied a duplex ODN with an ethyl cross-linked $I^{6S}/U^{4S}$ base pair as A/T pair analog within the 5'-TCGA-3' recognition sequence of the DNA MTase M.TaqI and determined the base flipping equilibrium constant $K_{flip}$ within the M.TaqI-DNA complex. DNA with a locked target base pair allows initial binding of the enzyme to the DNA but prevents subsequent flipping of the target base. $K_{flip}$ can therefore be calculated from two binding experiments, one with an unlocked $I^{6S}/U^{4S}$ pair, giving the overall binding constant that comprises both binding and base flipping, and one with the locked cross-linked target base pair (Figure 6), which reports the initial binding only. The resulting $K_{flip} = 5$ demonstrates that about 80% of the M.TaqI target base are found in the extrahelical conformation within the M.TaqI-DNA complex.

Since there are some chemical differences between an $I^{6S}/U^{4S}$ pair and the native A/T target base pair the value obtained for $K_{flip}$ should be interpreted as an approximation for the base flipping equilibrium with the native DNA target. In the $I^{6S}/U^{4S}$ base pair, both thiobases exist in their thio-keto form and therefore cannot form hydrogen bonds like the natural A/T base pair. Without having to expend the energy to break any hydrogen bonds, flipping the target base should be easier, but at the same time, the $I^{6S}$ base is not able to form the hydrogen bonds in the active site of the enzyme that are observed in the complex with the native adenine target [77]. Comparing binding of M.TaqI to its native hemi-methylated target DNA and the fully methylated DNA (Figure S1, Supporting Information File 1), it is expected that the methylated target base, which is the product of the DNA MTase reaction, sterically interferes with residues in the active site of the enzyme, resulting in favoring the innerhelical target base conformation which helps to release the reaction product from the enzyme. Under the assumption that the base flipping equilibrium lies completely on the innerhelical side for the fully methylated reaction product and using the dissociation constants for hemi-methylated DNA ($K_D = 2.9 \pm 1.5$ nM) and fully methylated DNA ($K_D = 56 \pm 30$ nM), one obtains $K_{flip} = 18$ for the hemi-methylated DNA-M.TaqI complex. This corresponds to 95% of the target adenine being in the extrahelical conformation which is similar to the 80% extrahelical target base obtained with the A/T base pair analog. Interestingly, the fully methylated duplex binds even worse than the cross-linked duplex (Figure S1, Supporting Information File 1) which suggests that the methyl group in the natural reaction product *N*6-methyladenine ($A^{Me}$) does not only destabilize the extrahelical conformation in the flipped but also in the non-flipped (initial) complex. This may further enhance product release and reduce product inhibition.

The cross-linked target base pair is missing the methyl group (Figure S2, Supporting Information File 1) and might therefore be a better mimic for the substrate with innerhelical target base than for the methylated reaction product.

Many other proteins locally open up the DNA duplex to perform their function, ranging from flipping a single nucleobase out of the DNA helix to melting whole stretches of base pairs. A base flipping mechanism is utilized by many enzymes that either chemically modify the flipped nucleobase or repair a lesion, e.g., other DNA methyltransferases [9,10], DNA demethylases like the AlkB family [11], DNA alkyltransferases that are essential to DNA alkylation damage repair [12,13], DNA glycosylases from the base excision repair pathway [14-16], and photolyases repairing UV damage in DNA [17]. DNA helicases, which locally separate the two DNA strands, are important to enable vital cellular processes like DNA replication, DNA repair, chromatin remodeling and telomere maintenance [78-81]. Cross-linked DNA will not only provide a useful tool to study DNA binding and base flipping thermodynamics, as we demonstrated for the DNA MTase M.TaqI, but could also be used to determine their site and mechanism of action by introducing cross-linked base pairs at different positions within the DNA, or stall proteins on the DNA in a pre-flipped complex for structure determination to reveal the initial contacts between the proteins and their DNA targets that are formed before the target base is flipped or strands are separated.

## Conclusion

In conclusion, we have developed a new efficient method to site-specifically and reversibly cross-link thionucleoside base pairs in duplex DNA via an ethylene bridge. Both linked A/T and G/C base pair analogs can easily be prepared which allows studying any base flipping enzyme regardless of its sequence specificity. We demonstrated that cross-linked DNA is a useful tool to study base flipping enzymes and proved that the base flipping equilibrium lies mostly on the extrahelical side and thus has an important contribution to the overall DNA-binding energy of M.TaqI.

## Experimental

Solid-phase DNA synthesis was performed on an ABI DNA Synthesizer 392 in 1 µmol scale following the standard cycles recommended by Applied Biosystems with a coupling time of 30 s for the natural nucleotides. For unnatural nucleotides, the coupling time was extended to 90 s. Fast deprotectable *tert*-butylphenoxyacetyl (TAC) protected A, C and G phosphoamidites and coupling reagents including TAC anhydride as capping reagent were purchased from Proligo. *N*6-Methyl-2'-deoxyadenosine ($A^{Me}$), 6-thio-2'-deoxyguanosine ($G^{6S}$) phosphoamidites, as well as the convertible thionucleoside precursor phosphoamidites *O*6-phenyl-2'-deoxyinosine ($I^{OPh}$) and 4-triazolyl-2'-deoxyuridine ($U^{Tri}$) were purchased from Glen Research. Oligodeoxynucleotides (ODN) with $G^{6S}$ and $A^{Me}$ were deprotected according to the manufacturer's protocol. ODN with convertible nucleosides $I^{OPh}$ and $U^{Tri}$ were deprotected and converted to their thionucleoside containing counterparts as described in [51]. Duplexes were annealed before use at a concentration of 100 µM in the respective experiment buffer by heating to 95 °C for 2 min followed by slow cooling to room temperature.

**Table 1:** ODN sequences used in this study. The M.TaqI recognition sequence is highlighted in bold. $A^{Me}$ = *N*6-methyladenine.

| ODN | Sequence |
|---|---|
| $\mathbf{1^{I6S}}$ | 5'- GCCGC **TCGI$^{6S}$** TGCCG -3' |
| $\mathbf{1^{G6S}}$ | 5'- GCCGC **TCGG$^{6S}$** TGCCG -3' |
| $\mathbf{1^{A}}$ | 5'- GCCGC **TCGA** TGCCG -3' |
| $\mathbf{1^{AMe}}$ | 5'- GCCGC **TCGA$^{Me}$** TGCCG -3' |
| $\mathbf{2^{U4S}}$ | 5'- CGGCA **U$^{4S}$CGA$^{Me}$** GCGGC -3' |
| $\mathbf{2^{T}}$ | 5'- CGGCA **TCGA$^{Me}$** GCGGC -3' |

The DNA MTase M.TaqI was overexpressed and purified as described before [57,77]. For binding experiments with cross-linked duplex ODN, M.TaqI was transferred into storage buffer without reducing agent.

### Denaturing anion exchange HPLC

Denaturing anion exchange HPLC was performed on a Perseptive Poros HQ 10 column (10 × 100 mm, 10 µm) at a flow rate of 2 mL/min. Buffer A consist of 10 mM Tris/HCl, pH 7 and 5 M urea. Buffer B consists of 10 mM Tris/HCl, pH 7, 1 M potassium chloride and 5 M urea. The column was heated in a water bath to 70 °C. Before entering the column, the buffer was pre-heated by passing it through a 0.5 m steel capillary placed in the 70 °C water bath. The DNA was eluted with 20% B (0–5 min), followed by linear gradients with 20–50% B (5–10 min), and 50–65% B (10–20 min). UV absorption was detected at 254 nm and 332 nm. The non-cross-linked parent duplexes dissociate and elute as their respective single strands, while the covalently linked duplexes elute significantly later. HPLC-purified cross-linked duplexes were desalted using a NAP-5 gel filtration column as described below.

### UV spectroscopy

UV spectroscopy was performed using a Varian CARY 3E UV–Vis spectrometer in a 1 cm quartz cuvette. The concentration of ODN was determined in water at 260 nm and 25 °C using the nearest neighbor method [82] to calculate the extinction coefficients. *N*6-Methyl-2'-deoxyadenosine was treated as

2'-deoxyadenosine. For $dG^{6S}$, an extinction coefficient of 8,000 L $mol^{-1}$ $cm^{-1}$ at 260 nm was derived from an experimental spectrum of the free nucleoside and the published extinction coefficient of 24,800 L $mol^{-1}$ $cm^{-1}$ at 342 nm [83]. For $dI^{6S}$ and $dU^{4S}$ the low absorption at 260 nm was neglected.

UV spectra of duplex ODN were recorded from 220 nm to 365 nm at a DNA concentration of 5 μM in phosphate buffer (10 mM NaPi, pH 7.0, 100 mM NaOAc).

Melting curves of duplex ODN (2.5 μM) in phosphate buffer (10 mM NaPi, pH 7.0, 100 mM NaOAc) were recorded at 260 nm between 40 °C and 95 °C in 0.1 K steps at a heating rate of 0.5 K/min. Melting temperatures $T_M$ were obtained as the maximum of the first derivative of the melting curve. Since the melting curves for the cross-linked DNA did not reach their plateau, amplitudes were normalized setting the lowest data point to 0 and the inflection point [84] to 0.5.

## Time course of cross-linking reactions with diiodoethane

Duplex ODN $\mathbf{1}^{\mathbf{I6S}}\cdot\mathbf{2}^{\mathbf{U4S}}$ or $\mathbf{1}^{\mathbf{G6S}}\cdot\mathbf{2}^{\mathbf{U4S}}$ (77 μM) were annealed in reaction buffer (50 μL; 20 mM Tris/HOAc pH 9.0, 10 mM $Mg(OAc)_2$, 50 mM KOAc) and a solution of 1,2-diiodoethane (20 μL, 33 mM) in DMF was added. Slight precipitate of 1,2-diiodoethane re-dissolved upon further mixing. The reaction mixture was incubated at room temperature in the dark. Samples (10 μL) were analyzed after 0 h, 0.5 h, 1 h and 2.3 h or 4 h reaction time by denaturing anion exchange HPLC. The non-cross-linked parent duplexes $\mathbf{1}^{\mathbf{I6S}}\cdot\mathbf{2}^{\mathbf{U4S}}$ or $\mathbf{1}^{\mathbf{G6S}}\cdot\mathbf{2}^{\mathbf{U4S}}$ dissociate into their respective single strands $\mathbf{1}^{\mathbf{I6S}}$ or $\mathbf{1}^{\mathbf{G6S}}$ and $\mathbf{2}^{\mathbf{U4S}}$ eluting at 10.5–10.7 min. Cross-linked duplexes $\mathbf{1}^{\mathbf{I6S\text{-}Et\text{-}S4U}}\mathbf{2}$ ($t_R$ = 13.4 min) and $\mathbf{1}^{\mathbf{G6S\text{-}Et\text{-}S4U}}\mathbf{2}$ ($t_R$ = 13.4 min) elute at significantly higher retention times. After 4 h or 2.3 h, no non-alkylated single strands could be observed, and cross-linked duplexes $\mathbf{1}^{\mathbf{I6S\text{-}Et\text{-}S4U}}\mathbf{2}$ and $\mathbf{1}^{\mathbf{G6S\text{-}Et\text{-}S4U}}\mathbf{2}$ were formed with 80% or 72% yield, respectively. Remaining amounts of ODN eluting around 10.5 min showed low 332 nm absorption and were attributed to S-alkylated single strands.

### Preparative cross-linking experiments

Duplex ODN $\mathbf{1}^{\mathbf{I6S}}\cdot\mathbf{2}^{\mathbf{U4S}}$ or $\mathbf{1}^{\mathbf{G6S}}\cdot\mathbf{2}^{\mathbf{U4S}}$ (77 μM) were annealed in reaction buffer (100 μL; 20 mM Tris/HOAc pH 9.0, 10 mM $Mg(OAc)_2$, 50 mM KOAc) and a solution of 1,2-diiodoethane (40 μL, 33 mM) in DMF was added. Slight precipitate of diiodoethane re-dissolved upon further mixing. The reaction mixture was incubated at room temperature in the dark. After 4 h, the cross-linked duplexes $\mathbf{1}^{\mathbf{I6S\text{-}Et\text{-}S4U}}\mathbf{2}$ ($t_R$ = 13.4 min) and $\mathbf{1}^{\mathbf{G6S\text{-}Et\text{-}S4U}}\mathbf{2}$ ($t_R$ = 13.4 min) were isolated by denaturing anion exchange HPLC. Purified cross-linked duplexes were desalted using illustra NAP-5 gel filtration columns (GE Healthcare). The columns were drained from storage solution by gravity flow and equilibrated with $H_2O$ (10 mL). Combined HPLC fractions (0.5 mL) were applied to the column and allowed to enter the gel bed. The DNA was eluted with $H_2O$ (1 mL), and the concentration of the eluate was determined by UV spectroscopy. The integrity of the cross-linked duplexes was confirmed by analytical denaturing anion exchange HPLC. The desalted duplexes were aliquoted (1 nmol), immediately lyophilized to dryness and stored at −80 °C.

### Linker removal from cross-linked duplex $\mathbf{1}^{\mathbf{I6S\text{-}Et\text{-}S4U}}\mathbf{2}$ with different thiol nucleophiles

Lyophilized $\mathbf{1}^{\mathbf{I6S\text{-}Et\text{-}S4U}}\mathbf{2}$ (1 nmol) was resuspended in buffer (400 μL; 20 mM Tris/HOAc pH 9.0, 10 mM $Mg(OAc)_2$, 50 mM KOAc). Dithiothreitol (DTT, 1 mM), β-mercaptoethanol (BME, 2 mM), ethanethiol (EtSH, 2 mM), or no thiol reagent were added. UV spectra were recorded (DTT: 0 min; BME: 0 min, EtSH: 0–25 min in 5 min increments) from 300 nm to 365 nm to monitor the reappearance of the band >300 nm which is characteristic for the non-alkylated thionucleobases. At the end of each time course, the reaction mixtures (200 μL each) were analyzed by denaturing anion exchange HPLC to confirm the opening of the cross-link.

### Determination of dissociation constants for M.TaqI-DNA complexes

Binding affinities of the non-fluorescent duplex $\mathbf{1}^{\mathbf{I6S}}\cdot\mathbf{2}^{\mathbf{U4S}}$ with an unlocked and the cross-linked duplex $\mathbf{1}^{\mathbf{I6S\text{-}Et\text{-}S4U}}\mathbf{2}$ with a locked target base pair, as well as the hemimethylated native substrate $\mathbf{1}^{\mathbf{A}}\cdot\mathbf{2}^{\mathbf{T}}$ and the fully methylated duplex $\mathbf{1}^{\mathbf{AMe}}\cdot\mathbf{2}^{\mathbf{T}}$ were determined in a competitive fluorescence binding assay using a 36mer duplex ODN containing the fluorescent base analog 2-aminopurine (2AP) at the target position within the recognition sequence of M.TaqI [73]. Fluorescence titrations were performed with a Varian Cary Eclipse fluorescence spectrophotometer using a 10 × 10 mm quartz cuvette at 25 °C with an excitation wavelength of 320 nm and an emission wavelength of 381 nm.

Binding of M.TaqI to duplexes $\mathbf{1}^{\mathbf{I6S}}\cdot\mathbf{2}^{\mathbf{U4S}}$ and $\mathbf{1}^{\mathbf{I6S\text{-}Et\text{-}S4U}}\mathbf{2}$: To a solution (600 μL) of 36mer duplex with 2AP at the target position (200 nM) and duplex $\mathbf{1}^{\mathbf{I6S}}\cdot\mathbf{2}^{\mathbf{U4S}}$ or $\mathbf{1}^{\mathbf{I6S\text{-}Et\text{-}S4U}}\mathbf{2}$ (400 nM) in M.TaqI binding buffer (20 mM TrisOAc, 10 mM $Mg(OAc)_2$, 50 mM KOAc, pH 7.9, 0.01% reduced Triton X-100) and either 1 mM DTT ($\mathbf{1}^{\mathbf{I6S}}\cdot\mathbf{2}^{\mathbf{U4S}}$) or no reducing agent ($\mathbf{1}^{\mathbf{I6S\text{-}Et\text{-}S4U}}\mathbf{2}$) were added stepwise increasing amounts of a solution containing M.TaqI (10 μM), 36mer duplex with 2AP (200 nM) and a duplex $\mathbf{1}^{\mathbf{I6S}}\cdot\mathbf{2}^{\mathbf{U4S}}$ or $\mathbf{1}^{\mathbf{I6S\text{-}Et\text{-}S4U}}\mathbf{2}$ (400 nM) in the same buffers. The relative fluorescence intensity was determined after each addition. A model with one binding site and two binding equilibria was fitted to the fluorescence data from the competi-

tive titrations using the software Scientist (Micromath), holding the known $K_D$ (20 nM) of the 2AP-containing 36mer and M.TaqI constant.

**Re-opening of the cross-linked target base pair during the binding experiment:** A binding experiment with the cross-linked duplex $\mathbf{1}^{\mathbf{I6S\text{-}Et\text{-}S4U}}\mathbf{2}$ in M.TaqI binding buffer without DTT was carried out as described above up to a M.TaqI concentration of 280 nM. Then, the cross-link was re-opened by adding a concentrated DTT solution (0.5 M) to reach a final DTT concentration of 1 mM in both solutions. The 2AP fluorescence decreased upon addition of DTT, indicating that re-opening of the target base pair has occurred and that the re-opened non-fluorescent duplex binds tighter. The titration with M.TaqI was then continued. The fitted binding curve from the titration of the cross-linked duplex $\mathbf{1}^{\mathbf{I6S\text{-}Et\text{-}S4U}}\mathbf{2}$ was overlaid onto the data points before the addition of DTT, and the binding curve from the titration of duplex $\mathbf{1}^{\mathbf{I6S}}\cdot\mathbf{2}^{\mathbf{U4S}}$ was overlaid onto the data points after adding DTT.

## Supporting Information

### Supporting Information File 1

Derivation of equation (1) and Figures S1 and S2.
[http://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-10-239-S1.pdf]

## Acknowledgements

We thank Kerstin Glensk for the DTT-free preparation of M.TaqI and Silke Harms for preparation of duplex $\mathbf{1}^{\mathbf{G6S\text{-}Et\text{-}S4U}}\mathbf{2}$. C.B. thanks the Studienstiftung des Deutschen Volkes for a Ph.D. fellowship.

## References

1. Tong, W. P.; Kohn, K. W.; Ludlum, D. B. *Cancer Res.* **1982,** *42,* 4460–4464.
2. Bodell, W. J. *Chem. Res. Toxicol.* **1999,** *12,* 965–970. doi:10.1021/tx980200c
3. Bodell, W. J.; Pongracz, K. *Chem. Res. Toxicol.* **1993,** *6,* 434–438. doi:10.1021/tx00034a008
4. Chen, F.-X.; Bodell, W. J.; Liang, G.; Gold, B. *Chem. Res. Toxicol.* **1996,** *9,* 208–214. doi:10.1021/tx950097g
5. Fischhaber, P. L.; Gall, A. S.; Duncen, J. A.; Hopkins, P. B. *Cancer Res.* **1999,** *59,* 4363–4368.
6. Wilds, C. J.; Xu, F.; Noronha, A. M. *Chem. Res. Toxicol.* **2008,** *21,* 686–695. doi:10.1021/tx700422h
7. Swenson, M. C.; Paranawithana, S. R.; Miller, P. S.; Kielkopf, C. L. *Biochemistry* **2007,** *46,* 4545–4553. doi:10.1021/bi700109r
8. Guainazzi, A.; Campbell, A. J.; Angelov, T.; Simmerling, C.; Schärer, O. D. *Chemistry* **2010,** *16,* 12100–12103. doi:10.1002/chem.201002041
9. Lloyd, R. S.; Cheng, X. *Biopolymers* **1997,** *44,* 139–151. doi:10.1002/(SICI)1097-0282(1997)44:2<139::AID-BIP3>3.0.CO;2-W
10. Jeltsch, A. *Curr. Top. Microbiol. Immunol.* **2006,** *301,* 203–225. doi:10.1007/3-540-31390-7_7
11. Sundheim, O.; Talstad, V. A.; Vågbø, C. B.; Slupphaug, G.; Krokan, H. E. *DNA Repair* **2008,** *7,* 1916–1923. doi:10.1016/j.dnarep.2008.07.015
12. Yang, C.-G.; Garcia, K.; He, C. *ChemBioChem* **2009,** *10,* 417–423. doi:10.1002/cbic.200800580
13. Tubbs, J. L.; Pegg, A. E.; Tainer, J. A. *DNA Repair* **2007,** *6,* 1100–1115. doi:10.1016/j.dnarep.2007.03.011
14. Stivers, J. T. *Prog. Nucleic Acid Res. Mol. Biol.* **2004,** *77,* 37–65. doi:10.1016/S0079-6603(04)77002-6
15. Brooks, S. C.; Adhikary, S.; Rubinson, E. H.; Eichman, B. F. *Biochim. Biophys. Acta* **2013,** *1834,* 247–271. doi:10.1016/j.bbapap.2012.10.005
16. Huffman, J. L.; Sundheim, O.; Tainer, J. A. *Mutat. Res.* **2005,** *577,* 55–76. doi:10.1016/j.mrfmmm.2005.03.012
17. Yang, W. *Protein Sci.* **2011,** *20,* 1781–1789. doi:10.1002/pro.723
18. Booth, J. D.; Murphy, S. P.; Noronha, A. M.; Wilds, C. J. *Nucleic Acids Symp. Ser.* **2008,** *52,* 431–432. doi:10.1093/nass/nrn219
19. da Silva, M. W.; Wilds, C. J.; Noronha, A. M.; Colvin, O. M.; Miller, P. S.; Gamcsik, M. P. *Biochemistry* **2004,** *43,* 12549–12554. doi:10.1021/bi0486435
20. Noll, D. M.; Noronha, A. M.; Wilds, C. J.; Miller, P. S. *Front. Biosci.* **2004,** *9,* 421–437. doi:10.2741/1246
21. Noronha, A. M.; Noll, D. M.; Wilds, C. J.; Miller, P. S. *Biochemistry* **2002,** *41,* 760–771. doi:10.1021/bi011610u
22. Noronha, A. M.; Wilds, C. J.; Miller, P. S. *Biochemistry* **2002,** *41,* 8605–8612. doi:10.1021/bi025671j
23. Sun, G.; Noronha, A. M.; Miller, S. P.; Wilds, C. J. Synthesis of Building Blocks and Oligonucleotides with {T}$N^3$-Alkylene-$N^3${T} Cross-Links. *Current Protocols in Nucleic Acid Chemistry;* Wiley, 2012; Unit 5. 11.
24. Wilds, C. J.; Booth, J. D.; Noronha, A. M. Synthesis of Building Blocks and Oligonucleotides with {G}O6-Alkyl-O6{G} Cross-Links. *Current Protocols in Nucleic Acid Chemistry;* Wiley, 2011; Unit 5. 9.
25. Wilds, C. J.; Noronha, A. M.; Robidoux, S.; Miller, P. S. *J. Am. Chem. Soc.* **2004,** *126,* 9257–9265. doi:10.1021/ja0498540
26. Huang, H.; Dooley, P. A.; Harris, C. M.; Harris, T. M.; Stone, M. P. *Chem. Res. Toxicol.* **2009,** *22,* 1810–1816. doi:10.1021/tx900225c
27. Kowalczyk, A.; Carmical, J. R.; Zou, Y.; Van Houten, B.; Lloyd, R. S.; Harris, C. M.; Harris, T. M. *Biochemistry* **2002,** *41,* 3109–3118. doi:10.1021/bi010450j
28. Alzeer, J.; Schärer, O. D. *Nucleic Acids Res.* **2006,** *34,* 4458–4466. doi:10.1093/nar/gkl587
29. Matteucci, M. D.; Webb, T. R. *Tetrahedron Lett.* **1987,** *28,* 2469–2472. doi:10.1016/S0040-4039(00)95443-1
30. Webb, T. R.; Matteucci, M. D. *J. Am. Chem. Soc.* **1986,** *108,* 2764–2765. doi:10.1021/ja00270a051
31. Cowart, M.; Benkovic, S. J. *Biochemistry* **1991,** *30,* 788–796. doi:10.1021/bi00217a032
32. Imoto, S.; Chikuni, T.; Kansui, H.; Kunieda, T.; Nagatsugi, F. *Nucleosides, Nucleotides Nucleic Acids* **2012,** *31,* 752–762. doi:10.1080/15257770.2012.726756
33. Nishimoto, A.; Jitsuzaki, D.; Onizuka, K.; Taniguchi, Y.; Nagatsugi, F.; Sasaki, S. *Nucleic Acids Res.* **2013,** *41,* 6774–6781. doi:10.1093/nar/gkt197
34. Kawasaki, T.; Nagatsugi, F.; Maeda, M.; Sasaki, S. *Nucleic Acids Symp. Ser.* **2000,** *44,* 129–130. doi:10.1093/nass/44.1.129

35. Stevens, K.; Claeys, D. D.; Catak, S.; Figaroli, S.; Hocek, M.; Tromp, J. M.; Schürch, S.; Van Speybroeck, V.; Madder, A. *Chem. – Eur. J.* **2011,** *17,* 6940–6953. doi:10.1002/chem.201100067
36. Shibata, T.; Dohno, C.; Nakatani, K. *Nucleic Acids Symp. Ser.* **2009,** *53,* 171–172. doi:10.1093/nass/nrp086
37. Angelov, T.; Guainazzi, A.; Schärer, O. D. *Org. Lett.* **2009,** *11,* 661–664. doi:10.1021/ol802719a
38. Kočalka, P.; El-Sagheer, A. H.; Brown, T. *ChemBioChem* **2008,** *9,* 1280–1285. doi:10.1002/cbic.200800006
39. Erlanson, D. A.; Wolfe, S. A.; Chen, L.; Verdine, G. L. *Tetrahedron* **1997,** *53,* 12041–12056. doi:10.1016/S0040-4020(97)00767-9
40. Coleman, R. S.; Kesicki, E. A. *J. Org. Chem.* **1995,** *60,* 6252–6253. doi:10.1021/jo00125a004
41. Coleman, R. S.; Pires, R. M. *Nucleic Acids Res.* **1997,** *25,* 4771–4777. doi:10.1093/nar/25.23.4771
42. Coleman, R. S.; McCary, J. L. *Bioorg. Med. Chem. Lett.* **1998,** *8,* 3039–3042. doi:10.1016/S0960-894X(98)00561-7
43. Coleman, R. S.; McCary, J. L.; Perez, R. J. *Tetrahedron* **1999,** *55,* 12009–12022. doi:10.1016/S0040-4020(99)00704-8
44. Coleman, R. S.; McCary, J. L.; Perez, R. J. *Tetrahedron Lett.* **1999,** *40,* 13–16. doi:10.1016/S0040-4039(98)80005-1
45. Coleman, R. S.; Pires, R. M. *Nucleosides Nucleotides* **1999,** *18,* 2141–2146. doi:10.1080/07328319908044871
46. Milton, J.; Connolly, B. A.; Nikiforovc, T. T.; Cosstick, R. *J. Chem. Soc., Chem. Commun.* **1993,** 779–780. doi:10.1039/c39930000779
47. Beddows, A.; Patel, N.; Finger, L. D.; Atack, J. M.; Williams, D. M.; Grasby, J. A. *Chem. Commun.* **2012,** *48,* 8895–8897. doi:10.1039/c2cc33400c
48. Coleman, R. S.; Siedlecki, J. M. *J. Am. Chem. Soc.* **1992,** *114,* 9229–9230. doi:10.1021/ja00049a089
49. Xu, Y.-Z. *Tetrahedron* **1996,** *52,* 10737–10750. doi:10.1016/0040-4020(96)00596-0
50. Xu, Y.-Z.; Zheng, Q.; Swann, P. F. *Nucleosides Nucleotides* **1995,** *14,* 929–933. doi:10.1080/15257779508012504
51. Beuck, C.; Weinhold, E. *Nucleosides, Nucleotides Nucleic Acids* **2003,** *22,* 635–639. doi:10.1081/NCN-120021969
52. Cheong, K.-K.; Fu, Y.-C.; Robins, R. K.; Eyring, H. *J. Phys. Chem.* **1969,** *73,* 4219–4227. doi:10.1021/j100846a031
53. Paeschke, K.; McDonald, K. R.; Zakian, V. A. *FEBS Lett.* **2010,** *584,* 3760–3772. doi:10.1016/j.febslet.2010.07.007
54. Ferentz, A. E.; Verdine, G. L. *J. Am. Chem. Soc.* **1991,** *113,* 4000–4002. doi:10.1021/ja00010a057
55. Surby, M. A.; Reich, N. O. *Biochemistry* **1996,** *35,* 2201–2208. doi:10.1021/bi951883n
56. Gowher, H.; Jeltsch, A. *J. Mol. Biol.* **2000,** *303,* 93–110. doi:10.1006/jmbi.2000.4127
57. Holz, B.; Dank, N.; Eickhoff, J. E.; Lipps, G.; Krauss, G.; Weinhold, E. *J. Biol. Chem.* **1999,** *274,* 15066–15072. doi:10.1074/jbc.274.21.15066
58. Allan, B. W.; Beechem, J. M.; Lindstrom, W. M.; Reich, N. O. *J. Biol. Chem.* **1998,** *273,* 2368–2373. doi:10.1074/jbc.273.4.2368
59. Cal, S.; Connolly, B. A. *J. Biol. Chem.* **1997,** *272,* 490–496. doi:10.1074/jbc.272.1.490
60. Cheng, X.; Roberts, R. J. *Nucleic Acids Res.* **2001,** *29,* 3784–3795. doi:10.1093/nar/29.18.3784
61. Jeltsch, A.; Roth, M.; Friedrich, T. *J. Mol. Biol.* **1999,** *285,* 1121–1130. doi:10.1006/jmbi.1998.2389
62. Larivière, L.; Moréra, S. *J. Biol. Chem.* **2004,** *279,* 34715–34720. doi:10.1074/jbc.M404394200
63. Lau, A. Y.; Schärer, O. D.; Samson, L.; Verdine, G. L.; Ellenberger, T. *Cell* **1998,** *95,* 249–258. doi:10.1016/S0092-8674(00)81755-9
64. McCullough, A. K.; Dodson, M. L.; Schärer, O. D.; Lloyd, R. S. *J. Biol. Chem.* **1997,** *272,* 27210–27217. doi:10.1074/jbc.272.43.27210
65. Mernagh, D. R.; Taylor, I. A.; Kneale, G. G. *Biochem. J.* **1998,** *336,* 719–725.
66. Roberts, R. J. *Cell* **1995,** *82,* 9–12. doi:10.1016/0092-8674(95)90046-2
67. Roberts, R. J.; Cheng, X. *Annu. Rev. Biochem.* **1998,** *67,* 181–198. doi:10.1146/annurev.biochem.67.1.181
68. Klimasauskas, S.; Kumar, S.; Roberts, R. J.; Cheng, X. *Cell* **1994,** *76,* 357–369. doi:10.1016/0092-8674(94)90342-5
69. Surby, M. A.; Reich, N. O. *Biochemistry* **1996,** *35,* 2209–2217. doi:10.1021/bi951884f
70. Holz, B.; Weinhold, E. In *Bioorganic Chemistry - Highlights and New Aspects,* 1st ed.; Diederichsen, U.; Lindhorst, T. K.; Westermann, B.; Wessjohann, L. A., Eds.; Wiley-VCH: Weinheim, 1999.
71. Allan, B. W.; Reich, N. O. *Biochemistry* **1996,** *35,* 14757–14762. doi:10.1021/bi9615708
72. Finger, L. D.; Patel, N.; Beddows, A.; Ma, L.; Exell, J. C.; Jardine, E.; Jones, A. C.; Grasby, J. A. *Nucleic Acids Res.* **2013,** *41,* 9839–9847. doi:10.1093/nar/gkt737
73. Holz, B.; Klimasauskas, S.; Serva, S.; Weinhold, E. *Nucleic Acids Res.* **1998,** *26,* 1076–1083. doi:10.1093/nar/26.4.1076
74. Jean, J. M.; Hall, K. B. *Proc. Natl. Acad. Sci. U. S. A.* **2001,** *98,* 37–41. doi:10.1073/pnas.98.1.37
75. Lycksell, P.-O.; Gräslund, A.; Claesens, F.; McLaughlin, L. W.; Larsson, U.; Rigler, R. *Nucleic Acids Res.* **1987,** *15,* 9011–9025. doi:10.1093/nar/15.21.9011
76. Su, T.-J.; Connolly, B. A.; Darlington, C.; Mallin, R.; Dryden, D. T. F. *Nucleic Acids Res.* **2004,** *32,* 2223–2230. doi:10.1093/nar/gkh531
77. Goedecke, K.; Pignot, M.; Goody, R. S.; Scheidig, A. J.; Weinhold, E. *Nat. Struct. Biol.* **2001,** *8,* 121–125. doi:10.1038/84104
78. Abdelhaleem, M. *Methods Mol. Biol.* **2010,** *587,* 1–12. doi:10.1007/978-1-60327-355-8_1
79. Li, Y.; Araki, H. *Genes Cells* **2013,** *18,* 266–277. doi:10.1111/gtc.12040
80. Brosh, R. M., Jr. *Nat. Rev. Cancer* **2013,** *13,* 542–558. doi:10.1038/nrc3560
81. Suhasini, A. N.; Brosh, R. M., Jr. *Adv. Exp. Med. Biol.* **2013,** *973,* 123–144. doi:10.1007/978-1-4614-5037-5_6
82. Cantor, C. R.; Warshaw, M. M.; Shapiro, H. *Biopolymers* **1970,** *9,* 1059–1077. doi:10.1002/bip.1970.360090909
83. Fox, J. J.; Wempen, I.; Hampton, A.; Doerr, I. L. *J. Am. Chem. Soc.* **1958,** *80,* 1669–1675. doi:10.1021/ja01540a041
84. Gottschling, D.; Seliger, H.; Tarrasón, G.; Piulats, J.; Eritja, R. *Bioconjugate Chem.* **1998,** *9,* 831–837. doi:10.1021/bc980074o

## License and Terms

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

# Specific DNA duplex formation at an artificial lipid bilayer: fluorescence microscopy after Sybr Green I staining

Emma Werz[1,2] and Helmut Rosemeyer[*1]

Full Research Paper

Open Access

Address:
[1]Organic Materials Chemistry and Bioorganic Chemistry, Institute of Chemistry of New Materials, University of Osnabrück, Barbarastr. 7, D-49069 Osnabrück, Germany and [2]Ionovation GmbH, Westerbreite 7 (CUT), D-49084 Osnabrück, Germany

Email:
Helmut Rosemeyer[*] - Helmut.Rosemeyer@uos.de

* Corresponding author

Keywords:
artificial lipid bilayers; lipo-oligonucleotide duplexes; nucleic acids; Sybr Green I

*Beilstein J. Org. Chem.* **2014,** *10,* 2307–2321.
doi:10.3762/bjoc.10.240

Received: 16 June 2014
Accepted: 16 September 2014
Published: 02 October 2014

This article is part of the Thematic Series "Nucleic acid chemistry" and is dedicated to Prof. Dr. Wolfgang Junge, University of Osnabrück, Germany.

Guest Editor: H.-A. Wagenknecht



## Abstract

The article describes the immobilization of different probe oligonucleotides (**4**, **7**, **10**) carrying each a racemic mixture of 2,3-bis(hexadecyloxy)propan-1-ol (**1a**) at the 5'-terminus on a stable artificial lipid bilayer composed of 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphoethanolamine (POPE) and 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine (POPC). The bilayer separates two compartments (*cis/trans* channel) of an optical transparent microfluidic sample carrier with perfusion capabilities. Injection of unlabeled target DNA sequences (**6**, **8**, or **9**), differing in sequence and length, leads in the case of complementarity to the formation of stable DNA duplexes at the bilayer surface. This could be verified by Sybr Green I double strand staining, followed by incubation periods and thorough perfusions, and was visualized by single molecule fluorescence spectroscopy and microscopy. The different bilayer-immobilized complexes consisting of various DNA duplexes and the fluorescent dye were studied with respect to the kinetics of their formation as well as to their stability against perfusion.

## Introduction

The post-biosynthetic lipophilization of various biomolecules such as of proteins and carbohydrates is of decisive importance for the correct function of the cell [1]. With the recent discovery of geranylated tRNAs in bacteria [2] the interest in so-called lipo-oligonucleotides (LONs) has grown tremendously [3,4]. The ability to form complex nano-architectures [5,6] as well as self-assembling aggregates such as micelles, vesicles [3] and bilayer formation of nucleolipids [7] offers numerous possibilities, e.g., for drug delivery [3]. Simultaneously, the interaction of such nanostructures with lipid membranes becomes an ever-greater focus [5,6,8-11].

The study of the interactions of single- and double-stranded nucleic acids with lipid bilayers is, therefore, of significant importance, particularly for the following reasons: 1. for the optimization of the in vivo delivery of lipophilic siRNAs [12-14],

2. for the development of analytical techniques for the detection of nucleic acids [15-17], 3. for structure elucidation of complex aggregates formed by such natural nanostructures [5,6] and 4. for cell-surface engineering [18].

In a preceding manuscript [19] we reported the lipid bilayer immobilization of lipo-oligonucleotides carrying a racemic bis(hexadecyloxy)propan-1-yl tag (**1a**) at the 5'-termini at an artificial lipid bilayer–water phase boundary. These were prepared using the cyanoethyl phosphoramidite **1b**. A specific duplex formation with complementary cyanine-5 (Cy5)-labelled DNA strands (**2b**) – prepared by using compound **2a** – was proven by fluorescence microscopy. Now, we simplify this technique by hybridizing an unlabelled DNA target strand to the bilayer-immobilized lipo-oligonucleotide and by using Sybr Green I (**3**, SG) [20-22] as a fluorescent double strand indicator [20-22]. This bears the advantage that the target DNA – the presence or absence of which is going to be analysed – should not be labelled separately with a fluorochrome tag such as cyanine-5 (Cy5) or TAMRA, neither by chemical synthesis nor by a polymerase chain reaction (PCR).

The chemical formulae (**1**–**3**) as well as the lipo-oligonucleotide sequences the paper is dealing with are shown in Figure 1.

## Results and Discussion

Figure 2 displays the six DNA duplexes which have been assembled at the artificial lipid bilayer and visualized by addition of Sybr Green I.

In the following the results of these experiments regarding the kinetics of incorporation into the bilayer (or its formation) as well as the release of the ternary complex (or its disaggregation) upon perfusion of the *cis* channel is described. Figure 3 illustrates a scheme of the experimental setup (for details see Experimental).

Experiment A: **4** + **5** + **SG**. In a preceeding manuscript we have studied the bilayer incorporation of the lipo-oligonucleotide **4** (5'-d(**1a**-TAG GTC AAT ACT)-3') and its successful and specific duplex formation with the cyanine-5-labeled complementary strand **5** (5'-d[(Cy5)-p-AGT ATT GAC CTA]-3') [19].

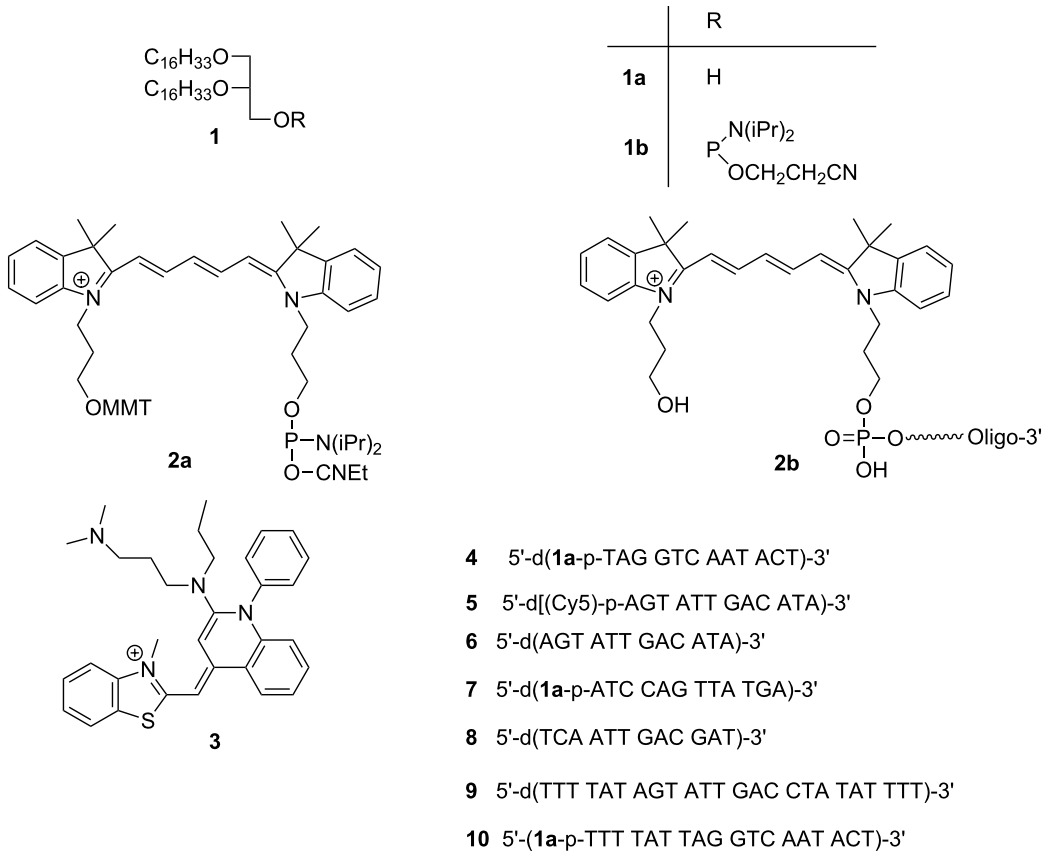



**Figure 1:** Chemical formulae and lipo-oligonucleotide sequences.

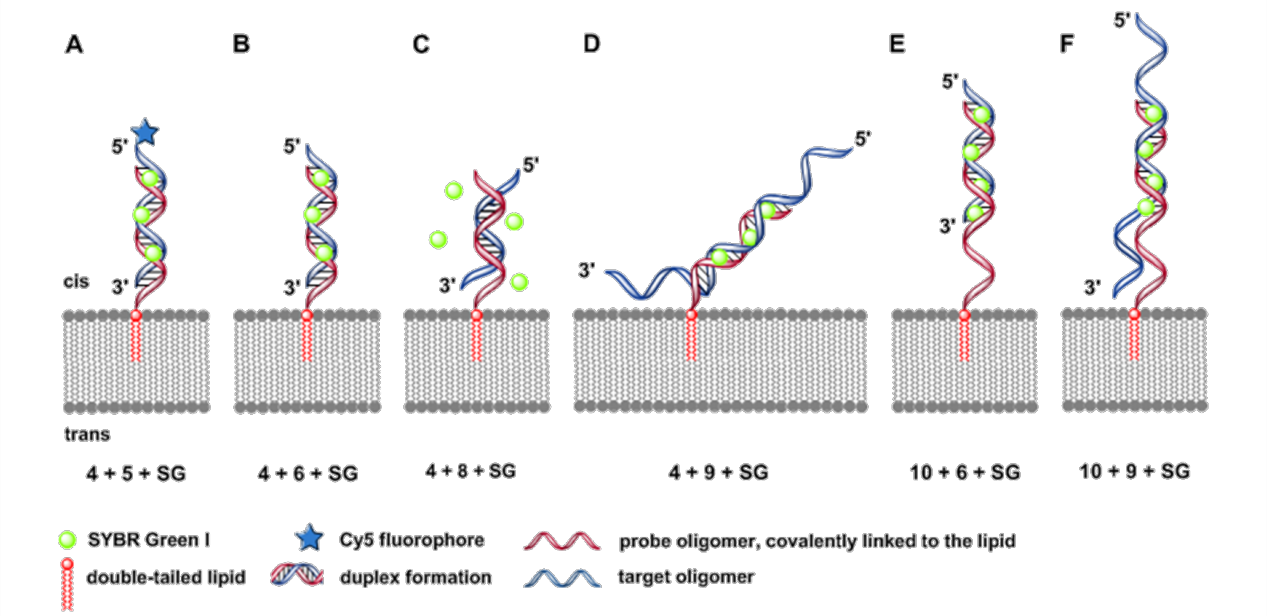



**Figure 2:** Schematic illustrations (experiments A–F) of the specific DNA duplex formation at artificial lipid bilayers forming structures of various membrane-bound nucleic acid–dye complexes. A) Covalent bound Cy5-fluorophore and intercalated SG with DNA duplex were irradiated simultaneously at 635 nm and 470 nm. B) 100% match of the duplex formation with 12 bp-target oligomer. C) 50% match (6 bp) of the duplex formation with 12 bp-target oligonucleotide. D) 24-mer target oligonucleotide with AT-rich tails, each tail with 6 AT-bases. E) 18-mer probe oligonucleotide with a free AT-rich tail of 6 bases as a spacer covalently linked to the double-tailed lipid, forming a 12 bps duplex with the target sequence. F) The same lipo-oligonucleotide as described in (E) forms a 12 bps duplex in combination with a target oligonucleotide, resulting in free overhanging AT-rich tails at the 5'- as well as 3'-ends, each tail containing 6 nucleobases.

As a negative control it had been shown that a membrane-bound duplex formation between the lipo-oligonucleotide **7** (5'-d(**1a**-ATC CAG TTA TGA)-3') and the oligomer **5** failed. Now, we repeated the first experiment (**4** + **5**) successfully and run a z-scan of the bilayer after 45 min of incubation with a laser irradiation of the Cy5 dye at 635 nm (Figure 2).

Then, a Sybr Green I (**3**) solution in dimethyl sulfoxide (≈1 µg/mL) was added to the *cis* compartment of the slide. In the first experiment (A) laser irradiation of the intercalated dye was performed at 470 nm. In orienting experiments Sybr Green I was irradiated at 470 nm, in order to see if this dye stains the complementary DNA strands **4** + **5**, which are known to form a duplex at the bilayer membrane [19]. Firstly, we evaluated the optimal measuring conditions such as the volume, the Sybr Green I concentration as well as of the lipo-oligonucleotides. At this point we first wanted to find out if a non-optimal wavelength of 470 nm would be sufficient for an efficient irradiation of the dye. It turned out, however, that an irradiation of Sybr Green I needs an irradiation wavelength which meets the optimal absorption wavelength to a higher extent. In subsequently performed experiments (B to F) the intercalated Sybr Green I was irradiated at 488 nm because this wavelength is close to its absorption maximum at 494 nm [20-22]. The experiments confirmed a duplex formation in an antiparallel mode on the bilayer surface being stable towards perfusion (Figure 4).

Figure 5 demonstrates the difference in the bilayer brightness upon irradiation of either cyanine-5 (635 nm) [19] or of the ternary complex with SG (**4** + **5** + **SG**) (470 nm). For the determination of the bilayer brightness a region-of-interest (ROI) was defined around the bilayer, and subsequently the density of the intensity counts were summarized (for details, see Experimental).

Besides the bilayer brightness, also the diffusion time, $t_D$, of the lipo-oligonucleotide duplex–Sybr Green I complex (i) within the bilayer (location 1) and (ii) above the bilayer (location 2) was determined (Figure 6).

Table 1 and Table 2 summarize the diffusion times [$t_D$ (ms)] of the complexes **4·5·SG** and **4·6·SG** before and after a perfusion at the two locations 1 and 2 as illustrated in Figure 6. It can be clearly seen that the diffusion time of the aggregates is significantly longer within the bilayer compared to the region slightly above (location 2), proving the strong immobilization of the lipophilized DNA duplex within the bilayer via the lipophilic head group. The fastest diffusion of the aggregates occurs in free solution without bilayer (Table 2) [19].

Experiment B: **4** + **6** + **SG** and control experiment: **7** + **6** + **SG**. In a second series of experiments we used an unlabelled target DNA (**6**) for duplex formation at the lipid bilayer and added

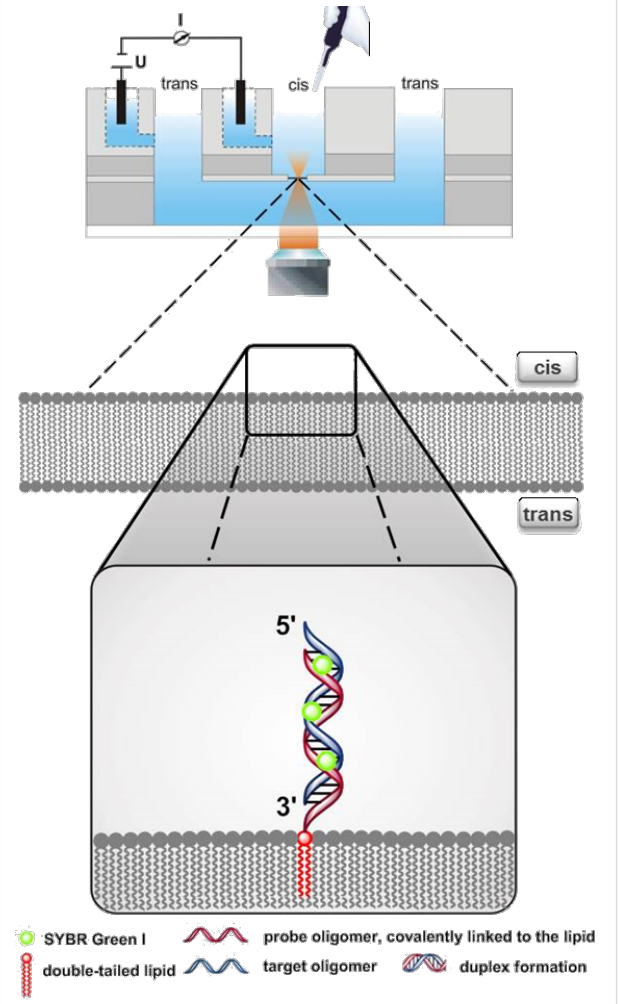



**Figure 3:** Experimental setup. Schematic drawing of the laser scanning microscope, the optical transparent microfluidic bilayer slide, and the lipid bilayer with the incorporated double-tailed lipo-oligonucleotides as well as of Sybr Green I, forming a membrane - bound nucleic acid duplex–dye complex.

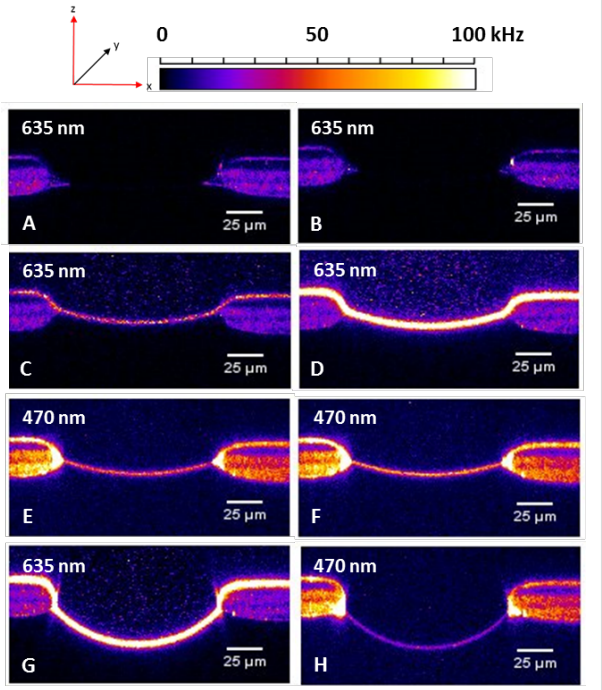



**Figure 4:** Chronological protocol of duplex formation of the lipo-oligonucleotide **4** with the complementary Cy5-labelled target strand **5** with simultaneous staining of the DNA duplex with Sybr Green I. A) Z-Scan of an empty bilayer; after the bilayer has been formed, both compartments (*cis* and *trans*) were perfused for 30 s (1.1 mL/min, each). Cy5 irradiation at 635 nm. The z-scan as in (A) with a Sybr Green I irradiation at 470 nm is not shown. As demonstrated in (A) there is no fluorescence in the bilayer. B) Z-Scan after addition of the oligomer **4** (8 µL, 500 nM) to the *cis* compartment, followed by 45 min of incubation, Cy5 irradiation at 635 nm. C) Z-Scan after addition of the oligomer **5** (8 µL, 50 nM) to the *cis* compartment and 45 min of incubation; Cy5 irradiation at 635 nm. D) Z-Scan after addition of Sybr Green I (8 µL), followed by 45 min of incubation; Cy5 irradiation at 635 nm. E) Z-Scan as in (D), but with Sybr Green I irradiation at 470 nm. F) Z-Scan after an incubation of 5 min and Sybr Green I irradiation at 470 nm. G) Z-Scan after 30 s of perfusion of the *cis* compartment (1.1 mL/min) and Sybr Green I irradiation at 635 nm. H) Z-Scan as in (G), but with Sybr Green I irradiation at 470 nm.

only a DMSO solution of Sybr Green I to the *cis* compartment of the bilayer slide. As can be seen from Figure 7 also the hybridization of an unlabelled target oligomer **6** with the lipo-oligonucleotide **4** can be proved by adding the dye and irradiation at 488 nm. In case of the incorporation of the lipo-oligonucleotide **7** into the bilayer and addition of the non-

**Table 1:** Diffusion times [$t_D$ (ms)] of the **4·5·SG** complex in the presence of a lipid bilayer, either by irradiation of the cyanine-5 dye or by irradiation of intercalated Sybr Green I within the **4·5·SG** complex, before and after perfusion of the *cis* compartment. Location 1: bilayer; location 2: solution in close proximity to the bilayer (Figure 6).

| | sample **4·5·SG** complex | | | |
|---|---|---|---|---|
| fluorescence signal | Cy5 – irradiation$_{(635\ nm)}$ | | **SG** – irradiation$_{(470\ nm)}$ | |
| $t_D$ at location 1 [ms] | before perfusion | after perfusion | before perfusion | after perfusion |
| | 17.23 ± 2.0[a] | 10.33 ± 2.0 | 0.60 ± 0.09 | 1.20 ± 0.36 |
| $t_D$ at location 2 [ms] | 2.47 ± 0.3[a] | 1.28 ± 0.05 | 0.72 ± 0.11 | 0.13 ± 0.05 |

[a]Data were taken from [19].

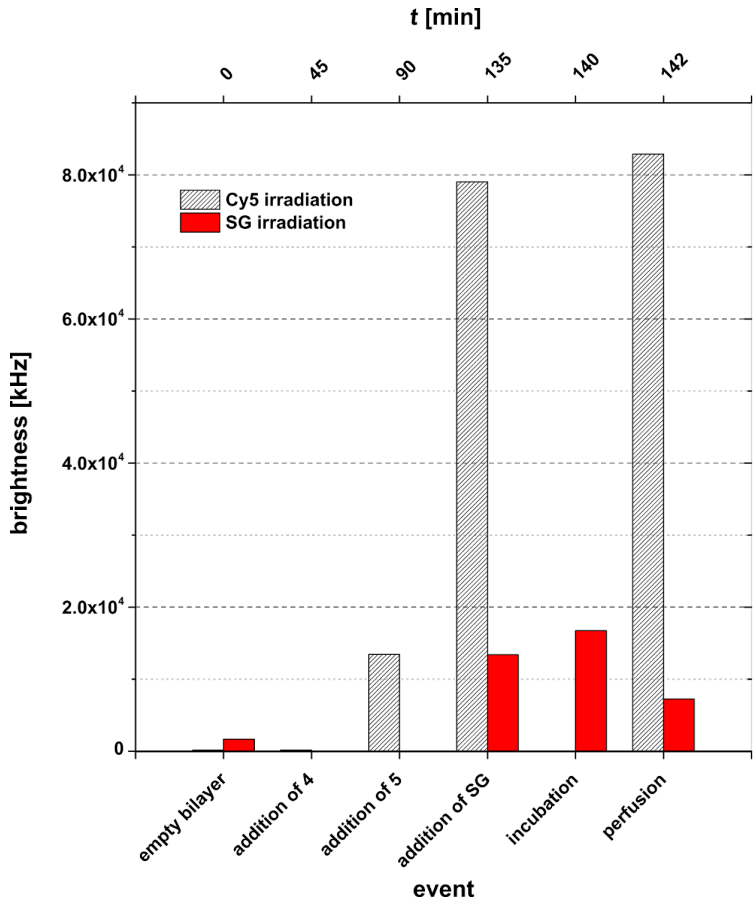



**Figure 5:** Comparison of the bilayer brightness intensity with either Cy5 (irradiation: 635 nm) or Sybr Green I staining (irradiation: 470 nm) of the **4** + **5** + **SG** complex.

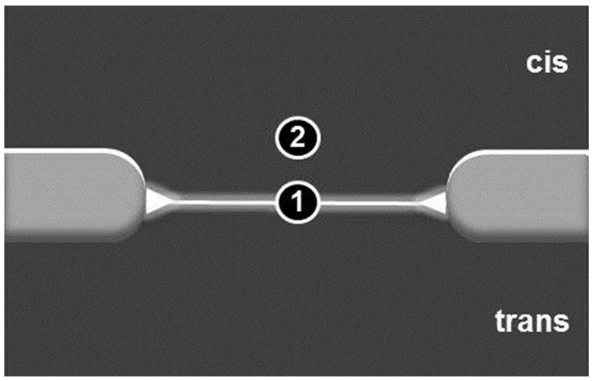



**Figure 6:** Scheme of a z-scan of a lipid bilayer showing two locations for measurements of the diffusion times. Location 1: bilayer; location 2: solution in close proximity to the bilayer surface.

**Table 2:** Diffusion times [$t_D$ (ms)] of the **4**·**6**·**SG** complex in the presence of a lipid bilayer in free solution without bilayer as well as in the presence of a bilayer after a 1. and 2. perfusion of the *cis* compartment. Location 1: bilayer; location 2: solution in close proximity to the bilayer (Figure 6).

| | sample **4**·**6**·**SG** complex | |
|---|---|---|
| $t_D$ in solution without bilayer [ms] | 0.03 ± 0.002 | |
| | perfusion | |
| $t_D$ at location 1 [ms] | 1. | 2. |
| | 1.30 ± 0.28 | 1.56 ± 0.22 |
| $t_D$ at location 2 [ms] | 0.48 ± 0.71 | 0.07 ± 0.03 |

complementary unlabelled oligomer **6** (negative control experiment), no duplex formation occurs (data not shown).

From Figure 7 it can be seen that a stepwise mixing of the lipo-oligonucleotide **4** and the complementary strand **6** with an intermediary incubation of 43 min for an optimal insertion of **4** into the bilayer, followed by a further waiting period (41 min) for duplex formation leads to a bilayer brightness of 70% upon addition of Sybr Green I. Two subsequent incubation periods of totally 40 min enhances the normalized brightness of 100%

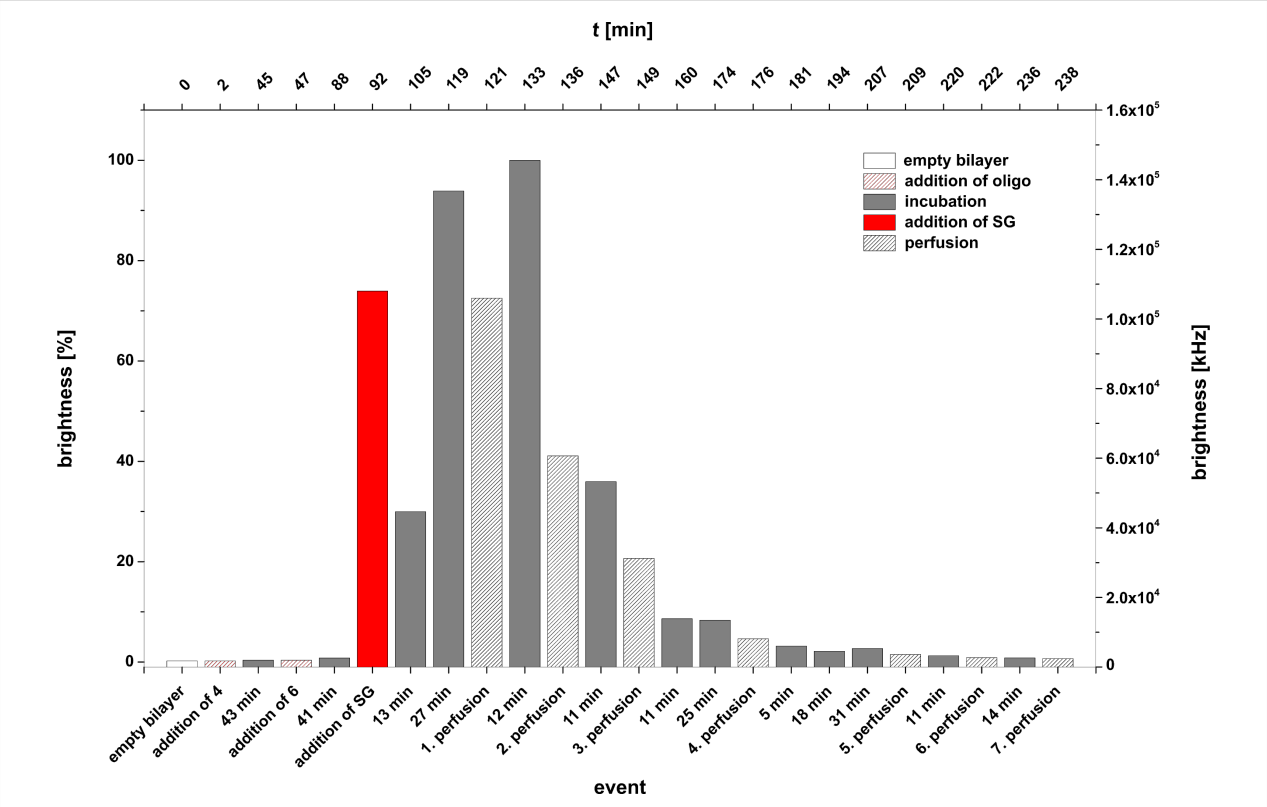



**Figure 7: Experiment B (4 + 6 + SG)**. Bilayer brightness as function of the various events (addition of oligonucleotides, addition of Sybr Green I, perfusion number as well as of time [min]) within one graph.

(≈15.2 × $10^4$ kHz). In the following 7 perfusion steps (see Experimental), interrupted by several incubation periods over a period of about 2 h, led to an almost total decrease of the brightness dropping to about 5%. This decrease might be due to either a release of the dye from the intact duplex or to a full disintegration of the ternary complex (**4** + **6** + **SG**).

Experiment C: **4** + **8** + **SG**. In this experiment the lipo-oligonucleotide **4** is inserted into the lipid bilayer, followed by a dodecamer (**8**) which matches the sequence of **4** only in its innermost part to 50%. An addition of Sybr Green I does not lead to a fluorescent bilayer indicating an unstable ternary complex at room temperature. This is in contrast to the experiment B where a 100% match of both dodecamers exists.

Experiment D: **4** + **9** + **SG**. In a further series of experiments we studied the duplex formation of the lipo-oligonucleotide **4** and the oligomer **9** at the lipid bilayer–water phase boundary layer. The oligomer **9** matches **4** to 100% within its innermost part but contains hexameric overhangs at both termini. Duplex formation was again indicated using Sybr Green I.

In this case an interesting phenomenon was observed: after immobilization of oligomer **4** within the bilayer and after addition of the complementary strand **9** 30 min later as well as of Sybr Green I, further 30 min later, a full development of maximal fluorescence (normalized maximal brightness of 100% ≈ 5.6 × $10^4$ kHz) of the bilayer could only be observed after ≈1 h (Figure 8). This time of formation of the ternary complex is significantly slower than in the experiments with blunt-ended duplex formation reactions described before, which show an almost spontaneous complex formation with the intercalating dye. Moreover, the complex between **4** and **9** as well as of Sybr Green I at the bilayer seems to be highly labile because already a single perfusion step of the *cis* compartment (1 min, 1.1 mL of buffer at $t$ = 143 min) leads to an almost 90% disappearance of the fluorescence. Interestingly, however, is the finding that a renewed addition of a Sybr Green I solution to the *cis* compartment (at $t$ = 168 min) leads again to the appearance of fluorescence 40 min later (brightness, 70%), indicating a partial reconstruction of the complex consisting of **4**, **9**, and the intercalating dye.

The results described above offer several possibilities of interpretation: The target strand **9** carries 3'-terminally to the recognition site of the lipo-oligonucleotide **4** an overhang of 6 nucleotides in length. Therefore, duplex formation between **4** and **9** over a full length of 12 base pairs leads inevitably to a

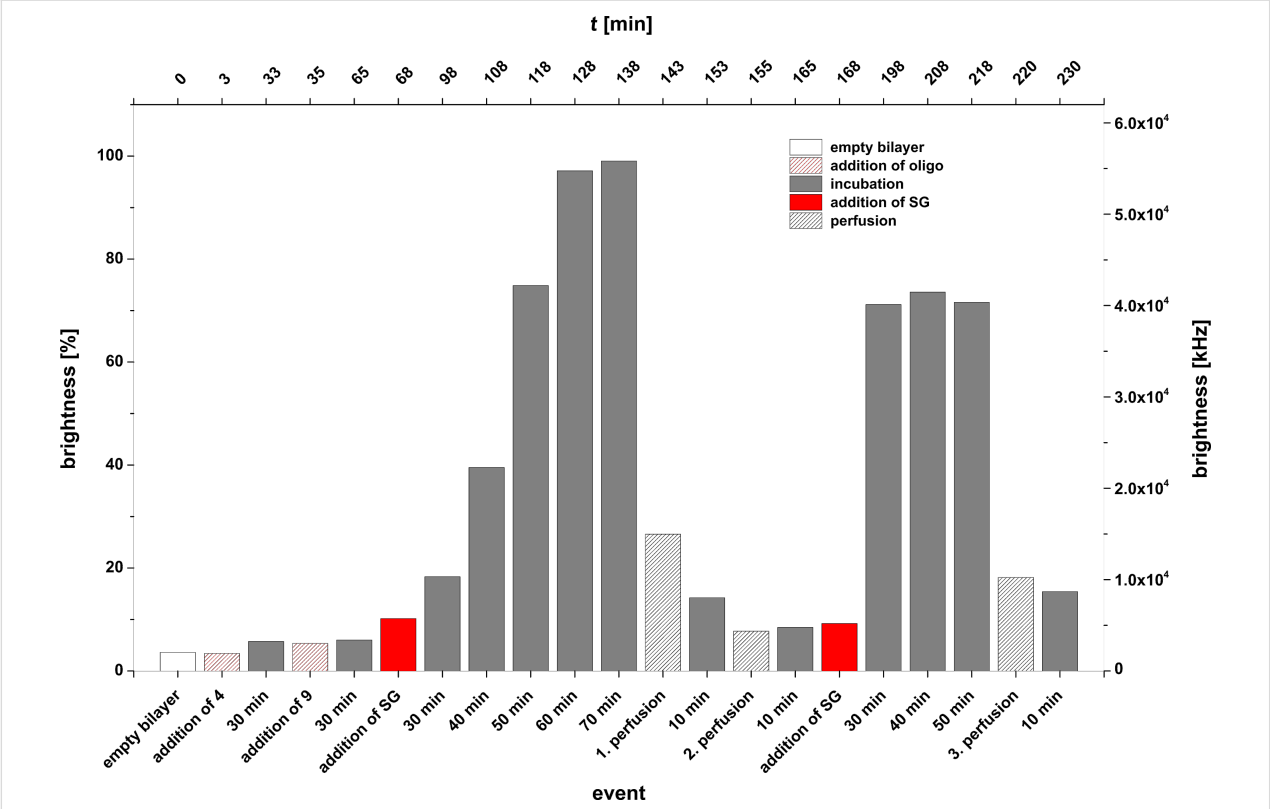



**Figure 8:** Experiment D (**4** + **9** + **SG**). Bilayer brightness as function of time and the various events [addition of **4** (at *t* = 3 min); addition of **9** (at *t* = 35 min), addition of Sybr Green I (at *t* = 68 min), incubation periods, perfusions].

clash of the overhang with the bilayer surface. As the head groups of the neighbouring bilayer molecules (POPC, POPE) are both positively charged and the internucleotide residues of the overhanging oligonucleotide are negatively charged, a strong Coulomb attraction should exist which can only be shielded by solvation of the lipid head groups and of the nucleic acid phosphodiester groups as well as by the metal cations of the surrounding buffer. The supramolecular assembly of all reaction partners might look like as shown in Figure 9. From this figure it can be deduced that two different tilt angles ($\theta_1$ and $\theta_2$) probably exit between the bilayer surface and the double helical part ($\theta_1$) as well as between the single stranded sequence and the surface ($\theta_2$). A stretched-out, parallel association of the full length nucleic acid with the bilayer surface can be most probably ruled out. In such a case a conformational constraint would arise between the lipid head group and the appending oligonucleotide. Moreover, a Cy-5-labelled oligonucleotide such as compound **5** should be bound tightly to the bilayer surface leading to a strong and stable fluorescence of the bilayer which could not be observed.

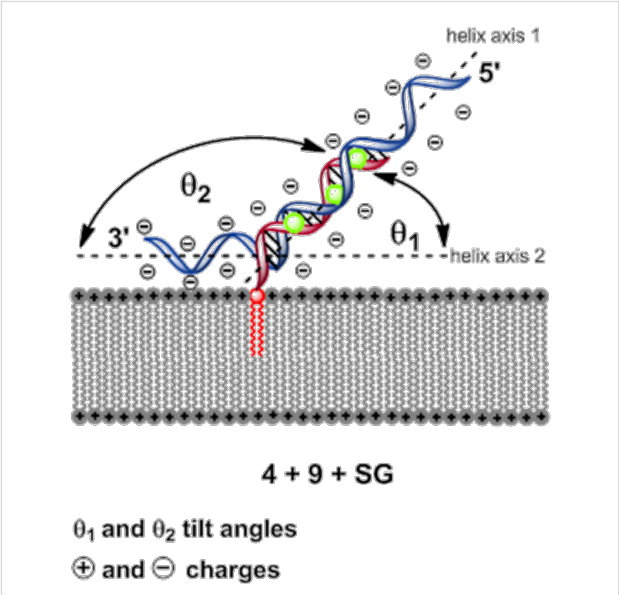



**Figure 9:** Conceivable geometry of the complex at the bilayer surface (*cis* compartment).

It is, however, obvious that a severe disturbance of either the complex formation between the probe lipo-oligonucleotide **4** and the target nucleic acid **9** and/or the intercalation of the Sybr Green dye by the overhang take place. This might be the reason for the instability of the assembled complex at the lipid bilayer.

Experiment E: **10** + **6** + **SG** as well as in a reversed order of addition (**SG** + **10** + **6**). Next, the oligomer **10** was prepared which contains the lipophilic 5' head group **1a** and the recognition sequence separated by an oligonucleotide spacer having the same length as the target sequence overhang (6 bp).

From Figure 10 it can be seen that in this case the ternary complex formation occurs spontaneously upon addition of a first portion of the dye solution. This is, however, not a stable situation because an incubation of only 6 min (without perfusion!) leads to a reduction of the bilayer brightness of just 20%. Several further waiting periods, interrupted by 6 perfusions, reduced the bilayer brightness from the original 100% (100% normalized, ≈6.6 × $10^5$ kHz) to less than 5% within 2.5 h. A further addition of Sybr Green I at $t$ = 212 min, however, enhances the bilayer brightness again to about 20%, indicating the presence of a substantial amount of the intact DNA duplex at the lipid bilayer surface on the *cis* side. Subsequent incubation periods, followed by single perfusion steps reduced the bilayer brightness each time to about 5%. A repeated Sybr Green I addition at $t$ = 234 min leads again to an enhancement of the brightness value of 18%. A 4$^{th}$ addition of Sybr Green I at 251 min results in an only very slight brightness enhancement of ≈5% (Figure 11).

In Figure 12 the z-scans of the Experiment E – before and directly after Sybr Green I additions – are displayed. As already shown in Figure 10 and Figure 11 the bilayer brightness increases with each Sybr Green I addition. After 6 perfusions and incubation steps the brightness is reduced to almost zero ($t$ = 211 min), however, with the next additions of the dye the brightness increases again. Only after seven perfusion and incubation periods the brightness does no longer increase significantly, which might be traced back to the fact that the immobilized DNA duplex is meanwhile dissociated, and that the non-lipophilized strand is has been washed out.

In a further experiment using the lipo-oligonucleotide **10** we changed the order of the addition of the three components of the ternary complex to the *cis* compartment of the bilayer slide (Figure 13). Within the first 11 min we studied the interaction between Sybr Green I and the empty bilayer. As can be seen no ponderable brightness of the bilayer could be detected after addition of the dye. However, after addition of the 24-mer **10** (at

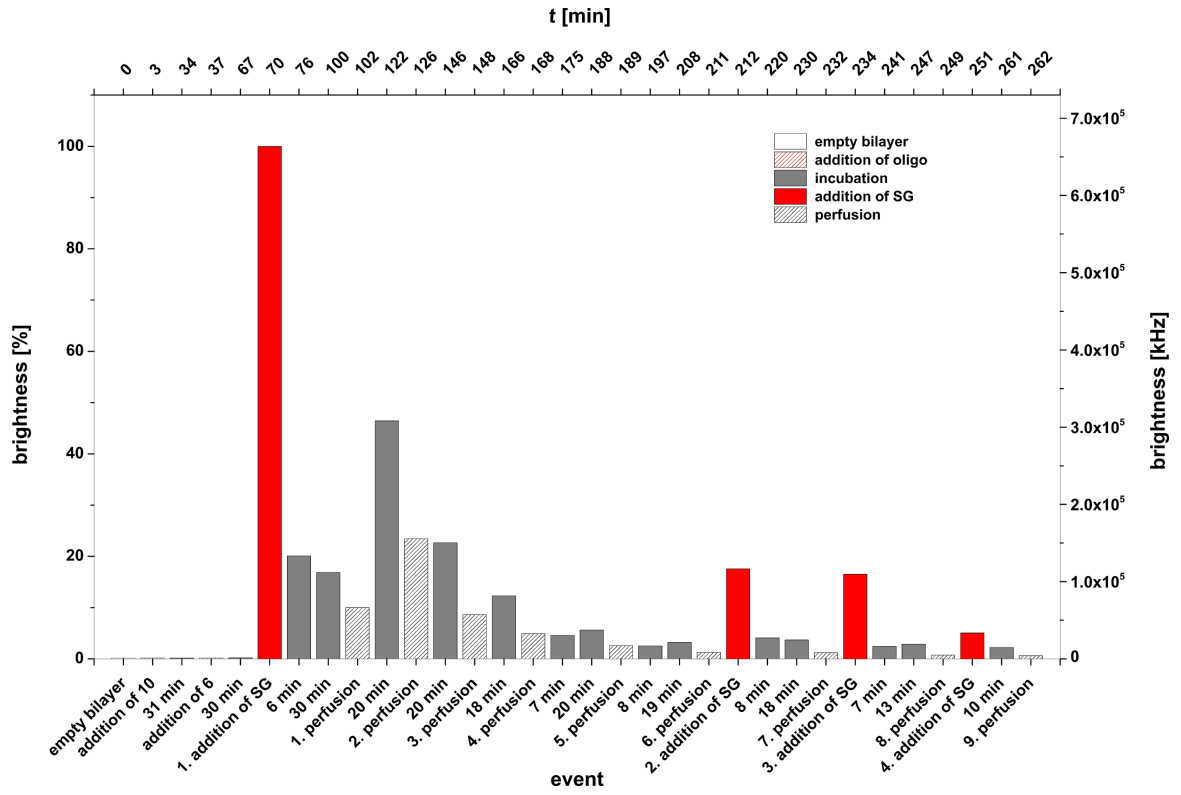



**Figure 10:** Experiment E (**10** + **6** + **SG**). Bilayer brightness as a function of time as well as of various events [addition of the lipo-oligonucleotide **10** (at $t$ = 3 min); addition of the oligomer **6** (at $t$ = 37 min); addition of Sybr Green I (at = 70 min, 212 min, 234 min, and 251 min); incubation periods, perfusions].

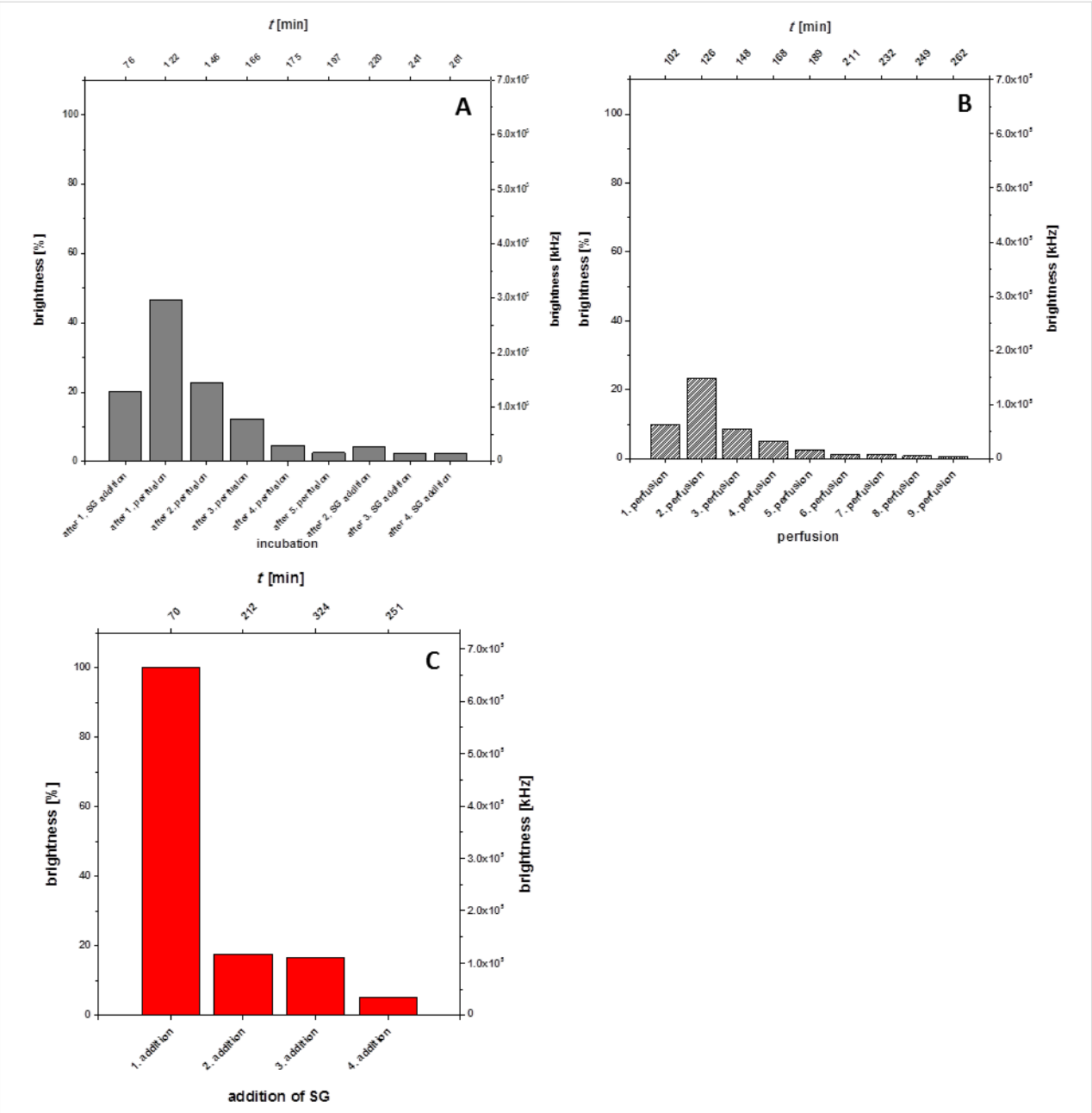



**Figure 11:** Experiment E (**10** + **6** + **SG**). Bilayer brightness as function of the incubation time (A), of the perfusion number (B) and of Sybr Green I addition (C).

$t$ = 13 min) the brightness increased to about 60% (normalized brightness, 100% ≈6.5 × $10^4$ kHz) and stays almost constant (35–50%) until $t$ = 28 min indicating a strong interaction between the dye and the single stranded 24-mer **10** [20]. Such an interaction could not be observed between the dye and the duplex of the 12-mers **4** and **7**. A subsequent perfusion reduces the brightness significantly to 5%. Interestingly, a subsequent waiting period of 10 min leads again to an increase of the bilayer brightness back to about 25%. This is probably due to an additional delivery of the complex **10**·**SG** from the *cis* compartment or from the *cis* side of the hydrophobic Teflon-made annulus into the lipid bilayer.

At $t$ = 63 min a second portion of the dye solution and at $t$ = 73 min the oligomer **6**, being partly complementary to the lipo-oligonucleotide **10**, were added. Within an incubation time of 22 min after addition of **6** the full bilayer brightness (100 % at $t$ = 95 min) was reached. Subsequent perfusion steps (1–4) interrupted by several incubation periods reduced the bilayer brightness from 100 to 10% (Figure 14). These results clearly

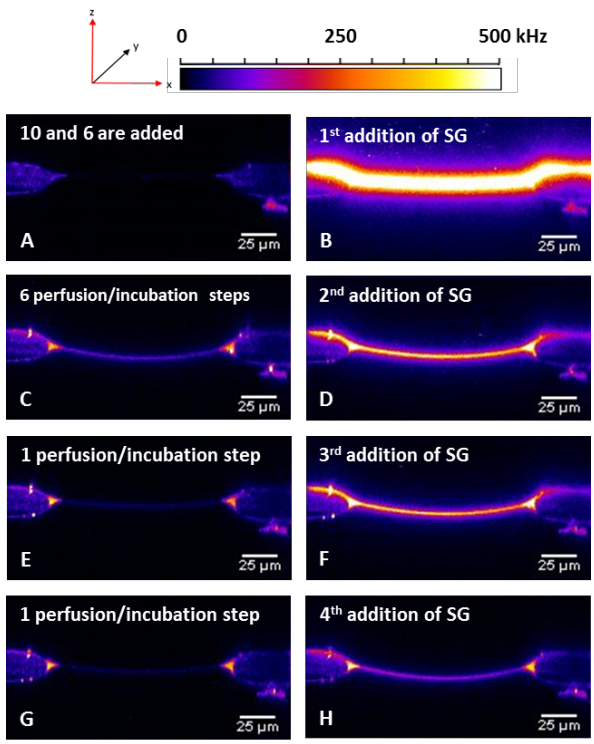



**Figure 12:** Z-scans of the experiment E before and after the addition of Sybr Green I. A) Z-scan after the addition and incubation of the lipo-oligonucleotide **10** and of the complementary target strand **6**, respectively. There is no brightness of the bilayer. B) Z-scan quite after the first addition of Sybr Green I. The brightness rises immediately because of the DNA duplex formation of the lipo-oligonucleotide **10** and of the complementary target strand **6**. C) Z-Scan after six perfusions each followed by an incubation step resulting in an decreasing of the brightness in the bilayer. D) Second addition of SG I leads to an increase of the brightness. E) Only one perfusion/incubation step removes the brightness of the bilayer completely. F) The third addition of Sybr Green I leads to an increase of the brightness in the bilayer on the same level as it was reached after the second SG addition. G) After a further perfusion/incubation step the brightness of the bilayer is removed again, H) but it does not rises as high as shown before.

indicate a significantly higher stability of the ternary complex at the bilayer compared to the complex of experiment D.

Experiment F: (**10** + **9** + **SG**). In the following the lipo-oligonucleotide **10** (24-mer) was immobilized on the lipid bilayer. After an incubation time of 30 min, an equimolar amount of the oligomer **9** – an oligomer of the same length and a central recognition sequence of 12 bases – was added, followed by addition of Sybr Green I after a further waiting period (30 min). Figure 15 displays the full protocol of mean brightness values (both, in % and kHz) as a function of time or the various events (incubation periods or perfusion). After the addition of the dye as last component of the ternary complex at $t$ = 66 min, a maximal brightness is slowly developed within about 2 h. In this case, however, the maximal brightness (normalized 100% ≈6 × $10^4$ kHz) does not reach the value which had been observed in the experiments A, B, and E. Only in case of experiment D the maximal brightness is even lower than in case F (2.5 × $10^4$ kHz).

18 Subsequent perfusions, interrupted by waiting periods of 10 min each, gave after 382 min a bilayer brightness of ≈5% (= 1 × $10^4$ kHz). This indicates that the formation of the ternary complex **10**·**9**·**SG** occurs slowly, but when it is once formed, it remains stable for more than 6 h.

## Conclusion

Figure 16 shows a comparison of the kinetics of complex formation at the lipid bilayer surface for 5 experiments within the first 60 min (brightness [kHz] vs time [min]).

As can be seen, the fastest ternary complex formation occurs for **10** + **6** + **SG** (Experiment E) where the complex formation occurs at a distance of a hexamer spacer between the bilayer surface and the DNA duplex region with intercalated Sybr Green I. It can also be observed that, in this case, a maximal brightness at ≈6.5 × $10^5$ kHz is reached.

If the spacer is missing, the complex (**4** + **6** + **SG**, Experiment B) formation occurs significantly slower; the maximal brightness is reached at 1.5 × $10^5$ kHz. The same results are found in case of experiment A (**4** + **5** + **SG**) with two different fluorescent dyes – intercalating Sybr Green I and pending cyanine-5 (Figure 5).

For the scenario of experiment F (Figure 2) the complex formation kinetics drops even further, but the brightness reaches a plateau at ≈2 × $10^5$ kHz. The by far slowest complex formation is observed in case of experiment D (Figure 1) in which a steric clash between the ternary complex and the bilayer surface occurs. Here, no plateau value of the bilayer brightness is reached, and the complex proved to be highly labile.

Further control experiments such as 1. **6** + **7** (non-complementary strands + **SG**); 2. **5** + **7** (non-complementary strands) + **SG**; 3. **6** (lipophilized 12-mer) + **SG**; 4. empty bilayer + **SG** do not result in fluorescent bilayers. However, addition of Sybr Green I followed by the 18-mer **10** (or in the reversed order) to the *cis* chamber of the bilayer slide gives almost spontaneously a fluorescent bilayer with brightness values of ≈5 × $10^4$ kHz.

Further experiments concerning a multiple compartment chamber with a common aqueous sub-phase, as well as a thermostated device for a suppression of nonspecific base pairing are underway [15]. Moreover, ab initio molecular-dynamics- (AIMD)- and ab initio Monte-Carlo- (AIMC)- calculations of

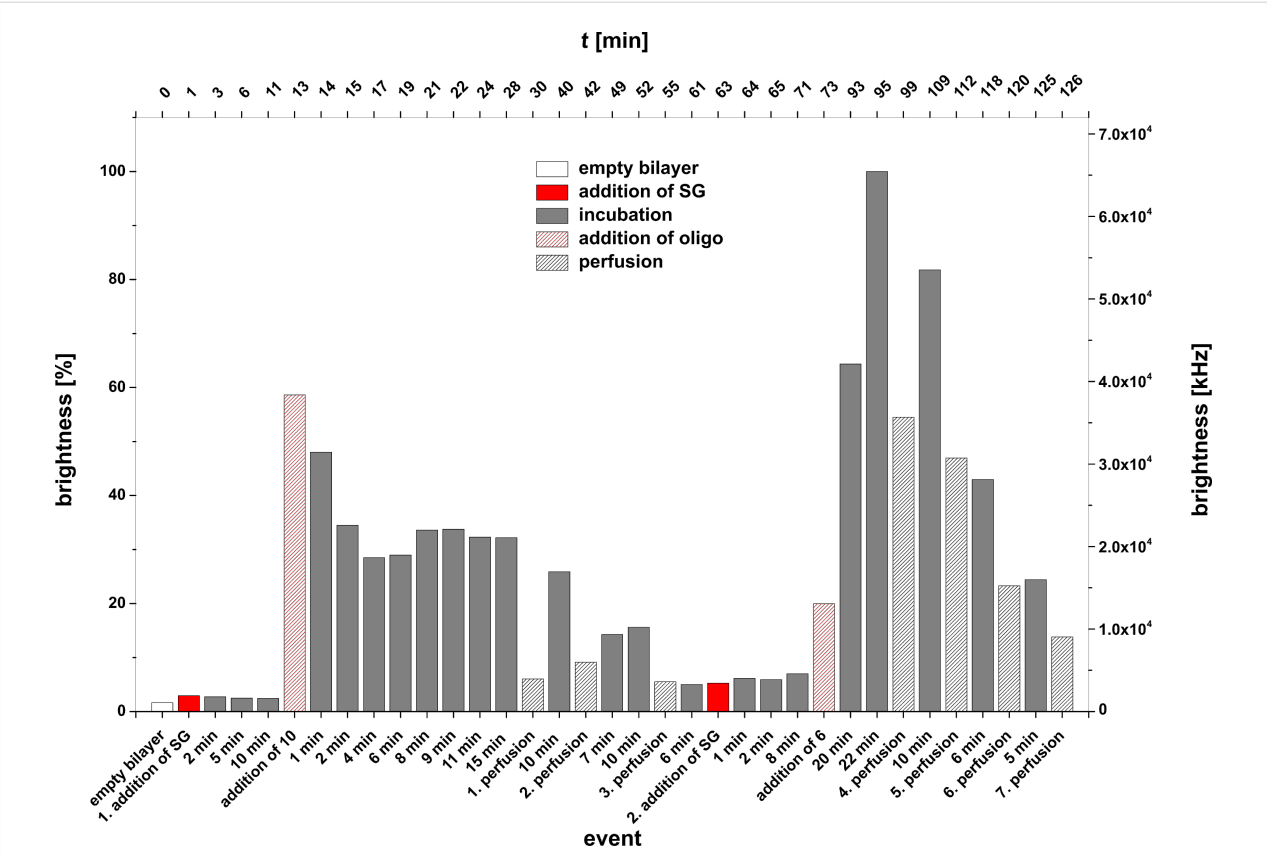



**Figure 13:** Experiment E in reversed order of component addition; **SG** + **10** + **6**.

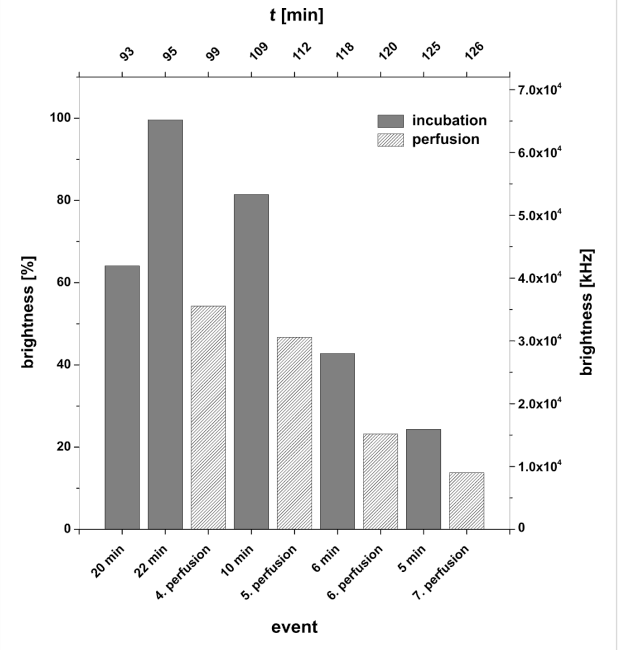



**Figure 14:** Bilayer brightness as a function of perfusion number and incubation periods for the experiment E in reversed order of component addition (**SG** + **10** + **6**).

lipid–nucleic acid complexes with MC and MD simulations of semi-quantitative, "course-grained" models (CGM) are going to be performed (Prof. Dr. Philipp Maass, Department of Statistical Physics, University of Osnabrück).

Studies described in this manuscript as well as of those of a forthcoming paper which deals with the interaction of a defined oligonucleotide single strand, carrying nucleolipid head groups of different lipophilicity with lipid bilayer membranes, might be of importance for the optimization of the in vivo delivery of lipophilic siRNA [13] (e.g., by DNA trafficking as shown in Figure 17), as well as of lipid derivatives of nucleoside antimetabolites [23,24]. Moreover, studies of a transdermal application of lipo-oligonucleotides through the human Stratum corneum by iontophoresis techniques are underway.

# Experimental

## Materials

1-Palmitoyl-2-oleyl-*sn*-glycero-3-phosphoethanolamine (POPE) and 1-palmitoyl-2-oleyl-*sn*-glycero-3-phosphocholine (POPC) were purchased from Avanti Polar Lipids (Alabaster, Al), *n*-decane from Alfa Aesar (Karlsruhe, Germany), dimethyl sulfoxide (DMSO), potassium chloride (KCl), 3-morpholino-

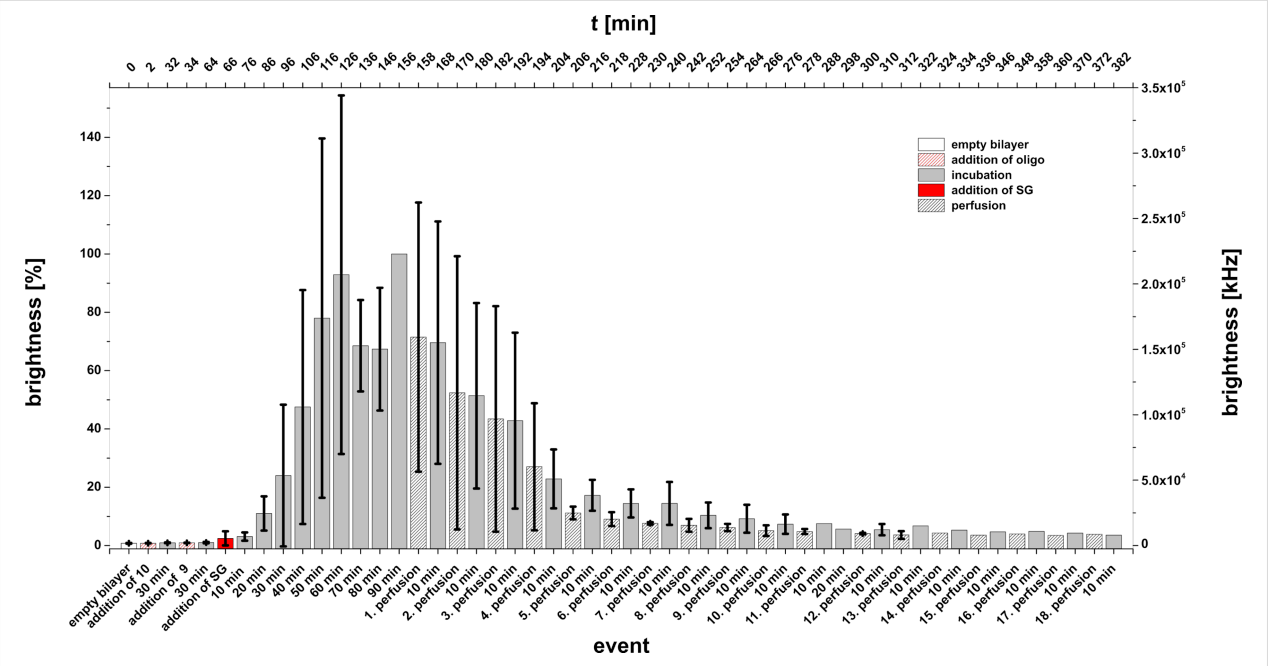



**Figure 15:** Bilayer brightness as a function of perfusion number and incubation periods for the experiment F (**SG** + **10** + **6**).

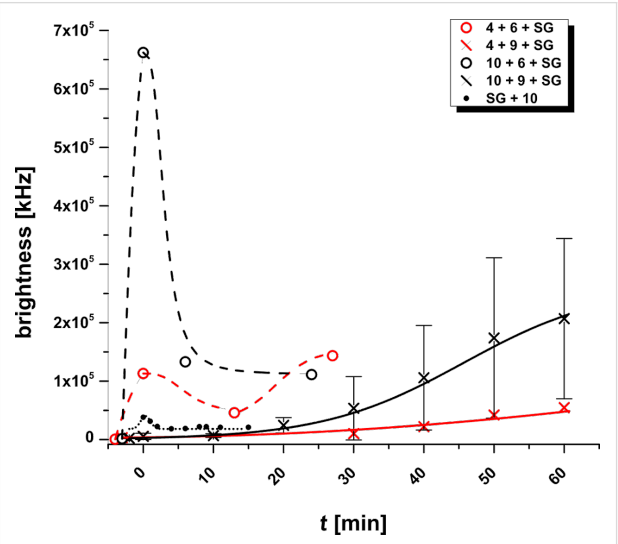



**Figure 16:** Kinetics of the tertiary complex formations of various lipo-DNA/DNA with Sybr Green I during incubation time *t* [min] after addition of the dye at $t_0$ until the first perfusion.

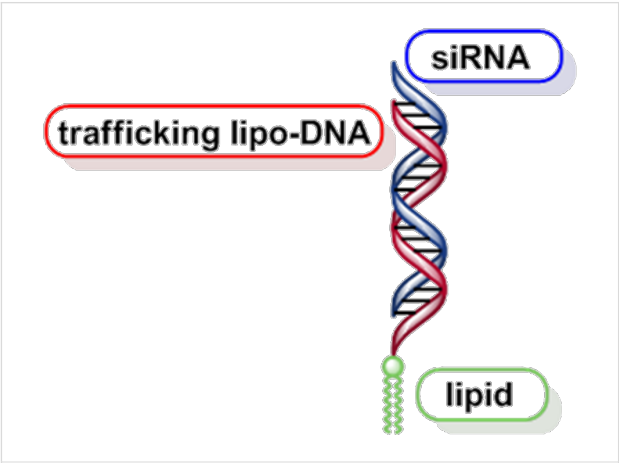



**Figure 17:** Trafficking of a siRNA by a lipophilized DNA.

propane-1-sulfonic acid (MOPS) and 2-amino-2-hydroxymethylpropane-1,3-diol (TRIS) were purchased from Roth (Karlsruhe, Germany).

All oligonucleotides **4–10** (Figure 1) were synthesized, purified and characterized by MALDI–TOF mass spectrometry by Eurogentec SA (Liege, Belgium). In each case the detected mass confirmed the corresponding calculated mass. MALDI–TOF–MS (*m*/*z*): 4247.7 (**4**, $[M + H]^+$; calcd 4247.4); 4178.4 (**5**, $[M + H]^+$; calcd 4178.1); 3645.5 (**6**, $[M + H]^+$; calcd 3644.5); 4248.1(**7**, $[M + H]^+$; calcd 4247.4); 3643.2 (**8**, $[M + H]^+$; calcd 3644.46); 7306.4 (**9**, $[M + H]^+$; calcd 7312.9); 6076.4 (**10**, $[M + H]^+$; calcd 6077.6).

**4** 5’-d(**1a**-p-TAG GTC AAT ACT)-3’

**5** 5’-d[(Cy5)-p-AGT ATT GAC CTA]-3’

**6** 5’-d(AGT ATT GAC CTA)-3’

**7** 5’-d(**1a**-p-ATC CAG TTA TGA)-3’

**8** 5'-d(TCA ATT GAC GAT)-3'

**9** 5'-d(TTT TAT AGT ATT GAC CTA TAT TTT)-3'

**10** 5'-d(**1a**-p-TTT TAT TAG GTC AAT ACT)-3'

We have chosen the particular sequences 5'-d(TAG GTC AAT ACT)-3' and 3'-d(ATC CAG TTA TGA)-5', because they do neither form hair pins nor self-complementary duplexes. The lipophilized strand and the complementary oligomer contain the same number of all canonical nucleotides, so that a mixture of 1 $A_{260}$ units, each of both strands gives an equimolar mixture. For nearly all base pairs nearly all nearest neighbour combinations exist. Both strands have the same composition; the unmodified duplex exhibits a melting point of 46 °C in a phosphate buffer solution.

A solution of Sybr Green I in DMSO with undisclosed concentration was purchased from Applichem (Darmstadt, Germany) and used as delivered. The concentration of the dye solution (≈1 μg/mL) was measured UV–vis spectrophotometrically using the published extinction coefficient of Sybr Green I at $\lambda_{max}$ of 494 nm (73.000 $M^{-1}cm^{-1}$) [20,21].

## Methods

**General.** A double-beam spectrophotometer (Specord 205, Analytik Jena GmbH, Jena, Germany) was used for spectrophotometric measurements. The UV–vis absorption spectra were carried out using ultra-micro quartz cuvettes type 105.202-QS from Hellma Analytics (Müllheim, Germany).

**Bilayer fabrication and incorporation of lipo-oligonucleotides therein.** Similar as described in [19], in the following the automated bilayer fabrication is described. Horizontal bilayers were fabricated automatically using a lipid mixture of 1-palmitoyl-2-oleyl-*sn*-glycero-3-phosphoethanolamine (POPE) and 1-palmitoyl-2-oleyl-*sn*-glycero-3-phosphocholine (POPC) (8:2, w/w, 10 mg/mL of *n*-decane) within the "Bilayer Slides" (Figure 18C,D) and an add-on for the inverted confocal microscope (Bilayer Slides and Ionovation Explorer, Ionovation GmbH, Osnabrück, Germany). After pre-filling with buffer (250 mM KCl, 10 mM MOPS/Tris, pH 7), the slide was inserted into the stage unit mounted on an inverted confocal microscope (Figure 18A). Ag/AgCl electrodes were mounted and after the addition of 0.2 μL of POPE/POPC lipid to the *cis* compartment using a 1 μL bended Hamilton syringe (CH-Bonaduz) (Figure 18E), the automated bilayer production was started; a modified painting technique, in which the air-water interface paints the lipid across the aperture was applied. The bilayer formation was monitored optically and electrically.

As soon as a stable bilayer has been established (C > 50 pF) the lipophilized oligonucleotide solution (8 μL, 500 nM) was injected into the *cis* compartment of the "Bilayer Slide". During the incubation time of 30–40 min the bilayer integrity was monitored by continuous capacitance measurements. After the incubation of the first sample the second sample of a complementary or non-complementary oligonucleotide was injected into the *cis* compartment and incubated for the next 30 minutes. Subsequently, the Sybr Green I solution (8 μL, undiluted) was injected and incubated until the highest brightness amount was reached. In all graphs the brightness was given either in kHz or it was normalized (0–100%). In the meantime every 10 minutes a z-scan was performed (see below). After the incubation steps the *cis* compartment was perfused repeatedly for 30 s (1.1 mL/min, each), and the bilayer was inspected by a confocal fluorescence microscope. Each perfusion step using the Ionovation perfusion unit (Figure 18B) was followed by an incubation step of 10 minutes. This procedure was repeated until the brightness of the bilayer finished decreasing continuously and turned out to be stable.

A confocal laser scanning microscope (Insight Cell 3D, Evotec Technologies GmbH, Hamburg, Germany) equipped with a 488 nm argon-ion laser, (Multiline, 2014-25MLYVW (max. total power 50 mW) from JDS Uniphase Corporation (Milpitas, CA,USA), as well as with a 635 nm emitting laser diode (LDH-P-635, PicoQuant, D-Berlin) und 470 nm (P-C-470B, PicoQuant, Berlin, Germany), a 40× water-immersion objective (UApo 340, 40×, NA = 1.15, Olympus, Tokyo, Japan), and an Avalanche photodiode detector (SPCM-AQR-13-FC, Perkin-Elmer Optoelectronics, Fremont, CA, USA) were used for the optical measurements. Fluorescence irradiation was obtained with an excitation laser power of 60 ± 5 μW (for 488 nm), 200 ± 20 μW (635 nm) and 50 ± 5 μW (470 nm) right in front of the objective. Z-scans were performed by scanning the confocal laser spot in XY direction with a rotating beam scanner and movement of the objective in Z direction. The movement in all directions is piezo-controlled, which allows a nanometer precise positioning.

In summary, each measuring protocol was carried out as follows: (I) a reference scan of the stable pure bilayer was performed; (II) then the lipophilized oligonucleotide sample was added and incubated for 30 min, (III) this step was followed by the addition of the complementary and non-complementary sample, respectively, and further incubation of 30 min; (IV) finally the cyanine-dye Sybr Green I solution (DMSO) was added to allow the quantification of double stranded DNA, and the solution was incubated until the brightness in the bilayer reached the highest amount; (V) afterwards additional scan series were performed after each perfusion of the *cis* compart-


**Figure 18:** A) Stage unit of the 'Ionovation Explorer' mounted on a standard inverted fluorescence microscope. B) The perfusion unit (on the left) during filling the syringes with buffer. C) Labelling of the single compartments and ports of the 'Bilayer Slide'. These are disposable, optically transparent microfluidic sample carriers with perfusion capabilities. D) Schematic demonstration of the *cis* chamber perfusion. E) Shot of the device from above during addition of the lipid mixture into the bilayer port of the slide. On the left, electrodes are visible, connected with the amplifier; on the right hand side tubes connected with the perfusion unit can be seen. F) Vertical cut of a buffer-filled bilayer slide, demonstrating its general design. The bilayer slide encloses two microfluidic channels (*cis* and *trans*) which are separated by a thin medical-grade PTFE (= polytetrafluoroethylene, Teflon) foil. This foil hosts a central 100 µm aperture which is located 120 µm above the cover slip and thus within the working distance of high NA (= numerical aperture) objectives. It is the only connection between the *trans* and the *cis* channels. When a lipid solution is painted across the aperture, a bilayer is formed spontaneously. The electrodes in the *cis* and *trans* channels allow an online monitoring of the bilayer integrity, as well as electrophysiological recordings.

ment (for 30 s; 1.1 mL/min) and (VI) followed by an incubation of 10 min.

In order to analyze the 2D images they were previously edited with Image JA 1.44 and the obtained data was evaluated with OriginPro 8 (OriginLab Corporation). For an evaluation of the z-scan data, the brightness within the bilayer was measured; the mean area of the bilayer cross section amounts to 353 ± 19 counts per pixel (N = 201). Next, the brightness was determined by summing up the number of pixels. 1 Pixel equals 1 kHz or 1 µm.

All devices (Figure 18) as well as the general techniques used in this paper have been described in detail in a preceding manuscript [19].

## Acknowledgements

The authors gratefully acknowledge financial support by the Ionovation GmbH, Osnabrück, Germany, as well as valuable discussions with Prof. Dr. Richard Wagner and Mr. Claudius Walter, University of Osnabrück, Department of Biophysics.

## References

1. Roskoski, R., Jr. *Biochem. Biophys. Res. Commun.* **2003,** *303,* 1–7. doi:10.1016/S0006-291X(03)00323-1
2. Dumelin, C. E.; Chen, Y.; Leconte, A. M.; Chen, Y. G.; Liu, D. R. *Nat. Chem. Biol.* **2012,** *8,* 913–919. doi:10.1038/NCHEMBIO.1070
3. Pokholenko, O.; Gissot, A.; Vialet, B.; Bathany, K.; Thiéry, A.; Barthélémy, P. *J. Mater. Chem. B* **2013,** *1,* 5329–5334. doi:10.1039/c3tb20357c
4. Gissot, A.; Camplo, M.; Grinstaff, M. W.; Barthélémy, P. *Org. Biomol. Chem.* **2008,** *6,* 1324–1333. doi:10.1039/b719280k

5. Langecker, M.; Arnaut, V.; List, J.; Simmel, F. C. *Acc. Chem. Res.* **2014,** *47,* 1807–1815. doi:10.1021/ar500051r
6. Taib, N.; Aimé, A.; Houmadi, S.; Castano, S.; Barthélémy, P.; Laguerre, M.; Bestel, I. *Langmuir* **2012,** *28,* 7452–7460. doi:10.1021/la300744x
7. Desbat, B.; Arazam, N.; Khiati, S.; Tonelli, G.; Neri, W.; Barthélémy, P.; Navailles, L. *Langmuir* **2012,** *28,* 6816–6825. doi:10.1021/la2047596
8. Loew, M.; Springer, R.; Scolari, S.; Altenbrunn, F.; Seitz, O.; Liebscher, J.; Huster, D.; Herrmann, A.; Arbuzova, A. *J. Am. Chem. Soc.* **2010,** *132,* 16066–16072. doi:10.1021/ja105714r
9. Johnson-Buck, A.; Jiang, S.; Yan, H.; Walter, N. G. *ACS Nano* **2014,** *8,* 5641–5649. doi:10.1021/nn500108k
10. Montis, C.; Baglioni, P.; Berti, D. *Soft Matter* **2014,** *10,* 39–43. doi:10.1039/c3sm52254g
11. Bunge, A.; Loew, M.; Pescador, P.; Arbuzova, A.; Brodersen, N.; Kang, J.; Dähne, L.; Liebscher, J.; Herrmann, A.; Stengel, G.; Huster, D. *J. Phys. Chem.* **2009,** *113,* 16425–16434. doi:10.1021/jp9067747
12. Peer, D.; Lieberman, J. *Gene Ther.* **2011,** *18,* 1127–1133. doi:10.1038/gt.2011.56
13. Wolfrum, C.; Shi, S.; Jayaprakash, K. N.; Jayaraman, M.; Wang, G.; Pandey, R. K.; Rajeev, K. G.; Nakayama, T.; Charrise, K.; Ndungo, E. M.; Zimmermann, T.; Koteliansky, V.; Manoharan, M.; Stoffel, M. *Nat. Biotechnol.* **2007,** *25,* 1149–1157. doi:10.1038/nbt1339
14. Kurz, A.; Bunge, A.; Windeck, A.-K.; Rost, M.; Flasche, W.; Arbuzova, A.; Strohbach, D.; Müller, S.; Liebscher, J.; Huster, D.; Herrmann, A. *Angew. Chem., Int. Ed.* **2006,** *45,* 4440–4444. doi:10.1002/anie.200600822
15. Honigmann, A.; Walter, C.; Erdmann, F.; Eggeling, C.; Wagner, R. *Biophys. J.* **2010,** *98,* 2886–2894. doi:10.1016/j.bpj.2010.03.033
16. Rosemeyer, H. Nucleolipids and use thereof, and devices for nucleic acid analysis. U.S. Patent 7,914,991 B2, March 29, 2011.
17. Rosemeyer, H. Nucleolipide und deren Verwendung, sowie Vorrichtung zur Nucleinsäureanalytik. Eur. Patent 1,893,773 B1, Jan 23, 2013.
18. Rode, A. B.; Endoh, T.; Tateishi-Karimata, H.; Takahashi, S.; Sugimoto, N. *Chem. Commun.* **2013,** *49,* 8444–8446. doi:10.1039/c3cc42990c
19. Werz, E.; Korneev, S.; Montilla-Martinez, M.; Wagner, R.; Hemmler, R.; Walter, C.; Eisfeld, J.; Gall, K.; Rosemeyer, H. *Chem. Biodiversity* **2012,** *9,* 272–281. doi:10.1002/cbdv.201100298
20. Zipper, H.; Brunner, H.; Bernhagen, J.; Vitzthum, F. *Nucleic Acids Res.* **2004,** *32,* e103. doi:10.1093/nar/gnh101
21. Zipper, H. Entwicklung analytisch-molekularbiologischer Verfahren zur Konstruktion einer Plasmid-Gendatenbank aus Boden-DNA in *Escherichia coli* und deren Durchmusterung nach neuen Enzymen für die technische Anwendung. Ph.D. Thesis, University of Stuttgart, Stuttgart, Germany, 2004; pp 17–30.
22. Dragan, A. I.; Pavlovic, R.; McGivney, J. B.; Casas-Finet, J. R.; Bishop, E. S.; Strouse, R. J.; Schenerman, M. A.; Geddes, C. D. *J. Fluoresc.* **2012,** *22,* 1189–1199. doi:10.1007/s10895-012-1059-8
23. Malecki, E.; Farhat, A.; Bonaterra, G. A.; Röthlein, D.; Wolf, M.; Schmitt, J.; Kinscherf, R.; Rosemeyer, H. *Chem. Biodiversity* **2013,** *10,* 2235–2246. doi:10.1002/cbdv.201300219
24. Farhat, A.; Malecki, E.; Bonaterra, G. A.; Röthlein, D.; Wolf, M.; Schmitt, J.; Rosemeyer, H.; Kinscherf, R. *Chem. Biodiversity* **2014,** *11,* 469–482. doi:10.1002/cbdv.201300347


## License and Terms

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

# Versatile synthesis of amino acid functionalized nucleosides via a domino carboxamidation reaction

Vicky Gheerardijn, Jos Van den Begin and Annemieke Madder*

Full Research Paper

Open Access

Address:
Department of Organic and Macromolecular Chemistry, Organic and Biomimetic Chemistry Research Group, Ghent University, Krijgslaan 281 S4, 9000 Ghent, Belgium

Email:
Annemieke Madder* - Annemieke.Madder@UGent.be

* Corresponding author

Keywords:
amino acids; bioorganic chemistry; carboxamidation reaction; chemically modified; DNA nucleosides; nucleic acids

*Beilstein J. Org. Chem.* **2014,** *10,* 2566–2572.
doi:10.3762/bjoc.10.268

Received: 18 July 2014
Accepted: 23 October 2014
Published: 04 November 2014

This article is part of the Thematic Series "Nucleic acid chemistry".

Guest Editor: H.-A. Wagenknecht



## Abstract

Functionalized oligonucleotides have recently gained increased attention for incorporation in modified nucleic acid structures both for the design of aptamers with enhanced binding properties as well as the construction of catalytic DNA and RNA. As a shortcut alternative to the incorporation of multiple modified residues, each bearing one extra functional group, we present here a straightforward method for direct linking of functionalized amino acids to the nucleoside base, thus equipping the nucleoside with two extra functionalities at once. As a proof of principle, we have introduced three amino acids with functional groups frequently used as key-intermediates in DNA- and RNAzymes via an efficient and straightforward domino carboxamidation reaction.

## Introduction

For decades DNA has been known as the carrier of the genetic information. Only recently, the use of synthetic oligonucleotides and their modified analogues for a range of therapeutic and diagnostic purposes [1], including antisense therapy [2,3], antigene therapy [4,5] and SNP (Single Nucleotide Polymorphism) detection [6] has gained major interest. Due to their predictable and well-investigated structure, the strength of nucleic acids for a series of applications such as DNA and RNA based drugs [7], drug delivery systems [8,9], DNA-biosensors [10,11] and potential catalysts has now firmly been recognized. Rather than using unmodified oligonucleotides, providing additional functional groups can lead to even higher activities and selectivities. Indeed, people have realized that equipping the nucleic acid scaffold with protein side chain-like moieties may assist in the design of a broad scope of functionalized oligonucleotides with various characteristics. This can be enabled by techniques such as solid phase synthesis, post synthetic modifications or enzymatic incorporation of modified analogues.

Originating from research into aptamers as strong and selective binders [12-14], several research groups are investigating the creation and synthesis of new DNA or RNA catalysts, also called DNAzymes [15-17] and RNAzymes [10,16,18-21], respectively. Most catalytic nucleic acids are generated using the

SELEX (Systematic Evolution of Ligands by EXponential enrichment) protocol [22-25] that generates a library of potential catalysts, which are then screened for catalytic potential using natural substrates or transition state analogues. Although examples [26-29] exist illustrating the introduction of modified residues and by that the development of new catalysts, the SELEX protocol becomes more labor intensive and reproducibility of results is often a problem. Moreover achieving control in terms of number, positioning and exact location of the desired catalytic moieties is far from straightforward. In contrast to the existing variety of rather complicated and unpredictable RNA based ribozyme-like structures, oligonucleotide duplexes have a stable and predictable structure allowing the design of engineered active sites through the carefully planned introduction of extra catalytic functionalities via solid phase DNA synthesis. We have further recently shown that the combination of solid phase synthesis and molecular modeling combined with advanced NMR techniques offers the possibility to predict and control the positioning of catalytic functions within a DNA duplex [30].

Chemical modification of nucleic acids can be performed on different positions, including the backbone, the sugar unit and the heterocyclic base, whereby base modification is the most common as it causes only minor disturbance in the helical structure [31]. Depending on the position of the incorporation on the nucleoside structure, the introduced functionalities can be pointed towards the major or the minor groove of the duplex. Modification of position 1' and 2' will orient substituents towards the minor groove while positions 5, 7 and 8 on the base allow directing incorporated substituents in the major groove. One can further imagine that introduction of extra functionalities can have an impact on the duplex stability and final structure of the double helix. This was illustrated in previous research in our laboratory where introduction of modified nucleosides on the 2'-position in a DNA double helix resulted in a destabilization of the duplex of 5 °C per modified unit [32]. Although this destabilization depends on many different aspects, such as the type and length of the linker and the position of the modification, it is important to minimize destabilization as much as possible, more specifically when introduction of more than one modification is desired. To date most research groups have focused on the introduction of only one extra functionality per nucleoside, mainly for synthetic reasons, and preferably making use of commercially available building blocks. We therefore considered the incorporation of moieties containing multiple functionalities on a nucleoside residue and present a straightforward method for the introduction of side chain functionalized amino acids onto nucleoside building blocks (Figure 1). Following the event of solid phase peptide synthesis, a large range of amino acids with different protection schemes and stereochemistry is currently commercially available. We here illustrate a methodology for direct incorporation of amino acids via a short synthetic pathway offering the added benefit of introducing multiple functional groups at the same time.

Indeed, whereas the α-amino group is used to ensure linkage to the nucleoside building block, both the side chain functional group and the α-carboxylate end up as extra functionalities on the nucleoside. As proof of principle and with the catalytic serine–histidine–aspartate catalytic triad of serine proteases in mind, we illustrate the preparation of three different nucleosides equipped with three functional amino acids containing hydroxy, imidazole, amine and carboxylate groups for later incorporation into nucleic acid duplexes via solid phase DNA synthesis [33].

## Results and Discussion

As mentioned earlier, a number of research groups have used the SELEX or in vitro selection protocol for the incorporation of amino acid-like modifications into DNA or RNA where there

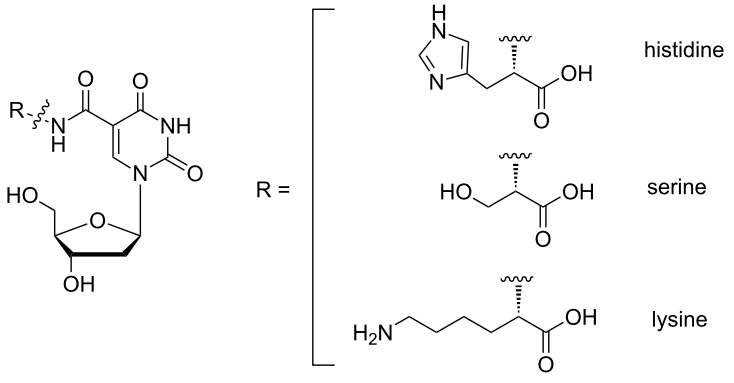
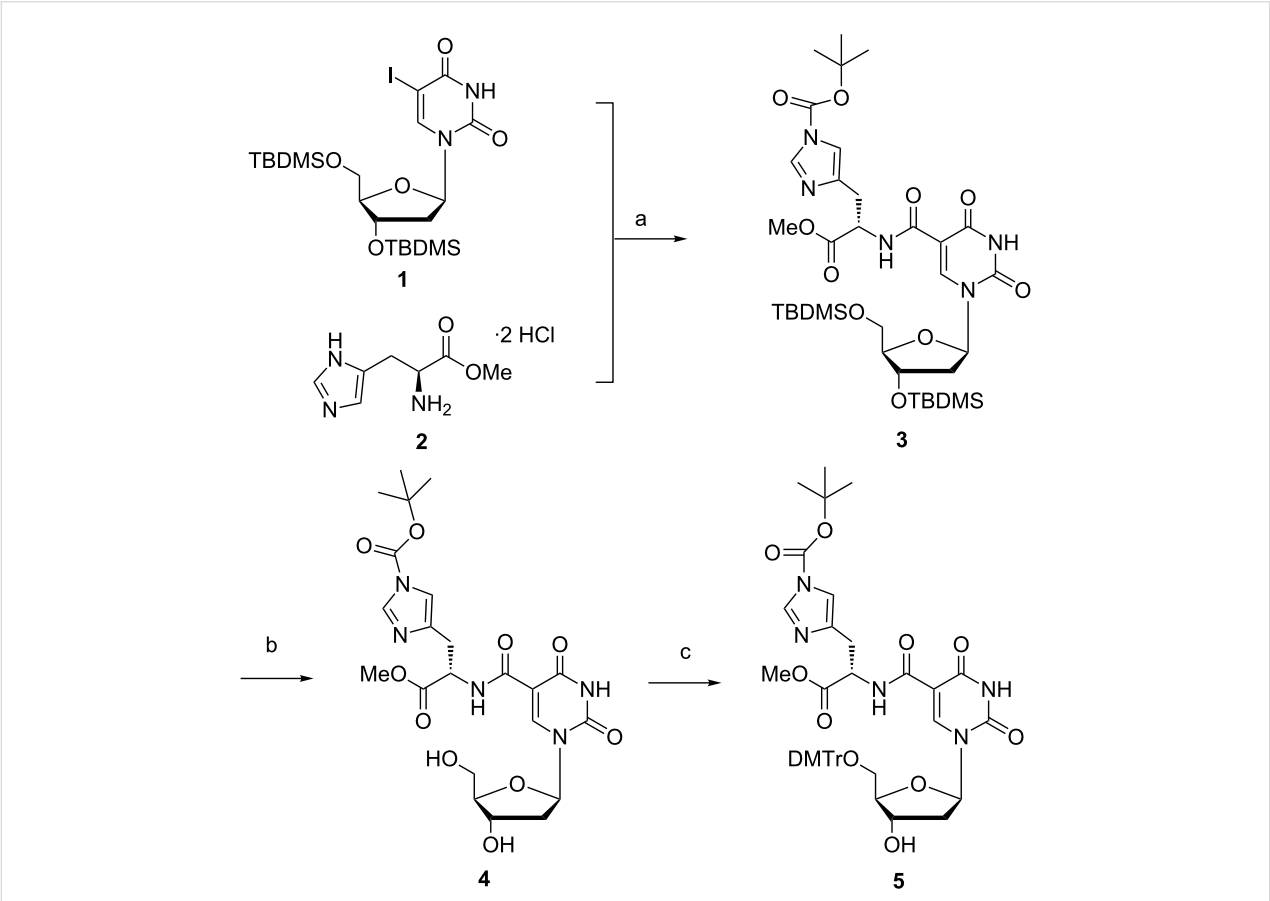



**Figure 1:** Amino acid functionalized nucleosides.

is no need for protection of the functional groups. For example, many imidazole and amine functionalities were incorporated to increase the structural diversity and catalytic potential of nucleic acids [34-37]. Most of these functionalities are introduced into the triphosphate building blocks using palladium-based coupling chemistry to the pyrimidine C5 or the 7-position of 7-deaza-2'-deoxyadenosine [38-40]. During enzymatic incorporation the extra functionalities do not need to be protected. However, solid phase DNA synthesis implies an appropriate protection of the extra functionalities on the corresponding nucleoside building blocks to avoid side-reactions during DNA synthesis. Not only should these protecting groups be stable under all chemical conditions used in the synthesis pathway, they should also be stable under the DNA synthesis conditions and easily removable after assembly of the desired chain without DNA damage.

In the current study, we have chosen to couple three amino acids, which contain functional groups commonly used in SELEX approaches to modify DNA or RNA, to 5-iodo-2'-deoxyuridine. We describe the direct and linker-less introduction of histidine, serine and lysine derivatives onto nucleosides via a straightforward and easy domino carboxamidation reaction. Previously the groups of Gait and Eaton [41-43] have used this reaction to couple histamine or simple amine derivatives to both 5-iodo-2'-deoxyuridine and purine nucleosides. Although a large number of imidazole modified pyrimidine and purine derivatives for solid phase synthesis have been described to date [41,44-47], we believe that the reactions described here serve as an ideal model system, which can be extended to other commercially available amino acid derivatives and nucleosides.

For the introduction of histidine onto the nucleoside as shown in Scheme 1, we chose to protect the free 3' and 5'-hydroxy groups with *tert*-butyldimethylsilyl (TBDMS) groups **1** to avoid side reactions during the next step, the carboxamidation reaction [48]. While protected histidine is commercially available, we have synthesized histidine as a methyl ester (**2**) with thionyl chloride in methanol in good yield according to literature procedures [49]. Although basic hydrolysis of alkyl esters has been shown to imply long reaction times, side-product formation and low reaction yields [50], we found that methyl esters within

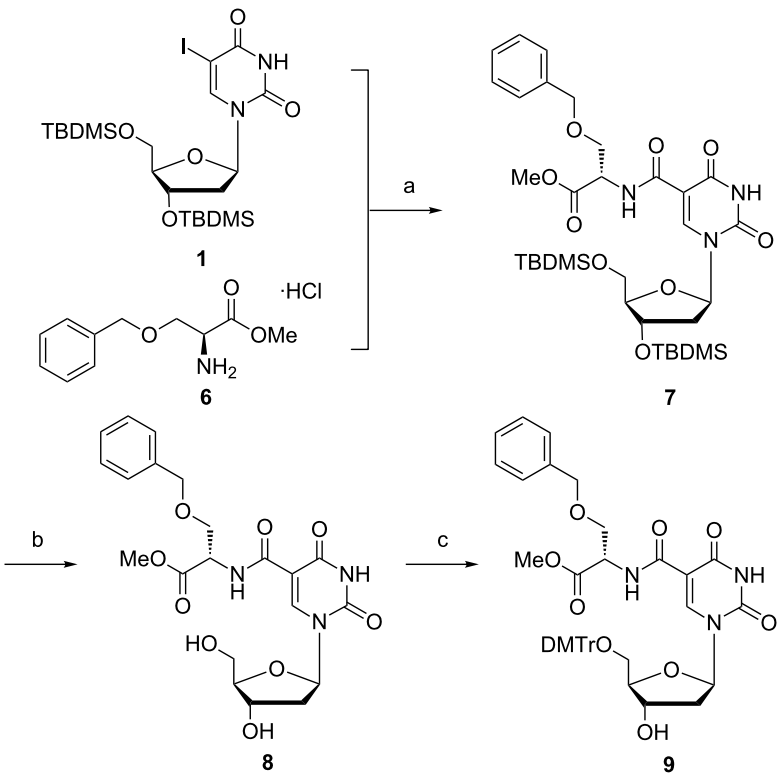



**Scheme 1:** Reagents and conditions: a) i. $Et_3N$, $Pd(PPh_3)_4$, THF, CO (4 bar), 70 °C, 48 h, ii. $Et_3N$, di-*tert*-butyl dicarbonate, DMF, rt, 15 min (79%); b) $Et_3N{\cdot}3HF$, $Et_3N$, THF, rt, overnight (72%); c) DMTr-Cl, pyridine, 0 °C, 7 h (52%).

DNA are hydrolysed under basic conditions even at room temperature using aqueous ammonia provided prolonged reaction times are respected [51]. Inspired by the results of Perrin and Joyce [26,52], who introduced a nucleoside containing an unprotected imidazole functionality into the DNA synthesis cycle without problems, we first tried to leave the imidazole unprotected in order to avoid extra protection and deprotection steps in the reaction sequence. However, in view of problems arising during the separation by column chromatography after the CO insertion reaction with an unprotected histidine methyl ester, we decided to protect the imidazole functionality immediately after the reaction with the *tert*-butyloxycarbonyl (*t*-Boc) protecting group [53], which is compatible with the reagents used during nucleoside- and oligonucleotide synthesis [45,54]. While the *t*-Boc group can be removed with a 10% TFA solution [54], the acid labile DNA can suffer from depurination after treatment with acid. According to literature, the *t*-Boc group can also be cleaved during standard deprotection procedures with saturated ammonia after oligonucleotide synthesis [41,55]. In view of all these considerations, coupling of the TBDMS-protected nucleoside **1** with histidine methyl ester **2** followed by in situ *t*-Boc protection of the free hydrogen on the imidazole functionality gave the desired compound **3** in satisfying yield.

In the next step of the synthesis, the TBDMS-protecting groups needed to be removed. First attempts using *tetra*-butylammonium fluoride (TBAF) resulted in degradation of the starting material (results not shown). Indeed, several reports in literature highlight the strong basic character of the fluoride anion ($F^-$) that can cause decomposition of or nucleophilic attack on the nucleoside [56,57]. Manfredini et al. have shown that selective deprotection of the TBDMS groups in the presence of a *t*-Boc protecting group is possible when using $Et_3N{\cdot}3HF$ [58] and also in our case using $Et_3N{\cdot}3HF$ in THF resulted in product **4** with good purity and yield (72%) [59]. Dimethoxytritylation could be accomplished under standard conditions and the corresponding derivative **5** was isolated and purified without problems.

As a second example we decided to employ serine as amino acid for introduction onto the nucleoside. In view of the earlier illustrated need for imidazole protection during introduction of histidine (vide supra), we used commercially available benzyl protected serine **6** to avoid problems due to the free hydroxy group of the serine during the carboxamidation reaction (Scheme 2). After DNA synthesis, the benzyl group can be cleaved using hydrogen and palladium on carbon. To avoid overreduction of the pyrimidine double bond, transfer hydro-



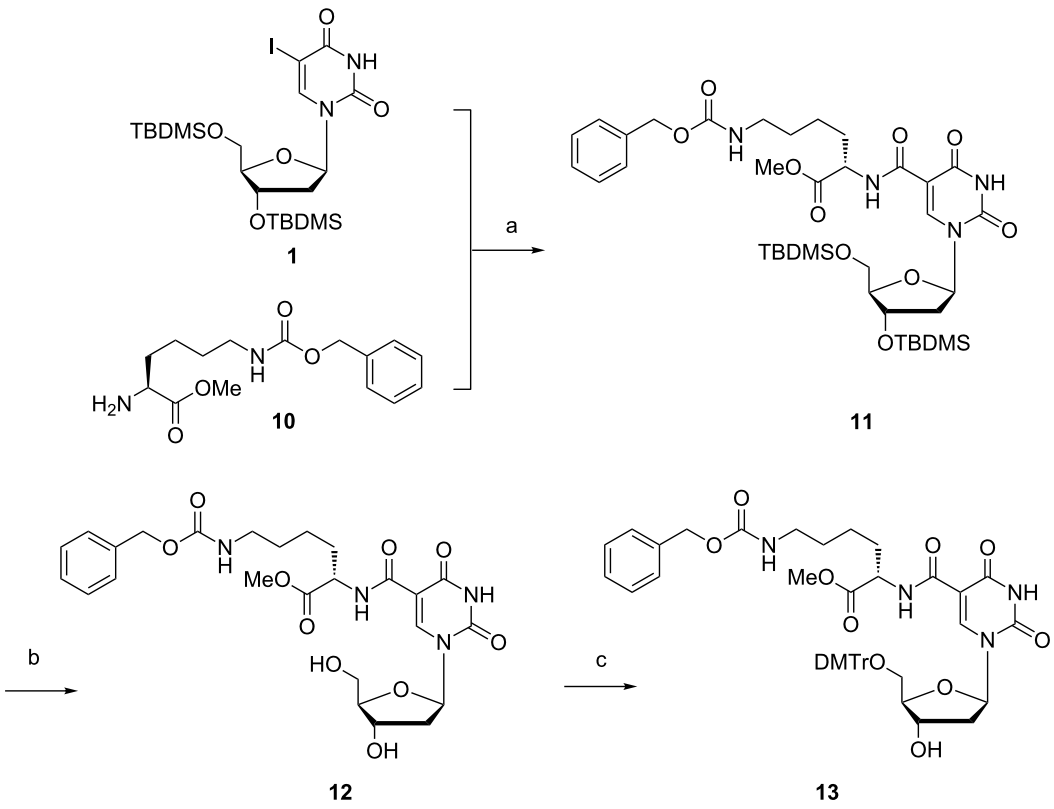

**Scheme 2:** Reagents and conditions: a) $Et_3N$, $Pd(PPh_3)_4$, THF, CO (4 bar), 48 h, 70 °C (68%); b) $Et_3N{\cdot}3HF$, $Et_3N$, THF, overnight, rt (86%); c) DMTr-Cl, pyridine, DCM, overnight, rt (82%).

genation with cyclohexene as hydrogen source and 10% palladium on carbon should be used [60-62]. As the methyl ester derivative of the amino acid is commercially available but rather expensive, we performed the esterification reaction on the benzyl protected serine to deliver product **6** in good yield and purity after recrystallization in cold diethyl ether (64%) [63]. Coupling of protected serine methyl ester **6** with TBDMS-protected nucleoside **1** in the presence of tetrakis(triphenylphosphine)palladium(0) and $Et_3N$ in THF under CO atmosphere (4 bar) gave the desired product **7**. After deprotection of the TBDMS groups to yield **8** under similar reaction conditions as used before to obtain product **4**, the free 5'-hydroxy group was selectively protected with DMTr-Cl to obtain the modified desired nucleoside **9** in good yield and purity.

As the functional groups of histidine (imidazole) and lysine (cationic amine) are known to be present in numerous enzyme active sites and capable of general acid–general base catalysis, we have finally chosen to modify a nucleoside with a lysine amino acid derivative as a third example of our current strategy, as shown in Scheme 3. Protected lysine **10** is commercially available and was coupled with the TBDMS-protected nucleoside **1** to afford product **11** in high yield. After deprotection of the *tert*-butyldimethylsilyl protecting groups using the above described protocol, modified nucleoside **12** could be selectively protected at the free 5'-hydroxy group with DMTr-Cl to afford **13**. Compared to the previous reaction schemes, this final reaction suffers more from the steric hindrance as compared to the other amino acids that show less bulky substituents.

## Conclusion

We here illustrated the efficient application of a simple carboxamidation procedure for the synthesis of modified nucleosides featuring two extra functional groups. Because hydrolysis of the methyl ester and removal of the *t*-Boc is performed during the standard cleavage of solid support using $NH_4OH$ and benzyl and benzoyl protecting groups can be removed by hydrogenation [64], only one extra deprotection step is needed after incorporation. After transformation of the modified nucleosides into the corresponding phosphoramidite building blocks **5**, **9** and **13** are amenable to incorporation in nucleic acids through standard DNA synthesis methodology. The described protocol relies on commercially available amino acid derivatives and should be applicable to a large series of natural and unnatural amino acids.




**Scheme 3:** Reagents and conditions: a) $Et_3N$, $Pd(PPh_3)_4$, THF, CO (4 bar), 48 h, 70 °C (90%); b) $Et_3N{\cdot}3HF$, $Et_3N$, THF, overnight, rt (43%); c) DMTr-Cl, pyridine, DCM, overnight, rt (9%).

## Supporting Information

### Supporting Information File 1

Experimental procedures, characterization data, and $^{1}H$ and $^{13}C$ NMR spectra of new compounds.
[http://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-10-268-S1.pdf]

## Acknowledgements

VG is indebted to UGent for a position as research assistant. JVDB is indebted to UGent for a position as technical assistant. Further support from the FWO-Flanders (G.0422.13) is gratefully acknowledged.

## References

1. Patil, S. D.; Rhodes, D. G.; Burgess, D. J. *AAPS J.* **2005,** *7,* E61–E77. doi:10.1208/aapsj070109
2. Uhlmann, E.; Peyman, A. *Chem. Rev.* **1990,** *90,* 543–584. doi:10.1021/cr00102a001
3. Crooke, S. T. *Annu. Rev. Med.* **2004,** *55,* 61–95. doi:10.1146/annurev.med.55.091902.104408
4. Buchini, S.; Leumann, C. J. *Curr. Opin. Chem. Biol.* **2003,** *7,* 717–726. doi:10.1016/j.cbpa.2003.10.007
5. Stull, R. A.; Szoka, F. C., Jr. *Pharm. Res.* **1995,** *12,* 465–483. doi:10.1023/A:1016281324761
6. Gray, I. C.; Campbell, D. A.; Spurr, N. K. *Hum. Mol. Genet.* **2000,** *9,* 2403–2408. doi:10.1093/hmg/9.16.2403
7. Radomska-Leśniewska, D.; Hevelke, A.; Jankowska-Steifer, E.; Malejczyk, J. *Cent. Eur. J. Immunol.* **2009,** *34,* 280–283.
8. Luo, D.; Saltzman, W. M. *Nat. Biotechnol.* **2000,** *18,* 33–37. doi:10.1038/71889
9. Shen, W. DNA Nanotechnology and Drug Delivery. In *DNA Nanotechnology: From Structure to Function;* Fan, C., Ed.; Springer: Heidelberg, New York, 2013; pp 321–334. doi:10.1007/978-3-642-36077-0_15
10. Zhai, J.; Cui, H.; Yang, R. F. *Biotechnol. Adv.* **1997,** *15,* 43–58. doi:10.1016/S0734-9750(97)00003-7
11. Teles, F. R. R.; Fonseca, L. P. *Talanta* **2008,** *77,* 606–623. doi:10.1016/j.talanta.2008.07.024
12. Jayasena, S. D. *Clin. Chem.* **1999,** *45,* 1628–1650.
13. Šmuc, T.; Ahn, I.-Y.; Ulrich, H. *J. Pharm. Biomed. Anal.* **2013,** *81–82,* 210–217. doi:10.1016/j.jpba.2013.03.014
14. Famulok, M.; Hartig, J. S.; Mayer, G. *Chem. Rev.* **2007,** *107,* 3715–3743. doi:10.1021/cr0306743
15. Santoro, S. W.; Joyce, G. F. *Biochemistry* **1998,** *37,* 13330–13342. doi:10.1021/bi9812221
16. Fiammengo, R.; Jäschke, A. *Curr. Opin. Biotechnol.* **2005,** *16,* 614–621. doi:10.1016/j.copbio.2005.10.006
17. Breaker, R. R.; Joyce, G. F. *Chem. Biol.* **1994,** *1,* 223–229. doi:10.1016/1074-5521(94)90014-0
18. Chapman, K. B.; Szostak, J. W. *Curr. Opin. Struct. Biol.* **1994,** *4,* 618–622. doi:10.1016/S0959-440X(94)90227-5
19. Famulok, M.; Jenne, A. *Top. Curr. Chem.* **1999,** *202,* 101–131. doi:10.1007/3-540-48990-8_4
20. Wilson, D. S.; Szostak, J. W. *Annu. Rev. Biochem.* **1999,** *68,* 611–647. doi:10.1146/annurev.biochem.68.1.611
21. Guerrier-Takada, C.; Gardiner, K.; Marsh, T.; Pace, N.; Altman, S. *Cell* **1983,** *35,* 849–857. doi:10.1016/0092-8674(83)90117-4
22. Ellington, A. D.; Szostak, J. W. *Nature* **1990,** *346,* 818–822. doi:10.1038/346818a0
23. Tuerk, C.; Gold, L. *Science* **1990,** *249,* 505–510. doi:10.1126/science.2200121
24. Klug, S. J.; Famulok, M. *Mol. Biol. Rep.* **1994,** *20,* 97–107. doi:10.1007/BF00996358
25. Joyce, G. F. *Annu. Rev. Biochem.* **2004,** *73,* 791–836. doi:10.1146/annurev.biochem.73.011303.073717
26. Santoro, S. W.; Joyce, G. F.; Sakthivel, K.; Gramatikova, S.; Barbas, C. F., III. *J. Am. Chem. Soc.* **2000,** *122,* 2433–2439. doi:10.1021/ja993688s
27. Perrin, D. M.; Garestier, T.; Hélène, C. *J. Am. Chem. Soc.* **2001,** *123,* 1556–1563. doi:10.1021/ja003290s
28. Sidorov, A. V.; Grasby, J. A.; Williams, D. M. *Nucleic Acids Res.* **2004,** *32,* 1591–1601. doi:10.1093/nar/gkh326
29. Kuwahara, M.; Nagashima, J.; Hasegawa, M.; Tamura, T.; Kitagata, R.; Hanawa, K.; Hososhima, S.; Kasamatsu, T.; Ozaki, H.; Sawai, H. *Nucleic Acids Res.* **2006,** *34,* 5383–5394. doi:10.1093/nar/gkl637
30. Buyst, D.; Gheerardijn, V.; Féher, K.; Van Gasse, B.; Van den Begin, J.; Martins, J. C.; Madder, A. *Nucleic Acids Res.* submitted.
31. Blackburn, G. M.; Gait, M. J.; Loakes, D.; Williams, D. M. *Nucleic Acids in Chemistry and Biology;* Royal Society of Chemistry: Cambridge, UK, 2006. doi:10.1039/9781847555380
32. Catry, M. A.; Gheerardijn, V.; Madder, A. *Bioorg. Chem.* **2010,** *38,* 92–97. doi:10.1016/j.bioorg.2010.01.002
33. Beaucage, S. L.; Caruthers, M. H. *Tetrahedron Lett.* **1981,** *22,* 1859–1862. doi:10.1016/S0040-4039(01)90461-7
34. Lee, S. E.; Sidorov, A.; Gourlain, T.; Mignet, N.; Thorpe, S. J.; Brazier, J. A.; Dickman, M. J.; Hornby, D. P.; Grasby, J. A.; Williams, D. M. *Nucleic Acids Res.* **2001,** *29,* 1565–1573. doi:10.1093/nar/29.7.1565
35. Vaught, J. D.; Bock, C.; Carter, J.; Fitzwater, T.; Otis, M.; Schneider, D.; Rolando, J.; Waugh, S.; Wilcox, S. K.; Eaton, B. E. *J. Am. Chem. Soc.* **2010,** *132,* 4141–4151. doi:10.1021/ja908035g
36. Hollenstein, M.; Hipolito, C. J.; Lam, C. H.; Perrin, D. M. *ACS Comb. Sci.* **2013,** *15,* 174–182. doi:10.1021/co3001378
37. Hipolito, C. J.; Hollenstein, M.; Lam, C. H.; Perrin, D. M. *Org. Biomol. Chem.* **2011,** *9,* 2266–2273. doi:10.1039/c0ob00595a
38. Matulic-Adamic, J.; Daniher, A. T.; Karpeisky, A.; Haeberli, P.; Sweedler, D.; Beigelman, L. *Bioorg. Med. Chem. Lett.* **2000,** *10,* 1299–1302. doi:10.1016/S0960-894X(00)00226-2
39. Čapek, P.; Cahová, H.; Pohl, R.; Hocek, M.; Gloeckner, C.; Marx, A. *Chem. – Eur. J.* **2007,** *13,* 6196–6203. doi:10.1002/chem.200700220
40. Jäger, S.; Rasched, G.; Kornreich-Leshem, H.; Engeser, M.; Thum, O.; Famulok, M. *J. Am. Chem. Soc.* **2005,** *127,* 15071–15082. doi:10.1021/ja051725b
41. Holmes, S. C.; Gait, M. J. *Eur. J. Org. Chem.* **2005,** 5171–5183. doi:10.1002/ejoc.200500413
42. Dewey, T. M.; Zyzniewski, M. C.; Eaton, B. E. *Nucleosides Nucleotides* **1996,** *15,* 1611–1617. doi:10.1080/07328319608002461
43. Tu, C.; Keane, C.; Eaton, B. *Nucleosides Nucleotides* **1997,** *16,* 227–237. doi:10.1080/07328319708001344
44. Xu, L.; Ji, C.; Bai, Y.; He, J.; Liu, K. *Biochem. Biophys. Res. Commun.* **2013,** *434,* 516–520. doi:10.1016/j.bbrc.2013.03.106
45. Bashkin, J. K.; Gard, J. K.; Modak, A. S. *J. Org. Chem.* **1990,** *55,* 5125–5132. doi:10.1021/jo00304a026

46. Ramasamy, K. S.; Zounes, M.; Gonzalez, C.; Freier, S. M.; Lesnik, E. A.; Cummins, L. L.; Griffey, R. H.; Monia, B. P.; Cook, P. D. *Tetrahedron Lett.* **1994,** *35,* 215–218. doi:10.1016/S0040-4039(00)76514-2
47. Prakash, T. P.; Kumar, R. K.; Ganesh, K. N. *Tetrahedron* **1993,** *49,* 4035–4050. doi:10.1016/S0040-4020(01)89916-6
48. Ogilvie, K. K. *Can. J. Chem.* **1973,** *51,* 3799–3807. doi:10.1139/v73-569
49. Abdo, M.-R.; Joseph, P.; Boigegrain, R.-A.; Liautard, J.-P.; Montero, J.-L.; Köhler, S.; Winum, J.-Y. *Bioorg. Med. Chem.* **2007,** *15,* 4427–4433. doi:10.1016/j.bmc.2007.04.027
50. Berthod, T.; Pétillot, Y.; Guy, A.; Cadet, J.; Molko, D. *J. Org. Chem.* **1996,** *61,* 6075–6078. doi:10.1021/jo960614f
51. Gheerardijn, V., Van den Begin, J., Madder, A. *unpublished results.*
52. Lermer, L.; Hobbs, J.; Perrin, D. M. *Nucleosides, Nucleotides Nucleic Acids* **2002,** *21,* 651–664. doi:10.1081/NCN-120015723
53. Agami, C.; Couty, F. *Tetrahedron* **2002,** *58,* 2701–2724. doi:10.1016/S0040-4020(02)00131-X
54. Bashkin, J. K.; McBeath, R. J.; Modak, A. S.; Sample, K. R.; Wise, W. B. *J. Org. Chem.* **1991,** *56,* 3168–3176. doi:10.1021/jo00009a044
55. Verbeure, B.; Lacey, C. J.; Froeyen, M.; Rozenski, J.; Herdewijn, P. *Bioconjugate Chem.* **2002,** *13,* 333–350. doi:10.1021/bc0155622
56. Westman, E.; Stromberg, R. *Nucleic Acids Res.* **1994,** *22,* 2430–2431. doi:10.1093/nar/22.12.2430
57. Clark, J. H. *Chem. Rev.* **1980,** *80,* 429–452. doi:10.1021/cr60327a004
58. Bazzanini, R.; Gouy, M.-H.; Peyrottes, S.; Gosselin, G.; Périgaud, C.; Manfredini, S. *Nucleosides, Nucleotides Nucleic Acids* **2005,** *24,* 1635–1649. doi:10.1080/15257770500267006
59. Pirrung, M. C.; Shuey, S. W.; Lever, D. C.; Fallon, L. *Bioorg. Med. Chem. Lett.* **1994,** *4,* 1345–1346. doi:10.1016/S0960-894X(01)80358-9
60. Felix, A. M.; Heimer, E. P.; Lambros, T. J.; Tzougraki, C.; Meienhofer, J. *J. Org. Chem.* **1978,** *43,* 4194–4196. doi:10.1021/jo00415a045
61. Anantharamaiah, G. M.; Sivanandaiah, K. M. *J. Chem. Soc., Perkin Trans. 1* **1977,** 490–491. doi:10.1039/p19770000490
62. Watkins, B. E.; Kiely, J. S.; Rapoport, H. *J. Am. Chem. Soc.* **1982,** *104,* 5702–5708. doi:10.1021/ja00385a026
63. Kantharaju; Babu, V. V. S. *Indian J. Chem., Sect. B: Org. Chem. Incl. Med. Chem.* **2006,** *45,* 1942–1944.
64. Catry, M. A.; Madder, A. *Molecules* **2007,** *12,* 114–129. doi:10.3390/12010114

## License and Terms

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

# Come-back of phenanthridine and phenanthridinium derivatives in the 21st century

Lidija-Marija Tumir, Marijana Radić Stojković and Ivo Piantanida*

Review

Open Access

Address:
Laboratory for Study of Interactions of Biomacromolecules, Division of Organic Chemistry and Biochemistry, Ruđer Bošković Institute, Bijenička cesta 54, PO Box 180, HR-10002 Zagreb, Croatia

Email:
Ivo Piantanida* - pianta@irb.hr

* Corresponding author

Keywords:
ds-DNA and ds-RNA binding; intercalation; minor groove binding; nucleic acids; organic synthesis; phenanthridine; phenanthridinium

*Beilstein J. Org. Chem.* **2014,** *10,* 2930–2954.
doi:10.3762/bjoc.10.312

Received: 17 July 2014
Accepted: 21 November 2014
Published: 10 December 2014

This article is part of the Thematic Series "Nucleic acid chemistry".

Guest Editor: H.-A. Wagenknecht



## Abstract

Phenanthridine derivatives are one of the most intensively studied families of biologically active compounds with efficient DNA binding capability. Attracting attention since DNA structure discovery (1960s), they were early recognized as a symbol of DNA intercalative binding, for many decades applied as gold-standard DNA- and RNA-fluorescent markers (ethidium bromide), probes for cell viability (propidium iodide), but also “ill-famed” for various toxic (genotoxic) and mutagenic effects. After two decades of low interest, the discovery of phenanthridine alkaloids and new studies of antiparasitic/antitumor properties of phenanthridine derivatives resulted in the strong increase of the scientific interest about the turn of this century. Here are summarized phenanthridine-related advances in the 21st century (2000-present period) with emphasis on the supramolecular interactions and bioorganic chemistry, as well as novel or improved synthetic approaches.

## Introduction

The search for therapeutic agents of the phenanthridine type has increased when the outstanding trypanocidal activity of some phenanthridinium compounds became apparent [1]. One of the most studied and used phenanthridine derivatives is 3,8-diamino-5-ethyl-6-phenylphenanthridinium known as ethidium bromide (EB), for many decades applied as gold-standard DNA- and RNA-fluorescent marker, and its close analogue (propidium iodide) as a probe for cell viability. Besides, an antiparasitic activity for EB was reported and it possesses significant antitumor activity [2-5] both in vivo and in vitro. Nevertheless, phenanthridine derivatives were rather neglected regarding their human medicinal applications due to potential carcinogenic and mutagenic properties of some derivatives (EB and analogues), which had negative influence on biomedically-oriented studies of the complete phenanthridine class till the end of the 20th century.

However, discovery of phenanthridine alkaloid analogues and in parallel new studies of antiparasitic properties of phenanthridine derivatives resulted in a strong increase of the scientific

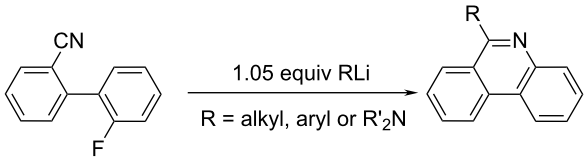
interest about the turn of this century, consequently yielding many publications at high impact chemical and biomedicinal journals, and patents covering various chemical, biochemical and biomedical uses. These results are up to our knowledge not summarized in any review within the last 10 years. Thus, taking advantage of our 20-year experience on phenanthridine derivatives (including very scarcely studied 4,9-diazapyrenium analogues with highly interesting biological effects), we summarized literature data (advances from 2000 to present) concerning supramolecular, bioorganic and medicinal chemistry, as well as novel or improved synthetic approaches.

# Review

## How to get phenanthridine: advances in synthetic pathways

Phenanthridine was first synthesized at the end of 19th century by Pictet and Ankersmit through pyrolysis of the condensation product of benzaldehyde and aniline [6]. The reaction conditions were improved by Morgan and Walls, based on a reaction including a cyclization of phenanthridine by dehydrative ring-closure with phosphorus oxychloride in boiling nitrobenzene [7]. Over the 20th century this method of phenanthridine preparation became the most common one. However, increased interest over the last decades resulted in many new and substantially different ways of phenanthridine synthesis with several different goals: to improve the reaction yield and to equip the phenanthridine moiety with various substituents, which were nicely summarized by Keller a decade ago [8]. We tried to survey the wide range of synthetic methods applied from 2000 on organizing them by similarity of reactants/catalysts or organic reactions; for instance the anionic ring-closure reactions using Grignard reagents (Scheme 1) [9], Bischler–Napieralski reactions [10], reduction of phenanthridones [11,12], free radical methodology, palladium/rhodium/iron-catalysed reactions, etc.

One of the approaches to the large variety of 6-arylphenanthridine derivatives was the synthesis starting from benzotriazole derivatives of diarylmethanes, acridine, xanthene, thioxanthene, etc. It was based on the generation of a benzotriazole-stabilized

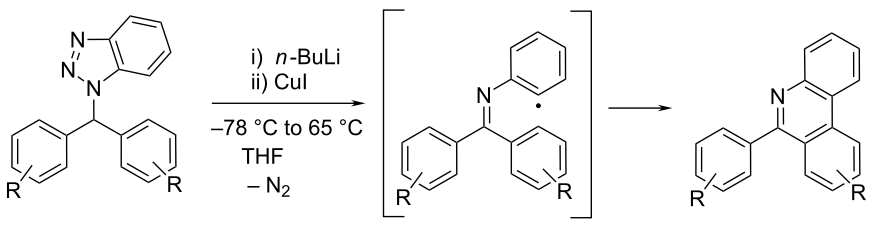



**Scheme 1:** The Grignard-based synthesis of 6-alkyl phenanthridine.

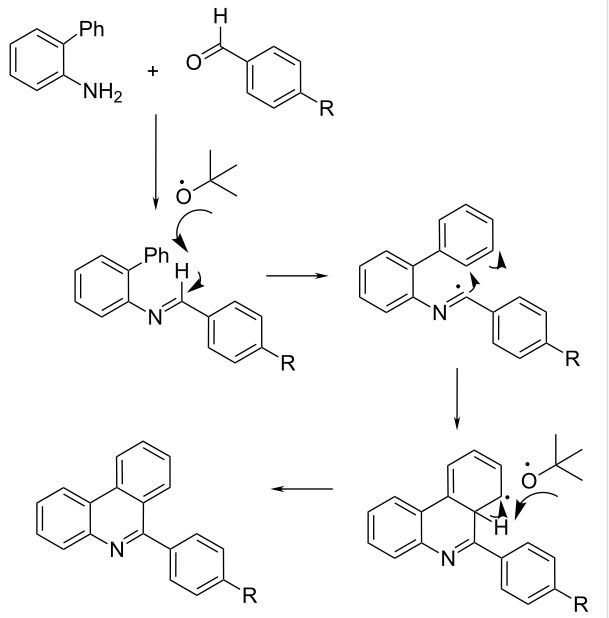
carbanion followed by oxidation of this carbanion by copper iodide to form a radical. Subsequent elimination of nitrogen followed by ring closure yielded phenanthridine (Scheme 2) [13,14].

In the 1980s, Leardini et al. [15] have shown that under radical conditions via a homolytic aromatic substitution route diarylimines were suitable precursors to a number of 6-arylphenanthridine derivatives. The reaction proceeded by initial imidoyl-H atom abstraction by the electrophilic iPrO$^{\bullet}$ radical, and subsequently the intermediate underwent intramolecular cyclization and oxidative aromatization to form the phenanthridine ring. Bowman et al. [16] modified this route for safety reasons by application of di(*tert*-butyl)peroxide as a source of the *t*-BuO$^{\bullet}$ radical (Scheme 3). The required arylimines were prepared from aminobiphenyl and arylaldehyde in dichloromethane in the presence of molecular sieves at room temperature. Radical cyclisation in the presence of (*tert*-butyl)peroxide in chlorobenzene at 140–150 °C for 48 h, yielded the corresponding phenanthridines in moderate yields. The *t*-BuO$^{\bullet}$ radical eliminated the imine-H and formed the imidoyl radical, added to the phenyl ring. The homolytic aromatic substitution was terminated by H-atom abstraction by another *t*-BuO$^{\bullet}$ radical.

Among very few routes targeting the synthesis of 5,6-unsubstituted phenanthridines, the here presented radical-based pathway used readily available *N*-(*o*-halobenzyl)arylamines as starting materials [17]. The *o*-haloarylbenzylamines (obtained by nucleophilic substitution of various anilines with 2-iodobenzyl chloride) gave the corresponding amide anions by an $S_{RN}1$ substitution reaction in $NH_3$ or DMSO as solvent under




**Scheme 2:** Radical-mediated synthesis of 6-arylphenanthridine [14].



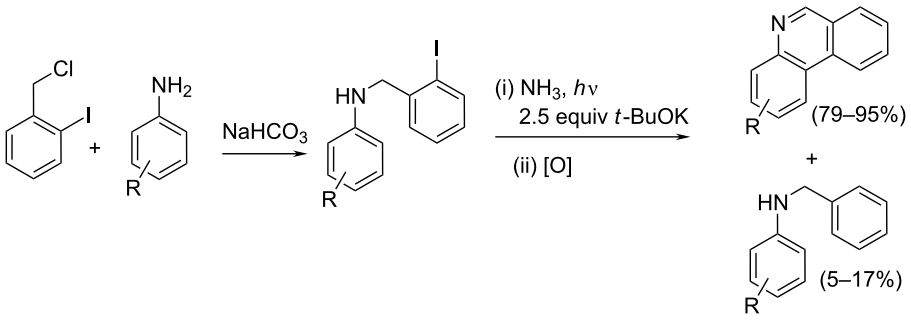
**Scheme 3:** A *t*-BuO$^{\bullet}$ radical-assisted homolytic aromatic substitution mechanism proposed for the conversion of diarylimine into the 6-arylphenanthridine derivatives [16].

photoinitiation in the presence of excess *t*-BuOK. The photoinduced ET to the amide-anion resulted in its radical anion. After fragmentation of the C–I bond, an intramolecular cyclization occurred, and after acidification of the reaction medium, the oxidized phenanthridine compounds were obtained in very good yields (up to 95%, Scheme 4).

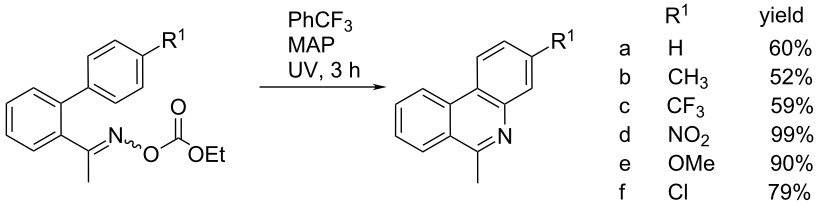
McBurney et al. prepared various N-heterocycles, using oxime carbonates as excellent precursors for the photoinduced generation of iminyl radicals, whereby at standard photolysis conditions, 3-substituted 6-methylphenanthridines were obtained in good to quantitative yields (52–99%, Scheme 5). Important advantages of the method are environmentally friendly and easily removable byproducts ($CO_2$ and ethanol or phenol), and the negligible impact of the electronic nature of the substituent on the reaction [18].

The oxidative $PhI(OAc)_2$-mediated cyclization of 2-isocyanobiphenyls with $CF_3SiMe_3$ under metal-free conditions showed to be a mild and efficient approach to 6-(trifluoromethyl)phenanthridines, characterised by good yields with high regioselectivity at ambient temperature (Scheme 6) [19,20].

Another radical-based route (targeting 6-perfluoroalkylphenanthridines), in which the transition metal is omitted, relied on the trifluoromethylation of isonitriles to yield trifluoromethylphenanthridines (Scheme 7) [21]. This approach employed the Togni reagent, and $Bu_4NI$ was applied as radical initiator; whereby phenanthridines were prepared in good to excellent yields [22]. Starting from the similar isonitrile structure, 6-aroylated

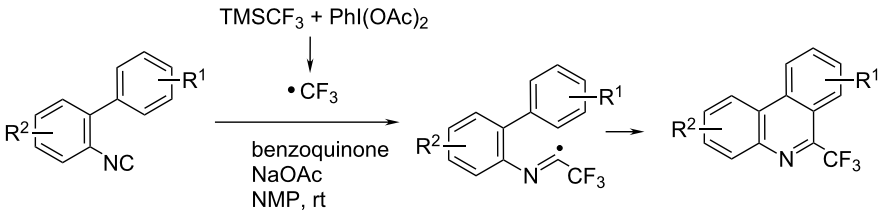


**Scheme 4:** Synthesis of 5,6-unsubstituted phenanthridine starting from 2-iodobenzyl chloride and aniline [17].

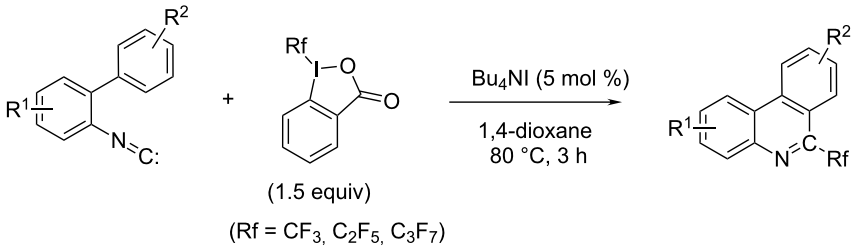


| | $R^1$ | yield |
|---|---|---|
| a | H | 60% |
| b | $CH_3$ | 52% |
| c | $CF_3$ | 59% |
| d | $NO_2$ | 99% |
| e | OMe | 90% |
| f | Cl | 79% |

**Scheme 5:** Phenanthridine synthesis initiated by UV-light irradiation photolysis of acetophenone *O*-ethoxycarbonyloxime derivatives at room temperature [18].



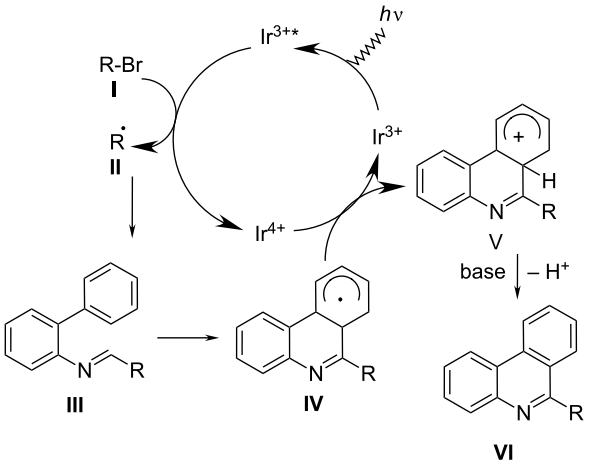

**Scheme 6:** $PhI(OAc)_2$-mediated oxidative cyclization of 2-isocyanobiphenyls with $CF_3SiMe_3$ [19,20].




**Scheme 7:** Targeting 6-perfluoroalkylphenanthridines [21,22].

phenanthridines via base promoted homolytic aromatic substitution (BHAS) can be prepared [23].

Several photoinduced synthetic procedures were also applied. For instance, the photochemical cyclization of *N*-benzylanilines was used for asymmetrically substituted derivatives at phenanthridine side-rings and unsubstituted central ring [24]. The recently reported photo-conversion of various isocyanide biphenyls into alkylated phenanthridine derivatives under rather mild reaction conditions introduced several novelties (Scheme 8) [25]. The most intriguing is the double role of the photocatalyst [fac-Ir(ppy)$_3$], consisting of photo-induced generation of alkyl radical **II** and oxidation of radical **IV** to cationic intermediate **V**, the latter process also regenerated the catalyst. Finally, the deprotonation assisted by base resulted in various 6-alkylated phenanthridines in excellent yields (>92%). The radical inhibitor 2,2,6,6-tetramethylpiperidin-1-oxyl (TEMPO) was applied to stop the transformation by a single electron transfer process.

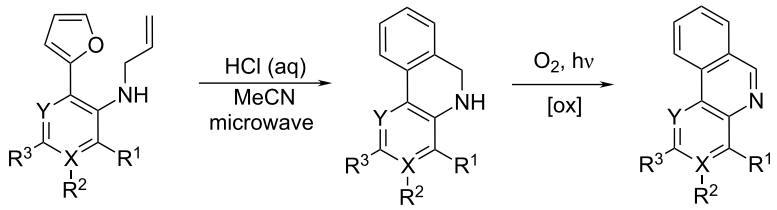



**Scheme 8:** Easily accessible biphenyl isocyanides reacting under mild conditions (room temp., visible light irradiation, blue LED light source, $N_2$, DMF, 10 h) with various common alkyl bromides by application the two-role catalyst [fac-Ir(ppy)$_3$], gave phenanthridines in good yields [25].

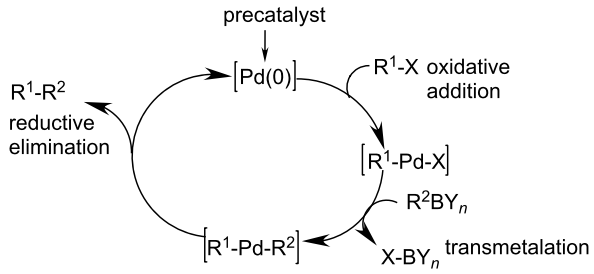
Intriguing combination of irradiation techniques (combined microwave-assisted and photochemical) offered a new route toward phenanthridines. Microwave-mediated intramolecular Diels–Alder cyclization of *o*-furyl(allylamino)arenes followed by spontaneous aromatization yielded dihydrophenanthridines, which upon exposure to UV light (315−400 nm) were oxidized into (aza)phenanthridines (Scheme 9) [26].

Synthetic pathways based on the transition metal-catalysed functionalization of carbon–hydrogen (C–H) bonds and formations of C–C bonds are often used to access phenanthridines [27-29]. The most common are high-yield, palladium-based methodologies under mild conditions, due to their applicability on a large variety of aryl substituents [16] as well as potential for stereo and regioselectivity (Scheme 10) [30,31].

Among many examples, very recently a two-step phenanthridine synthesis stands out as a new strategy, characterised by two roles of the Pd-catalyst in the 1$^{st}$ step, followed by simple and cost effective oxidation [32] (Scheme 11). This synthesis




**Scheme 9:** Microwave irradiation of Diels–Alder adduct followed by UV irradiation of dihydrophenanthridines yielded phenanthridines [26].

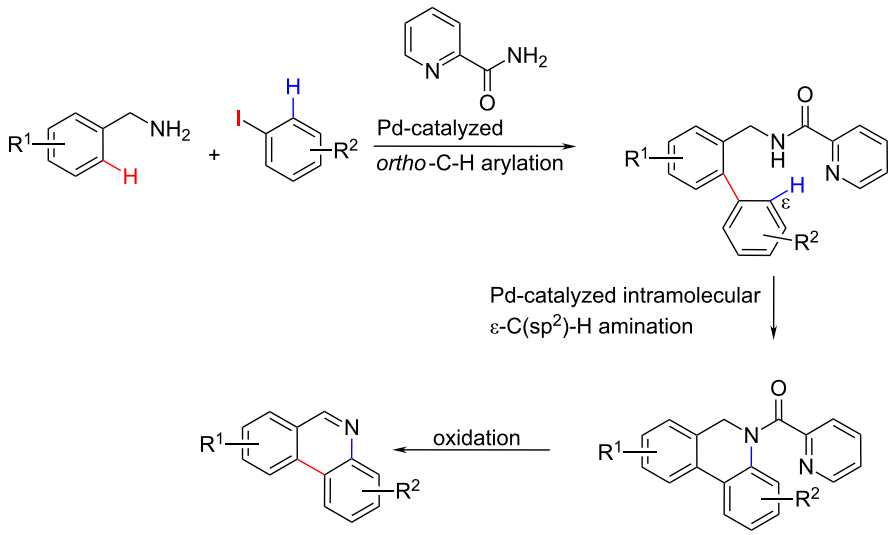



**Scheme 10:** A representative palladium catalytic cycle.

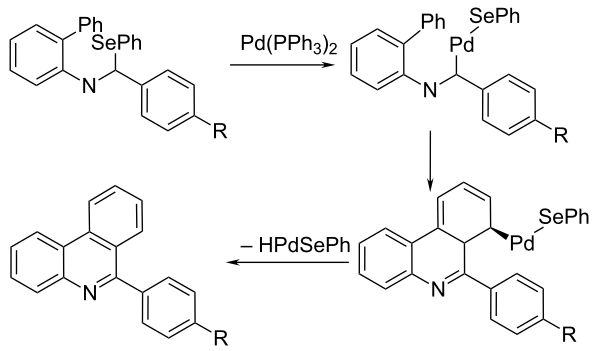
published by Pearson et al. was based on palladium-catalysed picolinamide-directed sequential C–H functionalization reactions, while readily available benzylamine and aryl iodide were used as precursors. In the first step the Pd-catalyzed reaction yielded a biaryl compound. The second step under the catalysis of $Pd(OAc)_2$ comprised both cyclisation and oxidation in a single step: a dehydrogenative C–H amination with $PhI(OAc)_2$ as oxidant and removal of the picolinamide group followed by oxidation with $Cu(OAc)_2$. This strategy afforded phenanthridines in moderate to good yields (up to 65% for the second step).

Bowman et al. reported a palladium-mediated route using imidoyl-selenides as precursors besides the radical route. Comparison of the cyclisation yields for the same set of phenanthridine derivatives revealed an overall better efficiency of the *t*-BuO$^{\bullet}$ radical-assisted homolytic aromatic substitution of diarylimine (Scheme 3) in respect to the Pd(0)-mediated cyclisation of imidoyl-selenides (Scheme 12) [16]. Authors proposed insertion of a Pd(0) species into the carbon–selenium bond, followed by carbo-palladation onto the phenyl ring. This intermediate then undergoes rapid rearomatization with the loss of HPdSePh to give the phenanthridine.

Candito et al. reported a new and highly efficient method for the synthesis of variously substituted phenanthridine derivatives employing *N*-unsubstituted imines or *N*-silylimines [33]. The method is limited to *ortho*-substituted aryl iodides as starting material and also requires a convenient imine derivative allowing the cleavage of the nitrogen-attached group ($R^5$) at




**Scheme 11:** The common Pd-catalyst for the biphenyl conjugation results simultaneously in picolinamide-directed cyclisation; obtained *N*-picolinamide dihydrophenanthridine is easily converted to phenanthridine [32].



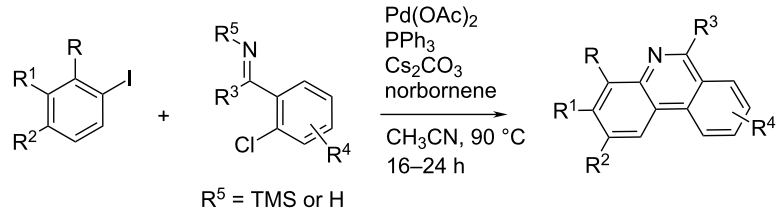

**Scheme 12:** Pd(0)-mediated cyclisation of imidoyl-selenides forming 6-arylphenanthridine derivatives [16]. The insertion of the Pd(0) species into the carbon selenium bond followed by fast rearomatisation to phenanthridine is involved with the loss of HPdSePh.

some point in the catalytic cycle (Scheme 13). It is noteworthy that polar solvents (DMF, *N*-methylpyrrolidone, and acetonitrile) had favourable impact on the reaction.

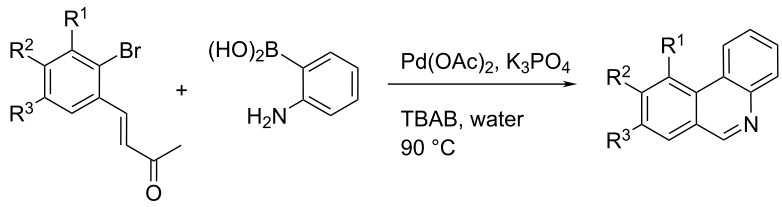
Ghosh, Dhara et al. also reported a synthesis of substituted phenanthridines based on palladium-mediated Suzuki coupling (Scheme 14) [34,35]. Aerobic ligand-free domino Suzuki coupling–Michael addition reaction in the presence of $Pd(OAc)_2$ and $K_3PO_4$ as a catalytic system in $H_2O$ was catalysed by palladium nanoparticles, that were generated in situ in water with the elimination of acetone.

One of the major issues is the preparation of polysubstituted phenanthridines, in particular asymmetrically positioned on one of phenyl side-rings. An intriguing approach over rhodium-catalysed alkyne [2 + 2 + 2] cycloaddition reaction [36] (Scheme 15) offered a highly efficient method with excellent regioselectivity (in case of bulky groups), with additional advantage of the C-6 fluorinated methyl substituent – promising for PET probe development.

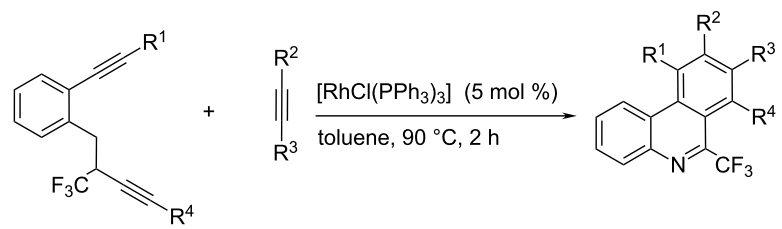
Most of the metal catalysts employed for phenanthridine synthesis are rather expensive; therefore efforts were made to replace them with cheaper analogues. One successful approach




**Scheme 13:** Palladium-catalysed phenanthridine synthesis.




**Scheme 14:** Aerobic domino Suzuki coupling combined with Michael addition reaction in the presence of a $Pd(OAc)_2/K_3PO_4$ catalytic system in water [34,35].




**Scheme 15:** Rhodium-catalysed alkyne [2 + 2 + 2] cycloaddition reactions [36].

included iron(III) acetylacetonate in acetic acid as catalytic agent (Scheme 16) [37], whereby *O*-acetyl oximes obtained from 2′-arylacetophenones underwent N–O bond cleavage and intramolecular *N*-arylation. The following conventional cross-coupling or directed C–H arylation resulted in substituted phenanthridines.

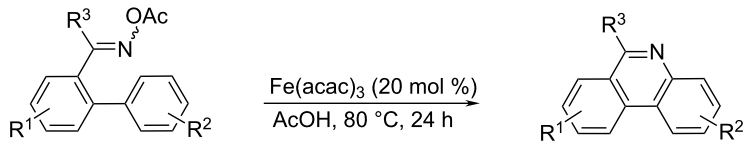
Homolytic aromatic substitution (HAS) by an aryl radical was used for the construction of biaryl motifs as alternative to transition metal-catalysed C–H bond arylation. That approach was also implemented in the two-component cyclization in the synthesis of phenanthridine derivatives [38]. The starting isocyanide biphenyl (similar to Scheme 8) reacts with the phenyl radical generated from phenylboronic acid and a manganese salt followed by spontaneous cyclisation and aromatisation.

Trying to omit the expensive metal catalysts, several successful attempts of a transition metal-free approach for phenanthridine synthesis were reported. For instance application of a simple diol combined with KO*t*-Bu resulted in intramolecular C–H arylation to give the respective phenanthridine derivatives (Scheme 17 top) [39]. More recently, a similar procedure worked just in the presence of KO*t*-Bu by intramolecular homolytic aromatic substitution (HAS), without the use of an organic molecule as ligand to give benzo[*c*]phenanthridine derivatives (Scheme 17 bottom) [40].

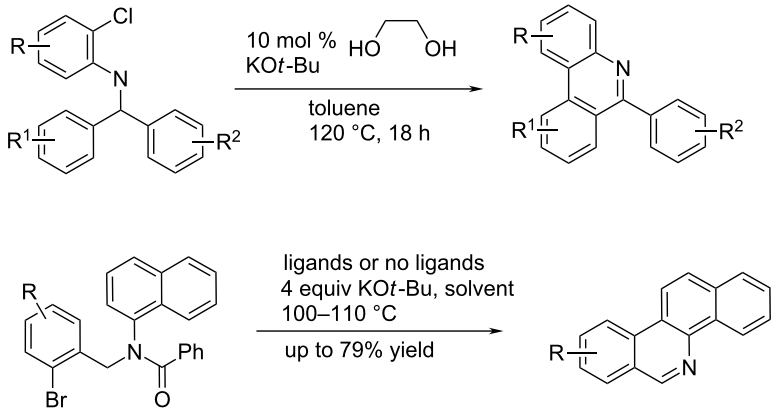
An unique approach to the phenanthridine core starting from a simple disubstituted aniline relied on the aza-Claisen rearrangement, ring-closing enyne metathesis and Diels–Alder reaction [41] (Scheme 18). The obtained phenanthridine derivatives were polysubstituted at the phenyl side-rings, while retaining the unsubstituted central heterocyclic double bond. The diversity of the aza-Claisen rearrangement allows the application of this approach to other related heterocyclic systems.

The preparation of a new variety of analogues, namely 6-phosphorylated phenanthridines was very recently reported, whereby central-ring cyclisation was accompanied with simultaneous phosphorylation [42] (Scheme 19). The particular importance of this economic and highly efficient synthetic method is the complementarity of the starting material, the easy availability of 2-isocyanobiphenyls, which could be converted to variously substituted phenanthridines in several ways (Scheme 7 and Scheme 8).

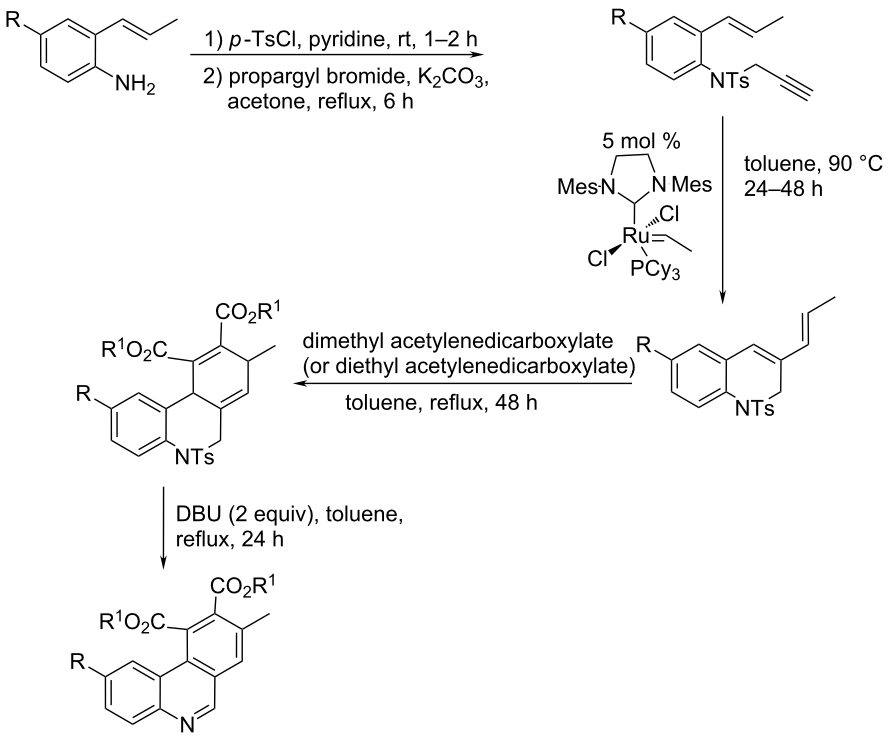
Because of the recent strong focus on benzophenanthridines due to their potent antitumor and antiinfectious activities [43], we have chosen one recent synthetic approach (differing from afore



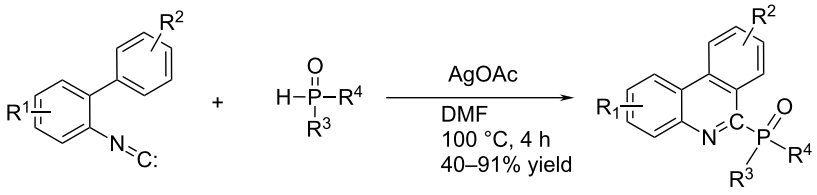

**Scheme 16:** The *O*-acetyloximes derived from 2′-arylacetophenones underwent N–O bond cleavage and intramolecular *N*-arylation, followed by cross-coupling or directed C–H arylation [37].




**Scheme 17:** C–H arylation with aryl chloride in the presence of a simple diol complex with KO*t*-Bu (top) [39]; for some cases it worked also in the absence of diol (bottom) [40].

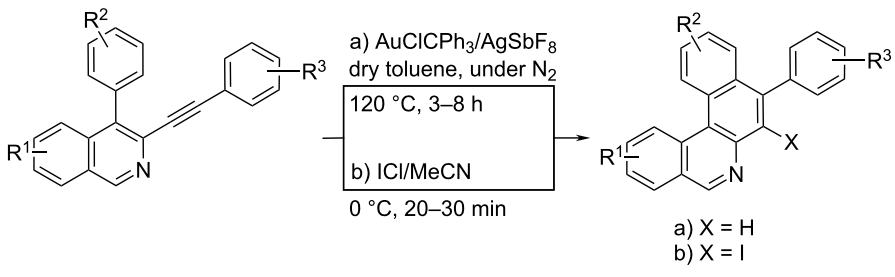



**Scheme 18:** The subsequent aza-Claisen rearrangement, ring-closing enyne metathesis and Diels–Alder reaction – a new "three-atom economic process" of phenanthridine synthesis [41].




**Scheme 19:** Phenanthridine central-ring cyclisation with simultaneous radical-driven phosphorylation [42].

listed examples) to benzo[*a*]phenanthridines as close analogues of phenanthridine (Scheme 20) [44]. The synthesis by multi-component tandem reaction/carbocyclization starts with the formation of a 4-aryl-3-arylethynylisoquinoline from 2-bromobenzaldehyde/*tert*-butylamine/1,3-diyne. The second (in situ) step is based on the ring closure, either via gold/silver-catal-




**Scheme 20:** Three component reaction yielding the benzo[*a*]phenanthridine core in excellent yields [44].

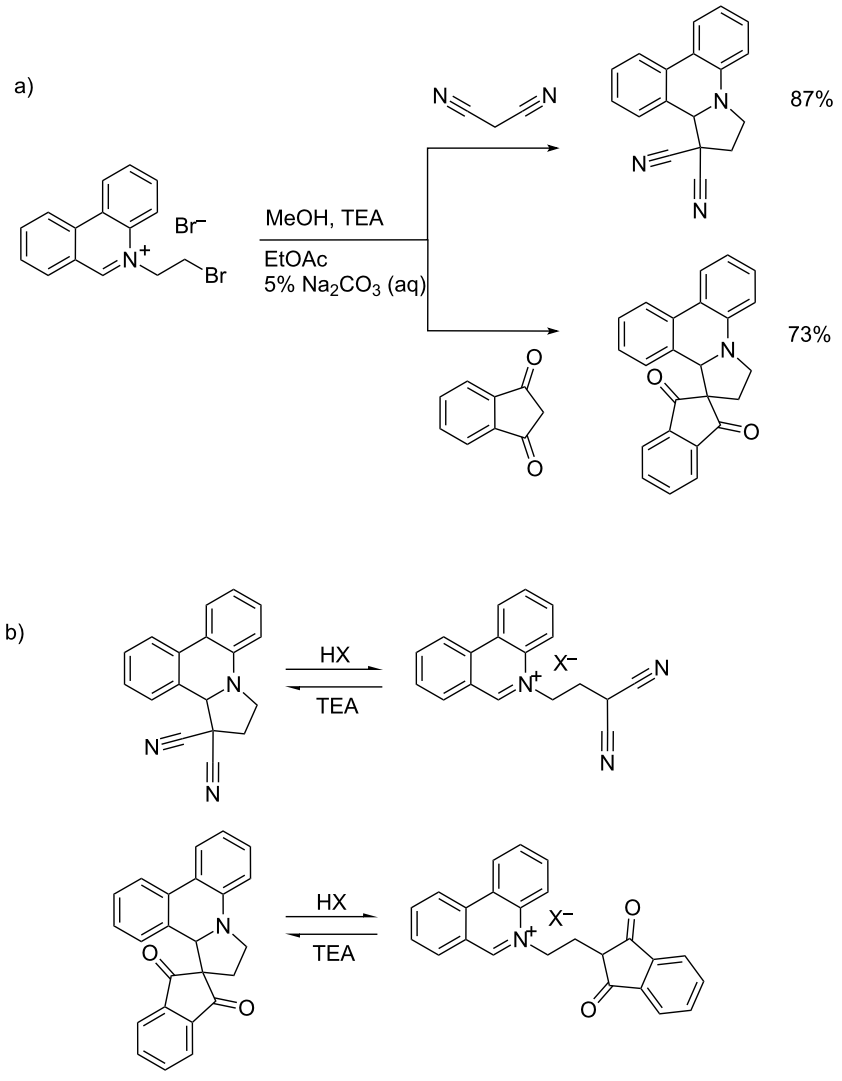
ysed intramolecular hydroarylation or via iodo-catalysed regioselective 6-endo-dig electrophilic cyclization.

Kitson et al. synthesized a class of 2,3-dihydro-12*H*-pyrrolo[1,2-*f*]phenanthridine (DPP) derivatives starting from malononitrile and 1,3-indandione as the initial nucleophiles, which reacted with *N*-bromoethylphenanthridinium bromide to give DPP-dicarbonitrile and DPP-indandione, respectively. Particularly an interesting property of these DPP products is the reversible, pH controlled ring-opening-cyclisation process, whereby at acidic conditions DPP undergoes rearomatisation of the phenanthridinium ring system (Scheme 21a), which at basic conditions (TEA) switches back to the initial DPP structure (Scheme 21b) [45].

## Summary of synthetic advances

Among many given examples of the phenanthridine synthesis, currently two most common routes are the synthesis under radical conditions [12-24] and the synthesis based on transition metal-catalysis [15,27-37].

The main advantage of the radical-based phenanthridine synthesis is easy available and generally cheap starting material (benzotriazole, aminobiphenyl, arylaldehyde, *N*-(*ortho*-halobenzyl)arylamines, oxime carbonates, isocyanobiphenyls, etc.). Phenanthridines are usually obtained within 2–3 reaction steps, by application of different radical initiators. An intriguing alternative is the radical generation by UV irradiation with or even without a photocatalyst. The major advantage of radical-




**Scheme 21:** a) Reaction of malononitrile and 1,3-indandione with BEP to form the cyclised DPP products; b) pH controlled reversible cyclisation process of DPP compounds [45].

based routes are usually mild reaction conditions, while reaction yields, after optimization of the synthesis parameters, span from 50–90%, mostly depending on the substituents attached to the starting material. The radical-based synthesis is typically used for the preparation 6-aryl or 6-alkylphenanthridine derivatives and 6-phosphorylated analogues, equipped with one or two additional substituents, usually positioned on the phenanthridine positions C1–4 or position C8.

Similarly to the radical-based synthesis, a synthetic approach based on transition metal-catalysis also allows the phenanthridine preparation from easily available starting material (benzylamine, aryl iodide, imines, etc.) in few reaction steps, under mild reaction conditions and with yields within the 50–90% range. The great advantage of this approach is the very broad versatility in preparation of phenanthridine derivatives, polysubstituted on the phenyl side-rings by a large variety of substituents, as well as stereo- and regioselectivity (particularly for the bulky groups). Nevertheless, due to the most common metal catalyst (palladium) this method is significantly more expensive and less environmentally friendly than radical-based methods. To address these disadvantages, in the last decade particular attention was given to the replacement of the expensive palladium catalyst, for instance by iron [37]. However, major impact was made by introduction of the cheap and environmentally friendly intramolecular homolytic aromatic substitution (HAS) reaction with the aid of the organo-catalysis; although it is currently applicable for the preparation of only a limited variety of phenanthridine derivatives and benzophenanthridines but future prospects are very promising.

Aside two most common ways to prepare the phenanthridine moiety, here are described several innovative approaches, with potential to be developed for a large versatility of phenanthridine derivatives or application of previously not used starting materials (for instance microwave-mediated intramolecular Diels−Alder cyclization of *o*-furyl(allylamino)arenes).

For the most of DNA or RNA targeted applications the phenanthridine is converted to the positively charged phenanthridinium cation by simple alkylation of the phenanthridine heterocyclic N5 nitrogen (thus giving permanent positive charge) or by the N5 nitrogen protonation at weakly acidic conditions (p*K*(N5) ca. 5.5–6) yielding reversible positive charge. Here are also described novel approaches to reversible positively charged (DPP and DIP derivatives [45]), which are related to remarkable structural features of the naturally occurring benzophenanthridine alkaloids – pH-dependent structural transition between the iminium (positively charged) and alkanolamine (neutral) form [46].

## Structural features of phenanthridines and phenanthridinium cations related to DNA and RNA binding

Structural studies on the phenanthridine system were mostly driven by its most widespread use as DNA and RNA intercalator (Figure 1) and/or fluorescent marker (ethidium bromide/propidium iodide) for ds-DNA and ds-RNA [47]. The phenanthridine structural features incorporate a unique set of properties related to the interaction with DNA and RNA (Figure 1): size and curvature of the aromatic surface corresponds to the

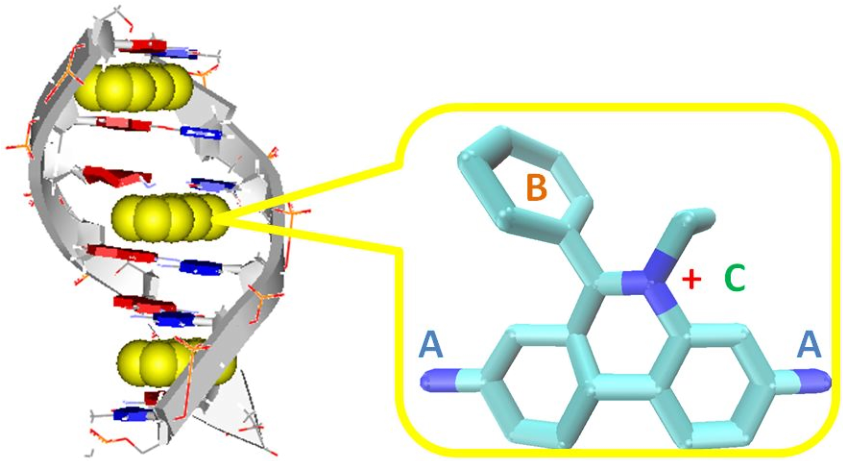



**Figure 1:** Schematic presentation of the intercalative binding mode by the neighbour exclusion principle and important structural features of ethidium bromide: A**)** amino substituents responsible for fluorescence increase upon DNA intercalation; B) phenyl substituent for steric control and also impact on fluorimetric properties; C) permanent positive charge for aqueous solubility and electrostatic attraction to the DNA or RNA phosphate backbone.

basepair shape, whereas the high polarizability (and permanent positive charge of N-5 alkylated derivatives) also plays an important role in aromatic and electrostatic interactions with polynucleotides. Moreover, non-covalent interactions with DNA and RNA can be reversibly controlled by a pH-induced positive charge at the heterocyclic nitrogen N-5, and strong electron affinity and polar groups at the 3 and/or 8 position of the phenanthridine can efficiently and predictably regulate the spectroscopic response (UV–vis and fluorescence) of the chromophore [48].

The understanding of the intercalation process requires a detailed knowledge of the energetics, thermodynamics and structural equilibrium – surprisingly few studies endeavoured to determine important parameters for such classical intercalator as ethidium bromide [49]. The most recent and very extensive theoretical study compared positively charged ethidium bromide and its neutral analogue, revealing detailed description of the forces included in the intercalation process, stressing the dispersion energy as a control factor [50]. Moreover, a number of kinetic measurements provided for the binding of ligands to DNA additionally clarify mechanistic details that are not apparent from equilibrium measurements [51].

Another, very comprehensive approach, relying mostly on the experimental data of X-ray crystallography, UV–vis, fluorescence and NMR spectroscopy, determined that the fine interplay between electron donating and electron withdrawing effects mediated by its nitrogen atoms defines the spectroscopic properties of ethidium bromide (EB) and its derivatives [48]. It turned out that, despite the positive charge of ethidium bromide, most of ethidium's aromatic carbon and hydrogen atoms have high electron densities compared to the 6-phenylphenanthridines. Thus, the electron-donor properties of the exocyclic amines, especially at 8-position have a stronger influence on the electron density of aromatic atoms than the electron withdrawing effects of endocyclic iminium. Fine tuning of electron properties of EB can be easily achieved via chemical modulation of its amino groups at 3 and 8 positions of the phenanthridine ring [52,53]. Systematic changing of the ethidium bromide exocyclic amines into guanidine, pyrrole, urea, and various substituted ureas revealed importance of electron-donor properties of substituents at the 3- and 8-position of the phenanthridinium relative to the unmodified primary amino groups. Namely, derivatives of EB having substituents with weaker electron-donor properties exhibited a stronger fluorescence emission than EB, while a stronger electron-donating substituent exhibited a much lower fluorescence emission. Such behaviour could be attributed to the ethidium exocyclic amines enabling by electron donation a non-radiative decay of phenanthridinium excited state, rather more likely than the previously proposed mechanism of water-induced deprotonation of phenanthridinium exocyclic amines, causing excited chromophore fluorescence quenching [54,55].

Taking into account the research results of several other groups, a general rule could be drawn that phenanthridines with no amino groups yield strong fluorescence in water but emission is totally quenched by DNA binding; one amino group at (usually) position 8 results in only a small fluorescence change in the complex with DNA, while two amino groups in 3,8-position result in a weak fluorescence with strong emission increase upon DNA binding [52,53,56].

A pronounced influence of the substituent at phenanthridine position 6 on the optical properties of the chromophore also had significant impact on the binding affinity toward ds-DNA. The comparison of three substituents in 6-position, 4-*N*,*N*-diethylaminophenyl, phenyl (EB) and methyl, revealed that the first one exhibits the strongest DNA binding affinity and the strongest fluorescence enhancement. That was related to the twist angle in the excited state between the 6-phenyl ring and the phenanthridinium chromophore, which controls the non-radiative relaxation [56].

## Substituted phenanthridine derivatives

In particular guanidine- and urea-substituted derivatives attracted a lot of attention due to the different interactions with various DNA and RNA. The ability of ethidium bromide analogues to inhibit the HIV-1 Rev–Rev Response Element (RRE) interaction, as well as their affinity to calf thymus (ct)DNA was analysed. One derivative (Figure 2, **1**) displayed

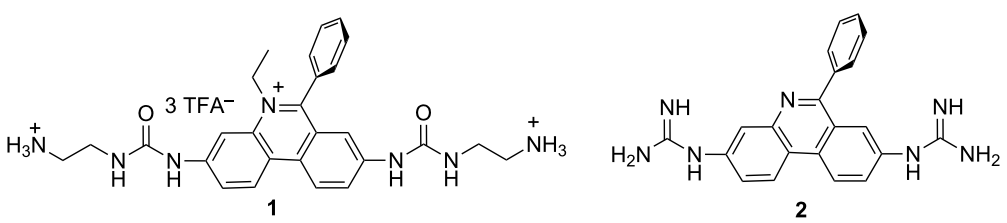



**Figure 2:** Urea and guanidine derivatives of EB with modified DNA interactions [57].

an enhanced affinity for HIV-1 RRE and a lower DNA affinity (i.e., lower mutagenic activities) compared with ethidium bromide. A recent study showed that substitution of both ethidium bromide (EB) and exocyclic amines by guanidines converted the classical intercalator (EB) into a DNA minor groove binder [57]. The most intriguingly, binding mode change did not weaken the DNA affinity, thus the affinity of guanidine derivative **2** (Figure 2) towards AT-rich DNA sequences was significantly stronger compared to ethidium and comparable to that of the known DNA minor groove binder furamidine.

The above mentioned guanidine-induced switch of the DNA- and RNA-binding mode [57] inspired a design of derivatives equipped with biguanide groups at 3 and/or 8 positions [58] (Figure 3), under the presumption that the extended H-bond-rich system should increase the ability of the chromophore to differ among various shapes of ds-DNA- and ds-RNA-grooves. Both, mono- (**3**) and bis-biguanide (**4**), efficiently discriminate between dAdT and dGdC polynucleotides by opposite changes of compound fluorescence, as well as opposite induced (I)CD bands (Figure 3). Moreover, both, **3** and **4**, show the binding to AU-RNA by a different fluorimetric and CD response in respect to DNA-binding. Observed recognition between various DNA and RNA polynucleotides was attributed to the switch of the binding mode (intercalation into dGdC-DNA and AU-RNA and minor groove binding into dAdT-DNA).

A common strategy for the modification of DNA- and RNA-targeting molecules by preparation of homo-dimers was also implemented on the phenanthridine moiety – many ethidium bromide-based dimers were prepared and reviewed in the last two decades of the 20th century, thus here will be presented results from 2000 on.

Systematic variation of steric and/or electrostatic effects by means of type, number, length and flexibility of linkers connecting two phenanthridine units is presented in Scheme 22.

The ability of switching on/off the charge of phenanthridine heterocyclic N5 via its protonation at weakly acidic pH (pK 5–6) was utilized in a design of phenanthridine derivatives to alter significantly their binding preferences toward polynucleotides. Among several examples, the most intriguing pH controlled binding of nucleotides and nucleic acids showed bis-phenanthridine triamine [59] (**8**, Scheme 22). Compound **8** intercalated with only one phenanthridinium subunit into all ds-DNA and ds-RNA, while additional interactions of the other subunit within the grooves finely tuned the recognition between various ds-polynucleotides. The sensitivity of spectroscopic response was particularly pronounced for ss-RNA, whereby at weakly acidic pH compound **8** exhibited specific fluorimetric sensing of poly(G) among other studied ss-polynucleotides.

Cyclic cage-like bis-phenanthridinium derivatives (Scheme 16; general structure **7**), with a rigid structure allowing accommodation of only one nucleobase, showed pronounced ss-RNA over ds-RNA/DNA selectivity [60], whereas more flexible cyclic (**6**) and acyclic analogues (**5**) [61] revealed opposite preference, stressing the importance of steric control over selectivity (Scheme 16). The selectivity of **7** was based on the switch

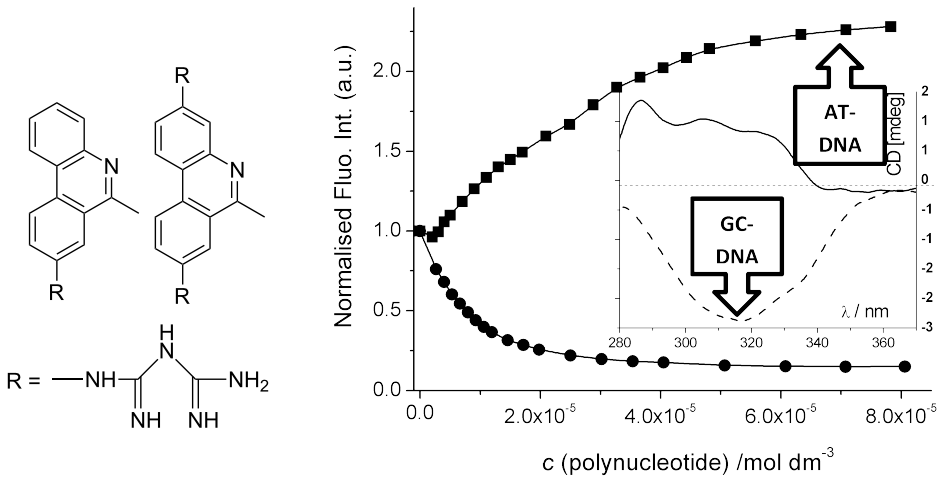
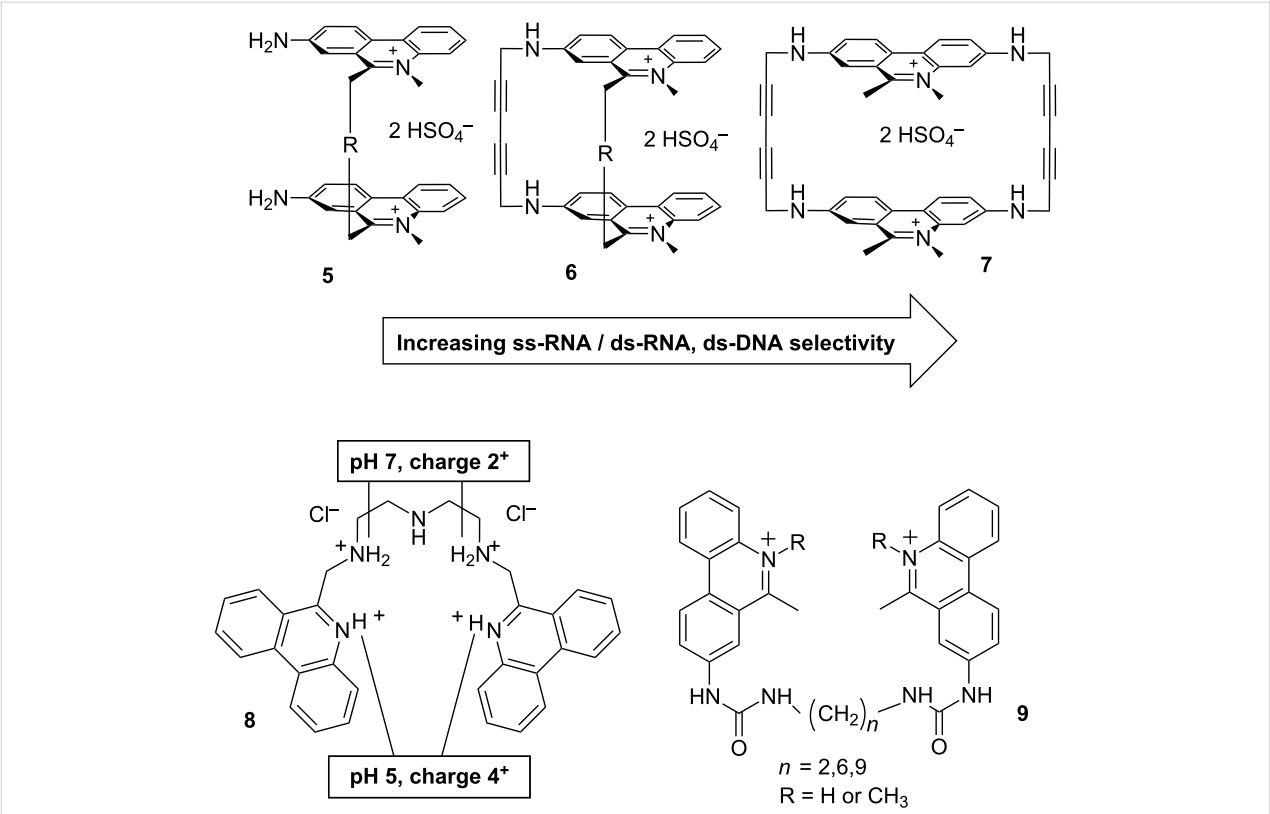



**Figure 3:** Structure of mono- (**3**) and bis-biguanide (**4**) derivative. Fluorescence (y-axis normalised to starting fluorescence of free **4**, $c$ = 1.0 × $10^{-6}$ mol $dm^{-3}$) was quenched by GC-DNA and increased for AT-DNA. Inset: induced (I)CD spectra λ > 280 nm of **4** (r(**4**)/DNA = 0.3; $c$(DNA) = 2 × $10^{-5}$ mol $dm^{-3}$) – strong positive ICD band for AT-DNA and negative ICD band for GC-DNA. Adapted with permission from [58]. Copyright 2011 The Royal Society of Chemistry.




**Scheme 22:** Bis-phenanthridinium derivatives (**5**–**7**; inert aliphatic linkers, R = –$(CH_2)_4$– or –$(CH_2)_6$–): rigidity of a "cage" – steric control of binding site. Triamine-linked bis-phenanthridine **8**, note reversible doubling of positive charges at pH 5 in respect to neutral conditions (pH 7). Bis-urea phenanthridines (general structure **9**): different from amino analogues (**5**) by fluorimetric response and DNA- and RNA-binding modes.

of binding mode; the very rigid pocket between two phenanthridinium moieties allows only bis-intercalation into single-stranded polynucleotides and only binding with double-stranded polynucleotides in non-intercalative mode (most likely within the DNA and RNA grooves). Moreover, the cage-like binding pocket of bisphenanthridiniums **7** showed to be sensitive to the minor structural differences between mononucleotides, yielding a very selective fluorimetric response upon binding of AMP in respect to other nucleotides. In addition, the observed selectivity towards poly(G) and poly(A) can be beneficial in biological applications for instance to influence the mRNA-function via binding to the poly(A) tail [62-64] and inhibition of the HIV-1 replication by targeting recognition of the polypurine tract by reverse transcriptase [65].

In a series of N5-protonated urea-substituted bis-phenanthridinium derivatives (Scheme 22, general structure **9**), the variation of the linker length connecting two urea-phenanthridinium conjugates significantly influenced the efficiency of intramolecular interactions between two phenanthridinium subunits and consequently their DNA- and RNA-binding mode (shorter linker–minor groove binding, the longest linker–intercalation) [66,67]. In addition, the derivative with the longest linker was, to the best of our knowledge, the first bis-phenanthridine-based intercalator able to differentiate between A–U(T) and G–C base pairs by sign of opposite fluorimetric response. An introduction of the permanent positive charge by methylation of the heterocyclic nitrogen changed the binding mode of the conjugates with shorter linkers from minor groove binding to intercalation and also resulted in significantly higher biological potency in respect to non-methylated analogues [67]. Moreover, the observed DNA and RNA interactions were also distinctively different from previously studied aliphatic-linker analogues (**5**), pointing out the decisive role of urea-linker interactions.

The common approach to complex small molecules targeting DNA and RNA usually required a number of consecutive synthetic steps, which made modification of the interesting structures a laborious and time-consuming task, quite often being the bottle-neck in the structure–activity relation research. With aim to facilitate structural modifications in DNA and RNA targeting by oligo-aryl derivatives, new amino acids with phenanthridine attached to the side chain were prepared and the solid phase synthesis of novel peptide-bridged bis-phenanthridine derivatives was developed (Figure 4) [68], whereby the position of the DNA-active chromophore in the peptide backbone as well as

the structural characteristics of the linker between them can easily be modified. In the first series of peptide-bridged bis-phenanthridine derivatives, derivative **11** with the shortest linker formed an intramolecular excimer, characterised by the specific fluorescence band sensitive to the pH as well as on the interactions with ds-DNA. Interestingly, all peptide-based phenanthridines revealed excellent water solubility combined with low in vitro toxicity, thus being good candidates for development of new safe fluorimetric DNA and RNA dyes.

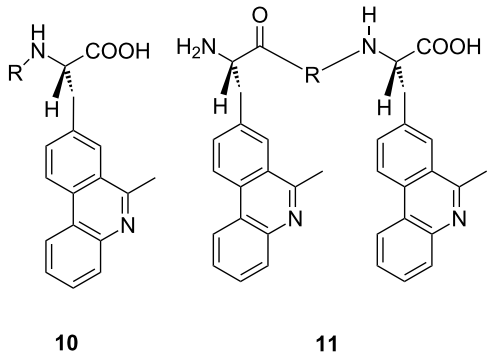



**Figure 4:** Series of amino acid–phenanthridine building blocks (general structure **10**; R = H; Gly) and peptide-bridged bis-phenanthridine derivatives (general structure **11**; R = X; Gly; Gly–Gly) [68].

Another large series of phenanthridinium-homodimers was constructed by linking two ethidium bromide subunits by peptide-like linkers of variable flexibility and rich in hydrogen-bonding possibilities within the DNA grooves (Figure 5). The resulting bis-intercalators (in comparison to the monomeric analogues) revealed significantly increased DNA-binding affinity and consequently enhanced telomerase and reverse transcriptase inhibition [69].

## Conjugates of phenanthridine with other DNA and RNA active moieties

Another common approach to increased selectivity of DNA- and RNA-targeting small molecules is the design of complex conjugates consisting of several DNA- and RNA-active parts (e.g., intercalator, groove binder, electrostatically binding component, various sterically directing parameters). The phenanthridine moiety was quite often used as presumably intercalating unit, although in some cases a switch of the binding site to the minor groove was reported.

In an effort to influence DNA sequence-selective recognition by small molecules (MW <1000), our group prepared a series of phenanthridine derivatives with one or two nucleobases covalently attached at the 3 and/or 8 positions of the phenanthridine ring (Scheme 23). The phenanthridinium–nucleobase conjugates did not show targeted selectivity toward complementary nucleotides in aqueous medium due to the strong competition of bulk water with the expected hydrogen bonds [70,71]. Fortuitously, the hydrophobic environment within the common DNA/RNA binding sites allowed H-bonding-based recognition of some complementary polynucleotide sequences. However, the recognition pattern was not straight-forward; for instance N5-protonated phenanthri-dinium–adenine derivative **12** successfully recognized a complementary poly(U) sequence [72] (Scheme 23), but this recognition was completely lost upon introduction of a permanent positive charge by methylation of phenanthridine-N5 **13** [71]. Intriguingly, N5-methylated phenanthridine–adenine conjugate **13** exhibited preferred

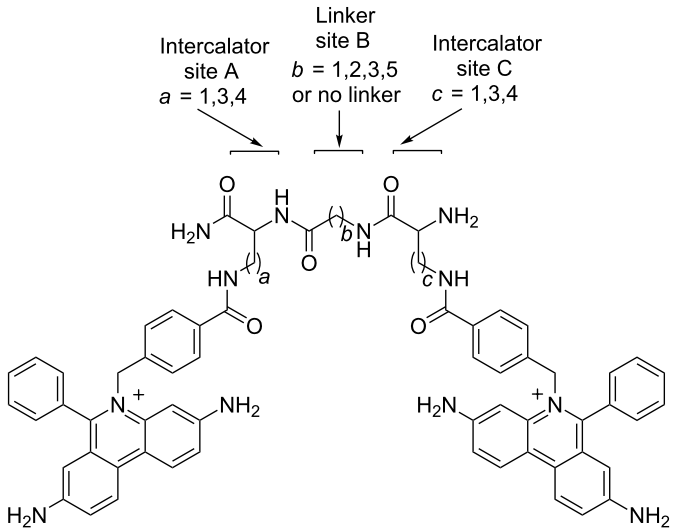
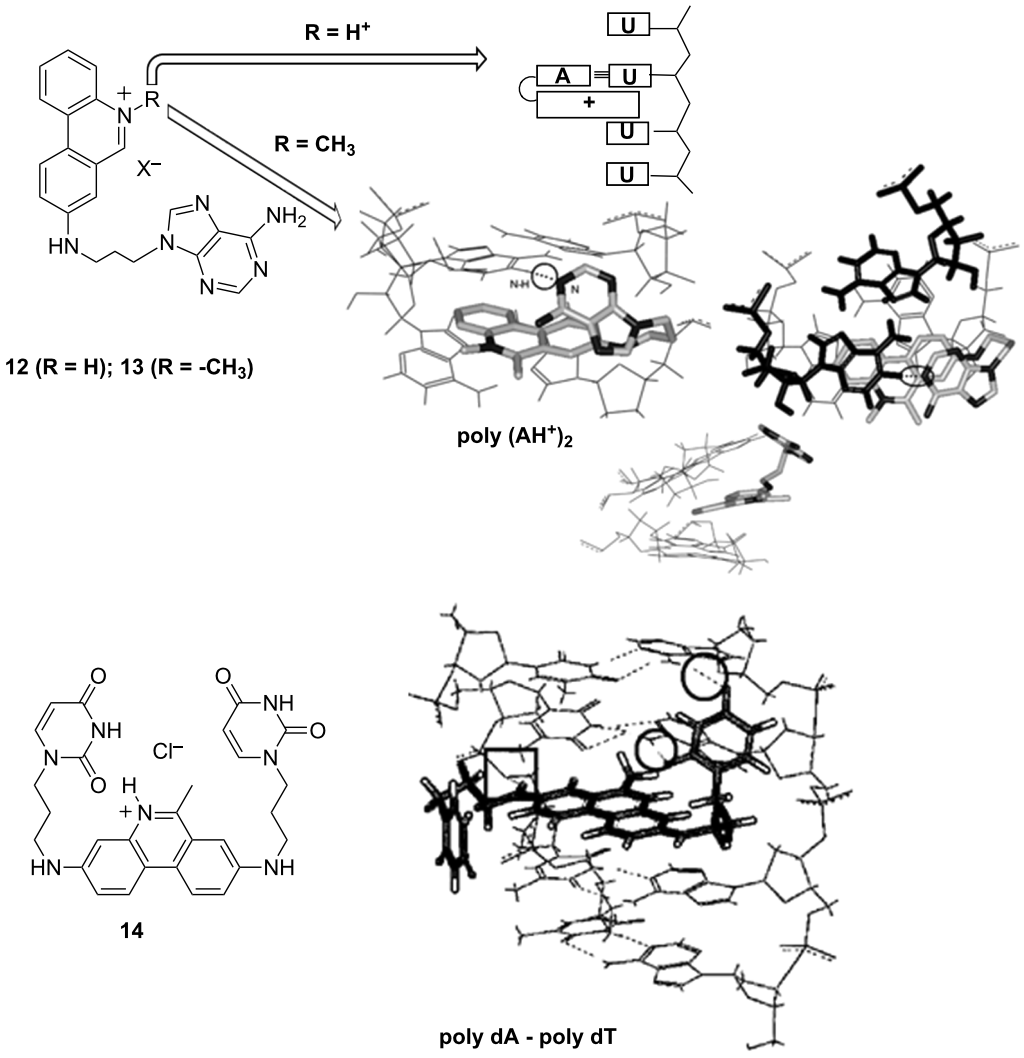



**Figure 5:** General structure of 45 bis-ethidium bromide analogues. Reproduced with permission from [69]. Copyright 2012 Elsevier Ltd.




**Scheme 23:** Top: Recognition of poly(U) by **12** and ds-polyAH$^+$ by **13**; bottom: Recognition of poly(dA)–poly(dT) by **14**, intramolecular H-bonds marked by circles. Reproduced with permission from [72,73], copyright 2002, 2005 The Royal Society of Chemistry and with permission from [71], copyright 2003 John Wiley & Sons, Inc.

binding to peculiar protonated poly AH$^+$ double stranded helix (Scheme 23) [71]. Attachment of two adenines to N5-protonated phenanthridinium completely abolished interactions with DNA and RNA due an extensively self-stacked structure but the bisuracil–phenanthridinium conjugate **14** was able to distinguish between alternating and consecutive AT sequences by peculiar combination of aromatic stacking and hydrogen-bonding interactions [73,74].

At variance to phenanthridinium–nucleobase conjugates (Scheme 23), which were not able to differentiate among mononucleotides, some bis-phenanthridinium–nucleobase conjugates provided a more convenient binding site for the nucleobase. For instance, adenine derivative **15** (Figure 6) selectively recognized the complementary nucleotide (UMP) by specific change in the UV–vis spectrum of phenanthridine subunits and high affinity [75]. Molecular modelling studies proposed a structure of the **15**–UMP complex stabilized by a set of intra- and intermolecular stacking interactions and intermolecular hydrogen bonds unique for derivative **15** interaction with UMP but not possible with other nucleotides. Moreover, mentioned bis-phenanthridinium–nucleobase conjugates also exhibited complex interactions with various ds- and ss-DNA and ds- and ss-RNA, whereby the thermal denaturation and ICD signal-based sensing was highly sensitive to the polynucleotide basepair composition and secondary structure [76]. However, the low solubility of the studied systems hampered NMR studies and the very complex set of possible interactions did not allow accurate structural explanation of observed ICD recognition.

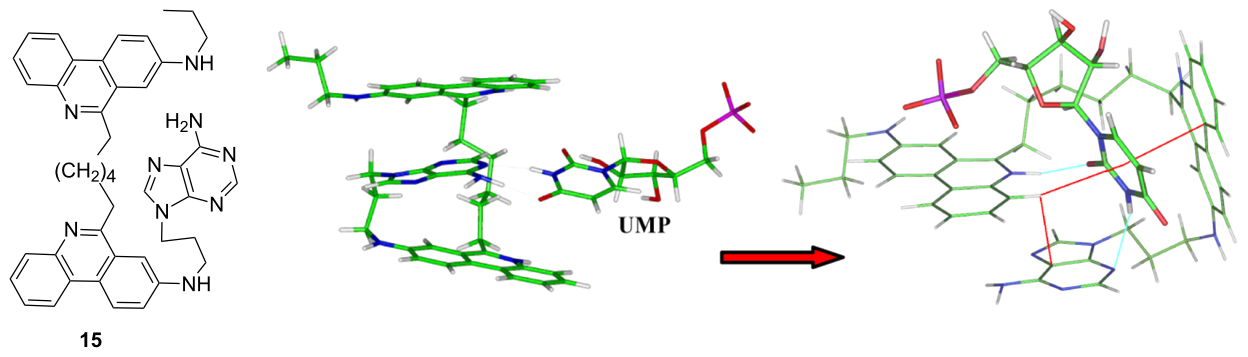



**Figure 6:** The bis-phenanthridinium–adenine derivative **15** (LEFT) showed selectivity towards complementary UMP; structure of the **15**–UMP complex (RIGHT) obtained by molecular modelling. Reproduced with permission from [75]. Copyright 2010 Elsevier Ltd.

Laborious synthetic procedures for the preparation of bis-phenanthridine–nucleobase conjugates initiated a novel, convergent and much more flexible approach relying on solid phase peptide synthesis described earlier (Figure 4). In such a manner prepared phenanthridine–thymine conjugates [77], intercalated into ds-DNA whereby binding was marginally influenced by attached thymine and the peptide backbone. More intriguing was the observed excimer fluorescence emission and the very specific CD spectrum of pentapeptide confirming the very efficient phenanthridine–thymine–phenanthridine stacking. The obtained results support efficient and predictable self-organisation of sterically crowded oligo-phenanthridine peptides (Figure 4, [68]) as well as analogues containing other (DNA and RNA binding) aromatic moieties [77], which as a proof of principle support future design of analogous peptide libraries for combinatorial approach to recognition of various DNA and RNA targets.

A structure–activity search revealed several phenanthridinium derivatives as promising binders to DNA:RNA hybrid structures [78]. Based on their previous work [79], Arya and coworkers designed neomycin–methidium conjugate **16** (Figure 7) [80], which selectively recognized the DNA:RNA hybrid duplex (poly(dA):poly(rU)) with sub-nanomolar affinity, much higher than the affinities shown for traditional aminoglycoside–nucleic acid targets. This joins the mentioned EB analogue to a small number of ligands that bind DNA:RNA hybrid structures. Latter play crucial roles in a number of biological processes (transcription, reverse transcription [79], the priming of DNA prior to replication [81], participating in different types of enzymatic activity, notably telomerases [82] and HIV RNase).

Ethidium bromide was introduced as a part of a heterogenic two-chromophore system, to take advantage of very efficient FRET energy transfer process (77%) from fluorescein to the

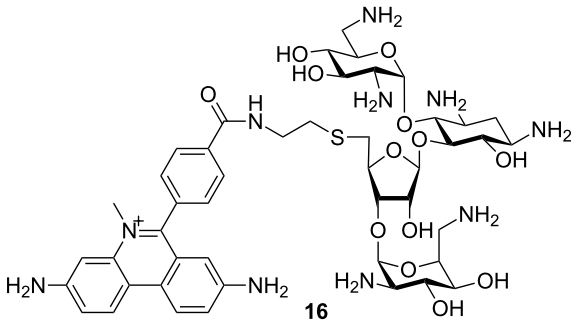



**Figure 7:** The neomycin–methidium conjugate targeting DNA:RNA hybrid structures [80].

RNA-intercalated phenanthridinium fluorophore (Figure 8, left) [83]. The resulting fluorescent dye exhibited improved ds-RNA-marker properties in comparison to other phenanthridinium analogues by means of signal brightness, signal-to-background noise and increased fluorescence half-lifetime. The same dye was also applied as convenient reporter for si-RNA (Figure 8, right) [84]. In parallel the designed and tested covalently linked ethidium bromide–ruthenium(II) complex also proved to be an imaging probe whose fluorescence intensity and lifetime changes substantially in the presence of RNA [85], thus supporting a strategy of phenanthridinium incorporation into the heterogenic two-chromophore system.

Phenanthridines are rarely combined with moieties covalently interacting with DNA and RNA. One of the most promising examples reported recently revealed that in a series of mono functional, cationic platinum(II) compounds, phenanthriplatin displayed a greater cytotoxic activity than either cisplatin or oxaliplatin despite a fact that binding to DNA induces a little distortion in the double helix (covalent adducts with DNA) [86]. The increased activity was attributed to improved cellular

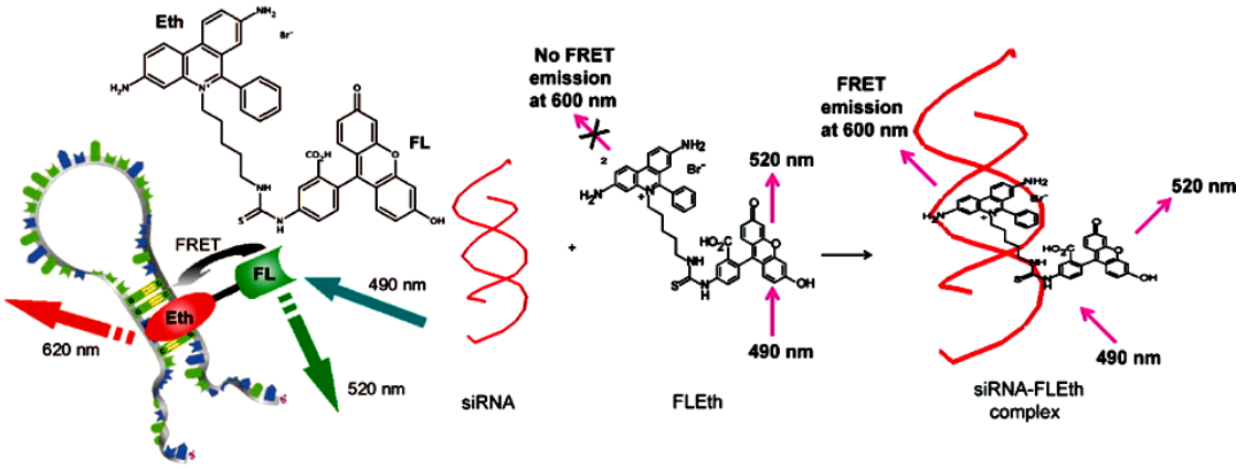



**Figure 8:** Two-colour RNA intercalating probe for cell imaging applications: Left: Chemical structure of EB-fluorescein conjugate (**FLEth**) and cartoon depicting the energy transfer process from fluorescein to the intercalated phenanthridine fluorophore. Reproduced with permission from [83]. Copyright 2008 American Chemical Society.; Right: Convenient Reporter for Small Interfering RNAs fluorophore. Reproduced with permission from [84]. Copyright 2009 American Chemical Society.

uptake and consequent inhibition of the cellular life cycle, whereby inhibition was additionally correlated to more expedient binding to nucleobases (5'-dGMP) in respect to less efficient binding of sulfur-containing nucleophiles present in resistance processes within the cell.

## Phenanthridine covalently bound to DNA and RNA

The phenanthridine aromatic moiety curvature nicely fits the shape of an average DNA and RNA basepair, while the length allows the incorporation of considerably long substituents at 3,8- positions available for attachment to DNA and RNA and/or various additional non-covalent interactions with the polynucleotide backbone.

The ethidium bromide incorporated as an artificial DNA base (**18**, Figure 9) at specific sites in duplex DNA was used to study photoinducible charge transfer processes [87]. Upon attachment to the DNA chain the phenanthridinium base (E, Figure 9) was efficiently intercalated into the DNA oligonucleotide, not disturbing the position of adjacent basepairs nor the complementary oligonucleotide strand (abasic site X). Though, ethidium 2'-deoxyribofuranoside (**17**) [88] revealed chemical instability and was therefore replaced with an acyclic linker system [87]. However, in a later work the acyclic linker was again modified to correspond by length to the deoxyribofuranoside, whereby it was proven that structural changes do not influence significantly the EB insertion into the double helix, nor EB spectroscopic properties [87,89].

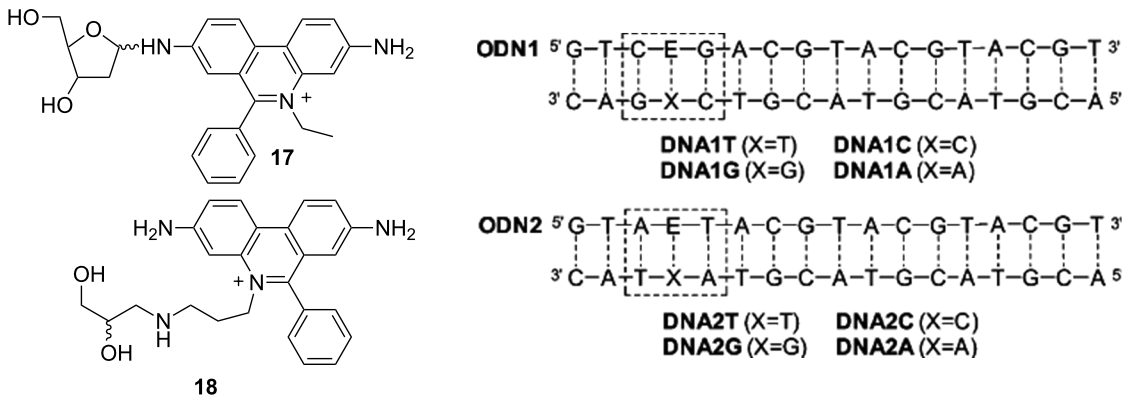



**Figure 9:** The ethidium bromide nucleosides **17** (top) and **18** (bottom). DNA duplex set 1 and 2 (E = phenanthridinium intercalation site). Reproduced with permission from [87]. Copyright 2004 American Chemical Society.

Further studies revealed that various adjacent base pairs (Figure 10, A–T in DNA1, G–C in DNA2) did not significantly influence the spectroscopic properties of the ethidium bromide [90], while usage of noncovalently bound electron-acceptor showed applicability of the phenanthridinium–DNA system for studies of electron transfer in DNA [90]. Thus the EB-nucleobase fluorescence was not sensitive to the type of naturally-occurring adjacent basepairs [90] but showed to be sensitive to major erroneous ds-DNA sites (e.g., abasic sites) [91]. Namely, by using the well-known system of EB-fluorescence quenching by 7-deazaguanine incorporated within modified oligonucleotides, it showed that the abasic site (S) either one base pair away (**DNA1-XY** and **DNA2-XY**) or two base pairs away (**DNA3-XY** and **DNA4-XY**) from the EB chromophore showed an enhanced fluorescence quenching compared to the matched duplexes [91].

Among many studies of charge transfer in DNA, several applying ethidium bromide, revealed an unexpected complexity of the process, pointing out the importance of the DNA/EB complex flexibility on the efficiency of the transfer. A study of comparatively flexible DNA/EB complex, EB covalently attached to the 5'-end of oligonucleotides, in detail described the rate and distance dependencies of charge transfer through DNA [92,93]. A more rigid type of EB-binding, whereby the EB-nucleobase was incorporated close to the centre of the DNA oligomer in combination with two different charge acceptors (7-deazaguanine as an electron hole acceptor and a 5-nitroindole as a suitable electron acceptor) [94], showed similar rates and distance dependencies for both, electron and hole transfer. The obtained results [92-94] stress the importance of DNA-basepair dynamics for the electronic transfer processes in DNA-stacks. The efficiency of transfer is rather more controlled by motions of chromophores involved in aromatic stacking of DNA-reporter complex than with rigid aryl-stacking, thus suggesting the presence of a base gating mechanism (for the here presented EB/DNA systems on the 10–100 ps scale).

## Phenanthridine analogues

One of the main incentives that increased interest in phenanthridines was the large family of naturally occurring close analogues, mostly of extended aromatic moieties (e.g., benzophenanthridines). Their distinctive biomedical properties resulted in a considerable amount of research and large number of publications, hampering their detailed description in this review. Nevertheless, several chosen examples of phenanthridine analogues will be presented.

The phenanthridine analogues, 4,9-diazapyrenium cations (very scarcely studied) [95], revealed a number of very intriguing properties upon binding to DNA and RNA. For instance, the closest analogue to ethidium bromide **19** (Figure 11) showed opposite fluorescence response upon binding to double-stranded GC-DNA and GC-RNA (quenching of emission) and AU(T) (emission increase) [96]. The only plausible structure of intercalated **19** requires parallel positioning of **19** and adjacent base pairs' long axes, consequently positioning the bulky phenyl substituents of **19** in opposite DNA grooves – thus **19** exhibits rare threading intercalation binding into double-stranded polynucleotides. Furthermore, derivative **19** formed two types of complexes with ss-RNA, a more stable one with a well organised, possibly helical structure (ICD evidence) close to

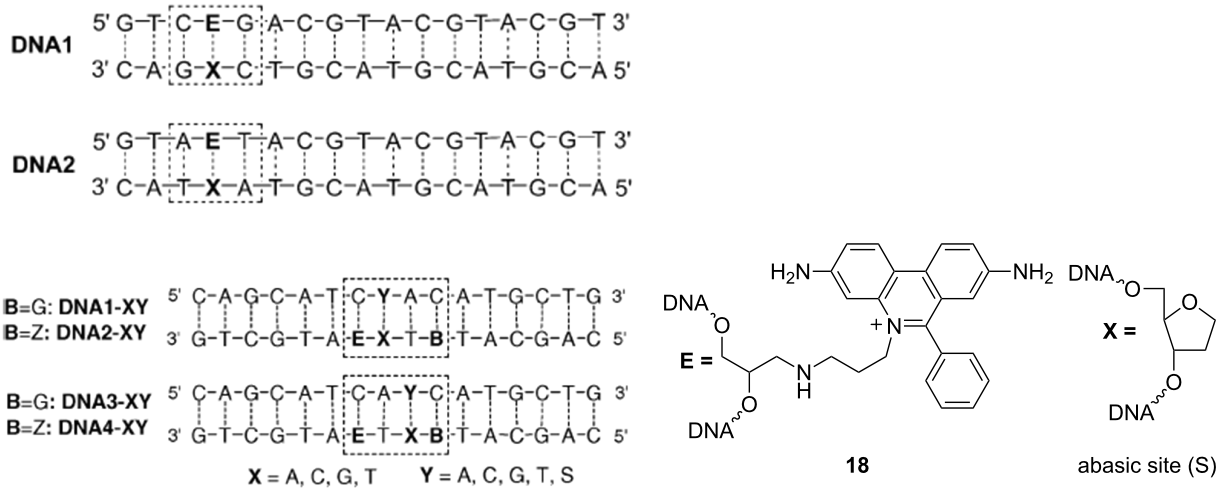


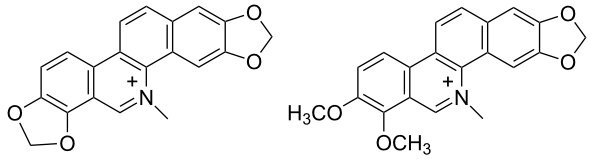

**Figure 10:** Left: various DNA duplexes; DNA1 and DNA2 used to study the impact on the adjacent basepair type on the EB fluorescence (reproduced with permission from [90], copyright 2004 John Wiley & Sons, Inc.) and **DNA1,2,3,4-XY** studying the EB fluorescence quenching by 7-deazaguanine (B) as a function of different position of abasic site (S). Reproduced with permission from [91] Copyright2005 Royal Society of Chemistry. Right: structure of incorporated EB (**18**) and of the abasic site (S).

saturation of poly(U) ($r \approx 1$) and less stable complexes with the other ss-RNA, characterised by decreased CD bands of polynucleotides. At variance to other 4,9-diazapyrenium compounds that lack the amino groups in positions 2 and 7, derivative **19** exhibited higher affinities and larger stabilisations of ds-DNA and ds-RNA probably due to additive interactions of its amino substituents within the polynucleotide binding site. All 4,9-DAP derivatives also showed considerable antiproliferative activity, interestingly only **19** having strong, micromolar activity in vitro but negligible in vivo toxic effects in mice [97]. Strong fluorescence of **19** allowed monitoring of the very efficient cellular uptake (Figure 11), upon which red colour of **19** accumulated in cell nuclei – intriguingly after only 2 hours fluorescence colour changed to yellow (Figure 11, right) and the dye distributed over the cytoplasm pointing out to the metabolic modification of the compound.

The new, easily accessible analogue, dihydroimidazophenanthridinium cation characterised by cyclic structure connecting positions 5 and 6, showed promising antiproliferative activity [3,98,99]. Molecular modelling results and some preliminary experiments suggest intercalative binding mode, however up till now interactions with various DNA and RNA were not studied in detail.

One of several reasons for the increased research on phenanthridines is the discovery of naturally occurring analogues, e.g., some protoberberine alkaloids (Figure 12, sanguinarine and chelerythrine), widely distributed in several botanical families exhibiting many therapeutic applications. Very extensive results would require a focused review, thus some examples are listed below as outline of the importance.

Most of the sanguinarine (**20**) and chelerythrine (**21**) derivatives were typical DNA and RNA intercalators [100], some of them showing also intriguing interactions with ss-RNA, poly(A) [46]. However, either intercalative binding mode or structural similarity to EB did not hamper their biomedical applications. For instance, in a series of 5-methylbenzo[*c*]phenanthridinium derivatives, based on combination of sanguinarine (**20**)/chelerythrine (**21**) structures [101], the presence of a 1-phenyl or 12-phenyl substituent on 2,3,8,9-tetramethoxy-5-ethylbenzo[*c*]phenanthridinium chloride [102] significantly enhances the antibacterial (*Staphylococcus aureus* and *Enterococcus faecalis*) activity relative to sanguinarine. Another example, using the strategy of bioactivity-guided fractionation, the bioactive compound chelerythrine (**21**, a quaternary benzo[*c*]phenanthridine alkaloid) was isolated from *Chelidonium majus L.* [103]. In addition to strong antihelmintic activity (against D. intermedius), chelerythrine also showed antimicrobial, antifungal and anti-inflammatory activity [104].

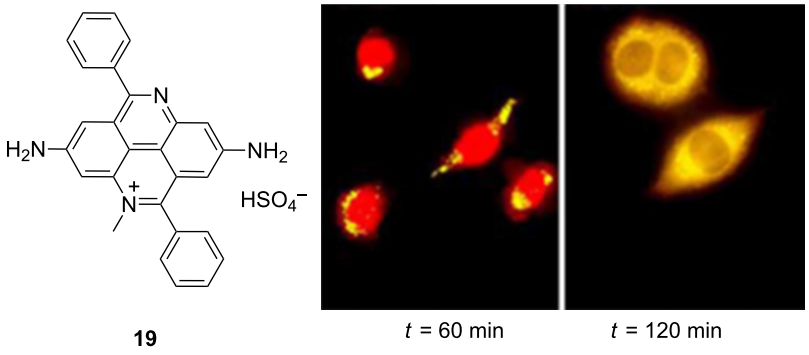



**20**, sanguinarine **21**, chelerythrine

**Figure 12:** Examples of naturally occurring phenanthridine analogues.

## Discussion of the presented results and perspectives

The data presented in this review endeavoured to stress the outstanding properties of the rather simple and, due to the substantial advance in synthetic approaches, now readily available moiety (phenanthridine). Within the last 15 years significant research efforts invested in the phenanthridine and phenanthridinium structure–DNA and RNA-binding relations resulted in a significantly advanced understanding of the chromophore system in the free form and even more important in complex with ds-DNA and ds-RNA.




**19**

**Figure 11:** Structure of 4,9-DAP derivative **19**; Rright: MIAPaCa-2 cells stained with 10 μM **19** after 60 and 120 min incubation, respectively. Magnification 630×. Reproduced with permission from [95]. Copyright 2000 Royal Society of Chemistry.

Detailed analysis of the DNA and RNA binding parameters (Table 1) revealed that contrary to the common paradigm about ethidium bromide derivatives as classical ds-DNA and ds-RNA intercalators, here presented results show a large variety of binding modes, very often the same molecule exhibiting more than one binding mode, depending on the ratio $r$ = [compound]/[polynucleotide]. Moreover, there is no set of rules which will accurately predict the dominant binding site of newly designed phenanthridine/phenanthridinium analogues. All aforementioned also hampers the prediction of the fluorimetric response, which is much more dependent on the binding mode than on the substituents attached to the chromophore. However, by the rule of thumb, if phenanthridine substituents at 3,8-positions sterically allow the intercalation into ds-DNA or ds-RNA, than a binding affinity within the micromolar range could be expected and the systematic research of Luedtke et al. [48] allows predictions of the fluorimetric response, while results of Wagenknecht et al. [88-94] as well as Turro et al. [83-

**Table 1:** Binding affinities (log *Ks*); thermal denaturation effects and proposed binding modes of chosen phenanthridine and phenanthridinium derivatives with natural and synthetic polynucleotides: ds-DNA and ds-RNA.

| Ligand | pH | log $K_s$[a]($\Delta T_m$[b]) / Binding mode[c] | | | | | |
|---|---|---|---|---|---|---|---|
| | | ctDNA | dA–dT | dAdT–dAdT | rA–rU | dG–dC | rG–rC |
| **1** | 7.5 | 5.6[d] ,(–) / IC | – | – | – | | – |
| **4** | 5 | –, (20.6) / MG | 7.4, (29.5) / MG | –, (25.0) / MG | 7.4, (35.4[e]/–1.8) / IC | 6.6, (–) / IC | – |
| | 7 | –, (5.6) / MG | 6.5, (9.8) / MG | –, (10.3) / MG | 5.7, (5.0) / NSA | 5.8, (–) / IC | |
| **5** R = $(CH_2)_4$ | 6.2 | – | 5.2, (–[f]/45.3) / IC + ESI | – | 5.7, (–[f]/21.5) / IC + ESI | – | 6.5,(–) / IC + ESI |
| **5** R = $(CH_2)_6$ | 6.2 | – | 5.7, (–[f]/42.5) / IC, NSA | – | 8.0, (–[f]/20.7) / IC, NSA | – | 6.3,(–) / IC, NSA |
| **6** R = $(CH_2)_6$ | 6.2 | – | –[g], (25.6) / GB | – | 5.3, (3.9) / GB | – | 6.5,(–) / GB |
| **7** | 6.2 | – | 6.0, (–) / GB, ESI | – | –[g] ,(–) / GB, ESI | – | 5.7,(–) / GB, ESI |
| **8** | 5 | 5.8, (25.8) / IC, ESI | –, (11.1) / IC, ESI | – | 5.1, (9.5) / IC, ESI | – | 5.8,(–) / IC, ESI |
| | 7 | 6.6, (5.7) / IC | –, (12.1) / IC | – | 6.2, (5.2) / IC | – | 5.2, (–) / IC |
| **9** (*n* = 2) | 5 | 5.5, (2[f]/22.6) / MG | 5.8 (3.3[f]/27.7) / MG | – | 5.2 (0[e]/1.1) / NSA | – | 6.6[h], (–) / NSA |
| **9** (*n* = 6) | 5 | 6.3, (2.9) / MG | 5.7, (3.4) / MG | – | 5.2 (0[e]/0.9) / NSA | – | 6.8[h], (–) / NSA |
| **9** (*n* = 9) | 5 | 6.0, (7.7) / IC | 6.1, (3.8) / IC | – | 6.0, (10.6[e]/–2.0) / IC | – | 5.0, (–) / IC |
| **10** R = Gly | 5 | 5.,5 (8.1) / IC | – | – | – | – | – |
| | 7 | 4.7, (–) / IC | – | – | – | – | – |
| **11** R = Gly | 5 | 6.9, (16.5) / IC | – | – | – | – | – |
| | 7 | 6.0, (–) / IC | – | – | – | – | – |
| **14** (bis-uracil analogue) | 5 | – | 4.6, (18.0) / PIC, HB | 5.1 / PIC | 5.0, (2[e]/12.4) / PIC | – | – |
| **14** (bis-adenine analogue) | 5 | – | 5.0, (6.0) / MG | 5.3, (1.1) / MG | 5.1, (1.5[e]/0.3) / GB | – | – |

**Table 1:** Binding affinities (log *Ks*); thermal denaturation effects and proposed binding modes of chosen phenanthridine and phenanthridinium derivatives with natural and synthetic polynucleotides: ds-DNA and ds-RNA. (continued)

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **16** | 5.5 | – | –, (7.7[i]) / IC, MB | 6, (7.4[i] ) / IC, MB | 9.3, (34.3 [i]) / IC, MB | – | – |
| **19** | 5 | 6.4, (19.6) / TIC | 5.4, (10[f]/20) / TIC, MB | 6.1 / TIC | 6.5, (35.6) / TIC | 6.1, (–) / TIC | 6.8, (–) / TIC |
| | 7 | –, (19.4) / TIC | 5.6, (11.3[f]/21.6) / TIC, MB | 5.9 / TIC | 6.5, (25.5) / TIC | 6.9[j], (–) / TIC | 6.1, (–) / TIC |
| **20** | 6.5 | – | – | – | 5.9, (19.0[k]) / IC | – | 5.5,(–) / IC |

[a]Binding constants calculated from titration data by processing according to the Scatchard equation. [b]r = [compound]/[polynucleotide] = 0.2 if not stated otherwise. [c]IC = intercalation; MG = minor groove; NSA = non-specific agglomeration; GB = undefined groove binding; BIC = bis-intercalation; ESI = electrostatic interaction; HB = hydrogen bonding; PIC = partial intercalation; MB = mixed binding mode; TIC = threading intercalation. [d]Binding constants calculated from ethidium bromide displacement experiments [52]. [e]Biphasic thermal denaturation transitions at pH 5 due to different RNA forms [62]. [f]Biphasic thermal denaturation transitions due to [b]MB, values for both transitions given when possible. [g]Not possible to calculate due to systematic deviation of experimental data from best-fitted Scatchard isotherm. [h]Cumulative binding constants for mixed binding mode. [i]r = [compound]/[polynucleotide] = 0.1. [j]Polynucleotide is poly(dG–dC)$_2$. [k]r = [compound]/[polynucleotide] = 0.5.

85] are applicable for the design of DNA charge transfer processes.

However, these rules do not apply for interactions of phenanthridine/phenanthridinium derivatives with significantly more flexible single stranded (ss-)polynucleotides, for instance ss-RNA (Table 2). The data about interactions with ss-DNA or ss-RNA are sparse and deficient, mostly determined for derivatives with substituents aiming toward particular nucleobase recognition, with very few referent compounds for any final conclusion about the binding properties of phenanthridine moiety alone. Nevertheless, binding data obtained for ethidium

**Table 2:** Binding affinities (log *Ks*) and proposed binding modes of chosen phenanthridine and phenanthridinium derivatives with synthetic polynucleotide ss-RNA.

| ligand | pH | log $K_s$[a] / Binding mode[b] | | | | |
|---|---|---|---|---|---|---|
| | | pA | pU | pG | pC | pA[c]+ ($\Delta T_m$)[d] |
| PHEN-Me[e] | 5 | 4.3 / IC | <3 / IC | – | – | |
| | 7 | 4.8 / IC | <3 | | | |
| PHEN-H$^{+}$[f] | 5 | 5.1 / IC | <3 / IC | – | – | |
| EB [96] | 5 | 3.3 / IC | <3 / IC | 3.8 / IC | 3.3 / IC | |
| | 7 | 3.9 / IC | <3 / IC | 3.1 / IC | <3 / IC | |
| **5** R = $(CH_2)_4$ | 6.2 | 4.2 / IC, NSA | 3.7 / IC, NSA | 5.3 / IC, NSA | – | – |
| **5** R = $(CH_2)_6$ | 6.2 | 3.8 / IC, NSA | 4.1 / IC, NSA | 5.8 / IC, NSA | – | – |
| **6** R = $(CH_2)_6$ | 6.2 | 4.1 / IC, NSA | 4.2 / IC, NSA | 5.4 / IC, NSA | – | – |
| **7** | 6.2 | 6.3 / BIC | 5 / BIC | 7.1 / BIC | – | – |
| **8** | 5 | 5.0 / IC | 4.5 / IC | 6.1 / IC | 5.4 / IC | – |
| | 7 | 4.6 / IC | 4.4 / IC | 5.1 / IC | 4.5 / IC | – |
| **12** (phenanthridinium-adenine) | 5 | 5.3 / IC | 4.5 / IC + HB | – | – | – |
| **12** (phenanthridinium-uracil) | 5 | 5.3 / IC | >3[g] / IC | – | – | – |

**Table 2:** Binding affinities (log *Ks*) and proposed binding modes of chosen phenanthridine and phenanthridinium derivatives with synthetic polynucleotide ss-RNA. (continued)

| | | | | | | |
|---|---|---|---|---|---|---|
| **13** | 5 | – | >3 [g]/ IC | – | – | 5.3, (3) / IC, ESI |
| | 7 | – | >3[g] / IC | – | – | 4.4 / IC |
| **14** (bis-uracil analogue) | 5 | – | – | – | – | 5.2, (8.0) / PIC, HB |
| **14** (bis-adenine analogue) | 5 | – | – | – | – | 5.1, (0.5) / MG |
| **19** | 5 | – | 3.9 / IC | 4.7 / IC | 2.9 / IC | – |
| | 7 | 4.1 / IC | 4.4 / IC | 4.5 / IC | 2.6 / IC | – |
| **20** | 6.5 | – | – | – | – | 4.5, (6.0) / IC |

[a]Titration data were processed according to the Scatchard equation. [b]IC = intercalation; MG = minor groove; NSA = non-specific agglomeration; BIC = bis-intercalation; ESI = electrostatic interaction; HB = hydrogen bonding; PIC = partial intercalation. [c]poly A at pH 5 is mostly protonated and forms ds-polynucleotide [62]. [d]r = [compound]/[polynucleotide] = 0.2, only for **20** r = 0.5. [e]8-(Propylamino)-5,6-dimethylphenanthridinium cation [71]. [f]8-(Propylamino)-6-methylphenanthridine [70]. [g]Estimated value due to less than 20% of complex formed.

bromide and 8-amino-substituted derivatives with methylated or protonated heterocyclic N5 (Table 2) show that the phenanthridine/phenanthridinium cation interacts with purine ss-sequences with affinity approximately one–two orders of magnitude lower in comparison to ds-DNA or ds-RNA, while interaction with pyrimidine ss-polynucleotides is even one order of magnitude lower. This agrees well with the aromatic stacking interactions between phenanthridine and nucleobase as dominant binding interaction (most likely intercalation), while differences between permanent (EB, PHEN-Me) and reversible (PHEN-$H^+$) positive charge do not play a significant role. Intriguingly, EB revealed an order of magnitude lower affinity toward poly(A) in comparison to PHEN-Me and PHEN-$H^+$, which could be attributed to the steric hindrance of EB at C6 and N5 positions to the optimal orientation of phenanthridinium within the intercalative binding site between adjacent nucleobases. As expected, bis-phenanthridine derivatives exhibited higher affinity due to the bis-intercalative binding mode, and in some cases show a fluorimetric recognition of a particular ss-polynucleotide (e.g., **8**) due to the fine interplay of binding interactions. Again, very scarce information about the complex structure did not allow accurate determination of binding contributions, which would clarify the observed selectivity.

Although only the current widespread biochemical application is focused on ethidium bromide/propidium iodide dyes for DNA dyeing and cell viability tests, results summarised in this review pointed out the intriguing potential of the phenanthridine/phenanthridinium system for chemical and biochemical research. Widely used fluorimetric dyes, such as cyanine derivatives, are non-fluorescent in the free state but give tremendous fluorescence emission upon binding to biomacromolecular targets. However, many of these dyes show photobleaching, a significant overlap of the absorption and emission spectrum (minor Stokes shift) and the chemical stability in stock solution is often declared by the producer to last only several months. Although the phenanthridine/phenanthridinium system in principle does not show the ideal combination of non-emissive form in the free state/very strong emission in the bound state, it has several advantages over cyanine dyes: phenanthridine/phenanthridinium fluorescence is characterised by a large Stokes shift (up to 100 nm) allowing the full use of absorption maxima as well as easy incorporation in FRET systems, high resistance to photobleaching and mostly very high chemical stability. Biomedical use in human medicine was deterred by the potential carcinogenic and mutagenic properties of some derivatives (EB and analogues) but this is recently reassessed due to the evidently innoxious treatment of African trypanosomiasis in livestock for more than 40 years (isometamidium chloride hydrochloride and ethidium bromide [105]), together with recent results on phenanthridine-based alkaloids and the promising bioactivity of phenanthriplatin [86].

All aforementioned gave the impetus to the phenanthridine/phenanthridinium system research, which made significant progress in the study of the most common phenanthridine substituent positions (3-, 5-, 6-, 8-). Nevertheless, there are still many promising targets, for instance systematic study of various substituents attached at rarely used positions (1-,2-,4-,7-,9-) would be of high interest, especially since natural phenanthridine alkaloids (Figure 12) are richly substituted on these positions and very likely owe a lot of biological activity to particular type of substituent. Several other phenanthridine characteristics such as reversible positive charge intro-

duction by protonation of the heterocyclic nitrogen (N5) were for the first time applied in designed DNA and RNA interactions, offering new biomedical applications – for instance, taking advantage of the significantly lower extracellular pH of many solid tumors [106], to which some antitumor drugs base their preferential accumulation in tumor tissue due to the weakly acidic p*Ka* value [107]. Furthermore, phenanthridine was very scarcely used as a ligand in metal coordination chemistry of biomedically oriented research, although heterocyclic nitrogen (N5) and/or various side-arm substituents offer many possibilities – as for example, very recently reported recognition of nucleotides by phenanthridine–lanthanide conjugates [108]. Finally, there are almost unlimited possibilities of phenanthridine incorporation into heterogenic fluorescent probes, taking advantage of the aforesaid phenanthridine spectrophotometric characteristics.

## Acknowledgements

Financial support by the Ministry of Science, Education, Sport of Croatia for our studies presented in this review is acknowledged.

## References

1. Browning, C. H.; Morgan, G. T.; Robb, J. V. M.; Walls, L. P. *J. Pathol. Bacteriol.* **1938,** *46,* 203–204. doi:10.1002/path.1700460121
2. Nishiwaki, H.; Miura, M.; Imai, K.; Ono, R.; Kawashima, K. *Cancer Res.* **1974,** *34,* 2699–2703.
3. Parenty, A. D. C.; Smith, L. V.; Guthrie, K. M.; Long, D.-L.; Plumb, J.; Brown, R.; Cronin, L. *J. Med. Chem.* **2005,** *48,* 4504–4506. doi:10.1021/jm050320z
4. Kock, I.; Heber, D.; Weide, M.; Wolschendorf, U.; Clement, B. *J. Med. Chem.* **2005,** *48,* 2772–2777. doi:10.1021/jm0490888
5. Krasiński, A.; Radić, Z.; Manetsch, R.; Raushel, J.; Taylor, P.; Sharpless, K. B.; Kolb, H. C. *J. Am. Chem. Soc.* **2005,** *127,* 6686–6692. doi:10.1021/ja043031t
6. Pictet, A.; Ankersmit, H. J. *Justus Liebigs Ann. Chem.* **1891,** *266,* 138–153. doi:10.1002/jlac.18912660107
7. Morgan, G. T.; Walls, L. P. *J. Chem. Soc.* **1931,** 2447–2456. doi:10.1039/jr9310002447
8. Keller, P. A. In *Science of Synthesis;* Black, D., Ed.; Georg Thieme Verlag KG: Stuttgart, 2004; Vol. 15, pp 1065–1095.
9. Lysén, M.; Kristensen, J. L.; Vedsø, P.; Begtrup, M. *Org. Lett.* **2002,** *4,* 257–259. doi:10.1021/ol0170051
10. Mamalis, P.; Petrow, V. *J. Chem. Soc.* **1950,** 703–711. doi:10.1039/jr9500000703
11. De Mayo, P.; Rigby, W. *Nature* **1950,** *166,* 1075. doi:10.1038/1661075b0
12. Narasimhan, N. S.; Chandrachood, P. S.; Shete, N. R. *Tetrahedron* **1981,** *37,* 825–827. doi:10.1016/S0040-4020(01)97703-8
13. Katritzky, A. R.; Du, W.; Matsukawa, Y.; Ghiviriga, I.; Denisenko, S. N. *J. Heterocycl. Chem.* **1999,** *36,* 927–932. doi:10.1002/jhet.5570360417
14. Katritzky, A. R.; Yang, B. *J. Heterocycl. Chem.* **1996,** *33,* 607–610. doi:10.1002/jhet.5570330313
15. Leardini, R.; Tundo, A.; Zanardi, G.; Pedulli, G. F. *Synthesis* **1985,** 107–110. doi:10.1055/s-1985-31129
16. Bowman, W. R.; Lyon, J. E.; Pritchard, G. J. *ARKIVOC* **2012,** No. vii, 210–227.
17. Budén, M. E.; Dorn, V. B.; Gamba, M.; Pierini, A. B.; Rossi, R. A. *J. Org. Chem.* **2010,** *75,* 2206–2218. doi:10.1021/jo9025918
18. McBurney, R. T.; Slawin, A. M. Z.; Smart, L. A.; Yu, Y.; Walton, J. C. *Chem. Commun.* **2011,** *47,* 7974–7976. doi:10.1039/c1cc12720a
19. Wang, Q.; Dong, X.; Xiao, T.; Zhou, L. *Org. Lett.* **2013,** *15,* 4846–4849. doi:10.1021/ol4022589
20. Wang, Y.-F.; Lonca, G. H.; Le Runigo, M.; Chiba, S. *Org. Lett.* **2014,** *16,* 4272–4275. doi:10.1021/ol501997n
21. Zhang, B.; Mück-Lichtenfeld, C.; Daniliuc, C. G.; Studer, A. *Angew. Chem., Int. Ed.* **2013,** *52,* 10792–10795. doi:10.1002/anie.201306082
22. Eisenberger, P.; Gischig, S.; Togni, A. *Chem. – Eur. J.* **2006,** *12,* 2579–2586. doi:10.1002/chem.200501052
23. Leifert, D.; Daniliuc, C. G.; Studer, A. *Org. Lett.* **2013,** *15,* 6286–6289. doi:10.1021/ol403147v
24. Linsenmeier, A. M.; Williams, C. M.; Braese, S. *Eur. J. Org. Chem.* **2013,** 3847–3856. doi:10.1002/ejoc.201300218
25. Jiang, H.; Cheng, Y.; Wang, R.; Zheng, M.; Zhang, Y.; Yu, S. *Angew. Chem., Int. Ed.* **2013,** *52,* 13289–13292. doi:10.1002/anie.201308376
26. Read, M. L.; Gundersen, L.-L. *J. Org. Chem.* **2013,** *78,* 1311–1316. doi:10.1021/jo3027033
27. Shabashov, D.; Daugulis, O. *J. Org. Chem.* **2007,** *72,* 7720–7725. doi:10.1021/jo701387m
28. Maestri, G.; Larraufie, M.-H.; Derat, E.; Ollivier, C.; Fensterbank, L.; Lacôte, E.; Malacria, M. *Org. Lett.* **2010,** *12,* 5692–5695. doi:10.1021/ol102509n
29. Peng, J.; Chen, T.; Chen, C.; Li, B. *J. Org. Chem.* **2011,** *76,* 9507–9513. doi:10.1021/jo2017108
30. Yan, L.; Zhao, D.; Lan, J.; Cheng, Y.; Guo, Q.; Li, X.; Wu, N.; You, J. *Org. Biomol. Chem.* **2013,** *11,* 7966–7977. doi:10.1039/c3ob41760c
31. Wang, W.-Y.; Feng, X.; Hu, B.-L.; Deng, C.-L.; Zhang, X.-G. *J. Org. Chem.* **2013,** *78,* 6025–6030. doi:10.1021/jo4007255
32. Pearson, R.; Zhang, S.; He, G.; Edwards, N.; Chen, G. *Beilstein J. Org. Chem.* **2013,** *9,* 891–899. doi:10.3762/bjoc.9.102
33. Candito, D. A.; Lautens, M. *Angew. Chem., Int. Ed.* **2009,** *48,* 6713–6716. doi:10.1002/anie.200902400
34. Ghosh, M.; Ahmed, A.; Dhara, S.; Ray, J. K. *Tetrahedron Lett.* **2013,** *54,* 4837–4840. doi:10.1016/j.tetlet.2013.06.089
35. Dhara, S.; Ghosh, M.; Ray, J. K. *Synlett* **2013,** *24,* 2263–2265. doi:10.1055/s-0033-1339793
36. Li, Y.; Zhu, J.; Zhang, L.; Wu, Y.; Gong, Y. *Chem. – Eur. J.* **2013,** *19,* 8294–8299. doi:10.1002/chem.201300288
37. Deb, I.; Yoshikai, N. *Org. Lett.* **2013,** *15,* 4254–4257. doi:10.1021/ol4020392
38. Tobisu, M.; Koh, K.; Furukawa, T.; Chatani, N. *Angew. Chem., Int. Ed.* **2012,** *51,* 11363–11366. doi:10.1002/anie.201206115
39. Wu, Y.; Wong, S. M.; Mao, F.; Chan, T. L.; Kwong, F. Y. *Org. Lett.* **2012,** *14,* 5306–5309. doi:10.1021/ol302489n
40. De, S.; Mishra, S.; Kakde, B. N.; Dey, D.; Bisai, A. *J. Org. Chem.* **2013,** *78,* 7823–7844. doi:10.1021/jo400890k
41. Mondal, P.; Thander, L.; Chattopadhyay, S. K. *Tetrahedron Lett.* **2012,** *53,* 1328–1331. doi:10.1016/j.tetlet.2011.12.095
42. Zhang, B.; Daniliuc, C. G.; Studer, A. *Org. Lett.* **2014,** *16,* 250–253. doi:10.1021/ol403256e

43. Nyangulu, J. M.; Hargreaves, S. L.; Sharples, S. L.; Mackay, S. P.; Waigh, R. D.; Duval, O.; Mberu, E. K.; Watkins, W. M. *Bioorg. Med. Chem. Lett.* **2005,** *15,* 2007–2010. doi:10.1016/j.bmcl.2005.02.074
44. Mandadapu, A. K.; Dathi, M. D.; Arigela, R. K.; Kundu, B. *Tetrahedron* **2012,** *68,* 8207–8215. doi:10.1016/j.tet.2012.07.067
45. Kitson, P. J.; Parenty, A. D. C.; Richmond, C. J.; Long, D.-L.; Cronin, L. *Chem. Commun.* **2009,** 4067–4069. doi:10.1039/b905383b
46. Giri, P.; Kumar, G. S. *Mol. BioSyst.* **2010,** *6,* 81–88. doi:10.1039/B910706A
47. Duhamel, J.; Kanyo, J.; Dinter-Gottlieb, G.; Lu, P. *Biochemistry* **1996,** *35,* 16687–16697. doi:10.1021/bi9610919
48. Luedtke, N. W.; Liu, Q.; Tor, Y. *Chem. – Eur. J.* **2005,** *11,* 495–508. doi:10.1002/chem.200400559
49. Jain, S. C.; Sobell, H. M. *J. Biomol. Struct. Dyn.* **1984,** *1,* 1179–1194. doi:10.1080/07391102.1984.10507511
50. Kubař, T.; Hanus, M.; Ryjáček, F.; Hobza, P. *Chem. – Eur. J.* **2006,** *12,* 280–290. doi:10.1002/chem.200500725
51. Pace, T. C. S.; Bohne, C. *Adv. Phys. Org. Chem.* **2007,** *42,* 167–223. doi:10.1016/S0065-3160(07)42004-4
52. Luedtke, N. W. *RNA Affinity and Specificity of Modified Aminoglycosides, Metal Complexes, and Intercalating Agents that Target the HIV-1 Rev Response Element*. Ph.D. Thesis, University of California, San Diego, USA, 2003.
53. Luedtke, N. W.; Liu, Q.; Tor, Y. *Bioorg. Med. Chem.* **2003,** *11,* 5235–5247. doi:10.1016/j.bmc.2003.08.006
54. Olmsted, J., III; Kearns, D. R. *Biochemistry* **1977,** *16,* 3647–3654. doi:10.1021/bi00635a022
55. Garbett, N. C.; Hammond, N. B.; Graves, D. E. *Biophys. J.* **2004,** *87,* 3974–3981. doi:10.1529/biophysj.104.047415
56. Prunkl, C.; Pichlmaier, M.; Winter, R.; Kharlanov, V.; Rettig, W.; Wagenknecht, H.-A. *Chem. – Eur. J.* **2010,** *16,* 3392–3402. doi:10.1002/chem.200902823
57. Bailly, C.; Arafa, R. K.; Tanious, F. A.; Laine, W.; Tardy, C.; Lansiaux, A.; Colson, P.; Boykin, D. W.; Wilson, W. D. *Biochemistry* **2005,** *44,* 1941–1952. doi:10.1021/bi047983n
58. Radić Stojković, M.; Miljanić, S.; Mišković, K.; Glavaš-Obrovac, L.; Piantanida, I. *Mol. BioSyst.* **2011,** *7,* 1753–1765. doi:10.1039/c1mb05030c
59. Malojčić, G.; Piantanida, I.; Marinić, M.; Žinić, M.; Marjanović, M.; Kralj, M.; Pavelić, K.; Schneider, H.-J. *Org. Biomol. Chem.* **2005,** *3,* 4373–4381. doi:10.1039/b509094f
60. Piantanida, I.; Palm, B. S.; Čudić, P.; Žinić, M.; Schneider, H.-J. *Tetrahedron Lett.* **2001,** *42,* 6779–6783. doi:10.1016/S0040-4039(01)01386-7
61. Piantanida, I.; Palm, B. S.; Čudić, P.; Žinić, M.; Schneider, H.-J. *Tetrahedron* **2004,** *60,* 6225–6231. doi:10.1016/j.tet.2004.05.009
62. Sucheck, S. J.; Wong, C.-H. *Curr. Opin. Chem. Biol.* **2000,** *4,* 678–686. doi:10.1016/S1367-5931(00)00142-3
63. Thomas, J. R.; Hergenrother, P. J. *Chem. Rev.* **2008,** *108,* 1171–1224. doi:10.1021/cr0681546
64. Islam, M. M.; Kumar, G. S. *Biochim. Biophys. Acta, Gen. Subj.* **2009,** *1790,* 829–839. doi:10.1016/j.bbagen.2009.05.015
65. Fitzgerald, M. E.; Drohat, A. C. *Chem. Biol.* **2008,** *15,* 203–204. doi:10.1016/j.chembiol.2008.03.001
66. Radić Stojković, M.; Piantanida, I. *Tetrahedron* **2008,** *64,* 7807–7814. doi:10.1016/j.tet.2008.05.142
67. Radić Stojković, M.; Marczi, S.; Glavaš-Obrovac, L.; Piantanida, I. *Eur. J. Med. Chem.* **2010,** *45,* 3281–3292. doi:10.1016/j.ejmech.2010.04.006
68. Dukši, M.; Baretić, D.; Čaplar, V.; Piantanida, I. *Eur. J. Med. Chem.* **2010,** *45,* 2671–2676. doi:10.1016/j.ejmech.2010.02.017
69. Jain, N.; Francis, S.; Friedman, S. H. *Bioorg. Med. Chem. Lett.* **2012,** *22,* 4844–4848. doi:10.1016/j.bmcl.2012.05.041
70. Tumir, L.-M.; Piantanida, I.; Novak, P.; Žinić, M. *J. Phys. Org. Chem.* **2002,** *15,* 599–607. doi:10.1002/poc.486
71. Tumir, L.-M.; Piantanida, I.; Juranović-Cindrić, I.; Hrenar, T.; Meić, Z.; Žinić, M. *J. Phys. Org. Chem.* **2003,** *16,* 891–899. doi:10.1002/poc.680
72. Juranović, I.; Meić, Z.; Piantanida, I.; Tumir, L.-M.; Žinić, M. *Chem. Commun.* **2002,** 1432–1433. doi:10.1039/b202615e
73. Tumir, L.-M.; Piantanida, I.; Juranović, I.; Meić, Z.; Tomić, S.; Žinić, M. *Chem. Commun.* **2005,** 2561–2563. doi:10.1039/b500617a
74. Tumir, L.-M.; Piantanida, I.; Žinić, M.; Juranović-Cindrić, I.; Meić, Z.; Kralj, M.; Tomić, S. *Eur. J. Med. Chem.* **2006,** *41,* 1153–1166. doi:10.1016/j.ejmech.2006.05.005
75. Tumir, L.-M.; Grabar, M.; Tomić, S.; Piantanida, I. *Tetrahedron* **2010,** *66,* 2501–2513. doi:10.1016/j.tet.2010.01.063
76. Grabar-Branilović, M.; Tomić, S.; Tumir, L.-M.; Piantanida, I. *Mol. BioSyst.* **2013,** *9,* 2051–2062. doi:10.1039/c3mb25578f
77. Dukši, M.; Baretić, D.; Piantanida, I. *Acta Chim. Slov.* **2012,** *59,* 464–472.
78. Shaw, N. N.; Arya, D. P. *Biochimie* **2008,** *90,* 1026–1039. doi:10.1016/j.biochi.2008.04.011
79. Arya, D. P.; Coffee, R. L., Jr.; Willis, B.; Abramovitch, A. I. *J. Am. Chem. Soc.* **2001,** *123,* 5385–5395. doi:10.1021/ja003052x
80. Shaw, N. N.; Xi, H.; Arya, D. P. *Bioorg. Med. Chem. Lett.* **2008,** *18,* 4142–4145. doi:10.1016/j.bmcl.2008.05.090
81. Ogawa, T.; Okazaki, T. *Annu. Rev. Biochem.* **1980,** *49,* 421–457. doi:10.1146/annurev.bi.49.070180.002225
82. Francis, R.; West, C.; Friedman, S. H. *Bioorg. Chem.* **2001,** *29,* 107–117. doi:10.1006/bioo.2000.1196
83. Stevens, N.; O'Connor, N.; Vishwasrao, H.; Samaroo, D.; Kandel, E. R.; Akins, D. L.; Drain, C. M.; Turro, N. J. *J. Am. Chem. Soc.* **2008,** *130,* 7182–7183. doi:10.1021/ja8008924
84. van der Wiel, I. M.; Cheng, J.; Koukiekolo, R.; Lyn, R. K.; Stevens, N.; O'Connor, N.; Turro, N. J.; Pezacki, J. P. *J. Am. Chem. Soc.* **2009,** *131,* 9872–9873. doi:10.1021/ja902636m
85. O'Connor, N. A.; Stevens, N.; Samaroo, D.; Solomon, M. R.; Marti, A. A.; Dyer, J.; Vishwasrao, H.; Akins, D. L.; Kandel, E. R.; Turro, N. J. *Chem. Commun.* **2009,** 2640–2642. doi:10.1039/b900290a
86. Park, G. Y.; Wilson, J. J.; Song, Y.; Lippard, S. J. *Proc. Natl. Acad. Sci. U. S. A.* **2012,** *109,* 11987–11992. doi:10.1073/pnas.1207670109
87. Huber, R.; Amann, N.; Wagenknecht, H.-A. *J. Org. Chem.* **2004,** *69,* 744–751. doi:10.1021/jo0355404
88. Amann, N.; Wagenknecht, H.-A. *Tetrahedron Lett.* **2003,** *44,* 1685–1690. doi:10.1016/S0040-4039(03)00025-X
89. Valis, L.; Wagenknecht, H.-A. *Synlett* **2007,** 2111–2115. doi:10.1055/s-2007-984899
90. Amann, N.; Huber, R.; Wagenknecht, H.-A. *Angew. Chem., Int. Ed.* **2004,** *43,* 1845–1847. doi:10.1002/anie.200353153
91. Valis, L.; Amann, N.; Wagenknecht, H.-A. *Org. Biomol. Chem.* **2005,** *3,* 36–38. doi:10.1039/b414672g
92. Kelley, S. O.; Barton, J. K. *Chem. Biol.* **1998,** *5,* 413–425. doi:10.1016/S1074-5521(98)90158-2
93. Wan, C.; Fiebig, T.; Kelley, S. O.; Treadway, C. R.; Barton, J. K.; Zewail, A. H. *Proc. Natl. Acad. Sci. U. S. A.* **1999,** *96,* 6014–6019. doi:10.1073/pnas.96.11.6014

94. Valis, L.; Wang, Q.; Raytchev, M.; Buchvarov, I.; Wagenknecht, H.-A.; Fiebig, T. *Proc. Natl. Acad. Sci. U. S. A.* **2006,** *103,* 10192–10195. doi:10.1073/pnas.0600957103
95. Piantanida, I.; Tomišić, V.; Žinić, M. *J. Chem. Soc., Perkin Trans. 2* **2000,** 375–383. doi:10.1039/a905306i
96. Piantanida, I.; Palm, B. S.; Žinić, M.; Schneider, H.-J. *J. Chem. Soc., Perkin Trans. 2* **2001,** 1808–1816. doi:10.1039/b103214n
97. Marczi, S.; Glavaš-Obrovac, L.; Belovari, T.; Stojković, R.; Ivanković, S.; Šerić, V.; Piantanida, I.; Žinić, M. *Cancer Chemother. Pharmacol.* **2008,** *62,* 595–604. doi:10.1007/s00280-007-0643-0
98. Richmond, C. J.; Eadie, R. M.; Parenty, A. D. C.; Cronin, L. *J. Org. Chem.* **2009,** *74,* 8196–8202. doi:10.1021/jo901622e
99. Smith, L. V.; Parenty, A. D. C.; Guthrie, K. M.; Plumb, J.; Brown, R.; Cronin, L. *ChemBioChem* **2006,** *7,* 1757–1763. doi:10.1002/cbic.200600205
100. Basu, P.; Bhowmik, D.; Kumar, G. S. *J. Photochem. Photobiol., B* **2013,** *129,* 57–68. doi:10.1016/j.jphotobiol.2013.09.011
101. Slaninová, I.; Pěnčiková, K.; Urbanová, J.; Slanina, J.; Táborská, E. *Phytochem. Rev.* **2014,** *13,* 51–68. doi:10.1007/s11101-013-9290-8
102. Parhi, A.; Kelley, C.; Kaul, M.; Pilch, D. S.; LaVoie, E. J. *Bioorg. Med. Chem. Lett.* **2012,** *22,* 7080–7083. doi:10.1016/j.bmcl.2012.09.097
103. Li, X.-L.; Yao, J.-Y.; Zhou, Z.-M.; Shen, J.-Y.; Ru, H.-s.; Liu, X.-L. *Parasitol. Res.* **2011,** *109,* 247–252. doi:10.1007/s00436-011-2320-9
104. Zdarilova, A.; Malikova, J.; Dvorak, Z.; Ulrichova, J.; Simanek, V. *Chem. Listy* **2006,** *100,* 30–41.
105. Boibessot, I.; Turner, C. M. R.; Watson, D. G.; Goldie, E.; Connel, G.; McIntosh, A.; Grant, M. H.; Skellern, G. G. *Acta Trop.* **2002,** *84,* 219–228. doi:10.1016/S0001-706X(02)00188-2
106. Gillies, R. J.; Robey, I.; Gatenby, R. A. *J. Nucl. Med.* **2008,** *49* (Suppl. 2), 245–425. doi:10.2967/jnumed.107.047258
107. Raghunand, N.; Gillies, R. J. *Drug Resist. Updates* **2000,** *3,* 39–47. doi:10.1054/drup.2000.0119
108. Weitz, E. A.; Chang, J. Y.; Rosenfield, A. H.; Morrow, E. A.; Pierre, V. C. *Chem. Sci.* **2013,** *4,* 4052–4060. doi:10.1039/c3sc51583d

## License and Terms

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

# A comparative study of the interactions of cationic hetarenes with quadruplex-DNA forming oligonucleotide sequences of the insulin-linked polymorphic region (ILPR)

Darinka Dzubiel, Heiko Ihmels*, Mohamed M. A. Mahmoud and Laura Thomas

Full Research Paper

Open Access

Address:
Department Chemie-Biologie, Universität Siegen, Adolf-Reichwein-Str. 2, 57068 Siegen, Germany

Email:
Heiko Ihmels* - ihmels@chemie.uni-siegen.de

* Corresponding author

Keywords:
DNA ligands; fluorescent probes; ILPR; nucleic acids; quadruplex DNA

*Beilstein J. Org. Chem.* **2014,** *10,* 2963–2974.
doi:10.3762/bjoc.10.314

Received: 31 August 2014
Accepted: 26 November 2014
Published: 11 December 2014

This article is part of the Thematic Series "Nucleic acid chemistry". Author names in alphabetical order does not reflect the specific contribution of each author.

Guest Editor: H.-A. Wagenknecht



## Abstract

The interactions of the ILPR sequence (ILPR = "insulin-linked polymorphic region") **a2** [d(ACAG$_4$TGTG$_4$ACAG$_4$TGTG$_4$)] with [2.2.2]heptamethinecyanine derivatives **1a–e** and with the already established quadruplex ligands coralyne (**2**), 3,3′-[2,6-pyridinediylbis(carbonylimino)]bis[1-methylquinolinium] (**3**), 4,4′,4″,4‴-(21*H*,23*H*-porphine-5,10,15,20-tetrayl)tetrakis[1-methylpyridinium] (**4**), naphtho[2,1-*b*:3,4-*b*′:6,5-*b*″:7,8-*b*‴]tetraquinolizinium (**5**) and thiazole orange (**6**) were studied. It is demonstrated with absorption, fluorescence and CD spectroscopy that all investigated ligands bind with relatively high affinity to the ILPR-quadruplex DNA **a2** (0.2–5.5 × $10^6$ $M^{-1}$) and that in most cases the binding parameters of ligand-ILPR complexes are different from the ones observed with other native quadruplex-forming DNA sequences.

## Introduction

The “insulin-linked polymorphic region” (ILPR) is a physiologically relevant G-rich DNA sequence that consists of repetitive DNA units with varying length and sequence [1]. The ILPR is located in the promoter region of the human insulin gene and is proposed to control the expression of the latter. It was shown that insulin-dependent diabetes mellitus (IDDM, type-I diabetes) is associated with the number of these minisatellites. Specifically, the number of these repeating units varies from ca. 160 in healthy humans to ca. 40 in IDDM patients [2]. From the 14 known repeating units of ILPR the variants **a** [d(ACAGGGGTGTGGGG)], **b** [d(ACAGGGGTCTGGGG)] and **c** [d(ACAGGGGTCCTGGGG)] occur most frequently; and **a** has the highest transcriptional activity [2,3]. It was shown that the sequences **a**–**c** form stable G-quadruplex structures in vitro

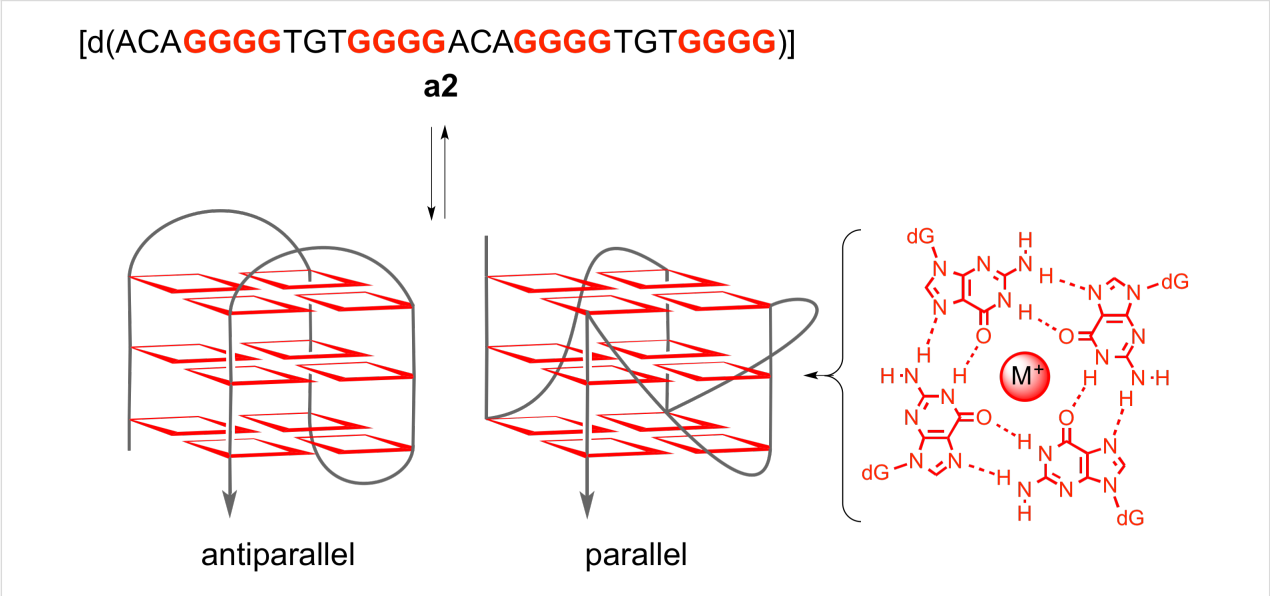
and that this tendency is even enhanced by binding of the insulin protein [4-8]. In addition, it was demonstrated that the nucleotide sequence **a2**, [d(ACAG$_4$TGTG$_4$ACAG$_4$TGTG$_4$)], has parallel and antiparallel quadruplex forms that coexist under physiological conditions (Scheme 1) [9]. Notably, the biological relevance of ILPR quadruplexes has been suggested considering recent discoveries on the influence of quadruplex formation on the function of nucleic acids and the observation that insulin associates with quadruplex DNA [10]. In particular, it was assumed that the formation of quadruplex structures affects the transcriptional activity, because the position of the ILPR in the promoter region of the gene represents a stake close to the transcription process [9,11,12]. Furthermore, quadruplex structures may interfere with the replication process because of their thermal and mechanical stability. Namely, the force required to unfold the quadruplex is larger than the one to block helicases [9,13], such that quadruplex structures reduce the activity of these enzymes to greater extent than the corresponding duplex DNA [14]. There is also evidence that longer ILPR sequences may form multiple quadruplex structures that interact with each other, and it was speculated that the observed increase in replication errors in minisatellite regions is related with these higher-order DNA structures [15].

It has been demonstrated throughout the last decade that quadruplex DNA-binding ligands have a large potential to increase its stability towards unfolding and to influence the equilibrium between different quadruplex forms by stabilizing one particular quadruplex conformation [16,17]. Therefore, it is proposed that such ligands may interfere with physiological processes that involve quadruplex DNA. In this context, it is remarkable that systematic studies on the interaction of ligands with ILPR-DNA are rather rare. To the best of our knowledge there exists one study to develop a fluorescence screening method for the identification of selective quadruplex ligands in which, along with other sequences, an ILPR sequence has been used [18,19]. To fill this gap of knowledge, we investigated the interactions of a series of quadruplex ligands with the representative ILPR sequence **a2**. As we have demonstrated already that the [2.2.2]heptamethinecyanine dye binds selectively to quadruplex DNA and that its interaction with the quadruplex is indicated by a drastic emission light-up effect [20,21], we chose the [2.2.2]heptamethinecyanine derivatives **1a–e** [22] as ligands for this study. For a better comparison of data we also included the already established quadruplex ligands **2**–**6** [23-27] in this study (Figure 1).

# Results

## Thermal DNA-denaturation experiments

The temperature-dependent structural changes of ILPR-DNA were analyzed by CD-spectroscopic and fluorimetric analysis. CD spectra were obtained from a solution of the ILPR quadruplex-forming oligonucleotide **a2** in $K^+$-containing buffer (95 mM) in a temperature range from 20 to 95 °C. The oligonucleotide **a2** showed CD signals that are characteristic of a mixture of parallel ($\lambda_{max}$ = 265 nm; $\lambda_{min}$ = 235 nm) and antiparallel ($\lambda_{max}$ = 295 nm) quadruplex structures (Figure 2A) [9,28]. With increasing temperature, the intensity of the maxima at 210 and 265 nm as well as the ones of the minimum at 235 nm decrease at $T$ > 60 °C. On the other hand, the intensity of the signal at 295 nm slightly increases at $T$ > 60 °C and then decreases at $T$ > 80 °C (Figure 1A and 1B). These results are in




**Scheme 1:** Equilibrium between single-stranded ILPR-DNA **a2** and its parallel and antiparallel quadruplex form.

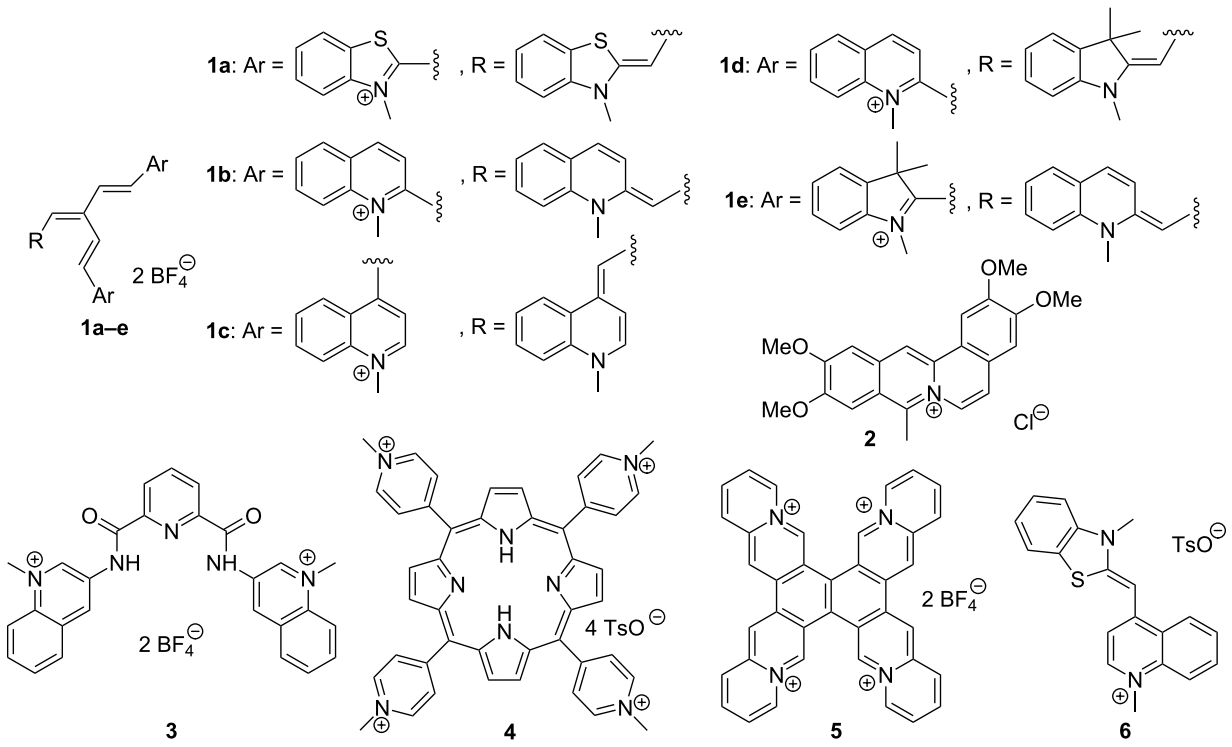



**Figure 1:** Structures of ligands **1–6** used in this study.

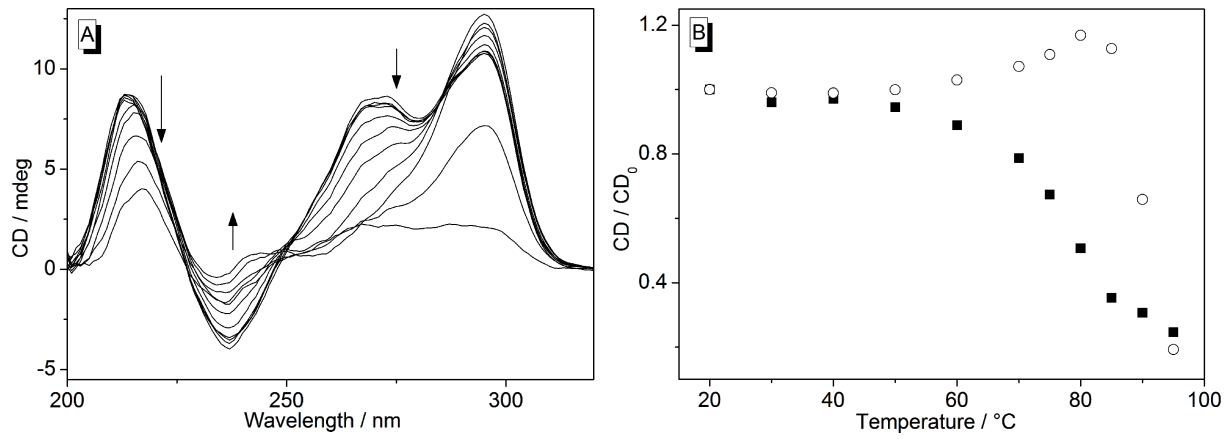



**Figure 2:** A) CD spectra of ILPR-DNA **a2** (20 μM) at different temperatures in potassium phosphate buffer (95 mM, pH 7.0). Arrows indicate the development of the bands with increasing temperature. B) Plot of the intensity changes ($CD/CD_0$) of the CD signals at 265 nm (■) and 295 nm (○) versus temperature.

agreement with literature data [9]; however, in the latter case the signal at 265 nm increased at 60 °C < $T$ > 80 °C instead of the signal at 290 nm, presumably because different buffer solutions were used in the latter study. It should be noted that the CD spectrum of **a2** in potassium phosphate buffer (95 mM $K^+$) changes slightly with time. Specifically, the intensity of the CD bands at 210, 235 and 265 nm decreases slowly within 48 h, whereas the maximum at 295 nm remains essentially the same in this time range (Figure 3).

The thermal stability of the ILPR-DNA with respect to unfolding was also examined by emission spectroscopy in $K^+$- and $Na^+$-containing buffer, i.e., the melting temperature, $T_m$, of the dye-labelled quadruplex-forming ILPR sequence **Fa2T**

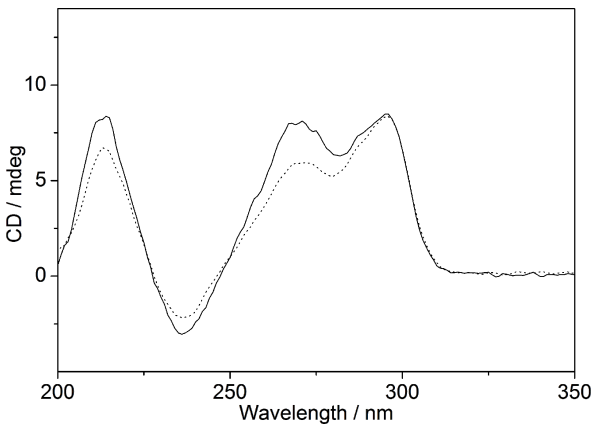



**Figure 3:** CD spectra of ILPR-sequence **a2** (20 μM) in potassium phosphate buffer (95 mM, pH 7.0; — : immediately after the preparation of the DNA stock solution; ···· : 2 d after the preparation).

[fluorescein-d(ACAG$_4$TGTG$_4$ACAG$_4$TGTG$_4$)-tetramethylrhodamine] was determined by fluorimetric monitoring of the temperature-dependent Förster resonance energy transfer (FRET) between the dyes [29]. In sodium cacodylate buffer (10 mM $Na^+$, 10 mM $K^+$, 90 mM $Li^+$) the melting curve of the DNA has a weak transition at 50 °C and a more pronounced one at $T_m$ = 71.0 °C (Figure 4). The results of the melting experiments always refer to the latter transition. In potassium phosphate buffer (95 mM $K^+$) the melting temperature of the DNA is $T_m$ = 87.8 °C. This temperature is relatively high, so that further stabilization by a ligand leads to a $T_m$ close to the boiling point of the solvent. Therefore, studies of ligand–DNA complexes were not performed in this medium.

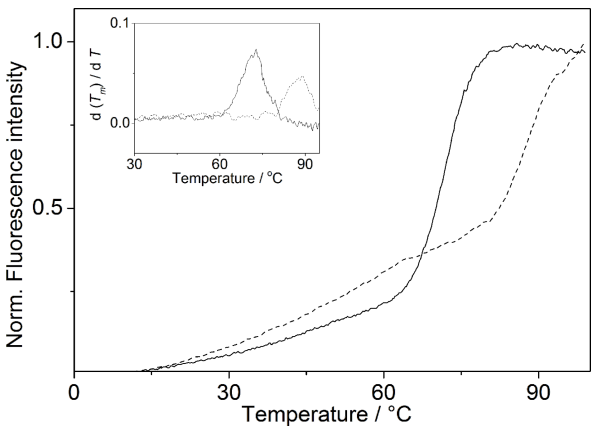



**Figure 4:** Fluorimetric monitoring of thermal DNA-denaturation of the ILPR quadruplex **Fa2T** (0.2 μM DNA concentration in oligonucleotide) in potassium phosphate buffer (····, 95 mM, pH 7.0) and KCl-LiCl-Na-cacodylate buffer (—, 10 mM $K^+$, 90 mM $Li^+$, pH 7.2). Inset: Plot of the first derivative *f'* (*T*) of the melting curves.

The influence of the ligands **1**–**6** on the melting temperature of ILPR-DNA was determined by fluorimetric monitoring of the DNA melting curves of **Fa2T** at different ligand–DNA ratios, LDR (Table 1, Figures S1 and S2 in Supporting Information File 1). Most of the tested ligands, i.e. **1a**, **1c**, **1d**, **2**, **4**, **5**, and **6**, induced only a moderate increase of the melting temperature of the ILPR-quadruplex ($\Delta T_m$ = ca. 2–6 °C). Nevertheless, the addition of the cyanine derivative **1b** or the bis-quinolinium derivative **3** induced a significantly larger shift of the quadruplex melting temperature (**1b**: $\Delta T_m$ = 14.5 °C; **3**: 14.9 °C; LDR = 5.0).

**Table 1:** Shifts of melting temperature, $\Delta T_m$, of quadruplex-DNA **Fa2T** induced by ligands **1**–**6** as determined from fluorimetric thermal DNA denaturation studies.

| | $\Delta T_m$ [°C] [a] | | |
|---|---|---|---|
| LDR | 1.3 | 2.5 | 5.0 |
| **1a** | 1.0 | 1.9 | 4.9 |
| **1b** | 0.8 | 7.1 | 14.5 |
| **1c** | 0.1 | 1.1 | 3.5 |
| **1d** | 0.5 | 4.0 | 3.7 |
| **1e** | 1.3 | 2.2 | 3.4 |
| **2** | 0.7 | 1.7 | 3.1 |
| **3** | 3.9 | 6.4 | 14.9 |
| **4** | 1.4 | 2.1 | 6.3 |
| **5** | 0.1 | 1.5 | 5.8 |
| **6** | 0.7 | 0.8 | 1.8 |

[a] $\Delta T_m$ of **Fa2T**; $c_{DNA}$ = 0.2 μM (in oligonucleotide); KCl-LiCl-Na-cacodylate buffer (10 mM $K^+$, 90 mM $Li^+$, pH 7.2); $\lambda_{ex}$ = 470 nm; $\lambda_{em}$ = 515 nm; estimated error: ±0.5 °C of the given data.

## Photometric and fluorimetric titrations

The interactions of the ligands **1**–**6** with ILPR-DNA **a2** were further analyzed with photometric and fluorimetric titrations (Figure 5 and Figure 6, Table 2). In general, a hypochromic effect and a bathochromic shift of the absorption maximum of the ligand were observed upon addition of DNA (Supporting Information File 1, Figures S3 and S5). In the case of coralyne (**2**) and the tetraazoniahetarene derivative **5** this effect is only weakly pronounced. Only during the titration of the bis-quinolinium derivative **3** an isosbestic point developed at 364 nm.

The addition of ILPR-DNA **a2** to the cyanine derivatives **1a**–**e** and **6** resulted in a strong increase of the emission intensity (light-up effect) with increasing concentration of **a2** (Figure 5 and Figure 6E). This effect was most pronounced in the case of thiazole orange (**6**) whose emission intensity increases by a factor of $I/I_0$ = 1766. Within the series of heptamethine cyanine dyes **1a**–**e**, the derivative **1d** exhibits the largest light-up factor

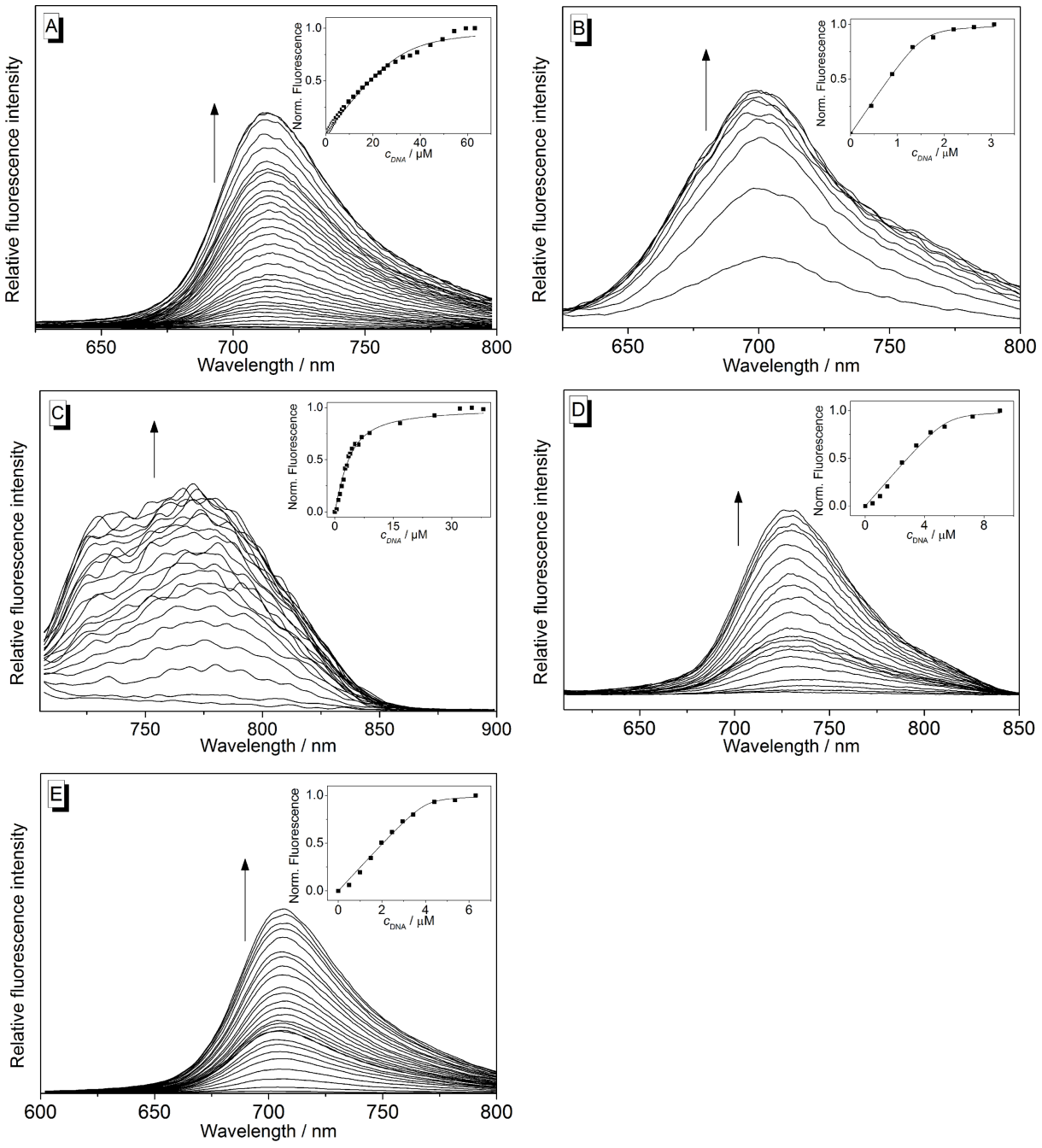



**Figure 5:** Fluorimetric titration of **1a** (A), **1b** (B), **1c** (C), **1d** (D) and **1e** (E) with **a2** in potassium phosphate buffer (95 mM, pH 7.0); A, B, E: $c_{Lig}$ = 5.0 µM; C, D: $c_{Lig}$ = 10 µM; A: $\lambda_{ex}$ = 385 nm; B: $\lambda_{ex}$ = 364 nm; C: $\lambda_{ex}$ = 431 nm; D, E: $\lambda_{ex}$ = 580 nm. Arrows indicate the development of the bands with increasing DNA concentration. Inset: Plot of the normalized emission intensity versus DNA concentration. Lines denote the best fit of experimental data to the theoretical model.

with $I/I_0$ = 128. In contrast, the emission of coralyne (**2**), bisquinolinium derivative **3** and the tetraazoniahetarene derivative **5** was quenched upon addition of **a2** (Figure 6A, 6B, and 6D; Table 2). The data of the fluorimetric titrations was used to estimate the binding constants $K_b^{a2}$ from a fit of the experimentally determined binding isotherms to the theoretical model [31] with resulting binding constants in the range of 0.2–5.5 × $10^6$ $M^{-1}$ (Table 2). Notably, the development of the emission intensity of ligands **1d** and **1e** showed two distinctly different trends during the titration with **a2**, which indicates two different binding modes (Supporting Information File 1, Figure S4). Specifically, at LDR values <1.1 (**1d**) or <0.8 (**1e**) the

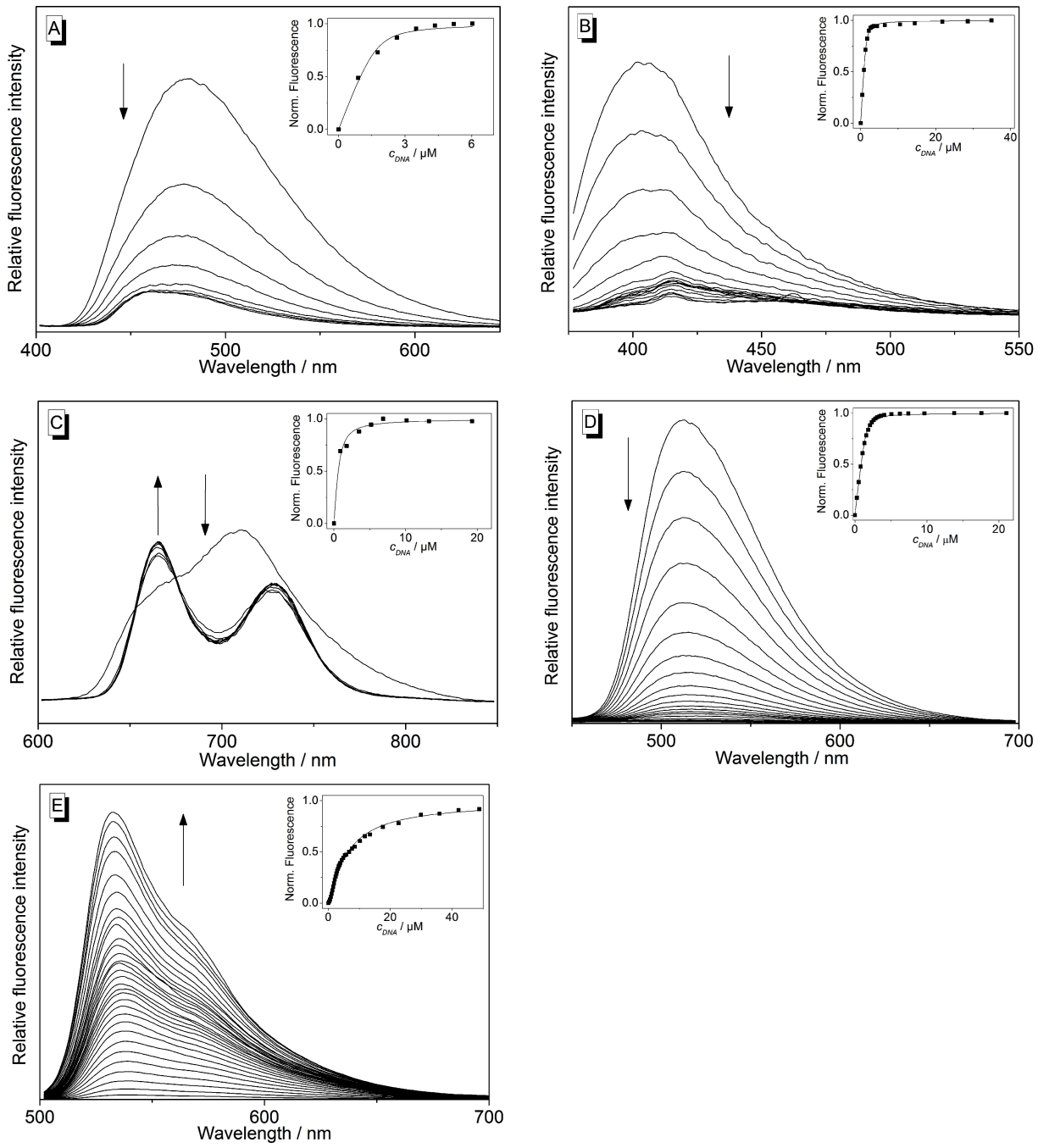



**Figure 6:** Fluorimetric titration of **2** (A), **3** (B), **4** (C), **5** (D) and **6** (E) with **a2** in potassium phosphate buffer (95 mM, pH 7.0); A, B, D, E: $c_{Lig}$ = 5.0 μM; C: $c_{Lig}$ = 10 μM; A, D: $\lambda_{ex}$ = 385 nm; B: $\lambda_{ex}$ = 364 nm; C: $\lambda_{ex}$ = 431 nm; E: $\lambda_{ex}$ = 490 nm. Inset: Plot of the change of the emission intensity versus DNA concentration. Lines denote the best fit of experimental data to the theoretical model. Arrows indicate the development of the bands with increasing DNA concentration.

emission intensity is significantly less quenched as compared to the beginning of the titration. The binding constants of the ligand–DNA complexes formed at the beginning of the titrations were estimated assuming that the emission intensities of **1d** und **1e** at LDR = 1.1 and 0.8 essentially represent the one of the bound ligand in this particular binding mode. Hence, the binding constants were determined from the respective binding isotherms (inset in Figure 5D and 5E; **1d**: $K_b^{\mathbf{a2}} = 5.5 \times 10^6\ M^{-1}$; **1e**: $K_b^{\mathbf{a2}} = 1.7 \times 10^7\ M^{-1}$). For comparison, fluorimetric titrations of the telomeric quadruplex **22AG** [5'-A(GGGTTA)$_3$GGG-3'] to the cyanine derivatives **1b–e** were performed and the binding constants were obtained from the

**Table 2:** Absorption and emission properties of ligands **1**–**6** in the absence and presence of ILPR-DNA **a2**, and binding constants $K_b$.

| Ligand | $\lambda_{Abs}$ [nm][a] | | $\lambda_{em}$ [nm][b] | $I/I_0$ [c] | $K_b^{a2}$ [$10^6$ $M^{-1}$][d] | $K_b^{22AG}$ [$10^6$ $M^{-1}$][e] |
|---|---|---|---|---|---|---|
| | free | bound | | | | |
| **1a** | 624 | 635 | 712 | 92 | 1.2 | 0.8[f] |
| **1b** | 577 | 648 | 700 | 17 | 2.7 | 0.2 |
| **1c** | 635 | 740 | 769 | 18 | 0.2 | 0.2 |
| **1d** | 626 | 648 | 718 | 128 | 5.5 | 0.9 |
| **1e** | 628 | 640 | 703 | 87 | 17 | 0.4 |
| **2** | 295 | 435 | 480 | <0.1 | 2.5 | n.d.[g] |
| **3** | 346 | 374 | 404 | <0.1 | 2.4 | n.d.[g] |
| **4** | 422 | 440 | 665, 728 | n.d.[g] | 2.2 | 2.9[h] |
| **5** | 387 | 395 | 513 | <0.1 | 3.5 | 1.8[h] |
| **6** | 473 | 513 | 530 | 1766 | 0.3 | 2.1[i] |

[a]Long-wavelength absorption maximum of the unbound and quadruplex-bound ligand. [b]Emission maximum of bound ligand. [c]Relative emission intensity, $I/I_0$ ($I$ = emission intensity of DNA-bound ligand at saturation, $I_0$ = emission of unbound ligand). [d]Binding constant of ligand–**a2** complex, $K_b^{a2}$, determined from fluorimetric titrations; DNA concentration in nucleotide bases. [e]Binding constant of ligand–**22AG** complex, $K_b^{22AG}$, determined either from fluorimetric titrations or literature data; DNA concentration in oligonucleotide. [f]Ref. [20]. [g]n.d. = not determined. [h]Ref. [23]. [i]Ref [30].

analysis of these data (Table 2; Figure S6 in Supporting Information File 1).

The stoichiometry of the complexes of the ILPR-DNA **a2** with the representative ligands **1e**, **4** and **6** was determined with the continuous variations method (Job plot analysis). For that purpose, the emission intensity was plotted versus the mole fraction of the ligand, $X_{Ligand}$, at constant total concentration of ligand and DNA (Figure 7). The graphically obtained maxima, i.e., the intercept of linearly fitted ascending and descending curve segments, were located at 0.48, 0.73 and 0.50 for ligands **1e**, **4** and **6**, respectively. Within the error margin these data correspond to a binding stoichiometry of 1:1 for complexes of **1e** and **6** with DNA **a2**, and a stoichiometry of 3:1 for the assembly of the porphyrin derivative **4** with the ILPR-DNA. However, it should be noted that strong fluctuations in the Job plot of the complex formation between ligand **1e** and ILPR-DNA **a2**, especially between $X_{Ligand}$ = 0–0.4, indicate strong heterogeneous binding under these conditions.

## CD spectroscopic analysis of ligand–DNA interactions

The interactions of the ligands **1d**, **1e**, **2**, **4**–**6** with ILPR-DNA **a2** were analyzed by CD spectroscopy (Figure 8, Figure S7 in Supporting Information File 1). In most cases, the weak CD signal at 265 nm decreases and the CD signal at 295 nm increases slightly when the ligands were added. Upon addition of the cyanine derivative **1e** to ILPR-DNA, the intensity of the peak at 295 nm increased, whereas the peak at 265 nm slightly decreased. On the other hand, only very small, insignificant changes were observed on titration of **1d** to ILPR-DNA (Figure 8, Figure S7 in Supporting Information File 1). Notably, the addition of the porphyrin **4** and tetraazoniahetarene **5** to ILPR-DNA **a2** has a strong effect on the positive CD signal at 265 nm, leading to the development of a negative band at 260 nm at high LDR ratio (Figure 8D and 8E). In addition, induced CD (ICD) signals were observed only in the absorption range of compounds **4**–**6**.

# Discussion

Generally, it is possible to distinguish between different G-quadruplex DNA forms with the aid of CD spectroscopy [32]. Hence, in agreement with literature data [9,10], the characteristic bands at 235, 265 and 295 nm (Figure 2) show that the employed ILPR-DNA **a2** exists as a mixture of parallel (265 nm) and antiparallel (295 nm) G-quadruplex structures in potassium phosphate buffer solution. The presence of different quadruplex forms is further supported by the temperature-dependent CD spectra that denote the different melting temperatures of the parallel and antiparallel forms (Figure 3). Although it has been reported that the population of the parallel form increases with a low concentration of $K^+$ ions [9]; in our hands it turned out to be dependent also on handling procedure. Hence, if the ILPR-quadruplex **a2** is heated to 97 °C and then cooled immediately to room temperature within 30 s, the CD spectrum of the solution shows a stronger CD signal at 265 nm than at 295 nm [9]. In contrast, if the ILPR-quadruplex **a2** is cooled slowly from 97 °C to room temperature with a rate of 0.5 °C/min the signal at 295 nm is higher than the one at 265 nm [10]. These results clearly demonstrate a delicate equilibrium between the different quadruplex forms and some kinetic barriers that result in rather slow interconversion

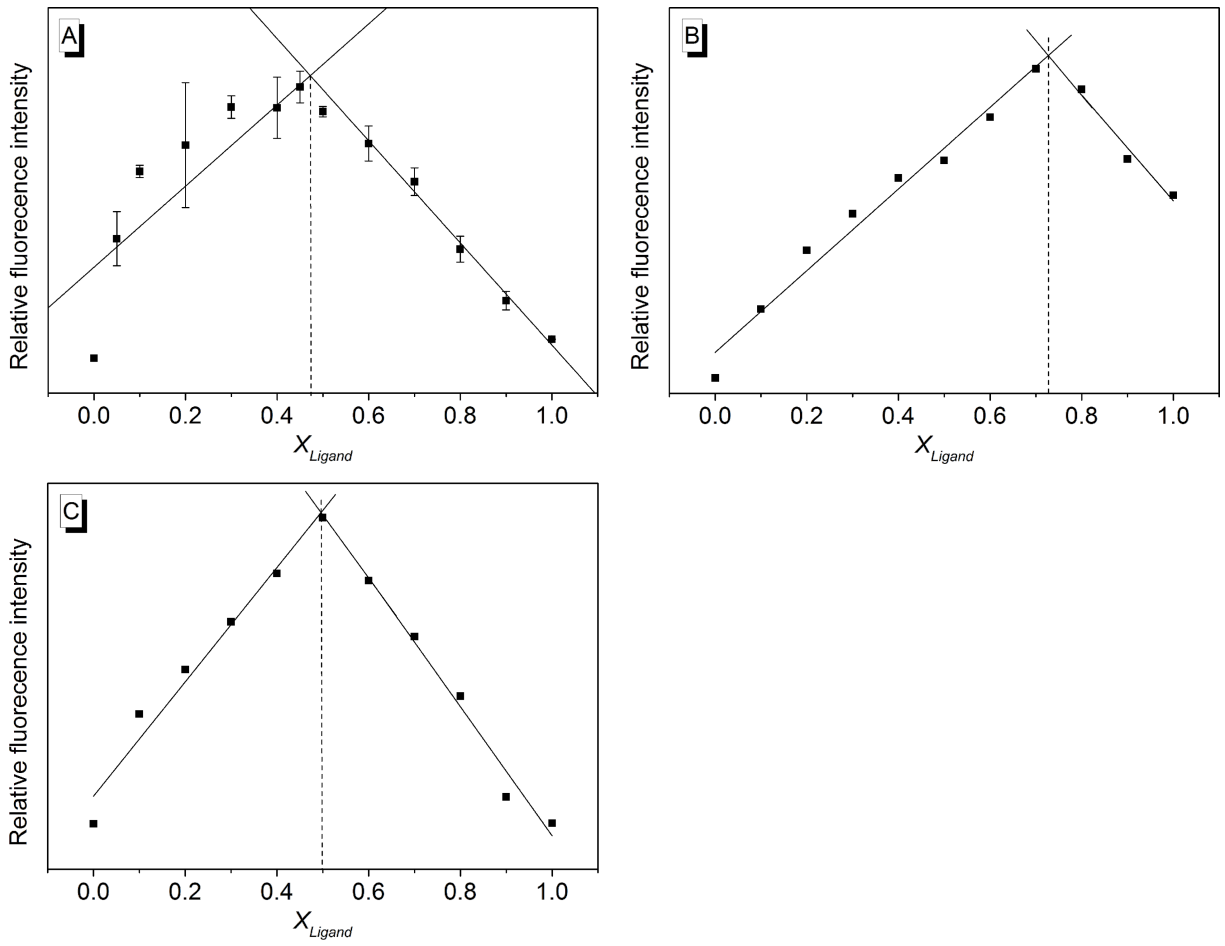



**Figure 7:** Job plot from fluorimetric analysis of mixtures of ligands **1e** (A), **4** (B) and **6** (C) with ILPR-DNA **a2** ($c_{Ligand}$ + $c_{DNA}$ = 10.0 μM) in potassium phosphate buffer (95 mM, pH 7.0); $X_{Ligand}$ = mole fraction of the ligand; A: $\lambda_{ex}$ = 580 nm; B: $\lambda_{ex}$ = 431 nm; C: $\lambda_{ex}$ = 490 nm; $T$ = 20 °C.

processes. Nevertheless, to provide comparable data, in this study we used slowly cooled ILPR-quadruplex **a2** to ensure that we have the same relative population of parallel and antiparallel forms in each experiment.

The fluorimetric thermal DNA denaturation experiments show that the ligands **1b** and **3** induce the highest stabilization of ILPR-quadruplex under the employed conditions ($\Delta T_m$ = 14.5 °C, 14.9 °C respectively, Table 1), whereas the ligands **1a**, **1c**–**e**, **2**, **4** and **5** stabilize this DNA just moderately ($\Delta T_m$ = ca. 3–6 °C). Although the cyanine derivatives **1d** and **1e** show the highest binding affinity towards ILPR-quadruplex, these ligands induce only a moderate shift of its melting temperature (Table 2). This apparent contradiction may be the result of different buffer solutions used in different types of experiments, because the buffer composition, especially the $K^+$ concentration, affects the population of parallel and antiparallel forms [9]. In the thermal denaturation experiment, the population of the parallel forms increases, because a low $K^+$ concentration buffer was used (vide supra). On the other hand, a larger concentration of $K^+$ was used during the fluorimetric titration leading to a higher fraction of the antiparallel form [9,10]. Apart from the derivatives **1c** and **6** the binding constants of the cyanine dyes with the ILPR-DNA are larger than those observed with the telomeric quadruplex **22AG** (Table 2). For example, the cyanine derivative **1e** shows the most significant difference between the binding affinity towards ILPR and the telomeric quadruplex (**a2**: $K_b^{\mathbf{a2}} = 1.7 \times 10^7$ $M^{-1}$; **22AG**: $K_b^{\mathbf{22AG}} = 3.9 \times 10^5$ $M^{-1}$). Although thiazole orange (**6**) has a relatively high binding constant with telomeric quadruplex DNA, it has a lower affinity towards the ILPR quadruplex under the same conditions. In addition, thiazole orange (**6**) shows the least pronounced stabilization effect in the thermal denaturation experiment (Table 2). As cyanine dyes bind to quadruplex DNA by terminal π-stacking [20], which is in agreement with the observed binding stoichiometry of 1:1 (Figure 7),

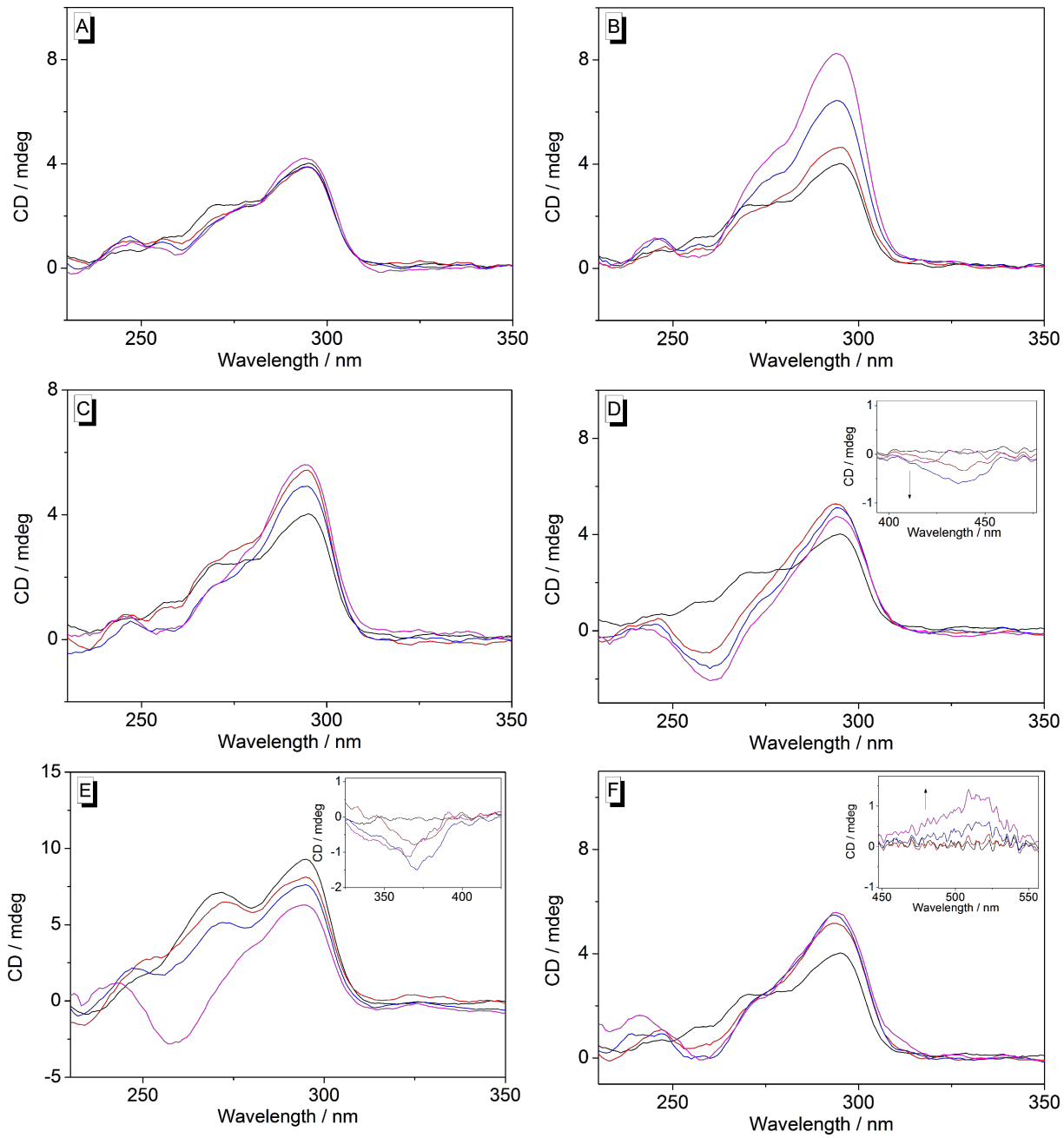



**Figure 8:** CD spectra of ILPR-DNA **a2** in the presence of **1d** (A), **1e** (B), **2** (C), **4** (D), **5** (E) and **6** (F) at LDR = 0 (black), 0.5 (red), 1 (blue), 2 (magenta); $c_{DNA}$ = 20 μM in potassium phosphate buffer (95 mM, pH 7.0); $T$ = 20 °C. Inset: Magnified ICD signals of the bound ligands (magnification factor: ca. 10).

the different composition of the loops in ILPR-quadruplex (ACA and TGT in ILPR versus TTA in telomeric quadruplex) may affect the binding affinity of ligands towards the ILPR-DNA. On the other hand, the porphyrin **4** and the tetraazonia-hetarene derivative **5** show similar binding affinities towards the ILPR-quadruplex thus resembling their interactions with the telomeric **22AG** quadruplex [23]. This observation indicates that these ligands may bind in similar types of binding sites. It should be noted that in the case of the porphyrin **4** the ligand–**a2** stoichiometry of 1:3 (Figure 7) indicates additional stacking of the ligand along the phosphate backbone of the ILPR-DNA.

The cyanine derivatives **1a**, **1d**, **1e** and **6** show significant light-up effects upon association with the ILPR quadruplex (Table 2).

Although thiazole orange (**6**) shows the highest light-up factor ($I/I_0$ = 1766), it has a low selectivity over different types of DNA [33]. Thus, the use of thiazole orange for the selective fluorimetric detection of quadruplex DNA is rather limited. On the other hand, the cyanine derivatives **1a**, **1d** and **1e** show high selectivity towards quadruplex as compared to double stranded DNA [20]. Therefore, these ligands may be used as probes for the detection of quadruplex DNA.

Although the CD-spectroscopic analysis only provides qualitative information about the binding between the ligand and the DNA, they may be used to identify some general trends. Thus, the CD-spectroscopic analysis revealed that some of the ligands induce a very small (**1d**, **6**), moderate (**2**), or even strong (**1e**) increase of the CD band of the DNA **a2** at 295 nm. As this particular band has been assigned to the antiparallel quadruplex structure of **a2** [9,10] the induced increase may indicate a stabilization of this DNA structure upon complex formation. Moreover, a comparison with the telomeric quadruplex **22AG** shows that the ligands **1d**, **1e** and **6** induce the same effect on the CD band of this quadruplex at 295 nm (Supporting Information File 1, Figure S8). At the same time, the ICD of thiazole orange (**6**) in the presence of **22AG** is stronger than that with **a2** (Figure 8F and Figure S8C, Supporting Information File 1) indicating a stronger interaction of the ligand towards the telomeric quadruplex. This observation is in agreement with the higher affinity of **6** towards **22AG** (Table 2). Notably, coralyne (**2**) has a significantly different influence on the CD spectrum of **22AG** as compared to **a2.** In contrast to ILPR-DNA **a2**, the band of **22AG** at 295 nm decreases upon addition of compound **2**, along with a strong ICD signal of the ligand [34]; thus indicating a different binding mode of coralyne (**2**) with **a2** and **22AG**.

Notably, the CD-spectroscopic analysis shows a marginal effect of **1d** and **1e** on the ILPR-DNA **a2**. Nevertheless, as the other experimental data confirm the association of this ligand to the quadruplex it may be concluded that the complex formation between **a2** and **1d** and **1e** does not lead to a significant change of the DNA structure.

In the case of ligands **2**, **4**, **5** and **6**, a negative CD band developed at 265 nm together with a weak positive band with a maximum at ca. 240 nm upon association with quadruplex **a2** (Figure 8). Along with the increase of the CD signal at 295 nm these developments may indicate the pronounced formation of an antiparallel quadruplex form [35]. Furthermore, the CD spectrum of **a2** in the presence of the ligands **2**, **4**, **5** and **6** displays roughly the same maxima as the antiparallel quadruplex with only two stacked G-quartet planes that is formed by the thrombin binding aptamer (TBA) [d($G_2T_2G_2TGTG_2T_2G_2$)], though with varying intensities and slightly different shifts [32]. In addition, it has been suggested that the porphyrin **4** also binds to an antiparallel quadruplex form that consists of just two stacked G-quartets and more extended loop structures [36], the so-called Form 3. Presumably, the sterically demanding ligand **4** fits well into the larger binding sites of the Form 3 quadruplex [37,38]. Considering all these observations it may be very carefully proposed that the ILPR-DNA **a2** also forms structures with only two stacked G-quartets that are stabilized upon association of the ligands **2**, **4**, **5** and **6**. Nevertheless, this proposed folding pattern of the oligonucleotide **a2** needs to be confirmed in a more detailed structure analysis by high-resolution NMR spectroscopy.

## Conclusion

This study represents the first comparative investigation of the interaction of different quadruplex ligands with an oligonucleotide sequence from the ILPR. In summary, it is demonstrated that in most cases the binding parameters of ligand-ILPR complexes are different from the ones observed with other native quadruplex-forming DNA sequences. All ligands under investigation show a high binding affinity towards the ILPR quadruplex. Cyanine derivatives – except thiazole orange (**6**) – have a higher affinity towards ILPR quadruplex than towards the telomeric quadruplex. From these results it may be concluded that, in principle, the association of exogenous ligands with the ILPR-DNA may assist or interfere with the biological activity of this physiologically relevant DNA sequence.

## Experimental

### Materials

Coralyne (**2**), porphyrin **4** and thiazole orange (**6**) were commercially obtained and used without further purification (**2**: Acros Organics, Geel, Belgium; **4**, **6**: Fluka Chemie AG, Buchs, Switzerland). The [2.2.2]heptamethine cyanine dyes **1a**–**e** [22], bis-quinolinium **3** [39] and tetraazoniahetarene **5** [23] were synthesized according to published procedures. Oligodeoxyribonucleotides (HPLC purified) **a2** [d($ACAG_4TGTG_4ACAG_4TGTG_4$)] and **$Fa_2T$** [d(Fluo-$ACAG_4TGTG_4ACAG_4TGTG_4$-Tamra)] were purchased from Metabion Int. AG (Planegg/Martinsried). The ILPR-DNA **a2** or **$Fa_2T$** were dissolved in the proper buffer, heated to 95 °C for 5 min, cooled slowly over 4 hours to room temperature and used two days after storing. Potassium phosphate buffer: 25 mM $K_2HPO_4$, 70 mM KCl; adjusted with 25 mM $KH_2PO_4$ to pH 7.0; sodium cacodylate buffer: 10 mM $Na(CH_3)_2AsO_2{\cdot}3H_2O$, 10 mM KCl, 90 mM LiCl; pH 7.2–7.3.

### Equipment

Absorption spectroscopy: Varian Cary 100 Bio spectrophotometer; emission spectroscopy: Varian Cary Eclipse;

CD spectroscopy: Chirascan CD-spectrometer, Applied Photophysics.

## Methods

All photometric and fluorimetric titrations were performed in thermostated quartz cuvettes at 20 °C. Titrant solutions were freshly prepared by dilution from the stock solution. Spectrophotometer slit widths were 2 nm for photometric experiments and 5 nm for fluorimetric experiments. In fluorimetric experiments, the spectra were smoothed with implemented moving-average function by a factor of 5. The binding constants were determined by fitting the binding isotherms from the fluorimetric titrations to the established theoretical model according to the independent-site model (Equation 1) [31].

$$y = \left( A + B + nx - \sqrt{(A + B + nx)^2 - 4Bnx} \right) / 2B \qquad (1)$$

Where $y$ is the normalized intensity; $A = 1/K_b$; $B = c_{Lig}$, and $n$ is the number of the binding sites per quadruplex DNA.

Thermal denaturation experiments were performed according to the published procedure [20].

For Job plot analysis (continuous variations method), a series of samples was prepared with a constant sum of concentrations at 10.0 μM, but with varying concentrations of ligand and ILPR quadruplex. The fluorescence spectra were recorded for each sample with $\lambda_{ex}$ = 580 nm for cyanine derivative **1e**, and $\lambda_{ex}$ = 431 nm for porphyrin derivative **4** and $\lambda_{ex}$ = 490 nm for thiazole orange (**6**). The maximum fluorescence intensity was plotted versus the molar fraction of the corresponding ligand, $X_{Ligand}$. For the determination of the maximum the ascending and descending segments of the curve were fitted to linear lines, respectively, and the intercept of both lines denotes the maximum and thus the stoichiometry of the complex.

For the CD experiments five samples were prepared with fixed ILPR-DNA concentration ($c_{DNA}$ = 20.0 μM). In four of the samples different amounts of ligand were added to obtain different ligand concentrations (5, 10, 20, 40 μM). CD signals were recorded with a band width of 1 nm, a recording speed of 1 nm $s^{-1}$ and a time per data point of 0.5 s. For the temperature dependent circular dichroism (CD) experiments, solutions of the ILPR-quadruplex **a2** (20 μM) and the ligand (20 μM) were prepared in potassium phosphate buffer (95 mM, pH 7.0) and measured after an equilibration time of 24 h. After each measurement, the temperature was increased by 5 °C or 10 °C, after the temperature has been kept to an equilibration time of 10 min.

## Supporting Information

Figures of fluorimetric DNA denaturation experiments, photometric titration of **a2** ILPR-DNA into ligands **1**–**6**, photometric and fluorimetric titration of **22AG** into ligands **1b**–**e**, plots of the CD intensity change of **a2** ILPR-DNA in presence of ligands **1**–**6** and CD spectra of **22AG** in presence of ligands **1d**, **1e** and **6** are provided.

### Supporting Information File 1

Additional experimental data.
[http://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-10-314-S1.pdf]

## References


1. Bell, G. I.; Horita, S.; Karam, J. H. *Diabetes* **1984,** *33,* 176–183. doi:10.2337/diab.33.2.176
2. Kennedy, G. C.; German, M. S.; Rutter, W. J. *Nat. Genet.* **1995,** *9,* 293–298. doi:10.1038/ng0395-293
3. Rotwein, P.; Yokoyama, S.; Didier, D. K.; Chirgwin, J. M. *Am. J. Hum. Genet.* **1986,** *39,* 291–299.
4. Catasti, P.; Chen, X.; Moyzis, R. K.; Bradbury, E. M.; Gupta, G. *J. Mol. Biol.* **1996,** *264,* 534–545. doi:10.1006/jmbi.1996.0659
5. Lew, A.; Rutter, W. J.; Kennedy, G. C. *Proc. Natl. Acad. Sci. U. S. A.* **2000,** *97,* 12508–12512. doi:10.1073/pnas.97.23.12508
6. Schonhoft, J. D.; Das, A.; Achamyeleh, F.; Samdani, S.; Sewell, A.; Mao, H.; Basu, S. *Biopolymers* **2010,** *93,* 21–31. doi:10.1002/bip.21289
7. Gerasimov, J. Y.; Schaefer, C. S.; Yang, W.; Grout, R. L.; Lai, R. Y. *Biosens. Bioelectron.* **2013,** *42,* 62–68. doi:10.1016/j.bios.2012.10.046
8. Xiao, J.; Carter, J. A.; Frederick, K. A.; McGown, L. B. *J. Sep. Sci.* **2009,** *32,* 1654–1664. doi:10.1002/jssc.200900060
9. Yu, Z.; Schonhoft, J. D.; Dhakal, S.; Bajracharya, R.; Hegde, R.; Basu, S.; Mao, H. *J. Am. Chem. Soc.* **2009,** *131,* 1876–1882. doi:10.1021/ja806782s
10. Timmer, C. M.; Michmerhuizen, N. L.; Witte, A. B.; Van Winkle, M.; Zhou, D.; Sinniah, K. *J. Phys. Chem. B* **2014,** *118,* 1784–1790. doi:10.1021/jp411293r
11. De Messieres, M.; Chang, J.-C.; Brawn-Cinani, B.; La Porta, A. *Phys. Rev. Lett.* **2012,** *109,* 058101. doi:10.1103/PhysRevLett.109.058101
12. Hammond-Kosack, M. C. U.; Dobrinski, B.; Lurz, R.; Docherty, K.; Kilpatrick, M. W. *Nucleic Acids Res.* **1992,** *20,* 231–236. doi:10.1093/nar/20.2.231
13. Perkins, T. T.; Li, H.-W.; Dalal, R. V.; Gelles, J.; Block, S. M. *Biophys. J.* **2004,** *86,* 1640–1648. doi:10.1016/S0006-3495(04)74232-0
14. Liu, J.-q.; Chen, C.-y.; Xue, Y.; Hao, Y.-h.; Tan, Z. *J. Am. Chem. Soc.* **2010,** *132,* 10521–10527. doi:10.1021/ja1038165
15. Schonhoft, J. D.; Bajracharya, R.; Dhakal, S.; Yu, Z.; Mao, H.; Basu, S. *Nucleic Acids Res.* **2009,** *37,* 3310–3320. doi:10.1093/nar/gkp181
16. Oganesian, L.; Bryan, T. M. *BioEssays* **2007,** *29,* 155–165. doi:10.1002/bies.20523
17. Murat, P.; Singh, Y.; Defrancq, E. *Chem. Soc. Rev.* **2011,** *40,* 5293–5307. doi:10.1039/c1cs15117g
18. Paramasivan, S.; Bolton, P. H. *Nucleic Acids Res.* **2008,** *36,* e106. doi:10.1093/nar/gkn487

19. Yan, J.-W.; Chen, S.-B.; Liu, H.-Y.; Ye, W.-J.; Ou, T.-M.; Tan, J.-H.; Li, D.; Gu, L.-Q.; Huang, Z.-S. *Chem. Commun.* **2014,** *50,* 6927–6930. doi:10.1039/c4cc01472c
20. Ihmels, H.; Thomas, L. *Org. Biomol. Chem.* **2013,** *11,* 480–487. doi:10.1039/c2ob26779a
21. Nanjunda, R.; Owens, E. A.; Mickelson, L.; Alyabyev, S.; Kilpatrick, N.; Wang, S.; Henary, M.; Wilson, W. D. *Bioorg. Med. Chem.* **2012,** *20,* 7002–7011. doi:10.1016/j.bmc.2012.10.008
22. Reichardt, C.; Mormann, W. *Chem. Ber.* **1972,** *105,* 1815–1837. doi:10.1002/cber.19721050606
23. Jäger, K.; Bats, J. W.; Ihmels, H.; Granzhan, A.; Uebach, S.; Patrick, B. O. *Chem. – Eur. J.* **2012,** *18,* 10903–10915. doi:10.1002/chem.201103019
24. Pennarun, G.; Granotier, C.; Gauthier, L. R.; Gomez, D.; Hoffschir, F.; Mandine, E.; Riou, J.-F.; Mergny, J.-L.; Mailliet, P.; Boussin, F. D. *Oncogene* **2005,** *24,* 2917–2928. doi:10.1038/sj.onc.1208468
25. Granzhan, A.; Ihmels, H.; Jäger, K. *Chem. Commun.* **2009,** 1249–1251. doi:10.1039/b812891j
26. Wang, P.; Leung, C.-H.; Ma, D.-L.; Yan, S.-C.; Che, C.-M. *Chem. – Eur. J.* **2010,** *16,* 6900–6911. doi:10.1002/chem.201000167
27. Monchaud, D.; Allain, C.; Teulade-Fichou, M.-P. *Bioorg. Med. Chem. Lett.* **2006,** *16,* 4842–4845. doi:10.1016/j.bmcl.2006.06.067
28. Williamson, J. R. *Annu. Rev. Biophys. Biomol. Struct.* **1994,** *23,* 703–730. doi:10.1146/annurev.bb.23.060194.003415
29. Renčiuk, D.; Zhou, J.; Beaurepaire, L.; Guédin, A.; Bourdoncle, A.; Mergny, J.-L. *Methods* **2012,** *57,* 122–128. doi:10.1016/j.ymeth.2012.03.020
30. Monchaud, D.; Allain, C.; Teulade-Fichou, M.-P. *Nucleosides, Nucleotides Nucleic Acids* **2007,** *26,* 1585–1588. doi:10.1080/15257770701548212
31. Stootman, F. H.; Fisher, D. M.; Rodger, A.; Aldrich-Wright, J. R. *Analyst* **2006,** *131,* 1145–1151. doi:10.1039/b604686j
32. Vorlíčková, M.; Kejnovská, I.; Sagi, J.; Renčiuk, D.; Bednářová, K.; Motlová, J.; Kypr, J. *Methods* **2012,** *57,* 64–75. doi:10.1016/j.ymeth.2012.03.011
33. Allain, C.; Monchaud, D.; Teulade-Fichou, M.-P. *J. Am. Chem. Soc.* **2006,** *128,* 11890–11893. doi:10.1021/ja062193h
34. Bhadra, K.; Kumar, G. S. *Biochim. Biophys. Acta* **2011,** *1810,* 485–496. doi:10.1016/j.bbagen.2011.01.011
35. Paramasivan, S.; Rujan, I.; Bolton, P. H. *Methods* **2007,** *43,* 324–331. doi:10.1016/j.ymeth.2007.02.009
36. Martino, L.; Pagano, B.; Fotticchia, I.; Neidle, S.; Giancola, C. *J. Phys. Chem. B* **2009,** *113,* 14779–14786. doi:10.1021/jp9066394
37. Lim, K. W.; Amrane, S.; Bouaziz, S.; Xu, W.; Mu, Y.; Patel, D. J.; Luu, K. N.; Phan, A. T. *J. Am. Chem. Soc.* **2009,** *131,* 4301–4309. doi:10.1021/ja807503g
38. Zhang, H.-J.; Wang, X.-F.; Wang, P.; Ai, X.-C.; Zhang, J.-P. *Photochem. Photobiol. Sci.* **2008,** *7,* 948–955. doi:10.1039/b809322a See also for stabilization of anti-parallel quadruplex by 4.
39. Dorazco-González, A.; Höpfl, H.; Medrano, F.; Yatsimirsky, A. K. *J. Org. Chem.* **2010,** *75,* 2259–2273. doi:10.1021/jo100037m

## License and Terms

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

# NAA-modified DNA oligonucleotides with zwitterionic backbones: stereoselective synthesis of A–T phosphoramidite building blocks

Boris Schmidtgall[1,2], Claudia Höbartner[3,4] and Christian Ducho[*1,2]

Full Research Paper

Open Access

Address:
[1]Department of Chemistry, University of Paderborn, Warburger Str. 100, 33 098 Paderborn, Germany, [2]Department of Pharmacy, Pharmaceutical and Medicinal Chemistry, Saarland University, Campus C2 3, 66 123 Saarbrücken, Germany, [3]Max-Planck-Institute for Biophysical Chemistry, Am Fassberg 11, 37 077 Göttingen, Germany and [4]Department of Chemistry, Institute of Organic and Biomolecular Chemistry, Georg-August-University Göttingen, Tammannstr. 2, 37 077 Göttingen, Germany

Email:
Christian Ducho[*] - christian.ducho@uni-saarland.de

* Corresponding author

Keywords:
backbone modifications; DNA; nucleic acids; oligonucleotides; stereoselective synthesis; zwitterions

*Beilstein J. Org. Chem.* **2015,** *11,* 50–60.
doi:10.3762/bjoc.11.8

Received: 24 October 2014
Accepted: 11 December 2014
Published: 13 January 2015

This article is part of the Thematic Series "Nucleic acid chemistry".

Guest Editor: H.-A. Wagenknecht



## Abstract

Modifications of the nucleic acid backbone are essential for the development of oligonucleotide-derived bioactive agents. The NAA-modification represents a novel artificial internucleotide linkage which enables the site-specific introduction of positive charges into the otherwise polyanionic backbone of DNA oligonucleotides. Following initial studies with the introduction of the NAA-linkage at T–T sites, it is now envisioned to prepare NAA-modified oligonucleotides bearing the modification at X–T motifs (X = A, C, G). We have therefore developed the efficient and stereoselective synthesis of NAA-linked 'dimeric' A–T phosphoramidite building blocks for automated DNA synthesis. Both the (*S*)- and the (*R*)-configured NAA-motifs were constructed with high diastereoselectivities to furnish two different phosphoramidite reagents, which were employed for the solid phase-supported automated synthesis of two NAA-modified DNA oligonucleotides. This represents a significant step to further establish the NAA-linkage as a useful addition to the existing 'toolbox' of backbone modifications for the design of bioactive oligonucleotide analogues.

## Introduction

Oligonucleotides are important agents for a number of biomedical applications [1]. Thus, they are employed to exert antigene [2] or antisense [3] mechanisms as well as to trigger or inhibit RNA interference [4]. The capability of sequence-specific molecular recognition is a striking feature of nucleic acids, but their high polarity represents a significant hurdle for cellular uptake and leads to problematic pharmcokinetics. Furthermore, they are prone to nuclease-mediated degradation. As a conse-

quence, it is of utmost importance to modify oligonucleotide structures using chemical or enzymatic methods in order to develop oligonucleotide-based drug candidates or biomedical chemical probes [5,6].

Native nucleic acids are connected by phosphate diesters as linking units, thus leading to a polyanionic backbone structure. Implications of this characteristic feature of nucleic acid architecture have been discussed by Westheimer [7] and Benner [8,9], among others. However, the accumulation of negative charges in the nucleic acid backbone is mainly responsible for their limited membrane penetration. Consequently, a significant number of artificial internucleotide linkages has been studied with the aim to manipulate the charge pattern in the backbone and to enhance nuclease stability. The electroneutral nucleic acid mimic 'peptide nucleic acid' (PNA) [10-12] is capable of sequence-specific binding to nucleic acids, but it was found to display limited water solubility and a peptide-like folding behaviour [8]. For the selective replacement of some phosphate linkages in otherwise native oligonucleotide structures, e.g., amide [13-20], triazole [21,22], phosphoramidate [23] and phosphate triester [24] moieties were reported alongside a considerable number of other modifications.

In comparison, the introduction of positive charges into the nucleic acid scaffold has found less attention. Positively charged units were attached to the 2'-hydroxy group or the nucleobase as a compensation for the presence of negative charges in the phosphate backbone [25-29]. In contrast to such zwitterionic, but densely charged systems, only very few attempts were made to replace the phosphate moiety by a positively charged motif [30], mainly by Bruice et al. [31-34]. Selective replacement of some phosphates with cationic motifs furnished oligonucleotide 'chimera' [30] with zwitterionic backbone structures as reported both by Bruice [35,36] and Letsinger [37].

We have found the approach to prepare oligonucleotides with zwitterionic backbones interesting both for its fundamental implications and also for its potential contribution to the existing 'toolbox' of backbone modifications for bioactive oligonucleotide analogues. These considerations have led to our recently reported design of a novel artificial internucleotide linkage named 'NAA-modification' (Figure 1) [38]. Ongoing synthetic and structure-activity relationship (SAR) studies on naturally occurring muraymycin antibiotics (e.g., muraymycin A1 (**1**)) [39-46] have led to our previously reported synthesis of 'nucleosyl amino acid' structures **2** [47,48] as simplified 5'-defunctionalized analogues of the muraymycin core motif. Formally merging the **n**ucleosyl **a**mino **a**cid (NAA) structure of type **2** with previously reported amide internucleotide linkages such as **3** and **4** provides an 'NAA-modified oligonucleotide' **5**, i.e., a nucleic acid strand with the NAA-linkage replacing a phosphate diester motif. The amino group of the NAA-modification is expected to display a positive charge at physiological pH values, thus leading to a (partially) zwitterionic backbone structure in NAA-modified oligonucleotides.

We have previously described that NAA-modified DNA oligonucleotides can be obtained by standard solid phase-supported automated DNA synthesis, using the 'dimeric' phosphoramidite building blocks **6** (Figure 1) [38]. Overall, 24 different oligonucleotide sequences with one to four NAA-modifications at various positions were synthesized. The stereochemistry of the NAA-motif was either (*S*) or (*R*) (obtained by application of the corresponding phosphoramidites (*S*)-**6** or (*R*)-**6** for DNA synthesis) in order to study the influence of the spatial orientation of the positive charge. Melting temperature measurements showed that NAA-modified DNA oligonucleotides formed stable duplexes with native unmodified DNA or RNA counterstrands, although moderate destabilization in comparison to native duplexes was observed (particularly for DNA/RNA duplexes). Further experiments with native counterstrands bearing one nucleobase mismatch were performed, and duplex structures were studied by CD spectroscopy. Overall, we found that NAA-modified DNA oligonucleotides (i) formed stable duplexes with complementary counterstrands; (ii) were fully capable of mismatch discrimination and (iii) formed duplexes without significant structural distortion, i.e., B-form helices (DNA/DNA duplexes) and A-form helices (DNA/RNA duplexes), respectively. It was concluded that typical chemical properties of nucleic acids are retained in NAA-modified DNA oligonucleotides [38], thus making the NAA-linkage an interesting structural motif for oligonucleotide analogues.

Using 'dimeric' T–T phosphoramidites **6**, it was only possible to introduce the NAA-modification at T–T motifs. While this was fully sufficient for initial studies, it will be of major importance to develop methods for the synthesis of NAA-modified oligonucleotides with the NAA-motif at more variable positions in the base sequence. The synthesis of 'dimeric' X–T phosphoramidites (X = A, C, G) would enable an introduction of the NAA-linkage at every position in an oligonucleotide sequence with a T in 3'-direction, thus significantly broadening the applicability of the modification. In this work, we describe the stereoselective synthesis of 'dimeric' NAA-linked A–T phosphoramidites (*S*)-**7** and (*R*)-**7** (Figure 1) as well as their application in automated DNA synthesis. This represents the first step towards a comprehensive set of 'dimeric' NAA-linked X–T phosphoramidites for the automated chemical synthesis of NAA-modified DNA oligonucleotides.

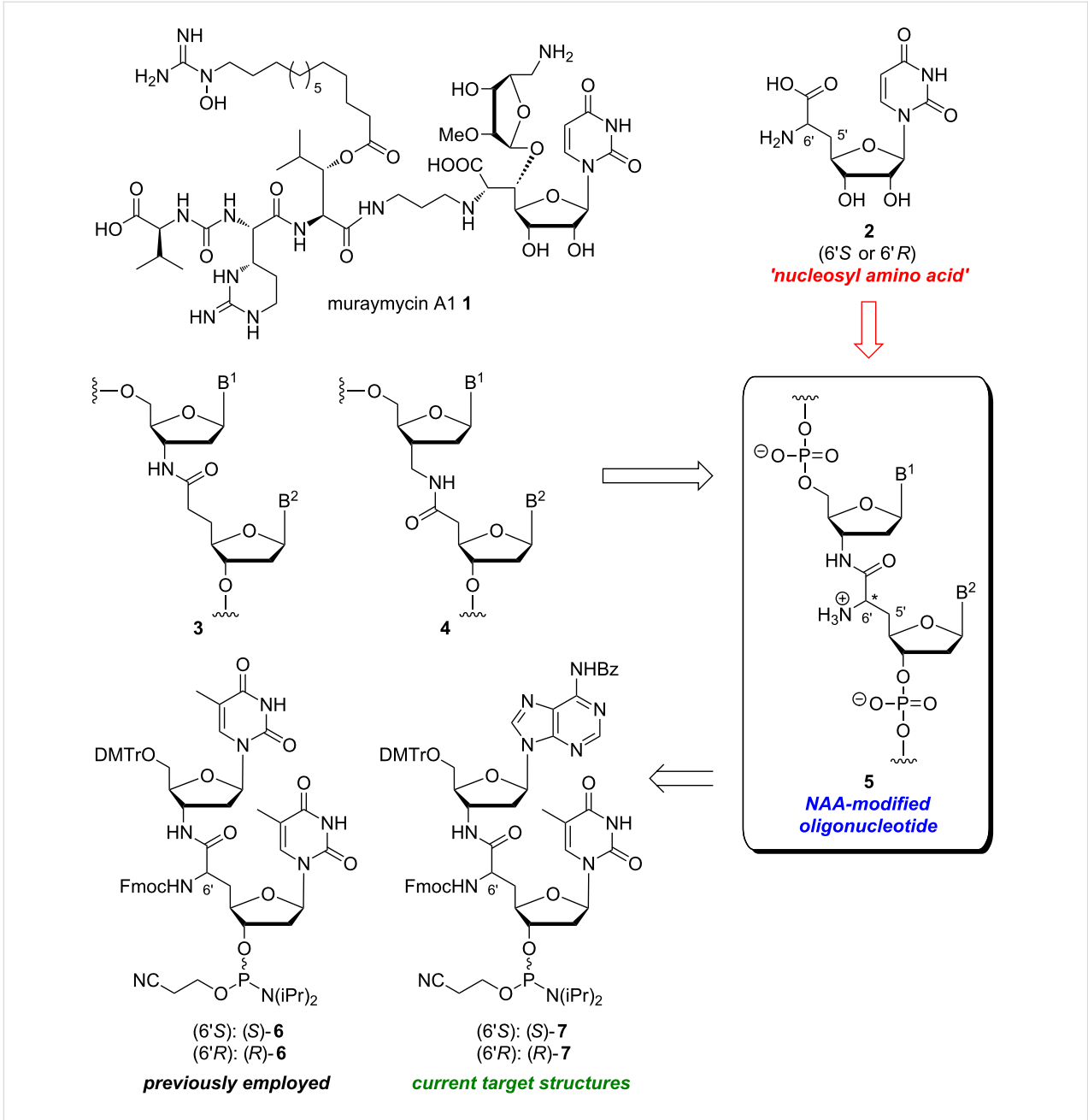
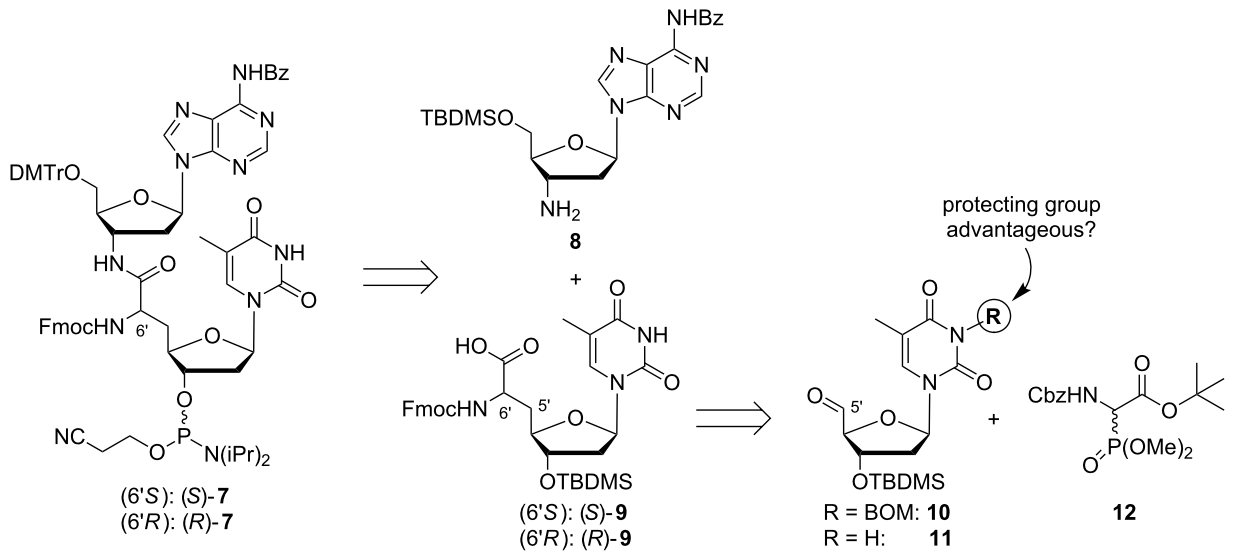



**Figure 1:** Design concept of nucleosyl amino acid (NAA)-modified oligonucleotides **5** formally derived from structures **1**–**4** ($B^1$, $B^2$ = nucleobases); previously employed 'dimeric' T–T phosphoramidites **6** [38] for the automated synthesis of NAA-modified oligonucleotides; new 'dimeric' A–T phosphoramidites **7** as target structures of this study (DMTr = 4,4'-dimethoxytrityl).

## Results and Discussion

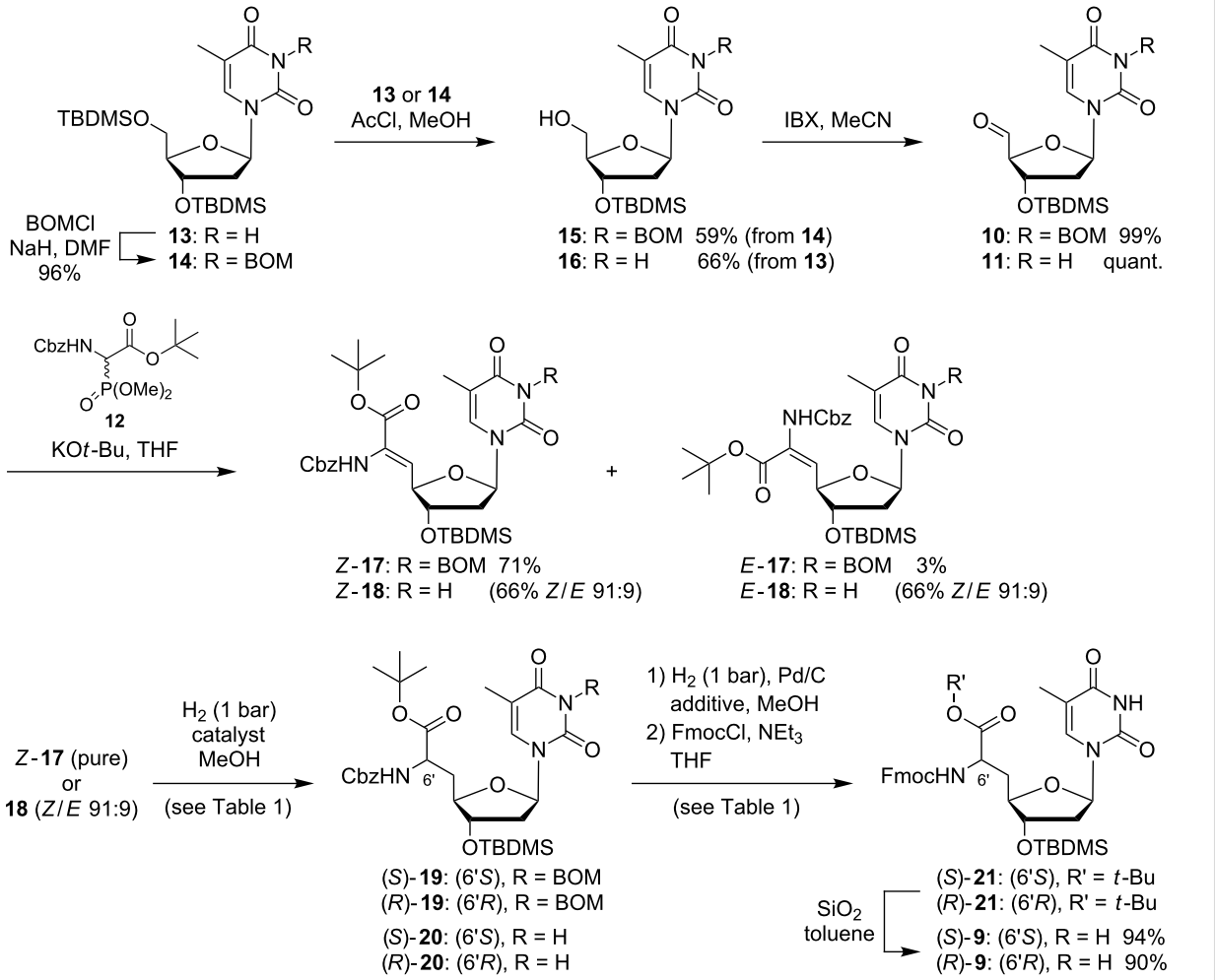
For the synthesis of target phosphoramidites **7**, it was planned to employ a similar synthetic strategy as previously described by us for T–T phosphoramidites **6** [38]. One important objective was to construct the 6'-stereocenter of the NAA-linkage in a controlled fashion and to retain the resultant (6'*S*)- or the (6'*R*)-configuration, respectively, on the way to dimeric structures of type **7** (Scheme 1). Thus, it was envisioned that target compounds (*S*)-**7** and (*R*)-**7** could be obtained from protected 3'-amino-2',3'-dideoxyadenosine **8** and the *N*-Fmoc-protected thymidine-derived nucleosyl amino acids (*S*)-**9** and (*R*)-**9**, respectively, via amide coupling, protecting group manipulation and phosphitylation. The stereoselective synthesis of both 6'-epimers of nucleosyl amino acid **9** has been reported before. We have used 3-(*N*-BOM)-protected thymidine-5'-aldehyde **10** (BOM = benzyloxymethyl), which can readily be obtained from thymidine, in a sequence of Wittig–Horner reaction, asymmetric hydrogenation of the resultant didehydro nucleosyl




**Scheme 1:** Retrosynthetic analysis of target phosphoramidites (*S*)-**7** and (*R*)-**7** (BOM = benzyloxymethyl).

amino acid and protecting group manipulations in order to obtain both 6'-epimers of **9** dependent on the choice of the hydrogenation catalyst [38]. Here, we report on further studies directed towards a possible synthesis of **9** without 3-(*N*-BOM)-protection of the thymine nucleobase. In the case of the corresponding uridine-derived nucleosyl amino acids, we have found that uracil protection was not advantageous for the aforementioned reaction sequence [48]. We have therefore decided to employ both the 3-(*N*-BOM)-protected thymidine-5'-aldehyde **10** and also its thymine-unprotected congener **11** in Wittig–Horner reactions with glycine-derived phosphonate **12** and to compare both possible routes towards **9**, i.e., with or without thymine protection (Scheme 1).

For the synthesis of the *N*-Fmoc-protected thymidine-derived nucleosyl amino acids (*S*)-**9** and (*R*)-**9**, 3',5'-bis-*O*-silylated thymidine **13** (which can be readily prepared from thymidine with TBDMS chloride and imidazole in pyridine as solvent in quantitative yield) was 3-*N*-protected by alkylation with benzyloxymethyl chloride (BOMCl), furnishing product **14** in 96% yield (Scheme 2). Although the route involving thymdine protection has been published before [38], it is also depicted in Scheme 2 in the interest of clarity and readability. Both bis-silyl ethers **13** and **14** then underwent selective acidic cleavage in the 5'-position using the conditions reported by Khan and Mondal [49], i.e., hydrochloric acid in methanol (generated by the reaction of acetyl chloride with methanol), thus providing 3'-*O*-TBDMS-protected derivatives **15** and **16** in yields of 59% and 66%, respectively. This method turned out to be advantageous compared to 5'-*O*-desilylation mediated by TFA, which had provided satisfying results in the case of the corresponding uridine derivatives [48]. The yield of the desired 5'-alcohols **15** and **16** was limited though by partial concomitant cleavage of the 3'-*O*-TBDMS group upon prolonged reaction times. Alcohols **15** and **16** were then oxidized to aldehydes **10** and **11** in quantitative yields using IBX in refluxing acetonitrile [50]. With respect to their limited stability, thymidine-5'-aldehydes **10** and **11** were not stored, but directly used for the subsequent Wittig–Horner reaction. They were therefore converted with glycine-derived phosphonate **12** [51-54] in the presence of potassium *tert*-butoxide as a base. As anticipated [47,48,55], these reactions showed pronounced stereoselectivity towards the *Z*-configured didehydro nucleosyl amino acids. In the case of the reaction of 3-(*N*-BOM)-protected thymidine-5'-aldehyde **10** with phosphonate **12**, isomer *Z*-**17** was isolated in 71% yield, with *E*-**17** representing a minor byproduct (3% yield) which could be separated by column chromatography. The assignment of the configuration of the newly formed trisubstituted C–C double bond was based on empirical $^1$H NMR criteria for didehydro amino acids [56], which were proven to be useful and reliable [38,47,48,57]. In the case of the conversion of 3-*N*-unprotected thymidine-5'-aldehyde **11** with phosphonate **12**, an unseparable mixture of *Z*-**18** and *E*-**18** was obtained in 66% overall yield (*Z*/*E* = 91:9). However, it is firmly established that the subsequent asymmetric hydrogenation proceeds significantly faster with the *Z*-configured didehydro amino acid substrate [58], and it was therefore decided to use the aforementioned mixture of double bond isomers (containing 9% of the unwanted *E*-isomer) as starting material for this transformation.

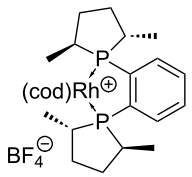



**Scheme 2:** Synthesis of *N*-Fmoc-protected thymidine-derived nucleosyl amino acids (*S*)-**9** and (*R*)-**9**; details on the reactions from **17** and **18** to **21** (asymmetric hydrogenation and subsequent protecting group manipulations) are given in Table 1.

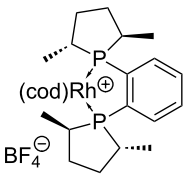
**Table 1:** Reactions from **17** and **18** to **21** (see Scheme 2).

| | | Asymmetric hydrogenation | | | Protecting group steps | |
|---|---|---|---|---|---|---|
| # | Starting material | Catalyst[a] | Reaction time | Yield[b] | Additive | Yield[c] |
| 1 | *Z*-**17**[d] | (*S*,*S*)-Me-DuPHOS-Rh | 2 d | 94% (*S*)-**19** | *n*-$BuNH_2$ | 90% (*S*)-**21** |
| 2 | *Z*-**17**[d] | (*R*,*R*)-Me-DuPHOS-Rh | 7 d | 99% (*R*)-**19** | *n*-$BuNH_2$ | 87% (*R*)-**21** |
| 3 | **18** (mix.)[e] | (*S*,*S*)-Me-DuPHOS-Rh | 9 d | 93% (*S*)-**20** | – | 84% (*S*)-**21** |
| 4 | **18** (mix.)[e] | (*R*,*R*)-Me-DuPHOS-Rh | 21 d | 77% (*R*)-**20** | – | 78% (*R*)-**21** |

[a]Homogeneous chiral hydrogenation catalysts:




(*S*,*S*)-Me-DuPHOS-Rh




(*R*,*R*)-Me-DuPHOS-Rh

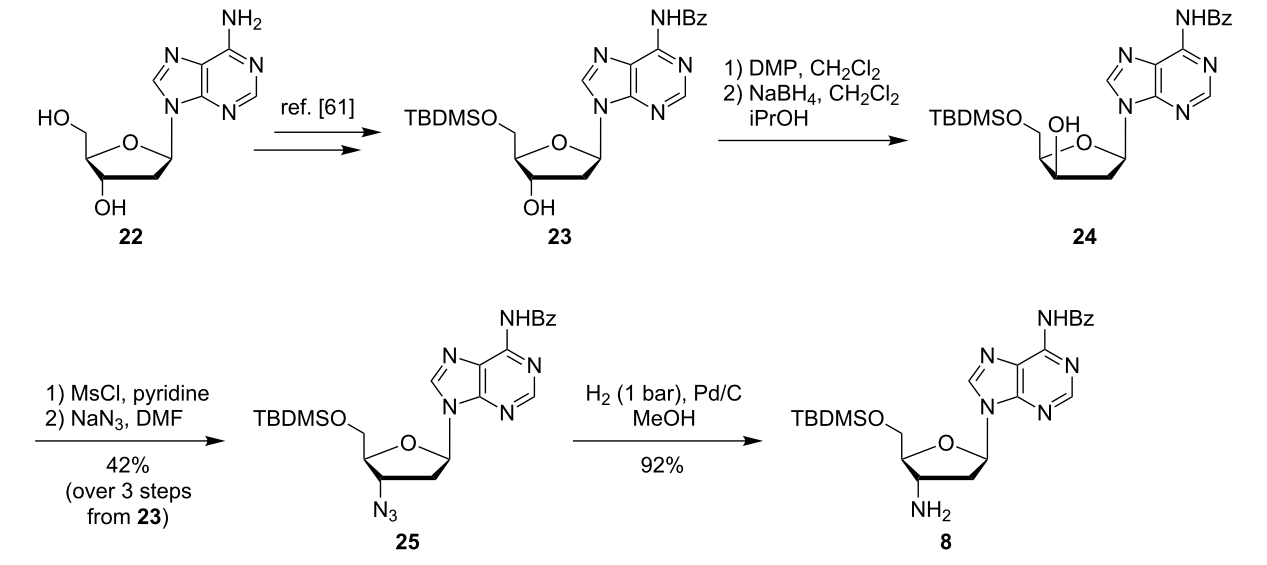
[b]d.r. >98:2 for all products; [c]over 2 steps from **19** or **20**; [d]pure *Z*-isomer; [e]unseparable mixture of *Z*- and *E*-isomers (*Z*/*E* 91:9).

Asymmetric hydrogenation reactions were performed under homogeneous conditions using the chiral catalysts (*S*,*S*)-Me-DuPHOS-Rh or (*R*,*R*)-Me-DuPHOS-Rh, respectively [59]. It is known that asymmetric hydrogenations of *Z*-configured didehydro amino acids catalyzed by (*S*,*S*)-Me-DuPHOS-Rh give L-amino acids and that analogous reactions catalyzed by (*R*,*R*)-Me-DuPHOS-Rh provide D-amino acids [57,60]. This has also been observed when these catalysts were applied for the synthesis of uridine-derived nucleosyl amino acids [47,48]. It was therefore possible to direct the stereochemical outcome of the hydrogenation reaction by the choice of either the (*S*,*S*)- or the (*R*,*R*)-catalyst. As reported previously [38], the hydrogenation of pure 3-(*N*-BOM)-protected *Z*-**17** in the presence of (*S*,*S*)-Me-DuPHOS-Rh thus furnished thymidine-derived nucleosyl amino acid (*S*)-**19** in 94% yield, and its 6'-epimer (*R*)-**19** was obtained from the same starting material in 99% yield using catalytic amounts of (*R*,*R*)-Me-DuPHOS-Rh (Table 1, entries 1 and 2). Both transformations displayed excellent diastereoselectivies (d.r. >98:2 for both products), which indicated that these reactions proceeded in a catalyst-controlled fashion. The observation that full conversion of *Z*-**17** to (*S*)-**19** was reached after 2 days with the (*S*,*S*)-Me-DuPHOS-Rh catalyst, while completion of the reaction took 7 days with the (*R*,*R*)-Me-DuPHOS-Rh catalyst, suggested that the former transformation represented the apparent 'matched' case and the latter the apparent 'mismatched' case. This was also in agreement with similar findings for the synthesis of the according uridine-derived congeners [47,48]. When the isomeric mixture (*Z*/*E* = 91:9) of the 3-*N*-unprotected didehydro amino acid **18** was employed as starting material, 3-*N*-unprotected products (*S*)-**20** (93% yield, with (*S*,*S*)-Me-DuPHOS-Rh) and (*R*)-**20** (77% yield, with (*R*,*R*)-Me-DuPHOS-Rh) were obtained, again with excellent diastereoselectivities (d.r. >98:2 for both products, Table 1, entries 3 and 4). However, it was much more difficult to drive these reactions to completion as reflected by the significantly prolonged reaction times (9 and 21 days, respectively). For the apparent 'mismatched' case, i.e., hydrogenation of **18** in the presence of (*R*,*R*)-Me-DuPHOS-Rh, it was also necessary to increase the catalyst load (4 mol % added portionwise in comparison to 1–2 mol % for the other reactions), and even under these modified conditions, full conversion could not be reached, resulting in a moderately reduced yield. For both transformations of the isomeric mixture of **18**, no hydrogenation of the much less reactive *E*-isomer was observed, which is proven by the high diastereoselectivities. If one takes into account that the purity of the reactive *Z*-configured starting material *Z*-**18** was only 91% (vide supra), it is possible to calculate yields of ca. 100% and 85% for products (*S*)-**20** and (*R*)-**20**, respectively. Overall, it can still be concluded though that 3-*N*-unprotected didehydro amino acid **18** represented a less reactive substrate for the asymmetric hydrogenation key step, which was found to be problematic particularly for reactions on a larger scale (>1 g starting material). It is interesting to note that such limitations were not encountered when the analogous nucleobase-unprotected uridine-derived didehydro nucleosyl amino acid underwent asymmetric hydrogenation with the two aforementioned catalysts [48]. This decreased reactivity might have been the result of the presence of the *E*-isomer *E*-**18** in the reaction mixture, which probably led to partial inhibition of the Rh(I) catalyst.

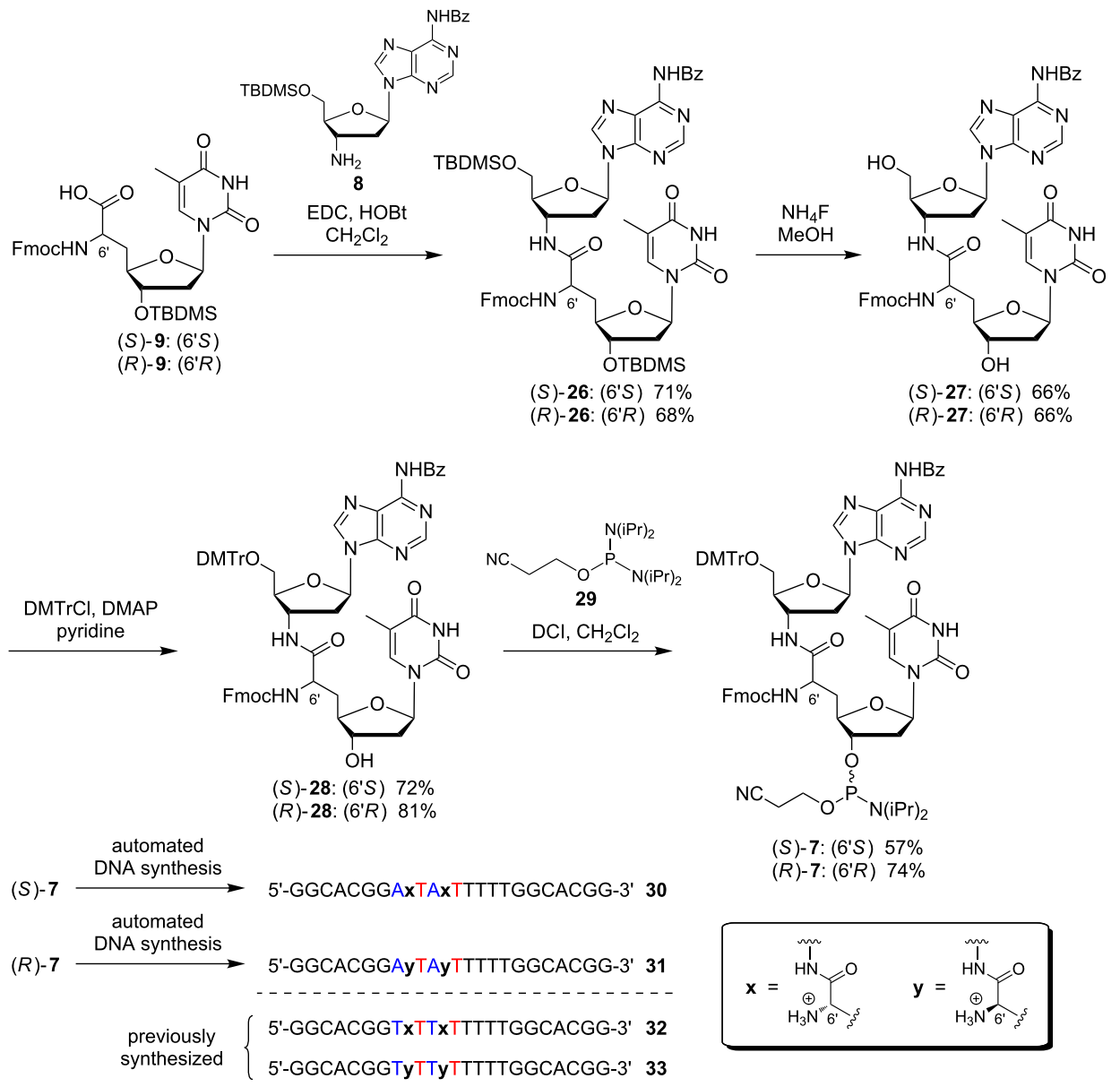
In order to convert the hydrogenation products **19** and **20** into the desired building blocks **9**, three further transformations were necessary: (i) hydrogenolytic cleavage of the Cbz and, in the case of **19**, also of the BOM group; (ii) Fmoc-protection of the 6'-amino functionality (furnishing intermediates (*S*)-**21** and (*R*)-**21**) and (iii) cleavage of the *tert*-butyl ester (Scheme 2). Using the two diastereomerically pure 6'-epimers of 3-(*N*-BOM)-protected **19** as starting material, one challenge was to avoid unwanted side reactions resulting from the generation of formaldehyde in the reaction mixture. Hydrogenolysis of the BOM group affords toluene and formaldehyde as byproducts, and the Cbz-deprotected 6'-amino group can undergo unwanted reductive amination, i.e., methylation, with the liberated formaldehyde. Our method to prevent this side reaction was to include an excess of *n*-butylamine as an additive in the reaction mixture of the hydrogenolysis step. This way, the formaldehyde methylated the added *n*-butylamine, furnishing a reasonably volatile byproduct [38,48]. Subsequent Fmoc-protection under standard conditions then afforded diastereomerically pure products (*S*)-**21** (90% yield over 2 steps from (*S*)-**19**) and (*R*)-**21** (87% yield over 2 steps from (*R*)-**19**), respectively (Table 1, entires 1 and 2). For the analogous transformation of 3-*N*-unprotected (*S*)-**20** and (*R*)-**20**, it was possible to omit the additive *n*-butylamine, but yields were not improved due to this simplification (84% yield of (*S*)-**21** over 2 steps from (*S*)-**20**, 78% yield of (*R*)-**21** over 2 steps from (*R*)-**20**, Table 1, entries 3 and 4). Finally, the *tert*-butyl ester was cleaved selectively in the presence of the acid-labile silyl group using silica in refluxing toluene, thus affording the desired nucleosyl amino acid building blocks (*S*)-**9** (94% yield) and (*R*)-**9** (90% yield), each in diastereomerically pure form (Scheme 2). Overall, it can be concluded that the omission of the 3-(*N*-BOM) protecting group did not lead to improvements in the synthesis of **9**, but rather made the sequence of Wittig–Horner olefination and asymmetric hydrogenation slightly less efficient. In contrast to uridine derivatives [48], thymidine analogues are significantly more robust towards unwanted reduction of the 5,6-double bond in heterogeneously catalyzed hydrogenolysis reactions with palladium catalysts. Therefore, the absence of the BOM group was not advantageous for the deprotection step following the asymmetric hydrogenation reaction.

The second challenge on the way to target phosphoramidites **7** was the synthesis of protected 3'-amino-2',3'-dideoxyadenosine **8** on a sufficient scale. Richert and Eisenhuth have published a comprehensive report on the synthesis of all four 3'-amino-2',3'-dideoxynucleosides with canonical bases [61]. We have decided to mainly follow their strategy for the synthesis of adenosine derivative **8**, though some modifications were applied (Scheme 3). Starting from 2'-deoxyadenosine (**22**), *N*-6-benzoyl-5'-*O*-TBDMS-2'-deoxyadenosine (**23**) was prepared [61] and subjected to oxidation of the 3'-hydroxy functionality to a keto group with Dess–Martin periodinane (DMP). This was followed by reduction with sodium borohydride via nucleophilic attack of the keto group from the sterically less hindered α-face, therefore resulting in the formation of the 3'-*xylo* derivative **24**. A tight control of the reaction conditions, i.e., amounts of reagents, reaction time and temperature, proved to be important for this transformation (see Supporting Information File 1). It was observed that the time period needed for the oxidation step was dependent on the scale of the reaction and that the sensitive keto intermediate decomposed when this time period was unreasonably exceeded. The obtained product **24** still contained minor amounts of aromatic byproducts from the oxidizing agent, which were difficult to remove by column chromatography at this stage. Therefore, this material was employed in subsequent transformations without further attempts to remove the aforementioned impurities. Several methods are known to perform the nucleophilic displacement at the 3'-position of the 3'-*xylo* intermediate with azide as a nucleophile [61-65]. We have found that robust results for this reaction could be achieved by activation of the 3'-hydroxy group as a mesylate, followed by treatment with sodium azide in DMF at elevated temperature (110 °C). This protocol furnished 3'-azido derivative **25** in a moderate, but reliably obtained yield of 42% over 3 steps from **23**. Azido nucleoside **25** was finally reduced to the 3'-amino analogue **8** by standard hydrogenation in 92% yield (Scheme 3).

With all building blocks **8**, (*S*)-**9** and (*R*)-**9** in hand, the dimeric target structures could be constructed (Scheme 4). Using a standard procedure for peptide coupling, suitably protected nucleosyl amino acids (*S*)-**9** and (*R*)-**9** were activated and reacted with amine **8** to give bis-*O*-silylated NAA-linked A–T dimers (*S*)-**26** and (*R*)-**26** in yields of 71% and 68%, respectively. For the subsequent desilylation reaction, several reaction conditions were tested, among them acidic silyl ether cleavage with hydrochloric acid in methanol or treatment with triethylamine trihydrofluoride ($3HF•NEt_3$). However, the only successful method was the conversion of both epimers of **26** with ammonium fluoride in methanol at elevated temperature, which afforded diols (*S*)-**27** and (*R*)-**27**, each in 66% yield. DMTr protection under standard conditions then provided intermediates (*S*)-**28** and (*R*)-**28** in yields of 72% and 81%, respectively. Finally, phosphitylation of the 3'-hydroxy group gave target phosphoramidites (*S*)-**7** (57% yield) and (*R*)-**7** (74% yield), each with defined stereochemistry at the 6'-position. For this reaction, 2-cyanoethyl *N*,*N*,*N*′,*N*′-tetraisopropylphosphordiamidite **29** was used under slightly acidic conditions, i.e., in the presence of the activator 4,5-dicyanoimidazole (DCI, Scheme 4). The alternative method for the introduction of the phosphoramidite functionality, i.e., treatment with the respective




**Scheme 3:** Synthesis of protected 3'-amino-2',3'-dideoxyadenosine **8**.




**Scheme 4:** Synthesis of target phosphoramidites (*S*)-**7** and (*R*)-**7** and of two NAA-modified DNA oligonucleotides **30**, **31** (DCI = 4,5-dicyanoimidazole); sites of NAA-modifications in the oligonucleotides are indicated as **x** (6'*S*) and **y** (6'*R*), all other linkages were native phosphates.

chlorophosphite in the presence of a base, resulted in unwanted concomitant cleavage of the Fmoc group.

In order to demonstrate their principle synthetic versatility, 'dimeric' phosphoramidites (*S*)-**7** and (*R*)-**7** were employed for the automated synthesis of two DNA oligonucleotides **30** and **31**, each bearing two NAA-modifications with defined stereochemistry at the NAA-linkage ((6'*S*) or (6'*R*)) and therefore displaying partially zwitterionic backbone structures. After assembly on the synthesizer and base-mediated cleavage from the solid support, the desired full-length products were purified, isolated and identified by ESI mass spectrometry. The sequence of **30** and **31** resembled the sequence of model oligonucleotides **32** and **33** prepared for our initial studies on the NAA-modification [38] (Scheme 4). This strategy will enable systematic future studies on the potential influence of the base sequence on the properties of NAA-modified oligonucleotides.

## Conclusion

In summary, we have successfully accomplished the stereoselective synthesis of 'dimeric' A–T phosphoramidite building blocks (*S*)-**7** and (*R*)-**7** for the preparation of novel NAA-modi-

fied DNA oligonucleotides with partially zwitterionic backbone structures. The required nucleosyl amino acid intermediates (*S*)-**9** and (*R*)-**9** were synthesized from thymidine in overall yields of 32% and 31%, respectively, over 9 steps when the nucleobase was 3-(*N*-BOM)-protected [38]. It was possible to slightly shorten this route to 8 steps by leaving the nucleobase unprotected, but the key steps of the synthesis, i.e., Wittig–Horner reaction and subsequent asymmetric hydrogenation, proceeded less efficiently in this case. Coupling of (*S*)-**9** and (*R*)-**9** with protected 3'-amino-2',3'-dideoxyadenosine **8** (obtained from protected 2'-deoxyadenosine derivative **23** in 39% overall yield over 4 steps) and some further transformations furnished target phosphoramidites (*S*)-**7** and (*R*)-**7** in overall yields of 19% and 27%, respectively, over 4 steps. Phosphoramidites (*S*)-**7** and (*R*)-**7** were then employed for solid phase-supported automated DNA synthesis, which afforded novel zwitterionic oligonucleotides **31** and **32** bearing the NAA-modification at A–T sites. Overall, it was therefore demonstrated that the synthesis of different 'dimeric' NAA-linked X–T phosphoramidites with X representing pyrimidine or purine nucleobases appears to be feasible. This will enable the preparation of NAA-modified oligonucleotides with significant variations in the base sequence. We are currently finishing the synthesis of a comprehensive set of corresponding X–T phosphoramidites (including C–T and G–T congeners). It is planned to use these reagents to perform a systematic study on the influence of the base sequence on the properties of NAA-modified oligonucleotides, particularly on duplex stability. This will set the stage for further investigations, particularly on oligonucleotides with fully zwitterionic backbone structures and also on the interaction of NAA-modified oligonucleotides with proteins such as nucleases and polymerases.

## Supporting Information

Supporting Information features preparation, analytical data and copies of $^{1}H$, $^{13}C$ and $^{31}P$ NMR spectra of compounds **7**, **8**, **11**, **16**, **18**, **20**, **21**, **24**–**28** as well as preparation and analytical data of oligonucleotides **30** and **31**.

### Supporting Information File 1

Experimental procedures and NMR spectra of compounds **7**, **8**, **11**, **16**, **18**, **20**, **21**, **24**–**28**, **30**, and **31**.
[http://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-11-8-S1.pdf]

## Acknowledgements

We thank the Deutsche Forschungsgemeinschaft (DFG, grant DU 1095/2-1) and the Fonds der Chemischen Industrie (FCI, Sachkostenzuschuss) for financial support. B. S. is grateful for a doctoral fellowship of the Studienstiftung des deutschen Volkes.

## References

1. Sharma, V. K.; Rungta, P.; Prasad, A. K. *RSC Adv.* **2014,** *4,* 16618–16631. doi:10.1039/c3ra47841f
2. Xodo, L. E.; Cogoi, S.; Rapozzi, V. *Curr. Pharm. Des.* **2004,** *10,* 805–819. doi:10.2174/1381612043452983
3. Wagner, R. W. *Nature* **1994,** *372,* 333–335. doi:10.1038/372333a0
4. Mello, C. C. *Angew. Chem.* **2007,** *119,* 7114–7124. doi:10.1002/ange.200701713
   *Angew. Chem., Int. Ed.* **2007,** *46,* 6985–6994. doi:10.1002/anie.200701713
5. Cobb, A. J. A. *Org. Biomol. Chem.* **2007,** *5,* 3260–3275. doi:10.1039/b709797m
6. Deleavey, G. F.; Damha, M. J. *Chem. Biol.* **2012,** *19,* 937–954. doi:10.1016/j.chembiol.2012.07.011
7. Westheimer, F. H. *Science* **1987,** *235,* 1173–1178. doi:10.1126/science.2434996
8. Benner, S. A.; Hutter, D. *Bioorg. Chem.* **2002,** *30,* 62–80. doi:10.1006/bioo.2001.1232
9. Benner, S. A. *Acc. Chem. Res.* **2004,** *37,* 784–797. doi:10.1021/ar040004z
10. Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. *Science* **1991,** *254,* 1497–1500. doi:10.1126/science.1962210
11. Zhilina, Z. V.; Ziemba, A. J.; Ebbinghaus, S. W. *Curr. Top. Med. Chem.* **2005,** *5,* 1119–1131. doi:10.2174/156802605774370892
12. Karkare, S.; Bhatnagar, D. *Appl. Microbiol. Biotechnol.* **2006,** *71,* 575–586. doi:10.1007/s00253-006-0434-2
13. Lebreton, J.; De Mesmaeker, A.; Waldner, A.; Fritsch, V.; Wolf, R. M.; Freier, S. M. *Tetrahedron Lett.* **1993,** *34,* 6383–6386. doi:10.1016/0040-4039(93)85051-W
14. De Mesmaeker, A.; Lebreton, J.; Waldner, A.; Fritsch, V.; Wolf, R. M.; Freier, S. M. *Synlett* **1993,** 733–736. doi:10.1055/s-1993-22588
15. De Mesmaeker, A.; Waldner, A.; Lebreton, J.; Hoffmann, P.; Fritsch, V.; Wolf, R. M.; Freier, S. M. *Angew. Chem.* **1994,** *106,* 237–240. doi:10.1002/ange.19941060230
    *Angew. Chem. Int. Ed.* **1994,** *33,* 226–229. doi:10.1002/anie.199402261
16. Lebreton, J.; Waldner, A.; Fritsch, V.; Wolf, R. M.; De Mesmaeker, A. *Tetrahedron Lett.* **1994,** *35,* 5225–5228. doi:10.1016/S0040-4039(00)77069-9
17. Rozners, E.; Katkevica, D.; Bizdena, E.; Stromberg, R. *J. Am. Chem. Soc.* **2003,** *125,* 12125–12136. doi:10.1021/ja0360900
18. Selvam, C.; Thomas, S.; Abbott, J.; Kennedy, S. D.; Rozners, E. *Angew. Chem.* **2011,** *123,* 2116–2118. doi:10.1002/ange.201007012
    *Angew. Chem., Int. Ed.* **2011,** *50,* 2068–2070. doi:10.1002/anie.201007012
19. Tanui, P.; Kennedy, S. D.; Lunstad, B. D.; Haas, A.; Leake, D.; Rozners, E. *Org. Biomol. Chem.* **2014,** *12,* 1207–1210. doi:10.1039/c3ob42532k
20. Mutisya, D.; Selvam, C.; Lunstad, B. D.; Pallan, P. S.; Haas, A.; Leake, D.; Egli, M.; Rozners, E. *Nucleic Acids Res.* **2014,** *42,* 6542–6551. doi:10.1093/nar/gku235
21. El-Sagheer, A. H.; Brown, T. *J. Am. Chem. Soc.* **2009,** *131,* 3958–3964. doi:10.1021/ja8065896
22. El-Sagheer, A. H.; Sanzone, A. P.; Gao, R.; Tavassoli, A.; Brown, T. *Proc. Natl. Acad. Sci. U. S. A.* **2011,** *108,* 11338–11343. doi:10.1073/pnas.1101519108

23. Ozaki, H.; Kitamura, M.; Yamana, K.; Murakami, A.; Shimidzu, T. *Bull. Chem. Soc. Jpn.* **1990,** *63,* 1929–1936. doi:10.1246/bcsj.63.1929
24. Moody, H. M.; van Genderen, M. H. P.; Koole, L. H.; Kocken, H. J. M.; Meijer, E. M.; Buck, H. M. *Nucleic Acids Res.* **1989,** *17,* 4769–4782. doi:10.1093/nar/17.12.4769
25. Strauss, J. K.; Roberts, C.; Nelson, M. G.; Switzer, C.; Maher, L. J., III. *Proc. Natl. Acad. Sci. U. S. A.* **1996,** *93,* 9515–9520. doi:10.1073/pnas.93.18.9515
26. Prakash, T. P.; Kawasaki, A. M.; Lesnik, E. A.; Sioufi, N.; Manoharan, M. *Tetrahedron* **2003,** *59,* 7413–7422. doi:10.1016/S0040-4020(03)01104-9
27. Prakash, T. P.; Kawasaki, A. M.; Lesnik, E. A.; Owens, S. R.; Manoharan, M. *Org. Lett.* **2003,** *5,* 403–406. doi:10.1021/ol027131k
28. Prhavc, M.; Prakash, T. P.; Minasov, G.; Cook, P. D.; Egli, M.; Manoharan, M. *Org. Lett.* **2003,** *5,* 2017–2020. doi:10.1021/ol0340991
29. Prakash, T. P.; Püschl, A.; Lesnik, E.; Mohan, V.; Tereshko, V.; Egli, M.; Manoharan, M. *Org. Lett.* **2004,** *6,* 1971–1974. doi:10.1021/ol049470e
30. Jain, M. L.; Bruice, P. Y.; Szabó, I. E.; Bruice, T. C. *Chem. Rev.* **2012,** *112,* 1284–1309. doi:10.1021/cr1004265
31. Blaskó, A.; Dempcy, R. O.; Minyat, E. E.; Bruice, T. C. *J. Am. Chem. Soc.* **1996,** *118,* 7892–7899. doi:10.1021/ja961308m
32. Linkletter, B. A.; Szabo, I. E.; Bruice, T. C. *J. Am. Chem. Soc.* **1999,** *121,* 3888–3896. doi:10.1021/ja984212w
33. Arya, D. P.; Bruice, T. C. *J. Am. Chem. Soc.* **1998,** *120,* 6619–6620. doi:10.1021/ja980629q
34. Arya, D. P.; Bruice, T. C. *J. Am. Chem. Soc.* **1998,** *120,* 12419–12427. doi:10.1021/ja9829416
35. Barawkar, D. A.; Bruice, T. C. *Proc. Natl. Acad. Sci. U. S. A.* **1998,** *95,* 11047–11052. doi:10.1073/pnas.95.19.11047
36. Challa, H.; Bruice, T. C. *Bioorg. Med. Chem. Lett.* **2001,** *11,* 2423–2427. doi:10.1016/S0960-894X(01)00455-3
37. Letsinger, R. L.; Singman, C. N.; Histand, G.; Salunkhe, M. *J. Am. Chem. Soc.* **1988,** *110,* 4470–4471. doi:10.1021/ja00221a089
38. Schmidtgall, B.; Spork, A. P.; Wachowius, F.; Höbartner, C.; Ducho, C. *Chem. Commun.* **2014,** *50,* 13742–13745. doi:10.1039/C4CC06371F
39. McDonald, L. A.; Barbieri, L. R.; Carter, G. T.; Lenoy, E.; Lotvin, J.; Petersen, P. J.; Siegel, M. M.; Singh, G.; Williamson, R. T. *J. Am. Chem. Soc.* **2002,** *124,* 10260–10261. doi:10.1021/ja017748h
40. Tanino, T.; Ichikawa, S.; Shiro, M.; Matsuda, A. *J. Org. Chem.* **2010,** *75,* 1366–1377. doi:10.1021/jo9027193
41. Tanino, T.; Al-Dabbagh, B.; Mengin-Lecreulx, D.; Bouhss, A.; Oyama, H.; Ichikawa, S.; Matsuda, A. *J. Med. Chem.* **2011,** *54,* 8421–8439. doi:10.1021/jm200906r
42. Takeoka, Y.; Tanino, T.; Sekiguchi, M.; Yonezawa, S.; Sakagami, M.; Takahashi, F.; Togame, H.; Tanaka, Y.; Takemoto, H.; Ichikawa, S.; Matsuda, A. *ACS Med. Chem. Lett.* **2014,** *5,* 556–560. doi:10.1021/ml5000096
43. Spork, A. P.; Koppermann, S.; Ducho, C. *Synlett* **2009,** 2503–2507. doi:10.1055/s-0029-1217742
44. Spork, A. P.; Koppermann, S.; Dittrich, B.; Herbst-Irmer, R.; Ducho, C. *Tetrahedron: Asymmetry* **2010,** *21,* 763–766. doi:10.1016/j.tetasy.2010.03.037
45. Spork, A. P.; Ducho, C. *Synlett* **2013,** *24,* 343–346. doi:10.1055/s-0032-1318117
46. Spork, A. P.; Büschleb, M.; Ries, O.; Wiegmann, D.; Boettcher, S.; Mihalyi, A.; Bugg, T. D. H.; Ducho, C. *Chem. – Eur. J.* **2014,** *20,* 15292–15297. doi:10.1002/chem.201404775
47. Spork, A. P.; Ducho, C. *Org. Biomol. Chem.* **2010,** *8,* 2323–2326. doi:10.1039/c003092a
48. Spork, A. P.; Wiegmann, D.; Granitzka, M.; Stalke, D.; Ducho, C. *J. Org. Chem.* **2011,** *76,* 10083–10098. doi:10.1021/jo201935w
49. Khan, A. T.; Mondal, E. *Synlett* **2003,** 694–698. doi:10.1055/s-2003-38360
50. More, J. D.; Finney, N. S. *Org. Lett.* **2002,** *4,* 3001–3003. doi:10.1021/ol026427n
51. Schmidt, U.; Lieberknecht, A.; Schanbacher, U.; Beuttler, T.; Wild, J. *Angew. Chem.* **1982,** *94,* 797–798. doi:10.1002/ange.19820941027 *Angew. Chem. Int. Ed.* **1982,** *21,* 776–777. doi:10.1002/anie.198207761
52. Schmidt, U.; Lieberknecht, A.; Wild, J. *Synthesis* **1984,** 53–60. doi:10.1055/s-1984-30730
53. Schmidt, U.; Wild, J. *Liebigs Ann. Chem.* **1985,** 1882–1894. doi:10.1002/jlac.198519850915
54. Hamzavi, R.; Dolle, F.; Tavitian, B.; Dahl, O.; Nielsen, P. E. *Bioconjugate Chem.* **2003,** *14,* 941–954. doi:10.1021/bc034022x
55. Schmidt, U.; Griesser, H.; Leitenberger, V.; Lieberknecht, A.; Mangold, R.; Meyer, R.; Riedl, B. *Synthesis* **1992,** 487–490. doi:10.1055/s-1992-26143
56. Mazurkiewicz, R.; Kuźnik, A.; Grymel, M.; Kuźnik, N. *Magn. Reson. Chem.* **2005,** *43,* 36–40. doi:10.1002/mrc.1496
57. Ducho, C.; Hamed, R. B.; Batchelar, E. T.; Sorensen, J. L.; Odell, B.; Schofield, C. J. *Org. Biomol. Chem.* **2009,** *7,* 2770–2779. doi:10.1039/b903312b
58. Vineyard, B. D.; Knowles, W. S.; Sabacky, M. J.; Bachman, G. L.; Weinkauff, D. J. *J. Am. Chem. Soc.* **1977,** *99,* 5946–5952. doi:10.1021/ja00460a018
59. Burk, M. J. *J. Am. Chem. Soc.* **1991,** *113,* 8518–8519. doi:10.1021/ja00022a047
60. Masquelin, T.; Broger, E.; Müller, K.; Schmid, R.; Obrecht, D. *Helv. Chim. Acta* **1994,** *77,* 1395–1411. doi:10.1002/hlca.19940770518
61. Eisenhuth, R.; Richert, C. *J. Org. Chem.* **2009,** *74,* 26–37. doi:10.1021/jo8018889
62. Nelson, J. S.; Fearon, K. L.; Nguyen, M. Q.; McCurdy, S. N.; Frediani, J. E.; Foy, M. F.; Hirschbein, B. L. *J. Org. Chem.* **1997,** *62,* 7278–7287. doi:10.1021/jo970801t
63. Krishnakumar, K. S.; Strazewski, P. *Synlett* **2010,** 1055–1058. doi:10.1055/s-0029-1219436
64. Iranpoor, N.; Firouzabadi, H.; Akhlaghinia, B.; Nowrouzi, N. *Tetrahedron Lett.* **2004,** *45,* 3291–3294. doi:10.1016/j.tetlet.2004.02.141
65. Herdewijn, P. A. M. *J. Org. Chem.* **1988,** *53,* 5050–5053. doi:10.1021/jo00256a027

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

# Synthesis of dinucleoside acylphosphonites by phosphonodiamidite chemistry and investigation of phosphorus epimerization

William H. Hersh

Full Research Paper

Open Access

Address:
Department of Chemistry and Biochemistry, Queens College and the Graduate Center of the City University of New York, Queens, NY 11367-1597, USA

Email:
William H. Hersh - william.hersh@qc.cuny.edu

Keywords:
acylphosphine; acyl phosphite; chiral; DFT; NMR; nucleic acids; oligonucleotides

*Beilstein J. Org. Chem.* **2015,** *11,* 184–191.
doi:10.3762/bjoc.11.19

Received: 02 November 2014
Accepted: 15 January 2015
Published: 30 January 2015

This article is part of the Thematic Series "Nucleic acid chemistry".

Guest Editor: H.-A. Wagenknecht



## Abstract

The reaction of the diamidite, $(iPr_2N)_2PH$, with acyl chlorides proceeds with the loss of HCl to give the corresponding acyl diamidites, $RC(O)P(N(iPr)_2)_2$ (R = Me (**7**), Ph (**9**)), without the intervention of sodium to give a phosphorus anion. The structure of **9** was confirmed by single-crystal X-ray diffraction. The coupling of the diamidites **7** and **9** with 5′-*O*-DMTr-thymidine was carried out with *N*-methylimidazolium triflate as the activator to give the monoamidites 3′-*O*-$(P(N(iPr)_2)C(O)R)$-5′-*O*-DMTr-thymidine, and further coupling with 3′-*O*-(*tert*-butyldimethylsilyl)thymidine was carried out with activation by pyridinium trifluoroacetate/*N*-methylimidazole. The new dinucleoside acylphosphonites could be further oxidized, hydrolyzed to the *H*-phosphonates, and sulfurized to give the known mixture of diastereomeric phosphorothioates. The goal of this work was the measurement of the barrier to inversion of the acylphosphonites, which was expected to be low by analogy to the low barrier found in acylphosphines. However, the barrier was found to be high as no epimerization was detected up to 150 °C, and consistent with this, density functional theory calculations give an inversion barrier of over 40 kcal/mol.

## Introduction

Antisense oligonucleotides, which are modified oligonucleotides that bind to mRNA in order to inhibit translation to proteins, have been intensively investigated for decades [1,2]. Modifications of native DNA or RNA are required to permit pharmaceutical use, for example, for conferring nuclease resistance, enhanced binding to mRNA, and recruitment of RNase H. A selection of recent publications [3-7] illustrates the changes that have occurred since the first antisense drug was approved in 1998 [8]. However, despite ongoing clinical trials, only recently has a second antisense drug [3] been approved [9].

Most (albeit not all) antisense reagents, including the two approved drugs, are phosphorothioates, in which one of the terminal oxygen atoms on the phosphodiester internucleotide

linker is replaced by sulfur. A potential issue with phosphorothioates is the creation of a new stereocenter at every phosphorus linker. This results in the formation of an enormous number of diastereomers that may have different properties, such as nuclease stability and RNase H interactions. Some [5,10] but not all [11,12] of the non-phosphorothioate antisense reagents avoid the diastereomer problem, but the extensive study of phosphorothioates is reason enough to investigate the synthesis of P-chiral phosphorothioates [13-19]. However, none of the reported methods seem to allow for routine, high-yield synthesis with high stereocontrol [20-23]. The extension to phosphorothioate RNA for applications of RNA interference has also been described [24].

We recently reported the synthesis of a number of chiral disulfide sulfurizing reagents [25] and the results of the sulfurization of phosphite triesters in order to look for stereoselectivity [26]. In order to change such a method into a practical synthesis route for P-chiral phosphorothioates, a trivalent dinucleoside phosphorus moiety is required that undergoes rapid epimerization under the sulfurization conditions – that is, the goal was to carry out a dynamic kinetic resolution with formation of a single epimeric phosphorothioate.

In order to carry out the required dynamic kinetic resolution, it was hypothesized that acylphosphonites **1** might undergo rapid epimerization by comparison to their known acylphosphine analogs **2** (Figure 1). Mislow previously showed that trialkyl and triarylphosphines slowly undergo epimerization only at temperatures above approximately 130 °C, with activation barriers near 33 kcal/mol [27], but that acylphosphines **2** undergo epimerization much more rapidly [28]. Although acylphosphines are pyramidal, the analogy to amides suggests that the planar transition state for epimerization was greatly stabilized. Literature data from several groups suggests that the inversion barrier is decreased to below 20 kcal/mol [28-30]. In order to extend the comparison from phosphines to phosphites, we recently showed that dinucleoside phosphite triesters, like phosphines, epimerize slowly – in this case at 150 °C with an inversion barrier of 33 kcal/mol [31]. Since the replacement of an alkyl or aryl group in phosphines by an acyl group lowered the epimerization barrier by >13 kcal/mol, it seemed reasonable to propose the same could happen by a similar replacement of an alkoxy group in phosphite triesters.

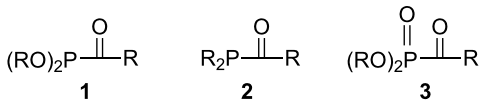



**Figure 1:** Acyl phosphorus compounds.

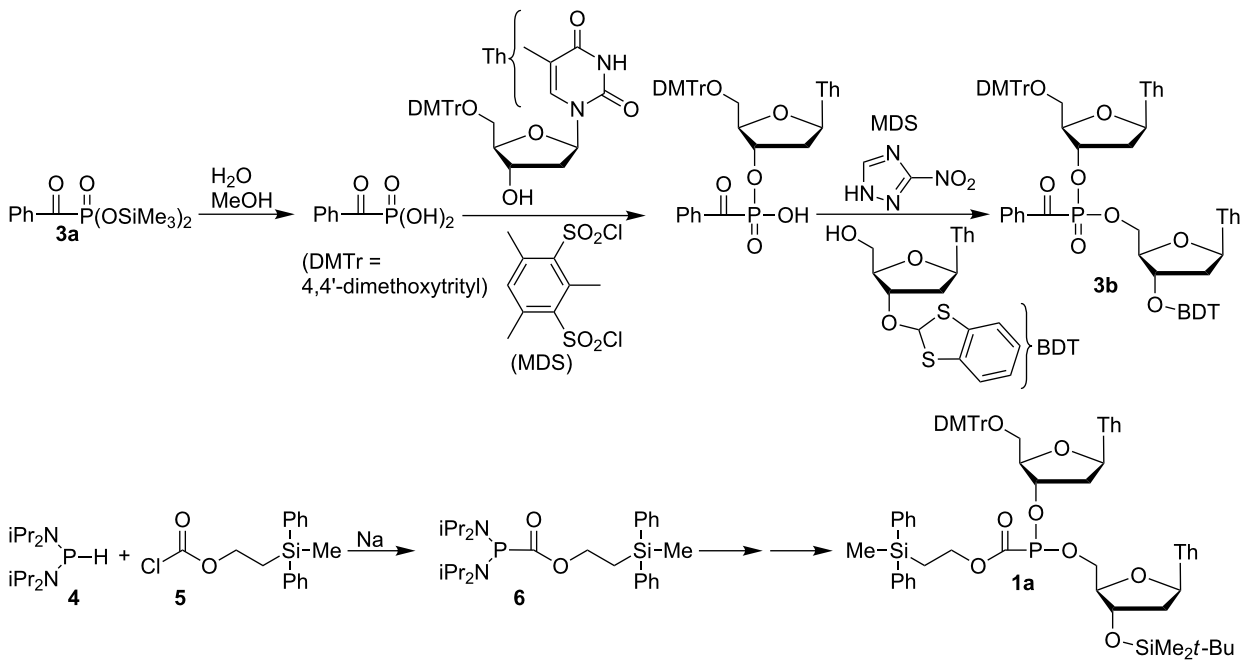
Although the required acylphosphonites are essentially unknown, acylphosphonates **3** (Figure 1) have been extensively studied over the past few decades [32], both for their synthetic utility [33-37] and as compounds of biological interest [38,39]. They have also been used for phosphorothioate synthesis, as will be described shortly. The straightforward route for the synthesis of **3** involves the Arbuzov reaction of a trialkyl phosphite ester with an acid chloride [40], necessarily giving the dialkyl acylphosphonate with identical alkoxy groups. Thus, no comparable synthesis of **1** could be carried out since the dinucleoside acylphosphonite would require different nucleoside moieties.

One example can be found in the literature of a non-Arbuzov route leading to the dinucleoside analog of **3** where as reported by Hata et al., **3a** [41] gives **3b** via sequential coupling [42,43] (Scheme 1). While the deoxygenation of phosphine oxides is known [44,45], deoxygenation of the functionally much more complex phosphonate **3b** to the dinucleoside analog of **1** was not attempted. Only one example of an acylphosphonite analog of **1** was found in the literature, namely, a dinucleoside formate reported by Caruthers et al. [46]. As shown in Scheme 1, the reaction of *H*-phosphonodiamidite **4** with chloroformate **5** gave acyl diamidite **6**, which by following standard nucleoside coupling protocols gave the formate phosphonite **1a**. The initial P–C bond formation yielding **6** was odd, since although the addition of sodium was intended to give the phosphorus anion $(iPr_2N)_2P^-$ from **4**, the authors noted that there was in fact no evidence of any reaction of **4** with sodium. The phosphonite **1a** was reported on the basis of $^{31}P$ NMR analysis to exist as a mixture of two diastereomers, but the possibility of epimerization seems not to have been considered. Instead, **1a** was immediately oxidized to give the phosphonoformate dinucleoside.

In this paper we report: (1) a more general synthesis of acyl diamidites, which most likely accounts for Caruthers' results, including the X-ray crystal structure of the benzoyl analog, (2) the diamidite conversion to dinucleoside acylphosphonites, (3) conversion of these compounds to phosphorothioates, (4) evidence that acylphosphonites, unlike their phosphine analogs, do not undergo facile inversion, and (5) confirmation of this observation by density functional theory (DFT) calculations.

## Results and Discussion

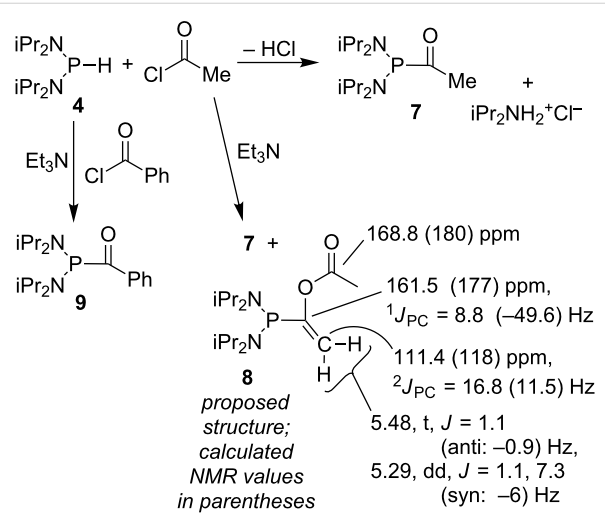
**Synthesis of acylphosphonodiamidites.** Initially we attempted to deprotonate *H*-phosphonodiamidite **4** with butyllithium, but in all cases, the major reaction product with acetic anhydride was simply the starting material. However, a minor product with a peak in the $^{31}P$ NMR spectrum near 63 ppm was observed that was comparable to that exhibited by **6** at 52 ppm. We repeated Caruthers' method [46], which involved a two




**Scheme 1:** Synthesis of a dinucleoside acylphosphonate (**3b**) and a formate diester (**1a**).

hour reflux of **4** with sodium, and similarly detected no visible disappearance of the sodium. No reaction of this mixture with acetic anhydride was detected, beyond the occasional appearance of the 63 ppm peak. Removal of the solvent from the **4**/Na mixture, followed by reaction of the resultant crude material with acetyl chloride, resulted in variable amounts of the 63 ppm peak.

Since there was no evidence that deprotonation of **4** gave rise to the product, the direct reaction of **4** with acetyl chloride in the absence of sodium was examined (Scheme 2). Surprisingly, an immediate reaction occurred with acetyl chloride, giving a yellow-orange solution accompanied by copious production of what was assumed to be gaseous HCl. The hexane-soluble product (**7**, Scheme 2) could be isolated as an orange oil exhibiting a single peak in the $^{31}$P NMR spectrum at 63 ppm, but only in 64% yield. Variable amounts of $iPr_2NH_2^+Cl^-$ also formed, which presumably accounted for the low yield. The characterization included a doublet for the carbonyl methyl group in both the $^1$H and $^{13}$C NMR spectra ($^1$H: 2.27 ppm, $^3J_{PH}$ = 8.8 Hz; $^{13}$C: 30.7 ppm, $^2J_{PC}$ = 49.7 Hz). However, a peak for the carbonyl carbon could not be found nor could any literature precedent be found. A DFT-NMR calculation using Gaussian [47] was carried out (see Supporting Information File 1 for details), with the surprising result that the carbonyl carbon peak was expected at 242 ppm – that is, not at all analogous to an amide chemical shift. Expansion of the standard sweep width beyond 220 ppm indeed led to observation of the missing carbon peak, at 227.9 ppm. The signal was observed as a doublet with $^1J_{PC}$ = 22.4 Hz, in poor agreement with the calculated value of −60 Hz, although both $CH_3$ calculations were in remarkably good agreement ($^{13}$C: 31 ppm, $^2J_{PC}$ = 54 Hz). Investigation of the poor agreement with the one-bond coupling constant has not been attempted, but the cause might be a phosphorus lone pair/carbonyl dihedral angle effect in which the true minimum energy conformation has not been found [48-50]. Once purified, **7** exhibited a strong carbonyl peak in the infrared spectrum at 1654 cm$^{-1}$, a surprisingly low frequency considering the down-




**Scheme 2:** Reaction of an *H*-phosphonodiamidite with acid chlorides.

field $^{13}C$ chemical shift. This inconsistency is perhaps due to an effect caused by the electron-donor diisopropylamide groups.

A rational approach to prevent the formation of the HCl and $iPr_2NH_2^+Cl^-$ that accompany the formation of **7** would involve the addition of a base. However, the use of $Et_3N$ gave a 43:57 mixture of **7** and a new byproduct, **8**, with a $^{31}P$ NMR peak at 49.7 ppm. Although efforts to isolate **8** have been unsuccessful, the proposed structure (Scheme 2; see Supporting Information File 1 for details) is consistent both with formation by base-induced enolization followed by trapping with acetyl chloride, and with the $^1H$, $^{13}C$, and DEPT NMR spectra. For instance, the alkene-like hydrogen signals at 5.48 and 5.29 ppm have a 1.1 Hz coupling constant, which is typical of geminal alkylidene hydrogen atoms. The difference between the 1.1 and 7.3 Hz phosphorus–hydrogen coupling constants is plausible for a vinylphosphine, although the values themselves are low [48]. The NMR spectrum calculated using Gaussian was also in good agreement with experimental observations (Scheme 2; also, $^{31}P$ NMR: observed 50 ppm, calculated 52 ppm, and for comparison, **7**: observed 64 ppm, calculated 70 ppm), although once again, the calculated one-bond coupling constant is in poor agreement.

The reaction of **4** with benzoyl chloride yielded a similar reaction, namely, copious bubbling of gaseous HCl with formation of a red-orange solution. In this case, unlike that with acetyl chloride, two peaks were observed in the $^{31}P$ NMR spectrum, one of which appeared to be due to the desired product, **9**, while the other was far downfield at 121 ppm, perhaps due to $Ph(CO)P(Cl)N(iPr)_2$. The use of $Et_3N$ resulted in a clean reaction presumably since no enolization is possible, leading to **9** in 89% yield after hexane extraction. A carbonyl peak was observed in the $^{13}C$ NMR spectrum, similar to that of **7**, as a doublet at 220.9 ppm ($^1J_{PC}$ = 23.7 Hz), in addition to a strong carbonyl band in the infrared spectrum at 1631 $cm^{-1}$.

**X-ray crystal structure of 9.** Large crystals of **9** could be obtained from hexane, allowing confirmation of the structure by single-crystal X-ray diffraction [51]. An ORTEP drawing is shown in Figure 2, illustrating the steric crowding around phosphorus due to the diisopropylamino groups. One result is that the amines are essentially planar, where the sum of the angles about each nitrogen atom is 360.0(1)°. However, the phosphorus, as expected, is pyramidal with an N–P–N angle of 112.51(5)° and average N–P–C angles of 100.2°. The dihedral angle between the carbonyl and the putative phosphorus lone pair is 108°, that is, only 18° from the angle required for the lone pair and the carbonyl p-orbital to be parallel to each other. The low CO stretching frequency in the infrared spectrum at 1631 $cm^{-1}$ is consistent with the overlap of the carbonyl and lone pair orbitals. However, as noted above, this does not result in an amide-like $^{13}C$ chemical shift.

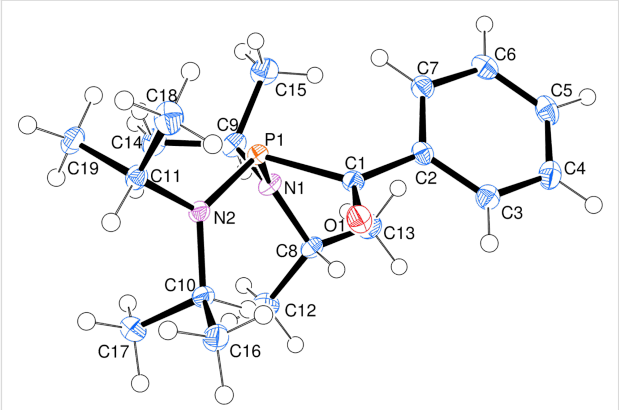
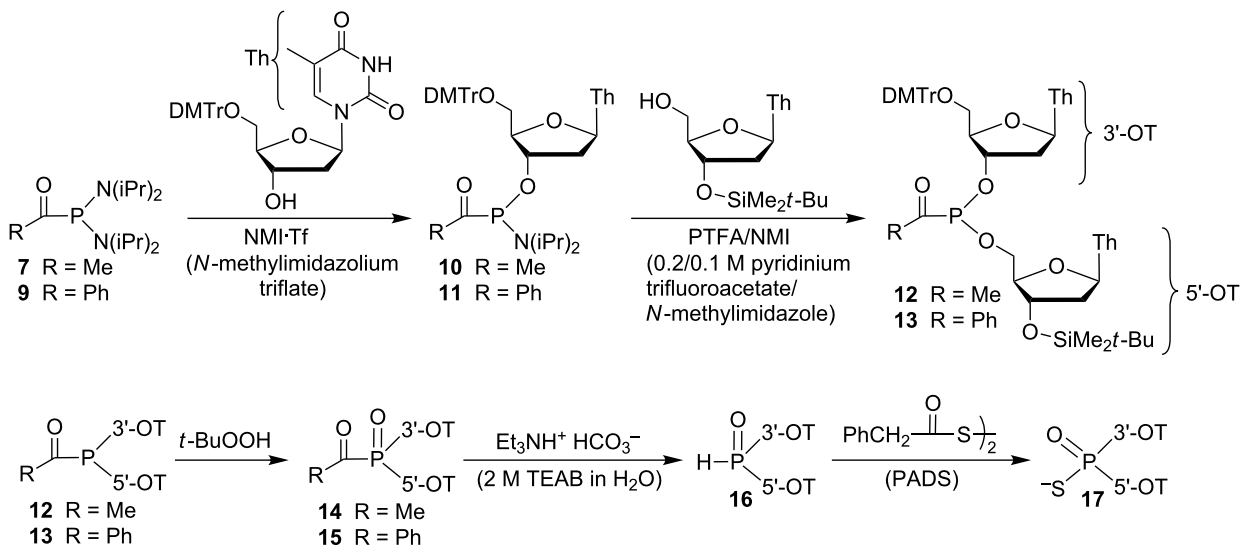



**Figure 2:** ORTEP [52] drawing of **9**. Selected distances (Å) and angles (°): P–N1 1.687(1), P–N2 1.679(1), P–C1 1.879(1), C1–O 1.221(1), C1–C2 1.495(1), C1–P–N1 99.90(5), C1–P–N2 100.59(5), P–C1–O 120.23(8), P–C1–C2 119.78(8), C2–C1–O 119.82(9).

**Nucleoside coupling.** The initial attempts at coupling **7** and **9** with 5′-*O*-DMTr-thymidine were carried out using the Caruthers procedure with 4,5-dicyanoimidazole (DCI) as the activator. While formation of the product was observed, it was not a clean reaction. The *N*-methylimidazolium triflate (NMI·Tf) method reported for the monosubstitution of diamidites was found to be effective [53], resulting in relatively clean formation of amidites **10** and **11** (Scheme 3). The initial reaction produced a yellow foam, which was assumed to be the color of the desired products due to its similarity to **7** and **9**. However, the presence of unreacted 5′-*O*-DMTr-thymidine, the starting acyls **7** and **9**, and downfield $^{31}P$ NMR peaks that are characteristic of diesters necessitated chromatographic separation. The isolation of **11** as a yellow solid as a 42.5:57.5 mixture of phosphorus epimers was carried out without difficulty. Little fractionation of diastereomers occurred during multiple chromatographies. In contrast, the isolation of **10** was somewhat tedious both due to the presence of impurities that were not easily separated and continuous fractionation of the phosphorus epimers that occurred during repeated chromatographies. However, **10** was ultimately obtained as a white foam.

Samples of the fast-moving diastereomer and of a mixture of the two diastereomers of **10** were fully characterized. They exhibited the characteristic carbonyl doublets at 226.9 and 226.8 ppm in the $^{13}C$ NMR spectra (both with $^1J_{PC}$ = 25.3 Hz). However, the infrared spectra could not be used to confirm the presence of the carbonyl since the thymidine ring exhibited a strong carbonyl absorption in the same region (i.e., 5′-*O*-DMTr-thymidine at 1684 $cm^{-1}$, see Supporting Information File 1). The absence of room-temperature epimerization was apparent




**Scheme 3:** Synthesis of dinucleosides.

from both the chromatographic separation as well as the distinct peaks in the NMR spectra, and is well-known (as is facile chromatographic separation) for the analogous β-cyanoethoxy phosphoramidites [54,55]. A similar characterization of the mixture of diastereomers of **11** also exhibited two sets of peaks in the NMR spectra. However, one peculiarity involving only the diisopropylamino moiety was noted, namely, a fluctional process that resulted in broad peaks for the isopropylmethyl signals of both **10** and **11** in the $^{1}$H and $^{13}$C NMR spectra. Since this was not relevant to phosphorus epimerization, it was not further investigated.

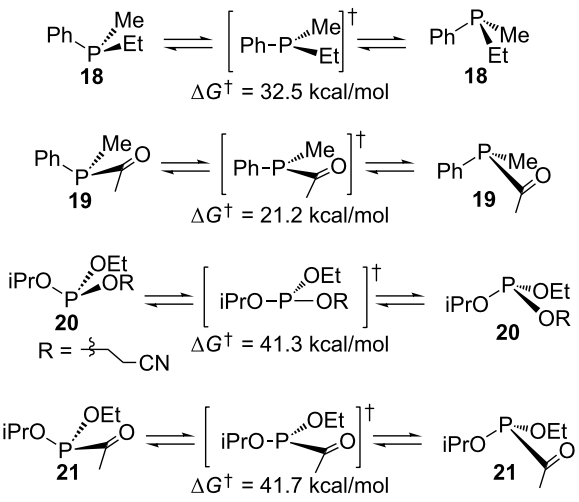
Dinucleoside coupling was carried out using the 2:1 pyridinium trifluoroacetate/*N*-methylimidazole (PTFA/NMI) activator developed at Isis to combine phosphonamidites **10** and **11** with 3′-*tert*-butyldimethylsilyl-protected thymidine to give **12** and **13** [56]. The reactions were monitored by $^{31}$P NMR, giving fairly clean conversion from starting materials to the *P*-epimeric products. The reaction of **10** to give **12** was significantly faster than that of **11** to give **13**. That is, while the former was nearly complete in 30 min, the latter was not complete even in 2.5 h. However, prolonged reaction led to formation of diester impurities that could not be chromatographically separated. While not identified, the impurities could be due to detritylation although the characteristic orange color of the trityl cation was not seen. The diesters were less robust than the precursor monoester amidites. For instance, rapid chromatographic separation of the products from the ammonium salts formed during the synthesis was required to prevent further decomposition. The best procedure was found to involve short reaction times to minimize formation of the diester impurities, and the use of excess phosphonamidite since both **10** and **11** could be chromatographically removed from the dinucleosides. However, even a 50% excess was not sufficient to consume all of the 3′-*O*-(*tert*-butyldimethylsilyl)thymidine, and separation of this alcohol impurity from the acyl dinucleosides could not be achieved.

The characterization of the acyl dinucleosides was accomplished by 1D and 2D $^{1}$H and $^{13}$C NMR spectra, with confirmation of the dinucleoside structures by high-resolution positive ion ESI mass spectrometry. The characteristic carbonyl doublets shifted a few ppm upfield with larger P–C coupling constants than those observed for **10** and **11** (223.28 and 223.25 ppm, $^{1}J_{PC}$ = 38.6 and 41.4 Hz for **12**, 211.6 and 211.1 ppm, $^{1}J_{PC}$ = 44.3 and 42.4 Hz for **13**). Particularly, for **13**, the two deoxyribose hydrogen atom networks could be correlated with the two H1′ muliplets at 6.2–6.4 ppm in the 2D-COSY NMR spectrum. The identity of the two networks is based on the assumption that the downfield H1′ multiplets are due to the DMTr-thymidine, and the upfield H1′ multiplets are due to the *tert*-butyldimethylsilyl thymidine, as is seen for the parent alcohols [31]. For **12**, this analysis could not be carried out with as much certainty due to the larger amount of the 3′-*O*-*tert*-butyldimethylsilyl thymidine impurity. As previously observed for the β-cyanoethoxy dithymidyl triesters [31], the characteristic two pairs of thymidyl methyl doublets (due to allylic coupling to the H6 alkene hydrogen) on the two diastereomers were seen, one pair downfield near 1.8 ppm for the *tert*-butyldimethylsilyl-thymidine, and one pair upfield near 1.4 ppm for the DMTr-thymidine.

**Conversion to phosphorothioates.** An additional confirmation of the structures was possible via oxidation and conversion to the known diastereomeric phosphorothioates. As shown by Hata et al., acylphosphonate **3b** (Scheme 1) was converted to the corresponding phosphorothioate by treatment with a base (to give the *H*-phosphonate via loss of carboxylate) followed by elemental sulfur [43]. As shown in Scheme 3, oxidation of **12** and **13** with anhydrous *tert*-butyl hydroperoxide gave the acylphosphonates **14** and **15**. These products were immediately hydrolyzed by addition of approximately 2 equiv of aqueous triethylammonium bicarbonate (TEAB) to give the *H*-phosphonate **16**. Treatment of **16** with approximately 4 equiv of phenylacetyl disulfide (PADS) [57] gave the known dinucleoside phosphorothioate **17** [26,58].

**Phosphorus epimerization.** During the many attempts to purify the acyl dinucleosides by chromatography, it became apparent that some fractionation of the diastereomeric mixtures occurred. Due to the sensitivity of the acyls, chromatography on long columns of silica (with or without 1% $NEt_3$ to deactivate the silica) gave rise to the loss of acyls, without any improvement in diastereomer separation. However, the individual samples of fractionated diastereomers of **12** or **13** showed no evidence of room-temperature epimerization. Heating of samples at temperatures from 50 to 150 °C over several hours was carried out in acetonitrile, as was done for the epimeric phosphite triesters [31]. Although this was not the most rigorous test (since available samples only ranged in diastereomer ratios from 1:1 to 1:1.4), no detectable change in the diastereomer ratios was observed. More importantly, sample decomposition occurred at all temperatures. Because of this, it would have been impossible to distinguish between small changes in diastereomer ratio and small differences in rates of diastereomer decomposition.

**Calculation of acylphosphonite inversion barrier.** As described in the Introduction, inversion barriers for phosphines, acylphosphines, and phosphites are known. These were calculated first for validation of the method. The structures of chiral phosphine **18**, chiral acylphosphine **19**, and chiral phosphite **20** (Scheme 4) were each minimized by DFT using Gaussian (see Supporting Information File 1 for details), and the vibrational calculation confirmed a local minimum for each structure. While the phosphites are complicated by multiple local minima, which depend on the conformations about the P–O bonds, the energies are comparable (separated by less than 1 kcal/mol) and do not have a significant impact on the inversion barriers. Transition states were located using routines implemented in Gaussian (see Supporting Information File 1 for details), and in all cases the vibrational calculation exhibited one negative frequency that corresponded to the phosphorus inversion vibration. Not surprisingly, the phosphine transition state exhibited a trigonal planar geometry, and the gas-phase energy barrier of 32.5 kcal/mol was in excellent agreement with the values reported by Mislow for reactions carried out in non-polar toluene as a solvent. Similarly, the acylphosphine transition state exhibited a trigonal planar geometry, and the gas-phase energy barrier of 21.2 kcal/mol was also in excellent agreement with the typical values noted above. The phosphite, in contrast, did not yield a trigonal planar transition state, but rather gave a T-shaped transition state. In fact, such a result has been reported for $PF_3$ [59,60] and is due to the presence of the highly electronegative atoms bound to phosphorus. While the alkoxy groups of the phosphite triester are considered to be π-electron donors, they are strongly electronegative. This results in an inversion geometry that is rather different from the phosphines, presumably mediated by oxygen–phosphorus lone pair–lone pair repulsion. The calculated barrier of 41.3 kcal/mol (from the slightly higher ground state geometry) was not in good agreement with that reported, being 8.3 kcal/mol higher. Unlike the other calculations, this one incorporated the polar acetonitrile solvent that was used for the epimerization kinetics [31], but it had essentially no effect on the calculated barrier.




**Scheme 4:** Calculated phosphine, acylphosphine, phosphite, and acylphosphonite inversion barriers.

The transition state calculations were then carried out on the chiral acylphosphonite **21** with the expectation that a high phosphite-like barrier might occur. In fact, the barrier of 41.7 kcal/mol was slightly higher than the phosphite triester calculation. Unlike the phosphite T-shaped transition state, this occurs via a trigonal planar transition state, although like the others with one negative frequency corresponding to phos-

phorus inversion. The high barrier is consistent with the absence of acylphosphonite epimerization of **12** and **13**, and this suggests that any amide-like stabilization of the phosphorus lone pair in a trigonal transition state is eliminated by oxygen–phosphorus lone pair–lone pair repulsion.

## Conclusion

High-yield syntheses have been developed for some novel phosphorus synthons, including the acyl and benzoyldiamidites via the surprising condensation of phosphonodiamidite **4** and the two acid chlorides with elimination of HCl. The nucleoside acylphosphonamidites were also readily prepared, although the benzoyl analog **11** was more easily isolated (albeit as the diastereomeric mixture of *P*-epimers), while the acetyl analog **10** allowed the diastereomers to be separated following repeated chromatographies. Final coupling to give the dinucleoside acylphosphonites proceeded smoothly, but separation of the pure compounds from the nucleoside precursors could not be completely achieved. If the pure acylphosphonites are to be obtained, solid-phase synthesis would probably be required. The chromatography results suggested that room-temperature epimerization at phosphorus did not occur, and heating at 50–150 °C also gave no indication of epimerization. Density functional theory calculations are consistent with this result, giving a calculated inversion barrier of 42 kcal/mol. While one might argue that given the results of such a calculation, the lack of epimerization was a foregone conclusion, it is still important to test theory with experiment. In conclusion, the analogy between phosphines and phosphites, in which the low barrier to acylphosphine inversion suggested that the same might occur with acylphosphonites, is in this case superficial, and the anticipated low-temperature inversion does not occur.

## Supporting Information

Supporting Information contains the experimental procedures, characterization of new compounds, copies of $^{1}H$, $^{13}C$ and $^{31}P$ NMR spectra, IR spectra, details of the X-ray structure determination, and details of the DFT calculations.

### Supporting Information File 1

Experimental.
[http://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-11-19-S1.pdf]

### Supporting Information File 2

CIF file of compound **9**.
[http://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-11-19-S2.cif]

## Acknowledgements

Mass spectra were obtained at the North Carolina State University Mass Spectrometry Facility located in the Department of Chemistry. Financial support from the City University of New York PSC-CUNY Research Award Program is gratefully acknowledged.

## References

1. Matsukura, M.; Shinozuka, K.; Zon, G.; Mitsuya, H.; Reitz, M.; Cohen, J. S.; Broder, S. *Proc. Natl. Acad. Sci. U. S. A.* **1987,** *84,* 7706–7710. doi:10.1073/pnas.84.21.7706
2. Uhlmann, E.; Peyman, A. *Chem. Rev.* **1990,** *90,* 543–584. doi:10.1021/cr00102a001
3. Crooke, S. T.; Geary, R. S. *Br. J. Clin. Pharmacol.* **2013,** *76,* 269–276. doi:10.1111/j.1365-2125.2012.04469.x
4. Lee, J. J. A.; Yokota, T. *J. Pers. Med.* **2013,** *3,* 144–176. doi:10.3390/jpm3030144
5. Šípová, H.; Špringer, T.; Rejman, D.; Šimák, O.; Petrová, M.; Novák, P.; Rosenbergová, S.; Páv, O.; Liboska, R.; Barvík, I.; Štěpánek, J.; Rosenberg, I.; Homola, J. *Nucleic Acids Res.* **2014,** *42,* 5378–5389. doi:10.1093/nar/gku125
6. Prakash, T. P.; Graham, M. J.; Yu, J.; Carty, R.; Low, A.; Chappell, A.; Schmidt, K.; Zhao, C.; Aghajan, M.; Murray, H. F.; Riney, S.; Booten, S. L.; Murray, S. F.; Gaus, H.; Crosby, J.; Lima, W. F.; Guo, S.; Monia, B. P.; Swayze, E. E.; Seth, P. P. *Nucleic Acids Res.* **2014,** *42,* 8796–8807. doi:10.1093/nar/gku531
7. Magen, I.; Hornstein, E. *Brain Res.* **2014,** *1584,* 116–128. doi:10.1016/j.brainres.2014.04.005
8. Marwick, C. *JAMA* **1998,** *280,* 871. doi:10.1001/jama.280.10.871-JMN0909-6-1
9. Jiang, K. *Nat. Med.* **2013,** *19,* 252. doi:10.1038/nm0313-252
10. Marshall, W. S.; Caruthers, M. H. *Science* **1993,** *259,* 1564–1570. doi:10.1126/science.7681216
11. Li, P.; Sergueeva, Z. A.; Dobrikov, M.; Shaw, B. R. *Chem. Rev.* **2007,** *107,* 4746–4796. doi:10.1021/cr050009p
12. Uehara, S.; Hiura, S.; Higashida, R.; Oka, N.; Wada, T. *J. Org. Chem.* **2014,** *79,* 3465–3472. doi:10.1021/jo500185b
13. Stec, W. J.; Grajkowski, A.; Kobylanska, A.; Karwowski, B.; Koziolkiewicz, M.; Misiura, K.; Okruszek, A.; Wilk, A.; Guga, P.; Boczkowska, M. *J. Am. Chem. Soc.* **1995,** *117,* 12019–12029. doi:10.1021/ja00154a001
14. Stec, W. J.; Karwowski, B.; Boczkowska, M.; Guga, P.; Koziołkiewicz, M.; Sochacki, M.; Wieczorek, M. W.; Blaszczyk, J. *J. Am. Chem. Soc.* **1998,** *120,* 7156–7167. doi:10.1021/ja973801j
15. Wang, J.-C.; Just, G. *J. Org. Chem.* **1999,** *64,* 8090–8097. doi:10.1021/jo990242l
16. Wilk, A.; Grajkowski, A.; Phillips, L. R.; Beaucage, S. L. *J. Am. Chem. Soc.* **2000,** *122,* 2149–2156. doi:10.1021/ja991773u
17. Yu, D.; Kandimalla, E. R.; Roskey, A.; Zhao, Q.; Chen, L.; Chen, J.; Agrawal, S. *Bioorg. Med. Chem.* **2000,** *8,* 275–284. doi:10.1016/S0968-0896(99)00275-8
18. Almer, H.; Szabo, T.; Stawinski, J. *Chem. Commun.* **2004,** 290–291. doi:10.1039/b311912b
19. Oka, N.; Yamamoto, M.; Sato, T.; Wada, T. *J. Am. Chem. Soc.* **2008,** *130,* 16031–16037. doi:10.1021/ja805780u
20. Lu, Y. *Mini-Rev. Med. Chem.* **2006,** *6,* 319–330. doi:10.2174/138955706776073439

21. Guga, P. *Curr. Top. Med. Chem.* **2007,** *7,* 695–713. doi:10.2174/156802607780487786
22. Oka, N.; Wada, T. *Chem. Soc. Rev.* **2011,** *40,* 5829–5843. doi:10.1039/c1cs15102a
23. Guga, P.; Koziołkiewicz, M. *Chem. Biodiversity* **2011,** *8,* 1642–1681. doi:10.1002/cbdv.201100130
24. Nukaga, Y.; Yamada, K.; Ogata, T.; Oka, N.; Wada, T. *J. Org. Chem.* **2012,** *77,* 7913–7922. doi:10.1021/jo301052v
25. Mukhlall, J. A.; Noll, B. C.; Hersh, W. H. *J. Sulfur Chem.* **2011,** *32,* 199–212. doi:10.1080/17415993.2011.580346
26. Mukhlall, J. A.; Hersh, W. H. *Nucleosides, Nucleotides Nucleic Acids* **2011,** *30,* 706–725. doi:10.1080/15257770.2011.597366
27. Baechler, R. D.; Mislow, K. *J. Am. Chem. Soc.* **1970,** *92,* 3090–3093. doi:10.1021/ja00713a028
28. Egan, W.; Mislow, K. *J. Am. Chem. Soc.* **1971,** *93,* 1805–1806. doi:10.1021/ja00736a053
29. Kostyanovskii, R. G.; Fomichev, A. A.; Zagurskaya, L. M.; Zakharov, K. S. *Izv. Akad. Nauk SSSR, Ser. Khim.* **1973,** 1915–1916.
30. Vedejs, E.; Donde, Y. *J. Am. Chem. Soc.* **1997,** *119,* 9293–9294. doi:10.1021/ja970369x
31. Mukhlall, J. A.; Hersh, W. H. *Inorg. Chim. Acta* **2011,** *369,* 62–70. doi:10.1016/j.ica.2010.11.022
32. McKenna, C. E.; Kashemirov, B. A. *Top. Curr. Chem.* **2002,** *220,* 201–238. doi:10.1007/3-540-45731-3_8
33. Evans, D. A.; Johnson, J. S. *J. Am. Chem. Soc.* **1998,** *120,* 4895–4896. doi:10.1021/ja980423p
34. Jang, K. P.; Hutson, G. E.; Johnston, R. C.; McCusker, E. O.; Cheong, P. H.-Y.; Scheidt, K. A. *J. Am. Chem. Soc.* **2014,** *136,* 76–79. doi:10.1021/ja410932t
35. Wen, Y.-Q.; Hertzberg, R.; Moberg, C. *J. Org. Chem.* **2014,** *79,* 6172–6178. doi:10.1021/jo500895u
36. Demir, A. S.; Esiringü, I.; Göllü, M.; Reis, Ö. *J. Org. Chem.* **2009,** *74,* 2197–2199. doi:10.1021/jo8026627
37. Jiang, H.; Paixão, M. W.; Monge, D.; Jørgensen, K. A. *J. Am. Chem. Soc.* **2010,** *132,* 2775–2783. doi:10.1021/ja9097803
38. Kluger, R.; Pike, D. C. *J. Am. Chem. Soc.* **1977,** *99,* 4504–4506. doi:10.1021/ja00455a052
39. Breuer, E.; Safadi, M.; Chorev, M.; Gibson, D. *J. Org. Chem.* **1990,** *55,* 6147–6153. doi:10.1021/jo00312a021
40. Berlin, K. D.; Taylor, H. A. *J. Am. Chem. Soc.* **1964,** *86,* 3862–3866. doi:10.1021/ja01072a053
41. Sekine, M.; Satoh, M.; Yamagata, H.; Hata, T. *J. Org. Chem.* **1980,** *45,* 4162–4167. doi:10.1021/jo01309a019
42. Kume, A.; Fujii, M.; Sekine, M.; Hata, T. *J. Org. Chem.* **1984,** *49,* 2139–2143. doi:10.1021/jo00186a012
43. Fujii, M.; Ozaki, K.; Sekine, M.; Hata, T. *Tetrahedron* **1987,** *43,* 3395–3407. doi:10.1016/S0040-4020(01)81630-6
44. Krenske, E. H. *J. Org. Chem.* **2012,** *77,* 1–4. doi:10.1021/jo202370x
45. Krenske, E. H. *J. Org. Chem.* **2012,** *77,* 3969–3977. doi:10.1021/jo300346g
46. Yamada, C. M.; Dellinger, D. J.; Caruthers, M. H. *J. Am. Chem. Soc.* **2006,** *128,* 5251–5261. doi:10.1021/ja060112b
47. *Gaussian 09,* Revision D.01; Gaussian, Inc.: Wallingford, CT, 2009.
48. Fedorov, S. V.; Krivdin, L. B.; Rusakov, Y. Yu.; Chernysheva, N. A.; Mikhailenko, V. L. *Magn. Reson. Chem.* **2010,** *48,* S48–S55. doi:10.1002/mrc.2614
49. Fedorov, S. V.; Krivdin, L. B.; Rusakov, Y. Yu.; Ushakov, I. A.; Istomina, N. V.; Belogorlova, N. A.; Malysheva, S. F.; Gusarova, N. K.; Trofimov, B. A. *Magn. Reson. Chem.* **2009,** *47,* 288–299. doi:10.1002/mrc.2386
50. Hersh, W. H.; Lam, S. T.; Moskovic, D. J.; Panagiotakis, A. J. *J. Org. Chem.* **2012,** *77,* 4968–4979. doi:10.1021/jo3003776
51. Cambridge Crystallographic Data Centre, deposition number CCDC 1030743.
52. Farrugia, L. J. *J. Appl. Crystallogr.* **1997,** *30,* 565. doi:10.1107/S0021889897003117
53. Xie, C.; Staszak, M. A.; Quatroche, J. T.; Sturgill, C. D.; Khau, V. V.; Martinelli, M. J. *Org. Process Res. Dev.* **2005,** *9,* 730–737. doi:10.1021/op050077d
54. Stec, W. J.; Zon, G. *Tetrahedron Lett.* **1984,** *25,* 5279–5282. doi:10.1016/S0040-4039(01)81583-5
55. Cheruvallath, Z. S.; Sasmor, H.; Cole, D. L.; Ravikumar, V. T. *Nucleosides, Nucleotides Nucleic Acids* **2000,** *19,* 533–543. doi:10.1080/15257770008035005
56. Eleuteri, A.; Capaldi, D. C.; Krotz, A. H.; Cole, D. L.; Ravikumar, V. T. *Org. Process Res. Dev.* **2000,** *4,* 182–189. doi:10.1021/op9900378
57. Cheruvallath, Z. S.; Carty, R. L.; Moore, M. N.; Capaldi, D. C.; Krotz, A. H.; Wheeler, P. D.; Turney, B. J.; Craig, S. R.; Gaus, H. J.; Scozzari, A. N.; Cole, D. L.; Ravikumar, V. T. *Org. Process Res. Dev.* **2000,** *4,* 199–204. doi:10.1021/op990077b
58. Bartoszewicz, A.; Kalek, M.; Stawinski, J. *J. Org. Chem.* **2008,** *73,* 5029–5038. doi:10.1021/jo8006072
59. Dixon, D. A.; Arduengo, A. J.; Fukunaga, T. *J. Am. Chem. Soc.* **1986,** *108,* 2461–2462. doi:10.1021/ja00269a063
60. Gilheany, D. G. *Chem. Rev.* **1994,** *94,* 1339–1374. doi:10.1021/cr00029a008


## License and Terms

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

# TEMPO-derived spin labels linked to the nucleobases adenine and cytosine for probing local structural perturbations in DNA by EPR spectroscopy

Dnyaneshwar B. Gophane and Snorri Th. Sigurdsson*

Full Research Paper

Open Access

Address:
University of Iceland, Department of Chemistry, Science Institute, Dunhaga 3, 107 Reykjavik, Iceland

Email:
Snorri Th. Sigurdsson* - snorrisi@hi.is

* Corresponding author

Keywords:
aminoxyl radical; ESR spectroscopy; nitroxide; nucleic acids; site-directed spin labeling (SDSL); spin labels; structure-dependent dynamics

*Beilstein J. Org. Chem.* **2015,** *11,* 219–227.
doi:10.3762/bjoc.11.24

Received: 18 November 2014
Accepted: 15 January 2015
Published: 09 February 2015

This article is part of the Thematic Series "Nucleic acid chemistry".

Guest Editor: H.-A. Wagenknecht



## Abstract

Three 2′-deoxynucleosides containing semi-flexible spin labels, namely $^{T}$A, $^{U}$A and $^{U}$C, were prepared and incorporated into deoxyoligonucleotides using the phosphoramidite method. All three nucleosides contain 2,2,6,6-tetramethylpiperidine-1-oxyl (TEMPO) connected to the exocyclic amino group; $^{T}$A directly and $^{U}$A as well as $^{U}$C through a urea linkage. $^{T}$A and $^{U}$C showed a minor destabilization of a DNA duplex, as registered by a small decrease in the melting temperature, while $^{U}$A destabilized the duplex by more than 10 °C. Circular dichroism (CD) measurements indicated that all three labels were accommodated in B-DNA duplex. The mobility of the spin label $^{T}$A varied with different base-pairing partners in duplex DNA, with the $^{T}$A•T pair being the least mobile. Furthermore, $^{T}$A showed decreased mobility under acidic conditions for the sequences $^{T}$A•C and $^{T}$A•G, to the extent that the EPR spectrum of the latter became nearly superimposable to that of $^{T}$A•T. The reduced mobility of the $^{T}$A•C and $^{T}$A•G mismatches at pH 5 is consistent with the formation of $^{T}$AH$^{+}$•C and $^{T}$AH$^{+}$•G, in which protonation of N1 of A allows the formation of an additional hydrogen bond to N3 of C and N7 of G, respectively, with G in a syn-conformation. The urea-based spin labels $^{U}$A and $^{U}$C were more mobile than $^{T}$A, but still showed a minor variation in their EPR spectra when paired with A, G, C or T in a DNA duplex. $^{U}$A and $^{U}$C had similar mobility order for the different base pairs, with the lowest mobility when paired with C and the highest when paired with T.

## Introduction

The knowledge about structures and conformational dynamics of nucleic acids, as well as other biomolecules, is essential to understand their biological functions, including interactions with other molecules. The exact atom-to-atom structural information can be obtained by X-ray crystallography [1-6] and nuclear magnetic resonance (NMR) spectroscopy [7-12]. Electron paramagnetic resonance (EPR) and fluorescence spectroscopies are nowadays routinely used to study global structures

and conformational changes under biologically relevant conditions through the determination of intermediate to long-range distances [13-30]. EPR spectroscopy can also give information about the relative orientation of two rigid spin labels [31-35]. Small angle X-ray scattering is also frequently used to study global structures of large molecules and molecular assemblies [36-39].

Local structural perturbations in nucleic acids can be studied with some of the aforementioned techniques. For example, NMR has been used to study hydrogen-bonding interactions [11,40-42], non-native base-pairing properties of nucleobases [43-46] and their dynamics [42,47,48]. Fluorescence spectroscopy, using environmentally sensitive fluorescent nucleosides has been used for detection of local structural perturbations [49-57], including the investigation of single-base mismatches [51,54,56,58,59], abasic sites [60] as well as nick sites in duplex DNA [61], and ligand-induced folding of riboswitches [62,63].

Continuous wave (CW) EPR spectroscopy can be used for the determination of structure-dependent dynamics based on the line-shape analysis of the EPR spectra [64-73]. The spin labels for such experiments are attached to the nucleotide via a flexible or a semi-flexible tether, which allows some motion of the spin label independent of the nucleic acid. Spin-label motion is affected by the local surroundings of the label, which in turn is reflected in the shape of the EPR spectra. We have previously used the spin label $^{T}$C [69], containing a 2,2,6,6-tetramethylpiperidine-1-oxyl (TEMPO) moiety conjugated to the exocyclic amino group of C (Figure 1A), to identify the base to which it is paired with in duplex DNA [69]. In other words, this label could not only distinguish between pairing with guanine and a mismatch but was also able to pinpoint the base-pairing partner. Furthermore, $^{T}$C revealed a flanking-base dependent variation in the EPR spectra, showing that minor structural variations in the local surroundings of a nucleic acid groove can be detected with spin labels by EPR spectroscopy.

Here, we describe the use of an analogous derivative of A, namely $^{T}$A, in which a TEMPO moiety is conjugated to the exocyclic amino group of 2′-deoxyadenosine (Figure 1B), to study local perturbations for a purine base in DNA. We show that $^{T}$A can indeed be used to differentiate between different base-pairing partners, albeit not as clearly as $^{T}$C. Lower pH causes noticeable changes in the EPR spectra for $^{T}$A, in particular for the $^{T}$A•G and $^{T}$A•C mismatches, presumably because of protonation of the base. We have also prepared urea-linked spin-labeled derivatives of 2′-deoxycytidine ($^{U}$C) and 2′-deoxyadenosine ($^{U}$A) and incorporated them into DNA duplexes (Figure 1C and D). These labels provide additional possibilities for hydrogen-bonding through the urea linkage but are also more flexible than $^{T}$A or $^{T}$C. In spite of the increased flexibility of $^{U}$C and $^{U}$A, inspection of the line-shape of their CW EPR

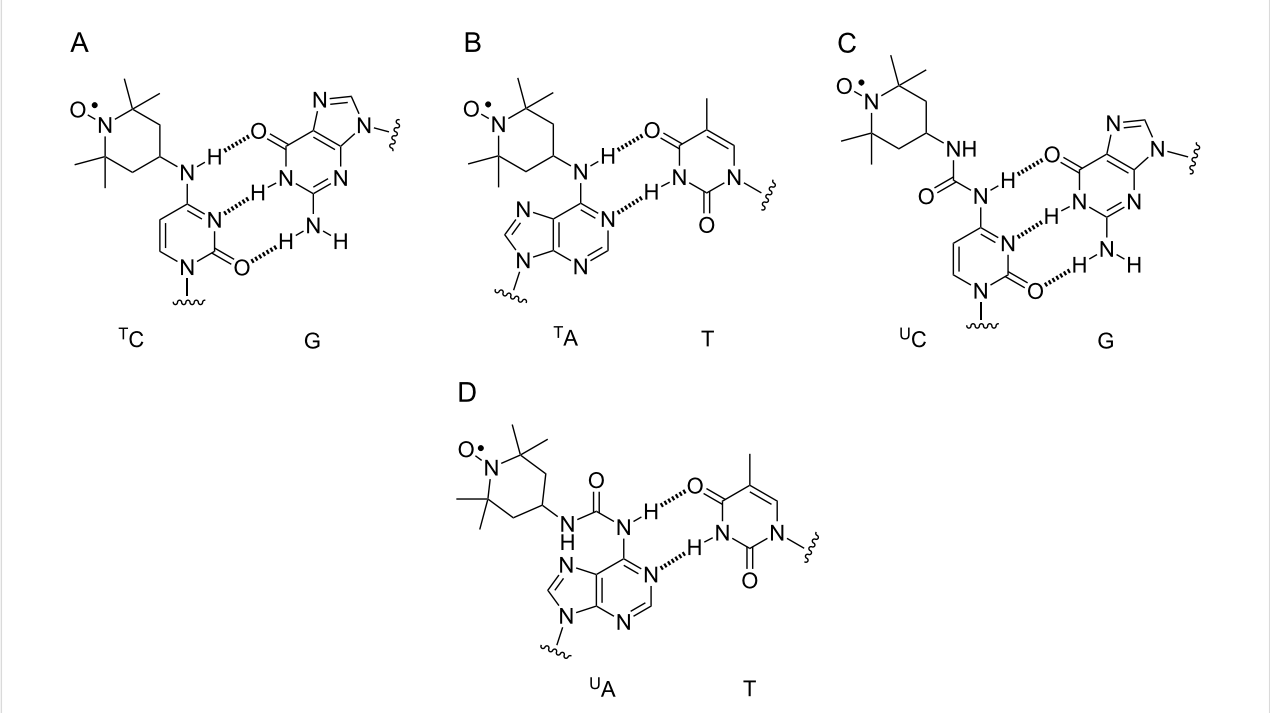



**Figure 1:** Base pairing of $^{T}$C with G (A), $^{T}$A with T (B), $^{U}$C with G (C) and $^{U}$A with T (D).

spectra reveals subtle differences between the four base-pairing partners A, T, G and C when placed in a DNA duplex.

## Results and Discussion

### Synthesis of $^{T}$A, $^{U}$A, $^{U}$C and their corresponding phosphoramidites

The spin-labeled nucleoside $^{T}$A and its corresponding phosphoramidite were prepared by a previously reported procedure [74] with minor modifications. The synthesis started with the reaction between 3′,5′-diacetyl-2′-deoxyinosine (**1**) and 2,4,6-triisopropylbenzenesulfonyl chloride to obtain compound **2** (Scheme 1). Coupling of **2** with 4-amino-TEMPO gave **3** in good yields and deprotection of the acetyl groups gave $^{T}$A, which was tritylated and phosphitylated to yield compounds **4** and **5**, respectively.

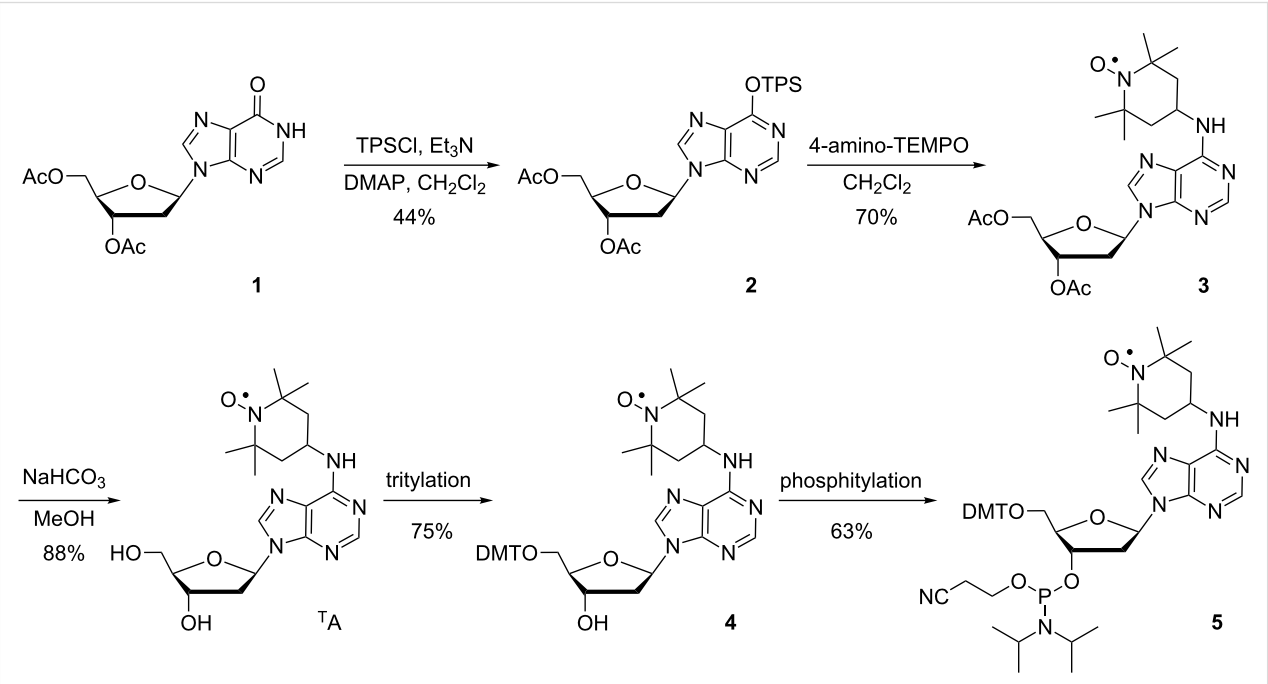
For the synthesis of $^{U}$A, 3′,5′-TBDMS-protected 2′-deoxyadenosine [75] (**6**) was reacted with 4-isocyanato-TEMPO (**7**) [64,76], which gave compound **8** in low yields (Scheme 2). Cleavage of the TBDMS groups using TBAF gave spin-labeled nucleoside $^{U}$A, which was tritylated to give compound **9** and phosphitylated to give phosphoramidite **10**.

The synthesis of $^{U}$C, the urea-cytidine analogue, started by reaction of 3′,5′-TBDMS-protected 2′-deoxycytidine (**11**) with 4-isocyanato-TEMPO (**7**) [64,76] to give **12** in good yields (Scheme 3). Removal of the TBDMS groups using TBAF gave spin-labeled nucleoside $^{U}$C, which was tritylated to give compound **13** and subsequent phosphitylation yielded phosphoramidite **14**.

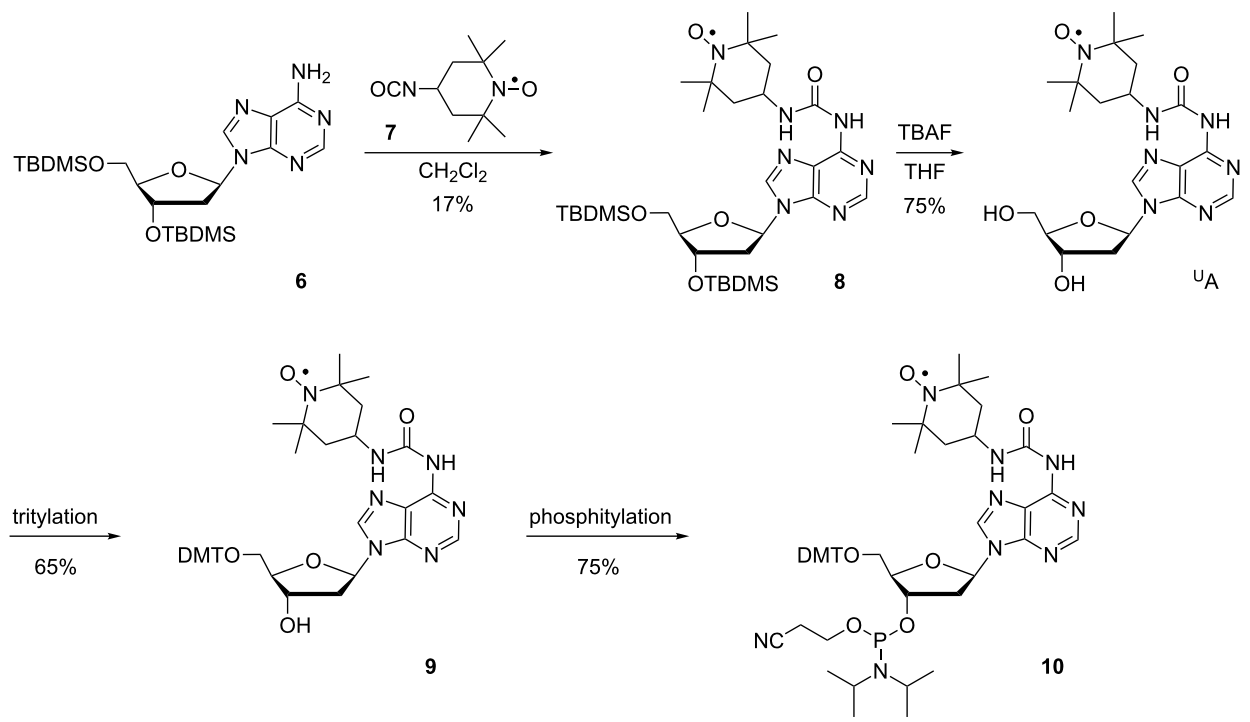
### Synthesis and characterization of $^{T}$A-, $^{U}$A- and $^{U}$C-containing oligonucleotides

The phosphoramidites of $^{T}$A (**5**), $^{U}$A (**10**) and $^{U}$C (**14**) were used to incorporate the spin-labeled nucleosides into DNA oligonucleotides using solid-phase synthesis [77]. The low stability of the nitroxide functional group in the TEMPO moiety towards acids lead to almost ca. 50% reduction of the nitroxide during oligonucleotide synthesis, which utilized dichloroacetic acid for the removal of the trityl groups. However, in spite of low yields, the spin-labeled oligonucleotides were readily separated from those containing the reduced spin label by denaturing polyacrylamide gel electrophoresis. The incorporation of $^{T}$A, $^{U}$A and $^{U}$C into DNA was confirmed by MALDI–TOF mass spectrometry (Table S1, Supporting Information File 1). Circular dichroism measurements showed that the incorporation of $^{T}$A, $^{U}$A and $^{U}$C does not alter the B-DNA conformation of DNA duplexes containing these modifications (Figure S1, Supporting Information File 1).

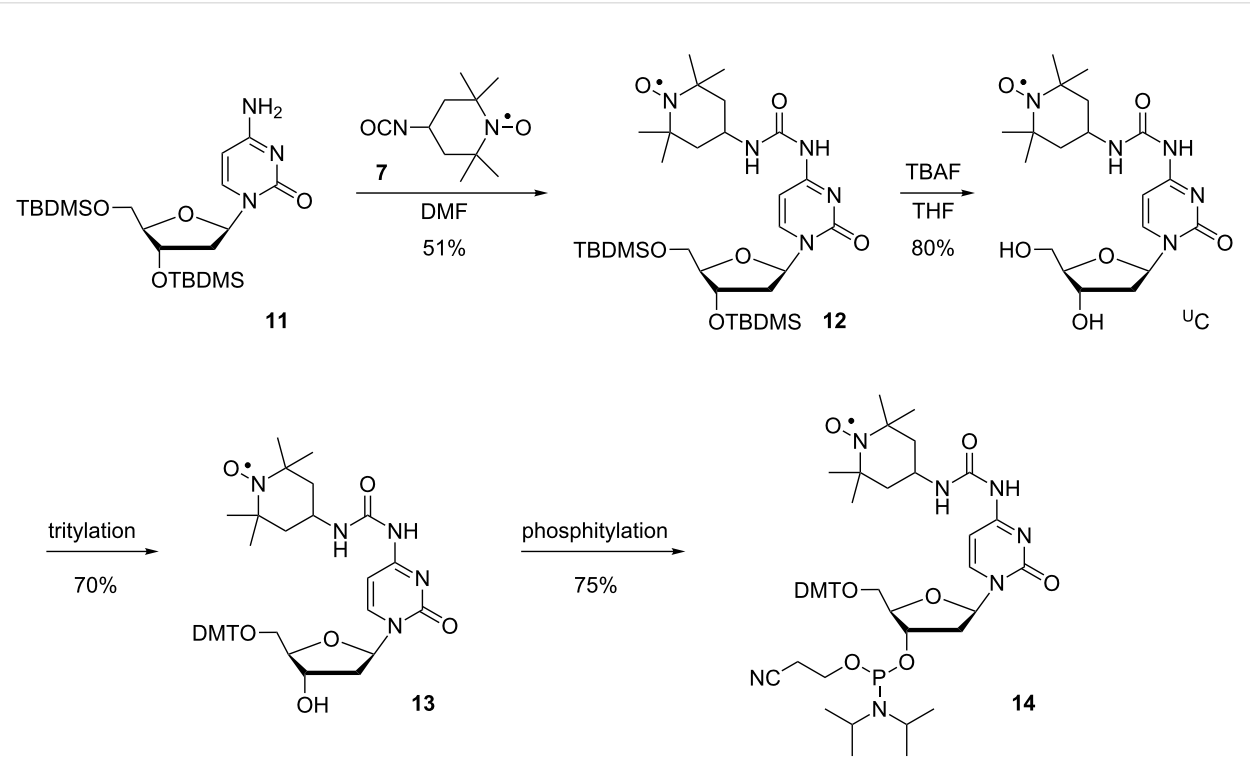
The thermal denaturation studies indicated only a minor destabilization of DNA the duplexes when $^{T}$A was paired with T ($T_m$ was only 2.7 °C lower), compared to the natural nucleoside A (Table S2, Supporting Information File 1). The least stable base pairing was observed with the $^{T}$A•A mismatch, which showed destabilization by 8.2 °C, compared to an A•A mismatch. The




**Scheme 1:** Synthesis of nucleoside $^{T}$A and its corresponding phosphoramidite **5**. TPS = 2,4,6-triisopropylbenzenesulfonyl.




**Scheme 2:** Synthesis of nucleoside $^{U}$A and its corresponding phosphoramidite **10**.




**Scheme 3:** Synthesis of nucleoside $^{U}$C and its corresponding phosphoramidite **14**.

duplex stability order as a function of base-paring with $^{T}$A was $^{T}$A•T > $^{T}$A•G > $^{T}$A•C > $^{T}$A•A, consistent with the order of stability previously reported for A [48]. Replacing C with $^{U}$C opposite G in a DNA duplex resulted in a 3.9 °C decrease in the melting temperature ($T_m$) and a stability order of $^{U}$C•G > $^{U}$C•A > $^{U}$C•T > $^{U}$C•C (Table S3, Supporting Information File 1). In contrast to $^{T}$A and $^{U}$C, $^{U}$A had a large destabilizing effect on DNA duplexes (ca. 11 °C, Table S4, Supporting Information File 1). Interestingly, the melting temperatures of the DNA duplexes containing the base pairs $^{U}$A•T, $^{U}$A•G and $^{U}$A•A were nearly identical, whereas the $^{U}$A•C mismatch showed a further decrease in melting temperature of ca. 9 °C.

## EPR analysis of $^{T}$A-, $^{U}$C-and $^{U}$A-labeled DNA duplexes

To investigate the mobility of $^{T}$A in duplex DNA, we analyzed the EPR spectra of the four 14-mer DNA duplexes 5′-d(GACCTCG$^{T}$AATCGTG)•5′-d(CACGATYCGAGGTC), where Y is either T, C, G or A (Figure 2). The EPR spectrum of the $^{T}$A•T pair was broadest, which is consistent with $^{T}$A forming a Watson–Crick base pair with T, and thereby restricting the rotation around the C6–N6 bond through hydrogen bonding to N6, and consequently slowing the motion of the label (Figure 1B). On the other hand, the spectrum of $^{T}$A•A was the narrowest and thereby indicating the highest mobility, while the EPR spectrum of $^{T}$A•G was slightly broader than $^{T}$A•A. Although base pairing schemes can be drawn for $^{T}$A•G and $^{T}$A•A that involve hydrogen bonding to N6 of $^{T}$A (Figure S2C and S2D, Supporting Information File 1) [44,48,78-80], the increased mobility could be the result of the label being pushed further into the major groove of the DNA duplex, due to the space-demanding purine–purine pairs [81], where the spin-label mobility would be less affected by the local surroundings.

The mobility of $^{T}$A•C at pH 7, as judged by its EPR spectrum (Figure 2A), was between that of $^{T}$A•T and $^{T}$A•G. Previous NMR studies of the $^{T}$A•C mismatch at pH 8.5 [47] and 8.9 [82] showed one hydrogen bond, located between N6 of $^{T}$A and N3 of cytidine (Figure S2B, Supporting Information File 1). If $^{T}$A is protonated on N1 to form $^{T}$AH$^{+}$, it could form a wobble-pair with C [47,83] (Figure 3A), which would be expected to decrease the mobility of the spin label. The apparent p$K_a$ of the proton on N1 has been determined by NMR studies to be 7.2 [47], which means that more than half of the $^{T}$A•C pairs would be protonated at pH 7. To explore if further reduction in mobility (due to conversion of the $^{T}$A•C pair to the $^{T}$AH$^{+}$•C pair) would be detected by EPR, its spectrum was also recorded at pH 5 (Figure 2B). Indeed, comparison of the EPR spectra of $^{T}$A•C at pH 5 and pH 7 clearly shows further broadening at the lower pH, almost to that of the $^{T}$A•T pair, and is consistent with the formation of the $^{T}$AH$^{+}$•C pair.

The $^{T}$A•T and $^{T}$A•A pairs had similar EPR spectra at pH 7 and 5 (Figure 2B). However, significant broadening was observed in the EPR spectrum of the $^{T}$A•G mismatch at pH 5. In fact, the spectrum of $^{T}$A•G is nearly superimposable to that of the $^{T}$A•T pair, indicating reduction in mobility due to formation of another hydrogen bond at pH 5. Studies by NMR and X-ray crystallography have shown that $^{T}$AH$^{+}$•G pairs form when $^{T}$A•G mismatches are incubated at pH below 5.6, in which the G has flipped into a syn conformation and the O6 and N7 form

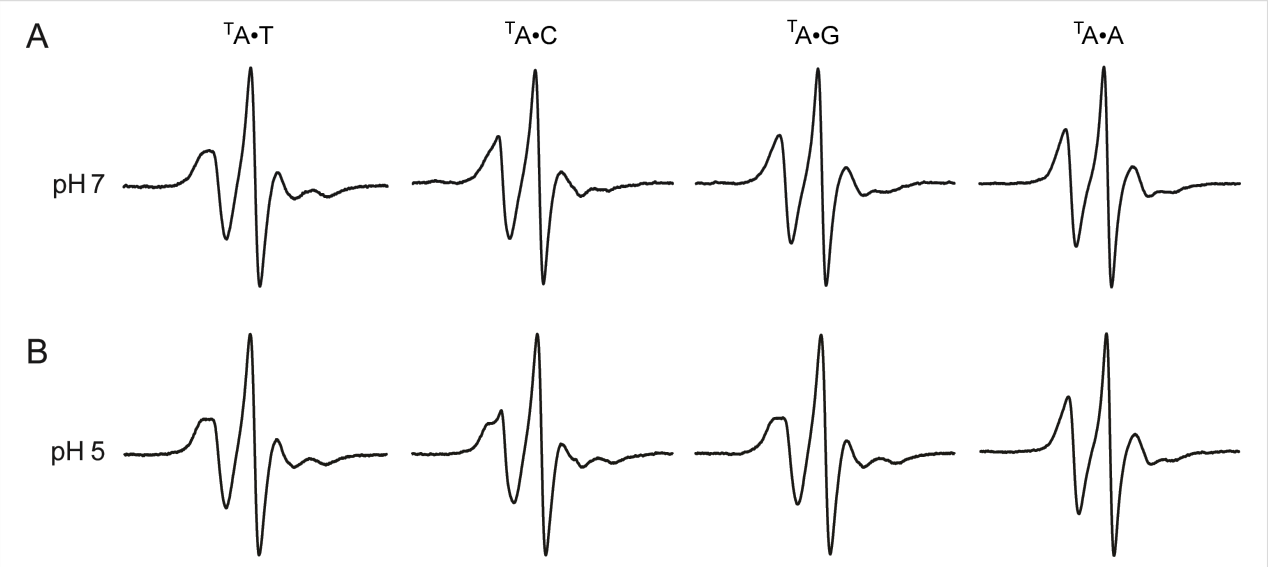



**Figure 2:** EPR spectra of 14-mer DNA duplexes 5′-d(GACCTCG$^{T}$AATCGTG)•5′-d(CACGATYCGAGGTC), (10 mM phosphate, 100 mM NaCl, 0.1 mM $Na_2EDTA$, pH 7.0 (A) and 5.0 (B)) at 10 °C, where $^{T}$A is paired with either Y = T, C, G or A.

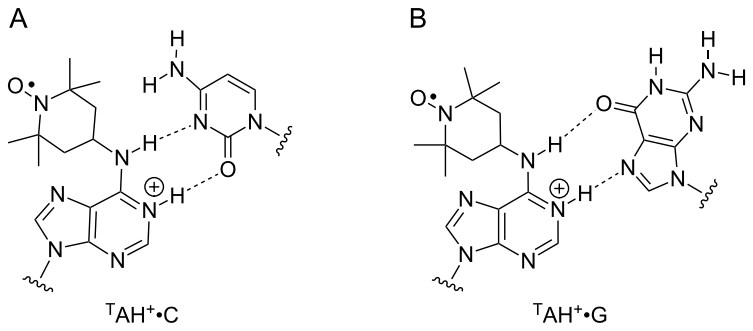



**Figure 3:** Possible base pairing of $^{T}AH^{+}$ with C (A) and G (B) at pH 5.

Hoogsteen hydrogen bonds (Figure 3B) [43,44,78]. Thus, the formation of the $^{T}AH^{+}$•G pair can be readily detected by EPR spectroscopy.

Nucleosides $^{U}A$ and $^{U}C$, the new and readily prepared spin labels that are described here, contain a stable urea linkage that provides additional hydrogen-bonding possibilities. In particular, pairing of $^{U}A$ and $^{U}C$ to C, which has been shown by NMR to form a hydrogen bond between its N3 and the proton on N6 of A [83] as well as N4 of C [69,84], should enable the oxygen in the urea moiety of both $^{U}A$ and $^{U}C$ to pair with a N4 proton of C (Figure 4A and B). Such hydrogen bonding should have an effect on the spin label mobility and be manifested in the line width of the EPR spectra.

Therefore, we recorded the EPR spectra of duplexes 5′-d(GACCTCGXATCGTG)•5′-d(CACGATYCGAGGTC), where X is either $^{U}A$ or $^{U}C$, and Y is either C, A, G or T (Figure 5). Indeed, the lowest mobility, i.e., the widest spectra was observed for both $^{U}A$•C and $^{U}C$•C, providing circumstantial evidence for hydrogen bonding of C to the urea linkage. Interestingly, the same mobility order was observed for both $^{U}A$ and $^{U}C$: X•C < X•A ≤ X•G < X•T. Pairing with T resulted in a high mobility for both $^{U}A$ and $^{U}C$, whereas pairing with either A or G, caused mobility intermediate between that of C and T. As expected, the EPR spectra of the duplexes containing the urea-linked spin labels ($^{U}A$ and $^{U}C$) were narrower than for the N6-TEMPO-dA ($^{T}A$) labeled duplexes. The extended urea linker not only contains more rotatable single bonds but also projects the spin label further out of the major groove, where it is less constrained sterically by the DNA.

## Conclusion

Three spin-labeled deoxynucleosides, $^{T}A$, $^{U}A$ and $^{U}C$, were prepared and incorporated into oligonucleotides by the phosphoramidite method. While $^{U}A$ resulted in a major decrease in the melting temperature of a DNA duplex when incorporated opposite to C, CD measurements revealed that all three spin labels were accommodated in a B-form DNA duplex. The mobility of the spin label $^{T}A$ was highly base-pair sensitive, allowing detection of its respective base-pairing partner in duplex DNA. Moreover, the mobility of $^{T}A$ was significantly reduced when paired with C or G in a DNA duplex at pH 5. This finding is consistent with protonation of $^{T}A$ and subsequent participation of the proton in hydrogen-bonding with C and G. The urea-linked $^{U}A$ and $^{U}C$ spin labels showed a similar

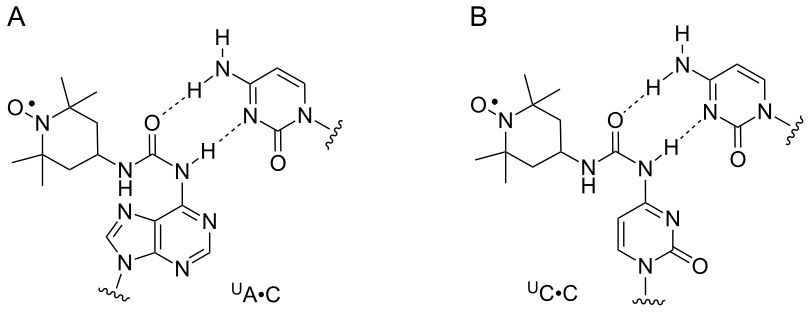



**Figure 4:** Possible base pairing of $^{U}A$ and $^{U}C$ with C.

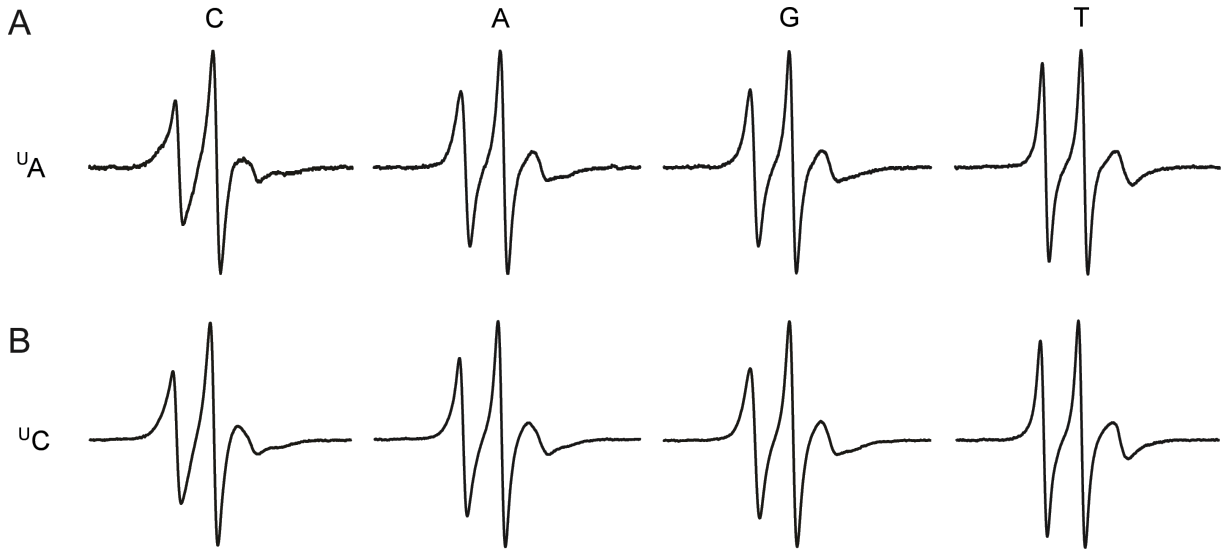



**Figure 5:** EPR spectra of 14-mer DNA duplex 5′-d(GACCTCG$^{U}$AATCGTG)•5′-d(CACGATYCGAGGTC) (A), and 5′-d(GACCTCG$^{U}$CATCGTG)•5′-d(CACGATYCGAGGTC) (B), (10 mM phosphate, 100 mM NaCl, 0.1 mM $Na_2EDTA$, pH 7.0) at 10 °C, where $^{U}$A or $^{U}$C is paired with either Y = T, C, G or A .

mobility order when paired with A, G, C or T in a DNA duplex, with the lowest mobility when paired with C and the highest for T. These results show that the three labels, in particular $^{T}$A, can report minor changes in their microenvironment, such as protonation, when placed in structured regions of nucleic acids.

## Supporting Information

### Supporting Information File 1

Experimental part.
[http://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-11-24-S1.pdf]

## Acknowledgements

We thank the Icelandic Research Fund and the University of Iceland Research Fund for financial support. We thank Dr. Pavol Cekan for technical assistance, Dr. Sigridur Jonsdottir for NMR and MS analyses and the members of the Sigurdsson research group for helpful discussions.

## References

1. Jacobo-Molina, A.; Ding, J.; Nanni, R. G.; Clark, A. D., Jr.; Lu, X.; Tantillo, C.; Williams, R. L.; Kamer, G.; Ferris, A. L.; Clark, P. *Proc. Natl. Acad. Sci. U. S. A.* **1993,** *90,* 6320–6324.
2. Laughlan, G.; Murchie, A.; Norman, D. G.; Moore, M. H.; Moody, P. C.; Lilley, D. M.; Luisi, B. *Science* **1994,** *265,* 520–524. doi:10.1126/science.8036494
3. Holbrook, S. R.; Holbrook, E. L.; Walukiewicz, H. E. *Cell. Mol. Life Sci.* **2001,** *58,* 234–243. doi:10.1007/pl00000851
4. Vassylyev, D. G.; Sekine, S.-i.; Laptenko, O.; Lee, J.; Vassylyeva, M. N.; Borukhov, S.; Yokoyama, S. *Nature* **2002,** *417,* 712–719. doi:10.1038/nature752
5. Edwards, T. E.; Klein, D. J.; Ferré-D'Amaré, A. R. *Curr. Opin. Struct. Biol.* **2007,** *17,* 273–279. doi:10.1016/j.sbi.2007.05.004
6. Mooers, B. H. M. *Methods* **2009,** *47,* 168–176. doi:10.1016/j.ymeth.2008.09.006
7. Clore, G. M.; Gronenborn, A. M. *Curr. Opin. Chem. Biol.* **1998,** *2,* 564–570. doi:10.1016/s1367-5931(98)80084-7
8. Brutscher, B.; Boisbouvier, J.; Pardi, A.; Marion, D.; Simorre, J.-P. *J. Am. Chem. Soc.* **1998,** *120,* 11845–11851. doi:10.1021/ja982853l
9. Tjandra, N. *Structure* **1999,** *7,* R205–R211.
10. Riek, R.; Pervushin, K.; Wüthrich, K. *Trends Biochem. Sci.* **2000,** *25,* 462–468. doi:10.1016/S0968-0004(00)01665-0
11. Latham, M. P.; Brown, D. J.; McCallum, S. A.; Pardi, A. *ChemBioChem* **2005,** *6,* 1492–1505. doi:10.1002/cbic.200500123
12. Campagne, S.; Gervais, V.; Milon, A. *J. R. Soc., Interface* **2011,** *8,* 1065–1078. doi:10.1098/rsif.2010.0543
13. Wu, P. G.; Brand, L. *Anal. Biochem.* **1994,** *218,* 1–13. doi:10.1006/abio.1994.1134
14. Prisner, T.; Rohrer, M.; MacMillan, F. *Annu. Rev. Phys. Chem.* **2001,** *52,* 279–313. doi:10.1146/annurev.physchem.52.1.279
15. Schiemann, O.; Piton, N.; Mu, Y.; Stock, G.; Engels, J. W.; Prisner, T. F. *J. Am. Chem. Soc.* **2004,** *126,* 5722–5729. doi:10.1021/ja0393877
16. Bowman, M. K.; Maryasov, A. G.; Kim, N.; DeRose, V. J. *Appl. Magn. Reson.* **2004,** *26,* 23–39. doi:10.1007/BF03166560
17. Kim, N.-K.; Murali, A.; DeRose, V. J. *Chem. Biol.* **2004,** *11,* 939–948. doi:10.1016/j.chembiol.2004.04.013
18. Piton, N.; Schiemann, O.; Mu, Y.; Stock, G.; Prisner, T.; Engels, J. W. *Nucleosides, Nucleotides Nucleic Acids* **2005,** *24,* 771–775. doi:10.1081/NCN-200060139
19. Sabanayagam, C. R.; Eid, J. S.; Meller, A. *J. Chem. Phys.* **2005,** *122,* 061103. doi:10.1063/1.1854120

20. Schiemann, O.; Piton, N.; Plackmeyer, J.; Bode, B. E.; Prisner, T. F.; Engels, J. W. *Nat. Protoc.* **2007,** *2,* 904–923. doi:10.1038/nprot.2007.97
21. Schiemann, O.; Prisner, T. F. *Q. Rev. Biophys.* **2007,** *40,* 1–53. doi:10.1017/S003358350700460X
22. Sicoli, G.; Mathis, G.; Delalande, O.; Boulard, Y.; Gasparutto, D.; Gambarelli, S. *Angew. Chem.* **2008,** *120,* 747–749. doi:10.1002/ange.200704133
23. Xie, Y.; Dix, A. V.; Tor, Y. *J. Am. Chem. Soc.* **2009,** *131,* 17605–17614. doi:10.1021/ja905767g
24. Sicoli, G.; Wachowius, F.; Bennati, M.; Höbartner, C. *Angew. Chem., Int. Ed.* **2010,** *49,* 6443–6447. doi:10.1002/anie.201000713
25. Kim, N.-K.; Bowman, M. K.; DeRose, V. J. *J. Am. Chem. Soc.* **2010,** *132,* 8882–8884. doi:10.1021/ja101317g
26. Ding, P.; Wunnicke, D.; Steinhoff, H.-J.; Seela, F. *Chem. – Eur. J.* **2010,** *16,* 14385–14396. doi:10.1002/chem.201001572
27. Stoller, S.; Sicoli, G.; Baranova, T. Y.; Bennati, M.; Diederichsen, U. *Angew. Chem., Int. Ed.* **2011,** *50,* 9743–9746. doi:10.1002/anie.201103315
28. Marko, A.; Denysenkov, V.; Margraf, D.; Cekan, P.; Schiemann, O.; Sigurdsson, S. T.; Prisner, T. F. *J. Am. Chem. Soc.* **2011,** *133,* 13375–13379. doi:10.1021/ja201244u
29. Hengesbach, M.; Kim, N.-K.; Feigon, J.; Stone, M. D. *Angew. Chem., Int. Ed.* **2012,** *51,* 5876–5879. doi:10.1002/anie.201200526
30. Pornsuwan, S.; Giller, K.; Riedel, D.; Becker, S.; Griesinger, C.; Bennati, M. *Angew. Chem., Int. Ed.* **2013,** *52,* 10290–10294. doi:10.1002/anie.201304747
31. Polyhach, Y.; Godt, A.; Bauer, C.; Jeschke, G. *J. Magn. Reson.* **2007,** *185,* 118–129. doi:10.1016/j.jmr.2006.11.012
32. Bode, B. E.; Plackmeyer, J.; Prisner, T. F.; Schiemann, O. *J. Phys. Chem. A* **2008,** *112,* 5064–5073. doi:10.1021/jp710504k
33. Schiemann, O.; Cekan, P.; Margraf, D.; Prisner, T. F.; Sigurdsson, S. T. *Angew. Chem., Int. Ed.* **2009,** *48,* 3292–3295. doi:10.1002/anie.200805152
34. Marko, A.; Margraf, D.; Cekan, P.; Sigurdsson, S. T.; Schiemann, O.; Prisner, T. F. *Phys. Rev. E* **2010,** *81,* 021911. doi:10.1103/PhysRevE.81.021911
35. Tkach, I.; Pornsuwan, S.; Höbartner, C.; Wachowius, F.; Sigurdsson, S. T.; Baranova, T. Y.; Diederichsen, U.; Sicoli, G.; Bennati, M. *Phys. Chem. Chem. Phys.* **2013,** *15,* 3433–3437. doi:10.1039/C3CP44415E
36. Svergun, D. I.; Koch, M. H. J. *Rep. Prog. Phys.* **2003,** *66,* 1735. doi:10.1088/0034-4885/66/10/R05
37. Lipfert, J.; Doniach, S. *Annu. Rev. Biophys. Biomol. Struct.* **2007,** *36,* 307–327. doi:10.1146/annurev.biophys.36.040306.132655
38. Bernadó, P.; Mylonas, E.; Petoukhov, M. V.; Blackledge, M.; Svergun, D. I. *J. Am. Chem. Soc.* **2007,** *129,* 5656–5664. doi:10.1021/ja069124n
39. Mertens, H. D. T.; Svergun, D. I. *J. Struct. Biol.* **2010,** *172,* 128–141. doi:10.1016/j.jsb.2010.06.012
40. Zıdek, L.; Štefl, R.; Sklenář, V. *Curr. Opin. Struct. Biol.* **2001,** *11,* 275–281. doi:10.1016/S0959-440X(00)00218-9
41. Fürtig, B.; Richter, C.; Wöhnert, J.; Schwalbe, H. *ChemBioChem* **2003,** *4,* 936–962. doi:10.1002/cbic.200300700
42. Al-Hashimi, H. M. *J. Magn. Reson.* **2013,** *237,* 191–204. doi:10.1016/j.jmr.2013.08.014
43. Gao, X.; Patel, D. J. *J. Am. Chem. Soc.* **1988,** *110,* 5178–5182. doi:10.1021/ja00223a045
44. Lane, A.; Jenkins, T.; Brown, D.; Brown, T. *Biochem. J.* **1991,** *279,* 269–281.
45. Maskos, K.; Gunn, B. M.; LeBlanc, D. A.; Morden, K. M. *Biochemistry* **1993,** *32,* 3583–3595. doi:10.1021/bi00065a009
46. Lane, A.; Ebel, S.; Brown, T. *Eur. J. Biochem.* **1994,** *220,* 717–727. doi:10.1111/j.1432-1033.1994.tb18672.x
47. Boulard, Y.; Cognet, J. A. H.; Gabarro-Arpa, J.; Le Bret, M.; Carbonnaux, C.; Fazakerley, G. V. *J. Mol. Biol.* **1995,** *246,* 194–208. doi:10.1006/jmbi.1994.0076
48. Peyret, N.; Seneviratne, P. A.; Allawi, H. T.; SantaLucia, J., Jr. *Biochemistry* **1999,** *38,* 3468–3477. doi:10.1021/bi9825091
49. Rist, M.; Marino, J. *Nucleic Acids Res.* **2001,** *29,* 2401–2408. doi:10.1093/nar/29.11.2401
50. Liu, C.; Martin, C. T. *J. Mol. Biol.* **2001,** *308,* 465–475. doi:10.1006/jmbi.2001.4601
51. Cekan, P.; Sigurdsson, S. T. *Chem. Commun.* **2008,** 3393–3395. doi:10.1039/b801833b
52. Sinkeldam, R. W.; Greco, N. J.; Tor, Y. *Chem. Rev.* **2010,** *110,* 2579–2619. doi:10.1021/cr900301e
53. Wilhelmsson, L. M. *Q. Rev. Biophys.* **2010,** *43,* 159–183. doi:10.1017/S0033583510000090
54. Gardarsson, H.; Sigurdsson, S. T. *Bioorg. Med. Chem.* **2010,** *18,* 6121–6126. doi:10.1016/j.bmc.2010.06.060
55. Shin, D.; Sinkeldam, R. W.; Tor, Y. *J. Am. Chem. Soc.* **2011,** *133,* 14912–14915. doi:10.1021/ja206095a
56. Gardarsson, H.; Kale, A. S.; Sigurdsson, S. T. *ChemBioChem* **2011,** *12,* 567–575. doi:10.1002/cbic.201000478
57. Dierckx, A.; Miannay, F.-A.; Ben Gaied, N.; Preus, S.; Björck, M.; Brown, T.; Wilhelmsson, L. M. *Chem. – Eur. J.* **2012,** *18,* 5987–5997. doi:10.1002/chem.201103419
58. Saito, Y.; Miyauchi, Y.; Okamoto, A.; Saito, I. *Tetrahedron Lett.* **2004,** *45,* 7827–7831. doi:10.1016/j.tetlet.2004.09.003
59. Xie, Y.; Maxson, T.; Tor, Y. *Org. Biomol. Chem.* **2010,** *8,* 5053–5055. doi:10.1039/C0OB00413H
60. Sankaran, N. B.; Sato, Y.; Sato, F.; Rajendar, B.; Morita, K.; Seino, T.; Nishizawa, S.; Teramae, N. *J. Phys. Chem. B* **2009,** *113,* 1522–1529. doi:10.1021/jp808576t
61. Gislason, K.; Gophane, D. B.; Sigurdsson, S. T. *Org. Biomol. Chem.* **2013,** *11,* 149–157. doi:10.1039/c2ob26536b
62. Haller, A.; Rieder, U.; Aigner, M.; Blanchard, S. C.; Micura, R. *Nat. Chem. Biol.* **2011,** *7,* 393–400. doi:10.1038/nchembio.562
63. Haller, A.; Altman, R. B.; Soulière, M. F.; Blanchard, S. C.; Micura, R. *Proc. Natl. Acad. Sci. U. S. A.* **2013,** *110,* 4188–4193. doi:10.1073/pnas.1218062110
64. Edwards, T. E.; Okonogi, T. M.; Robinson, B. H.; Sigurdsson, S. T. *J. Am. Chem. Soc.* **2001,** *123,* 1527–1528. doi:10.1021/ja005649i
65. Qin, P. Z.; Butcher, S. E.; Feigon, J.; Hubbell, W. L. *Biochemistry* **2001,** *40,* 6929–6936. doi:10.1021/bi010294g
66. Qin, P. Z.; Hideg, K.; Feigon, J.; Hubbell, W. L. *Biochemistry* **2003,** *42,* 6772–6783. doi:10.1021/bi027222p
67. Sowa, G. Z.; Qin, P. Z. *Prog. Nucleic Acid Res. Mol. Biol.* **2008,** *82,* 147–197. doi:10.1016/s0079-6603(08)00005-6
68. Zhang, X.; Cekan, P.; Sigurdsson, S. T.; Qin, P. Z. *Methods Enzymol.* **2009,** *469,* 303–328. doi:10.1016/s0076-6879(09)69015-7
69. Cekan, P.; Sigurdsson, S. T. *J. Am. Chem. Soc.* **2009,** *131,* 18054–18056. doi:10.1021/ja905623k
70. Ricci, A.; Marinello, J.; Bortolus, M.; Sánchez, A.; Grandas, A.; Pedroso, E.; Pommier, Y.; Capranico, G.; Maniero, A. L.; Zagotto, G. *J. Med. Chem.* **2011,** *54,* 1003–1009. doi:10.1021/jm101232t

71. Höbartner, C.; Sicoli, G.; Wachowius, F.; Gophane, D. B.; Sigurdsson, S. T. *J. Org. Chem.* **2012,** *77,* 7749–7754. doi:10.1021/jo301227w
72. Nguyen, P.; Qin, P. Z. *Wiley Interdiscip. Rev.: RNA* **2012,** *3,* 62–72. doi:10.1002/wrna.104
73. Gophane, D. B.; Sigurdsson, S. T. *Chem. Commun.* **2013,** *49,* 999–1001. doi:10.1039/c2cc36389e
74. Giordano, C.; Fratini, F.; Attanasio, D.; Cellai, L. *Synthesis* **2001,** 565–572. doi:10.1055/s-2001-12355
75. Chen, L. S.; Bahr, M. H.; Sheppard, T. L. *Bioorg. Med. Chem. Lett.* **2003,** *13,* 1509–1512. doi:10.1016/S0960-894X(03)00204-X
76. Zakrzewski, J.; Krawczyk, M. *Heteroat. Chem.* **2006,** *17,* 393–401. doi:10.1002/hc.20228
77. Cekan, P.; Smith, A. L.; Barhate, N.; Robinson, B. H.; Sigurdsson, S. T. *Nucleic Acids Res.* **2008,** *36,* 5946–5954. doi:10.1093/nar/gkn562
78. Brown, T.; Leonard, G. A.; Booth, E. D.; Chambers, J. *J. Mol. Biol.* **1989,** *207,* 455–457. doi:10.1016/0022-2836(89)90268-4
79. Brown, T.; Leonard, G. A.; Booth, E. D.; Kneale, G. *J. Mol. Biol.* **1990,** *212,* 437–440. doi:10.1016/0022-2836(90)90320-L
80. Leonard, G. A.; Booth, E. D.; Brown, T. *Nucleic Acids Res.* **1990,** *18,* 5617–5623. doi:10.1093/nar/18.19.5617
81. Brown, T. *Aldrichimica Acta* **1995,** *28,* 15–20.
82. Boulard, Y.; Cognet, J. A. H.; Gabarro-Arpa, J.; Le Bret, M.; Sowers, L.; Fazakerley, G. V. *Nucleic Acids Res.* **1992,** *20,* 1933–1941. doi:10.1093/nar/20.8.1933
83. Allawi, H. T.; SantaLucia, J., Jr.. *Biochemistry* **1998,** *37,* 9435–9444. doi:10.1021/bi9803729
84. Boulard, Y.; Cognet, J. A. H.; Fazakerley, G. V. *J. Mol. Biol.* **1997,** *268,* 331–347. doi:10.1006/jmbi.1997.0975

## License and Terms

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

# C-5'-Triazolyl-2'-oxa-3'-aza-4'a-carbanucleosides: Synthesis and biological evaluation

Roberto Romeo[*1,§], Caterina Carnovale[1], Salvatore V. Giofrè[*1,¶], Maria A. Chiacchio[2], Adriana Garozzo[3], Emanuele Amata[2], Giovanni Romeo[1] and Ugo Chiacchio[2]

Full Research Paper

Open Access

Address:
[1]Dipartimento Scienze del Farmaco e dei Prodotti per la Salute, University of Messina, Via S.S. Annunziata, 98168 Messina, Italy, [2]Dipartimento di Scienze del Farmaco, University of Catania, Via A. Doria 6, 95125-Catania, Italy and [3]Dipartimento di Scienze Bio-Mediche, University of Catania,Via Androne 81, 95124 Catania, Italy

Email:
Roberto Romeo[*] - robromeo@unime.it; Salvatore V. Giofrè[*] - sgiofre@unime.it

* Corresponding author
§ Tel.: +39-090-356230; fax: +39-090-6766474
¶ Tel.: +39-090-6766566; fax: +39-090-6766474

Keywords:
antitumor activity; click chemistry; 1,3-dipolar cycloaddition; nucleic acids; 2'-oxa-3'-aza-4'a-carbanucleoside analogs

*Beilstein J. Org. Chem.* **2015,** *11,* 328–334.
doi:10.3762/bjoc.11.38

Received: 04 December 2014
Accepted: 20 February 2015
Published: 09 March 2015

This article is part of the Thematic Series "Nucleic acid chemistry".

Guest Editor: H.-A. Wagenknecht



## Abstract

A novel series of 2'-oxa-3'-aza-4'a-carbanucleosides, featured with a triazole linker at the 5'-position, has been developed by exploiting a click chemistry reaction of 5'-azido-2'-oxa-3'-aza-4'a-carbanucleosides with substituted alkynes. Biological tests indicate an antitumor activity for the synthesized compounds: most of them inhibit cell proliferation of Vero, BS-C-1, HEp-2, MDCK, and HFF cells with a $CC_{50}$ in the range of 5.0–40 μM. The synthesized compounds do not show any antiviral activity.

## Introduction

Synthetic modified nucleosides are of great interest as potential new lead structures in particular as antiviral or anticancer agents [1-8]. As analogues these compounds can interfere in nucleic acid synthesis or block nucleosides- and/or nucleotide-dependent biological processes by mimicking natural nucleosides and serving as inhibitors or building units [9-12]. Many structural variations of the natural nucleosides have been exploited. In general, the performed modifications included the replacement of the furanose moiety by other carbon or heterocyclic systems [13,14] or even acyclic fragments [15,16], the substitution of pyrimidine or purine natural nucleobases with unnaturally-substituted heteroaromatics or homoaromatic systems, or the modification of the phosphate P(O)–O–C bond with the non–hydrolyzable phosphonate P(O)–C linkage [17,18].

In this context, nucleoside analogues, where different carbon or heterocyclic systems replace the furanose ring, have been reported as anticancer or antiviral agents [19,20]. In particular,

2'-oxa-3'-aza-4'a-carbanucleosides **1**–**4**, characterized by the presence of an isoxazolidine ring, represent a scaffold of modified dideoxynucleosides endowed with interesting physiological features (Figure 1) [21-27].

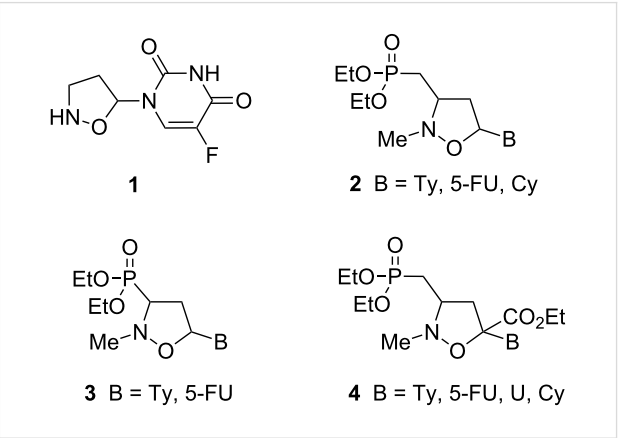



**Figure 1:** 2'-Oxa-3'-aza-modified nucleosides and 2'-oxa-3'-aza-modified nucleotides.

2'-Oxa-3'-aza-4'a-carbanucleosides **1**–**4** can be considered as mimics of natural nucleosides and act as terminators of the viral DNA chain. Their antiviral activity is linked to the competitive reversible inhibition of the reverse transcriptase. Furthermore, as antimetabolites, they can interact with intracellular targets to induce cytotoxicity [28-32].

Several functionalities have been inserted as linkers on the 2'-oxa-3'-aza-4'a-carbanucleoside skeleton in order to confer novel mechanisms of action for nucleoside mimics: in this context, the 1,2,3-triazole unit assumes particular interest according to its easily access and the well-known biological activity of many derivatives. In these last years, in fact, triazoles have gained considerable attention in medicinal chemistry, bioconjugation, drug-delivery, and materials science [33-38]. Moreover, the 1,2,3-triazole motif is exceedingly stable to basic or acidic hydrolysis and interacts strongly with biological targets through hydrogen bonding to nitrogen atoms as well as through dipole–dipole and π-stacking interactions [39].

Recently, a synthetic approach towards 3-hydroxymethyl-5-(1*H*-1,2,3-triazol)-isoxazolidines **5** has been described [40]: the obtained compounds inhibit the growth of anaplastic and follicular human thyroid cancer cell lines, with $IC_{50}$ values in the range of 3.87–8.76 μM. In the same context, novel 1,2,3-triazole-appended 2'-oxa-3'-azanucleoside analogs **6** were developed [41]: Some of these compounds show a good anticancer activity against the anaplastic (8305C) and the follicular (FTC-133) human thyroid cancer cell lines, and especially on the U87MG human primary glioblastoma cell line (Figure 2).

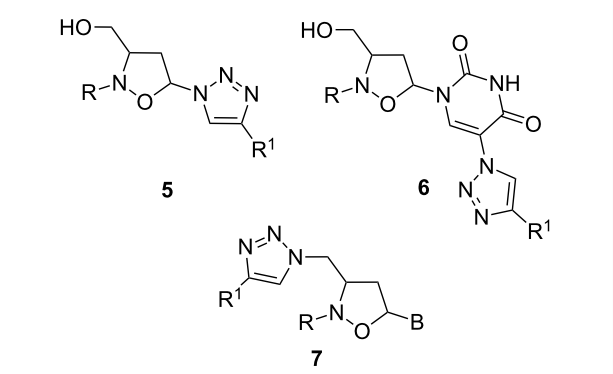



**Figure 2:** Triazolyl-2'-oxa-3'-aza-4'a-carbanucleosides.

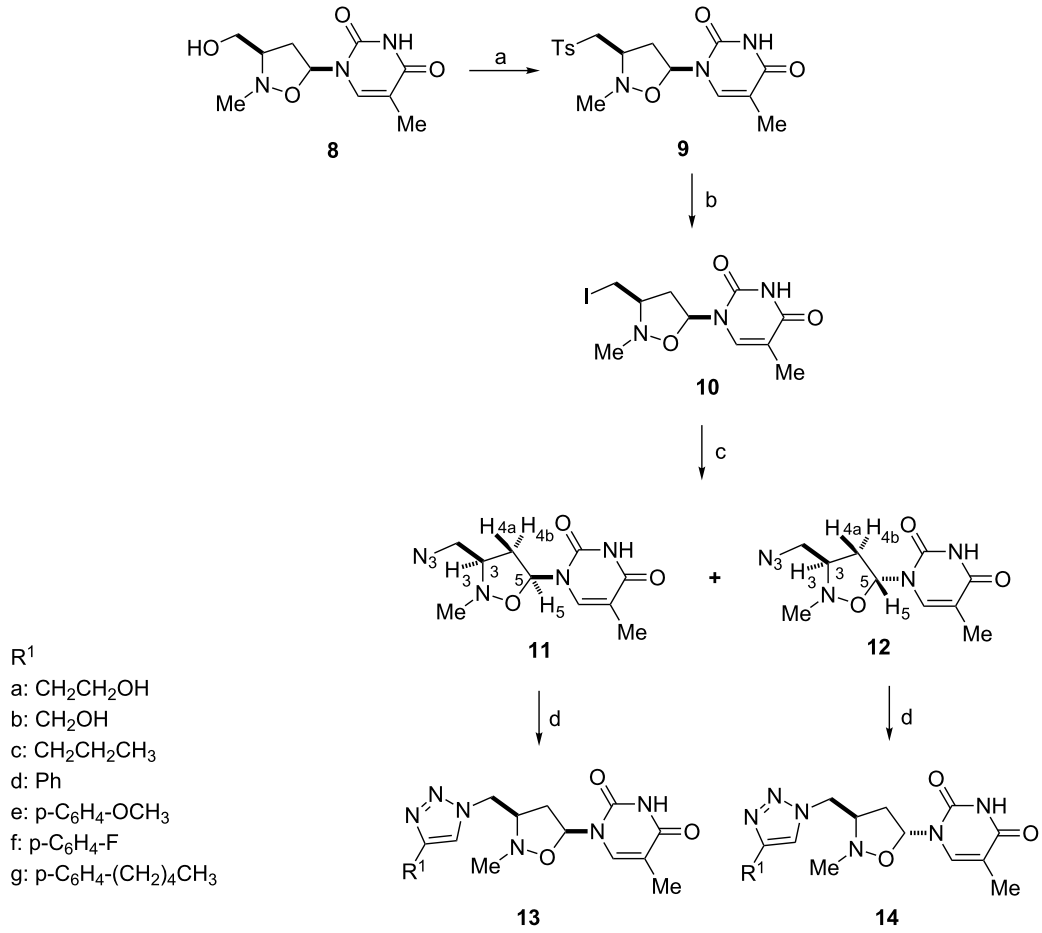
Accordingly, considering that the incorporation of the triazole moiety can lead to interesting biological properties, we report in this paper the preparation of a small library of nucleoside analogues **7** (Figure 2), where the furanose ring is substituted by an isoxazolidine system and a triazole unit replaces the phosphodiester linker at 5' position of the 2'-oxa-3'-aza-4'a-carbanucleoside. However, in order to maintain the six-bond periodicity of the oligonucleotides and thus the flexibility of the oligonucleotide chain the methylene bridge at the pseudo-5'-position was retained. The obtained compounds have shown to be endowed with an interesting antitumor activity: most of them inhibit cell proliferation of Vero, BS-C-1**,** HEp-2, MDCK, and HFF cells by 50% ($CC_{50}$) at concentrations in the range of 5.0–40.0 μM. No antiviral activity against both RNA and DNA viruses was observed.

# Results and Discussion

## Chemistry

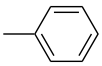
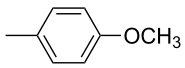
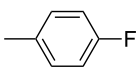
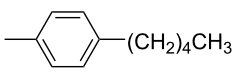
The synthetic route to 5'-triazolyl-2'-oxa-3'-aza-4'a-carbanucleosides **13** and **14** is described in Scheme 1 (and Table 1). (3′*RS*,5′*SR*)-2′-*N*-methyl-3′-hydroxymethyl-1′,2′-isoxazolidin-5′-ylthymine **8**, obtained as the main compound, in a two-step process, by 1,3-dipolar cycloaddition of vinyl acetate to *C*-[(*tert*-butyldiphenylsilyl)oxy]-*N*-methylnitrone, followed by Hilbert–Jones nucleosidation using silylated thymine and TBAF [42-44], was converted into the corresponding iodo-derivative **10** by sequential tosylation and iodination.

The subsequent reaction of **10** with sodium azide, performed at 50 °C in $CH_3CN/H_2O$ (1:10) in the presence of $NH_4Cl$ for 48 h afforded two azides, **11** and **12**, epimeric at C-5, in a relative ratio 2:1 with a global yield of 85%. Two azides were separated by flash chromatography ($CH_2Cl_2$/MeOH 98:2 as eluent). Compound **12** originates from **11**: its formation can be rationalized by considering that the acidic medium of the reaction, linked to the presence of $NH_4Cl$, promotes an equilibrium process which starts from **11** and leads to a mixture of α- and




**Scheme 1:** Synthesis of triazolyl isoxazolidinyl-nucleosides **13** and **14**. Reagents and conditions: a) Tosyl chloride, TEA, $CH_2Cl_2$, rt, 24 h; b) NaI, acetone, reflux, 72 h; c) $NaN_3$, $CH_3CN/H_2O$ (1:10) in the presence of $NH_4Cl$, 50 °C for 48 h; d) substituted alkynes, **17a–g**, $CuSO_4{\cdot}5H_2O$, sodium ascorbate, TEA, rt, 5 h.

**Table 1:** C-5'-Triazolyl-2'-oxo-3'-aza-4'a-carbanucleosides **13a–g** and **14a–g** produced via click chemistry.

| Alkyne | $R^1$ | Product | Yield[a] | Product | Yield[a] |
|---|---|---|---|---|---|
| **17a** | $-CH_2CH_2OH$ | **13a** | 88 | **14a** | 79 |
| **17b** | $-CH_2OH$ | **13b** | 84 | **14b** | 81 |
| **17c** | $-CH_2CH_2CH_3$ | **13c** | 80 | **14c** | 83 |
| **17d** |  | **13d** | 78 | **14d** | 82 |
| **17e** |  | **13e** | 78 | **14e** | 82 |
| **17f** |  | **13f** | 85 | **14f** | 84 |
| **17g** |  | **13g** | 89 | **14g** | 85 |

[a]Isolated yield by flash chromatography.

β-anomers, via the intermediate oxonium ion **15** (path a) or **16** (path b) (Figure 3). As reported in similar systems [45], in the equilibrium mixture the β-anomer **11**, thermodynamically more stable, predominates.

The structure of the obtained compounds was determined by spectroscopic data and MS analysis: the main product of the reaction was the *cis* derivative. NOE measurements confirm the assigned stereochemistry. For compound **11**, the *cis* isomer, irradiation of the H-5 resonance at 5.99 ppm (as doublet of doublets) induced a positive NOE effect on H-3 resonance at 3.85–4.00 ppm (as a multiplet) and on H-4b proton (2.34–2.42 ppm, multiplet) (Scheme 1). Accordingly, in the *trans* derivative **12**, on irradiating H-5 resonance (6.14 ppm; doublet of doublets), a positive NOE effect was detected only for the H-4a proton that resonates at 2.18 ppm as a doublet of doublet of doublets.

5'-Azido-2'-oxa-3'-aza-4'a-carbanucleosides **11** and **12** were independently engaged in a CuI-catalyzed Huisgen [3 + 2] cycloaddition reaction with a series of substituted alkynes **17**, according to the procedure described by Sharpless [46] (Scheme 1 and Table 1). The click chemistry process, carried out with equimolar amounts of the respective dipolarophiles, afforded in all the cases the corresponding C-5'-triazolyl-2'-oxa-3'-aza-4'a-carbanucleosides **13** and **14** in good yields (79–89%). According to other copper-catalyzed azide–alkyne cycloadditions, no traces of 1,5-regioisomers were observed [47,48].

The structure of the obtained compounds was assessed according to $^{1}$H NMR, $^{13}$C NMR and MS data. In particular, the $^{1}$H NMR spectra of 5-methyl-1-[(3*RS,5SR*)-2-methyl-3-(1*H*-1,2,3-triazol-1-ylmethyl)isoxazolidin-5-yl]pyrimidine-2,4(1*H,3H*)diones **13** and 5-methyl-1-[(3*RS,5RS*)-2-methyl-3-(1*H*-1,2,3-triazol-1-ylmethyl)isoxazolidin-5-yl]pyrimidine-2,4(1*H,3H*)diones **14** show, besides the resonances of the protons of the isoxazolidine unit, diagnostic resonances at 7.25–7.75 ppm, as a singlet, for the proton of the triazole system, and at 4.50–5.10 and 4.25–4.75 ppm, respectively in **13** and **14**, as a doublet of doublets, for the methylene group at C-4' position.

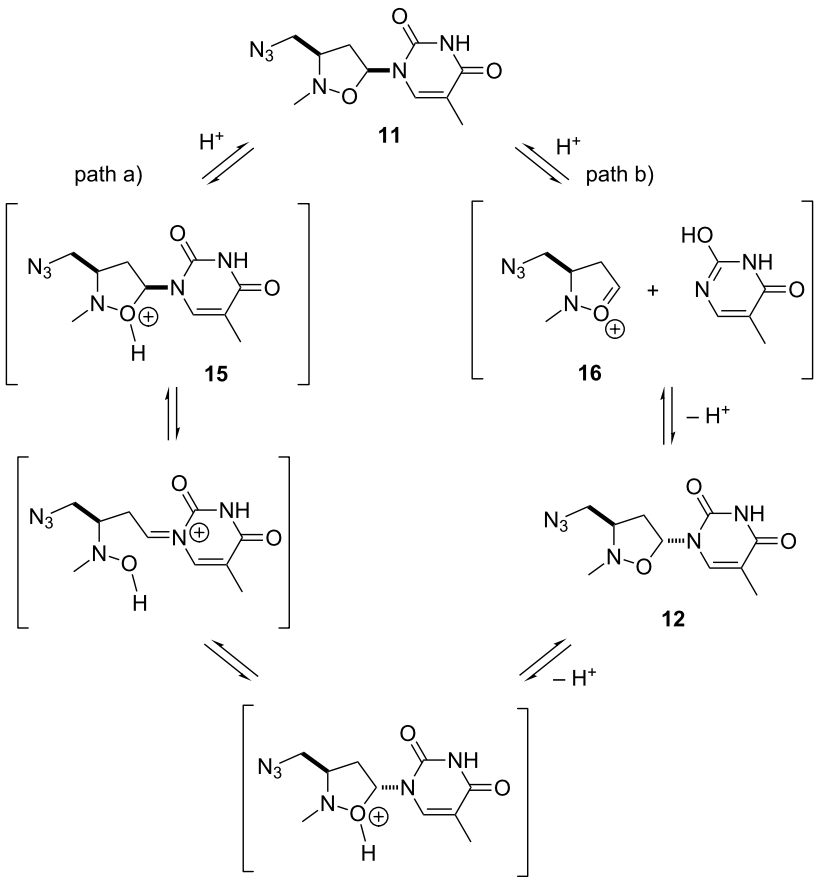



**Figure 3:** α–β Epimerization.

## Biological tests

The antiproliferative effect of the obtained derivatives was tested on a panel of cell lines: african green monkey kidney cells (Vero and BS-C-1), human epidermoid carcinoma larynx cells (HEp-2), Madin–Darby canine kidney (MDCK), and human foreskin fibroblast cells (HFF). In these assays the cells were in the logarithmic phase of growth.

Inhibition of cell proliferation, with a $CC_{50}$ ranging from 5 to 40 μM (Table 2), has been observed for all the new synthesized compounds. In particular, compound **14d** showed a high level of inhibitory activity with $CC_{50}$ values of 5 μM for all the utilized cell lines, while compounds **13c**, **13e**, **13d**, **14c**, **14e**, **14f** and **14g** show the same $CC_{50}$ values only for HFF cells.

**Table 2:** Biological activity of C-5'-triazolyl-2'-oxa-3'-aza-4'a-carbanucleosides **13a–g** and **14a–g**.

| | $CC_{50}$ μM[a] | | | | |
|---|---|---|---|---|---|
| Compound | VERO | HEp2 | MDCK | HFF | BS-C-1 |
| **13a** | 10 | 40 | 10 | 10 | 10 |
| **14a** | 20 | 40 | 20 | 20 | 20 |
| **13b** | 40 | 40 | 30 | 30 | 30 |
| **14b** | 20 | 40 | 20 | 20 | 10 |
| **13c** | 20 | 40 | 20 | 5 | 20 |
| **14c** | 20 | 40 | 20 | 5 | 20 |
| **13d** | 20 | 20 | 20 | 5 | 20 |
| **14d** | 5 | 5 | 5 | 5 | 5 |
| **13e** | 20 | 40 | 20 | 5 | 20 |
| **14e** | 20 | 40 | 20 | 5 | 20 |
| **13f** | 10 | 40 | 10 | 40 | 10 |
| **14f** | 20 | 40 | 20 | 5 | 20 |
| **13g** | 10 | 20 | 10 | 5 | 10 |
| **14g** | 10 | 20 | 10 | 5 | 10 |

[a]$CC_{50}$: Concentration which inhibited cell growth by 50% as compared with control cultures. Values are mean ± 0.5 S.D. (estimated maximal standard deviation) of three separate assays.

Noteworthy, the relative *cis, trans* configuration of **13** and **14** does not seem to affect the biological effect. The cytostatic activity of the compounds was particularly exploited against HFF cell proliferation.

According to our initial hypothesis, the presence of the triazole linker at C-5' position in the 2'-oxa-3'-aza-4'a-carbanucleoside skeleton induces a different biological effect with respect to 2'-oxa-3'-aza-4'a-carbanucleosides devoid of the triazole unit, such as compounds **2** and **8**, which are endowed with antiviral activity, but do not show any cytotoxicity

The ability of compounds **13a–g** and **14a–g** to interfere with the replication of different DNA and RNA viruses was also evaluated, by using the subsequent cell-virus tests: (a) Vero cell for poliovirus 1, human echovirus 9**,** herpes simplex type 1 (HSV-1); (b) HEp-2 cell for Coxsackievirus B1, adenovirus type 2; (c) human foreskin fibroblast cells (HFF) for cytomegalovirus (CMV); (d) BS-C-1 cell (African green monkey kidney) for varicella-zoster virus (VZV); (e) Madin–Darby canine kidney (MDCK) for influenza virus A/Puerto Rico/8/34 H1N1 (PR8). Acyclovir was used as the reference compound. For the synthesized compounds, no inhibitory activity against any virus was detected until 250 μM.

## Biological assays

**Cells.** Biological assays have been performed on African green monkey kidney cells (Vero and BS-C-1), human epithelial type 2 cells (HEp-2), human foreskin fibroblast cells (HFF), Madin-Darby canine kidney (MDCK). All cell lines were obtained from the American Type Culture Collection. The cell cultures were maintained at 37 °C in a humidified atmosphere with 5% $CO_2$ and grown in D-MEM (Dulbecco's modified Eagle's Minimum Essential medium) supplemented with 10% FCS (fetal calf serum, 2 mM/L glutamine, 0.1% sodium bicarbonate, 200 μg/mL of streptomycin and 200 units/mL of penicillin G. The maintenance medium (DMEM with 2% heat inactivated FCS) was used to culture the viruses.

**Cell viability.** The cytotoxicity of the tested compounds was evaluated by measuring the effect created on cell morphology and/or cell growth (cytostatic activity). Cell monolayers were prepared in 24-well tissue culture plates and exposed to various concentrations of the compounds. Cytotoxicity was recorded as morphological variations (such as rounding up, shrinking and detachment) at 24, 48, 72 and 96 h, using light microscopy. Cytotoxicity was expressed as the minimum cytotoxic concentration (MCC) that caused a microscopically detectable variation of cell morphology. The extent of cytostatic activity was measured as inhibition of cell growth using the MTT method, as previously described [49,50]. The 50% cytotoxic dose ($CC_{50}$) is the compound concentration required to reduce cell proliferation by 50% relative to the absorbance of the untreated control. $CC_{50}$ values were estimated from graphic plots of the percentage of control as a function of the concentration of the test compounds.

**Test compounds**. Compounds **13** and **14** were dissolved in DMSO and diluted in maintenance medium to achieve the final required concentration. The final dilution of test compounds contained a maximum concentration of 0.01% DMSO, which had no effect on the viability of the cell lines. Stock solutions of acycloguanosine (Sigma, USA) were prepared in distilled water, filtered through 0. 2 μm filter and stored at 4 °C until use.

**Viruses**. In the antiviral assays the following viruses were used: Poliovirus 1 (Sabin strain: VR-1562), Human echovirus 9 (VR-1050), Herpes simplex type 1 (HSV-1: VR-260), Coxsackievirus B1 (VR-28), adenovirus type 2 (VR-1080), Cytomegalovirus (CMV: VR-538), varicella-zoster virus (VZV: VR-1367), influenza virus A/Puerto Rico/8/34 H1N1 (PR8). Viruses were obtained from the American Type Culture Collection. The tests on the antiviral activity were carried out by the 50% plaque reduction assay or by 50% virus-induced cytopathogenicity, as previously described [51]. The concentration of the compound that inhibit the formation of viral plaques or virus-induced cytopathogenicity by 50% is expressed as EC50.

## Conclusion

In summary, starting from 5'-azido-2'-oxa-3'-azanucleosides, a new series of C-5'-triazolyl-2'-oxo-3'-aza-4'a-carbanucleosides has been synthesized by using a CuI-catalyzed Huisgen [3 + 2] cycloaddition with substituted alkynes. The biological assays indicate that these compounds inhibit the cell proliferation of Vero, BS-C-1, HEp-2, MDCK, and HFF cells by 50% ($CC_{50}$) at concentrations in the range of 5.0–40.0 μM. No antiviral activity at subtoxic concentrations was observed.

## Supporting Information

### Supporting Information File 1

Preparation and analytical data of compounds **9**–**14**. Copies of $^{1}H$ and $^{13}C$ NMR spectra of all new compounds.
[http://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-11-38-S1.pdf]

## Acknowledgements

We gratefully acknowledge the Italian Ministry of Education, Universities, and Research (MIUR), the Universities of Messina and Catania, and the Interuniversity Consortium for Innovative Methodologies and Processes for Synthesis (CINMPIS) for partial financial support.

## References

1. Mehellou, Y.; De Clercq, E. *J. Med. Chem.* **2010,** *53,* 521–538. doi:10.1021/jm900492g
2. Štambaský, J.; Hocek, M.; Kočovský, P. *Chem. Rev.* **2009,** *109,* 6729–6764. doi:10.1021/cr9002165
3. Galmarini, C. M.; Popowycz, F.; Joseph, B. *Curr. Med. Chem.* **2008,** *15,* 1072–1082. doi:10.2174/092986708784221449
4. Balestrieri, E.; Matteucci, C.; Ascolani, A.; Piperno, A.; Romeo, R.; Romeo, G.; Chiacchio, U.; Mastino, A.; Macchi, B. *Antimicrob. Agents Chemother.* **2008,** *52,* 54–64. doi:10.1128/AAC.00470-07
5. De Clercq, E. *Nat. Rev. Microbiol.* **2004,** *2,* 704–720. doi:10.1038/nrmicro975
6. Galmarini, C. M.; Mackey, J. R.; Dumontet, C. *Lancet Oncol.* **2002,** *3,* 415–424. doi:10.1016/S1470-2045(02)00788-X
7. Pathak, T. *Chem. Rev.* **2002,** *102,* 1623–1668. doi:10.1021/cr0104532
8. Ferrero, M.; Gotor, V. *Chem. Rev.* **2000,** *100,* 4319–4348. doi:10.1021/cr000446y
9. Saag, M. S. *Top. Antivir. Med.* **2012,** *20,* 162–167.
10. Bonate, P. L.; Arthaud, L.; Cantrell, W. R.; Stephenson, K.; Secrist, J. A.; Weitman, S. *Nat. Rev. Drug Discovery* **2006,** *5,* 855–863. doi:10.1038/nrd2055
11. Hatse, S.; De Clercq, E.; Balzarini, J. *Biochem. Pharmacol.* **1999,** *58,* 539–555. doi:10.1016/S0006-2952(99)00035-0
12. Lauria, F.; Benfenati, D.; Raspadori, D.; Rondelli, D.; Zinzani, P. L.; Tura, S. *Leuk. Lymphoma* **1993,** *11,* 399–404. doi:10.3109/10428199309067932
13. Romeo, G.; Chiacchio, U.; Corsaro, A.; Merino, P. *Chem. Rev.* **2010,** *110,* 3337–3370. doi:10.1021/cr800464r
14. Merino, P. *Curr. Med. Chem.* **2006,** *13,* 539–545. doi:10.2174/092986706776055779
15. Hirota, K.; Monguchi, Y.; Sajiki, H. Synthesis of Purine Acyclonucleosides via Ribofuranose-Ring Cleavage of Purine Nucleosides by Diisobutylaluminum Hydride. In *Recent Advances in Nucleosides: Chemistry and Chemotherapy;* Chu, C. K., Ed.; Elsevier: Amsterdam, 2002; pp 57–70. doi:10.1016/B978-044450951-2/50003-5
16. Littler, E.; Zhou, E. E. In *Comprehensive Medicinal Chemistry II;* Taylor, J. B.; Triggle, D. J., Eds.; Elsevier, 2006; Vol. 7, pp 295–327.
17. Sharma, P. L.; Nurpeisov, V.; Hernandez-Santiago, B.; Beltran, T.; Schinazi, R. F. *Curr. Top. Med. Chem.* **2004,** *4,* 895–919. doi:10.2174/1568026043388484
18. Bortolini, O.; Mulani, I.; De Nino, A.; Maiuolo, L.; Nardi, M.; Russo, B.; Avnet, S. *Tetrahedron* **2011,** *67,* 5635–5641. doi:10.1016/j.tet.2011.05.098
19. Piperno, A.; Chiacchio, M. A.; Iannazzo, D.; Romeo, R. *Curr. Med. Chem.* **2006,** *13,* 3675–3695. doi:10.2174/092986706779026110
20. Maiuolo, L.; Bortolini, O.; De Nino, A.; Russo, B.; Gavioli, R.; Sforza, F. *Aust. J. Chem.* **2014,** *67,* 670–674. doi:10.1071/CH13511
21. Merino, P.; Tejero, T.; Unzurrunzaga, F. J.; Franco, S.; Chiacchio, U.; Saita, M. G.; Iannazzo, D.; Piperno, A.; Romeo, G. *Tetrahedron: Asymmetry* **2005,** *16,* 3865–3876. doi:10.1016/j.tetasy.2005.11.004
22. Chiacchio, U.; Genovese, F.; Iannazzo, D.; Librando, V.; Merino, P.; Rescifina, A.; Romeo, R.; Procopio, A.; Romeo, G. *Tetrahedron* **2004,** *60,* 441–448. doi:10.1016/j.tet.2003.11.007
23. Chiacchio, U.; Corsaro, A.; Pistarà, V.; Rescifina, A.; Iannazzo, D.; Piperno, A.; Romeo, G.; Romeo, R.; Grassi, G. *Eur. J. Org. Chem.* **2002,** 1206–1212. doi:10.1002/1099-0690(200204)2002:7<1206::AID-EJOC1206>3.0.CO;2-0
24. Chiacchio, U.; Corsaro, A.; Iannazzo, D.; Piperno, A.; Procopio, A.; Rescifina, A.; Romeo, G.; Romeo, R. *Eur. J. Org. Chem.* **2001,** 1893–1898. doi:10.1002/1099-0690(200105)2001:10<1893::AID-EJOC1893>3.0.CO;2-K
25. Romeo, R.; Carnovale, C.; Giofrè, S. V.; Monciino, G.; Chiacchio, M. A.; Sanfilippo, C.; Macchi, B. *Molecules* **2014,** *19,* 14406–14416. doi:10.3390/molecules190914406
26. Romeo, R.; Navarra, M.; Giofrè, S. V.; Carnovale, C.; Cirmi, S.; Lanza, G.; Chiacchio, M. A. *Bioorg. Med. Chem.* **2014,** *22,* 3379–3385. doi:10.1016/j.bmc.2014.04.047

27. Romeo, R.; Giofrè, S. V.; Garozzo, A.; Bisignano, B.; Corsaro, A.; Chiacchio, M. A. *Bioorg. Med. Chem.* **2013,** *21,* 5688–5693. doi:10.1016/j.bmc.2013.07.031
28. Romeo, R.; Carnovale, C.; Giofrè, S. V.; Romeo, G.; Macchi, B.; Frezza, C.; Marino-Merlo, F.; Pistarà, V.; Chiacchio, U. *Bioorg. Med. Chem.* **2012,** *20,* 3652–3657. doi:10.1016/j.bmc.2012.03.047
29. Piperno, A.; Giofrè, S. V.; Iannazzo, D.; Romeo, R.; Romeo, G.; Chiacchio, U.; Rescifina, A.; Piotrowska, D. G. *J. Org. Chem.* **2010,** *75,* 2798–2805. doi:10.1021/jo902485m
30. Chiacchio, U.; Borrello, L.; Iannazzo, D.; Merino, P.; Piperno, A.; Rescifina, A.; Richichi, B.; Romeo, G. *Tetrahedron: Asymmetry* **2003,** *14,* 2419–2425. doi:10.1016/S0957-4166(03)00525-1
31. Chiacchio, U.; Corsaro, A.; Iannazzo, D.; Piperno, A.; Rescifina, A.; Romeo, R.; Romeo, G. *Tetrahedron Lett.* **2001,** *42,* 1777–1780. doi:10.1016/S0040-4039(00)02325-X
32. Romeo, R.; Giofrè, S. V.; Iaria, D.; Sciortino, M. T.; Ronsisvalle, S.; Chiacchio, M. A.; Scala, A. *Eur. J. Org. Chem.* **2011,** 5690–5695. doi:10.1002/ejoc.201100767
33. Singhal, N.; Sharma, P. K.; Kumar, N.; Duhe, R. *Chem. Biol. Interface* **2011,** *1,* 338–348.
34. Singh, R. J.; Singh, D. K. *E-J. Chem.* **2009,** *6,* 796–800. doi:10.1155/2009/419214
35. Moorhouse, A. D.; Moses, J. E. *ChemMedChem* **2008,** *3,* 715–723. doi:10.1002/cmdc.200700334
36. Lutz, J.-F. *Angew. Chem., Int. Ed.* **2007,** *46,* 1018–1025. doi:10.1002/anie.200604050
37. Angell, Y. L.; Burgess, K. *Chem. Soc. Rev.* **2007,** *36,* 1674–1689. doi:10.1039/b701444a
38. Tome, A. C. In *Product class 13: 1,2,3-triazoles;* Stor, R.; Gilchrist, T., Eds.; *Science of Synthesis*, Vol. 13; Thieme: New York, 2004; pp 415–601.
39. Rowan, A. S.; Nicely, N. I.; Cochrane, N.; Wlassoff, W. A.; Claiborne, A.; Hamilton, C. J. *Org. Biomol. Chem.* **2009,** *7,* 4029–4036. doi:10.1039/b913066g
40. Romeo, R.; Giofrè, S. V.; Carnovale, C.; Campisi, A.; Parenti, R.; Bandini, L.; Chiacchio, M. A. *Bioorg. Med. Chem.* **2013,** *21,* 7929–7937. doi:10.1016/j.bmc.2013.10.001
41. Romeo, R.; Giofrè, S. V.; Carnovale, C.; Chiacchio, M. A.; Campisi, A.; Mancuso, R.; Cirmi, S.; Navarra, A. *Eur. J. Org. Chem.* **2014,** 5442–5447. doi:10.1002/ejoc.201402106
42. Carnovale, C.; Iannazzo, D.; Nicolosi, G.; Piperno, A.; Sanfilippo, C. *Tetrahedron: Asymmetry* **2009,** *20,* 425–429. doi:10.1016/j.tetasy.2009.02.026
43. Chiacchio, U.; Rescifina, A.; Iannazzo, D.; Piperno, A.; Romeo, R.; Borrello, L.; Sciortino, M. T.; Balestrieri, E.; Macchi, B.; Mastino, A.; Romeo, G. *J. Med. Chem.* **2007,** *50,* 3747–3750. doi:10.1021/jm070285r
44. Iannazzo, D.; Piperno, A.; Pistarà, V.; Rescifina, A.; Romeo, R. *Tetrahedron: Asymmetry* **2002,** *58,* 581–587. doi:10.1016/S0040-4020(01)01161-9
45. Ward, D. I.; Jeffs, S. M.; Coe, P. L.; Walker, R. T. *Tetrahedron Lett.* **1993,** *34,* 6779–6782. doi:10.1016/S0040-4039(00)61700-8
46. Kolb, H. C.; Finn, M. C.; Sharpless, K. B. *Angew. Chem., Int. Ed.* **2001,** *40,* 2004–2021. doi:10.1002/1521-3773(20010601)40:11<2004::AID-ANIE2004>3.0.CO;2-5
47. Spiteri, C.; Moses, J. E. *Angew. Chem., Int. Ed.* **2010,** *49,* 31–33. doi:10.1002/anie.200905322
48. Tornøe, C. W.; Christensen, C.; Meldal, M. *J. Org. Chem.* **2002,** *67,* 3057–3064. doi:10.1021/jo011148j
49. Denizot, F.; Lang, R. *J. Immunol. Methods* **1986,** *89,* 271–277.
50. Cutrì, C. C. C.; Garozzo, A.; Siracusa, M. A.; Sarvà, M. C.; Tempera, G.; Geremia, E.; Pinizzotto, M. R.; Guerrera, F. *Bioorg. Med. Chem.* **1998,** *6,* 2271–2280. doi:10.1016/S0968-0896(98)80007-2
51. Garozzo, A.; Cutrì, C. C. C.; Castro, A.; Tempera, G.; Guerrera, F.; Sarvà, M. C.; Geremia, E. *Antiviral Res.* **2000,** *45,* 199–210. doi:10.1016/S0166-3542(00)00072-3

## License and Terms

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

# A procedure for the preparation and isolation of nucleoside-5’-diphosphates

Heidi J. Korhonen[1,2], Hannah L. Bolt[1] and David R. W. Hodgson[*1]

Letter

Open Access

Address:
[1]Department of Chemistry, Science Laboratories, Durham University, South Road, Durham, DH1 3LE, United Kingdom, and [2]Department of Chemistry, University of Turku, Vatselankatu 2, 20014 Turku, Finland

Email:
David R. W. Hodgson[*] - d.r.w.hodgson@durham.ac.uk

* Corresponding author

Keywords:
NDP synthesis; nucleic acids; nucleoside-5’-diphosphate; phosphorylation

*Beilstein J. Org. Chem.* **2015,** *11,* 469–472.
doi:10.3762/bjoc.11.52

Received: 28 November 2014
Accepted: 25 March 2015
Published: 10 April 2015

This article is part of the Thematic Series "Nucleic acid chemistry".

Guest Editor: H.-A. Wagenknecht



## Abstract

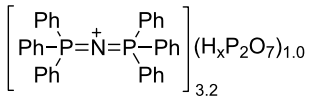
Tris[bis(triphenylphosphoranylidene)ammonium] pyrophosphate (PPN pyrophosphate) was used in the $S_N2$ displacements of the tosylate ion from 5’-tosylnucleosides to afford nucleoside-5’-diphosphates. Selective precipitation permitted the direct isolation of nucleoside-5’-diphosphates from crude reaction mixtures.

## Introduction

Nucleoside-5'-phosphates are key to many mechanistic studies and chemical biology applications [1-3]. Synthetic approaches towards nucleoside-5'-phosphates are well-established [4], however, the methods tend to be cumbersome. Newer approaches have started to become available over the last decade, where a key driver in these approaches is limiting the effects of moisture on the reaction outcome [5-11].

Nucleoside-5'-triphosphates (NTPs) are frequently accessed via the Ludwig modification [12] of the Yoshikawa procedure [13,14], where an unprotected ribonucleoside is first phosphophorylated, then the resulting phosphodichloridate is captured using pyrophosphate ions. Unfortunately, similar attempts towards nucleoside-5'-diphosphates (NDPs), with phosphate ion as nucleophile, tend to produce greater quantities of NTP than NDP [15,16]. Nuceloside-5’-diphosphates can, however, be accessed via $S_N2$ displacement at 5’-activated nucleosides using pyrophosphate as the nucleophile. This approach, developed by Poulter and co-workers [17], forms a cornerstone in the armoury towards the synthesis of nucleoside-5’-polyphosphates [4]. The key element to the success of this process is the availability of dry, organic-solvent-soluble pyrophosphate. Tris(tetrabutylammonium) pyrophosphate is generally employed in this role [18], however, isolating this material in a dry condition, and maintaining its dryness are significant problems [4,19]. In addition, excess pyrophosphate is usually employed in order to drive the kinetics of the displacement process, however, this excess material tends to co-elute with nucleoside-5’-diphosphate products during anion exchange chromatographic purifications.

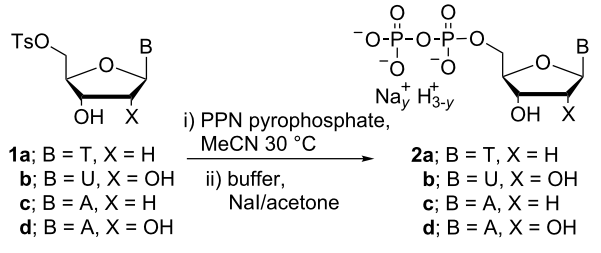
We have recently reported the successful synthesis of PPN pyrophosphate (Scheme 1) and its application in the syntheses of nucleoside-5'-triphosphates (NTPs) [11]. PPN pyrophosphate is straightforward to prepare, and, unlike alkylammonium salts, shows limited levels of moisture uptake. In this letter we explore the use of PPN pyrophosphate as a replacement for tris(tetrabutylammonium) pyrophosphate in the preparation of NDPs via the Poulter approach [17]. Thereafter, we consider the removal of the PPN cation from reaction mixtures alongside removal of excess pyrophosphate to facilitate the isolation of NDP products.



**Scheme 1:** PPN pyrophosphate.

## Results and Discussion

Preliminary experiments focused on 5'-tosylthymidine **1a** as a test substrate (Scheme 2), using conditions similar to those employed by Poulter and co-workers [17]. An excess of pyrophosphate was used to drive the kinetics of the displacement process in concentrated solutions in acetonitrile as the reaction solvent. The reaction progress was monitored via $^{31}$P NMR spectroscopy, and ~62% conversion to TDP **2a** was attained after 30 h of reaction based on pyrophosphate consumption. We found that slightly elevated temperature (30 °C) accelerated the displacement, with no adverse effect on conversion levels to NDP. We believe that our use of PPN pyrophosphate supports the use of elevated temperatures because it limits (or even eliminates) the possibility of water ingress into reaction mixtures and the resulting hydrolysis processes that lead to the formation of monophosphate byproducts. Additional 5'-tosylnucleosides **1b–d** were then exposed to the same reaction conditions and reaction progress was monitored by $^{31}$P NMR spectroscopy (Table 1).



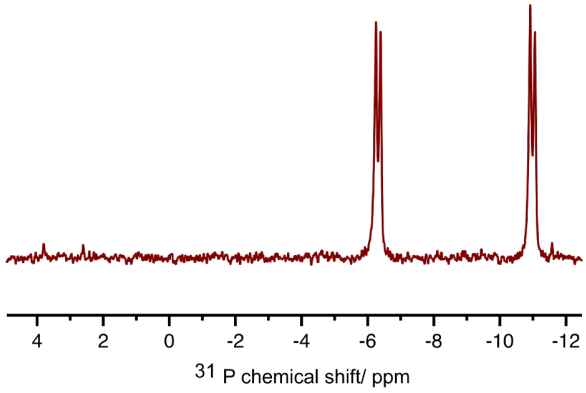

**Scheme 2:** Preparation of NDPs.

**Table 1:** Reaction times, yields and purities of 5'-diphosphates.

| Nucleoside | Time [h] | Yield [%] | Purity [%][a] | Product |
|---|---|---|---|---|
| **1a** | 30 | 45 | 98 | **2a** |
| **1b** | 91 | 40 | 98 | **2b** |
| **1c** | 91 | 44 | 92 | **2c** |
| **1d** | 91 | 20 | 87 | **2d** |

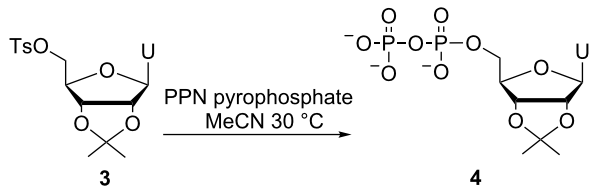
[a]Determined by $^{31}$P NMR spectroscopy.

In order to remove inorganic pyrophosphate from the crude reaction mixtures we developed selective precipitation procedures. To ensure that these procedures were readily repeatable, we used triethylammonium bicarbonate (TEAB) buffer to control the pH value, and we used ice baths to ensure constant conditions of temperature. Owing to its charge-dense nature, we hoped to be able to precipitate pyrophosphate as its sodium salt through the use of sodium iodide solution in acetone, where acetone served as the precipitating solvent. A systematic study of conditions was performed in order to allow selective precipitation of pyrophosphate ions. Variables included; TEAB buffer concentration and volume, the concentration of the sodium iodide solution in acetone and the volume of this solution that was added. The effectiveness of conditions was assessed using $^{31}$P NMR spectroscopic analyses of both precipitated and supernatant materials, and, conveniently, we found conditions that removed the pyrophosphate effectively. Thereafter, NDPs were isolated through the addition of further volumes of sodium iodide solution in acetone. Gratifyingly, the sequence of selective precipitations allowed for the removal of both pyrophosphate and PPN ions, and isolated materials showed high levels of purity, and reasonable levels of mass recovery without chromatographic purifications (Scheme 3, Table 1).




**Scheme 3:** $^{31}$P NMR spectrum of TDP **2a** after precipitation from reaction mixture.

Protected substrate **3** showed conversion to diphosphate **4** after 91 h of reaction, however, we were unable to precipitate this material after removal of excess pyrophosphate ions (Scheme 4). We believe the isopropylidene protecting group decreases the polarity of diphosphate **4** to such an extent that it remains soluble even after the addition of excess sodium iodide and acetone.




**Scheme 4:** Attempted use of an isopropylidene-protected 5'-tosylnucleoside.

## Conclusion

PPN pyrophosphate is a convenient, effective reagent for the preparation of NDPs from 5'-tosylnucleosides. Post-reaction removal of PPN cations and elimination of excess pyrophosphate followed by isolation of polar NDPs as sodium salts can be readily achieved through selective precipitation.

## Experimental

### General procedure for the synthesis of NDPs 2a–d

PPN pyrophosphate was prepared as described previously [11] and was dried in a vacuum desiccator over $P_2O_5$ (we have also found freeze-drying to be effective in other experiments), then dissolved in dry acetonitrile, and stored over activated 4 Å molecular sieves. Tosylated nucleoside **1a–d** (0.18 mmol) and PPN pyrophosphate (0.36 mmol) in dry acetonitrile (0.8 mL) were stirred at 30 °C under an inert atmosphere. The reaction progress was monitored by $^{31}$P NMR spectroscopy until no further reaction was observed. Sodium iodide (0.149 g) was dissolved into 9:1 acetone/water (10 mL) and the solution was added dropwise with stirring to the crude reaction mixture which had been transferred to a centrifuge tube. The solution was mixed for 30 minutes at 0 °C and then centrifuged at 4,000 rpm for 10 minutes. The supernatant was removed by a pipette and the solid residue was dissolved in TEAB buffer (3 mL, 0.1 M, pH 8.0) [4]. Sodium iodide in acetone (5 mL, 0.1 M) was added slowly to the solution and the mixture was stirred for 30 minutes at 0 °C. Solids were collected by centrifugation at 4,000 rpm for 10 minutes, and the liquid layer was analysed by $^{31}$P NMR spectroscopy to ensure that selective removal of inorganic pyrophosphate had been achieved. In cases where pyrophosphate ions were still present, an additional small volume (<1 mL) of sodium iodide in acetone was added and the mixture was stirred and solids were collected by centrifugation as described above. When all pyrophosphate had been removed from the supernatant layer (determined by $^{31}$P NMR spectroscopy), sodium iodide in acetone (5 mL, 0.1 M) was added in order to precipitate the NDP product. The mixture was stirred for 30 minutes at 0 °C and then centrifuged at 4,000 rpm for 10 minutes. The precipitated NDP was washed with acetone (2 × 2 mL) in order to remove remaining sodium iodide and dried in a vacuum desiccator to yield a white solid.

## Supporting Information

### Supporting Information File 1

Experimental procedures for the preparation of 5'-tosylates and their $^1$H and $^{13}$C NMR spectra, and $^{31}$P and $^1$H NMR spectra of NDPs.
[http://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-11-52-S1.pdf]

## Acknowledgements

This work was supported by the Jenny and Antti Wihuri Foundation (H.J.K), Durham University (H.L.B.) and the Royal Society (chromatography system).

## References

1. Hodgson, D. R. W.; Schröder, M. *Chem. Soc. Rev.* **2011,** *40,* 1211–1223. doi:10.1039/c0cs00020e
2. Korhonen, H. J.; Conway, L. P.; Hodgson, D. R. W. *Curr. Opin. Chem. Biol.* **2014,** *21,* 63–72. doi:10.1016/j.cbpa.2014.05.001
3. Lee, S. E.; Elphick, L. M.; Anderson, A. A.; Bonnac, L.; Child, E. S.; Mann, D. J.; Gouverneur, V. *Bioorg. Med. Chem. Lett.* **2009,** *19,* 3804–3807. doi:10.1016/J.Bmcl.2009.04.028
4. Williams, D. M.; Harris, V. H. Phosphorus methods in nucleotide chemistry. In *Organophosphorus Reagents: a Practical Approach in Chemistry;* Murphy, P. J., Ed.; Oxford University Press: Oxford, 2004; pp 237–272.
5. Caton-Williams, J.; Lin, L.; Smith, M.; Huang, Z. *Chem. Commun.* **2011,** *47,* 8142–8144. doi:10.1039/c1cc12201k
6. Cremosnik, G. S.; Hofer, A.; Jessen, H. J. *Angew. Chem., Int. Ed.* **2014,** *53,* 286–289. doi:10.1002/anie.201306265
7. Jessen, H. J.; Ahmed, N.; Hofer, A. *Org. Biomol. Chem.* **2014,** *12,* 3526–3530. doi:10.1039/C4OB00478G
8. Mohamady, S.; Desoky, A.; Taylor, S. D. *Org. Lett.* **2012,** *14,* 402–405. doi:10.1021/ol203178k
9. Ravalico, F.; Messina, I.; Berberian, M. V.; James, S. L.; Migaud, M. E.; Vyle, J. S. *Org. Biomol. Chem.* **2011,** *9,* 6496–6497. doi:10.1039/c1ob06041d
10. Strenkowska, M.; Wanat, P.; Ziemniak, M.; Jemielity, J.; Kowalska, J. *Org. Lett.* **2012,** *14,* 4782–4785. doi:10.1021/ol302071f
11. Korhonen, H. J.; Bolt, H. L.; Vicente-Gines, L.; Perks, D. C.; Hodgson, D. R. W. *Phosphorus, Sulfur Silicon Relat. Elem.*, in press. doi:10.1080/10426507.2014.984032
12. Ludwig, J. *Acta Biochim. Biophys. Hung.* **1981,** *16,* 131–133.

13. Yoshikawa, M.; Kato, T.; Takenishi, T. *Tetrahedron Lett.* **1967,** *8,* 5065–5068. doi:10.1016/S0040-4039(01)89915-9
14. Yoshikawa, M.; Kato, T.; Takenishi, T. *Bull. Chem. Soc. Jpn.* **1969,** *42,* 3505–3508.
15. Mishra, N. C.; Broom, A. D. *J. Chem. Soc., Chem. Commun.* **1991,** 1276–1277. doi:10.1039/C39910001276
16. Hoffmann, C.; Genieser, H.-G.; Veron, M.; Jastorff, B. *Bioorg. Med. Chem. Lett.* **1996,** *6,* 2571–2574. doi:10.1016/0960-894x(96)00461-1
17. Davisson, V. J.; Davis, D. R.; Dixit, V. M.; Poulter, C. D. *J. Org. Chem.* **1987,** *52,* 1794–1801. doi:10.1021/jo00385a026
18. Dixit, V. M.; Laskovics, F. M.; Noall, W. I.; Poulter, C. D. *J. Org. Chem.* **1981,** *46,* 1967–1969. doi:10.1021/jo00322a060
19. McMullan, R.; Jeffrey, G. A. *J. Chem. Phys.* **1959,** *31,* 1231–1234. doi:10.1063/1.1730574

## License and Terms

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

# Sequence-specific RNA cleavage by PNA conjugates of the metal-free artificial ribonuclease tris(2-aminobenzimidazole)

Friederike Danneberg[1], Alice Ghidini[2], Plamena Dogandzhiyski[1], Elisabeth Kalden[1], Roger Strömberg[2] and Michael W. Göbel[*1]

Full Research Paper

Open Access

Address:
[1]Institute for Organic Chemistry and Chemical Biology, Goethe Universität Frankfurt, Max-von-Laue-Straße 7, D-60438 Frankfurt am Main, Germany and [2]Karolinska Institutet, Department of Biosciences and Nutrition, Novum, SE-141 83 Huddinge, Sweden

Email:
Michael W. Göbel[*] - m.goebel@chemie.uni-frankfurt.de

* Corresponding author

Keywords:
antisense; fluorescence correlation spectroscopy; guanidine; miRNA 20a; peptide nucleic acids

*Beilstein J. Org. Chem.* **2015,** *11,* 493–498.
doi:10.3762/bjoc.11.55

Received: 29 January 2015
Accepted: 02 April 2015
Published: 16 April 2015

This article is part of the Thematic Series "Nucleic acid chemistry".

Guest Editor: H.-A. Wagenknecht



## Abstract

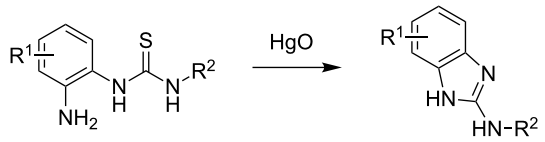
Tris(2-aminobenzimidazole) conjugates with antisense oligonucleotides are effective site-specific RNA cleavers. Their mechanism of action is independent of metal ions. Here we investigate conjugates with peptide nucleic acids (PNA). RNA degradation occurs with similar rates and substrate specificities as in experiments with DNA conjugates we performed earlier. Although aggregation phenomena are observed in some cases, proper substrate recognition is not compromised. While our previous synthesis of 2-aminobenzimidazoles required an HgO induced cyclization step, a mercury free variant is described herein.

## Introduction

Sequence specific artificial ribonucleases are an attractive research target for several reasons. On the one hand, they can improve our mechanistic understanding of natural ribonucleases and of phosphodiester reactivity in general. On the other hand, a wide range of practical applications is conceivable for such RNA cleavers ranging from tools in molecular biology to chemotherapeutics targeting mRNAs or miRNAs. Although most site-selective artificial ribonucleases consist of a catalytic unit (both with and without metal ions) attached to a DNA oligonucleotide or 2'-*O*-methyloligoribonucleotide complementary to the targeted RNA [1-3], possible future in cell applications suggest the conjugation of RNA cleavers to oligonucleotide analogues such as peptide nucleic acids (PNA) [4]. The benefits of oligonucleotide analogues are higher affinity towards RNA [5] as well as improved resistance against biodegradation [6]. Furthermore, PNA can be applied to block miRNA functions in cell culture experiments [7].

PNA conjugates of both metal-containing and metal-free RNA cleavers have been effectively used as artificial nucleases towards different substrates [8-11]. Most notable is a copper-based PNAzyme which cleaves RNA site-specifically with substrate half-lifes as low as 30 minutes [12]. A PNA-10mer attached to diethylenetriamine showed a higher cleavage activity

[13] when compared to the corresponding 22mer DNA conjugate [14]. We therefore decided to investigate the activity of tris(2-aminobenzimidazole) [15] – a metal-free ribonuclease developed by us formerly – as a conjugate with PNA. Previous results for the corresponding DNA conjugates had shown sequence specific RNA cleavage with substrate half-lives in the range of 12 to 17 h [16]. Tris(2-aminobenzimidazole) **7** can be obtained in a six-step reaction sequence starting from tris(2-aminoethyl)amine (**1**) [15]. However, the use of toxic mercury(II) oxide for the formation of the benzimidazole moieties (Scheme 1) is a disadvantage of this synthesis especially with regard to future biological applications of the final product. Furthermore, conditions for these reactions are rather harsh and the yields may differ considerably. Here we report a new and mercury-free synthesis of tris(2-aminobenzimidazole) and its conjugation to PNA oligomers. In addition, the results of cleaving experiments with three different RNA substrates are presented.

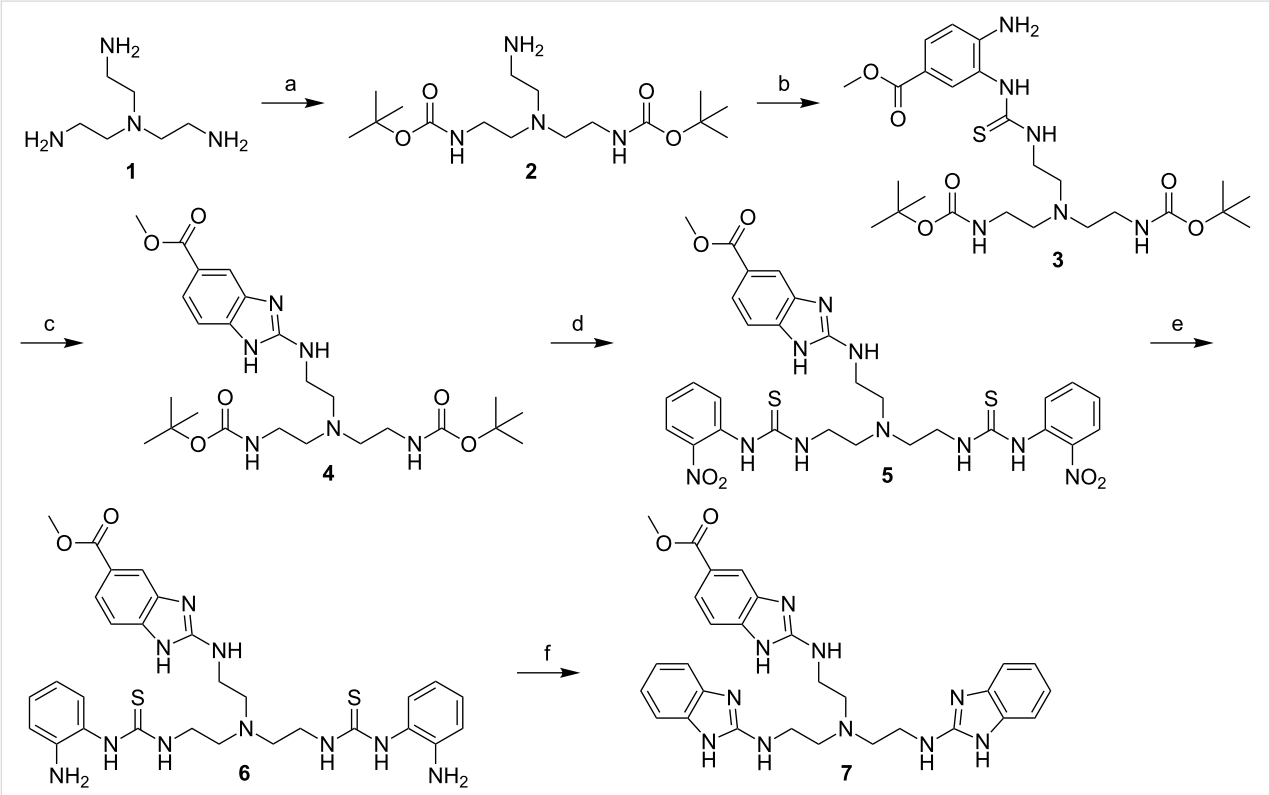



**Scheme 1:** Formation of the 2-aminobenzimidazole moiety.

## Results and Discussion

Alternative reagents to convert thioureas into guanidines include metal salts like $CuCl_2$ or $HgCl_2$ and carbodiimides such as 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDCI) or *N,N'*-diisopropylcarbodiimide (DIC) [17]. Another method developed by Lipton et al. uses Mukaiyama's reagent (2-chloro-1-methylpyridinium iodide) to prepare guanidines in high yields [18]. The reaction proceeds at ambient temperature. In contrast to ureas and thioureas formed as byproducts when using carbodiimides, the *N*-methylpyridine-2(1*H*)-thione resulting from Mukaiyama's reagent is soluble and can be easily removed by chromatography. As we were searching for a heavy metal-free approach towards tris(2-aminobenzimidazole), we adopted Lipton's method to the synthesis of aminobenzimidazoles.

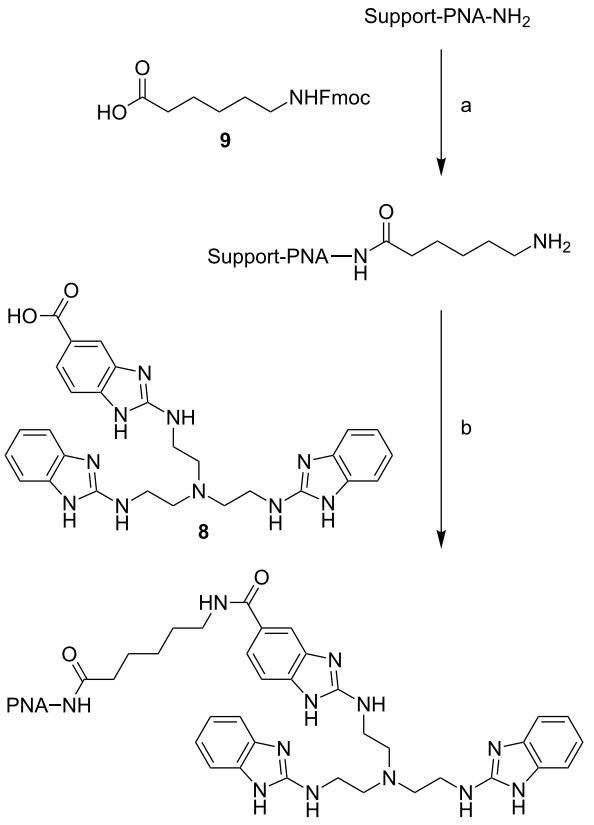
Thiourea **3** was prepared as described before [15], starting with the Boc-protection of tris(2-aminoethyl)amine (**1**) followed by the stepwise reaction of the free $NH_2$ group of product **2** with thiocarbonyldiimidazole and methyl 3,4-diaminobenzoate (Scheme 2). The cyclisation to form the first benzimidazole unit




**Scheme 2:** Synthesis of tris(2-aminobenzimidazole). Conditions: a: Boc-ON, THF, 0 °C to rt, 46 h, 45%; b: 1) 1,1'-thiocarbonyldiimidazole, imidazole, MeCN, 0 °C to rt, 1 h; 2) methyl 3,4-diaminobenzoate, MeCN, 50 °C, 4 h, rt, overnight, 79%; c: Mukaiyama's reagent, $NEt_3$, DMF, 20 h, rt, 88%; d: 1) $SOCl_2$, MeOH, 0 °C to rt, 3 h; 2) 2-nitrophenylisothiocyanate, $NEt_3$, EtOH, 0 °C to rt, overnight, 90%; e: 45 bar $H_2$, Pd/C (10%), MeOH sat. with $NH_3$, 60 °C, 5 h, rt, overnight, 37%; f: Mukaiyama's reagent, $NEt_3$, DMF, 22 h, rt, 56%. Boc-ON: 2-(*tert*-butoxycarbonyloxyimino)-2-phenylacetonitrile.

was then achieved by reacting thiourea **3** with Mukaiyama's reagent in the presence of $NEt_3$ at rt yielding 88% of **4**. Although the yield is slightly lower compared to the cyclisation with HgO (97% yield), the mild and metal-free reaction conditions are a big advantage of this new method. For the subsequent two conversions, again the known procedures were used. Thiourea **6** was prepared by deprotection of **4** with $SOCl_2$ in MeOH and addition of 2-nitrophenylisothiocyanate followed by reduction of the nitro groups with $H_2$. In the last step of the synthesis, again Mukaiyama's reagent was used for the formation of the two benzimidazole moieties. Stirring at rt for 22 h in the presence of $NEt_3$ yielded 56% of tris(2-aminobenzimidazole) **7** after HPLC purification. Ester hydrolysis with aq HCl led to the free carboxylic acid **8**, which after removal of the solvent was used for conjugation to PNA without further purification (Scheme 3).

Conjugation of tris(2-aminobenzimidazole) **8** was performed with fully protected PNA-oligomers still bound to the solid support. To increase the solubility of the final product, 10mers were attached to one, 15mers to two lysine units placed at the C-terminus. For conjugation, the Fmoc-protected 6-aminohexanoic acid **9** was bound to the PNA at the terminal amino group as a linker. After deprotection of the linker's amino function, tris(2-aminobenzimidazole) was attached. In both steps, DIC and HOBt were used as reagents. Coupling of the cleaver proceeded in high yields with no unconjugated PNA detectable in the MALDI mass spectra of the crude products. All conjugates were purified by RP-HPLC and isolated as trifluoroacetate salts. Starting from 20–30 mg of Rink amide MBHA resin (0.66 mmol/g) even manual synthesis routinely yielded milligram amounts of the purified final product. This compares favorably to the yield of DNA conjugates we have studied earlier [16].




**Scheme 3:** Synthesis of PNA conjugates. Conditions: a: 1) **9**, HOBt, DIC, DMF, rt, 24 h; 2) piperidine, DMF, rt, 30 min; b: 1) **8**, HOBt, DIC, $NEt_3$, DMF, rt, 72 h; 2) TFA/$H_2O$/triisopropysilane (95:2.5:2.5), rt, 3 h.

Cleavage experiments were run with Cy5-labeled RNA, the fluorescent dye permitting the detection and quantification of fragments by gel electrophoresis in a DNA sequencer (Figure 1). Substrates **15** and **16** have been used for comparison with our previous studies [16]. The third substrate **17** has the

Conjugates with trisbenzimidazole:

**Lys-aacagtcctc**-linker-cleaver **10**

**Lys-gctgacggct**-linker-cleaver **11**

**LysLys-tatggaacagtcctc**-linker-cleaver **12**

**LysLys-gatcggctgacggct**-linker-cleaver **13**

**LysLys-atttcacgaatatca**-linker-cleaver **14**

RNA substrates and accessory oligonucleotides:

**Cy5-$T_{10}$-ribo($5'$AUACCUUGUCAGGAGAAGAGAGGCCGUUA)-$T_4{}^{3'}$ 15**

**Cy5-$T_{10}$-ribo($5'$CUAGCCGACUGCCGAUCUCGCUGACUGAC)-$T_4{}^{3'}$ 16**

**Cy5-$T_{10}$-ribo($5'$UAAAGUGCUUAUAGUGCAGGUAG)-$T_4{}^{3'}$ 17**

**Cy5-TTTTTTTTTTTTTTTTTTTTU$^{3'}$ 18**

**TGTGGAATTGTGAGCGGATA$^{3'}$ 19**

**ribo($5'$GGCGCUAGCCGACUGCCGAUCUCGCUGACUGAC)$^{3'}$ 20**

**Figure 1:** Sequences of PNA conjugates **10–14** and oligonucleotides **15–20**. Lysines are attached to the *C*-terminus, linker and cleaver to the *N*-terminus. The parts with base complementarity to 15mer PNAs, i.e., the binding sites, are underlined. When 10mer PNAs bind to substrates **15** and **16**, the first five ribonucleotides at the 5' end will remain single stranded.

sequence of miRNA 20a, a member of the oncogenic miRNA 17–92 cluster [19] that is a promising target for site-specific RNA cleavers. All RNA parts of substrates **15**–**17** are embedded into a stretch of nonreacting deoxythymidine residues to improve the separation and resolution of substrate and fragment peaks in the sequencer.

Incubation of RNA **15** with conjugate **12**, which has the same 15mer sequence as a previously tested DNA conjugate [16], led to 8 dye labeled fragments, all located in the non-hybridized part of the substrate. Most prominent is cleavage after G(18) (Figure 2, lane b; note the slight shift and duplication of the substrate peak compared to the hydrolysis ladder shown in lane f). The product distribution is considerably broader compared to degradation of RNA **15** by the corresponding DNA conjugate. Similar results were obtained by reaction of substrate **17** with 15mer-conjugate **14** (lane e). Here, A(18) and G(19) are the main cleaving sites. Unexpectedly, no distinct cleavage pattern could be observed with the 15mer-conjugate **13** and its cognate RNA **16**. Instead, the electropherogram showed a shift of several signals including the substrate peak (lane d, compare to hydrolysis ladder of **16** in lane g), making it difficult to assign the signals to individual cleavage fragments. However, the position of the signals relative to the substrate peak suggests that cleavage occurred only in the non-hybridized part of the RNA. We explain this effect by incomplete PNA–RNA strand separation during electrophoresis due to stronger hybridization resulting from the G/C-rich sequence of the conjugate (10 G/C-base pairs in contrast to 7 and 4 for conjugates **12** and **14**, respectively). The fact that incubation of RNA **15** with PNA **12** also leads to a slight, but much weaker shift of signals (lane b), supports this assumption. The first four nucleotides in the single stranded part of RNA **16** where cleavage takes place have only pyrimidine bases, while RNA **15** and **17** have mainly purine bases in the corresponding regions. If the tris(2-aminobenzimidazole) cleaver interacts with the purine bases (e.g., by stacking) it is not impossible that this could also be related to the apparent lower rate and lower selectivity in cleavage of RNA **16** with both **11** and **13**. Consistent with that idea, we previously had observed similar effects when RNAs **15** and **16** were cleaved by DNA-catalyst conjugates analogous to **12** and **13** [16].

As it is known that PNA binds more tightly to RNA than does DNA [5], 10mer-conjugates **10** and **11** complementary to substrates **15** and **16** were also tested in cleavage experiments. Compared to the 15mers, the first 5 monomers were omitted in order to keep the cleavage site unchanged. Results for the incubation of **15** with **10** were as expected (lane a). In contrast to incubation of substrate **16** with 15mer-conjugate **13**, no signal shift is seen with the 10mer-conjugate **11**. However, the

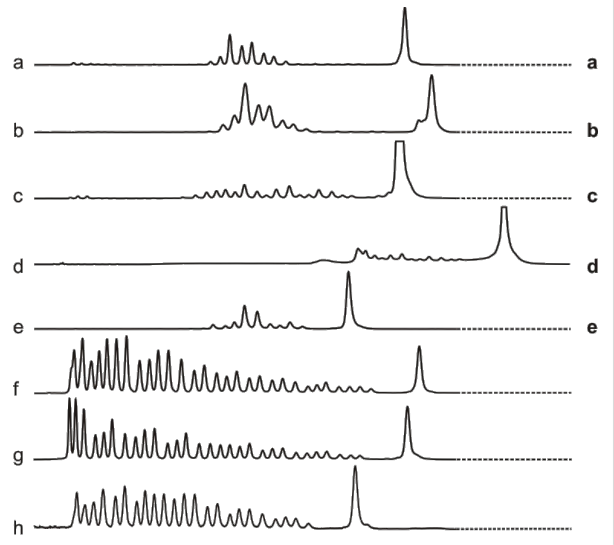



**Figure 2:** Cleavage of RNA by their corresponding PNA conjugates (150 nM substrate, 750 nM conjugate, 50 mM Tris-HCl, pH 8, 37 °C, 20 h). Lane a: conjugate **10** and substrate **15**. Lane b: conjugate **12** and substrate **15**. Lane c: conjugate **11** and substrate **16**. Lane d: conjugate **13** and substrate **16**. Lane e: conjugate **14** and substrate **17**. Lanes f, g and h: hydrolysis patterns of **15**, **16** and **17** ($Na_2CO_3$).

cleavage occurs in a nonspecific way in nearly all positions not protected by hybridization with PNA (lane c). Though site specificity is low in this case, conjugate **13** is nevertheless substrate specific and does not degrade RNAs **15** or **17** under identical conditions (Supporting Information File 1). The relatively low site specificity, as discussed above, may be related to the pyrimidine rich sequence of RNA **16**.

To examine whether PNA conjugates **10**–**14** would form molecular aggregates with noncognate oligonucleotides, the diffusion time of Cy5-labeled DNA **18** (19 nM, diluted with 131 nM unlabeled DNA **19**) in absence and presence of different PNA conjugates (750 nM) was studied by fluorescence correlation spectroscopy (FCS) [15,20]. No significant change in diffusion time was observed upon addition of 10mer conjugate **10** to noncognate DNAs **18**/**19**, indicating that no aggregates were formed. In contrast, addition of 10mer conjugate **11** or of 15mer conjugates **12**, **13**, and **14** to DNAs **18**/**19** led to considerable changes in diffusion times, which was most evident for PNAs **11** and **13**. This clearly shows that molecular aggregates of oligonucleotides and PNA conjugates are formed in all of these cases [21]. Although such effects were absent in FCS experiments with analogous DNA conjugates [16], this result is not surprising as PNA, with its uncharged backbone, is more likely to form aggregates. In the presence of 10% of DMSO, no change in diffusion times was observed for conjugates **12** and **14**. Conjugates **11** and **13** still caused an effect but less pronounced than in the absence of DMSO, indicating that DMSO prevents aggregation partially. We also investigated

RNA **16** by FCS (19 nM **16**, diluted with 131 nM of the unlabeled analog **20**). Upon addition of 10mer PNA **11**, a large increase of diffusion times was observed that can only be explained by aggregation.

Aggregation of PNA conjugates with complementary RNA but even with noncognate oligonucleotides [21] might cause a loss of sequence specificity. It was of critical importance, therefore, to conduct cross reaction experiments of PNA conjugates **10**–**14** in all possible combinations with substrates **15**–**17**. Exemplarily, Figure 3 shows tests of conjugates **12** and **14** with RNAs **15** and **17**. Consistent data was obtained for all other possible combinations of conjugates and substrates (Supporting Information File 1), which proves that cleavage in the examined concentration range fully depends on Watson–Crick base pairing.

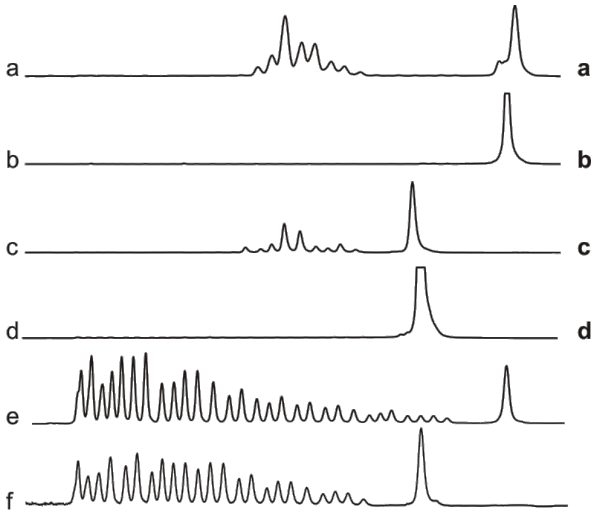



**Figure 3:** Substrate specificity of conjugates **12** and **14** (150 nM substrate, 750 nM conjugate, 50 mM Tris-HCl, pH 8, 37 °C, 20 h). Lane a: conjugate **12** and substrate **15**. Lane b: conjugate **14** and substrate **15**. Lane c: conjugate **14** and substrate **17**. Lane d: conjugate **12** and substrate **17**. Lanes e and f: hydrolysis patterns of **15** and **17** ($Na_2CO_3$).

Cleavage experiments with varying amounts of conjugates (Figure 4) show that – at concentrations below 1.5 µM – RNA cleavage by PNAs **10**–**12** and **14** obeys saturation kinetics. Full activity is reached at concentrations of 500–750 nM. In comparison to the shortened analog **10**, 15mer conjugate **12** shows higher cleavage yields of RNA **15** at concentrations from 62.5 to 1000 nM. Larger amounts of PNA, however, decrease the effectiveness of RNA degradation. This drop is quite considerable for the 15mer conjugates, whereas it is less significant for the 10mer analogs. The decreasing cleavage yields at high PNA concentrations are likely to result from increased aggregation, which might prevent the cleaver from interacting with the RNA backbone.

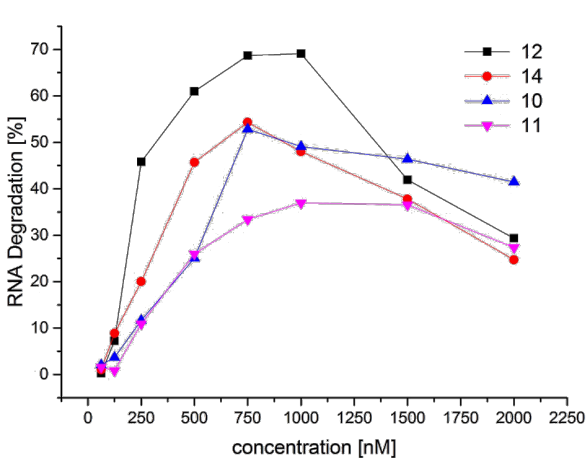



**Figure 4:** Cleavage of RNA substrates **15**, **16**, and **17** by their matching conjugates as a function of conjugate concentration (150 nM substrate, 62.5–2000 nM conjugate, 50 mM Tris-HCl, pH 8, 37 °C, 20 h). Data points are connected by lines for the sake of clarity.

For the reaction of substrate **15** with conjugate **12**, cleavage kinetics were studied in detail. Figure 5 shows RNA decay under saturation conditions as a function of time. While in the absence of a cleaver the substrate proved to be quite stable for several days, it was cleaved significantly by conjugate **12** within a few hours. Almost complete degradation was achieved after 60 h. As seen before [16], a small percentage of RNA remained intact even after one week. This could be due to structural imperfections in the synthetic substrate RNA preventing the PNA conjugate from hybridizing with its target molecule. The best fit of data to a first order rate law is shown as a solid curve: $k_1$ = 0.062 $h^{-1}$, $t_{1/2}$ = 11.2 h. This number comes close to the

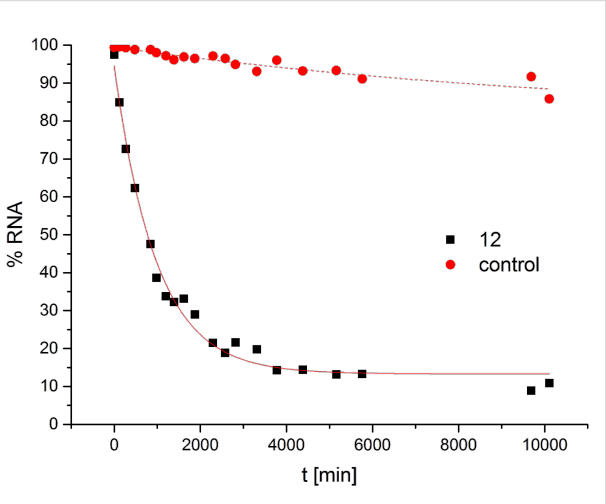



**Figure 5:** Cleavage kinetics of **15** in the presence and absence of conjugate **12**. Conditions: 150 nM substrate, 0 or 750 nM conjugate, 50 mM Tris-HCl, pH 8, 100 mM NaCl, 37 °C. The solid curves are calculated assuming first order kinetics.

half-life of **15** when cleaved by the analogous DNA conjugate ($t_{1/2}$ = 12.4 h) [16].

## Conclusion

We have shown that PNA conjugates of tris(2-aminobenzimidazole) are potent hydrolytic cleavers of RNA, exhibiting similar efficiency and substrate specificity as the corresponding DNA conjugates tested previously. Our optimized synthesis of tris(2-aminobenzimidazole) provides a new way to obtain the catalyst avoiding mercury salts with all their disadvantages (toxicity, harsh reaction conditions, variable yields). Conjugation to PNA oligomers can be readily achieved in high yields, providing an easy way to synthesize sequence specific metal-free artificial nucleases for a wide range of RNA substrates. Due to the tendency of PNA conjugates to form aggregates, a phenomenon not seen with their DNA analogs, optimization of the PNA oligomer length is necessary, especially for strands rich in G/C base pairs. Reduced activity of long PNA conjugates caused by aggregation can be a minor disadvantage in comparison to DNA conjugates. However, the higher affinity of PNA towards RNA allows using shorter PNA oligomers. In addition, the stability of PNA against biodegradation is another advantage especially with respect to possible in cell applications.

## Supporting Information

### Supporting Information File 1

Experimental procedures and characterization data.
[http://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-11-55-S1.pdf]

## Acknowledgements

Financial support by the European 7th framework program (238679, PhosChemRec), the Swedish Research Council and by the state of Hesse (LOEWE, SynChemBio) is gratefully acknowledged.

## References

1. Lönnberg, H. *Org. Biomol. Chem.* **2011,** *9,* 1687–1703. doi:10.1039/c0ob00486c
2. Åström, H.; Williams, N. H.; Strömberg, R. *Org. Biomol. Chem.* **2003,** *1,* 1461–1465. doi:10.1039/b212216b
3. Åström, H.; Strömberg, R. *Org. Biomol. Chem.* **2004,** *2,* 1901–1907. doi:10.1039/b403652b
4. Nielsen, P. E.; Egholm, M. *Curr. Issues Mol. Biol.* **1999,** *1,* 89–104.
5. Egholm, M.; Buchardt, O.; Christensen, L.; Behrens, C.; Freier, S. M.; Driver, D. A.; Berg, R. H.; Kim, S. K.; Norden, B.; Nielsen, P. E. *Nature* **1993,** *365,* 566–568. doi:10.1038/365566a0
6. Demidov, V. V.; Potaman, V. N.; Frank-Kamenetskii, M.; Egholm, M.; Buchardt, O.; Sönnichsen, S. H.; Nielsen, P. E. *Biochem. Pharmacol.* **1994,** *48,* 1310–1313. doi:10.1016/0006-2952(94)90171-6
7. Torres, A. G.; Fabani, M. M.; Vigorito, E.; Williams, D.; Al-Obaidi, N.; Wojciechowski, F.; Hudson, R. H. E.; Seitz, O.; Gait, M. J. *Nucleic Acids Res.* **2012,** *40,* 2152–2167. doi:10.1093/nar/gkr885
8. Whitney, A.; Gavory, G.; Balasubramanian, S. *Chem. Commun.* **2003,** 36–37. doi:10.1039/b210135a
9. Riguet, E.; Tripathi, S.; Chaubey, B.; Désiré, J.; Pandey, V. N.; Décout, J.-L. *J. Med. Chem.* **2004,** *47,* 4806–4809. doi:10.1021/jm049642d
10. Gaglione, M.; Milano, G.; Chambery, A.; Moggio, L.; Romanelli, A.; Messere, A. *Mol. BioSyst.* **2011,** *7,* 2490–2499. doi:10.1039/c1mb05131h
11. Murtola, M.; Strömberg, R. *Org. Biomol. Chem.* **2008,** *6,* 3837–3842. doi:10.1039/b810106j
12. Murtola, M.; Wenska, M.; Strömberg, R. *J. Am. Chem. Soc.* **2010,** *132,* 8984–8990. doi:10.1021/ja1008739
13. Petersen, L.; de Koning, M. C.; van Kuik-Romeijn, P.; Weterings, J.; Pol, C. J.; Platenburg, G.; Overhand, M.; van der Marel, G. A.; van Boom, J. H. *Bioconjugate Chem.* **2004,** *15,* 576–582. doi:10.1021/bc034219p
14. Komiyama, M.; Inokawa, T.; Yoshinari, K. *J. Chem. Soc., Chem. Commun.* **1995,** 77–78. doi:10.1039/c39950000077
15. Scheffer, U.; Strick, A.; Ludwig, V.; Peter, S.; Kalden, E.; Göbel, M. W. *J. Am. Chem. Soc.* **2005,** *127,* 2211–2217. doi:10.1021/ja0443934
16. Gnaccarini, C.; Peter, S.; Scheffer, U.; Vonhoff, S.; Klussmann, S.; Göbel, M. W. *J. Am. Chem. Soc.* **2006,** *128,* 8063–8067. doi:10.1021/ja061036f
17. Katritzky, A. R.; Rogovoy, B. V. *ARKIVOC* **2005,** No. iv, 49–87. doi:10.3998/ark.5550190.0006.406
18. Yong, Y. F.; Kowalski, J. A.; Lipton, M. A. *J. Org. Chem.* **1997,** *62,* 1540–1542. doi:10.1021/jo962196k
19. Lagos-Quintana, M.; Rauhut, R.; Lendeckel, W.; Tuschl, T. *Science* **2001,** *294,* 853–858. doi:10.1126/science.1064921
20. Müller, J. D.; Chen, Y.; Gratton, E. *Methods Enzymol.* **2003,** *361,* 69–92. doi:10.1016/S0076-6879(03)61006-2
21. Sahu, B.; Sacui, I.; Rapireddy, S.; Zanotti, K. J.; Bahal, R.; Armitage, B. A.; Ly, D. H. *J. Org. Chem.* **2011,** *76,* 5614–5627. doi:10.1021/jo200482d

## License and Terms

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

# Terminal lipophilization of a unique DNA dodecamer by various nucleolipid headgroups: Their incorporation into artificial lipid bilayers and hydrodynamic properties

Emma Werz[1,2] and Helmut Rosemeyer[*1]

Full Research Paper

Open Access

Address:
[1]Organic Chemistry I - Bioorganic Chemistry, Institute of Chemistry of New Materials, University of Osnabrück, Barbarastr. 7, 49069 Osnabrück, Germany and [2]Ionovation GmbH, Westerbreite 7, 49078 Osnabrück, Germany

Email:
Helmut Rosemeyer[*] - Helmut.Rosemeyer@uos.de

* Corresponding author

Keywords:
lipid bilayer; lipo-oligonucleotides; nucleic acids; nucleolipids

*Beilstein J. Org. Chem.* **2015,** *11,* 913–929.
doi:10.3762/bjoc.11.103

Received: 23 February 2015
Accepted: 08 May 2015
Published: 01 June 2015

This article is part of the Thematic Series "Nucleic acid chemistry" and is dedicated to Prof. Dr. Richard Wagner, University of Osnabrück, Germany.

Guest Editor: H.-A. Wagenknecht



## Abstract

A series of six cyanine-5-labeled oligonucleotides (LONs **10**–**15**), each terminally lipophilized with different nucleolipid head groups, were synthesized using the recently prepared phosphoramidites **4b**–**9b**. The insertion of the LONs within an artificial lipid bilayer, composed of 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine (POPC) and 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphoethanolamine (POPE), was studied by single molecule fluorescence spectroscopy and microscopy with the help of an optically transparent microfluidic sample carrier with perfusion capabilities. The incorporation of the lipo-oligonucleotides into the bilayer was studied with respect to efficiency (maximal bilayer brightness) as well as stability against perfusion (final stable bilayer brightness). Attempts to correlate these parameters with the log *P* values of the corresponding nucleolipid head groups failed, a result which clearly demonstrates that not only the lipophilicity but mainly the chemical structure and topology of the head group is of decisive importance for the optimal interaction of a lipo-oligonucleotide with an artificial lipid bilayer. Moreover, fluorescence half-live and diffusion time values were measured to determine the diffusion coefficients of the lipo-oligonucleotides.

## Introduction

For proper cell function, the biosynthetic lipophilization of proteins and carbohydrates, such as palmitoylation and farnesylation, is of decisive importance [1]. The same seems to be true for bacterial tRNAs [2], which was postulated recently and substantiated by the finding of geranylated tRNAs. Therefore a continuously growing interest in so-called lipo-oligonucleotides (LONs) [3-6] can be observed. In a preceding manuscript [7] we described the base-specific DNA duplex formation at a

lipid bilayer–water phase boundary using SYBR Green I staining. For this purpose, various nucleic acids of different length and sequence were lipophilized by appending a terminal non-nucleosidic lipid head group and then bound on the surface of an artificial lipid bilayer. Subsequently, equimolar amounts of complementary and non-complementary DNA strands as well as a SYBR Green I solution were added to the bilayer-bound lipo-oligonucleotide. The kinetics of the ternary complex formation at the bilayer surface as well as the stability of the complexes against perfusion was measured by determination of the bilayer brightness using fluorescence microscopy [8].

In the following, we describe the insertion of six unique single-stranded DNA dodecamers carrying different nucleolipid head groups as well as cyanine-5 (Cy5) at the 5'-($n$−1) position in artificial lipid bilayers. In this case we varied the amphiphilic head group but preserved the oligonucleotide DNA sequence. This sequence [5'-d(NL-(Cy5) TAG GTC AAT ACT); NL = nucleolipid] was chosen because it forms neither a hairpin structure nor a self-complementary duplex [9]. Particularly, for a successful delivery of siRNAs [10-12] their lipophilization and studies on the interactions of lipophilized nucleic acids with lipid bilayers is of growing importance. As lipophilic head groups we used a series of nucleolipid cyanoethyl phosphoramidites that were recently synthesized [13-17].

## Results and Discussion

Figure 1 indicates the positions of the pyrimidine β-D-ribonucleosides uridine (**1**, R = H), 5-methyluridine (**2**, R = $CH_3$) and 5-fluorouridine (**3**, R = F) which were lipophilized by different hydrophobic residues.

The following formulae (Figure 2) show the six nucleolipids **4a**–**9a** [13-17], which designate that mono-, sesqui-, and diterpenes as well as single and double-chained alkyl groups (completed by an alicyclic alkyl group), have been introduced for the lipophilization of the pyrimidine nucleosides. The resulting nucleolipids were converted into their corresponding 2-cyanoethyl phosphoramidites **4b**–**9b**.

Figure 3 shows the energy-minimized 3D structures of the lipophilic nucleoside headgroups **4a**–**9a**.

The reactive building blocks **4b**–**9b** were subsequently used to synthesize the following lipo-oligonucleotides (LONs), which have an identical nucleotide sequence in common as well as a cyanine-5 (Cy5) fluorophore at the ($n$−1) position. The integrity of all oligonucleotides was proven by MALDI–TOF mass spectrometry (see Experimental).

## Insertion of the six LONs into an artificial lipid bilayer and their specific characteristics

The different lipo-oligonucleotides (LON **10**–**15**) were incorporated into stable artificial lipid bilayers, which create two compartments (cis/trans channel) of an optically transparent microfluidic sample carrier with perfusion capabilities, as shown in Figure 4. Both the cis and the trans channel were filled with saline buffered solution. The injection of the different oligonucleotides into the cis channel followed by cycles of incubation and perfusion leads to the binding of the corresponding oligomer on the membrane, as detected by microscopy and single molecule fluorescence spectroscopy. The kinetics of the binding of the different bilayer-bound oligonucleotides were studied as well as their stability against perfusion.

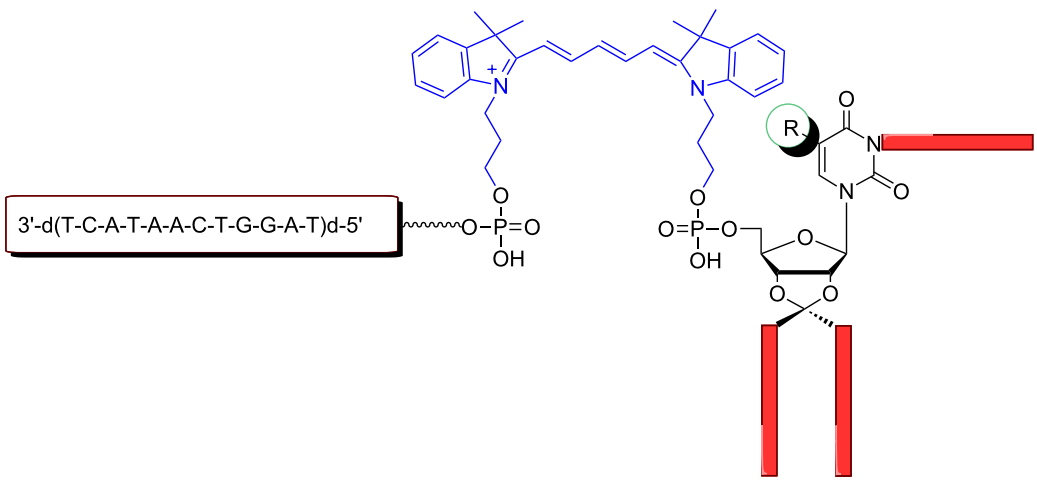



**Figure 1:** General structure of the fluorophore-labeled lipo-oligonucleotides and positions of lipophilization of the pyrimidine β-D-ribonucleoside head groups **1**–**3** [18].

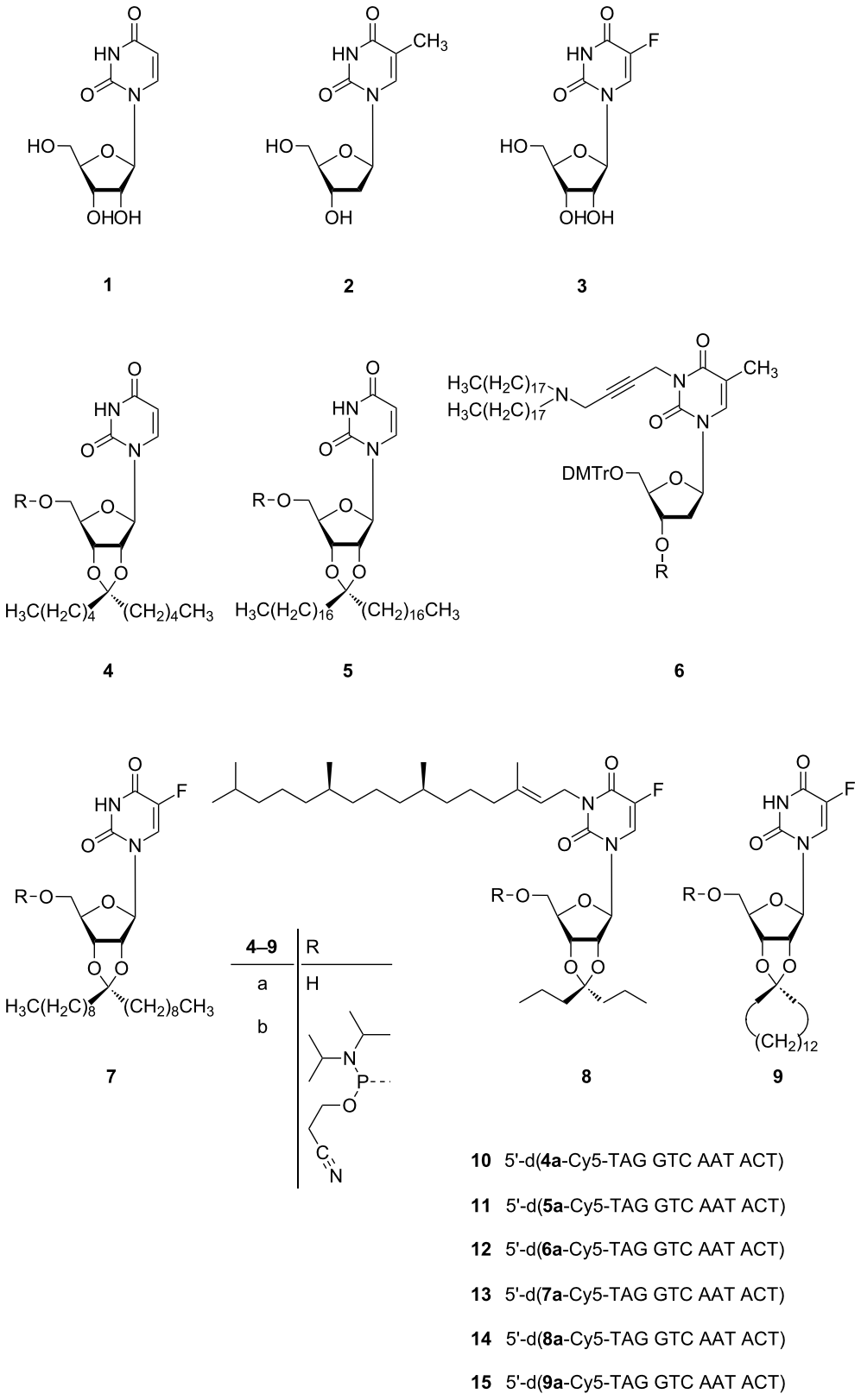



**Figure 2:** Chemical formulae **1**–**9** and lipo-oligonucleotide sequences **10**–**15**.

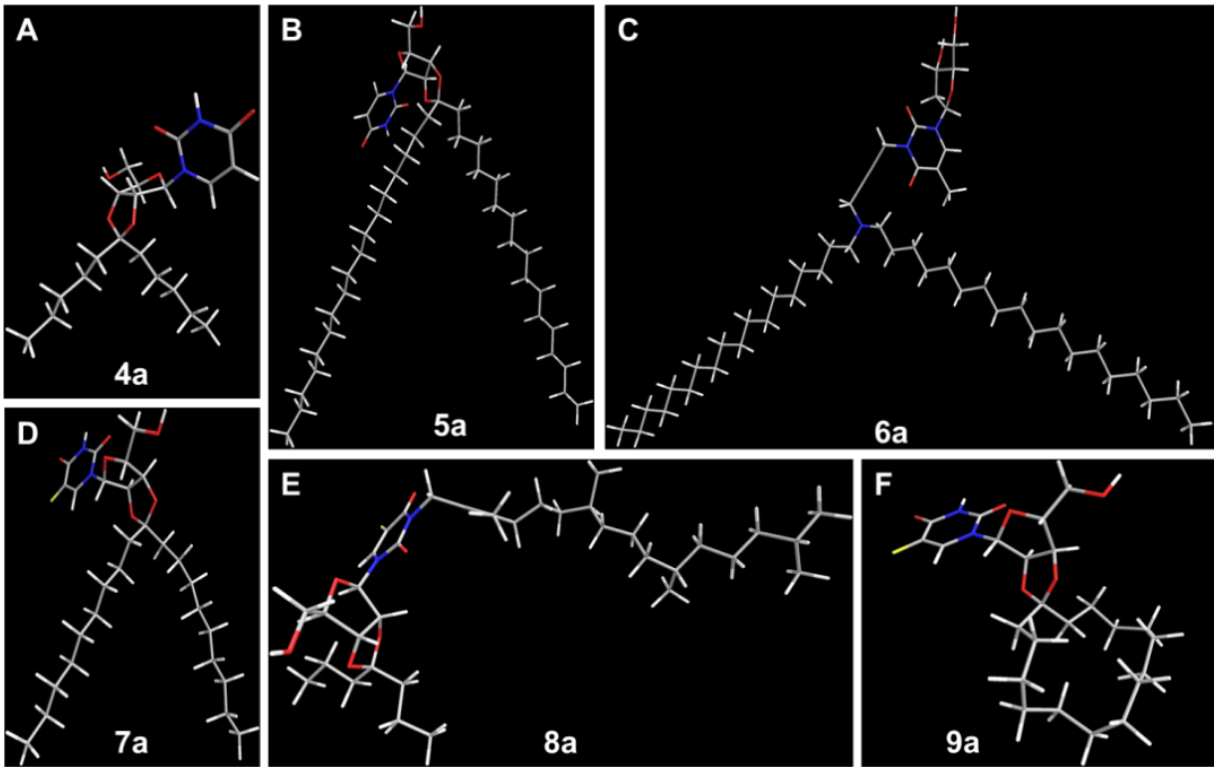



**Figure 3:** Energy-minimized 3D structures of the lipophilic nucleoside headgroups **4a**–**9a**. All 3D structures were calculated using the program ChemBio3D Ultra v. 12.0 (number of iterations, 387 ± 147, standard deviation of the mean).

Figure 4 illustrates the three main equilibria: the lipid bilayer–$H_2O$ boundary, the air–$H_2O$ interface boundary, and at the Teflon surface, which exist on the cis side of the bilayer chamber. This was proven experimentally as shown later in Figure 9.

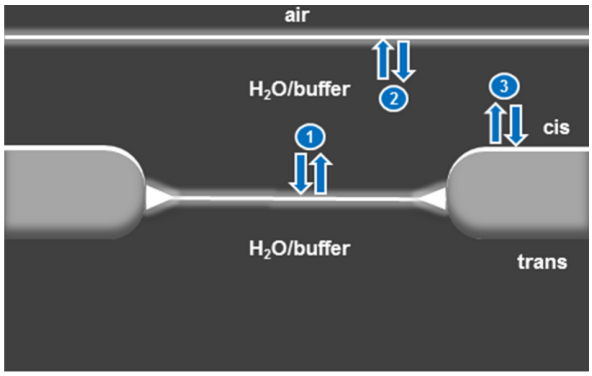



**Figure 4:** Schematic display of the three main equilibria existing on the cis side of the bilayer chamber.

The technique and some of the results are visualized in a video which can be downloaded free from the repository (RepOSitorium) of the library of the University of Osnabrück, Germany [18].

Figure 5 displays the bilayer brightness (both normalized to 100% as maximum as well as in kHz) as a function of the various events (perfusion and incubation time) for the lipo-oligomer **10**.

As can be seen, the incorporation of the lipo-oligomer **10** carrying an *O*-2',3'-undecylidene-modified uridine head group into the artificial lipid bilayer occurs spontaneously after 25 min, corresponding to the maximum brightness. However, even after a single perfusion, the brightness is reduced to about 15%, and after further incubation, it is again enhanced to 30% brightness. Subsequent perfusion and incubation periods over the course of roughly 1 h resulted in a constant brightness of 5–10%.

In contrast to this, an *O*-2',3'-pentatriacontanylidene-modified uridine head group shows a different behavior (Figure 6): the corresponding lipo-oligonucleotide **11** is bound directly after addition into the cis compartment of the sample carrier within the lipid bilayer (5 min). Once bound it exhibits a significantly higher resistance against perfusion of the cis chamber, and after approximately 100 min (8 perfusions), a constant bilayer brightness (40%, $1 \times 10^4$ kHz) is reached. This demonstrates the significant influence of the *O*-2',3'-alkylidene chain length of the nucleolipid head group on the

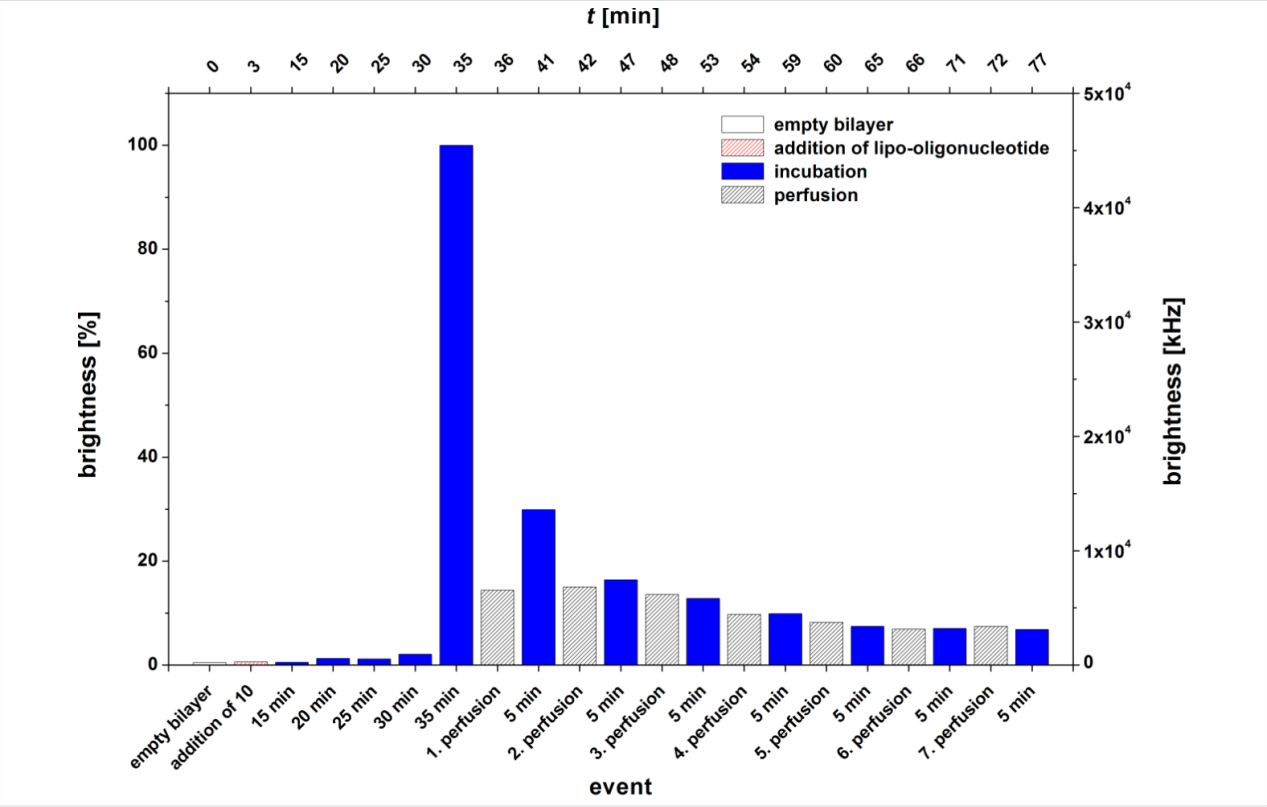



**Figure 5:** Bilayer brightness vs event graph of 5'-d(**4a**-Cy5-TAG GTC AAT ACT) (**10**).

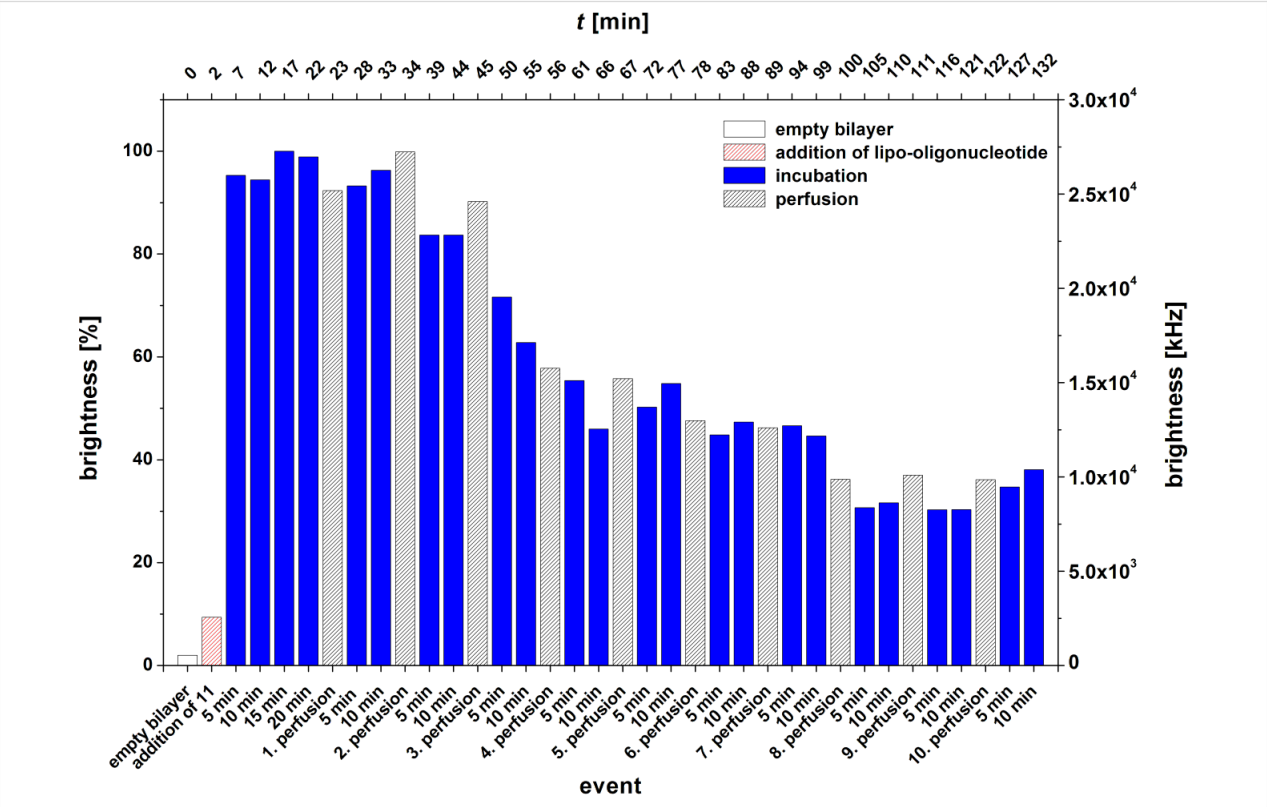



**Figure 6:** Bilayer brightness vs event graph of 5'-d(**5a**-Cy5-TAG GTC AAT ACT) (**11**).

stability of the lipo-oligonucleotide within the lipid bilayer against perfusion.

Figure 7 presents the bilayer brightness as a function of the perfusion/incubation events for the oligomer **12**. This oligomer is 5'-terminated with a nucleolipid head group which carries a double-chained lipid with a clickable group at N(3) of thymidine (**6a**). Among all nucleolipid head groups, this is the only one which can be incorporated once or more at each position of a growing oligonucleotide chain. As was found for the oligomer **10** (Figure 5), a maximum brightness was not immediately reached, but rather after a long incubation time (42 min). However, in contrast to the aforementioned experiment, the oligomer **12** is highly resistant against perfusions. After 3 perfusions, interrupted by incubation periods of 10 min each, a constant normalized bilayer brightness of about 60–65% is reached, which stays constant for more than 1 h and 6 perfusions.

Figure 8 displays the adsorption and desorption results for the lipo-oligonucleotide **13** carrying an *O*-2',3'-nonadecylidene-lipophilized 5-fluorouridine at the 5'-terminus. This oligomer shows a somewhat unusual behavior upon binding on the lipid bilayer. After 70 min the bilayer brightness reaches a first maximum at about 45%, which stays constant for about 10 min, even during 2 perfusions. Subsequently, however, between $t$ = 81 min and $t$ = 99 min, the bilayer brightness significantly increases from 40 to 100% (third to fifth perfusion). Further perfusions (up to the 22nd) and incubation periods (until $t$ = 208 min) lead to an exponential decrease of the brightness value to almost zero. In order to study this further, the level of the buffer phase in the cis compartment was lowered until it could be visualized within the focus of the fluorescence microscope (Figure 9).

From the z-direction scan shown in Figure 9E (for which the air–buffer boundary was lowered until it could be observed within the focus), it can be recognized that the lipo-oligonucleotide is concentrated at various phase boundaries: (1) at the lipid bilayer–$H_2O$ boundary, (2) at the air–$H_2O$ phase boundary, as well as (3) at the Teflon surface. Therefore, the further enhancement of the bilayer brightness (as in Figure 4 at $t$ = 81 min) might be due to a redelivery of the lipo-oligonucleotide from (2) or (3).

Transport of the oligomer through the aqueous phase most likely occurs in the form of micelles. For example, the infiltration of such micelles was demonstrated after the 9th perfusion step (Figure 4). Moreover, from Figure 9E it can be seen that the bilayer is significantly arched upwards towards the trans

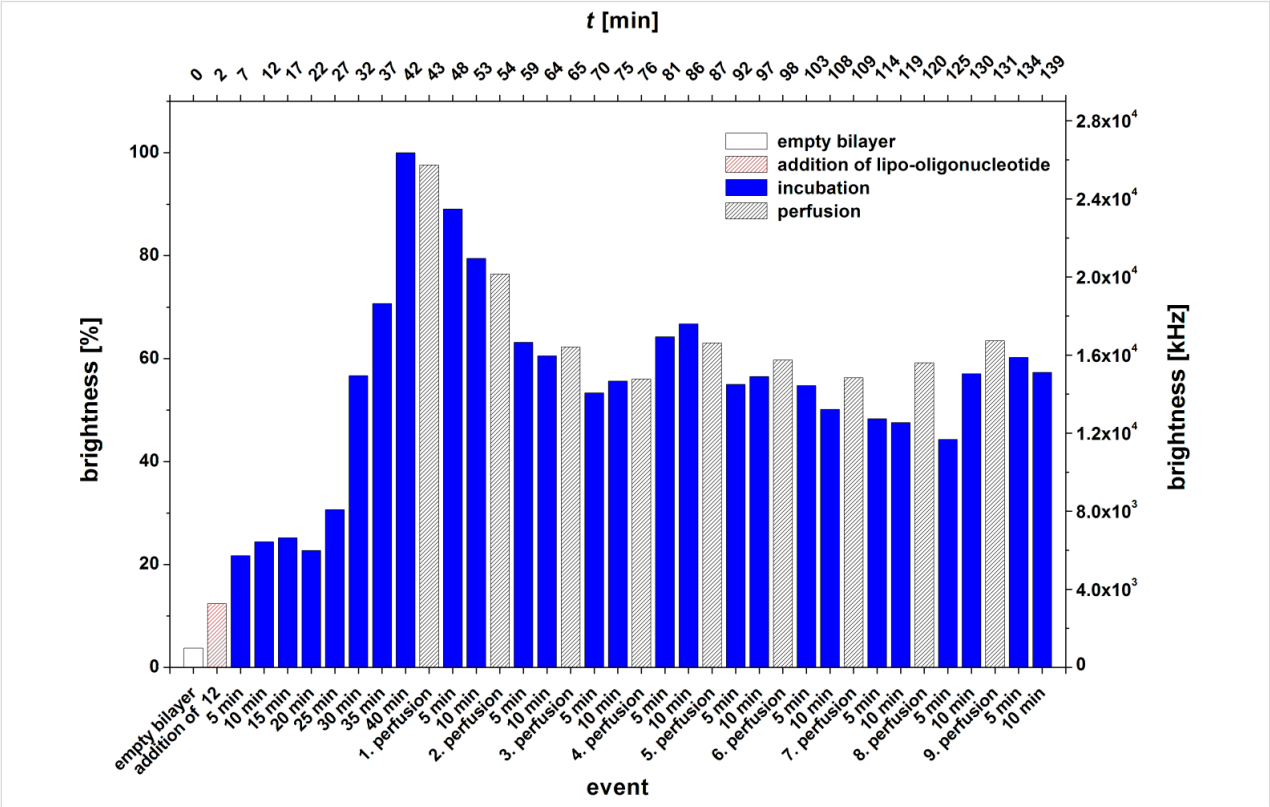



**Figure 7:** Bilayer brightness vs event graph of 5'-d(**6a**-Cy5-TAG GTC AAT ACT) (**12**).

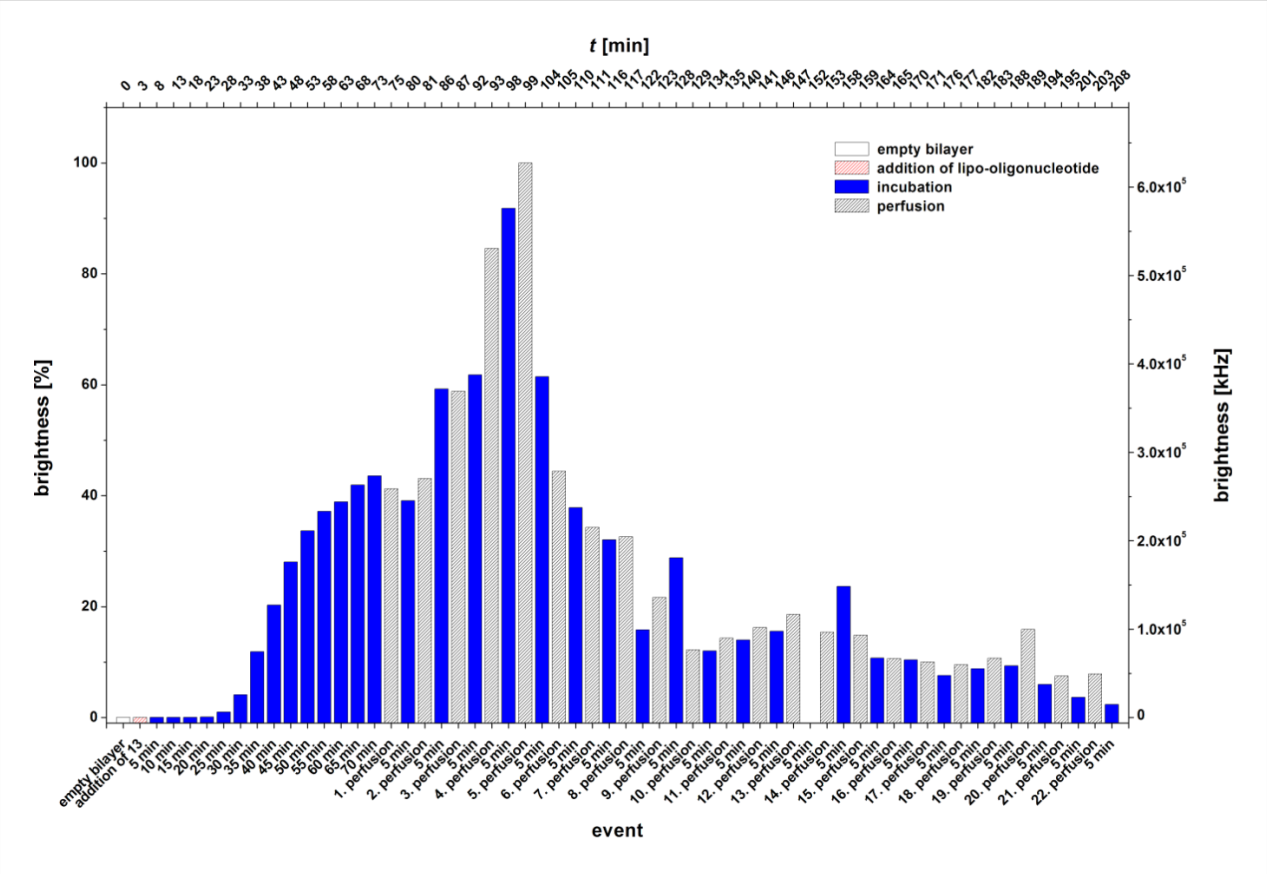



**Figure 8:** Bilayer brightness vs event graph of 5'-d(**7a**-Cy5-TAG GTC AAT ACT) (**13**).

chamber so that it is out of focus. This indicates the development of a Donnan pressure caused by the existence of significantly different osmotic pressures within the cis and the trans compartment, which is due to the fact that the polyanionic nucleic acid is unable to cross the artificial bilayer.

Figure 10 shows the bilayer brightness vs the various events for the lipo-oligonucleotide **14**. This oligomer carries a phytyl residue at the N(3) atom of 5-fluorouridine (**8a**). Nature uses this diterpenoid group for a membrane binding of chlorophyll, tocopherol and other biomolecules.

The lipo-oligonucleotide is characterized by being very quickly bound at the lipid bilayer. After 15 min the maximum brightness is reached followed by an exponential decrease over the course of six perfusion/incubation periods. A stable bilayer brightness is reached after approximately 1 h after, which remains constant for at least another 90 min.

The incorporation kinetics and stability of the lipo-oligonucleotide **15** are shown in Figure 11. As the nucleolipid head group, it contains 5-fluorouridine carrying an alicyclic *O*-2',3'-cyclopentadecanylidene ring. As can be seen, the binding within or (more likely) at the lipid bilayer occurs relatively slowly within 55–60 min until the maximum brightness of the bilayer is reached. A single perfusion reduces the brightness to 55%. Further perfusion, followed by incubation periods of 5 min, leads to an almost complete removal of the lipo-oligonucleotide ($t \leq 80$ min). We therefore assume that the lipophilic moiety of the head group is too bulky (Figure 3F) to be fully inserted into the bilayer and that the lipo-oligonucleotide is only loosely adsorbed at the bilayer surface.

From these results, it can be stated that the insertion rate of a lipo-oligonucleotide into an artificial lipid bilayer as well as its stability within the bilayer does not exclusively depend on the lipophilicity of the head group but also (i) results from the three-dimensional structure of the latter, (ii) from the length of the lipophilic moiety and (iii) from the position of the lipid residue at the nucleoside.

For the LONs **10** and **11** we observed that the elongation of the C-atom chain length from 5 to 17 does not necessarily lead to an enhancement of the amount of the LON incorporated into the bilayer. LON **10** is on one hand side incorporated very fast and in a high amount, and on the other hand, 85–90% is released

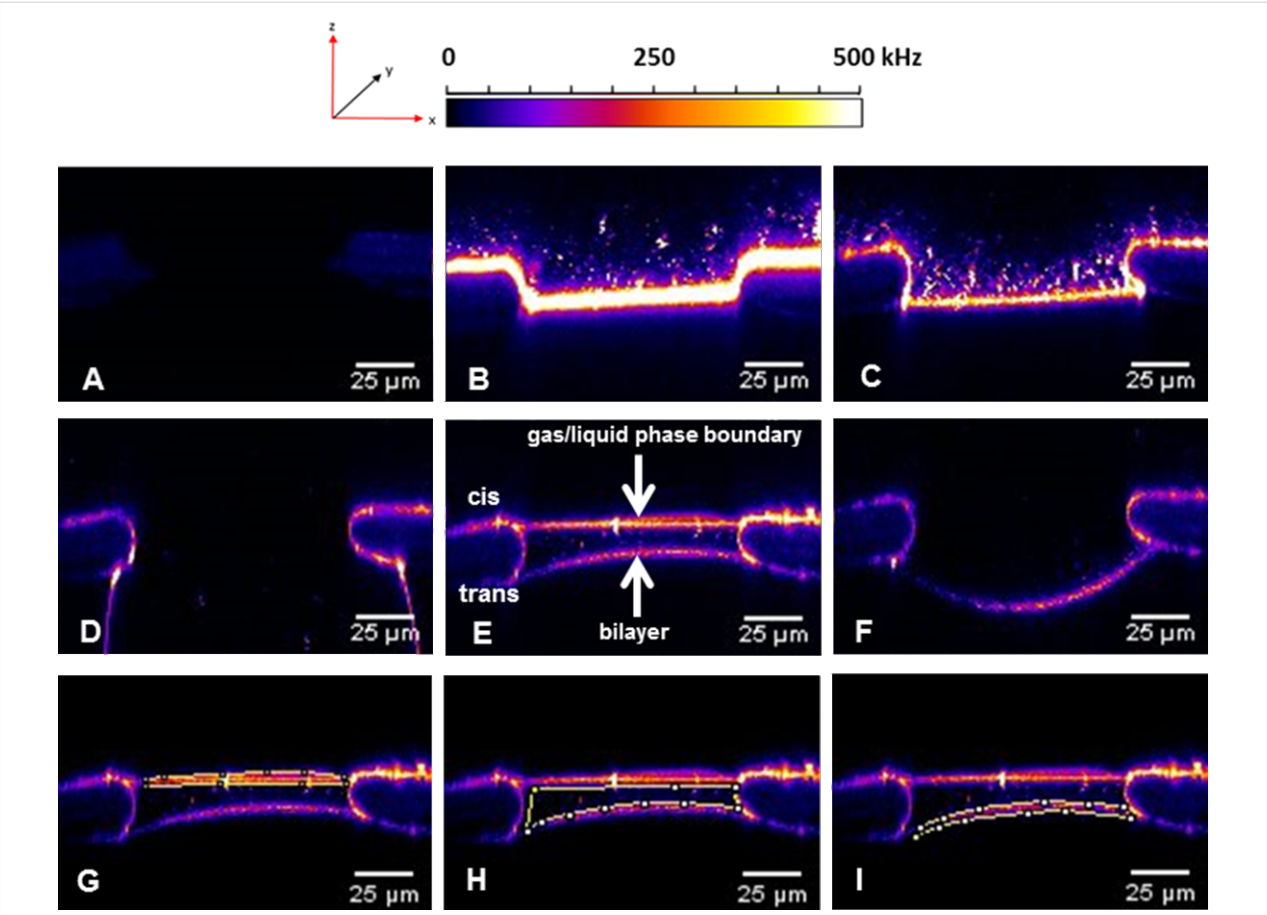



**Figure 9:** Infiltration of lipo-oligonucleotide (**13**) aggregates from the buffer into the lipid bilayer. Z-direction scan A) of a pure lipid bilayer without fluorescence, B) after the fourth perfusion/incubation step with high bilayer fluorescence and an accumulation of fluorescent aggregates/micelles above the bilayer. C) Z-direction scan of the bilayer after the ninth perfusion/incubation step, resulting in decreased bilayer brightness and an increased amount of micelles. D) After the 13th perfusion/incubation step, the bilayer sags into the trans compartment. After the 14th perfusion the bilayer returns to the starting position. E) After the 22nd perfusion/incubation step, the air–buffer boundary of the cis compartment was lowered over the bilayer. F) Z-direction scan after raising the buffer out of the focus. Brightness of the air–buffer boundary and of the lipid bilayer, each determined within the region of interest. G) The brightness of the liquid–air boundary is approximately 100,000 kHz. H) The brightness of the area between the two boundaries is approximately 28,000 kHz, and the brightness of the bilayer is approximately 50,000 kHz. I) The ratio of the brightness values of the liquid–air boundary and that of the bilayer is 2.

from the bilayer very easily by just a single perfusion (Figure 12A,B). In contrast, LON **11**, which differs from LON **10** only in the C-atom chain length of the lipid moiety, is incorporated only at a relatively low amount but stays bound to a significantly greater extent (40–50%), even after 10 perfusions (Figure 6).

Other lipophilic head groups such as LONs **13**–**15** lead to a fast and efficient incorporation into the bilayer (Figure 12A). Here, we observed LON **14** carrying a single-stranded phytyl residue at N(3) of 5-fluorouridine plus a short *O*-2',3'-ketal group. This was also observed for LON **13** with an *O*-2',3'-nonadecylidene ketal group which had a very strong binding within the bilayer, much stronger than that observed for LON **15** with a cyclic *O*-2',3'-ketal, which was washed out completely from the bilayer after only three perfusion steps (Figure 11).

LON **12** represents an oligomer with a long double-tailed lipid moiety at the N(3) position of thymidine. Interestingly, this LON is incorporated into the bilayer at a relatively low rate but proved highly stable against perfusion.

## Diffusion time measurements of LON 10 inserted within the lipid bilayer

Additionally, the diffusion times of LON **10** at positions 1, 2, and 3 (Figure 13A–C, Table 1) were measured in order to characterize its binding at and within the bilayer. Right above the bilayer (position 2), the diffusion of the oligomer is significantly reduced as compared to that in free solution (position 1).

From the fitted and normalized fluorescence correlation curves (Figure 13C), the corresponding diffusion times could be deduced as summarized in Table 1. The decreasing mobility of

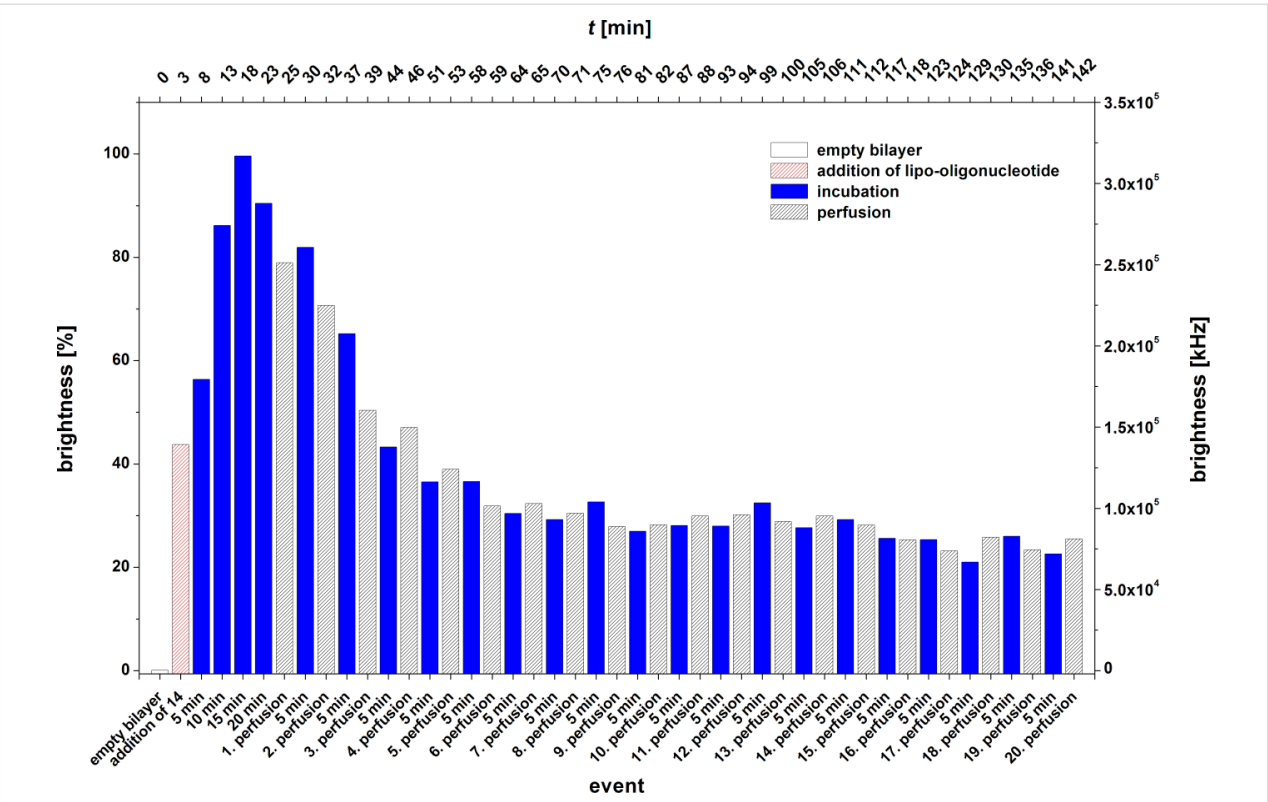



**Figure 10:** Bilayer brightness vs event graph of 5'-d(**8a**-Cy5-TAG GTC AAT ACT) (**14**).

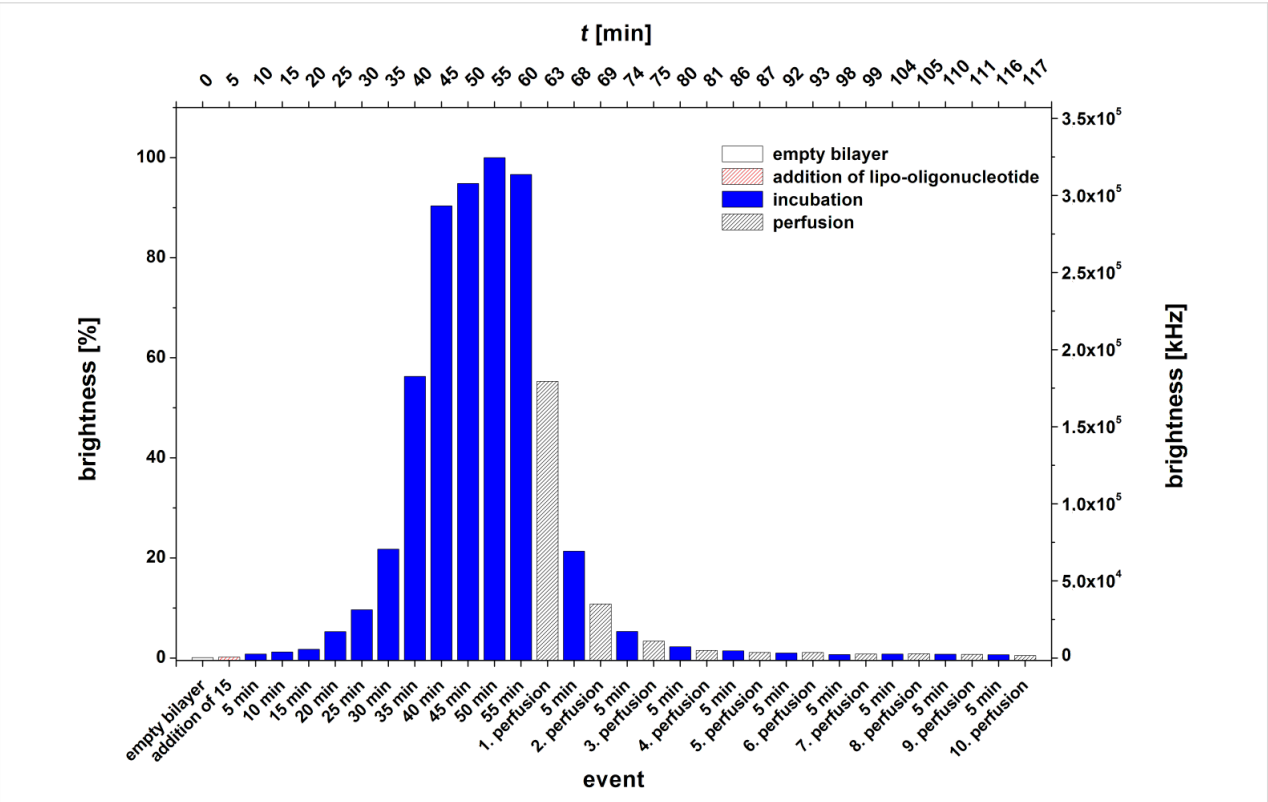



**Figure 11:** Bilayer brightness vs event graph of 5'-d(**9a**-Cy5-TAG GTC AAT ACT) (**15**).

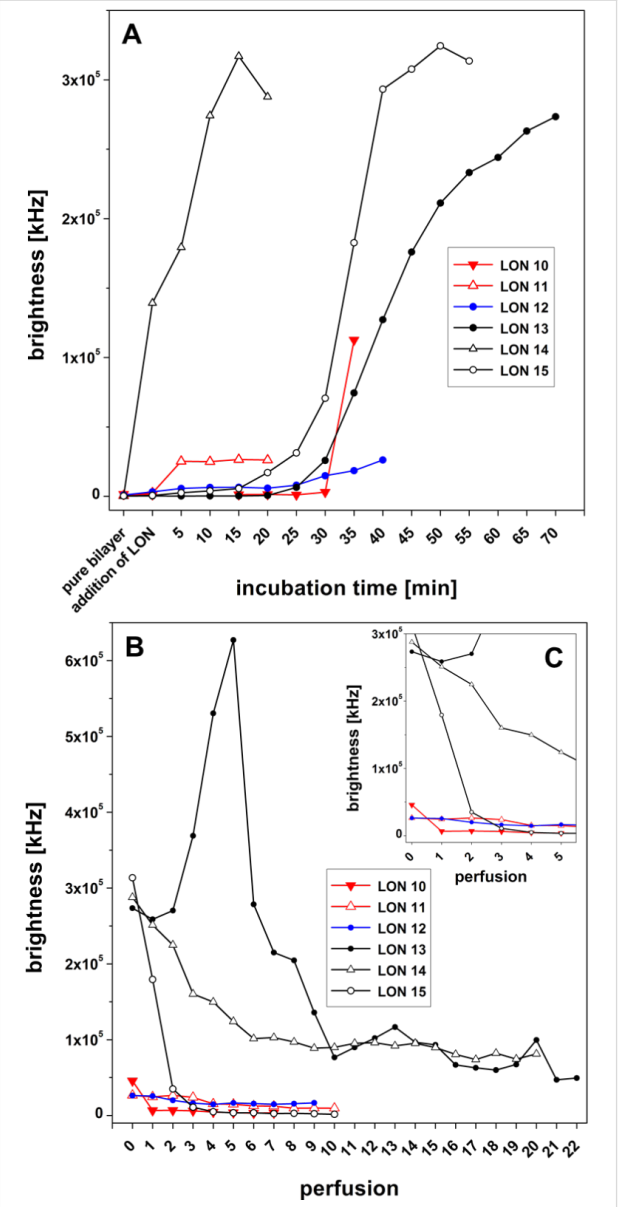



**Figure 12:** Summarizing the bilayer brightness for LONs **10–15** depending on the incubation time up to the maximum brightness (A), as well as on the number of perfusions (B) with details shown in (C).

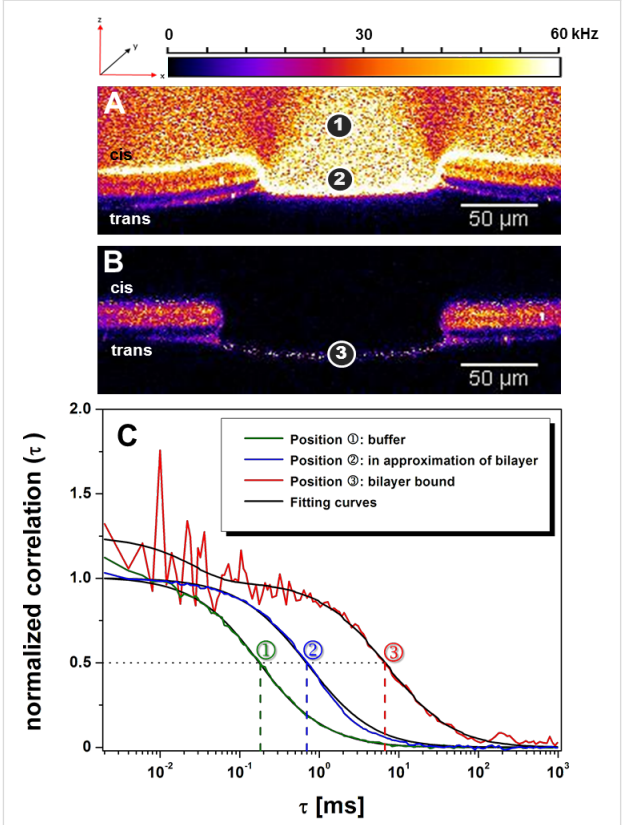



**Figure 13:** Diffusion time measurements of LON **10** (5 µL, 50 nM). A) Z-direction scan directly after the addition of the oligomer probe into the cis compartment of the bilayer chip. B) Z-direction scan after 7 perfusion/incubation cycles upon which unbound oligomer was completely washed out. This allows a diffusion time measurement of bilayer-bound LON **10**. C) Normalized and fitted correlation curves of the measured diffusion times at positions 1, 2 and 3.

**Table 1:** Diffusion times, $\tau_D$ (ms), of LON **10** in the presence of a lipid bilayer. Position: 1 - buffer; 2 - solution in close proximity to the bilayer; 3 - bilayer-bound.

| Position | | $\tau_D$ [ms] |
|---|---|---|
| **1** | in buffer | 0.25 ± 0.08 |
| **2** | close to the bilayer | 0.73 ± 0.01 |
| **3** | bilayer-bound | 4.25 ± 0.98 |

the oligomer from position 1 < position 2 < position 3 clearly indicates a loose adsorption at the bilayer (position 2) followed by incorporation within the bilayer (position 3).

Obviously, at position 2 (which is only ≈5–10 µm away from the bilayer), an attractive interaction exists between a LON and the lipids of the bilayer, which leads to a decrease in diffusion time. However, it is not clear if the diffusion slows down gradually. For this, it would be necessary to measure the diffusion at various positions above or below the bilayer. This is, however, difficult because of the floating of the bilayer.

## Correlation between the structure of the lipophilic head groups of the lipo-oligonucleotides and their stability within the artificial lipid bilayer

Next, we tried to correlate various parameters taken from the "brightness vs event" graphs such as the absolute maximum bilayer brightness values (in MHz) as well as the final stable brightness data (in MHz and %) with the log *P* values of the corresponding nucleolipid headgroup using Tetko's ALOGPS method (v. 3.01) for the calculation (Table 2). As can be seen in Table 2, no obvious correlation could be deduced. This means

that the lipophilicity of the nucleolipid head group is of minor importance for the effectiveness of lipo-oligonucleotide binding within the artificial lipid bilayer with respect to (i) the maximal bilayer brightness, (ii) the binding rate to reach this maximum brightness, (iii) the stability of the LON within the bilayer (expressed in number of perfusions to reach a stable brightness, and (iv) the final constant bilayer brightness. The log *P* values can only be correlated with the number of C-atoms of the lipid moieties, which are probably responsible for the binding to the bilayer. Of decisive importance for an effective insertion of a LON is obviously the chemical structure and topology of the corresponding nucleolipid head group. An acceptable compromise seems to be the phytol-labeled 5-fluorouridine head group, which has a mean log *P* value (6.43 ± 0.74) of the nucleolipid. The corresponding LON **14** reaches a high maximum bilayer brightness (316.9 MHz) within 15 min and proves also to be relatively stable within the bilayer (final brightness, 88.5 MHz). The log *P* values of the nucleotides uridine, 2-deoxythymidine and 5-fluorouridine are as follows: −1.90, −0.96 and −1.30 (±0.38 with a confidence interval of 66%). Thus they prove significantly smaller than the value of the corresponding lipo-oligonucleotides. The log *P* data gives insight into the augmentation of the lipophilicity of the nucleosides **4a**–**9a**.

## Calculations and experimental measurements of hydrodynamic properties of lipo-oligonucleotides

### Simulation and modeling of the LONs and calculation of their hydrodynamic parameters

For the calculation of the hydrodynamic properties of LONs **10**–**15**, first the 3D structure of the single-stranded oligomer sequence 5'-d(TAG GTC AAT ACT)-3' was calculated using the HyperChem software (Hypercube, ON, Canada) and then 5'-linked to the structures of the lipophilic head groups. The resulting 3D structures of the LONs were subsequently energy-minimized by applying the MMFF94 force field by using the freely available software package, Avogadro v 1.1.1 [19] (Figure 14).

For the subsequent calculations of the viscous drag and the diffusion coefficient, *D,* LONs **10**–**15** were adapted to the form of a stretched rotation ellipsoid, moving at random (Figure 15). For this purpose, from the 3D structures (Figure 14) of the LONs, the half axes *a* and *b* were deduced. The half axis *a* amounts to half of the sum $a_1 - a_3$ (Figure 15, Equation 2). The half axis *b* is half of the diameter of the DNA single strand of the LON and is approximated using Pythagoras' theorem. For this assessment, the distances between the phosphate groups $P_1$ and $P_5$, $P_5$ and $P_9$ as well as between $P_9$ and $P_1$ were taken from the 3D structures from the 12-mer DNA sequence using the Avogadro or the RasMol software. Half of the diameter corresponds to the half axis, *b*.

Because nucleic acids bind water molecules and are, therefore, covered by a hydration shell, 2.8 Å is added to the radius *b* which is the diameter of a water molecule. The mean value of the semi-minor axis, *b*, of a hydrated nucleic acid single strand is 12.03 ± 0.15 Å.

The frictional coefficient, *f*, of an ellipsoid, moving at random, with the semi-major axis *a* and semi-minor axis *b* [20], is described by Equation 1:

**Table 2:** Experimental results of the incorporation of the LONs into artificial lipid bilayers with respect to the log *P* values of the nucleolipid head groups.

| | LON | | | | | |
|---|---|---|---|---|---|---|
| | **10** | **11** | **12** | **13** | **14** | **15** |
| Nucleolipid head group (number of C-atoms of the lipid moieties) | **4a** (11) | **5a** (35) | **6a** (36) | **7a** (19) | **8**a (20) | **9a** (15) |
| log *P*[a] of head groups | 2.27[b] | 9.07[b] | 9.04[b] | 5.34[b] | 6.43[b] | 3.99[b] |
| Maximum bilayer brightness [MHz][c] | 45.77 | 26.52 | 26.22 | 273.49 (sh)/ 627.14[d] | 316.97 | 324.46 |
| Time to reach maximum brightness [min] | 35 | 15 | 40 | 70[e]/ 99[f] | 15 | 55 |
| Number of perfusion steps to reach a stable bilayer brightness | 4 | 7 | 3 | 1/ 16 | 6 | 4 |
| Final stable bilayer brightness [MHz] | 3.64 ± 0.56 | 9.78 ± 1.60 | 14.87 ± 1.52 | 258.16 ± 1.52/ 53.95 ± 20.68 | 88.51 ± 9.71 | 2.98 ± 0.99 |
| Final normalized stable bilayer brightness [%] | 7.96 ± 1.22 | 36.88 ± 6.04 | 56.73 ± 5.80 | 41.16 ± 2.00/ 8.60 ± 3.00 | 27.93 ± 3.06 | 0.92 ± 0.31 |

[a]log *P* values were calculated at the http://eadmet.com/de/physprop.php site with ePhysChem that uses ALOGPS v. 3.01 for log *P* prediction. [b]±0.74 with a confidence interval of 66%. [c]The brightness of the empty bilayer amounts to 425 ± 300 kHz (*N* = 6). [d]Reached after 5 perfusions. [e]First plateau, reached after 1 perfusion. [f]Constant final brightness, reached after 16 perfusions.

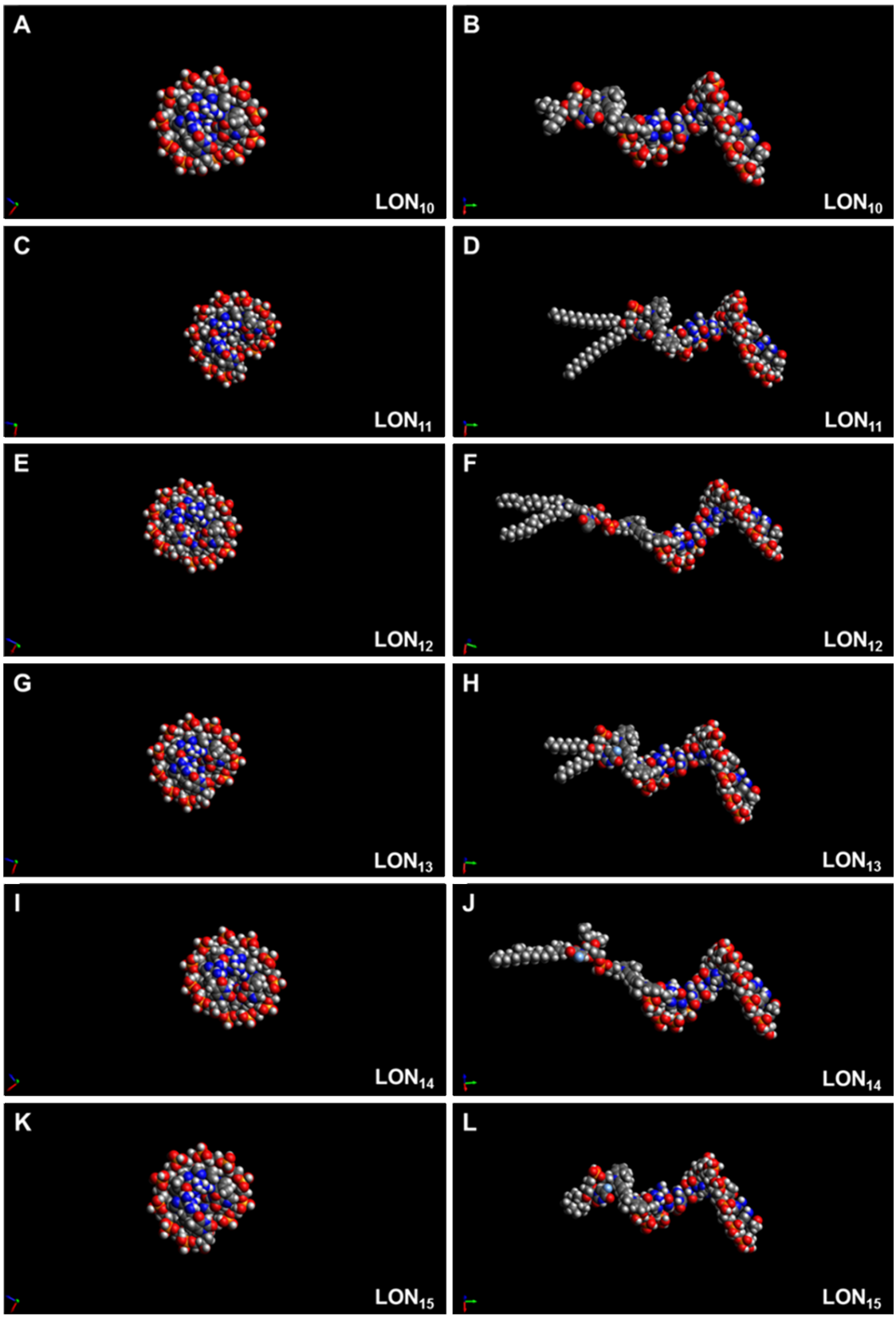



**Figure 14:** Hydrodynamic models of lipophilic oligonucleotides **10**–**15**. Views are perpendicular (right) or parallel (left) to the helix axis.

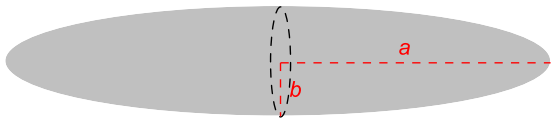



**Figure 15:** Scheme of a stretched rotation ellipsoid with the semi-major axes *a* and *b*.

$$f = \frac{6\pi\eta a}{\ln\frac{2a}{b}} \quad (1)$$

Figure 16 schematically displays the structure of a Cy5-labeled lipo-oligonucleotide for the calculation of the semi-major axis *a* according to Equation 2. According to the Stockes–Einstein relation for ellipsoidal particles, the diffusion coefficient, *D*, can be expressed as

$$D = \frac{k_B \cdot T}{f} \quad (2)$$

where $k_BT$ is the thermal energy, $f$ is the frictional coefficient of an ellipsoid, moving at random and η (Equation 1) is the solvent viscosity.

Table 3 summarizes the results of the theoretically calculated hydrodynamic properties of the lipo-oligonucleotides **10–15**. The mean diffusion coefficient was found to be $1.2 \cdot 10^{-6} \pm 0.08$ cm$^2$ s$^{-1}$.

We do not exclude the formation of LON aggregates, but under our experimental conditions with respect to concentration and temperature, we do not observe a strong aggregation tendency – a finding corroborated by the agreement between the calculated and measured hydrodynamic properties. Any aggregates are obviously not bound within the lipid bilayer, but only in free solution. They act to resupply LONs for the incorporation of LON monomers into the lipid bilayer (Figure 9) by disintegration.

## Fluorescence correlation spectroscopy (FCS) measurements and experimental determination of hydrodynamic properties

Using fluorescence correlation spectroscopy (FCS), we monitored the diffusion of Cy5-labeled lipophilized oligonucleotides **10–15** in aqueous solution within a confocal detection volume of 0.9 fL. The diffusion time, $\tau_D$ (Equation 3), depends on the mass and the shape of the molecule as well as on the dimensions of the detection focus $\omega_{x,y}$ in the x,y-dimension, in addition to the diffusion coefficient *D* as

**Table 3:** Hydrodynamic properties of the lipo-oligonucleotides **10–15**.

| LON | Nucleo-lipid head group[a] | $MW_{LON}$ [g·mol$^{-1}$] | $a_1$ DNA length [Å] | $a_2$ Cy5-base distance [Å] | $a_3$ lipid length [Å] | $2 \cdot b_{anh}$ DNA diameter [Å] | *A* Semi-major axis [Å][b] | $b_{hyd}$ Semi-minor axes [Å][c] | $f_{calc}$[d] × 10$^{-8}$ [g·s$^{-1}$] | $D_{calc}$[e] × 10$^{-6}$ [cm$^2$·s$^{-1}$] |
|---|---|---|---|---|---|---|---|---|---|---|
| **10** | **4a** (11)[a] | 4637.8 | 33.17 | 13.85 | 7.35 | 18.55 | 27.19 | 12.08 | 3.41 | 1.19 |
| **11** | **5a** (35)[a] | 4974.4 | 33.13 | 13.53 | 22.13 | 18.15 | 34.40 | 11.87 | 3.69 | 1.10 |
| **12** | **6a** (36)[a] | 5054.4 | 34.48 | 22.33 | 25.85 | 19.00 | 41.33 | 12.30 | 4.09 | 0.99 |
| **13** | **7a** (19)[a] | 4767.8 | 33.07 | 13.51 | 11.85 | 18.45 | 29.22 | 12.02 | 3.48 | 1.16 |
| **14** | **8a** (20)[a] | 4875.9 | 32.92 | 18.78 | 21.23 | 18.37 | 36.47 | 11.98 | 3.81 | 1.06 |
| **15** | **9a** (15)[a] | 4707.6 | 33.68 | 13.59 | 6.31 | 18.29 | 26.49 | 11.94 | 3.36 | 1.20 |

[a]Number of C-atoms of the lipid moieties. [b]From an approximation of an ellipsoid, moving at random [20]. [c]$b_{hyd}$, semi-major axis of hydrated ssDNA. [d]Calculated viscous drag coefficient of an ellipsoid, moving at random [20]. [e]Calculated diffusion coefficient.

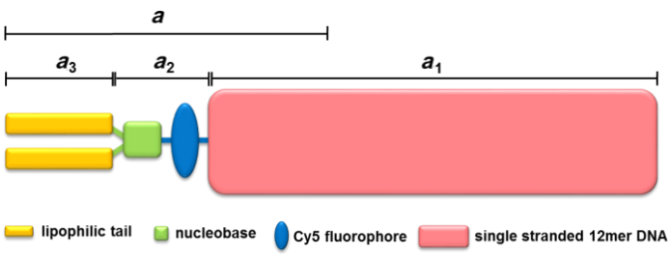



**Figure 16:** Schematic structure of a Cy5-labeled lipo-oligonucleotide.

$$\tau_D = \frac{\omega^2}{4 \cdot D} \quad (3)$$

Assuming a diffusion coefficient of $D = 4.3 \times 10^{-6}$ cm$^2$ s$^{-1}$ for MR121 [21,22] and a measured diffusion time of 60 ± 2.05 µs, we calculated the diffusion coefficients for the LONs to be 1.61 ± 0.06 × $10^{-6}$ cm$^2$ s$^{-1}$ (Table 4). This means that the ratio between the experimental and the calculated diffusion coefficients (Table 5, last column) is 1.45 ± 0.09 on average. Our data are comparable with those reported by Fernandes et al. in 2002 [23], which was also determined for DNA 12mers (1.22 and 1.45 × $10^{-6}$ cm$^2$ s$^{-1}$).

## Hydration of the lipo-oligonucleotides

In an aqueous environment, water molecules form a hydration shell around the nucleic acid. Consequently, the surface of the nucleic acid chain is smoothed, and the partial volume of the hydrated molecule ($V_{hyd}$) is increased by the volume of the water shell. By determining the $V_{hyd}$ of the LONs, one can calculate the degree of hydration. The anhydrous volume, $V_{anh}$, can be expressed by Equation 4:

$$V_{anh} = \frac{M}{N_A} \cdot \hat{v} \quad (4)$$

where $M$ is the molecular weight of the lipo-oligonucleotide, $N_A$ is Avogadro's number, and $\hat{v}$ is the partial specific volume.

For the calculation of $V_{anh}$ and $V_{hyd}$, the volume of the LON is adapted to be an elongated rotation ellipsoid as:

$$V_{ellipsoid} = \frac{4}{3} \cdot \pi \cdot a \cdot b^2 \quad (5)$$

where $b$ is the shorter semi-major axis. For the calculations, the parameters $a$ and $b_{anh}$ and $b_{hyd}$ (Table 3), respectively, were used.

The difference $V_{hyd}$−$V_{anh}$ describes the volume of the hydrodynamically bound water. The degree of hydration is often described as the ratio, δ, of bound water in grams per gram of molecule as described by

$$\delta = \left( \frac{V_{hyd}}{V_{anh}} - 1 \right) \hat{v} \cdot \rho \quad (6)$$

Our calculations demonstrate slight differences between the degree of hydration values depending on the different methods

**Table 4:** Experimental hydrodynamic data from FCS measurements.

| LON | Nucleolipid headgroup[a] | $\tau_D$ [ms][b] | $\chi^2$ [ms][c] | $D_{exp}$ × $10^{-6}$ [cm$^2$·s$^{-1}$][d] | $D_{exp}/D_{calc}$[e] |
|---|---|---|---|---|---|
| **10** | **4a** (11)[a] | 0.16 ± 0.007 | 0.12 | 1.62 | 1.37 |
| **11** | **5a** (35)[a] | 0.17 ± 0.015 | 0.27 | 1.54 | 1.41 |
| **12** | **6a** (36)[a] | 0.16 ± 0.005 | 0.23 | 1.62 | 1.61 |
| **13** | **7a** (19)[a] | 0.16 ± 0.006 | 0.16 | 1.63 | 1.40 |
| **14** | **8a** (20)[a] | 0.16 ± 0.010 | 0.28 | 1.57 | 1.48 |
| **15** | **9a** (15)[a] | 0.15 ± 0.011 | 0.14 | 1.68 | 1.42 |

[a]Number of C-atoms of the lipid moieties. [b]Measured diffusion time [ms]. [c]$\chi^2$: indicates the quality of the dataset. [d]Diffusion coefficient calculated from measured diffusion time, $\tau_D$. [e]Ratio of the experimentally determined diffusion coefficients and the calculated data.

**Table 5:** $V_{hyd}$, $V_{anh}$ and partial specific volume $\hat{v}$ as well as the degree of hydration (δ) for LONs **10**–**15**.

| LON | Hydration involving the whole ellipsoid, calculated with major semi-axis *a* | | | | | Hydration of DNA strand only, calculated with major semi-axis $a_1$ | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $V_{anh}$ × $10^{-20}$ [cm$^3$] | $V_{hyd}$ × $10^{-20}$ [cm$^3$] | $V_{hyd}$−$V_{anh}$ × $10^{-21}$ [cm$^3$] | $\hat{v}$ [cm$^3$ g$^{-1}$] | δ [cm$^3$ g$^{-1}$] | $V_{anh}$ × $10^{-21}$ [cm$^3$] | $V_{hyd}$ × $10^{-20}$ [cm$^3$] | $V_{hyd}$−$V_{anh}$ × $10^{-21}$ [cm$^3$] | $\hat{v}$ [cm$^3$ g$^{-1}$] | δ [cm$^3$ g$^{-1}$] |
| **10** | 0.98 | 1.66 | 6.81 | 1.3 | 0.9 | 5.97 | 1.01 | 4.15 | 0.8 | 0.5 |
| **11** | 1.19 | 2.03 | 8.45 | 1.4 | 1.0 | 5.71 | 0.98 | 4.07 | 0.7 | 0.5 |
| **12** | 1.56 | 2.62 | 10.56 | 1.9 | 1.2 | 6.51 | 1.09 | 4.41 | 0.8 | 0.5 |
| **13** | 1.04 | 1.77 | 7.28 | 1.3 | 0.9 | 5.89 | 1.00 | 4.12 | 0.7 | 0.5 |
| **14** | 1.29 | 2.19 | 9.05 | 1.6 | 1.1 | 5.81 | 0.99 | 4.08 | 0.7 | 0.5 |
| **15** | 0.94 | 1.60 | 6.62 | 1.2 | 0.8 | 5.90 | 1.00 | 4.16 | 0.8 | 0.5 |

of volume calculation: (i) calculation using $a$ or (ii) using $a_1/2$ (Figure 16). Assuming an increase in the hydrated shell with an increase in the total LONs, the hydration values amount to 0.8–1.2 g of $H_2O$ per gram of LON. Assuming, however, only hydration of the nucleic acid moiety, the hydration values amount to 0.5 g of $H_2O$ per g of molecules.

From IR measurements it could be concluded that 20 water molecules hydrate one nucleotide residue within a nucleic acid [24,25]. This results in a hydration of 1 g of $H_2O$ per gram. However, Bloomfield et al. in 2000 [26] as well as Fernandes et al. in 2002 [23] reported hydration degrees of 0.35 [26] or 0.3 g of water per gram of nucleotide (Table 5).

### Confocal fluorescence lifetime analysis (cFLA)

Next, the fluorescence lifetime values of the LONs **10**–**15** were measured and are listed in Figure 17A–C as well as in Table 6.

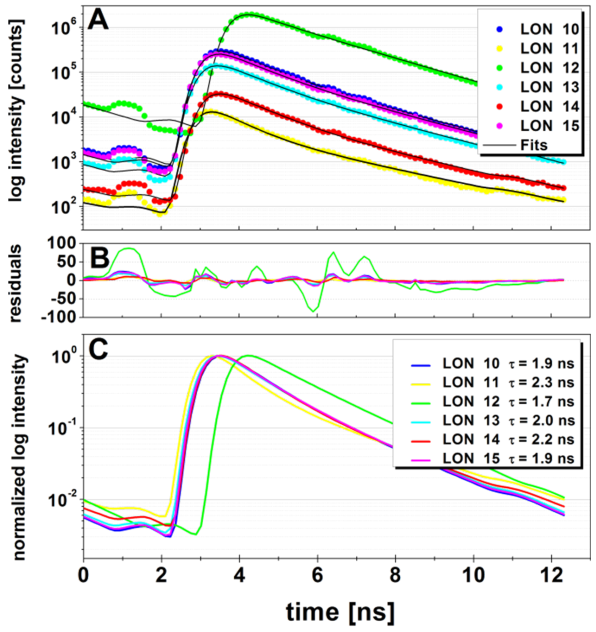



**Figure 17:** Fluorescence lifetime $\tau_F$ values of the LONs **10**–**15**. A) Logarithmic imaging of fluorescence intensity–time decay of LONs **10**–**15** and the corresponding fitting curves (–). B) The residuals show some minor systematic errors (the color legend is the same as in A and C). C) Normalized logarithmic fluorescence intensity of the fitting curves.

The fluorescence lifetime of the free Cy5 dye is reported to be 0.91 ns [27], 1.0 ns [28,29], and 1.0–1.4 ns [30]. In the literature, fluorescence lifetimes of free Cy5 dye attached to ssDNA as well as to dsDNA, respectively, also exist. The fluorescence lifetime, $\tau_F$, of two 5'-Cy5-ssDNA conjugates were measured as 1.8 ± 0.1 ns [31] for a 36mer and a 23mer [32]. For a Cy5-labeled 25mer, a value of 1.26 ns [27] was reported.

**Table 6:** Fluorescence lifetime values of LONs **10**–**15**.

| LON | $\tau_F$ [ns] |
|---|---|
| **10** | 1.92 ± 0.01 |
| **11** | 2.29 ± 0.08 |
| **12** | 1.73 ± 0.00 |
| **13** | 1.99 ± 0.01 |
| **14** | 2.23 ± 0.04 |
| **15** | 1.92 ± 0.01 |

For the LONs **10**–**15** described here, longer lifetime values between 1.7 and 2.3 ns were measured. These data are in accordance with corresponding data of Cy5–ssDNA, and also with those of Cy5–dsDNA.

## Conclusion

The studies described in this manuscript complement the research of various scientific groups whose work is focused on the formation of complex nanoarchitectures [33,34], the interaction of such nanostructures with lipid membranes [35-38], as well as the optimization of the in vivo delivery of lipophilic siRNA [10-12] (e.g., by DNA trafficking [7]).

Further studies from our group will include the binding of terminally lipophilized oligonucleotides at water–air or water–decane boundaries [39] as well as the transdermal application of such LONs through the human Stratum corneum by iontophoresis techniques as well as the transport across the blood–brain barrier [40].

## Experimental

**Materials.** 1-Palmitoyl-2-oleyl-*sn*-glycero-3-phosphoethanolamine (POPE) and 1-palmitoyl-2-oleyl-*sn*-glycero-3-phosphocholine (POPC) were purchased from Avanti Polar Lipids (Alabaster, Al), *n*-decane from Alfa Aesar (D-Karlsruhe), potassium chloride (KCl), 3-morpholinopropane-1-sulfonic acid (MOPS) and 2-amino-2-(hydroxymethyl)propane-1,3-diol (TRIS) were purchased from Roth (D-Karlsruhe). All lipo-oligonucleotides **10**–**15** (Figure 2) were synthesized, purified and characterized by MALDI–TOF mass spectrometry by Eurogentec SA (B-Liege). In each case, the detected mass confirmed the corresponding calculated mass. MALDI–TOF–MS (*m*/*z*): 4636.8 (**10**, $[M + H]^+$; calcd 4637.8); 4976.3 (**11**, $[M + H]^+$; calcd 4974.4); 5043.4 (**12**, $[M + H]^+$; calcd 5054.4); 4755.0 (**13**, $[M + H]^+$; calcd 4767.8); 4866.5 (**14**, $[M + H]^+$; calcd 4875.9); 4696.9 (**15**, $[M + H]^+$; calcd 4707.6).

Similar to that described in [9,18], we have chosen the particular sequence 5'-d(TAG GTC AAT ACT)-3', because it does not form hair pins or self-complementary duplexes. The lipophilized strand and the complementary oligomer contain the same number of canonical nucleotides, so that a mixture of the 1 $A_{260}$ units, from both strands, gives an equimolar mixture. For nearly all base pairs, amongst all nearest neighbor combinations exist. Both strands have the same composition; the unmodified duplex exhibits a melting point of 46 °C in a phosphate buffered solution.

**Methods.** The technique of horizontal lipid bilayer fabrication and incorporation of lipo-oligonucleotides is described therein. Analogous to the description in [7,8,41], in the following, the automated bilayer fabrication is described. Horizontal bilayers were automatically fabricated using a lipid mixture of POPE/POPC 8:2 (w/w), 10 mg/mL of *n*-decane within the "bilayer slides" [7,8,41] and observed with an add-on for the inverted confocal microscope (bilayer slides and Ionovation Explorer, Ionovation GmbH, D-Osnabrück). A bent Hamilton syringe (CH-Bonaduz) was used to fill the bilayer slide with 250 mM KCl, 10 mM MOPS/Tris, pH 7 buffer and 0.2 µL of the lipid mixture. The automated bilayer production was started and monitored optically and electrochemically. For the electrochemical measurements, Ag/AgCl electrodes embedded in agarose were used to connect the cis and the trans compartments. As described in previous works [7,8,15,18,41], the measurement protocol is as follows: When a bilayer has been established (C > 50 pF ± 20), a solution of the cyanine-labeled lipophilized oligonucleotide (5 µL, 50 nM, solubilized in water) was injected over the bilayer. The probe was incubated until a stable brightness was reached. During the incubation time, the bilayer integrity was monitored by continuous capacitance measurements. In all graphs, the brightness was given either in kHz or it was normalized (0–100%). The z-direction scan was performed in 5 min intervals (see below). After the incubation steps, the cis compartment was perfused repeatedly for 30 s (1.1 mL/min, each), and the bilayer was inspected by confocal fluorescence microscopy. Each perfusion step using the Ionovation perfusion unit was followed by an incubation step. This procedure was repeated until the brightness of the bilayer stabilized.

All confocal imaging and lifetime measurements were performed on a confocal laser scanning microscope (Insight Cell 3D, Evotec Technologies GmbH, Hamburg, Germany) equipped with a 635 nm emitting laser diode (LDH-P-635, PicoQuant, Berlin, Germany), a 40× water immersion objective (UApo 340, 40×, NA = 1.15, Olympus, Tokyo, Japan), and an Avalanche photodiode detector (SPCM-AQR-13-FC, Perkin-Elmer Optoelectronics, Fremont, CA, USA). Fluorescence irradiation was achieved with an excitation laser power of 200 ± 20 µW directly in front of the objective. Z-direction scans were performed by scanning the confocal laser spot in the *X*–*Y* direction with a rotating beam scanner and movement of the objective in the *Z* direction. The movement in all directions was piezo-controlled, which allows for nanometer precision positioning.

As in [7], each measurement protocol was carried out as follows: (a) a reference scan of the stable bilayer was performed, (b) the lipophilized oligonucleotide sample was added, (c) the solution was incubated until the brightness in the bilayer reached a maximum, (d) followed by an additional series of scans performed after each perfusion of the cis compartment (for 30 s, 1.1 mL/min), and (e) followed by a given incubation time.

In order to analyze the 2D images they were processed with Image J A 1.44 software and the obtained data was evaluated with OriginPro 8 (OriginLab Corporation). For an evaluation of the z-direction scan data, the brightness within the bilayer was measured. The mean area of the bilayer cross section was 400 ± 115 counts per pixel ($N$ = 230). Next, the brightness was determined by summing up the number of pixels, where one pixel equals 1 kHz or 1 µm.

All devices used as well as the general techniques used in this paper have been described in detail in preceding manuscripts as well as in an mp4 video [18].

## References


1. Roskowski, R., Jr. *Biochem. Biophys. Res. Commun.* **2003,** *303,* 1–7. doi:10.1016/S0006-291X(03)00323-1
2. Dumelin, C. E.; Chen, Y.; Leconte, A. M.; Chen, Y. G.; Liu, D. R. *Nat. Chem. Biol.* **2012,** *8,* 913–919. doi:10.1038/nchembio.1070
3. Pokholenko, O.; Gissot, A.; Vialet, B.; Bathany, K.; Thiéry, A.; Barthélémy, P. *J. Mater. Chem. B* **2013,** *1,* 5329–5334. doi:10.1039/c3tb20357c
4. Gissot, A.; Camplo, M.; Grinstaff, M. W.; Barthélémy, P. *Org. Biomol. Chem.* **2008,** *6,* 1324–1333. doi:10.1039/b719280k
5. Schade, M.; Berti, D.; Huster, D.; Herrmann, A.; Arbuzova, A. *Adv. Colloid Interface Sci.* **2014,** *208,* 235–251. doi:10.1016/j.cis.2014.02.019
6. Gangadhara, K. L.; Srivastava, P.; Rozenski, J.; Mattelaer, H.-P.; Leen, V.; Dehaen, W.; Hofkens, J.; Lescrinier, E.; Herdewijn, P. *J. Syst. Chem.* **2014,** *5,* No. 5. doi:10.1186/s13322-014-0005-3
7. Werz, E.; Rosemeyer, H. *Beilstein J. Org. Chem.* **2014,** *10,* 2307–2321. doi:10.3762/bjoc.10.240
8. Werz, E.; Korneev, S.; Montilla-Martinez, M.; Wagner, R.; Hemmler, R.; Walter, C.; Eisfeld, J.; Gall, K.; Rosemeyer, H. *Chem. Biodiversity* **2012,** *9,* 272–281. doi:10.1002/cbdv.201100298
9. Seela, F.; Münster, I.; Lüchner, U.; Rosemeyer, H. *Helv. Chim. Acta* **1998,** *81,* 1139–1155. doi:10.1002/hlca.19980810527
10. Peer, D.; Lieberman, J. *Gene Ther.* **2011,** *18,* 1127–1133. doi:10.1038/gt.2011.56

11. Wolfrum, C.; Shi, S.; Jayaprakash, K. N.; Jayaraman, M.; Wang, G.; Pandey, R. K.; Rajeev, K. G.; Nakayama, T.; Charrise, K.; Ndungo, E. M.; Zimmermann, T.; Koteliansky, V.; Manoharan, M.; Stoffel, M. *Nat. Biotechnol.* **2007,** *25,* 1149–1157. doi:10.1038/nbt1339
12. Kurz, A.; Bunge, A.; Windeck, A.-K.; Rost, M.; Flasche, W.; Arbuzova, A.; Strohbach, D.; Müller, S.; Liebscher, J.; Huster, D.; Herrmann, A. *Angew. Chem., Int. Ed.* **2006,** *45,* 4440–4444. doi:10.1002/anie.200600822
13. Werz, E.; Viere, R.; Gassmann, G.; Korneev, S.; Malecki, E.; Rosemeyer, H. *Helv. Chim. Acta* **2013,** *96,* 872–888. doi:10.1002/hlca.201200573
14. Korneev, S.; Rosemeyer, H. *Helv. Chim. Acta* **2013,** *96,* 201–216. doi:10.1002/hlca.201100410
15. Malecki, E.; Ottenhaus, V.; Werz, E.; Knies, C.; Montilla-Martinez, M.; Rosemeyer, H. *Chem. Biodiversity* **2014,** *11,* 217–232. doi:10.1002/cbdv.201300127
16. Malecki, E.; Knies, C.; Werz, E.; Rosemeyer, H. *Chem. Biodiversity* **2013,** *10,* 2209–2220. doi:10.1002/cbdv.201300107
17. Malecki, E.; Rosemeyer, H. *Helv. Chim. Acta* **2010,** *93,* 1500–1512. doi:10.1002/hlca.201000121
18. Zentrum virtUOS für Informationsmanagment und virtuelle Lehre, Universität Osnabrück, 2014, Movie “Directors, Incorporation of Lipophilic Single Strands with Different Nucleolipid Head Groups into an Artificial Lipid Bilayer.” https://repositorium.uni-osnabrueck.de/handle/urn:nbn:de:gbv:700-2014091012807
19. Hanwell, M.; Curtis, D. E.; Lonie, D. C.; Vandermeersch, T.; Zurek, E.; Hutchison, G. R. *J. Cheminf.* **2012,** *4,* No. 17. doi:10.1186/1758-2946-4-17
20. Berg, H. C. Diffusion with Drift. *Random Walks in Biology;* Princeton University Press: Prinston, 1983; pp 48–62.
21. Doose, S.; Barsch, H.; Sauer, M. *Biophys. J.* **2007,** *93,* 1224–1234. doi:10.1529/biophysj.107.107342
22. Bollmann, S.; Löllmann, M.; Sauer, M.; Doose, S. *Phys. Chem. Chem. Phys.* **2011,** *13,* 12874–12882. doi:10.1039/c1cp21111k
23. Fernandes, M.; Ortega, A.; López Martínez, M.; García de la Torre, J. *Nucleic Acids Res.* **2002,** *30,* 1782–1788. doi:10.1093/nar/30.8.1782
24. Wolf, B.; Hanlon, S. *Biochemistry* **1975,** *14,* 1661–1670. doi:10.1021/bi00679a018
25. Falk, M.; Hartman, K. A.; Lord, R. C. *J. Am. Chem. Soc.* **1963,** *85,* 387–391. doi:10.1021/ja00887a004
26. Bloomfield, V. A.; Crothers, D. M.; Tinoco, I., Jr. Interactions of Nucleic Acids with Water and Ions. *Nucleic Acids. Structures, Properties and Functions;* University Science Books: Sausalito, CA, 2000; pp 475–534.
27. Carlsson, N.; Jonsson, F.; Wilhelmsson, L. M.; Nordén, B.; Åkerman, B. *Soft Matter* **2013,** *9,* 7951–7959. doi:10.1039/c3sm50982f
28. Marmé, N.; Knemeyer, J.-P.; Sauer, M.; Wolfrum, J. *Bioconjugate Chem.* **2003,** *14,* 1133–1139. doi:10.1021/bc0341324
29. Klehs, K.; Spahn, C.; Endesfelder, U.; Lee, S. F.; Fürstenberg, A.; Heilemann, M. *ChemPhysChem* **2014,** *15,* 637–641. doi:10.1002/cphc.201300874
30. Widengren, J.; Schwille, P. *J. Phys. Chem. A* **2000,** *104,* 6416–6428. doi:10.1021/jp000059s
31. Zhao, M.; Huang, R.; Peng, L. *Opt. Express* **2012,** *20,* 26806–26827. doi:10.1364/OE.20.026806
32. Ray, K.; Zhang, J.; Lakowicz, J. R. *Anal. Chem.* **2008,** *80,* 7313–7318. doi:10.1021/ac8009356
33. Langecker, M.; Arnaut, V.; List, J.; Simmel, F. C. *Acc. Chem. Res.* **2014,** *47,* 1807–1815. doi:10.1021/ar500051r
34. Taib, N.; Aimé, A.; Houmadi, S.; Castano, S.; Barthélémy, P.; Laguerre, M.; Bestel, I. *Langmuir* **2012,** *28,* 7452–7460. doi:10.1021/la300744x
35. Loew, M.; Springer, R.; Scolari, S.; Altenbrunn, F.; Seitz, O.; Liebscher, J.; Huster, D.; Herrmann, A.; Arbuzova, A. *J. Am. Chem. Soc.* **2010,** *132,* 16066–16072. doi:10.1021/ja105714r
36. Johnson-Buck, A.; Jiang, S.; Yan, H.; Walter, N. G. *ACS Nano* **2014,** *8,* 5641–5649. doi:10.1021/nn500108k
37. Montis, C.; Baglioni, P.; Berti, D. *Soft Matter* **2014,** *10,* 39–43. doi:10.1039/c3sm52254g
38. Bunge, A.; Loew, M.; Pescador, P.; Arbuzova, A.; Brodersen, N.; Kang, J.; Dähne, L.; Liebscher, J.; Herrmann, A.; Stengel, G.; Huster, D. *J. Phys. Chem. B* **2009,** *113,* 16425–16434. doi:10.1021/jp9067747
39. Rosemeyer, H. Nucleolipids and use thereof, and devices for nucleic acid analysis. U.S. Patent 7,914,991 B2, March 29, 2011.
40. Seelig, A.; Gottschlich, R.; Devant, R. M. *Proc. Natl. Acad. Sci. U. S. A.* **1994,** *91,* 68–72. doi:10.1073/pnas.91.1.68
41. Köstler, K.; Werz, E.; Malecki, E.; Montilla-Martinez, M.; Rosemeyer, H. *Chem. Biodiversity* **2013,** *10,* 39–61. doi:10.1002/cbdv.201100338


## License and Terms