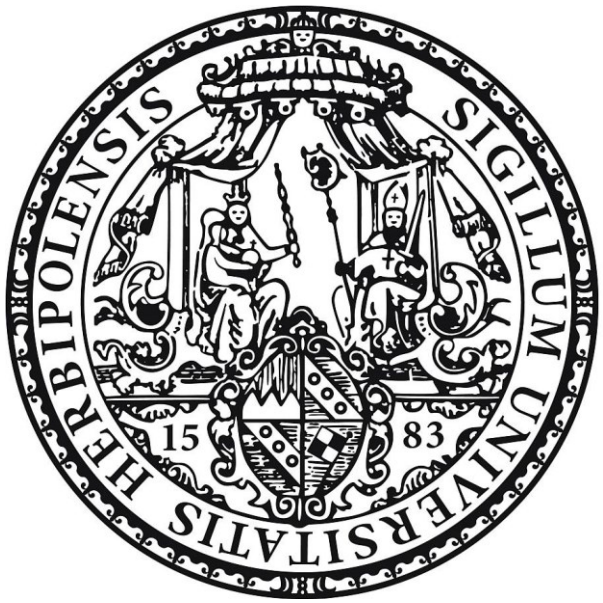
Hydrophobic recognition motifs in functionalized DNA



Dissertation zur Erlangung des naturwissenschaftlichen Doktorgrades der

Julius-Maximilians-Universität Würzburg

Vorgelegt von

Hermann Neitz

aus Göttingen

Würzburg, 2023

Eingereicht bei der Fakultät für Chemie und Pharmazie am

24.10.2023

Gutachter der schriftlichen Arbeit

1. Gutachter: Prof. Dr. Claudia Höbartner
2. Gutachter: Prof. Dr. Ann-Christin Pöppler

Prüfer des öffentlichen Promotionskolloquiums

1. Prüfer: Prof. Dr. Claudia Höbartner
2. Prüfer: Prof. Dr. Ann-Christin Pöppler
3. Prüfer: Prof. Dr. Holger Helten
4. Prüfer: Prof. Dr. Frank Würthner
5. Prüfer: Prof. Dr. Caroline Kisker

Datum des öffentlichen Promotionskolloquiums

19.12.2023

Doktorurkunde ausgehändigt am

_____

Acknowledgements

Für ihre ständige Hilfsbereitschaft und unermüdliche Betreuung möchte ich mich bei Prof. Dr. Claudia Höbartner bedanken. Durch sie hatte ich die Möglichkeit, verschiedene Themenbereiche in der Welt der biologischen Chemie kennen zu lernen und so mein Wissen zu erweitern. Dabei hat sich mein Repertoire zur Lösung wissenschaftlicher Fragestellungen ständig weiterentwickelt.

Bei Prof. Dr. Ann-Christin Pöppler möchte ich mich für die Übernahme des Zweitgutachtens bedanken.

An dieser Stelle möchte ich mich auch bei Dr. Jochen Kuper und Prof. Dr. Caroline Kisker für die erfolgreiche Zusammenarbeit bedanken. Es hat mich immer sehr gefreut, wenn ein von mir entwickeltes System in anderen Forschungsbereichen eingesetzt werden konnte und anderen Projekten geholfen hat.

Von entscheidender Bedeutung während meiner Promotion war neben der Zusammenarbeit mit anderen Arbeitskreisen auch die Zusammenarbeit innerhalb der eigenen Gruppe. In diesem Zusammenhang möchte ich mich bei Dr. Irene Bessi für die Messung meiner NMR-Proben bedanken. Die gemeinsame Auswertung der Spektren hat mir viel Spaß gemacht und wieder einmal gezeigt, welch großartige Methode die NMR-Spektroskopie ist. Ich möchte mich bei Manuela Michel für die Unterstützung bei den verschiedenen Synthesen bedanken. Die gemeinsame Arbeit als eingespieltes Team hat mir das Leben sehr erleichtert.

Für die Unterstützung bei der MS- und NMR-Analytik danke ich Dr. Matthias Grüne, Dr. Juliane Adelmann, Sebastian Mayer, Michaela Schraut, Stefanie Schmitt und Patricia Altenberger.

Ich möchte mich bei meinen Masterstudenten, Bachelorstudenten und Praktikanten Valentin Kachler, Justus Müller, Ricarda Preiss, Kersten Ulrich und Jonas Hofmann für ihren Beitrag bedanken. Es war immer eine Freude, so engagierte Studenten zu haben und ihnen neue Dinge zu zeigen.

Wichtig im Laboralltag sind die Menschen, die im Hintergrund den Laden am Laufen halten. Neben Manuela möchte ich hier meine Lyo-Partnerin Ann-Kathrin Lenz nennen. Danke für die unzähligen Puffer. Das gilt auch für die Auszubildenen, die mitgeholfen haben.

Ein wichtiger Teil meines Lebens in den letzten Jahren (vor allem während der Corona-Pandemie) waren die Mitglieder des AK Höbartner. Danke an alle für die ständige Unterstützung und Hilfsbereitschaft! Mein besonderer Dank gilt Julia und Carina, meinen Platznachbarn. Für die lustigen Unternehmungen zusammen mit Florian möchte ich mich auch bedanken. Es war ein einsames Jahr ohne euch im Büro. Ich danke Taku für seine ständige Bereitschaft, über alle möglichen Themen zu diskutieren und Irene für das Zuhören, wenn ich frustriert war über verschiedene Dinge. Für die nicht-wissenschaftlichen Diskussionen im letzten Jahr war Konstantin für mich da. Ich danke Caro für ihre Unterstützung bei der Dissertation. Außerdem danke ich Dr. Doris Feineis und Lisa Weidner für ihre organisatorische Hilfe.

Für die Korrekturen danke ich Julia, Hauke, Irene, Henrik und Taku.

Freunde und Familie waren mir während meines gesamten Studiums wichtige Begleiter und Unterstützer. Danke, dass es euch gibt!

List of Publications

Parts of this thesis have been published:

Neitz, H.; Bessi, I.; Kachler, V.; Michel, M.; Höbartner, C. Tailored Tolane-Perfluorotolane Assembly as Supramolecular Base Pair Replacement in DNA. *Angew. Chem., Int. Ed.,* **2023**, *62*, e202214456.

Neitz, H.; Bessi, I.; Kuper, J.; Kisker, C.; Höbartner, C. Programmable DNA Interstrand Crosslinking by Alkene–Alkyne [2 + 2] Photocycloaddition. *J. Am. Chem. Soc.* **2023**, *145*, 9428–9433.

Neitz, H.; Höbartner, C. A tolane-modified 5-ethynyluridine as a universal and fluorogenic photochemical DNA crosslinker. *Chem. Commun.* **2023**, *59*, 12003–12006.

Table of Contents

1 Introduction

As a carrier of genetic information, deoxyribonucleic acid (DNA) is one of the most important biomolecules in living organisms and is characterized by its simple sequential structure and high stability. The research carried out on nucleic acids has made it possible to extend the artificial functions of DNA, which are based on structures that are different from those of natural B-DNA. However, the known structure and programmability of the B-DNA double helix has also led to breakthroughs in DNA-based nanotechnology and laid the foundation for new structures and dynamic systems based on the DNA double helix. By the introduction of chemical modifications, the functionality of the DNA can be extended and lead to new structures, reactivities or spectroscopic properties.

As a foundation for all following chapters of this thesis, this introduction will cover the fundamentals of the DNA duplex with special attention to interactions that stabilize the duplex structure. Additionally, alternative DNA structures will be discussed using additional features of DNA like the Hoogsteen base pairing.

Parts of the introduced alternative structures are required for functional DNA. For a deeper understanding of nucleic acid chemistry, examples of functional DNA will be presented. With their complex and often unpredictable structures, aptamers and deoxyribozymes take on different functions and are among the highlights in the field of functional DNA and will be discussed here. In contrast, DNA nanotechnology follows well-defined rules to create larger functional structures. Examples of structural and dynamic DNA nanotechnology are given to illustrate the possibilities of DNA programmability. In particular, DNA origami will be discussed as one of the greatest achievements in the field in the recent years. These examples of functional DNA provide the basis for further applications in this thesis and future applications derived by synthetically modified DNA.

Synthetic chemistry further expands the functionality of DNA. By introducing chemical modifications, new supramolecular DNA assemblies are created allowing DNA to adopt new structures. To showcase the great potential of the introduction of synthetic modifications into

natural DNA, different approaches are presented here that can be used to generate supramolecular structures. Afterwards artificial nucleosides containing modified backbones or nucleobases are described, which enables novel properties like recognition, reactivity, or spectroscopic properties. Particular emphasis is put on the development of unnatural base pairs as a fundamental part of this thesis.

1.1 Nucleic acid structures and functions

1.1.1 DNA duplex

DNA consists of four different nucleotides. Each nucleotide consists of a phosphate, a sugar, and a nucleobase. All nucleotides in DNA contain deoxyribose sugar, but their nucleobases differ. The nucleobases can be divided into the purine bases adenine (A) and guanine (G) and the pyrimidine bases cytosine (C) and thymine (T). Based on previous findings and results in the scientific community, Watson and Crick postulated a three-dimensional structure of a double helix in 1953.[1] A closer look at the interactions that drive this assembly reveals a prime example of supramolecular chemistry.[2-3] The stabilization of the DNA double helix is driven by a cooperative interplay of hydrogen bonding, base stacking and electrostatic interactions. Precise recognition of the two strands within the double-stranded DNA molecule is primarily due to hydrogen bonding. The two Watson-Crick base pairs, A/T and G/C, enable recognition of both strands, while base stacking stabilizes the structure of the DNA duplex (see Figure 1.1).[4]

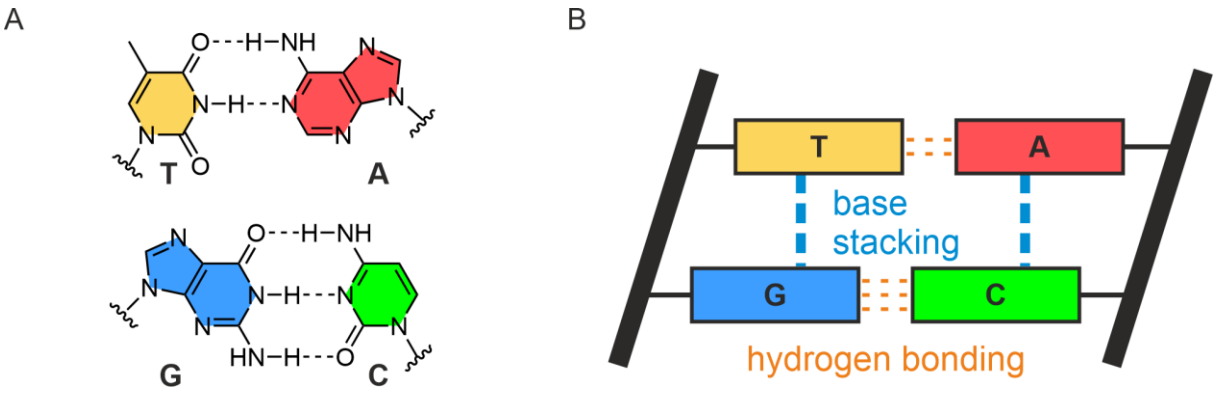


Figure 1.1 (A) Watson-Crick base pairs in DNA. (B) Illustration of interstrand (hydrogen bonding) and intrastrand (base stacking) stabilizing forces inside the DNA duplex.

Free interaction enthalpies of −14.6 kcal mol^{-1} for A/T and −27.5 kcal mol^{-1} for G/C were calculated for the Watson-Crick base pairs.[5] However, nucleobases are surrounded by water molecules, which leads to the formation of hydrogen bonds between the water molecules and

the Watson-Crick base pairing side of the nucleobase. Therefore, desolvation energy must be supplied before hydrogen bonding can occur between two nucleobases.[6] One way to experimentally estimate the contribution of hydrogen bonding is to substitute an A/T Watson-Crick base pair with nonpolar base surrogates such as difluorotoluene and 4-methylbenzimidazole. Incorporation of these building blocks causes destabilization of 3.2 kcal mol^{-1}, suggesting a hydrogen bond free energy of −1.6 kcal mol^{-1}.[7] The constrained structural flexibility of the nucleobases is further associated with an entropic penalty, which is reversed by the exclusion of water molecules during base pair formation.[4]

While hydrogen bonding is characterized by its programmable nature, base stacking is the more critical factor for the overall structural stabilization of the DNA double helix. Base stacking contains (i) dispersive, (ii) electrostatic and (iii) solvent effects as driving forces.[4, 8]

Nucleobases are aromatic systems, which means they are planar molecules with delocalized π-electrons, resulting in high polarizabilities.[9] The stacked conformation of the nucleobases results from the interaction of their induced dipoles, which allows the flat surfaces of the nucleobases to stack on top of each other to maximize van der Waals contacts. The importance of dispersive interaction is shown by hydrophobic base pair surrogates, which lack hydrogen bonding properties, but can even increase the thermal stability of the DNA duplex.[10]

Electrostatic effects play a role in the arrangement of the DNA duplex, in addition to dispersive effects. As an asymmetric heteroaromatic system, nucleobases possess quadrupole and dipole moments that allow for electrostatic interactions to occur. These are important for positioning stacked nucleobases relative to each other, although dispersion energies are the main reason for base stacking.[9]

In addition to the electrostatic effects, the solvent contributes to the stabilization of the DNA structure. While the sugar phosphate backbone is well soluble in polar solvents such as water, the aromatic bases do not prefer solvation with water. Due to the hydrophobic effect, stacking of nucleobases is preferred to exclude water molecules. The addition of salts to the buffer system further increases this stabilization and reduces the repulsive forces of the negatively charged backbones.[8, 11-12]

The interaction of these forces forms the structure of the DNA duplex to act as a carrier of genetic information.

1.1.2 Alternative structures of the DNA

The right-handed B-DNA, described by Watson and Crick, is the most common secondary structure adopted by the DNA. A DNA helix can also adopt other helical structures like right-handed A-DNA or left-handed Z-DNA, depending on its sequence, hydration state or salt concentration.[13-15] Along with the double helix, alternative structures that use Watson-Crick base pairing, such as hairpins or Holliday junctions, are possible. The use of other base pairing schemes, such as Hoogsteen base pairs, allows DNA strands to adopt other structural conformations. The best-known non-canonical structures are triplexes (H-DNA) and tetraplexes, such as G-quadruplexes or i-motifs.[16]

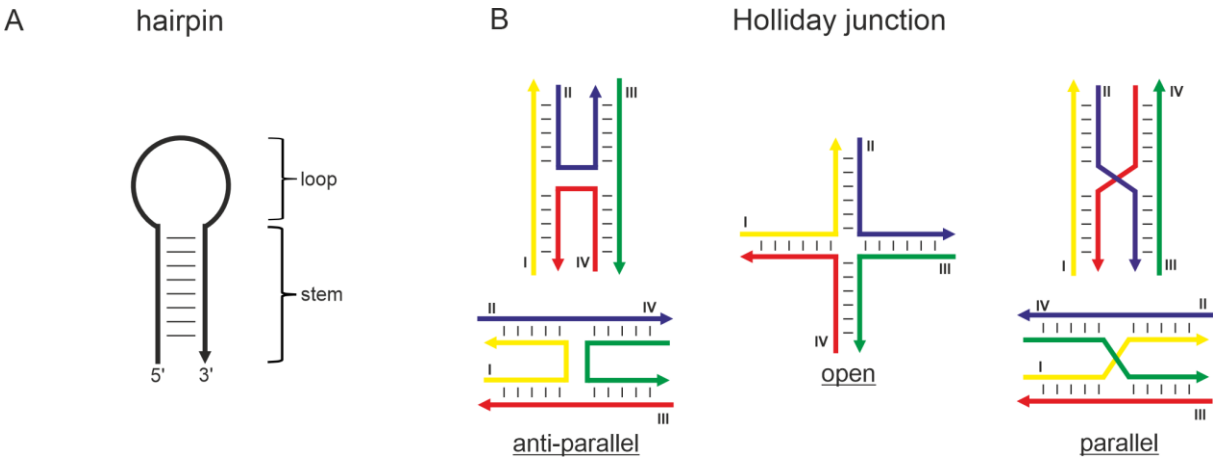


Figure 1.2 Different structures derived from Watson-Crick base pairing. (A) Structure of a DNA-hairpin and (B) different conformation which can be adopted by an immobilized Holliday junction.[17]

Hairpins are single-stranded and consist of an intramolecular base paired stem connected by a single-stranded loop (see Figure 1.2A). However, there are several aspects to consider for stability and folding. Hiraos group were able to create DNA mini hairpins containing only two G/C base pairs in their stem and a small loop with a G-N-A sequence, where N is one of the canonical nucleobases. Remarkable melting temperatures of around 70 °C were achieved. Depending on the arrangement of the G/C base pairs in the stem and the sequence in the loop, they found differences in the thermal stability.[18-19] Further studies suggest that high stability is achieved through the interaction between the loop integrated G/A pair and the neighboring C/G base pair.[20] Steady-state measurements using a fluorophore and a quencher on the 5'- and 3'-end of a hairpin assumed that the initial folding of a hairpin intermediate occurs on the microsecond time scale. However, subsequent experiments showed that the correct folding is established within milliseconds.[21]

Another structural motif derived from Watson-Crick base pairing is a four-way DNA junction, also known as a Holliday junction. It is a key intermediate in the repair of a double-strand break by homologous recombination.[22] The use of two identical double-strands during repair allows the branch point of the Holliday junction to migrate. The asymmetric design of the sequence immobilizes the branch point, leading to different conformations of the Holliday junction.[23] Coaxial stacking of the two helical domains occurs in the presence of magnesium, which allows the helical domains to adopt a parallel or anti-parallel conformation. The anti-parallel conformation is energetically favored. However, which of the two possible conformations is formed depends on the sequence. The parallel and anti-parallel conformation are mediated by an open structure, which is also formed in the absence of ions (see Figure 1.2B).[17, 24]

The addition of a triplex-forming oligonucleotide (TFO) enables the formation of a DNA triplex, also called H-DNA. Hoogsteen base pairs within the major groove of the DNA duplex drive the association of the TFO. Due to the nature of Hoogsteen base pairing, triplex formation can only occur on oligopurine sequences. TFOs can bind in a parallel orientation under acidic conditions and in presence of stabilizing cations to their counter strands, forming TxAT, C^+xGC or GxGC base pairs or in an anti-parallel orientation forming GxGC, AxAT or TxAT (see Figure 1.3A).[25] Compared to the DNA duplex, H-DNA is less stable and shows a pH and salt concentration dependence.[26]

G- or C-rich DNA strands can form tetraplex structures called G-quadruplexes or i-motifs (see Figure 1.3B-C). Hoogsteen base pairs of four guanine nucleobases can form planar G-tetrads, in presence of monovalent or divalent cations that stabilize the structure. The stacking of multiple G-tetrads forms the G-quadruplex, which is present in various structures and topologies. G-quadruplexes can be classified as intrastrand or interstrand and parallel or anti-parallel, depending on the number and orientation of the strands involved in G-quadruplex formation.[16, 27] C-C^+ base pairs in a C-rich DNA strand lead to the formation of an i-motif under slightly acidic conditions. It consists of two pairs of parallel oriented strands. These double-strands are arranged anti-parallel to each other. In analogy to G-quadruplexes, these structures can be distinguished between intrastrand or interstrand. The two different topologies are 3'E and 5'E, depending on the side of the outer C-C^+ base pair. Even though a protonated cytosine is required, i-motifs can be found under neutral conditions. [28-29]

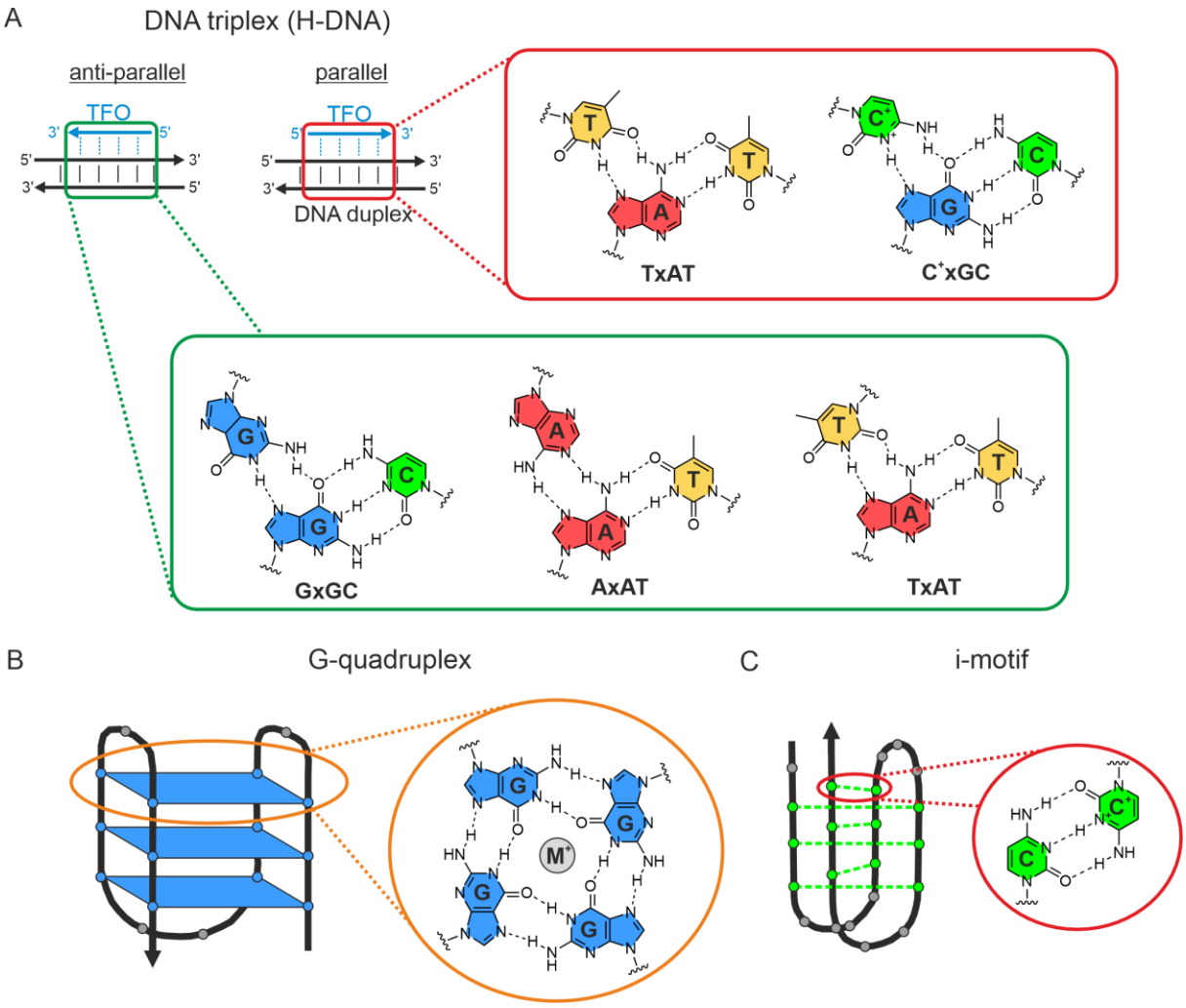


Figure 1.3 Examples that use alternative base pairing. (A) DNA triplex using additional Hoogsteen base pairing. Tetraplexes can be formed by (B) G-quadruplex and (C) i-motif. The example shown is one possible topology of these structures.

These alternative structures show that DNA can occur in a wide variety of different architectures in addition to the B-DNA duplex. Importantly, these structures have an *in vivo* function. Repetitive sequences can be found in DNA that adopt these alternative structures. They can block helicase activity, recruit enzymes that regulate genetic code readout, or initiate DNA repair.[30]

1.1.3 Alternative functions of DNA

DNA is also a good candidate for artificial functions because of its huge repertoire of alternative structures. The potential of DNA strands to fold into complex 3D structures allows them to perform additional tasks, such as binding small molecules or enabling chemical reactions.[31] The *in vitro* evolution of nucleic acids is a key to find sequences that perform the desired

function (see Figure 1.4). Nucleic acids are privileged because methods such as the polymerase chain reaction allow them to be replicated without altering their genotype. In 1990, Szostak and Gold independently showed how functional sequences could be selected from a random library of RNA sequences.[32-33] A functional RNA was obtained through a cycle of selection and amplification. The method for obtaining nucleic acid oligomers binding to ligands is also known as SELEX (systematic evolution of ligands by exponential enrichment).

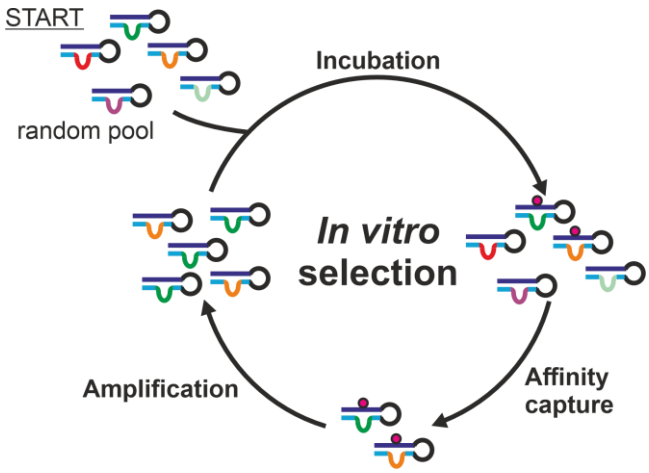


Figure 1.4 Illustration of the *in vitro* selection concept. An incubation step is carried out under preferred binding or reaction conditions using a random pool of sequences. The active sequences (introduced modification shown as magenta circle) are captured and then amplified. The process can be repeated to obtain the desired aptamer or DNAzyme.

Under the assumption that the capability of folding into more complex structures would facilitate ligand binding, most selections were made for RNA.[30, 34] As discussed in the previous section, DNA can also adopt complex tertiary structures and has other advantages over RNA, such as higher chemical and biological stability. On this basis, examples of functional DNA sequences obtained by selection are discussed in the following.

1.1.3.1 DNA aptamers

Aptamers are short single-stranded nucleic acids that bind small molecules, proteins, or even cells. With binding affinities in the pico- to micromolar range, they are comparable to antibodies.[35] The production of DNA aptamers can be easily upscaled with acceptable prices due to the possibility of chemical synthesis.[36] Higher stability, small size and biocompatibility are further advantages compared to protein antibodies.[37] A thrombin-binding aptamer and a dye-binding aptamer were the first examples of aptamer selections in 1992.[38-39] NMR structure solution showed that the thrombin-binding aptamer has an anti-parallel G-quadruplex structure, but it is well known that aptamers can adopt other complex tertiary structures.[40-41]

Aptamers can be used for therapeutic and sensing purposes. One possible application is shown for the detection of SARS-CoV-2 infections. Using aptamers that bind to the SARS-CoV-2 spike protein or receptor binding domain, detection methods could be developed. [35, 42] Of particular interest are aptamers with fluorogenic properties. Light-up aptamers, which demonstrate fluorogenic properties only after binding to a ligand, are of particular interest because of their low background emission when used. Using chromophores derived from the green fluorescent protein (GFP) fluorophore 4-hydroxybenzylidene imidazolinone, a series of different GFP-mimicking light-up aptamers have been obtained.[43] Binding reduces the rotational freedom of the aromatic ring systems, and the chromophore becomes fluorescent. Jaffrey and coworkers developed the fluorogenic DNA aptamer Lettuce, which could be used for RNA detection.[44] X-ray crystallography and cryogenic electron microscopy revealed that the Lettuce aptamer forms a four-way junction to bind the chromophore, demonstrating that DNA can adopt complex structures like RNA.[45]

1.1.3.2 DNAzymes

In addition to ligand binding activity, DNA strands can catalyze reactions. To find catalytically active sequences, so-called DNAzymes or deoxyribozymes, the selection method has to be adapted. Therefore, most of the known DNA enzymes have simple strand cleavage or ligation activity. Active variants can be selected more easily because they have a different size after amplification compared to their inactive variants. No active natural DNA enzyme has been discovered yet.[46] The first catalytically active DNA sequence was discovered by Joyce. It cleaves specifically one RNA strand in the presence of Pb^{2+}.[47] Moreover, a DNAzyme has accomplished DNA cleavage, which is more difficult because of the absence of a 2'-hydroxyl group capable of attacking the bridging phosphate.[48] The high specificity for substrates or cofactors of DNAzymes, like metals, has led to a wide range of applications, such as biosensors, logic gates or even mRNA-targeted gene silencing.[49] Biosensors mostly use a fluorophore quencher system that becomes fluorescent when the strand is cleaved, depending on the presence of the essential cofactor. Other readout like electrochemical signals are also possible.[49] Recently DNAzymes were used to detect RNA modifications that play an essential role in the epitranscriptomics.[50-51] It was even possible to discriminate between different methylated cytidine isomers, demonstrating how sensitive DNAzymes can be selected.[52]

DNAzymes can also ligate RNA strands. One example was the deoxyribozyme 9DB1 selected in 2005, which can ligate two RNA strands using a triphosphate at the 5' end of one of the RNA strands.[53] It was the first DNAzyme whose catalytically relevant structure was solved by X-ray crystallography (see Figure 1.5A).[54] Tertiary interactions stabilize the folding of DNA, resulting in a compact three-dimensional structure. For example, two thymidines of the 9DB1 interact with the nucleotides A and G of the two RNAs at the ligation junction. Using X-ray crystallography and NMR spectroscopy, only two more structures of the RNA-cleaving deoxyribozymes 8-17 DNAzyme[55] and 10-23 DNAzyme[56] could be solved, supporting the necessity of tertiary interactions that allow DNA to assume the correct folding to act as a catalyst (see Figure 1.5).

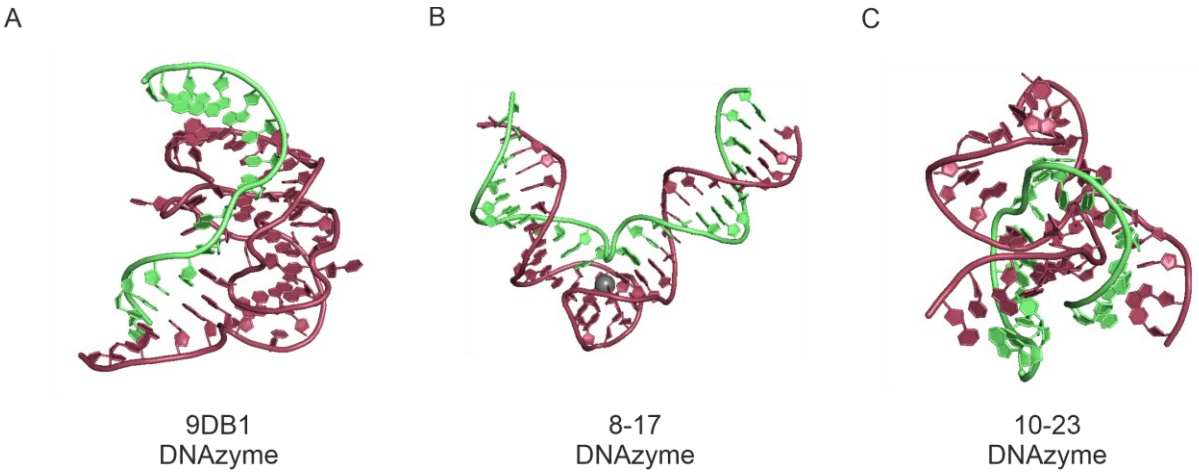


Figure 1.5 The structures of the DNAzymes (A) 9DB1 (PDB: 5CKK), (B) 8-17 and (PDB: 5XM8), and (C) 10-23 (PDB: 7PDU) show the complex structure that DNA must adopt to act as a catalyst. DNAzymes are presented in red, and DNA/RNA substrates are shown in green.

In addition to cleavage and ligation of DNA/RNA strands, artificial reactions using DNAzymes are also possible. Tyr-RNA nucleopeptide linkage,[57]amide hydrolysis[58] or even thymine dimer photo reversion using photolyase activity[59] are some examples. Based on the known ability of deoxyribozymes to catalyze phosphorylation reactions,[60] The Curtis lab selected for sequences that phosphorylate their 5' end with the 1,2-dioxetane substrate CDP-Star. Dephosphorylation of CDP-Star results in a chemically initiated electron-exchange luminescence reaction. The deoxyribozyme Supernova was discovered to produce light, with a 6500-fold increase in brightness.[61]

1.2 DNA nanotechnology

The examples of aptamers and deoxyribozymes show the power behind a simple biopolymer that can adopt different three-dimensional structures. However, these structures are difficult to predict and can only be solved with the help of laborious three-dimensional structural analysis. On the other hand, the B-DNA is well understood. The formation of B-DNA duplexes follows well-defined rules of base pairing, making it predictable and programmable. In addition, there is an enormous number of possible DNA sequences. Seeman realized the advantage of these programmable interactions to create nanostructures based on DNA. He assumed that single-stranded overhangs, called sticky ends, on a Holliday junction could create nanostructures in a bottom-up approach.[62] Based on this idea, he developed double crossover (DX) molecules as basis for 2D-DNA crystals.[63] The design of two parallel double helices connected by two strand exchanges provides the necessary rigidity and stability for 2D-DNA crystal formation.[64] Controlled interaction of the DX molecules is achieved by adding sticky ends (see Figure 1.6A). The basic design of the DX molecule has been expanded to allow for the creation of different types of DNA crystals. Triple crossover (TX) molecules were introduced, where the design of three helices connected by two crossovers allowed the formation of a flat lattice structure that could be programmed for structure formation.[65-66] Introduction of additional thiol groups even allows the formation of DNA nanotubes (see Figure 1.6B).[67] The lattice shape could be changed by using different building blocks, such as Y-shaped DNA motifs, which form a hexagonal lattice (see Figure 1.6C) or a 6x4 DNA tile which forms a triangular lattice.[68-69] A square lattice was realized with 4x4 DNA tiles (see Figure 1.6D). By placing biotin molecules in the center of the tiles, it was possible to create an array of streptavidin proteins, which was Seeman's vision in the early days of DNA nanotechnology.[70-71] A fluorescence resonance energy transfer (FRET) reporter system was used to observe the temperature-dependent disassembly of the lattice and the 4x4 DNA tile, allowing the thermodynamic parameters of the system to be calculated.[72]

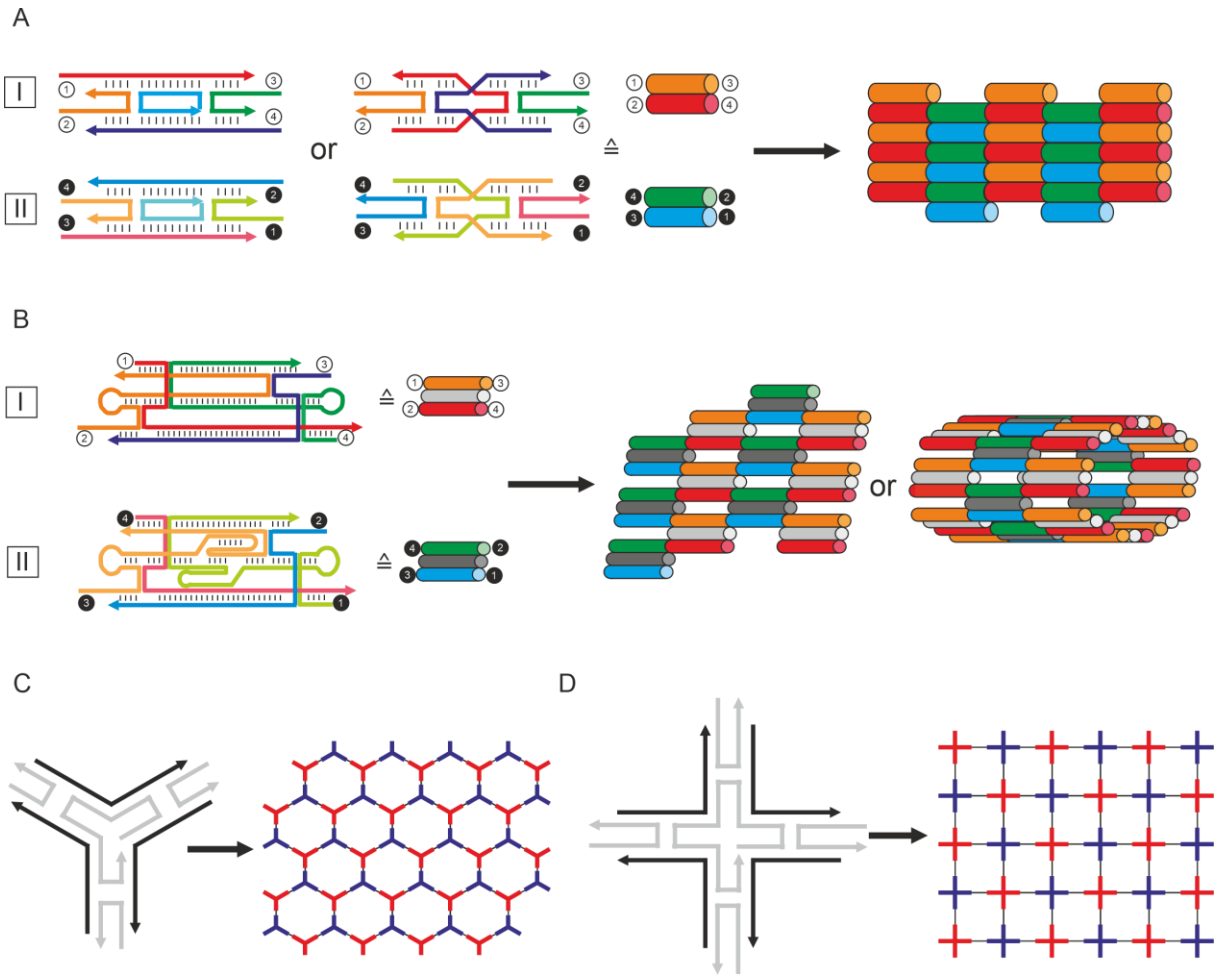


Figure 1.6 Design of the molecules and schematic representation of the 2D DNA structures that they form. The sticky ends of the (A) DX[64] and (B) TX molecules[67] are labeled with black numbers and the complementary strand is labeled with white numbers. The general structure of the building block that can form the corresponding 2D DNA structure is shown for (C) Y-shaped motifs[68] and (D) 4×4 DNA tiles.[70]

In addition to two-dimensional structures, three-dimensional structures could also be realized using DNA as a material. The first example was given by Seeman, who used three-way junctions to fold circular DNA into a cube (see Figure 1.7A).[73] The structure of a truncated octahedron[74] and a tetrahedron[75] could be formed by utilizing single-stranded DNA. An approach to create different DNA polyhedra was presented by Mao and coworkers. The Y-shaped or three-pointed star motif, consisting of single strands, can form a tetrahedron, dodecahedron or Buckminsterfullerene-like structure, depending on the concentration of the DNA (see Figure 1.7B).[76] This work was extended by Krishnan and coworkers, who used a five-way junction to form an icosahedron that could also encapsulate gold nanoparticles.[77]

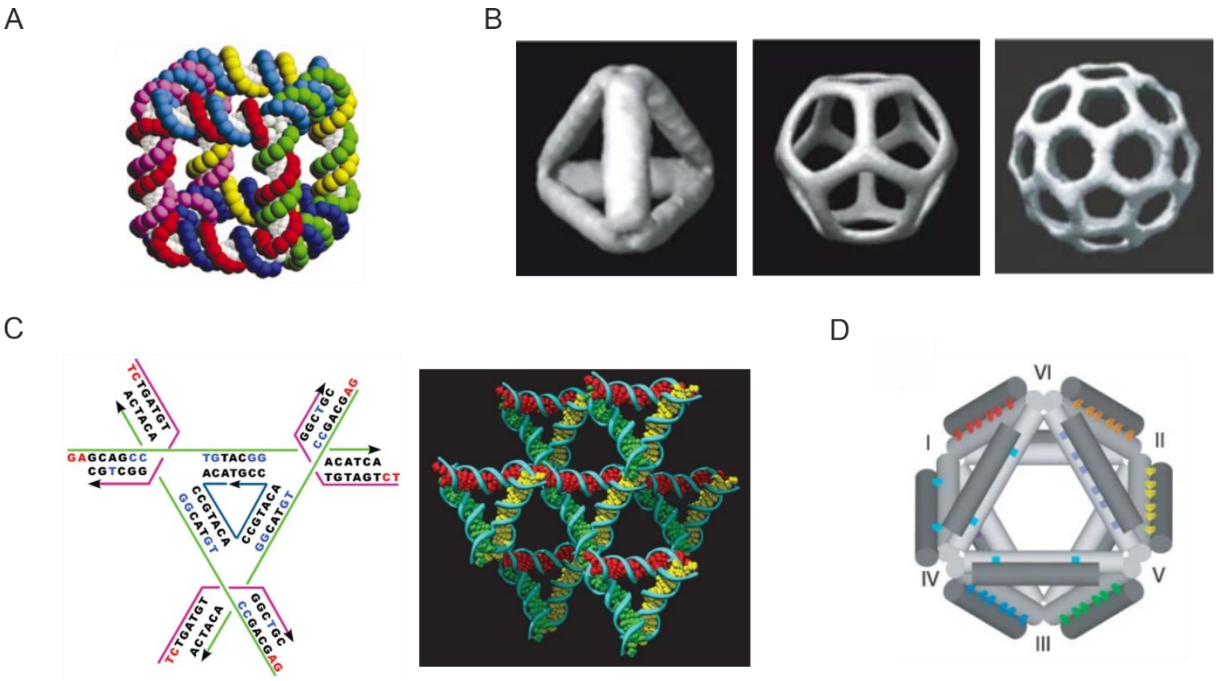


Figure 1.7 Examples of various 3D DNA structures derived from DNA nanotechnology. (A) DNA cube consisting of cyclic single strands.[78] Adapted with permission. Copyright 2003 Springer Nature. (B) Reconstructed structures of a DNA tetrahedron (left), dodecahedron (middle) and Buckminsterfullerene-like structure (right).[76] Adapted with permission. Copyright 2008 Springer Nature. (C) Sequence design (left) and illustration of a tensegrity triangle (right) with the ability to form 3D-DNA crystals with short sticky ends.[79] Adapted with permission. Copyright 2009 Springer Nature. (D) A schematic representation of the octahedron structure created by folding a 1700-base DNA strand with shorter DNA strands.[80] Adapted with permission. Copyright 2004 Springer Nature.

The design of a tensegrity triangle was the key to create the first three-dimensional DNA crystal. The three DNA double strands of the triangle were not assembled in a plane but rather in an over-and-under structure, which allows the formation of a three-dimensional structure using sticky ends that consist of only two nucleotides (see Figure 1.7C).[79] Different structures can also be obtained by varying these sticky ends.[81] Joyce and his colleagues took a completely different approach to create three-dimensional structures. Instead of using single strands to create building blocks for a complex structure in a bottom-up approach, they used an about 1700 nucleotides long single-stranded DNA and controlled the correct folding with five 40mer oligonucleotides to form an octahedron structure (see Figure 1.7D).[80] This system is considered the early version of DNA origami.

1.2.1 DNA origami

In 2006, Rothemund presented a way to fold a long single-stranded DNA with other shorter DNA strands in a two-dimensional shape. He called this approach DNA origami. The basis of this approach is a circular genomic DNA from the virus M13mp18 with about 7000

nucleotides. Once a desired shape has been selected, the structure is filled with an even number of parallel double helices. A computer program designs the desired complementary DNA single strands to fold the long DNA strand with periodic crossovers to hold the DNA double helices together. After the combination of all DNA strands and an annealing step, the desired structure could be observed (see Figure 1.8).[82]

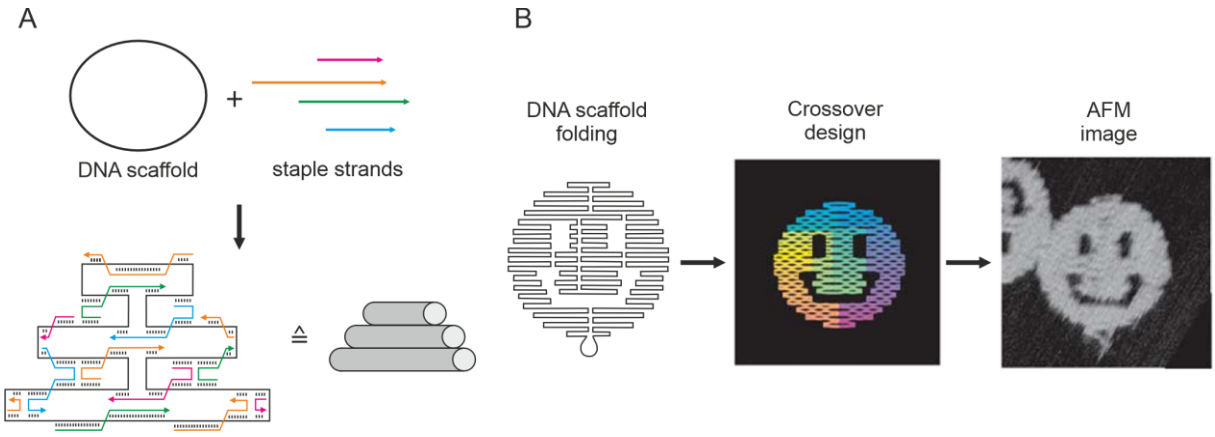


Figure 1.8 (A) Principle of DNA origami. A single-stranded cyclic DNA scaffold is folded into the desired shape by smaller synthetic DNA strands that specifically bind different scaffold regions to connect them via crossovers. (B) Application of the DNA origami using the smiley shape as an example. After putting the DNA scaffold in the shape of a smiley face to design the single strands for the crossovers, the DNA strands are combined with each other to create the desired structure. The product is then visualized by AFM spectroscopy (AFM image size: 165 nm × 165 nm).[82] Adapted with permission. Copyright 2006 Springer Nature.

This approach has been adopted and was further elaborated by other research groups. The area of the two-dimensional structure is limited to around 10000 nm^2.[82] Several groups came up with the idea to extend DNA origami with DNA sticky ends. These sticky ends connect the individual DNA origamis, resulting in a polymer. First attempts in Liu's group to combine DNA origami plates with DNA sticky ends at two opposite sides resulted in the formation of linear arrays and tubular structures.[83] To create a two-dimensional DNA origami array, Seeman and co-workers used a double-layered DNA origami tile that can be extended in four directions, resulting in a two-dimensional structure with 2-3 µm long sides (see Figure 1.9A).[84] By programming the interaction of the ends of a rectangular DNA origami piece made of four triangular DNA origami tiles, complex two-dimensional structures can be formed. Apart from using short sticky ends, Qian's team used blunt-end stacking interactions, which are known to affect the assembly of DNA nanostructures.[85-86] Due to the weaker interaction of blunt-end stacking, kinetic traps that cause misassembled tiles are avoided. This approach could align up to 5x5 arrays of rectangular DNA origami pieces.[86] Fractal assembly of the DNA oligo tiles allows exact

placement in an array, which enables the drawing of a pattern of the Mona Lisa or other images at nanometer scale (see Figure 1.9B).[87]

There are multiple approaches to generate three-dimensional structures with DNA origami.[88] Introducing single-stranded spacers in a planar DNA origami design resulted in the controlled folding of the DNA origami into a three-dimensional structure. This approach has been applied to assemble a DNA cube using six DNA origami sheets (see Figure 1.9C).[89] Other concepts utilize a multilayer strategy. In conventional DNA origami design, the helical axes of the DNA duplexes are parallel to each other to obtain a planar structure. This arrangement is due to the crossover distances of 1.5 turns, about 15-16 base pairs. A change in the crossover distances to 7 base pairs results in an orientation of three helices with an attachment angle of 120°, forming a stapled three-dimensional structure with a honeycomb-shaped lattice cross-section.[90] Shih and Dietz extended this system to control twist and curvature in DNA bundles.[91] Later, Yan and coworkers were able to stack concentric rings of DNA in a curvature to form, for example, spherical shells or nanoflasks.[92] A new approach was presented 2015 from the Högberg group. The conversion of three-dimensional structures to a polyhedral triangulation mesh allowed the creation of wireframes of the desired structure. The next step was to find an optimal path in which the scaffold strand passes along each edge at least once and to create staple strands that would stabilize the structure. The group was able to form with this concept complex three-dimensional structures like a waving stickman or a Stanford bunny (see Figure 1.9D).[93]

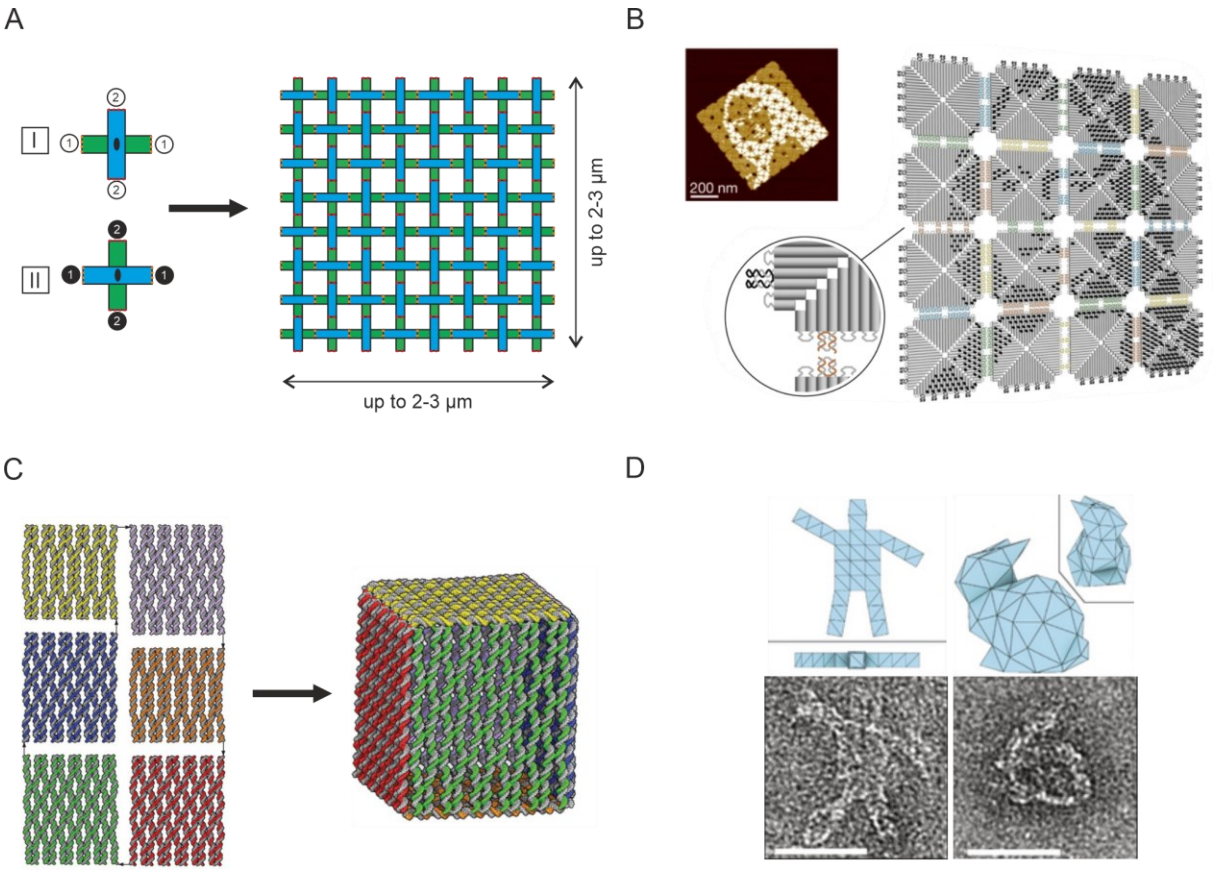


Figure 1.9 Examples of 2D and 3D structures that have been derived from DNA origami. (A) 2D grid of double-layered DNA origami structures. The black-labeled numbers indicate the sticky ends, and the white-labeled numbers indicate the complementary strands.[84] Adapted with permission. Copyright 2010 John Wiley and Sons. (B) AFM image and concept for a controlled 2D DNA origami array by fractal assembly, showing a pattern of the Mona Lisa. The inset shows the interaction of the rectangular DNA origami tiles by sticky and blunt end interaction.[87] Adapted with permission. Copyright 2017 Springer Nature. (C) DNA origami cube formed by six DNA origami sheets.[89] Adapted with permission. Copyright 2009 Springer Nature. (D) Wireframe DNA origami structures and negative-stain dry-state transmission electron microscopy images of waving stickman or a Stanford bunny (image size: 100 nm × 100 nm).[93] Adapted with permission. Copyright 2015 Springer Nature.

In recent years, alternatives to classical DNA origami have been presented. 2012 Yin and coworkers used a single-stranded-tile method to assemble molecular shapes. Each strand consists of four domains. Two domains contain 21 nucleotides, allowing them to form a two-dimensional DNA lattice in a brick-wall arrangement. Programming the assembly for unique recognition allows the formation of complex shapes (see Figure 1.10A).[94] The functionality was surprising, as DNA self-assembly of uniquely addressed tiling approaches is challenging because it is experimentally difficult to adjust the same concentration of individual strands. It was assumed that if each individual strand was not present in the same amount, the formation of the shape would be stuck in various unfinished structures.[95] However, the practicality of this method was demonstrated. The concept of the single-stranded tiles was extended to create three-dimensional structures. Using single strands with four domains, where each domain

contains eight nucleotides, allows the formation of three-dimensional structures. The individual strands were connected with a 90° left-handed turn, resulting in brick layers shifted 90° to each other (see Figure 1.10B).[96-97] Yield, kinetic and thermodynamic stability could be optimized using domains with a length of 13 nucleotides.[98] The disadvantage of this method is that many different single strands are required. In addition to multi-strand approaches, techniques for folding single-stranded DNA and RNA into complex shapes have been reported.[99] The combination with self-cleaving DNA enzymes also allows the upscaling of DNA single strands needed for DNA origami.[100]

Structures derived from DNA origami were used in a huge variety of applications.[88] The advantage of create rigid programmable structures of known size makes DNA origami perfect for using it as a scaffold to arrange, for example, gold nanoparticles.[101] Biophysical measurements also benefit from DNA origami. A "DNA force spectrometer or "nanoscopic force clamp" has been built using DNA origami to measure forces between nucleosomes or stacking forces.[102-104] Even the formation of a motor or a track for cargo transport made of DNA origami was reported.[105-106]

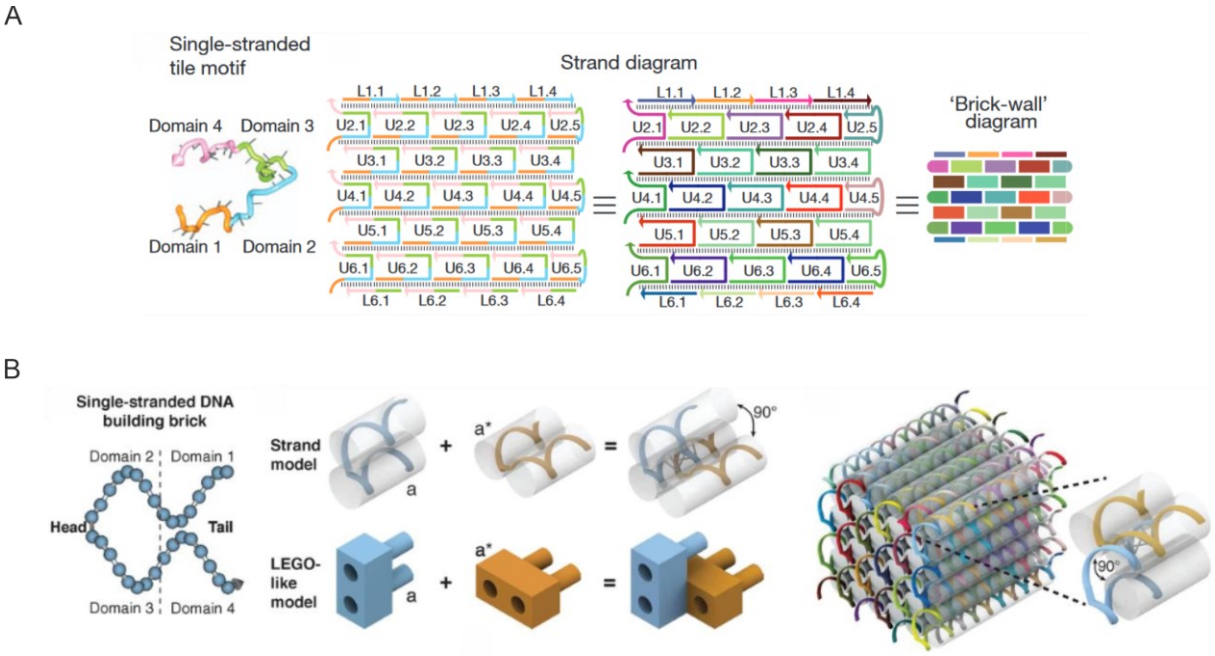


Figure 1.10 Alternative to traditional DNA origami. Designed single-stranded DNA strands with four domains can form (A) 2D lattices[94] and (B) 3D bricks[96]. Adapted with permission. Copyright 2012 Springer Nature and American Association for the Advancement of Science.

1.2.2 Dynamic DNA nanotechnology

In addition to the structural achievements based on the programmable character of DNA, the DNA duplex can be used for dynamic purposes. One of the first examples of a dynamic DNA device is a pair of linked DX molecules with a DNA helix that can switch between B-DNA and Z-DNA depending on the salt concentration, resulting in a conformational change of the entire structure.[107] Further developments of a dynamic DNA devices rely on programmed hybridization of complementary strands.[108-109] The most basic case of strand displacement is the exchange of a single strand in a DNA duplex with another complementary strand (see Figure 1.11A). Spontaneous dissociation of the duplex for strand displacement is required in this case. This dissociation is dominant at the DNA duplex melting temperature.[110] One of the first examples where strand displacement was used is the molecular beacon. The molecular beacon consists of a hairpin structure containing a fluorescence quencher system at the 5' and 3' ends of the stem region. Adding a single strand with the complementary sequence to the molecular beacon opens the hairpin structure driven by additional base pairs with the loop region, resulting in a detectable increase in fluorescence (see Figure 1.11B).[111] The concept has been used for a variety of targets, such as DNAzyme-based molecular beacons, molecular aptamer beacons, or molecular peptide beacons.[112-114]

A landmark in the field of dynamic nanotechnology is the concept of toehold mediated strand displacement. DNA toeholds are single-stranded domains that can initiate the strand displacement process. The unpaired toehold nucleotides facilitate strand displacement by binding to the target. The strand displacement proceeds through a three-way branch migration (see Figure 1.11C). The length and sequence complementarity of the toehold can significantly control the strand displacement rate.[108, 115] Yurke *et al.* presented the first example using toehold mediated strand displacement to control the structure of a DNA tweezer. They were able to switch the molecular tweezers between open and closed states by using different DNA strands (see Figure 1.11D).[116]

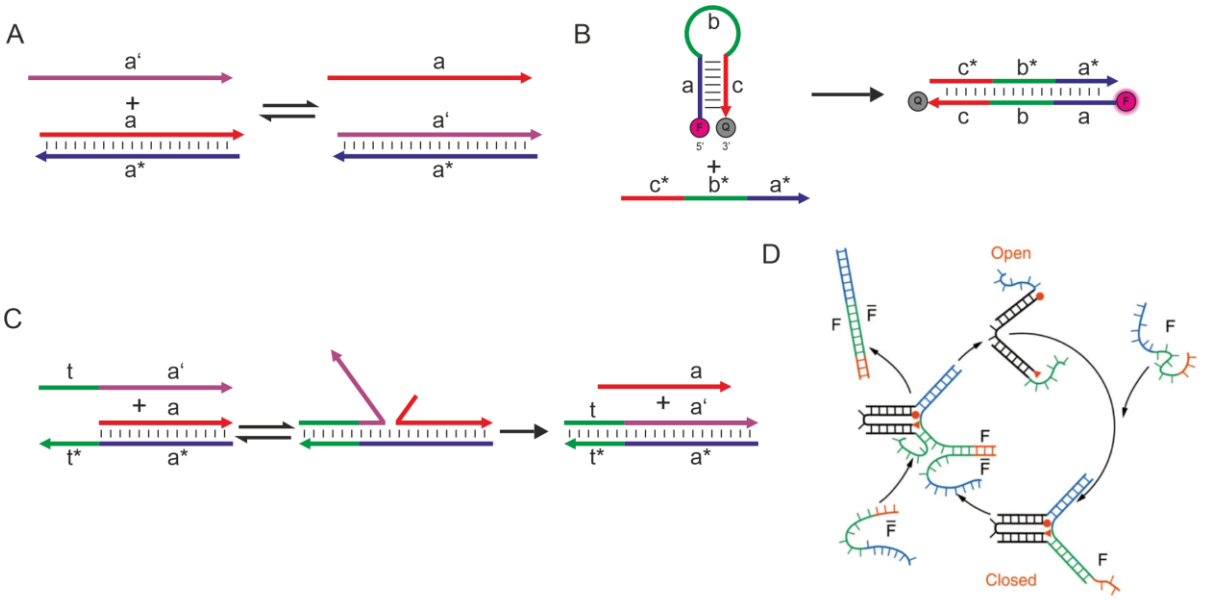


Figure 1.11 (A) Displacement of strand "a" with a strand "a'" that contains the same sequence. (B) Molecular beacon labeled with a fluorophore (magenta sphere) and a quencher (gray sphere). After binding to the target, the fluorophore lights up. (C) Toehold-mediated strand displacement using a three-way branch migration. (D) DNA tweezer capable of switching between open and closed conformation using DNA strands as input.[116] Adapted with permission. Copyright 2012 Springer Nature.

Strand displacement can also be used to control the configuration of larger structures or even two-dimensional DNA lattices.[117-118] The first DNA walkers were created based on DNA molecular machines, which can walk along a defined path using two leglike DNA strands. A DNA biped that can take a step was demonstrated by Sherman and Seeman. The addition of input strands that allow the walker to take a step by releasing the leg from its position on the path triggers the movement.[119] The design of DNA walkers has been improved to allow a greater number of steps.[120] A great advantage was the development of DNA origami, which allows to create a two-dimensional track for the DNA walker. A molecular random walker or molecular spider, consisting of a streptavidin molecule as an inert body and three deoxyribozymes as catalytic legs, can walk along a defined path on a DNA origami landscape. Essential for the movement of the molecular spider is the incorporated 8-17 DNAzyme sequence, which can cleave the substrate on the DNA origami landscape to allow the dissociation of the molecular spider.[121] Besides movement, cargo transport can be realized with a DNA walker.[122] Qian and coworkers created a cargo-sorting DNA robot. This DNA robot uses a random walk approach to move freely on the DNA origami landscape where different cargos are placed. Depending on the DNA tag of the cargo, the DNA robot can transport it to a defined area on the DNA

origami. The DNA robot can move approximately 300 steps, which is one to two orders of magnitude greater than other previously reported DNA robots.[123]

Controlling the motion of molecular machines is not the only research field where strand displacement plays an important role. As a result of strand displacement, a new single strand is released. This can be used to create programmed strand displacement cascades. Short DNA strands were used as inputs and outputs to implement Boolean logic gates.[124] These have been further involved using thresholding and catalysis within every logical operation.[125-126] A DNA-encoded enzymatic neural network was shown by Okumura *et al.* to detect mixtures of DNAs at different concentrations.[127]

Controlled structure formation by adding a single strand as an input is another area of research. For example, Yin *et al.* designed self-assembly pathways based on DNA hairpins that can form three- or four-armed branched junctions by adding an initiator strand.[128] Qian and coworkers developed a system in which a mixture of hairpins selectively forms DNA rings of different sizes depending on the initiator strand that is added.[129]

1.2.2.1 Hybridization chain reaction

In 2004, Dirks and Pierce presented an enzyme-free polymerization technique that can be initiated by an initiator strand, known as hybridization chain reaction (HCR).[130] The system includes two hairpins that coexist because of their kinetically trapped structure. A specific initiator can open one of the hairpins (H2, see Figure 1.12) by a toehold-mediated strand displacement, which exposes a single-stranded domain of the hairpin that acts as an initiator for opening the second hairpin (H1, see Figure 1.12). The exposed single strand sequence after opening the second hairpin is identical to the initiator sequence, enabling another hairpin (H2) to open, which was initially opened by the initiator. This process can be repeated until all hairpins have been used up in the procedure. The outcome is a DNA nanowire polymer, whose assembly is driven by the gain of free energy due to the additional base pair formed by of the bases in the former hairpin loop and toehold regions (see Figure 1.12).

The size of the polymeric DNA depends on the ratio of hairpins to initiator. Using 0.1 equivalents of initiator relative to the hairpin concentrations results in nanowires larger than 2000 base pairs. In addition, in the first design by Dirks and Pierce, ATP-aptamers were formed along the length of the nanowire after successful HCR.[130] Ang and Yung have tried to come up

with a simple set of rules to facilitate the rational design of hairpins in the HCR. The toehold may not exceed a length of 12 nucleotides with a GC content that is not higher than 30-40%. Furthermore, the length of the stem should be greater than the length of the toehold, and if the lengths are comparable, the content of CG in the stem should be at least 60%.[131] The kinetics of the HCR can be controlled by mismatches in the stem or by the use of pH-dependent triplex formation of the toehold and the stem.[132-133] Besides linear nanowires, alternative structures can be created by redesigning of the system, including dendrimers and superbranched structures. [134-135]

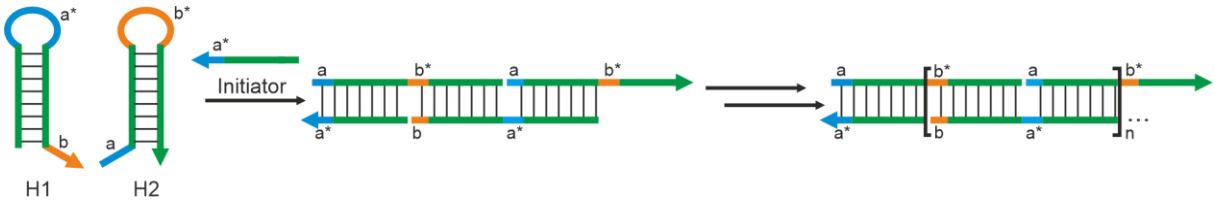


Figure 1.12 Concept of HCR. The hairpins (H1/H2) are kinetically trapped in their structure. The addition of an initiator will open the hairpin structure through toehold-mediated strand displacement, enabling the exposed single strand to initiate the opening of the other hairpin, leading to a polymerization-like reaction.

The HCR is mainly used in biosensing assays due to its amplifying, enzyme-free and isothermal properties. HCR approaches have been used to detect small molecules, nucleic acids, proteins, and even cells. To visualize HCR products, mainly optical methods such as fluorescence or phosphorescence are used, but atomic force microscopy or electrochemical methods can also be applied.[136] Another approach is to build a functional moiety, as Dirks and Pierce did with the ATP aptamer.[130] Sequences were detected by creating a functional DNAzyme that cleaved a reporter strand with a fluorophore quencher system, resulting in an increase in fluorescence.[137]

1.3 Supramolecular DNA assembly

Solid-phase oligonucleotide synthesis is an important method for DNA nanotechnology. It allows the synthesis of sequence-defined DNA strands that are needed for the formation of DNA nanostructures.[71] The underlying phosphoramidite chemistry introduced by Caruthers is not limited to the four canonical nucleotides.[138] The introduction of modifications within the DNA strand provides chemists with a powerful tool to obtain a variety of new structures and functions in the DNA. Supramolecular DNA assemblies can be divided into (i) DNA structure expansion through modification incorporation, (ii) DNA-templated supramolecular self-assemblies

and (iii) DNA-organic amphiphiles. The different classes will be discussed in the following section.

1.3.1 DNA structure expansion through modification incorporation

One of the first examples to control folding with artificial modifications was the use of an organic linker, such as hexaethylene glycol[139] or terephthalamide[140], that connects complementary strands to form a stable hairpin structure. Even electron transfer in DNA could be analyzed when using stilbene or diphenylacetylene as a cap to the hairpin structure.[141-142] A well-known concept in supramolecular chemistry is the use of predefined geometric molecules to define the assembly of a structure.[143] Shi and Bergstrom used two *p*-(2-hydroxyethyl)phenylethynylphenyl spacers as a rigid tetrahedral ligand for two DNA strands. The use of self-complementary sequences resulted in a mixture of cyclic and oligomeric structures of up to seven units (see Figure 1.13A).[144] Using six different DNA sequence designs and 1,3-bis-(4-hydroxyphenyl)benzene as a rigid 120° linker, it was possible to form a cyclic DNA hexamer structure that enables templated gold nanoparticle arrangements (see Figure 1.13B).[145] The same linker can be incorporated into cyclic DNA triangles or squares with the correct design. The combination with strand displacement allows the exchange of the DNA strands with gold nanoparticle functionalized DNA strands.[146] Three-dimensional structures were achieved by the same group. They used a DNA helix, which connects the cyclic 2D DNA building blocks to form structures like a cube or a biprism.[147]

In addition to connecting two strands to implement a specific bending angle in the strand, multiple strands can be connected on a linker to mimic a three-way or four-way junction. A three-branched trisoligonucleotide resulted in the formation of di- and tetramers. However, linear and branched oligomers could be obtained using a rigid 1,3,5-trisubstituted benzene molecule connected to phenylacetylene as a linker.[148-149] In analogy to the unmodified Y-shaped building block using a three-way junction, it was possible to assemble a dodecahedral motif with the synthetic trisoligonucleotide using a 1,3,5-trisubstituted benzene as a linker (see Figure 1.13C).[150] Four-way-branched DNA connectors consisting of a porphyrin moiety were used to connect two-dimensional DNA lattices built from DX molecules to induce tube formation of the DNA (see Figure 1.13D).[151] Various branched organic molecules were used as a platform to organize triangular DNA origami, leading to a template for gold nanoparticles.[152]

A tetrahedral organic molecule could also be used as the basis for DNA-based nanoparticles (see Figure 1.13E).[153]

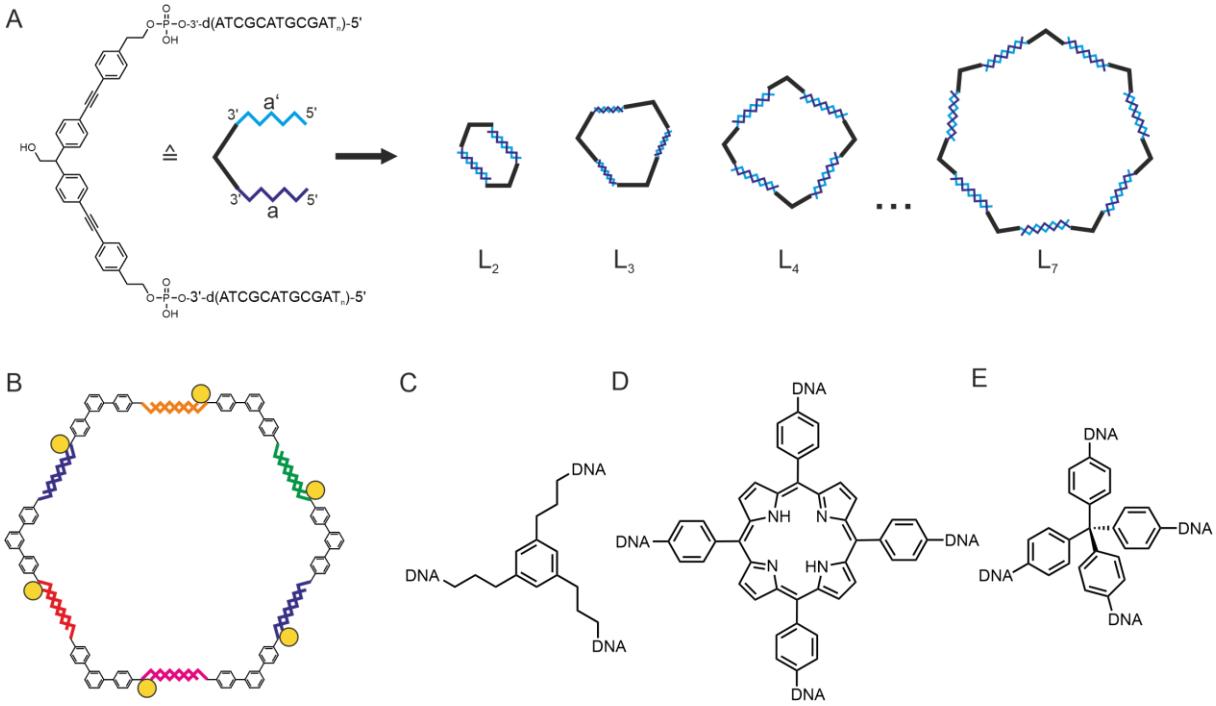


Figure 1.13 (A) System using a tetrahedral linker attached to self-complementary DNA sequences. Cyclic structures of up to seven units have been derived.[144] Adapted with permission. Copyright 1997 John Wiley and Sons. (B) A hexagon formed by programmed DNA sequences connected by a rigid linker. Incorporated gold nanoparticles are shown as yellow spheres.[145] Adapted with permission. Copyright 2006 John Wiley and Sons. (C) 1,3,5-Trisubstituted benzene, (D) porphyrin, or (E) tetraphenyl methane as examples of linkers for multi-strand linkage.

In addition to organic molecules, metal complexes are useful for DNA alignment because of their known coordination geometry.[154] Transition metal complexes could be incorporated directly during synthesis. McLaughlin and coworkers used this approach to synthesize a six-oligonucleotide-armed ruthenium(II) tris(bipyridyl)-centered complex.[155] Transition metal complexes must be oxidatively stable, which can be problematic. Therefore, only ligands such as bipyridine[156], terpyridine[157], phenylphenanthroline[158] or salen[159] were used in solid-phase oligonucleotide synthesis and the transition metal was introduced afterwards. Ligand specificity was used to distinguish the metals Fe^{II}, Cu^{I} and Cu^{II}[160] or to create chiral four-way junctions.[161] The controlled formation of G-quadruplexes[162] or i-motifs[163] could also be achieved by transition metals.

1.3.2 DNA-templated supramolecular self-assemblies

The ability to interact with the DNA controls the hybridization of small molecules and allows the DNA to be used as a template. There are distinct target sites where DNA hybridization can take place. Hydrogen bonds, which can form at the Watson-Crick or Hoogsteen base pairing side, are the best-known interaction options. Interactions with the charged phosphate backbone or alignment in the major groove are also feasible.[164]

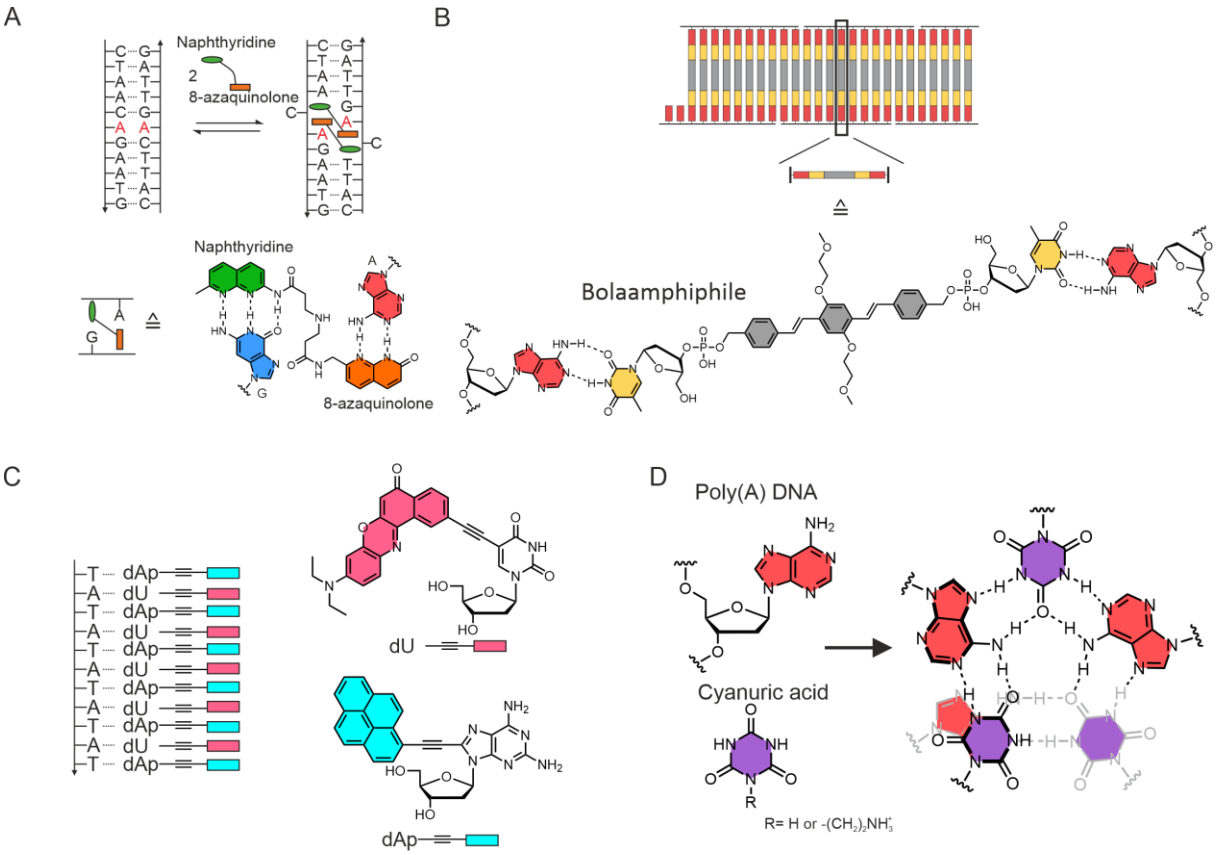


Figure 1.14 (A) Base pairing of a naphthyridine-azaquinolone ligand with canonical purine nucleobases induces base-flipping in an A-A mismatch (highlighted in red).[165] Adapted with permission. Copyright 2005 Springer Nature. (B) Schematic representation of the nanofiber structure that forms between a nucleotide bolaamphiphile and poly(A) DNA by A/T base pairs.[166] Adapted with permission. Copyright 2006 American Chemical Society. (C) Programmable arrangement of the ethynyl pyrene diaminopurine 2'-deoxynucleoside (blue) and the ethynyl Nile red-modified 2'-deoxyuridine (red) on a DNA single-strand.[167] Adapted with permission. Copyright 2018 John Wiley and Sons. (D) Illustration of the helical structure of a DNA triplex by continuous cyanuric acid-adenine hydrogen bonds.[168-169]

A naphthyridine-azaquinolone ligand was able to induce base flipping in an A-A mismatch by hydrogen bonding twice using the Watson-Crick base pairing side of the A and a neighboring G (see Figure 1.14A).[165] This concept has been used to stabilize a DNA duplex containing multiple mismatches[170] or even to control the assembly of a DNA nanostructure as molecular glue.[171] Artificial supramolecular assemblies using a DNA strand as a template have been made

using so-called nucleotide bolaamphiphiles. These are two nucleotides connected by a hydrophobic linker, which can either be two thymidine monophosphates connected by a C20 alkyl chain or oligo(*p*-phenylene vinylene)s. Using poly(A) DNA strands, the bolaamphiphil can build nanofibers through A/T base pairing (see Figure 1.14B).[166, 172] This assembly is made of two strands of DNA, but building blocks could also be hybridized on a single strand. This single-stranded assembly requires a moiety that can form hydrogen bonds to nucleobases and an aromatic ring that allows sufficient π-stacking to stabilize the supramolecular structure. The combination of diamino triazine as a unit for hydrogen-bonding and naphthalene or oligo(*p*-phenylene vinylene) for π-stacking allows the formation of a new DNA hybrid with a poly-T strand.[173-174] Wagenknecht and coworkers showed that single-strand hybridization is also possible with nucleosides that connect the chromophore via an ethynyl bridge on the nucleobase.[175] A programmable arrangement of the ethynyl pyrene diaminopurine 2'-deoxynucleoside and the ethynyl Nile red-modified 2'-deoxyuridine was achieved using designed sequences consisting of dA and dT (see Figure 1.14C).[167] Moreover, the non-planar tetraphenylethylene attached to 2'-deoxyuridine was able to hybridize, showing aggregation-induced emission.[176] Sleiman and coworkers observed the formation of nanofibers using cyanuric acid and a poly(A) DNA strand, resulting in a helical structure of a DNA triplex by continuous cyanuric acid-adenine hydrogen bonds (see Figure 1.14D).[168, 177-178] Formation can be controlled by photoacids, which allow depolymerization and subsequent repolymerization by light irradiation, resulting in robust fibers.[179] A hydrogel that can release an antisense oligonucleotide for gene silencing was created by connecting the fibers using a double strand.[169]

1.3.3 DNA-organic amphiphiles

Amphiphilic molecules consist of a hydrophilic and a hydrophobic unit. One of the most prominent representatives are lipid molecules, which can form biomembranes. Depending on the critical packing parameter, different morphologies such as micelles, spherical vesicles, cylindrical micelles or planar bilayers can be adapted in water.[180]

DNA is a hydrophilic molecule, but its structure relies also on hydrophobic interactions such as π-stacking of blunt ends,[104] which are important for controlling the assembly of DNA origami structures.[85] By incorporating artificial hydrophobic building blocks, the hydrophobic character of DNA can be enhanced.[181] The Wagenknecht group used perylene bisimide units

at the three ends of a three-way junction to act as a glue to hold the ends together and form aggregates.[182] Similar behavior was previously observed for perylene bisimides when used as linkers for a hairpin structure, which can form dimers by stacked perylene bisimide units.[183] The importance of hydrophobic interactions between chromophores for the folded structure was demonstrated with 1,3,5-phenylacetylene-substituted benzenes. Two different conformations could be observed depending on the stiffness of the linker between the units. Using a T_6-linker between the hydrophobic moiety and the complementary DNA sequence induced the formation of a DNA cage dimer, whereas no linker led to a face-to-face dimer (see Figure 1.15A).[184] In addition to efforts introducing hydrophobic modifications into DNA to allow biomembrane incorporation, hydrophobic moieties can control the morphology of DNA-organic amphiphiles.[185] The Gianneschi group developed amphiphilic DNA brush copolymers with programmable morphology. Initially, the copolymers form a micelle with a spherical morphology. The DNA strands were shortened by a DNAzyme, resulting in a cylindrical morphology. This was reversed by the addition of a complementary DNA strand, increasing the size of the hydrophilic DNA part.[186] The Sleiman lab presented a way to synthesize amphiphilic DNA polymers based on dodecanediol or hexaethylene glycol linked by a phosphate moiety via solid-phase synthesis using phosphoramidite chemistry.[187] The addition of Cy3 units at different positions of 12 dodecanediol units changed the morphology of the assembly. Cy3 units at the end of the hydrophobic unit induced the formation of DNA nanofibers, which could be cleaved using a photolabile linker to obtain again a spherical assembly (see Figure 1.15B). The spherical assembly is also obtained when the Cy3 units are placed between the DNA and the hydrophobic part of the molecule.[188] Sequence and length of the DNA in the amphiphilic polymer play an important role in assembling.[189] Additional control is provided by attaching alkyl chains on defined sides of DNA nanostructures. Placing the chains on one side of DNA cubes allows controlled formation of a hydrophobic core surrounded by the DNA cubes[190], and placing cholesterol molecules on one side of a DNA origami sheet results in the folding of the structure.[191]

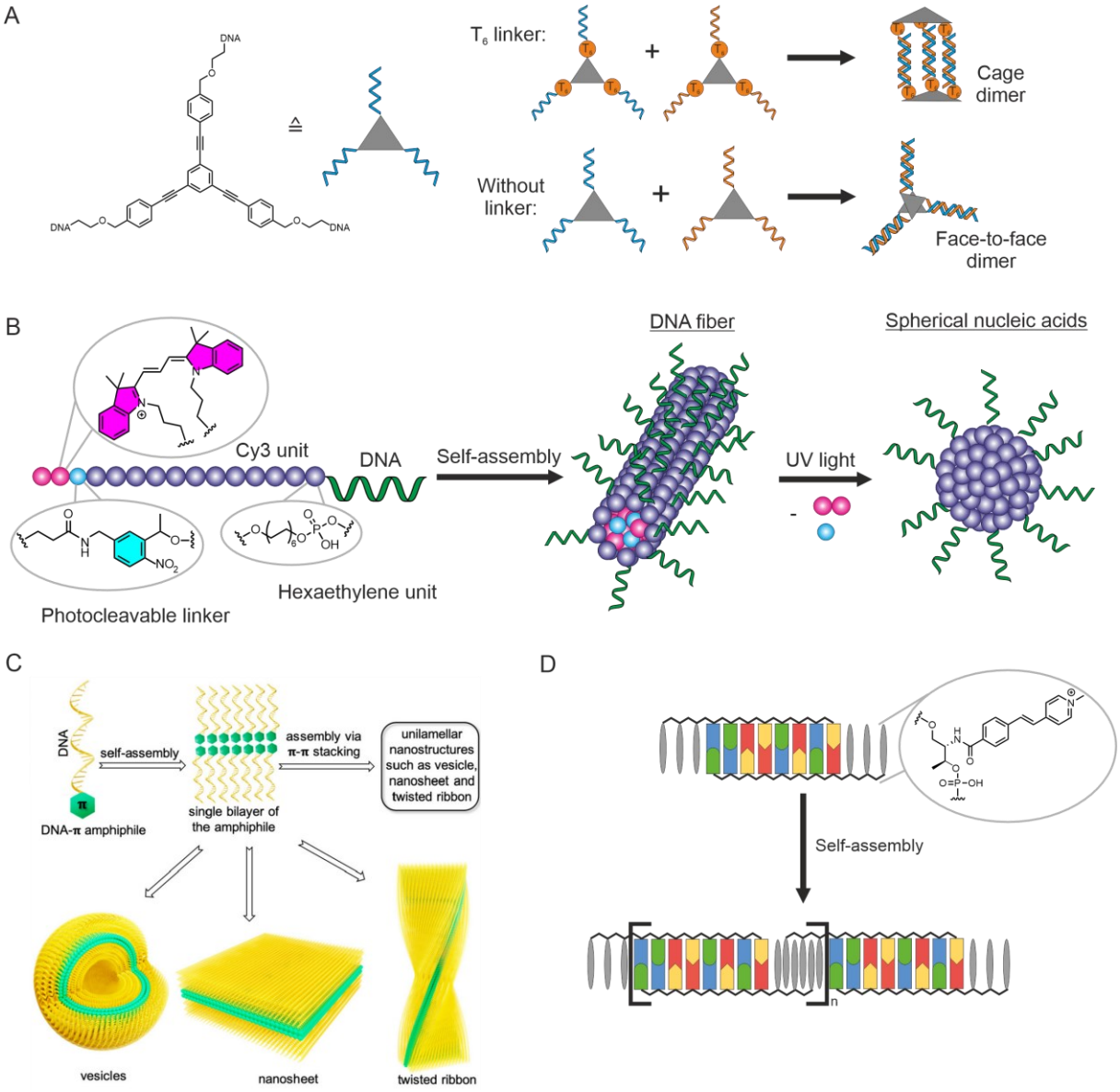


Figure 1.15 Examples of various structures formed by DNA amphiphiles. (A) DNA functionalized 1,3,5-phenylacetylene-substituted benzenes that form a cage and face-to-face dimer based on hydrophobic interactions and sequence design.[184] Adapted with permission. Copyright 2012 American Chemical Society. (B) Self-assembly of DNA fibers from amphiphilic DNA polymers based on hexaethylene units with Cy3 chromophores. After cleavage by a photocleavable linker, the morphology changes to a spherical micelle.[188] Adapted with permission. Copyright 2018 American Chemical Society. (C) Different morphologies that can be derived by DNA-π amphiphiles.[192] Adapted with permission. Copyright 2020 American Chemical Society. (D) *p*-Methylstilbazoles function as a molecular glue to form a polymeric DNA nanowire.[193] Adapted with permission. Copyright 2011 John Wiley and Sons.

Flexible hydrocarbon chains mostly lead to the formation of spherical or cylindrical micelles, and the formation is driven only by the hydrophobic effect. The use of aromatic rings extends the control of the interaction by π-π stacking, allowing the formation of more complex assemblies like vesicles, nanosheets or twisted ribbons (see Figure 1.15C).[192] Incorporating multiple

perylene bisimide units into DNA strands leads to a nanostructure formation driven by the stacked assembly of the aromatic rings.[194] Asanuma and coworkers introduced *p*-methylstilbazoles as sticky ends at both ends of a DNA duplex. The aromatic stacking, as well as the positive charge of the building block, which can interact with the negative charge of the phosphate backbone, results in self-recognition of the triplex unit and the formation of a DNA nanowire (see Figure 1.15D).[193] The assembly of oligoarenotides, which are oligophosphates of aromatic moieties, was the focus of Häner's lab. In 2007, they succeeded in embedding 14 consecutive achiral pyrene building blocks into a DNA strand.[195] High salt concentrations allow hybridization of stacked pyrenes into a supramolecular structure. The arrangement of the stacked pyrenes can be influenced by the chiral information of 1,2-diaminocyclohexane or deoxycytidine.[196-197] Functionalized DNA with seven pyrenes at its 5' end creates a stair-like folding of the pyrene units, forming a DNA-grafted linear supramolecular polymer as a ribbon-shaped helical aggregate.[198] These kind of wires were also obtained using short amphiphilic oligomers consisting of three phenanthrene units.[199] Their light-harvesting behavior makes them ideal for functionalization with a DNA photonic wire to transfer excitation energy absorbed by the phenanthrene molecules to an acceptor dye and further to other chromophores in the DNA photonic wires.[200]

Häner's group realized that, depending on the position of the linker on the pyrene or phenanthrene unit, the trimers can form multidimensional structures such as nanotubes.[201-202] A vesicle-like structure was obtained with oligo(*p*-phenylene-ethynylene)s as hydrophobic units on a DNA single strand. By adding a complementary single strand, this structure was converted into a light-harvesting cylindrical micelle.[203] Attaching a C-rich sequence to the oligo(*p*-phenylene-ethynylene) motif allowed the morphologies to switch in a pH-dependent manner between vesicles and an entangled network due to i-motif formation.[204] Sticky ends consisting of three phenanthrene or tetraphenylethylene units at each end of a DNA duplex also resulted in the formation of vesicles.[205-206] Especially in the case of tetraphenylethylene, different vesicle shapes of ribbons or star-like nanoobjects could be observed if residues like a PEG chain or a branched tricarbohydrate moiety was attached at the 5'-end of a DNA duplex.[207] Vargehese and coworkers also observed the formation of nanosheets using hexa-*peri*-benzocoronene or tetraphenylethylene units as a dendritic wedge when attached at the 5' end of a single strand. High-resolution transmission electron microscopy revealed a distance of 3.4 Å between the separating stripes, the optimal distance for stacking interactions.[208-209] With

hexaphenylbenzene and its adopted propeller conformation, even helically twisted ribbons were realized.[210] Due to its amphiphilic character and biocompatibility these structures could be optimal for use as a nanocarrier for drug delivery.[192]

1.4 Artificial nucleic acids

In the previous chapters, it was shown that DNA, in addition to its well-known role in living organisms as the carrier of genetic information in the form of B-DNA, can adopt a variety of different structures and functions based on its base pairing and folding capabilities. To expand the function of the canonical nucleobases, the bases are chemically modified in nature. Modifications such as methylation are involved in maintaining cellular function and genomic stability.[211] Chemical synthesis offers the possibility to extend the function of nucleotides. Besides the connecting phosphate unit, the sugar backbone and the nucleobase are the primary targets for modification and exchange.[212-213] Artificial nucleic acids, also known as xenonucleic acids (XNA), were initially developed to overcome the *in vivo* instability of RNA, mainly through ribose modification. The field of application has been greatly expanded by an enormous number of possible modifications.[214] Artificial backbones can stabilize the hybridization and allow the introduction of reporter molecules into the backbone. Simplification of the chemical synthesis by artificial backbones allows for the incorporation of nucleobase surrogates.[215-216] These artificial nucleobases can function as new base pairs orthogonal to the canonical base pairs,[217] as reporter molecules, such as fluorescent nucleobases,[218] or can carry functional groups that extend its reactivity, such as crosslinking reactions[219]. The use of artificial backbones and nucleobases will be the subject of further discussion.

1.4.1 Artificial nucleic acid backbones

Naturally occurring nucleic acids have a ribose or deoxyribose backbone in common. To optimize the strand for hybridization and stability, the backbone was modified. The backbone modifications can be divided into ribose-, sugar-modified, and acyclic backbones.

1.4.1.1 Ribose-modified backbones

Ribose-modified backbones are derived from the ribose sugar moiety and can cause a distortion of the helical structure. 2'OH modified sugars should also show resistance to nucleases.[216] Methylation of 2'OH of the ribose is a common post-transcriptional modification found in

tRNAs, rRNAs and mRNAs of living systems and has a regulating function.[220] For example, it was shown that mRNA methylation can regulate gene expression *in vivo*.[221] In addition, thermal stabilization of an RNA duplex was observed due to its preference for the C3'-endo conformation (see Figure 1.16A).[222-223] Higher affinity to complementary DNA and RNA sequences due to its fixed C3'-endo conformation was also observed for the locked nucleic acid (LNA).[224] The locked conformation is achieved by a methylene bridge between 2'-oxygen and 4'-carbon.[225] The tolerance of 2'-*O*-Me RNA and LNA *in vivo* makes these modifications optimal scaffolds for antisense oligos or DNAzymes.[226-227] Another artificial backbone can be derived from the ribose sugar if its 2'OH is replaced by a fluorine atom. Depending on the 2'C substituent of the sugar, there are different sugar conformations. 2'-Fluoro ribonucleic acid (FRNA) shows a *C*3'-endo conformation like RNA[228], whereas its 2'-epimer 2'-fluoro-arabinose nucleic acid (FANA) prefers a *O*4'-endo conformation like the 2'-epimer of RNA arabino nucleic acid (ANA) (see Figure 1.16B).[229] Due to the absence of fluorine atoms in DNA, fluorinated building blocks are helpful as tags for NMR spectroscopy to observe different structures of nucleic acids.[230] As an anomalous scatterer for multiwavelength anomalous dispersion, 2'-methyl seleno-modified oligoribonucleotides are also useful in X-ray crystallography.[231]

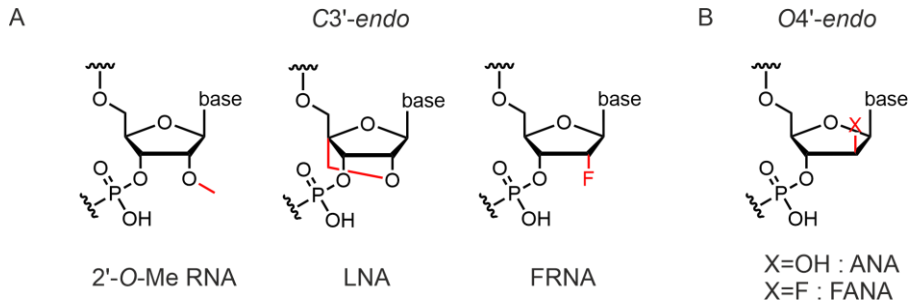


Figure 1.16 Examples of ribose-modified backbones. (A) Structure of 2'-*O*-Me RNA, LNA and FRNA preferring a C3'-endo conformation and (B) ANA and FANA preferring a *O*4'-endo conformation.

1.4.1.2 Sugar-modified backbones

In a systematic experimental study, Eschenmoser and coworkers synthesized several alternative nucleic acid systems based on pyranose or furanose rings. Homo DNA was presented as the first XNA with a hexopyranosyl-(4'→6') backbone, which can form a more stable duplex than DNA due to a stronger pre-organized single-strand. Watson-Crick base pairing formation was not observed with DNA or RNA as the second strand. The use of β-allo-, β-altro- and β-glucopyranosyl as backbone, which contain two additional hydroxy groups compared to homo DNA, does also not lead to Watson-Crick base pairing, mainly due to steric hindrance caused by the introduced hydroxy groups (see Figure 1.17A).[232-233] By inserting an additional carbon

atom between the *O*4' and *C*1' of the β-D-2'-deoxyribose moiety, Herdewijn and coworkers created hexitol nucleic acids (HNA) that exhibit stable duplex formation with DNA, RNA, and themselves.[234-235] The use of a cyclohexene ring instead of a pyranose ring in HNA results in cyclohexene nucleic acids (CeNA) (see Figure 1.17B), which form even more stable duplexes compared to HNA with DNA and RNA and RNA cleavage mediated by RNase H is observed in RNA/CeNA hybrids.[236-237]

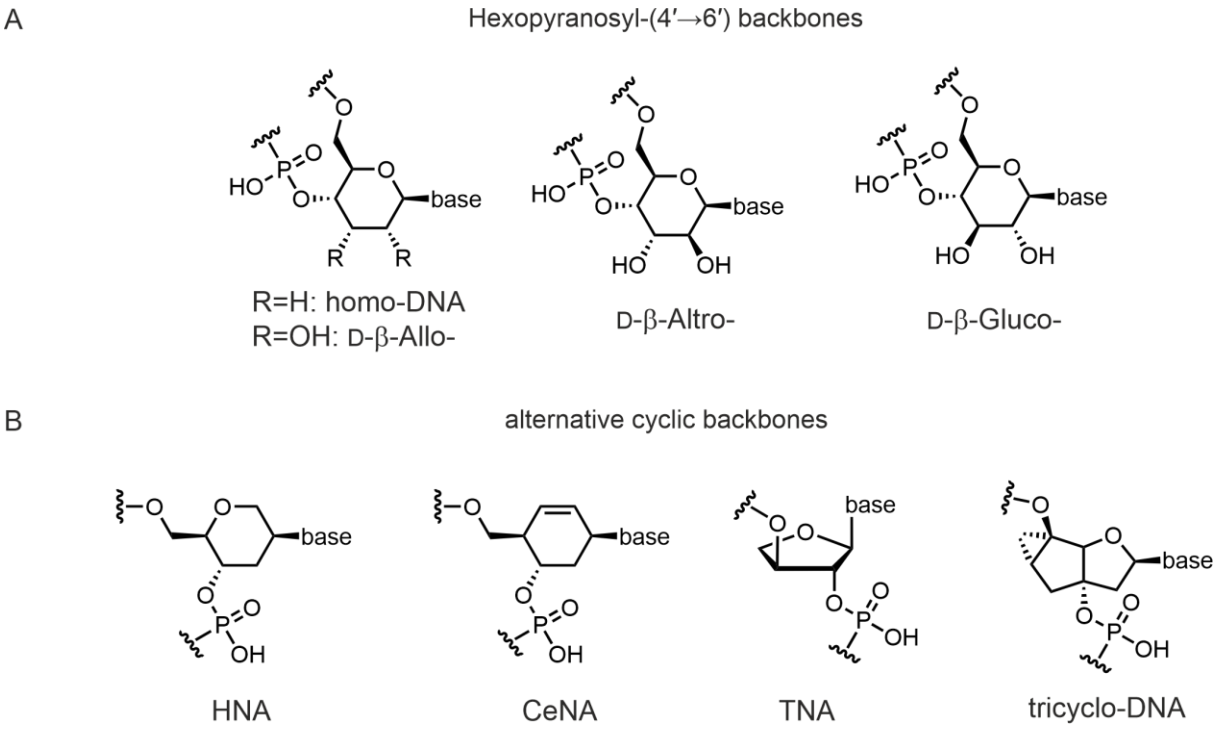


Figure 1.17 Examples of sugar-modified backbones. (A) Pyranose derived and (B) alternative cyclic artificial backbones.

With the the pyranosyl isomer of RNA (p-RNA), Eschenmoser and coworkers were able to form pentopyranosyl (2'→4') oligonucleotides containing β-ribo-, β-xylo-, α-lyxo-, and α-arabinopyranosyl. Each duplex of these systems has a higher thermal stability due to self-pairing than a RNA duplex, and they can even cross-pair with each other.[232] The L-α-threofuranosyl oligonucleotides (TNA), linked by (3'→2') phosphodiester bridges, are derived from the tetrose L-threose as a backbone sugar substitute (see Figure 1.17B). It contains one covalent bond less than the six bonds per repeating mononucleotide unit found in DNA and RNA. Interestingly, TNA has the ability to create stable duplexes not only with itself but also with a DNA or RNA strand.[238]

A tricyclo-DNA was introduced by Leumann and coworkers which reduces the flexibility of the β-D-2'-deoxyribose by an ethylene bridge between C3' and C5' and an additional cyclopropyl unit. As in LNA, the entropic cost of duplex formation is less, resulting in high melting temperatures when base paired with itself or DNA.[239]

1.4.1.3 Acyclic backbones

Tricylo-DNA and LNA achieved stronger hybridization of the duplex by fixing its backbone conformation due to additional cyclic bridges into the sugar moiety. On the other hand, acyclic backbones with higher flexibility can be obtained by removing some covalent bonds of the ribose backbone.[240] One of the first examples was the flexible nucleic acid (FNA), which lacks the C2' of the ribose (see Figure 1.18A). Incorporation of a single nucleotide led to strong destabilization compared to the native DNA duplex, due to the higher flexibility in the strand.[241] While a poly(A) strand with FNA is unable to base pair with its T analog, a melting behavior has been observed with a poly(T) DNA strand.[242] The unlocked nucleic acid (UNA), which only lacks the bond between C2' and C3' of ribose, shows a similar destabilizing behavior when incorporated into a DNA duplex (see Figure 1.18B).[243]

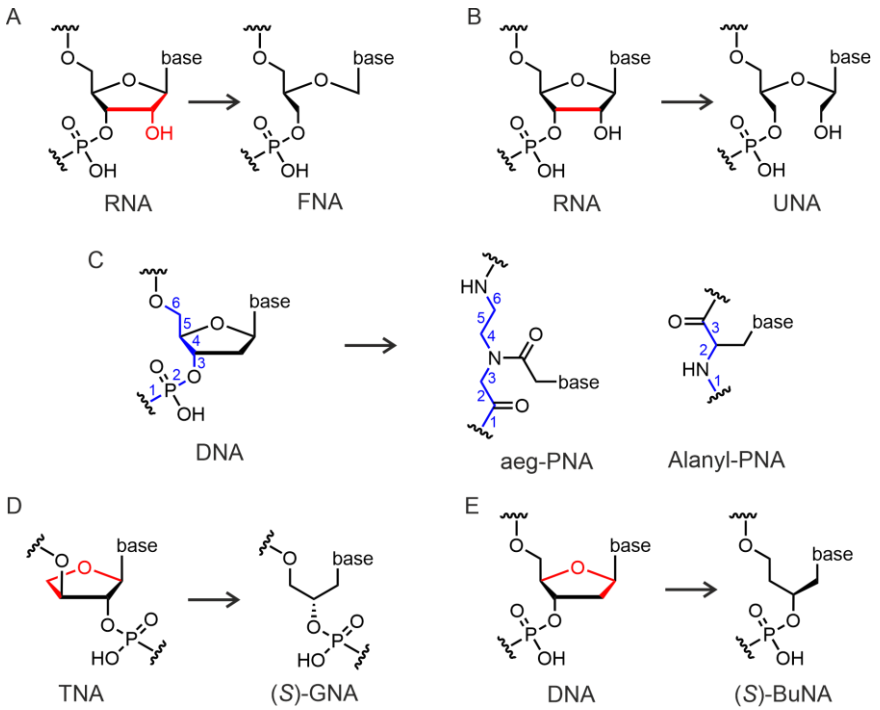


Figure 1.18 Examples of acyclic backbones. (A) FNA and (B) UNA are derived from RNA. (C) Peptide nucleic acids aeg-PNA and alanyl-PNA. The number of bonds per repeating unit is highlighted in blue. (D) (*S*)-GNA and (E) (*S*)-BuNA derived from TNA or DNA. Removed bonds are highlighted in red.

In 1991, Nielsen described *N*-(2-aminoethyl)glycine peptide nucleic acid (aeg-PNA), the first XNA with an acyclic backbone capable of forming duplex hybrids with DNA and RNA that have

even higher melting temperatures than the corresponding DNA and RNA duplex.[244-245] The backbone of aeg-PNA consists of *N*-(2-aminoethyl)glycine units connected by a peptide bond and the nucleobase attached through a methylene carbonyl group.[244] The lack of a charged phosphate linkage and achirality makes it unique compared to the DNA backbone analogs, but aeg-PNA still contains six covalent bonds per repeating unit like in DNA (see Figure 1.18C). This is the reason why aeg-PNA can form helical duplex structures with itself, unlike other peptide nucleic acids such as alanyl-PNA, which contains three covalent bonds per repeating unit (see Figure 1.18C).[246-247] The high stability of the hybrid duplexes with DNA and RNA makes aeg-PNA interesting for antigens or antisense agents and can be explained by the lack of repulsive electrostatic interactions of the phosphate groups. Due to its poor membrane permeability further modifications are needed for cellular applications.[245, 248]

The first self-complementary acyclic XNA with a charged phosphodiester backbone showing higher homoduplex stability than DNA was the glycol nucleic acid (GNA). The GNA design is based on the TNA introduced by Eschenmoser and consists of three carbon atoms with a chiral center (see Figure 1.18D).[249] The synthesis of the first building block was described in 1971 by Takemoto and coworkers using glycidol or 3-chloro-1,2-propanediol.[250] Meggers group used phosphoramidite chemistry to synthesize oligonucleotides for (*S*)-GNA and (*R*)-GNA.[251] Structural analysis revealed an existing pre-organization of the single strand, resulting in diminished entropic cost for duplex formation. This explains the high thermal stability of the GNA homoduplex compared to the DNA duplex.[252] Cross-hybridization was only observed with (*S*)-GNA and RNA containing a sufficient number of A/T(U) base pairs.[251] X-ray structures revealed that (*S*)-GNA can form two different right-handed helices with 10 or 16 residues per turn.[253-254] The torsion angle between the *C*3'-*O* and *C*2'-*O* bonds of the propylene glycol backbone can adopt gauche- and anti-conformations, which were observed alternately between nucleotides in one strand. Base pair formation was observed for nucleotides with opposite conformations.[253] The incorporation of a single GNA base pair into a DNA or RNA duplex resulted in strong destabilization. Schlegel *et al.* analyzed structures in which a single RNA nucleotide was exchanged for a GNA nucleotide and (*S*)-GNA seems to be better tolerated as a substitution than (*R*)-GNA. X-ray structures of duplexes containing a GNA thymidine nucleotide showed a rotation of the GNA nucleobase such that the C5 methyl group of GNA-T points to the minor groove instead of the major groove. This rotation was demonstrated using isoC and isoG,

which show a stronger duplex stability than C and G when base paired with their GNA counterparts.[255] Due to the synthetic simplification of the necessary phosphoramidites, the GNA scaffold is often used to introduce various building blocks into DNA.[256-257]

Butyl nucleic acid (BuNA) is an expansion of GNA by extending the acyclic backbone by one carbon unit to restore the six covalent bonds per repeating unit as in DNA (see Figure 1.18E). Using (*S*)-malic acid or (*R*)-aspartic acid as backbone precursors, the corresponding phosphoramidites for (*S*)-BuNA have been synthesized. BuNA can form a left-handed homo-duplex, which is thermally less stable than the corresponding DNA duplex. The incorporation of single BuNA base pairs into a DNA duplex also resulted in destabilization. Heteroduplex formation has not been observed with DNA or RNA.[258] Under low pH conditions, a poly(A) sequence can form an organized assembly held together by Hoogsteen base pairing of protonated adenines.[259]

1.4.2 Artificial base pairs

The synthesis of additional bases that can mediate base pair formation orthogonal to canonical base pairs provides access to the incorporation of specific functionalities that affect structural stability, programmability, or even to genetic code expansion. Successful recognition of the artificial base pair can be showed by enzymatic incorporation of polymerases or thermal stabilization of the DNA duplex. Various strategies have been used to create artificial base pairs, which will be discussed below.

1.4.2.1 Hydrogen bonding

A way to create new artificial base pairs is to generate alternative hydrogen bonding patterns. As early as 1962, Rich proposed isoC and isoG as artificial base pairs with an alternative hydrogen bonding pattern compared to natural base pairs.[260] In 1989, Benner showed the successful incorporation of isoG triphosphate derivatives by enzymes with the use of an isoC in the template.[261] The system showed drawbacks such as isoC decomposition and isoG tautomerization, where the presence of the enol form of isoG allows incorporation of T.[262] Optimization of the system leads to the P/Z base pair. Use of imidazo[1,2-a][1,3,5]triazine as the aromatic ring for P resulted in the removal of the imino proton and tautomeric stabilization and allowed efficient base pairing with Z, a base that uses a nitro group for reduced epimerization.[263-264] After 30 years, the Benner group introduced the base pair isoG (named B) and S to complete an eight-letter system called hachimoji, which means eight letters in Japanese (see

Figure 1.19A). The creation of a fluorescent RNA aptamer based on the spinach aptamer with the hachimoji DNA and RNA was possible.[265]

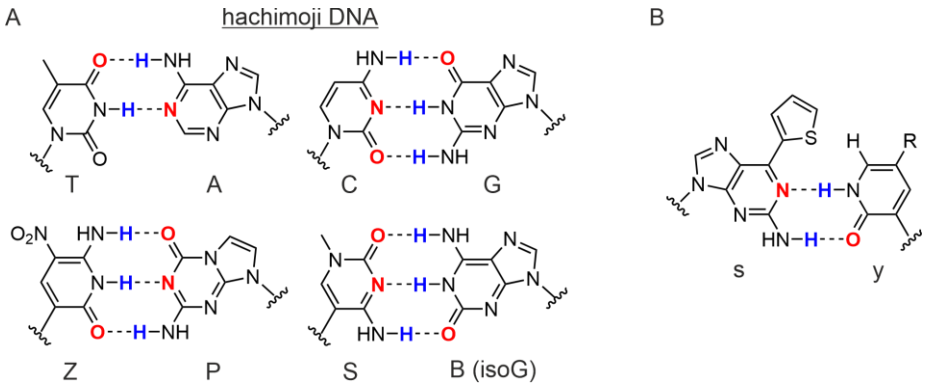


Figure 1.19 (A) The hachimoji DNA base pairs and (B) artificial base pair of the sterically demanding base s and the pyridone analog y.

The Hirao group introduced a sterically demanding group, such as a thiophene at the 6-position of the 2-amino purine to create the base s, which have a steric and electrostatic clash with the 4-keto group of T. To complete the base pair, the pyridone analog y (see Figure 1.19B) was developed and further optimized with a dehydroimidazolone scaffold to prevent mispairing of y with A.[266-267] Other approaches, such as hydrogen bonding with a nucleotide in the noncanonical syn-glycosidic conformation[268] or alkynylated purine and pyridazine[269] have been used to extend hydrogen bonding recognition.

1.4.2.2 Hydrophobic shape complementary

The hydrophobic character of the nucleobases is essential for the stabilization of the DNA duplex due to the hydrophobic effect. In contrast, the specific incorporation of nucleotides by polymerases depends on the complementary hydrogen bonding of the nucleobases. Kool and coworkers showed that difluorotoluene (F), a thymine shape mimic, could be specifically incorporated by the Klenow Fragment opposite to A.[270] 9-Methyl-1-*H*-imidazo[(4,5)-b]pyridine (Q) is an optimized version of 4-methylindole. Q acting as an adenine shape mimic and can form a selective unnatural base pair with F. In addition, the interaction with the polymerases has been improved compared to 4-methylindole. [271] Hirao went on to develop unnatural base pairs based on the size of a natural purine-pyrimidine pair (see Figure 1.20A).[272] An assumed steric clash of hydrogen bonds in the Q/F base pair was the basis for the first modification. The F base was replaced by an aldehyde-containing pyrrole, resulting in higher selectivity.[273]

Further optimization has been achieved by the use of the unnatural nucleobase Ds, which is derived from the more sterically demanding base s.[274] An improved version of the pyrrole nucleobase was generated by replacing the aldehyde with a nitro group and attaching a propynyl group to the aromatic ring, resulting in the nucleobase analog Px. The Px/Ds pair has a selectivity of 99.9% and can be labeled with an additional group on the alkyne moiety of the Px.[275-276] This behavior could be used to directly functionalize RNA, such as the fluorogenic aptamer Spinach, where the DFHBI chromophore is specifically attached to the strand.[277] Structural characterization showed that the base pair Px-Ds pairs in an edge-to-edge manner.[278]

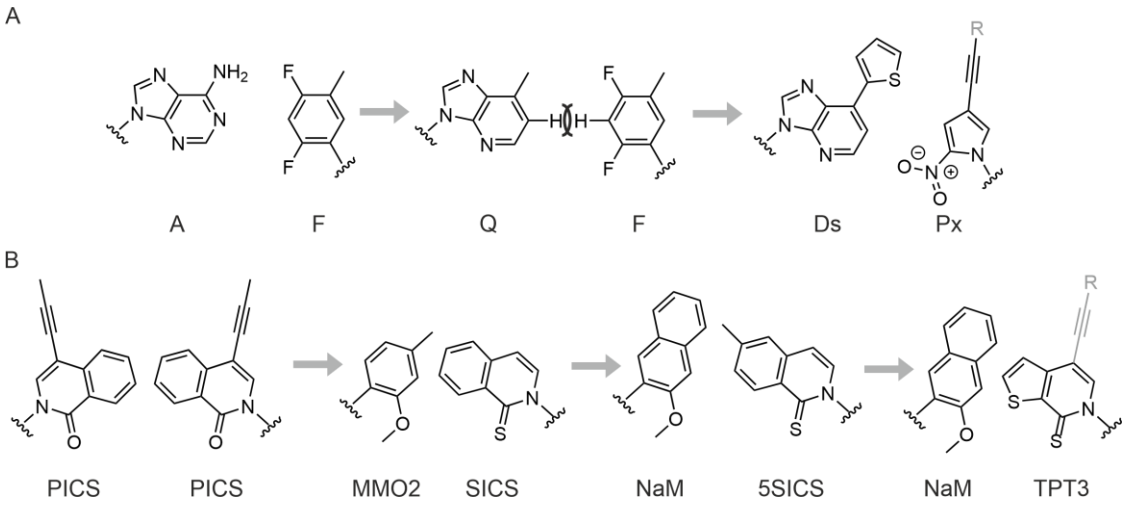


Figure 1.20 Development of the unnatural base pair (A) Ds/Pa and (B) NaM/TPT3. Residues shown in gray are optional for labeling purposes.

Romesberg began the development of unnatural base pairs by synthesizing and evaluating various hydrophobic analogs (see Figure 1.20B).[217] In 1999, his group presented the 7-propynyl isocarbostyril nucleoside (PICS), which could be introduced as homopair by the Klenow Fragment into a DNA duplex.[279] Structural analysis revealed cross-strand intercalation rather than an edge-to-edge pairing as in hydrogen-bonded base pairs or the unnatural base pairs of Kool and Hirao.[280-281] The use of nucleobase analogs with weaker intercalating properties, such as benzene, and the introduction of hydrogen bond acceptors with reduced polar character, such as methoxy groups or sulfur atoms, which facilitate the interaction with the polymerase, resulted in the formation of the MMO2/SICS base pair, which was the basis for further developments.[282] Optimization of SICS self-pairing and MMO2 incorporation resulted in the artificial base pair 5SICS-NaM, which is efficiently amplified by PCR using a variety of polymerases.[283] This unnatural base pair was tolerated by *Escherichia coli*, generating the first

semi-synthetic organism with an expanded genetic alphabet.[284] Further optimization yielded TPT3, which forms a base pair with NAM and, interestingly, with *N*3-methylated thymidine.[285-286] The incorporation of a modified amino acid *in vivo* could be realized by using an extended genetic code with dTPT3.[287] The Px, TPT3 and 5SICS also allow the attachment of functional groups via an alkyne linker, which can be used to label RNA.[288]

1.4.2.3 Stacking interaction

Aromatic interactions are crucial for biological systems. Besides the already mentioned stacking interaction of nucleobases to stabilize the DNA duplex, aromatic interactions are also important for the structure and function of proteins and for recognizing ligands in a ligand-protein complex.[289-290] The benzene dimer, which has been studied in detail by theoretical and experimental methods, is the simplest case of arene-arene interaction.[291-292] The benzene dimer can adopt two different arrangements, T-shaped (edge to face) and slip-stacked arrangement (see Figure 1.21A). Dispersion, electrostatic and solvation effects can explain these arrangements. [293] Aromatic rings have a large planar surface area, which is optimal for van der Waals contacts. These van der Waals contacts can be optimized in a slip-stacked arrangement.[294] The T-shaped formation can be explained by the electrostatic interactions of the aromatic rings. Hunter and Sander described aromatic rings as consisting of a positively charged σ-framework between two negatively charged π-electron clouds. The favorable electrostatic quadrupole-quadrupole interaction between the π-electrons and the σ-framework leads to the T-shaped formation of the aromatic rings, while a face-to-face arrangement is disfavored due to the repulsive interaction of the negatively charged π-electrons.[295]

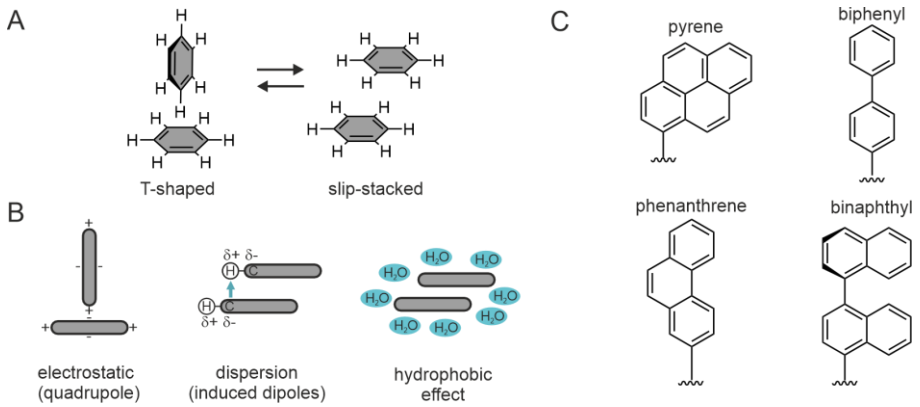


Figure 1.21 (A) Preferred arrangements of a benzene dimer. (B) Interactions that control the arrangement of benzene dimers in solution.[296] Adapted with permission. Copyright 2021 American Chemical Society. (C) Structures of artificial bases forming homopairs on the basis of arene-arene interactions.

Solvation effects play an important role, especially in the case of water, due to the hydrophobic effect, which leads to reduced alignment of the polar water molecules on the hydrophobic surface of the aromatic ring (see Figure 1.21B).

The PICS structure suggests that stacking interactions may play an important role in the formation of artificial base pairs.[280] Based on arene-arene interactions of aromatic molecules, other base pair analogs have been developed (see Figure 1.21C). The Kool group replaced nucleobases with a series of polycyclic aromatic hydrocarbons to study their fluorescence properties and their influence on the DNA duplex structure.[297] A pyrene triphosphate has been selectively incorporated opposite an abasic site.[298] The pyrene deoxynucleoside is approximately the size of an A/T base pair and incorporation opposite of an abasic site showed little thermal stabilization.[299] Christensen and coworkers used a glycerol linker to incorporate two opposing pyrene units into a DNA duplex. NMR structural analysis revealed a stacked conformation of the pyrene units, but distortion of the adjacent base pairs was also observed, resulting in missing imino signals. The incorporation of the pyrene units showed only a slight reduction in thermal destabilization compared to the unmodified duplex.[300] Developing artificial base pairs with interstrand stacking properties, was the focus of Leumann and coworkers. They incorporated *C*-nucleosides containing a bipyridine and biphenyl moiety and observed that a bipyridine homodimer inside a DNA duplex leads to stronger stabilization than an A/T base pair, whereas the biphenyl homodimer destabilizes the DNA duplex around 8 °C compared to the bipyridine homodimer. When the number of consecutive biphenyl homodimers was increased, the melting temperature was higher by 3.0 to 4.4 °C per homodimer. A helical structure of the DNA duplex was still observed after incorporation of three biphenyl homodimers.[301] The same group also introduced 7-functionalized 2-phenanthrenyl-*C*-nucleosides, which show a stacked arrangement of the phenanthrenes in an X-ray structure.[302] The homodimer of the nitro-functionalized phenanthrene has a remarkable melting temperature increase of 11.7 °C compared to its unmodified phenanthrene analog, which is also 5.6 °C higher than a G/C base pair in the DNA context.[303] The Seitz group incorporated binaphthyl pairs into the strands, allowing for intra- and extra-helical stacking interactions. However, the stabilization of the DNA duplex was low.[304]

The disadvantage of these pairs is the formation of homopairs. In order to introduce an additional level of programmability into the base pair, the capability to form a heteropair is

required. This heteropair formation can be achieved by using polar π-stacking interactions. The use of electron withdrawing or donating groups, or the introduction of a heteroatom into the aromatic ring, alters the quadrupole moment of the ring or introduces a stronger dipole moment, which can lead to attractive electrostatic interactions. The Leumann group modified their biphenyl with nitro and methoxy groups and analyzed the melting temperatures of the duplex, with the result that the heteropair of a double functionalized biphenyl moiety showed a slightly higher melting temperature than the homopair containing nitro groups (see Figure 1.22A).[305] The group solved the NMR structure of a duplex containing a methoxy/nitro biphenyl heteropair, showing that the distal phenyl ring of one biphenyl unit and the proximal phenyl ring of the other biphenyl unit are stacked.[306] In primer extension experiments, Rothlisberger *et al.* demonstrated the successful incorporation of the biphenyl triphosphate opposite other biphenyl derivatives by several polymerases.[307] Using acyclic backbones, other groups established systems that stabilize the DNA duplex through the interaction of electron-rich and electron-poor aromatic rings. Häner and coworkers used four electron-rich 1,8-dialkynylpyrenes or electron-poor perylenediimides at the 3' and 5' ends of a DNA duplex, which resulted in higher stabilization compared to the homopairs (see Figure 1.22B).[308] Asanuma's team successfully created stable heteropairs in a DNA duplex that can lead to the formation of a completely artificial heteroduplex without the use of DNA. Pyrene, with its electron-rich aromatic ring, and anthraquinone, with its electron-poor aromatic ring, were used in the system (see Figure 1.22C).[309] Iverson and coworkers were able to successfully implement their system of 1,5-dialkoxynaphthalene and 1,4,5,8-naphthalenetetracarboxylic diimide with a GNA backbone into DNA (see Figure 1.22D).[310] This system was also used to stabilize G-quadruplex structures[311] or hairpins.[312]

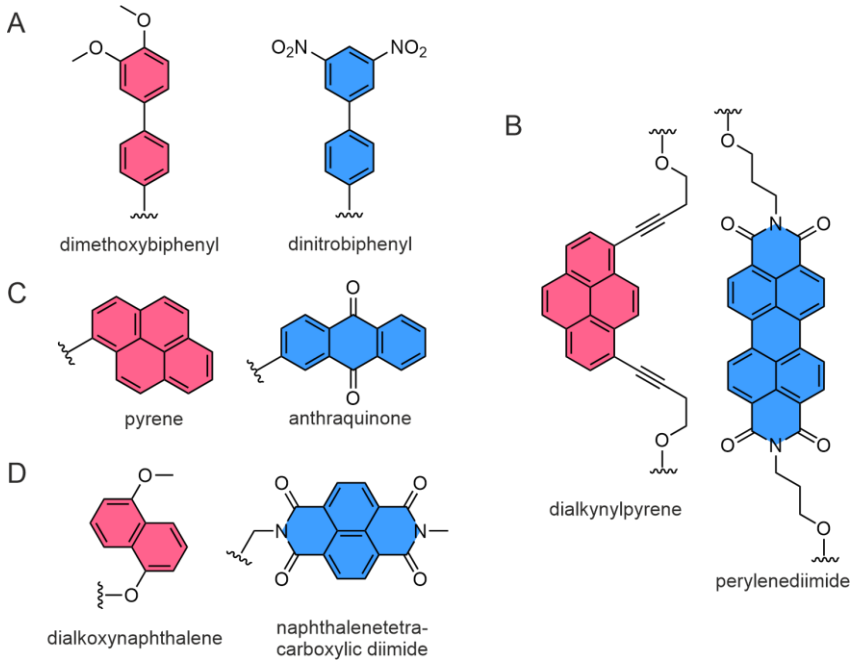


Figure 1.22 Artificial base pairs (A) dimethoxybiphenyl-dinitrobiphenyl, (B) dialkynylpyrene-perylenediimide, (C) pyrene-anthraquinone and (D) dialkoxynaphthalene-naphthalenetetracarboxylic diimide which use polar π-stacking interactions. Electron-rich rings are highlighted in red, while electron-deficient rings are depicted in blue.

The electrostatic properties of an aromatic ring are also altered by the introduction of fluorine atoms, which results in a positive quadrupole moment. Arene-fluoroarene interactions were first observed by Patrick and Prosser in 1960. They reported that benzene and hexafluorobenzene, which are liquids at room temperature, form a solid when mixed in equal amounts at room temperature.[313] Due to their different quadrupole moments and dispersion interactions, benzene and hexafluorobenzene form a slip-stacked arrangement (see Figure 1.23A–B).[314-315] Arene-fluoroarene interactions have been used in a wide variety of applications in chemistry, like organic materials[316-317], polymers[318-319] or crystals.[320-321] Marder used pentafluorophenyl and phenyl residues that are connected for crystal engineering. An equal amount of diphenylacetylene and its fluorinated analog forms crystals in which the fluoroarene and arene adapt a slip-stacked arrangement. The same arrangement was observed for 1-pentafluorophenyl-2-phenylacetylene, where the fluorinated and non-fluorinated phenyl rings of two molecules are stacked on each other.[322] Functionalization with an alkyl aryl ether on the fluorinated and non-fluorinated ring also allows the stacked arrangement in the crystal.[322-323] The precise placement of the molecules is the reason for reactions inside the crystal, like the photopolymerization of diynes[324] or tetrafluoroanthracen.[321] Using arene-fluoroarene interactions as a recognition motif in water is challenging due to the hydrophobic effect, which is the primary driving force for the aggregation of hydrophobic molecules[296]. However, Gao's group

was able to control the dimer assembly of the designed protein α2D through arene-fluoroarene interactions by fluorinating the aromatic ring of one phenylalanine. This leads to the stacking of a phenylalanine with its fluorinated analog in the protein dimer, providing protein stabilization through arene-fluoroarene interactions by 1 kcal mol^{-1}.[325] Partial fluorination of the phenyl ring could further program the recognition of the protein dimers.[326] The polar π-interaction has been used for labeling and purifying proteins, as demonstrated by Pentelute and coworkers. They used a cysteine and proline residue flanked by two phenylalanines to selectively label the proteins with a fluorinated biphenyl.[327] In addition, functionalized resins with pentafluorophenylalanine have been used to purify *N*-terminal tryptophane-containing proteins through cucurbit[8]uril-mediated arene-fluoroarene interactions.[328]

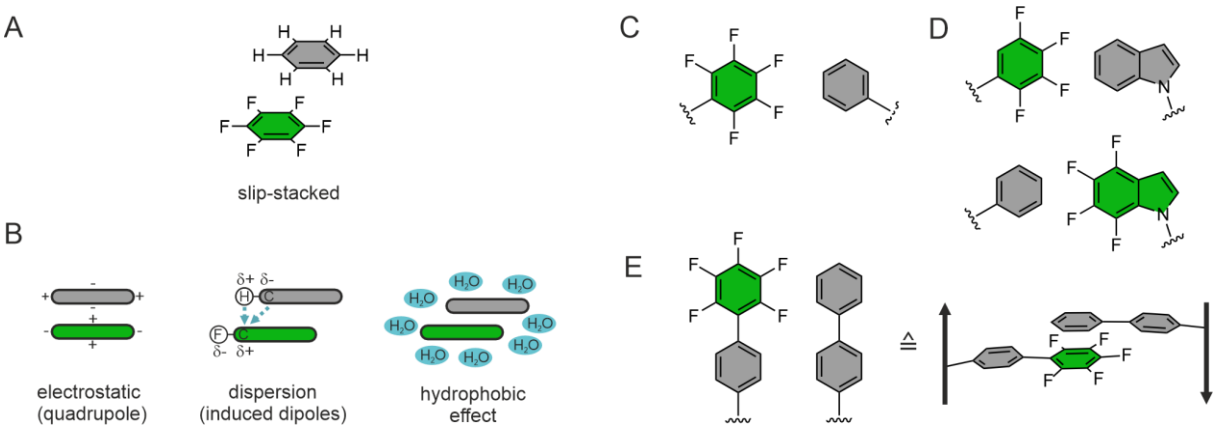


Figure 1.23 (A) Preferred slip-stacked arrangement of a benzene and hexafluorobenzene. (B) Interactions that control the arrangement of benzene and hexafluorobenzene in solution.[296] Adapted with permission. Copyright 2021 American Chemical Society. Structures of artificial pairs (C) pentafluorophenyl-phenyl, (D) tetrafluorophenyl-indole or phenyl-tetrafluoroindole and (E) fluorinated biphenyl-biphenyl based on arene-fluoroarene interactions.

Mathis and Hunziker were the first to report the use of a fluoroarene arene system within DNA.[329] The phosphoramidites of pentafluorophenyl-β-D-deoxyriboside and phenyl-β-D-deoxyriboside were synthesized and incorporated into DNA. No significant stabilization was observed when incorporating phenyl and pentafluorophenyl groups in opposite positions for base pair formation (see Figure 1.23C). However, when four phenyl residues were incorporated consecutively with alternating order of fluorinated and non-fluorinated phenyl rings inside a self-complementary sequence, duplex formation occurred. This formation was not observed when the alternating pattern was changed.[329] Kool and colleagues used 2,3,4,5-tetrafluorophenyl and 4,5,6,7-tetrafluoroindole as nucleobase analogs and their non-fluorinated

analogs but did not observe any arene-fluoroarene interactions. Instead, the base pair of both fluorinated building blocks showed the highest melting point, probably due to their higher hydrophobicity.[330] The biphenyl structure of Leumann's group was modified and its distal ring was fluorinated. This system allows a stacked arrangement of fluorinated and non-fluorinated nucleobase analogs in two different strands (see Figure 1.23D). A stronger stabilization was observed once again in presence of the fluorinated biphenyl homopair due to its higher hydrophobicity.[331] This indicates that the hydrophobic effect plays a dominant role in nonpolar base analogs, but the recognition needs to be adjusted by dispersive and local electrostatic interactions to enable the precise formation of base pairs using artificial building blocks.[217]

1.4.2.4 Metal-mediated base pairs

As early as 1963, Katz proposed that mercury ions may be involved in the formation of non-canonical base pairs between two thymines.[332] Miyake *et al.* demonstrated the existence of a metal-mediated base pair by directly measuring the thermal stabilization of a DNA duplex containing a T-T mismatch and the shifts in the imino proton region of the 1D ^{1}H NMR, achieved by addition of a mercury salt.[333] Same was observed for a cytosine-cytosine mismatch after adding silver ions.[334] Mercury and silver are also capable of mediating other base pairs, such as thymine-cytosine (see Figure 1.24A).[335-336] This was impressively demonstrated by a DNA duplex consisting entirely of silver-mediated base pairs.[337]

Introducing metal-binding ligands can mediate an unnatural base pair using metal ions as artificial nucleobase replacements. Several different non-natural base pairs relying on metal-mediated bonds are known.[338-339] The Shionoya lab presented 5-carboxyuracil, which has the ability to form Watson-Crick base pairs with adenine, as well as metal-mediated base pairs with other canonical nucleobases and with itself using mercury, silver and copper ions (see Figure 1.24B).[340]

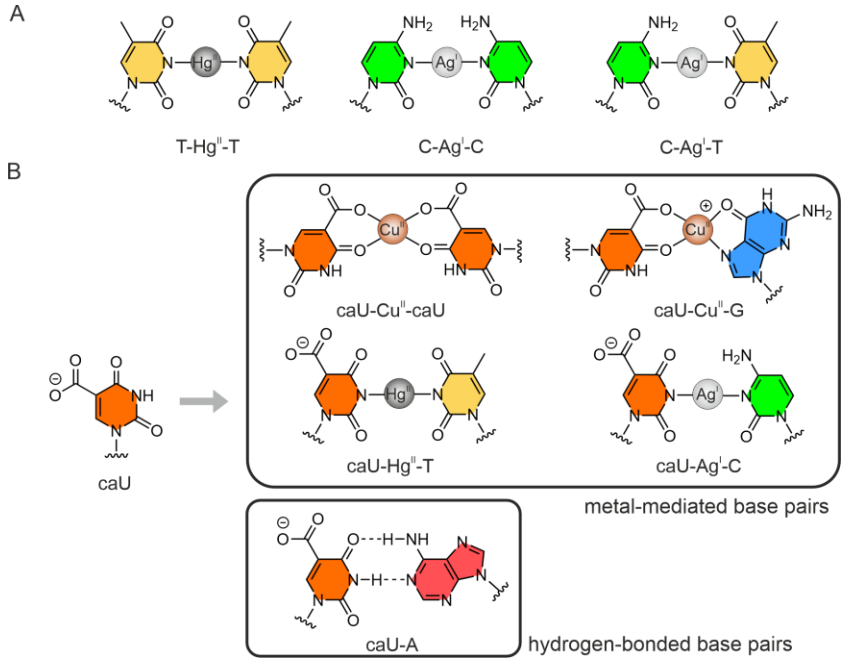


Figure 1.24 (A) Metal-mediated base pairs by mercury and silver ions using natural nucleobases. (B) 5-Carboxyuracil as a modified nucleobase allows specific base pairing to all canonical nucleobases and to itself through metal-mediated base pairs.[340] Adapted with permission. Copyright 2020 American Chemical Society.

Metal-mediated base pair formation can be controlled by the addition and removal of metal ions and is able to adapt to different coordination geometries. These advantageous properties make them optimal for the use in DNA nanotechnology.[341] For instance, a metal-mediated base pair can adapt the function of a toehold by initiating the strand-replacement reaction by adding of the correct metal ion.[342] These properties make metal-mediated base pair formation important for the programmed assembly of DNA structures and conformations.

1.4.3 Functionalized nucleotides

The properties of nucleotides can be expanded by incorporating functional groups and replacing the backbone or nucleobase with a functional moiety, allowing the artificial nucleotide to act as a label or carry a reactive group for further modification.[185, 343] To acquire information about the structure and function of the oligonucleotide, it is crucial to introduce tags that are suitable for respective analytical methods and are not found within natural DNA or RNA.[343] Additionally, these tags can be used as a reporter signal in various applications. Such labels can be fluorine[344] or isotopes[345] for NMR spectroscopy, spin labels for EPR spectroscopy,[346] heavy atoms for X-ray analysis,[343] redox active moieties for electrochemical analysis[347] or, fluorescent labels[218]. Modified nucleotides can be introduced either by solid-phase or enzymatic

synthesis using modified phosphoramidites or NTPs, respectively. Furthermore, by using a nucleotide with a reactive moiety, the modification can be introduced post-synthetically through substitution reactions, addition reactions, cycloadditions or cross-coupling reactions.[185, 348] Apart from labeling with small molecules, crosslinking reactions can be achieved, such as DNA-protein crosslinking or crosslinking of two strands, resulting in interstrand crosslinking or ligation.[219, 349] The following sections discuss the application of functionalized groups in response to photoirradiation, specifically fluorescent labels and photoreactions.

1.4.3.1 Fluorescent labels

Natural oligonucleotides are not fluorescent. This makes fluorescent labels particularly interesting because they can be used without interfering with signals from the system under study.[218] Fluorescent labels can be introduced as fluorescent nucleoside analogs, which can be divided into expanded, isomorphic, chromophoric and extended nucleosides (see Figure 1.25A).[350] A further approach is the direct incorporation of fluorescent dyes using an artificial linker.[351]

Fluorophores can be used for a variety of applications, such as sequence detection. As mentioned before, molecular beacons can be modified by a fluorophore quencher system. Opening the molecular beacon by binding to a complementary sequence results in an increase in fluorescence.[111] This system was further optimized by using different fluorogenic systems.[352] Häner group presented an approach using their electron rich 1,8-dialkynylpyrenes and electron-poor perylenediimide system. Incorporation of these molecules into the stem resulted in a stacked arrangement of the chromophores, preventing dimer formation of the pyrenes. Upon binding to a complementary sequence, the two different chromophores are separated and the pyrenes can form an excited dimer, also called an excimer, which can be detected by its fluorescence.[353]

Pyrene excimer formation was furthermore applied for sequence detection in HCR. The hairpins were labeled with terminal pyrenes. This resulted in an HCR product that, upon addition of an initiator strand, can form the pyrene excimers. This approach allowed detection in picomolar concentrations of the initiator.[354] Sleiman and coworkers reported an HCR approach using thiol-disulfide exchange after HCR product formation to release a fluorophore that showed higher fluorescence after it is unbound.[355]

Point mutations, known as single nucleotide polymorphisms (SNPs), are also detected by changes in fluorescence. Saito and coworkers used extended pyrene moieties as base-discriminating fluorescent nucleobases, which show higher fluorescence when base paired with the correct nucleobase. In the case of a mismatch, the pyrene chromophore intercalates into the DNA and the fluorescence is quenched.[356] Luedtke and coworkers presented a uracil-derived nucleobase analog, which is highly fluorescent if base paired with adenine (see Figure 1.25B). A 28-fold brighter fluorescence intensity was observed compared to a mismatch with C or T. This strong discrimination was achieved by the addition of trans-stilbene as a molecular rotor.[357]

By using the molecular rotor on a cytosine nucleobase, it was possible to distinguish single-stranded, duplex, and i-motif based on brightness and emission wavelength.[358] Srivatsan established the benzofuran extended 2'-deoxyuridine at position C5 as environment-sensitive nucleoside probe. With this moiety, it is possible to distinguish different G-quadruplex topologies or i-motif formation (see Figure 1.25C).[359-360] The fluorescence change of this building block was further used as a probe for DNA polymerase activity.[361]

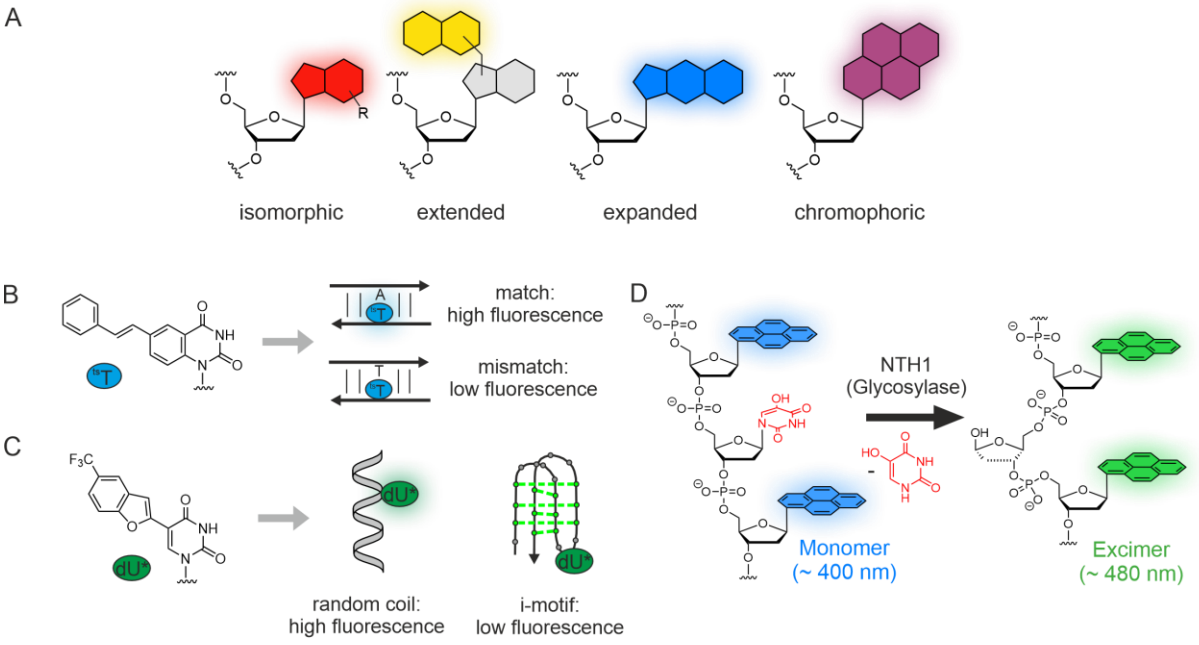


Figure 1.25 (A) Classification of fluorescent nucleoside analogs. (B) Detecting SNPs using a fluorescent molecular rotor.[357] (C) Benzofuran-modified uridine analog capable of detecting multiple DNA structures, including random coil and i-motif.[360] Adapted with permission. Copyright 2022 American Chemical Society. (D) Probe for characterization of glycosylase NTH1 activity based on pyrene excimer formation.[362] Copyright 2020 John Wiley and Sons.

Characterization of DNA processing enzymes using fluorescent reporters is advantageous because it allows direct readout and does not require time-consuming methods such as polyacrylamide gel electrophoresis (PAGE). Wilson and Kool developed an assay to analyze the demethylation of 1-methyladenine (m1A) by ALKBH3. m1A quenches the fluorescence of a pyrene chromophore when both are incorporated next to each other. In this way, an oligonucleotide strand can be generated as a probe for the visualization of ALKBH3 activity *in vivo*.[363] DNA glycosylases were analyzed using the universal base excision reporter (UBER). Nucleobase excision produces a hemiacetal that reacts with a UBER probe containing an aminooxy group, allowing rapid oxime formation. The otherwise non-fluorescent UBER probe becomes highly fluorescent due to the restricted movement caused by the DNA intercalation, allowing characterization of the DNA enzyme by fluorescence change.[364] Another approach using a damaged nucleobase between two pyrene chromophores has been reported for the characterization of DNA glycosylases. Removal of the nucleobase resulted in the formation of a pyrene excimer, which could be monitored by fluorescence changes. This probe has been used in a screening process to find an inhibitor for DNA glycosylase NTH1 (see Figure 1.25D).[362]

In addition to environment-sensitive nucleosides, structures and sequences can be detected by forming of fluorophores in DNA-templated reactions. G-quadruplex formation detection was achievable by using PNA strands functionalized with cyanine dye precursors. When the PNA strands hybridize on a quadruplex structure, it allows the precursors to react to a cyanine dye and produce a fluorescence signal.[365] Prusty and Herrmann utilized a monoiodinated BODIPY compound as a profluorophore. Upon a successful templated Heck reaction, the compound recovered its fluorescent properties, allowing template detection at low picomolar concentration.[366] Seitz and colleagues demonstrated that it is possible to generate a fluorophore with multiple turnovers using a DNA-templated reaction. PNA strands functionalized with positive charges allow unspecific binding mediated by the negatively charged DNA strand. After cleavage of a linker by a photocatalyst, a fluorophore is released and the cleaved PNA strand is replaced by a new probe.[367] Photoreactions are of particular interest because they provide additional control since irradiation is required. The upcoming section will discuss other photoreactions in DNA.

1.4.3.2 Photoreactions in DNA

The need for irradiation in photoreactions provides an additional level of control in biological systems. For instance, studies have demonstrated that a 5'-photocaged mRNA has the capability to regulate protein translation in cells. Irradiation provides both spatial and temporal control, allowing only the irradiated cells to express the target protein.[368]

UV-induced reactions that cause DNA damages are well-known. When exposed to UV irradiation, two pyrimidine nucleobases can undergo a photoreaction to form cyclobutane pyrimidine dimers (CPDs) and 6,4-photoproducts (see Figure 1.26A).[369] By incorporating a photosensitizer into the DNA, Wagenknecht achieved T-T dimerization via energy transfer within the DNA, resulting in a ligated DNA strand.[370-371] Gerling *et al.* stabilized their DNA nanostructures by forming thymine CPDs. This shows that the introduction of CPD crosslinks may be beneficial in the field of DNA nanotechnology.[372] Optimizing CPD formation on terminal thymine oligomers using photosensitizers as additives revealed the potential applications of CPD formation in various fields such as gene silencing, DNA aptamers, or DNA nanotechnology.[373]

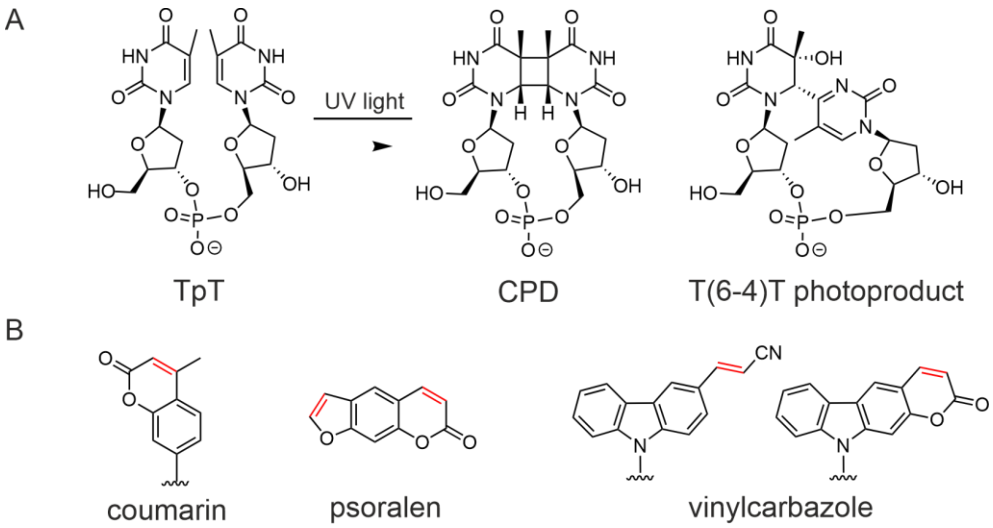


Figure 1.26 (A) Naturally occurring UV-induced DNA damages. (B) Moieties that can react with pyrimidine nucleobases in a [2+2] cycloaddition. Reactive double bonds are highlighted in red.

Rather than utilizing only naturally occurring nucleobases, photo-reacting groups can be inserted into the DNA strands to allow photo-triggered reactions. Upon irradiation, diazirine,[374-375] phenyl azide[376] and phenyl selenide[377] can create reactive species that lead to unspecific covalent crosslinks between DNA strands or other proteins.

Instead of unspecific bond formation, the photoreaction should target a specific functional group of nucleobases or amino acids. An example is the use of *o*-nitrobenzyl alcohol for selective profiling of specific DNA-transcription factor interactions. When exposed to light, an aldehyde that can selectively react with the primary amine of lysine is formed.[378] According to Madder, the furan moiety can form an aldehyde through a [4+2] cycloaddition reaction with singlet oxygen. This aldehyde can then react with a cytosine on the opposite strand, creating an interstrand crosslink.[379] An additional popular method for photo-crosslinking involves using a [2+2] cycloaddition with pyrimidine nucleobases. Known reactive molecules for this are coumarin[380], psoralen[381] or vinylcarbazole,[382-383] allowing interstrand crosslinks (see Figure 1.26B). The 3-cyanovinylcarbazole nucleoside is useful for the attachment of barcodes for cells in sequencing methods[384] and for the stabilization of DNA origami structures[385]. The 5-vinyldeoxyuridine molecule with a reactive motif was utilized for ligation reactions. It was achieved through a [2+2] cycloaddition with the vinyl moiety and C5-C6 double bond of cytosine.[386]

Another approach for photo-crosslinking involves using photocaging groups. Hentschel *et al.* demonstrated the ability to crosslink with other nucleobases by using a photocaged nucleobase containing a crosslink motif based on nucleophilic substitution. After irradiation, the photolabile protecting group is removed, enabling a substitution reaction.[387]

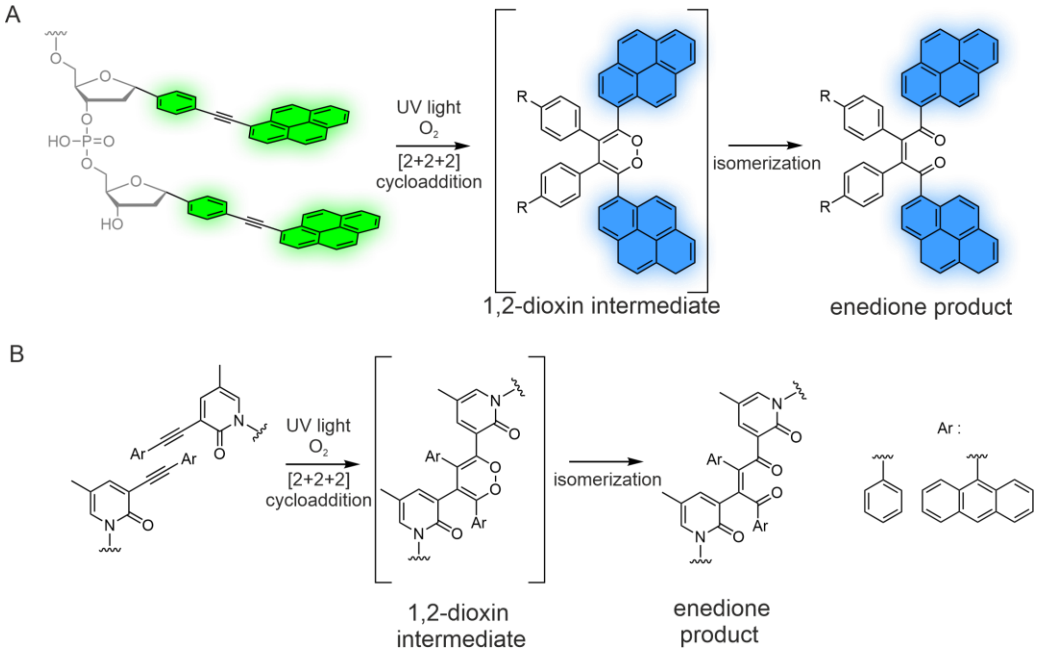


Figure 1.27 Artificial nucleobases that can react in a [2+2+2] cycloaddition. (A) Intrastrand reaction results in fluorescence change due to pyrene excimer splitting.[388] Copyright 2017 John Wiley and Sons. (B) The interstrand reaction leads to the formation of a covalent bond.[389] Adapted with permission. Copyright 2019 American Chemical Society.

DNA can also serve as a template, bringing reactive moieties into the correct position to react with each other. Kashida *et al.* used DNA to assemble two units of a styryl-modified aromatic moiety with an acyclic backbone in a stacked arrangement, allowing an efficient [2+2] cycloaddition.[390] The Nagatsugi group utilized a styryl-functionalized pyridone with a DNA backbone following the same approach.[391] Kool and colleagues demonstrated that a photoinduced [2+2+2] cycloaddition with oxygen can be performed by two consecutive phenylethynylpyrene units that contain a DNA scaffold (see Figure 1.27A). The formation of the enedione product was verified through successful adduct formation with arylhydrazines. Artificial DNA strands containing the bisalkyne moiety were used as a fluorescent dye with color-change properties in cells. This fluorescence change is due to the fact that pyrene excimer formation is not possible after cyloaddition.[388] The Nagatsugi group later used the [2+2+2] cycloaddition to establish interstrand crosslinking (see Figure 1.27B). A stacked arrangement of the used 3-arylethynyl-5-methyl-2-pyridone nucleobases was achieved by bulge formation of the nucleobases located opposite the modification sites. The formation of the product was observed after a few minutes. Mass spectrometric analysis reveals the addition of oxygen, proving the proposed [2+2+2] cycloaddition discovered by Kool. However, species without additional oxygen were also observed. It was suggested that there was an intramolecular cyclization, but the reaction product could not be identified.[389] This circumstance opens the door to further characterization of this type of reaction.

2 Research objectives and outline

DNA is a well-established biomolecule, which can be further exploited in a variety of applications. Beyond its biological function of storing genetic information, it can adapt different structures to perform specialized functions, for example as a catalyst or biomaterial. The well-characterized structure of the DNA duplex, consisting of four nucleotides, was critical to the success of DNA, allowing it to replicate and become programmable. The capability of tuning structural and functional features of DNA would be beneficial in a variety of applications. This gap can be filled by chemically modified DNA. The introduction of bioorthogonal modifications provides an additional level of control over DNA. These modifications can be new reactive functionalities, moieties for spectroscopic characterization, or supramolecular recognition motifs. By introducing these modifications into DNA, a variety of new functions of DNA can be established. Artificial base pairs, enzyme characterization, or the formation of new nanostructures are just a few examples.

The aim of this thesis is to use hydrophobic units to expand the function of DNA. In aqueous environments, hydrophobic interactions play a major role for DNA stability. A controllable quadrupole moment is an additional advantage of aromatic rings, allowing polar π-stacking interactions. The combination of the tolane (*i.e.* diphenylacetylene) motif and arene-fluoroarene interaction creates a new set of recognition motifs orthogonal to the canonical Watson-Crick base pairs. Detailed analysis of the incorporated novel base pair in DNA helps to understand the design principles, allowing further extension of the system independent of the DNA context in the future. In addition to the implementation of a novel hydrophobic base pair, canonical nucleobases can be extended to allow them to adapt a base paired and a stacked conformation with the opposite nucleobases. The programmability of the different conformations helps in the arrangement of functional groups to control their reaction.

A better understanding of stacking interactions through hydrophobic modification in DNA is crucial. Optimizing the requirements for stacking interactions, novel base pair and reaction systems can be designed, resulting in an expansion of DNA properties. This leads to functionalized DNA that can be used to create substrates that enable the formation of biomaterials or

aid in the characterization of enzymes, demonstrating that the understanding of hydrophobic interactions in DNA opens a variety of new tools in nucleic acid chemistry.

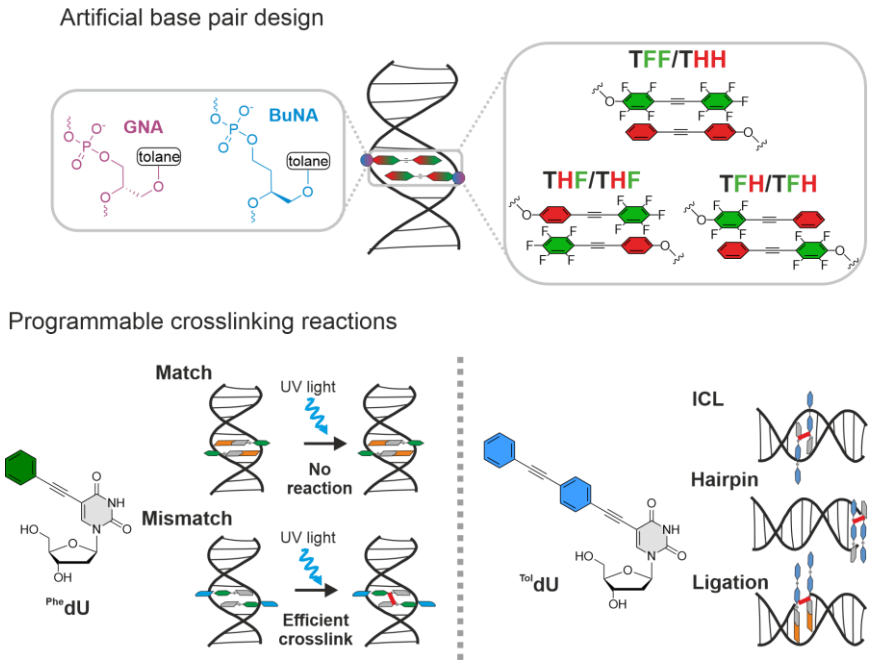


Figure 2.1 An overview of the two projects that are covered in this thesis using aromatic stacking interactions in aqueous solution. The first project uses the tolane moiety as an artificial base pair in the DNA context based on arene-fluoroarene interactions (Chapters 3 and 4). In the second project, a photochemical reaction was programmed by stacking interaction. The novel cross-linking reaction can be carried out in three different topologies. (Chapters 5 and 6).

In the first part of the thesis, the tolane motif is chosen to introduce novel base pairs into a DNA double helix based on polar π-interactions (see Figure 2.1). Chapter 3 describes the development of an artificial base pair using arene-fluoroarene interactions. For this purpose, a perfluorinated tolane is used which possesses an opposite quadrupole moment compared to the non-fluorinated tolane. By adding different acyclic backbones, the tolane building blocks are integrated into DNA strands using phosphoramidite chemistry. Thermodynamic and NMR spectroscopic analyses are used to characterize the differences using two acyclic backbones, revealing the butyl nucleic acid backbone as a privileged backbone surrogate for the implementation of polar π-interactions in the DNA context by a tolane-perfluorotolane pair.

The concept of the use of tolanes as nucleobase surrogates will be further developed in Chapter 4. Partial fluorination of the tolane units provides additional self-recognizing artificial base pairs, complementing a supramolecular language based on arene-fluoroarene interactions in the DNA double helix. It turns out that the homopair of BTHF, which is a combination of the

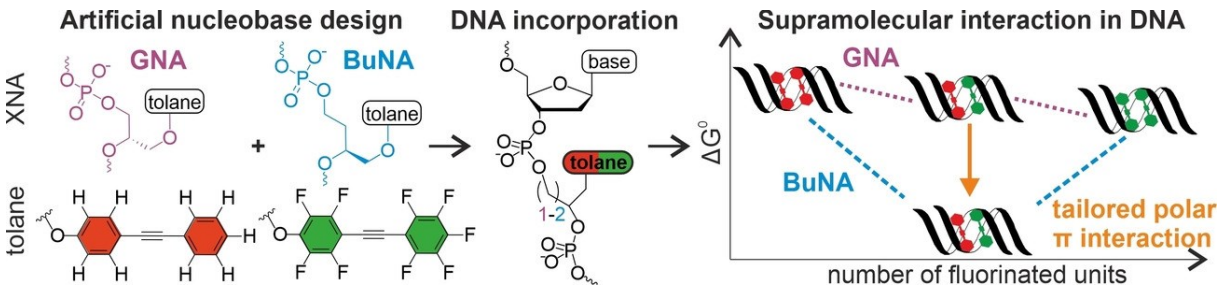
butyl nucleic acid backbone and the tolane analog with an outer fluorinated ring, can strongly stabilize a DNA double helix. The properties of BTHF as a recognition moiety in water were tested by multiple incorporation into a DNA strand.

In the second part of the thesis, a hydrophobic recognition unit was added to natural nucleobase uracil to enable stacking interactions between two opposite modified nucleobases. Such a molecule has been realized by the artificial nucleoside ^{Phe}dU, which combines a hydrophobic phenylacetylene unit with 2'-deoxyuridine (see Figure 2.1). This approach is described in Chapter 5. Irradiation with UV light induced the formation of a covalent bond. The programmability and characterization of the crosslinking reaction is shown, as well as its use as a blocker of helicase DNA unwinding. A detailed analysis of the reaction is performed, which reveals an alkene-alkyne [2+2] cycloaddition for the formation of the covalent bond.

Expansion of the hydrophobic aromatic ring of ^{Phe}dU with another phenylacetylene moiety yielded the artificial nucleoside 5-[4-(phenylethynyl)phenylethynyl]-2'-deoxyuridine (^{Tol}dU, Figure 2.1), which also exhibited fluorescent properties and is described in Chapter 6. ^{Tol}dU is capable of crosslinking two different DNA strands in three architectures. This reaction leads to an increase in fluorescence, which enables the detection of specific sequences by a hybridization chain reaction.

Finally, the results of this thesis will be summarized. Furthermore, an outlook on the future application of the developed systems will be given.

3 Tailored tolane-perfluorotolane assembly as supramolecular base pair replacement in DNA



This chapter and the corresponding supporting information were published in:

H. Neitz, I. Bessi, V. Kachler, M. Michel, C. Höbartner, *Angew. Chem., Int. Ed.,* 2023, **62**, e202214456. (https://onlinelibrary.wiley.com/doi/10.1002/anie.202214456)

Published by Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim.

Abstract: Arene-fluoroarene interactions offer outstanding possibilities for engineering of supramolecular systems, including nucleic acids. Here, we implement the tolane-perfluorotolane interaction as base pair replacement in DNA. Tolane (THH) and perfluorotolane (TFF) moieties were connected to acyclic backbone units, comprising glycol nucleic acid (GNA) or butyl nucleic acid (BuNA) building blocks, that were incorporated via phosphoramidite chemistry at opposite positions in a DNA duplex. Thermodynamic analyses by UV thermal melting revealed a compelling stabilization by THH/TFF heteropairs only when connected to the BuNA backbone, but not with the shorter GNA linker. Detailed NMR studies confirmed the preference of the BuNA backbone for enhanced polar π-stacking. This work defines how orthogonal supramolecular interactions can be tailored by small constitutional changes in the DNA backbone, and it inspires future studies of arene-fluoroarene-programmed assembly of DNA.

3.1 Introduction

Noncovalent interactions are crucial in chemical and biological systems for driving structural organization as well as for controlling molecular recognition. Xeno nucleic acids (XNA) with artificial backbones[240, 392] and unnatural base pairs (UBPs)[217, 272, 281] that exploit alternative hydrogen bonding patterns,[265] metal coordination[338, 393] or hydrophobic interactions[10, 282] are heavily investigated for applications in nucleic acid nanotechnology as well as chemical and synthetic biology.[394-395] Examples reported by Leumann,[305] Häner,[308] Iverson,[310] and Asanuma[309] used electron-rich and electron-deficient aromatic systems to control the assembly of DNA strands. In contrast to the majority of known unnatural base pairs, which rely on hydrogen bonding or hydrophobic shape complementarity, in this work we exploit the stabilizing electrostatic interaction of opposite-sign quadrupole moments in arenes and fluoroarenes, which has previously been harnessed for the controlled assembly of organic materials,[317, 396-397] polymers,[318-319, 398] and crystals.[320-321, 399-400] Using arene-fluoroarene interactions in aqueous solution is challenging due to the weak affinity and the hydrophobic effect, which can lead to uncontrolled aggregation.[296, 401] Nevertheless, several supramolecular systems using arene-fluoroarene interactions in water have been reported,[402] including peptides[326-327, 403] and nucleic acids. In DNA, Kool,[330, 404] Hunziker,[329] Leumann[331] and others explored various fluorinated analogs, in which observed stabilizations were mostly of entropic origin due to the increased hydrophobicity of the fluorinated compound. Alternative structural scaffolds are therefore needed to fully exploit the potential of opposite-sign quadrupole moments and enhanced electrostatic interactions.[325] Thus, the construction of an artificial arene-fluoroarene based recognition element in DNA that integrates well in the DNA double helix remained to be explored.

3.2 Results and discussion

3.2.1 Design of the study

Here, we combine acyclic XNA backbones with aromatic and fluorinated aromatic hydrocarbons and report a bioorthogonal supramolecular recognition motif that serves as a base pair replacement in DNA. We chose the tolane (diphenylacetylene) moiety (THH, see Figure 3.1A) and its perfluorinated analogue (TFF, see Figure 3.1B) as substitutes for Watson-Crick base pairs. The length of the tolane unit provides an excellent fit to the diameter of a DNA double helix,[405-406] and the phosphodiester linkage can be tailored for optimizing the stacking

geometry. We compared glycol nucleic acid (GNA, see Figure 3.1C) with butyl nucleic acid (BuNA, see Figure 3.1D) units connected to the tolane/perfluorotolane. GNA contains a 1,2-propanediol phosphodiester backbone[249, 251] and formally represents an acyclic version of the threofuranosyl nucleic acid (TNA). Butyl nucleic acid (BuNA)[258] was introduced as an acyclic mimic of the ribose backbone and contains one additional methylene unit in the linker compared to GNA. To comprehensively evaluate the arene-fluoroarene interactions in the DNA context, the XNA tolane units were placed within the synthetic DNA dodecamer duplex shown in Figure 3.1E at position X7 and Y18 (see Table 3.1 and S3.1).

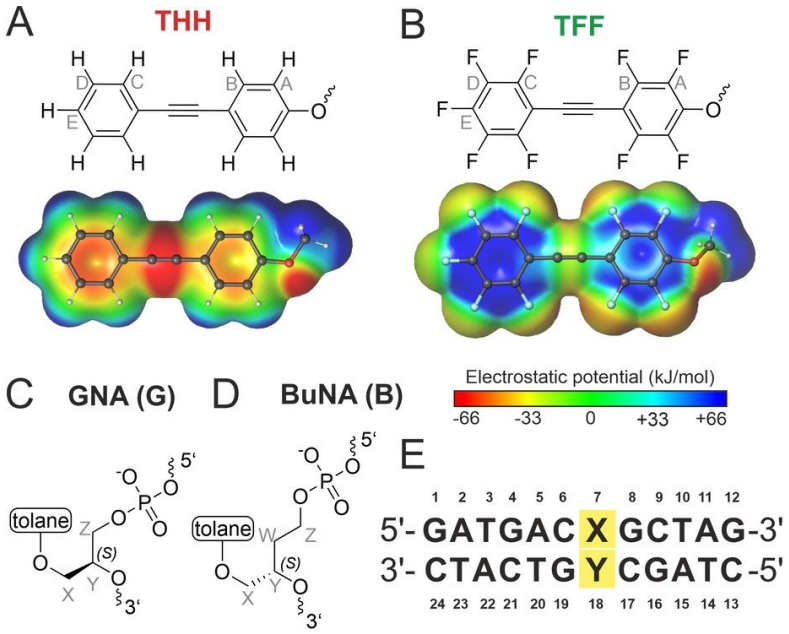
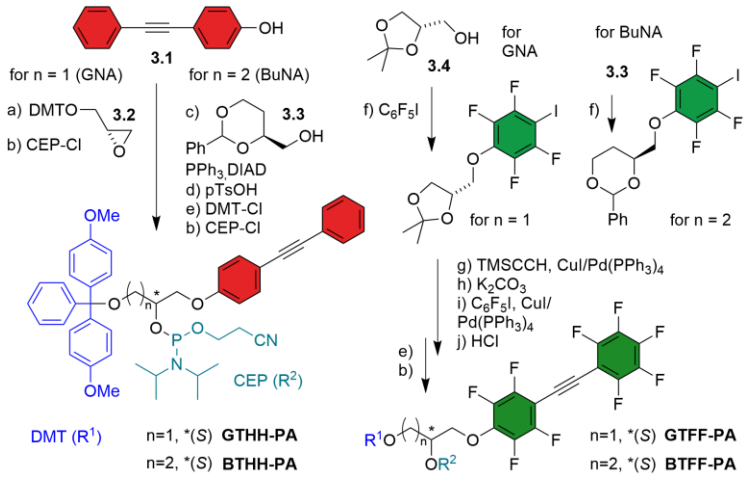


Figure 3.1 Structure of (A) tolane (THH) and (B) perfluorotolane (TFF) ether, with respective electrostatic potential surface. Structure of the acyclic backbones (C) GNA and (D) BuNA. Nomenclature for NMR assignment is given in grey letters. (E) Sequence of the DNA duplex used in this study, modified position X7 and Y18 highlighted in yellow.

3.2.2 Synthesis

Phosphoramidites (PA) of the THH building blocks were prepared from 4-(2-phenylethynyl)phenol (**3.1**) (see Scheme 3.1). The GNA analogue GTHH was synthesized via a base-catalyzed ring-opening reaction of DMT-protected (*R*)-glycidol (**3.2**). The BuNA variant BTHH was obtained via a Mitsunobu reaction with 4-(*S*)-hydroxymethyl-2-phenyl-1,3-dioxan (**3.3**). The syntheses of the fluorotolanes BTFF and GTFF started with the nucleophilic aromatic substitution of iodopentafluorobenzene with **3.3** or D-((+))-solketal (**3.4**), followed by two Sonogashira reactions to install the acetylene unit and attach the second pentafluorophenyl ring.

The building blocks were transformed to the DMT-protected 2-cyanoethyl *N*,*N*-diisopropyl phosphoramidites, which were used for solid-phase DNA synthesis (analytical anion exchange HPLC and HR-ESI data in Supporting Information).



Scheme 3.1 Synthesis of phosphoramidite building blocks. DIAD=diisopropyl azodicarboxylate, DMT-Cl=4,4'-dimethoxytrityl chloride, CEP-Cl=2-cyanoethyl *N*,*N*-diisopropyl chlorophosphoramidite. For details see Scheme S3.1 and S3.2.

3.2.3 Thermodynamic analysis

The thermodynamic stability of the tolane-containing DNA duplexes was characterized by concentration-dependent UV thermal melting experiments (see Figure 3.2A-B and Table 3.1). The tolane homopair GTHH/GTHH destabilized the duplex by 4.1 °C compared to a T/A base pair at positions X7/Y18. In contrast, the duplex containing the perfluorinated homopair GTFF/GTFF had the same melting temperature (T_m) as the Watson-Crick reference duplex. Surprisingly, the T_m of the duplexes containing GNA heteropairs were in between the T_m of the two homopairs, suggesting that the connection to the GNA backbone was suboptimal and did not allow favorable arene-perfluoroarene interactions to occur. Within the BuNA series, the T_m values were higher than for the corresponding GNA-containing analogs. For the homopairs BTHH/BTHH and BTFF/BTFF, the relative trends in enthalpy and entropy were similar as with the GNA backbone. However, upon incorporation of the arene-fluoroarene heteropairs the expected stabilization was observed: BTHH/BTFF and BTFF/BTHH showed the highest T_m in the series with 46.9 °C and 48.1 °C, respectively. This corresponds to a thermal stabilization of >3 °C compared to the Watson-Crick reference duplex, which is also reflected in the enthalpy and entropy values. With ΔH^0 of −81.8 kcal mol^{-1} for BTHH/BTFF and −83.3 kcal mol^{-1} for BTFF/BTHH, ΔH^0 was circa 10 kcal mol^{-1} more favorable in the BuNA heteropair series,

while the entropic stabilization decreased. As expected, incorporation of a single BTHH or BTFF opposite to a nucleobase showed a destabilization of the duplex (see Figure S3.1).

Table 3.1 Melting temperatures and thermodynamic parameters of DNA duplexes shown in Figure 3.1E containing THH and TFF homo- and hetero-pair combinations.

Name[a] X7/Y18	T_m[b] [°C] (ΔT_m[c])	ΔH^{0}[d] [kcal mol^{-1}]	ΔS^{0}[d] [cal mol^{-1} K^{-1}]	ΔG^{298}[e] [kcal mol^{-1}]
Ref DNA T/A	44.8 (0.0)	−92.9 ± 1.5	−263 ± 4.3	−14.5 ± 2.0
GTHH/GTHH	40.7 (−4.1)	−74.5 ± 0.6	−209 ± 1.6	−12.2 ± 0.7
GTHH/GTFF	42.4 (−2.4)	−69.6 ± 0.4	−192 ± 1.2	−12.3 ± 0.6
GTFF/GTHH	43.5 (−1.3)	−73.3 ± 0.5	−203 ± 1.3	−12.9 ± 0.6
GTFF/GTFF	44.8 (0.0)	−68.5 ± 0.5	−187± 1.6	−12.8 ± 0.8
BTHH/BTHH	42.0 (−2.8)	−76.7 ± 0.7	−214 ± 2.1	−12.8 ± 1.0
BTHH/BTFF	46.9 (+2.1)	−81.8 ± 1.1	−227 ± 3.2	−14.1 ± 1.5
BTFF/BTHH	48.1 (+3.3)	−83.3 ± 1.2	−231 ± 3.3	−14.6 ± 1.5
BTFF/BTFF	45.2 (+0.4)	−71.7 ± 0.8	−197 ± 2.3	−13.1 ± 1.1

[a] G (GNA), B (BuNA). [b] T_m at 1 µM DNA, in 100 mM NaCl, 10 mM phosphate buffer, pH 7.0. [c] Difference to T_m of reference. [d] Derived from van't Hoff analyses with five concentrations (1, 2, 5, 10, 20 µM; Table **S3.2**). [e] $\Delta G^{298}=\Delta H^{0}-T\Delta S^{0}$. Calculated at 25 °C.

Using these data in a chemical double mutant cycle[407] allowed us to disentangle the thermodynamic contributions of the interactions between the tolane units and the neighboring DNA (see Figure 3.2C). Starting from the THH/THH homopair, we focused on the changes of free energy upon fluorination and defined $\Delta\Delta G$ for the tolane units in the heteropair as $\Delta\Delta G_{het}$, and for each fluorinated tolane unit with the neighboring DNA as $\Delta\Delta G_{X7}$ and $\Delta\Delta G_{Y18}$, respectively. The fluorinated tolane at position X7 had a stabilizing effect on the DNA, but not at position Y18: $\Delta\Delta G_{X7}$ (−0.6 kcal mol^{-1} for GNA and −0.4 kcal mol^{-1} for BuNA) in comparison to $\Delta\Delta G_{Y18}$ (0.0 kcal mol^{-1} for GNA or +0.1 kcal mol^{-1} for BuNA). For the GNA variant a negligible $\Delta\Delta G_{het}$ of −0.1 kcal mol^{-1} was found (see Figure S3.2). The gain in stability upon introduction of the BuNA heteropair was significantly larger, with $\Delta\Delta G_{het}$ of −1.4 kcal mol^{-1} (see Figure 3.2C). A similar stabilization was observed in a duplex containing a neighboring A/T base pair instead of G/C (see Figure S3.3). These values are on the same order of magnitude as previously estimated for the contribution of a phenylalanine-pentafluorophenylalanine interaction in a peptide-based system.[325] Thus, the thermodynamic analysis shows that the BuNA backbone is the favored connection for the tolane-fluorotolane heteropair.

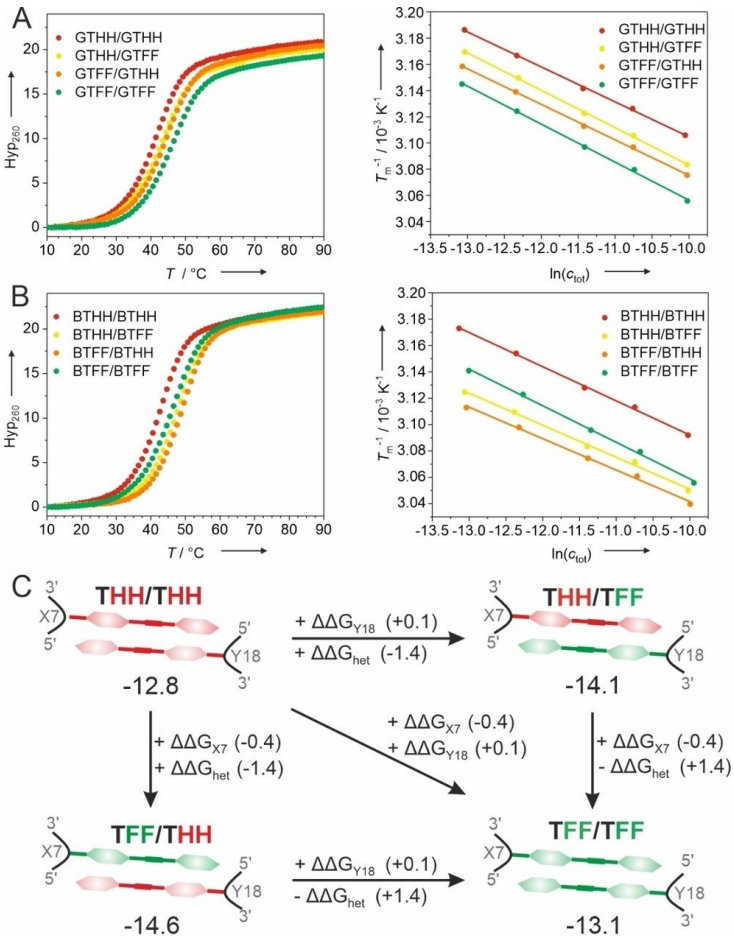


Figure 3.2 UV-melting curves (left) and van't Hoff analysis (right) for DNA duplexes containing (A) GNA and (B) BuNA tolane homopairs and heteropairs. Hyperchromicity at 260 nm, 1 µM DNA in 10 mM sodium phosphate pH 7.0, 100 mM NaCl. (C) Schematic representation of the double mutant cycle. ΔΔG values are given for the BuNA series.

3.2.4 Kinetic analysis

FRET-based DNA strand displacement experiments[408] were designed to investigate the preference for heteropairing in competition with homopairing (see Figure 3.3A, Figure S3.4, S3.5 and Table S3.3). First, an unlabeled strand was added to a Cy3/Cy5 labeled homopair-containing duplex, resulting in a partial strand displacement. The displacement was tracked by monitoring the increase of the Cy3 fluorescence emission intensity, caused by the reduced FRET to Cy5. Then, an excess of a DNA strand with an unsubstituted propyl linker (C_3) was added, allowing the determination of the maximal Cy3 fluorescence and the fraction of total displacement (F_{norm}). The duplexes containing a THH/THH homopair reached a stronger displacement upon addition of a TFF strand than with a THH strand (see Figure 3.3B-C). The BuNA variant showed a higher strand displacement (67.2±4.6 %) than the GNA analog (55.5±2.4 %). The experiments that were performed in the opposite direction, *i.e.* started with a labeled duplex containing the TFF/TFF homodimer, confirmed this observation. The results in Figure S3.5

show that in GNA the displacement with a TFF strand was slightly more effective; however, changing to BuNA led to a higher exchange with THH (70.5±2.4 %) than with TFF (55.1±2.2 %). This again confirms that the BuNA backbone is more favorable for the heteropair formation.

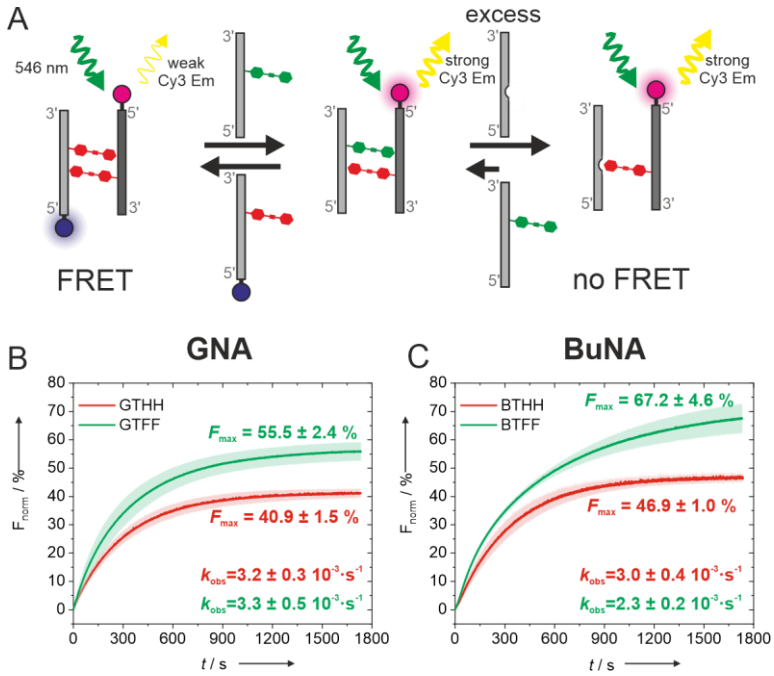


Figure 3.3 (A) Schematic representation of a strand displacement experiment with a THH/THH (red) homopair and an unlabeled strand containing TFF (green). Strand 1 (light grey) is labeled with Cy5 (blue) and strand 2 (dark grey) is labeled with Cy3 (magenta). DNA strand containing C3 linker is represented with a gap. (B,C) Fraction of total displacement (F_{norm}) as a function of time, upon addition of an unlabeled THH (red curve) or TFF (green curve) containing strand to THH/THH homomopair in GNA (B) or BuNA (C).

3.2.5 Structural analysis

A systematic NMR analysis of chemical shift perturbations (CSPs), imino-water exchange rates (k_{EX}) and NOESY cross peaks was conducted on the duplexes GTHH/GTHH, GTFF/GTHH, BTHH/BTHH and BTFF/BTHH, as well as on the reference T/A DNA duplex (see Figure S3.6-S3.19). All the non-terminal imino protons were observed in the ^{1}H 1D NMR spectrum of the modified duplexes at 25 °C, and assigned as indicated in Figure 3.4A. For both GNA and BuNA backbones, THH/THH induced a downfield shift of G8 H1 (blue arrow, see Figure 3.4A) and an upfield shift of G19 H1 (violet arrow, see Figure 3.4A) compared to the reference duplex. In contrast, G8 H1 and G19 H1 were both upfield shifted for the TFF/THH heteropairs in both GNA and BuNA. Importantly, all imino signals near the modification site were detected as sharp signals even at 45 °C (see Figure S3.7), suggesting that no major disruption of the base pairs is required to accommodate the tolane units. This is in strong contrast to the results

observed by Christensen *et al.* for a DNA duplex containing pyrenes linked to a GNA backbone, where the modification induced an overall severe perturbation of the DNA duplex, including base pair disruption.[300] Also comprehensive CSP analyses of aromatic base protons and sugar protons suggest only local perturbation of the duplex (see Figure S3.12). Analysis of the tolane resonances revealed a single peak for pairs even at 10 °C of rotationally symmetric tolane ring protons, which showed that the tolane moieties in the modification site are flipping fast on the NMR time scale (see Figure S3.13A-D). This fast rotation has been reported also for two biphenyl moieties embedded as *C*-nucleosides in a similar DNA duplex.[306] The combined $^{1}H/^{13}C$ CSPs for THH18 next to THH7 or TFF7 were analyzed (see Figure S3.13E-F), and showed comparable trends for changing from a homo- to a heteropair within the GNA and BuNA backbones.

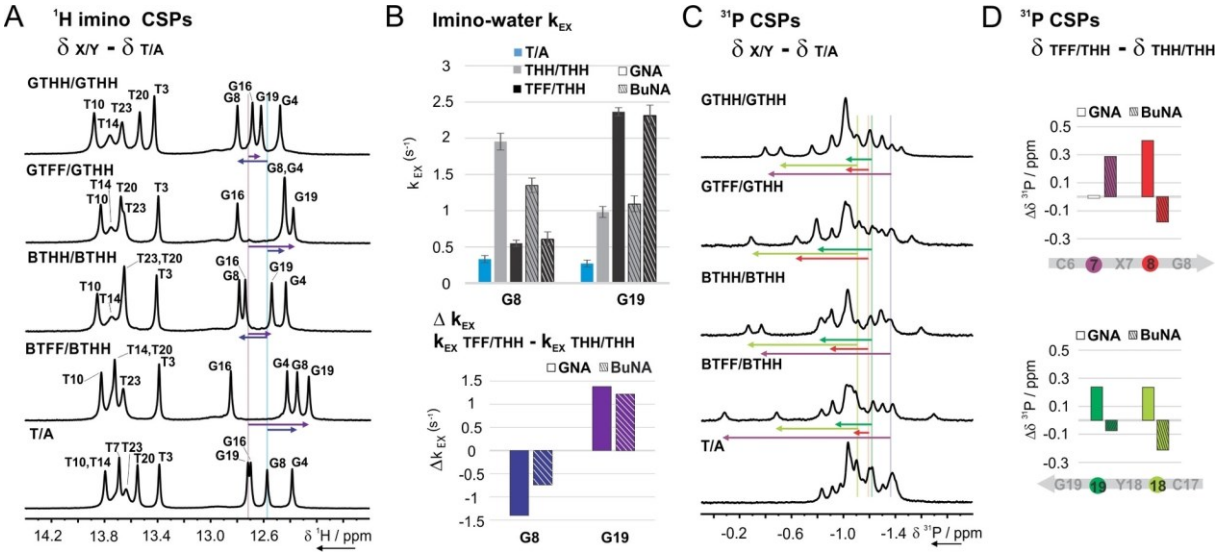


Figure 3.4 (A) Imino region of the 1D ^{1}H NMR spectrum at 25 °C of the duplexes indicated on the left with assignment. Chemical shift perturbation (CSP) of imino protons belonging to residues G8 and G19 is indicated with blue and violet arrows, respectively. B) Imino-water exchange rates (k_{EX}) of imino protons belonging to residues G8 and G19, as determined by CLEANEX-PM experiments at 25 °C (B, top). Comparison of G8 and G19 k_{EX} for homo- and heteropairs embedded in GNA (full bars) or BuNA (dashed bars) backbone (B, bottom). (C) 1D ^{31}P NMR spectra of the DNA duplexes with CSP (vs reference T/A duplex) of selected phosphate units indicated with arrows (see panel D for color code and numbering). (D) ^{31}P CSP (TFF/THH vs THH/THH containing duplexes) of phosphate units belonging to residues adjacent to the modification site in GNA (full bars) and BuNA (dashed bars).

CLEANEX-PM experiments were performed to obtain imino-water exchange rates,[409-410] in order to analyze the duplex stability at the base pair level (see Figure S3.14). Focusing on the C6/G19 and G8/C17 Watson-Crick base pairs flanking the tolane pairs, comparable effects were observed for both GNA and BuNA backbones (see Figure 3.4B). However, the relative influences of THH and TFF were quite distinct. A THH/THH homopair destabilizes both neighboring G−C base pairs to a similar extent, while the TFF/THH heteropair has an asymmetric

effect. A reduced exchange rate of the G8 stacking on TFF comes along with a stronger disturbance (enhanced exchange rate) of G19 flanking the THH moiety.

Next, we evaluated the impact of the tolane modifications on the phosphate backbone by analyzing the ^{31}P chemical shifts (see Figure 3.4C-D and Figure S3.16). The phosphate groups connected to the Z-end of the tolane units (P7 and P18, violet and light green, respectively, in Figure 3.4C-D) showed the largest ^{31}P CSP compared to the reference T/A in all the duplexes (see Figure 3.4C). Upon replacing THH7 with TFF7, the ^{31}P CSPs of the phosphate groups P8, P18 and P19 showed opposite trends depending on the length of the acyclic backbone (see Figure 3.4D): downfield shifts with GNA and upfield shifts with BuNA. Since downfield shifts can be interpreted as an increase of BII/BI population ratio,[411-412] we can conclude that a THH/THH to TFF/THH substitution results in an increased backbone distortion in presence of GNA, but not in presence of BuNA. Thus, the ^{31}P CSP data are in line with the lower global thermal stability observed for the GTFF/GTHH containing duplex, compared to the BTFF/BTHH.

Further insights into the local architecture were obtained by analysis of NOESY cross peaks for the samples BTHH/BTHH and BTFF/BTHH (see Figure 3.5, Figure S3.17-S3.19). The dense network of ^{1}H,^{1}H homonuclear and ^{19}F,^{1}H heteronuclear NOE interactions supports a head-to-tail arrangement of the tolane units in both homo- and heteropair combinations. Inter-tolane NOESY cross peaks were detected between the outer ring (C/D/E) of one unit and the inner ring (A/B) of the other unit. Furthermore, the outer ring of each tolane unit showed several NOE contacts to the butyl linker of the opposite strand as well as to the sugar unit Z-end of the opposite tolane unit, indicating that the tolane moiety spans the complete neighboring base pair. Interestingly, the NOESY spectra of the BTHH/BTHH duplex revealed that each tolane unit had several cross peaks with both GC base pairs flanking the modification site, rather than to only one expected preferred neighboring base pair (see Figure S3.17A, S3.18). As example, Figure 3.5A,C shows NOE contacts between each tolane unit and the imino protons of both G8 and G19. In contrast, such imino-tolane inter-strand cross peaks were not detected for the heteroduplex BTFF/BTHH (see Figure 3.5B,D). Consistent with a polar-π-stacked orientation of the arene-fluoroarene heteropair, medium to strong NOE contacts with only one base pair were observed (see Figure S3.17B, S3.19).

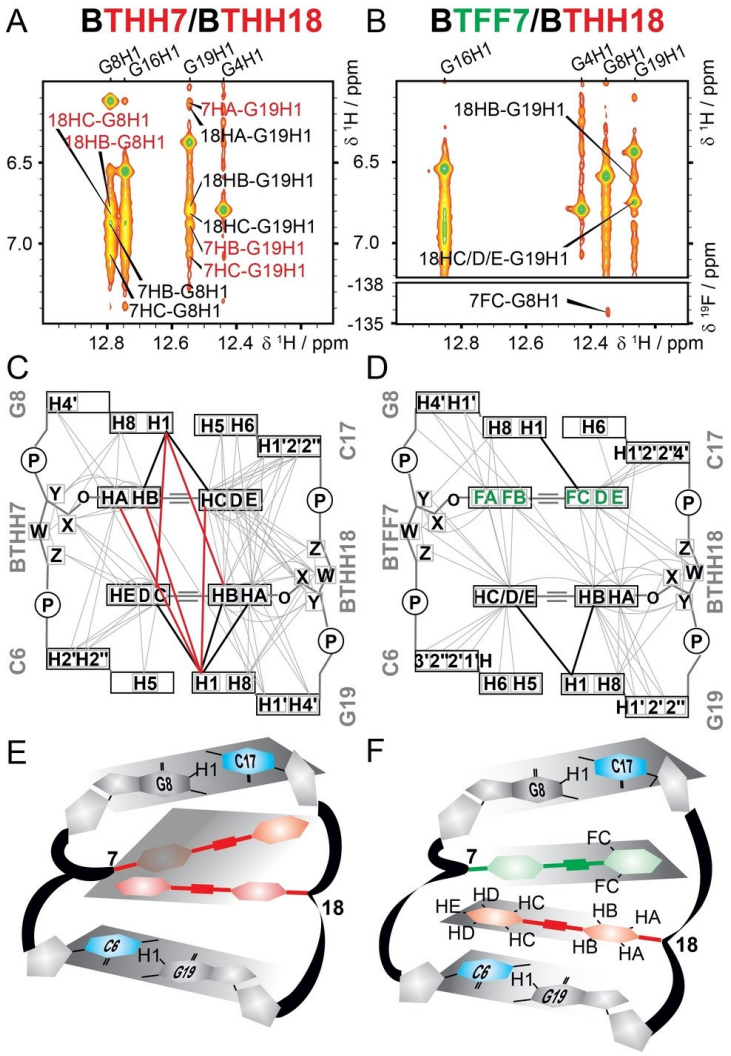


Figure 3.5 (A,B) Imino/aromatic region of ^{1}H,^{1}H NOESY spectrum of BTHH/BTHH (A) and BTFF/BTHH (B, top) and ^{1}H,^{19}F HOESY (B, bottom), with cross peaks between the tolane units and the imino protons of G8 and G19. (C,D) Schematic representation of all the homo- and heteronuclear NOE correlations with the tolane groups observed in ^{1}H,^{19}F HOESY and ^{1}H,^{1}H NOESY spectra. Imino-tolane NOE contacts assigned in panels A–B are shown as bold lines, inter-strand contacts are bold red. Further details are given in Figure S3.17. E,F) Sketches of the corresponding homopair (E) and heteropair (F) regions.

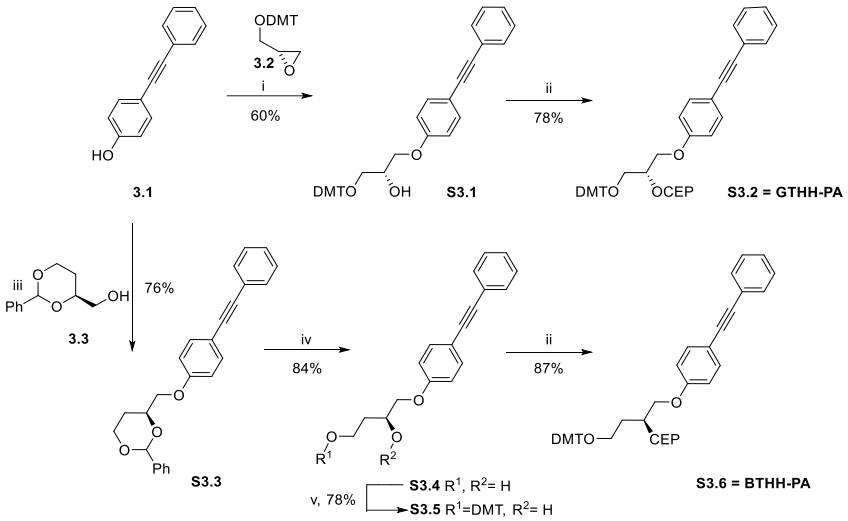
Taken together, the NOE contact map as well as the intensities of the NOESY cross peaks (see Figure S3.17-S3.19) are consistent with the interpretation that a TFF/THH duplex forms a stacked heteropair, while the THH/THH homopair is not arranged in a preferred π-stacked orientation but may experience a more flexible environment (sketch in Figure 3.5E). The combined data for the BTFF/BTHH pair suggest that the inner ring of each tolane unit is stacked to the neighboring guanine base and the outer ring is stacked mostly to the corresponding base paired cytosine (sketch in Figure 3.5F).

3.3 Conclusions

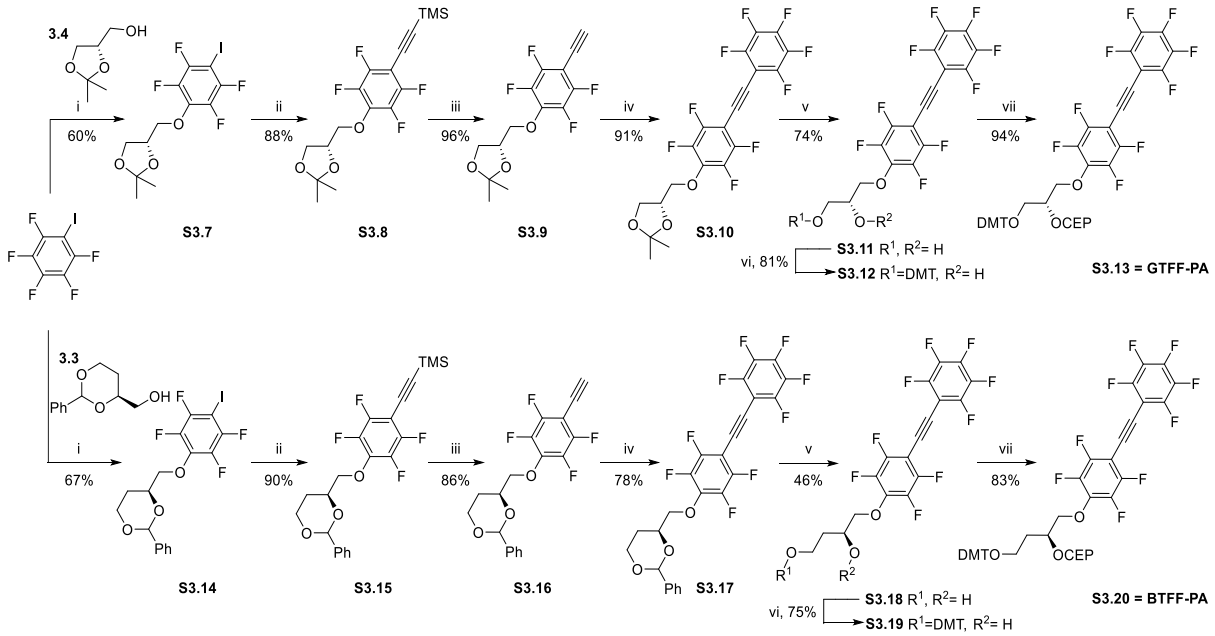
In summary, we introduced a bioorthogonal supramolecular recognition motif based on the tolane moiety, which is driven by polar π-stacking and serves as a base pair replacement in DNA. The interaction energy of a TFF/THH heteropair contributes 1.4 kcal mol^{-1} to DNA duplex stabilization when it is incorporated via a BuNA backbone, but not when it is attached to a GNA backbone. Analysis of a double mutant cycle and imino exchange rate analyses allowed us to disentangle the directional stacking contributions of the fluorotolanes to the neighboring Watson–Crick base pairs. Comprehensive NMR CSP analyses together with NOESY data support the model, in which the tolane moieties of a heteropair are engaged in π-stacking onto each other in a head-to-tail fashion. Consistent with the thermodynamic data, the NMR ^{31}P CSP data revealed that a stronger perturbation of the DNA backbone is required to accommodate a GNA heteropair than for the BuNA heteropair. Thus, even an apparently small modification of the backbone constitution has a large impact on the stacking geometry and the overall duplex stability. While a continuous BuNA backbone is too flexible for stable Watson–Crick base pairing,[258] it is clearly a privileged scaffold for exploiting the supramolecular tolane-fluorotolane interaction in a DNA duplex. Further modifications of the tolane moiety may be used to refine the modes of aromatic interactions, including complementary partially fluorinated tolanes, that could lead to arene-fluoroarene programmed assembly of DNA structures in the future.

3.4 Supporting information

3.4.1 Additional data



Scheme S3.1 Synthesis of the GTHH and BTHH phosphoramidites. i) NaH, DMF, 110°C, 1 d; ii) CEP-Cl, DIPEA, DCM, r.t., 6 h; iii) PPh_3, DIAD, THF, r.t., 30 min; iv) *p*-toluenesulfonic acid, MeOH, r.t., 40 h; v.) DMT-Cl, pyridine, r.t.; 3 h.



Scheme S3.2 Synthesis of the GTFF and BTFF phosphoramidites. i) NaH, THF, r.t., 3 h; ii) trimethylsilylacetylene, CuI, $Pd(PPh_3)_4$, 60°C, 3 h for S3.8; 80°C, 20 h for S3.15; iii) K_2CO_3, THF/MeOH, r.t., 2 h; iv) C_6F_5I, CuI, $Pd(PPh_3)_4$, 80°C, 3 h for S3.10; 80°C, 20 h for S3.17; v) 2 M HCl, THF, r.t., 4-5 d; vi) DMT-Cl, pyridine, r.t.; 23 h; vii) CEP-Cl, DIPEA, DCM, r.t., 2 h.

Table S3.1 Sequences and ESI-MS results of the used DNA oligonucleotides.

Oligo	5'-Sequence-3'	Chemical formula	Calc. Mass	Measured Mass
Reference1	GATGACTGCTAG	$C_{118}H_{148}N_{47}O_{70}P_{11}$	3683.65798	3683.63696
Reference2	CTAGCAGTCATC	$C_{116}H_{148}N_{43}O_{70}P_{11}$	3603.64569	3603.63199
AbasicLinker1	GATGAC(C3)GCTAG	$C_{111}H_{142}N_{45}O_{67}P_{11}$	3517.62069	3517.57859
AbasicLinker2	CTAGC(C3)GTCATC	$C_{111}H_{142}N_{45}O_{67}P_{11}$	3428.59683	3428.57390
GTHH1	GATGAC(GTHH)GCTAG	$C_{125}H_{150}N_{45}O_{68}P_{11}$	3709.67766	3709.68018
Alk-GTHH1	Alk-GATGAC(GTHH)GCTAG	$C_{131}H_{159}N_{45}O_{71}P_{12}$	3869.70659	3869.72739
GTHH2	CTAGC(GTHH)GTCATC	$C_{123}H_{151}N_{38}O_{70}P_{11}$	3620.65434	3620.67986
Alk-GTHH2	Alk-CTAGC(GTHH)GTCATC	$C_{129}H_{160}N_{38}O_{73}P_{12}$	3780.68272	3780.69911
GTFF1	GATGAC(GTFF)GCTAG	$C_{125}H_{141}F_9N_{45}O_{68}P_{11}$	3871.59286	3871.59038
Alk-GTFF1	Alk-GATGAC(GTFF)GCTAG	$C_{131}H_{150}F_9N_{45}O_{71}P_{12}$	4031.62234	4031.62269
GTFF2	CTAGC(GTFF)GTCATC	$C_{123}H_{142}F_9N_{38}O_{70}P_{11}$	3782.56900	3782.58448
Alk-GTFF2	Alk-CTAGC(GTFF)GTCATC	$C_{129}H_{151}F_9N_{38}O_{73}P_{12}$	3942.59848	3942.61320
BTHH1	GATGAC(BTHH)GCTAG	$C_{126}H_{152}N_{45}O_{68}P_{11}$	3723.69331	3723.70657
Alk-BTHH1	Alk-GATGAC(BTHH)GCTAG	$C_{132}H_{161}N_{45}O_{71}P_{12}$	3883.72224	3883.71342
BTHH2	CTAGC(BTHH)GTCATC	$C_{124}H_{153}N_{38}O_{70}P_{11}$	3634.66944	3634.66984
Alk-BTHH2	Alk-CTAGC(BTHH)GTCATC	$C_{130}H_{162}N_{38}O_{73}P_{12}$	3794.69837	3794.68923
BTFF1	GATGAC(BTFF)GCTAG	$C_{126}H_{143}F_9N_{45}O_{68}P_{11}$	3885.60851	3885.61827
Alk-BTFF1	Alk-GATGAC(BTFF)GCTAG	$C_{132}H_{152}F_9N_{45}O_{71}P_{12}$	4045.63799	4045.63546
BTFF2	CTAGC(BTFF)GTCATC	$C_{124}H_{144}F_9N_{38}O_{70}P_{11}$	3796.58520	3796.62324
Alk-BTFF2	Alk-CTAGC(BTFF)GTCATC	$C_{130}H_{153}F_9N_{38}O_{73}P_{12}$	3956.61413	3956.62462
T_BTHH	GATGAT(BTHH)GCTAG	$C_{127}H_{153}N_{44}O_{69}P_{11}$	3738.69297	3738.71257
BTHH_A	CTAGC(BTHH)ATCATC	$C_{124}H_{153}N_{38}O_{69}P_{11}$	3618.67453	3618.67141
T_BTFF	GATGAT(BTFF)GCTAG	$C_{127}H_{144}N_{44}O_{69}P_{11}F_9$	3900.60873	3900.62418
BTFF_A	CTAGC(BTFF)ATCATC	$C_{124}H_{144}N_{38}O_{69}P_{11}F_9$	3780.59028	3780.60927

Table S3.2 Thermodynamic data for DNA duplexes in phosphate buffer (100 mM NaCl, 10 mM sodium phosphate, pH 7.0).

Duplex	Oligo	Sequence	c_{total}[a] [µM]	T_m [°C]	ΔH^0 [kcal mol^{-1}]	ΔS^0 [cal mol^{-1} K^{-1}]	ΔG^{298}[b] [kcal mol^{-1}]
T/A	Reference1 Reference2	5'...CTG...3' 3'...GAC...5'	2.0 4.0 10.0 19.7 40.2	44.8 46.3 48.5 49.6 51.5	–92.9±1.5	–263±4.3	–14.5±2.0
GTHH/GTHH	GTHH1 GTHH2	5'...C**(GTHH)**G...3' 3'...G**(GTHH)**C...5'	2.2 4.4 10.8 21.2 43.2	40.7 42.6 45.2 46.7 48.8	–74.5±0.6	–209±1.6	–12.2±0.7

GTHH/GTFF	GTHH1 GTFF2	5'...C(GTHH)G...3' 3'...G(GTFF)C...5'	2.2 4.5 11.0 21.4 44.2	42.4 44.3 47.1 48.9 51.2	–69.6±0.4	–192±1.2	–12.3±0.6
GTFF/GTHH	GTFF1 GTHH2	5'...C(GTFF)G...3' 3'...G(GTHH)C...5'	2.1 4.3 11.0 21.4 44.3	43.5 45.4 48.1 49.8 52.0	–73.3±0.5	–203±1.3	–12.9±0.6
GTFF/GTFF	GTFF1 GTFF2	5'...C(GTFF)G...3' 3'...G(GTFF)C...5'	2.1 4.4 11.0 21.7 44.4	44.8 46.9 49.8 51.6 54.1	–68.5±0.6	–187±1.6	–12.8±0.8
BTHH/BTHH	BTHH1 BTHH2	5'...C(BTHH)G...3' 3'...G(BTHH)C...5'	2.0 4.3 10.9 21.4 44.3	42.0 43.9 46.5 48.1 50.3	–76.7±0.7	–214±2.1	–12.8±1.0
BTHH/BTFF	BTHH1 BTFF2	5'...C(BTHH)G...3' 3'...G(BTFF)C...5'	2.1 4.2 11.3 21.5 44.4	46.9 48.4 51.1 52.4 54.7	–81.8±1.1	–227±3.2	–14.1±1.5
BTFF/BTHH	BTFF1 BTHH2	5'...C(BTFF)G...3' 3'...G(BTHH)C...5'	2.2 4.5 11.4 22.2 45.6	48.1 49.6 52.1 53.6 55.8	–83.3±1.2	–231±3.3	–14.6±1.5
BTFF/BTFF	BTFF1 BTFF2	5'...C(BTFF)G...3' 3'...G(BTFF)C...5'	2.3 4.7 11.8 23.1 48.0	45.2 47.1 49.9 51.6 54.1	–71.7±0.8	–197±2.3	–13.1±1.1
T_BTHH/BTHH_A	T_BTHH BTHH_A	5'...T(BTHH)G...3' 3'...A(BTHH)C...5'	2.0 4.1 10.5 20.1 41.7	40.2 42.0 44.5 45.9 48.0	–78.5±0.5	–222±1.5	–12.4±0.7
T_BTHH/BTFF_A	T_BTHH BTFF_A	5'...T(BTHH)G...3' 3'...A(BTFF)C...5'	2.1 4.4 11.3 21.4 44.2	41.2 43.2 45.6 47.7 49.6	–72.6±0.6	–202±1.8	–12.4±0.8
T_BTFF/BTHH_A	T_BTFF BTHH_A	5'...T(BTFF)G...3' 3'...A(BTHH)C...5'	2.2 4.4 11.2 21.7 44.9	44.1 45.8 48.3 50.1 52.1	–77.1±0.3	–215±0.8	–13.1±0.4
T_BTFF/BTFF_A	T_BTFF BTFF_A	5'...T(BTFF)G...3' 3'...A(BTFF)C...5'	2.0 4.4 11.4 22.3 45.8	39.3 41.5 44.3 46.3 48.5	–67.3±0.1	–186±0.2	–11.7±0.1

[a] Total concentration of DNA strands. Values were calculated with the absorption at 260 nm of the melted duplex [b] Calculated for T = 25 °C.

Table S3.3 Total strand displacement and displacement rate of the FRET exchange experiments. Values were obtained from a fit to first-order reaction kinetics.

Labeled duplex	Unlabeled strand	F_{max} [%]	k_{obs} [10^{-3} s^{-1}]
GTHH/GTHH	GTHH	40.9 ± 1.5	3.2 ± 0.3
GTHH/GTHH	GTFF	55.5 ± 2.4	3.3 ± 0.5
GTFF/GTFF	GTHH	49.5 ± 4.5	1.2 ± 0.1
GTFF/GTFF	GTFF	60.1 ± 4.1	0.8 ± 0.1
BTHH/BTHH	BTHH	46.9 ± 1.0	3.0 ± 0.4
BTHH/BTHH	BTFF	67.2 ± 4.6	2.3 ± 0.2
BTFF/BTFF	BTHH	70.5 ± 2.4	0.7 ± 0.1
BTFF/BTFF	BTFF	55.1 ± 2.2	1.2 ± 0.1

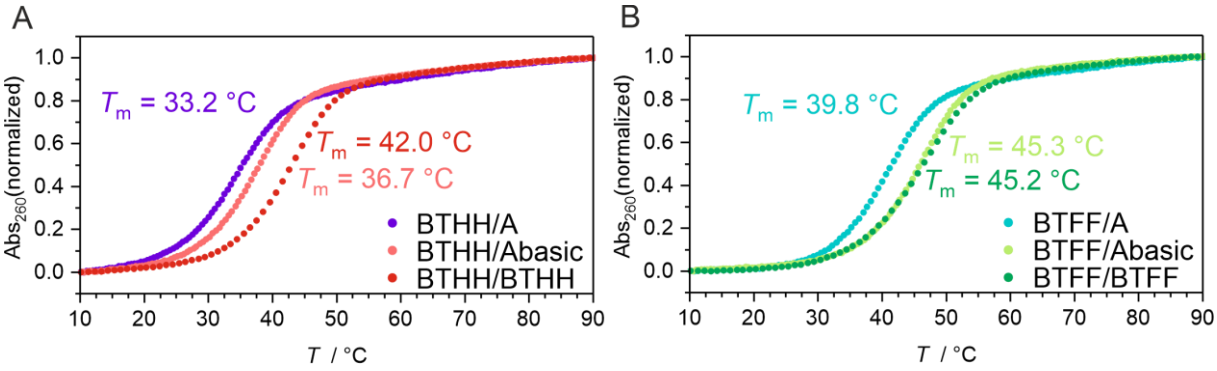


Figure S3.1 UV-melting curves of a DNA duplex with a A) BTHH or B) BTFF unit opposite to the nucleobase A, an abasic linker and another tolane unit.

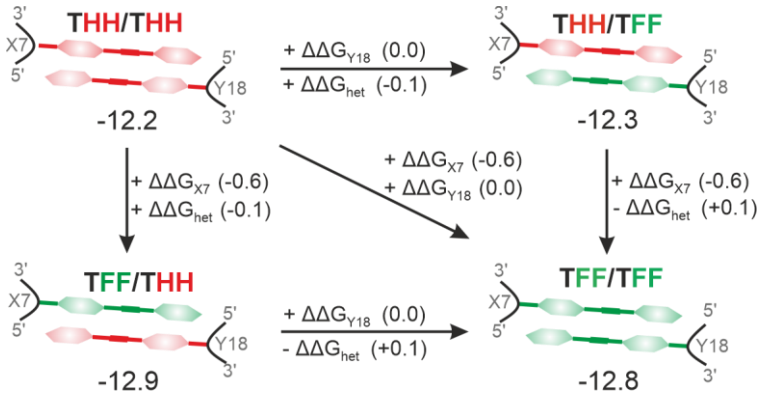


Figure S3.2 Double mutant cycle for GNA series. See Figure 3.2C in the manuscript for BuNA data.

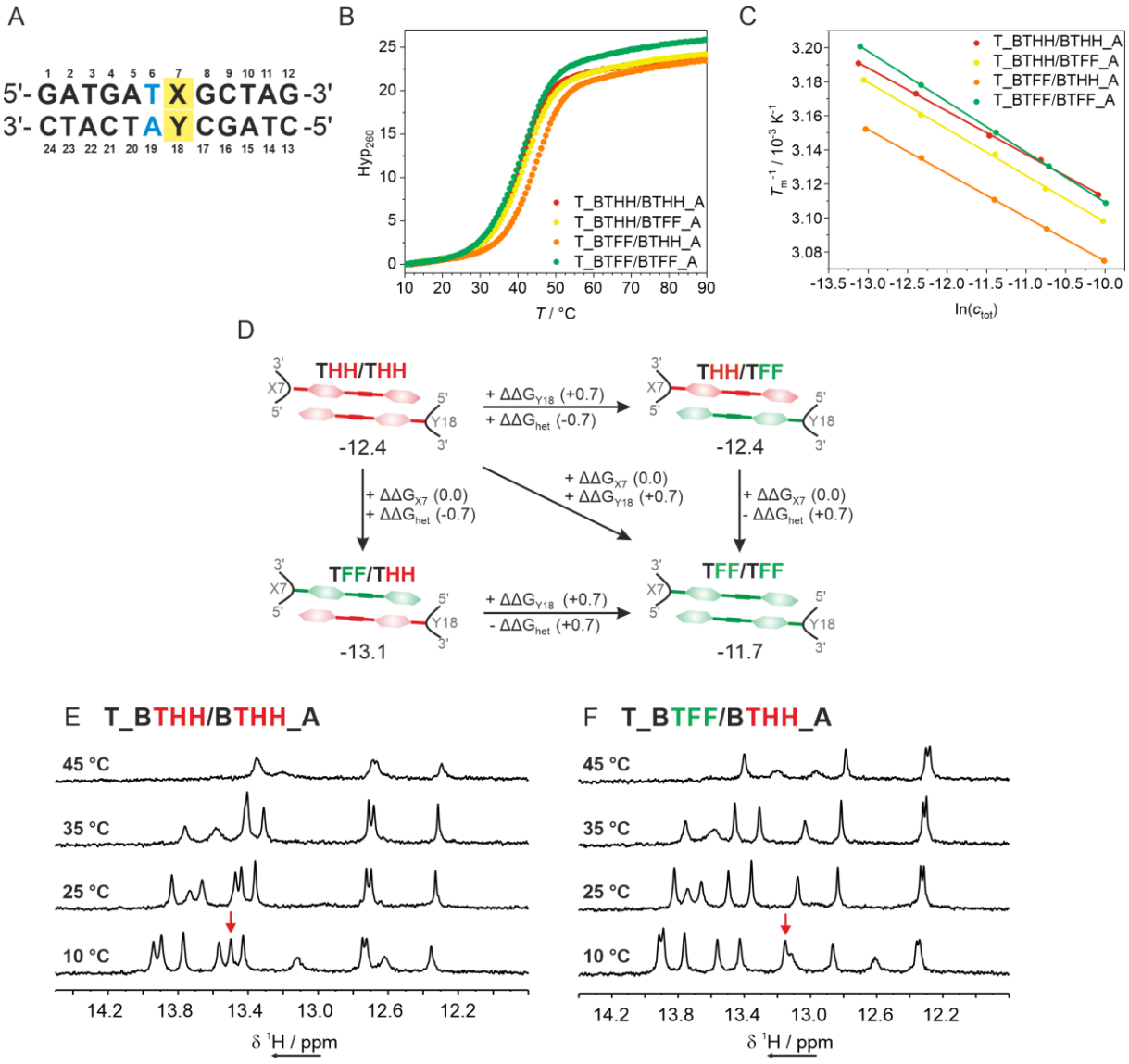


Figure S3.3 (A) Sequence of the DNA duplex with a A/T-base pair next to the modification site (blue letters), modified position X7 and Y18 highlighted in yellow. UV-melting curves at 1 µM duplex concentration (B) and van't Hoff analysis (C) for DNA duplexes containing BuNA tolane homopairs and heteropairs. D) Double mutant cycle for the series with an A/T-base pair replacement. Imino region of the ^{1}H NMR spectrum of T_BTHH/BTHH_A (E) and T_BTFF/BTHH_A (F) at various temperatures. Spectra recorded on 50 µM DNA duplex sample in NMR buffer, 90% H_2O, 10% D_2O, 600 MHz. Red arrow indicates tentative assignment of T6H1 on the spectra recorded at 10 °C.

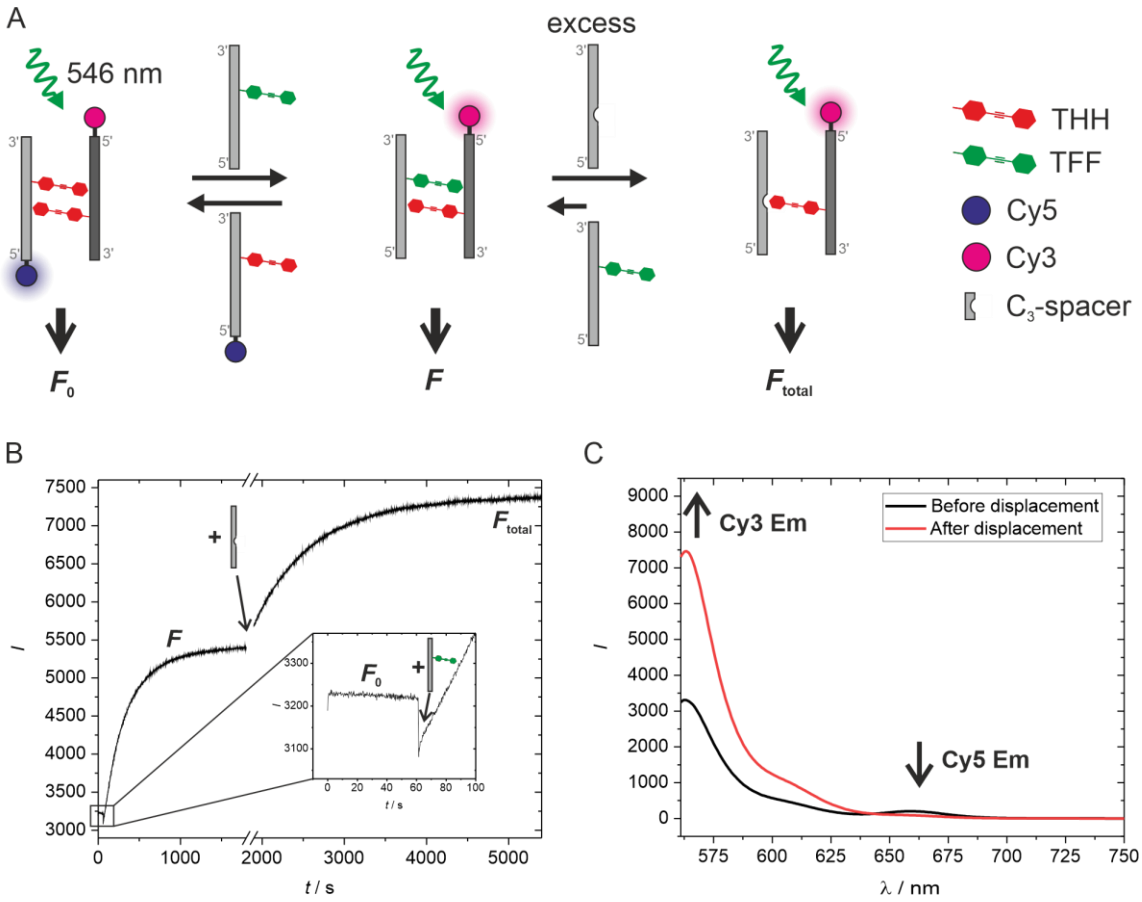


Figure S3.4 (A) Schematic representation of a FRET strand displacement experiment starting with a THH homopair and adding an unlabeled TFF containing single strand. All variations of starting homopairs (THH/THH with TFF/TFF) with different unlabeled single strands (TFF/THH) were tested for both backbones. B) Change in Cy3 fluorescence during the FRET exchange experiment. Injection points of the single strands with TFF or THH modification and C_3-spacer and regions for the estimation of F_0, F and F_{total} are highlighted in the graph. F was normalized according to equation 3.10 and fitted with equation 3.11 to obtain the kinetic parameters F_{max} and k_{obs}. C) Emission spectra before and after complete strand displacement. Change of Cy3 and Cy5 Emission is highlighted.

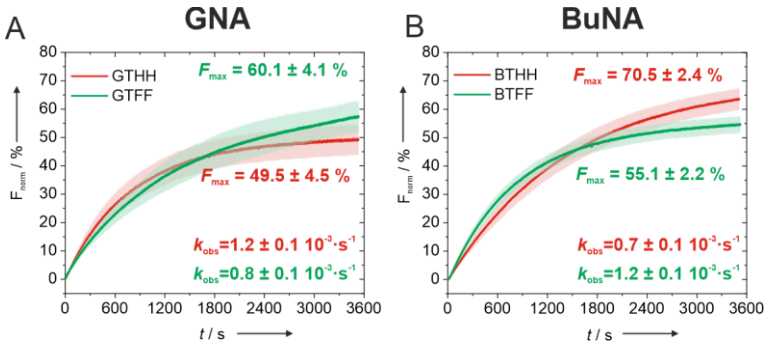


Figure S3.5 Fraction of total displacement (F_{norm}) as a function of time, upon addition of an unlabeled THH (red curve) or TFF (green curve) containing strand to TFF/TFF homomopair in GNA (A) or BuNA (B). See Figure 3.3B,C in the manuscript for THH/THH homomopair data.

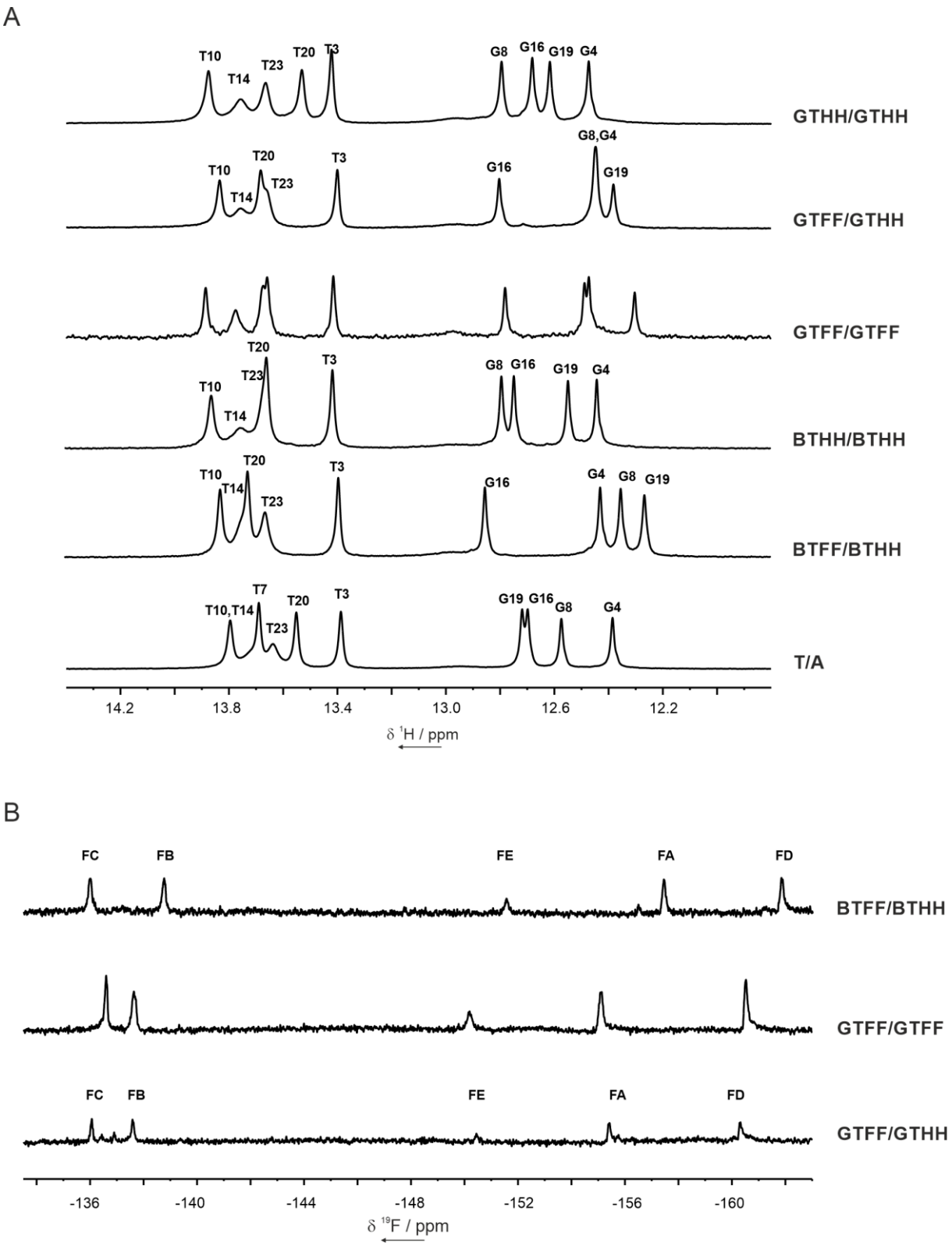


Figure S3.6 (A) Imino region of the 1D ^{1}H NMR spectrum and (B) 1D ^{19}F NMR spectrum of the duplexes indicated on the right. Assignment confirmed by 2D NMR experiment is shown (not available for GTFF/GTFF). Spectra recorded on 0.2-1.0 mM DNA duplex in NMR buffer, 90%H_2O/10%D_2O, 600 MHz, 25 °C.

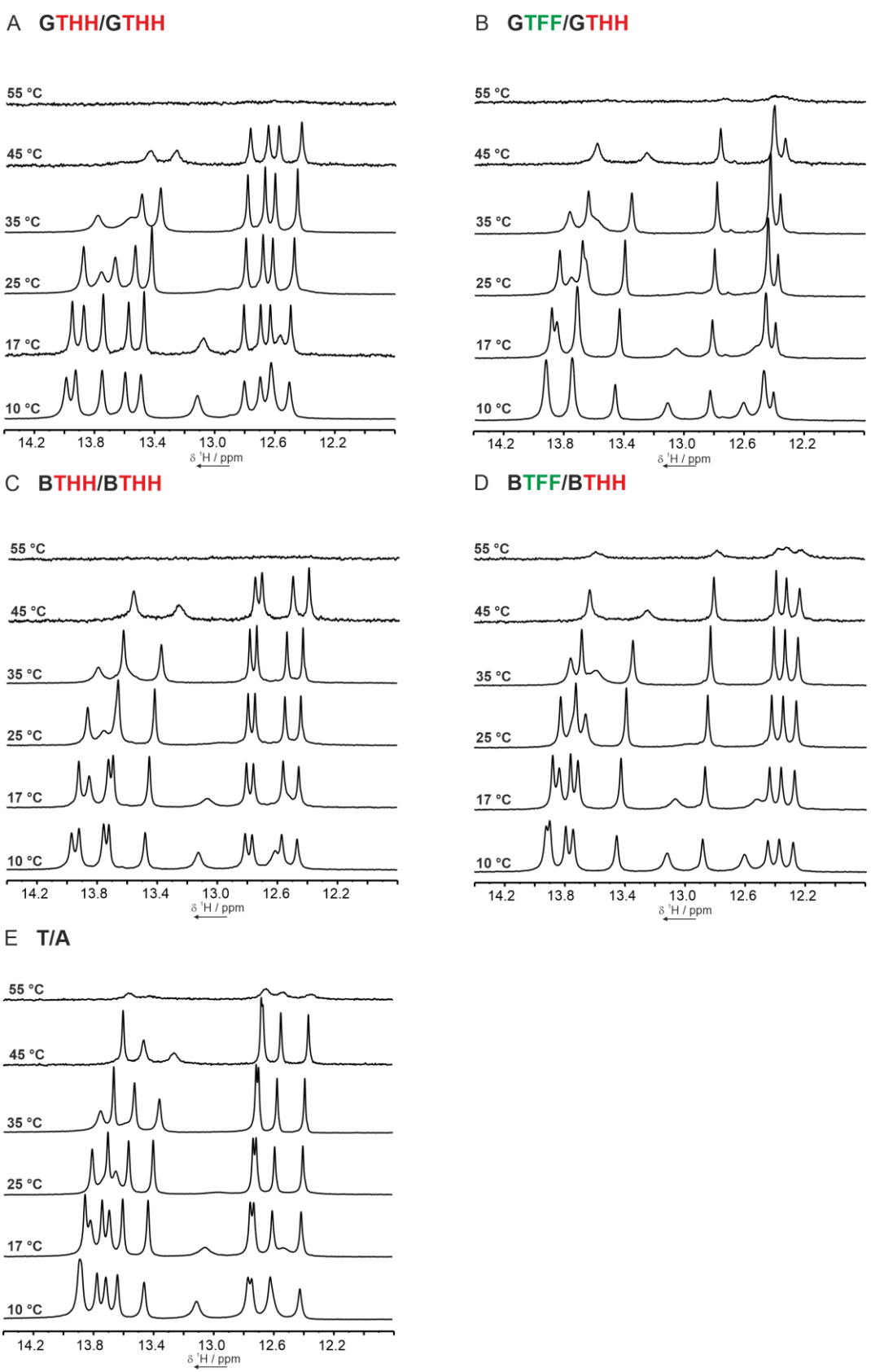


Figure S3.7 Imino region of the 1D ^{1}H NMR spectrum of the duplexes (A) GTHH/GTHH, (B) GTFF/GTHH, (C) BTHH/BTHH, (D) BTFF/BTHH and (E) T/A at various temperatures. Spectra recorded on 0.2-1.0 mM DNA duplex in NMR buffer, 90%H_2O/10%D_2O, 600 MHz.

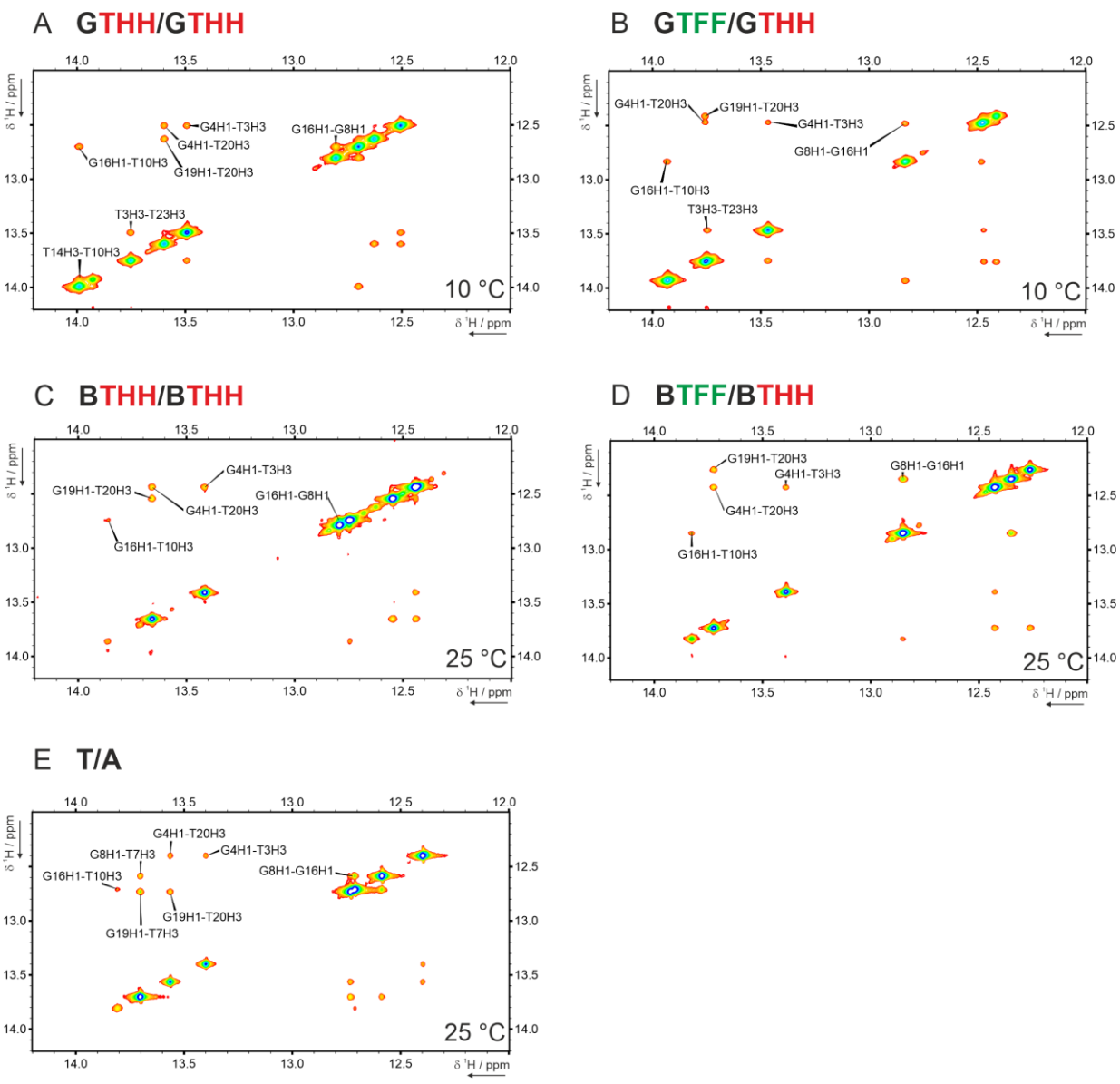


Figure S3.8 Imino region of the 2D ^{1}H,^{1}H NOESY NMR spectrum of the duplexes (A) GTHH/GTHH, (B) GTFF/GTHH, (C) BTHH/BTHH, (D) BTFF/BTHH and (E) T/A at the indicated temperature. Assignment of the imino-imino connectivities is shown. Spectra recorded on 1.0 mM DNA duplex in NMR buffer, 90%H_2O/10%D_2O, 600 MHz.

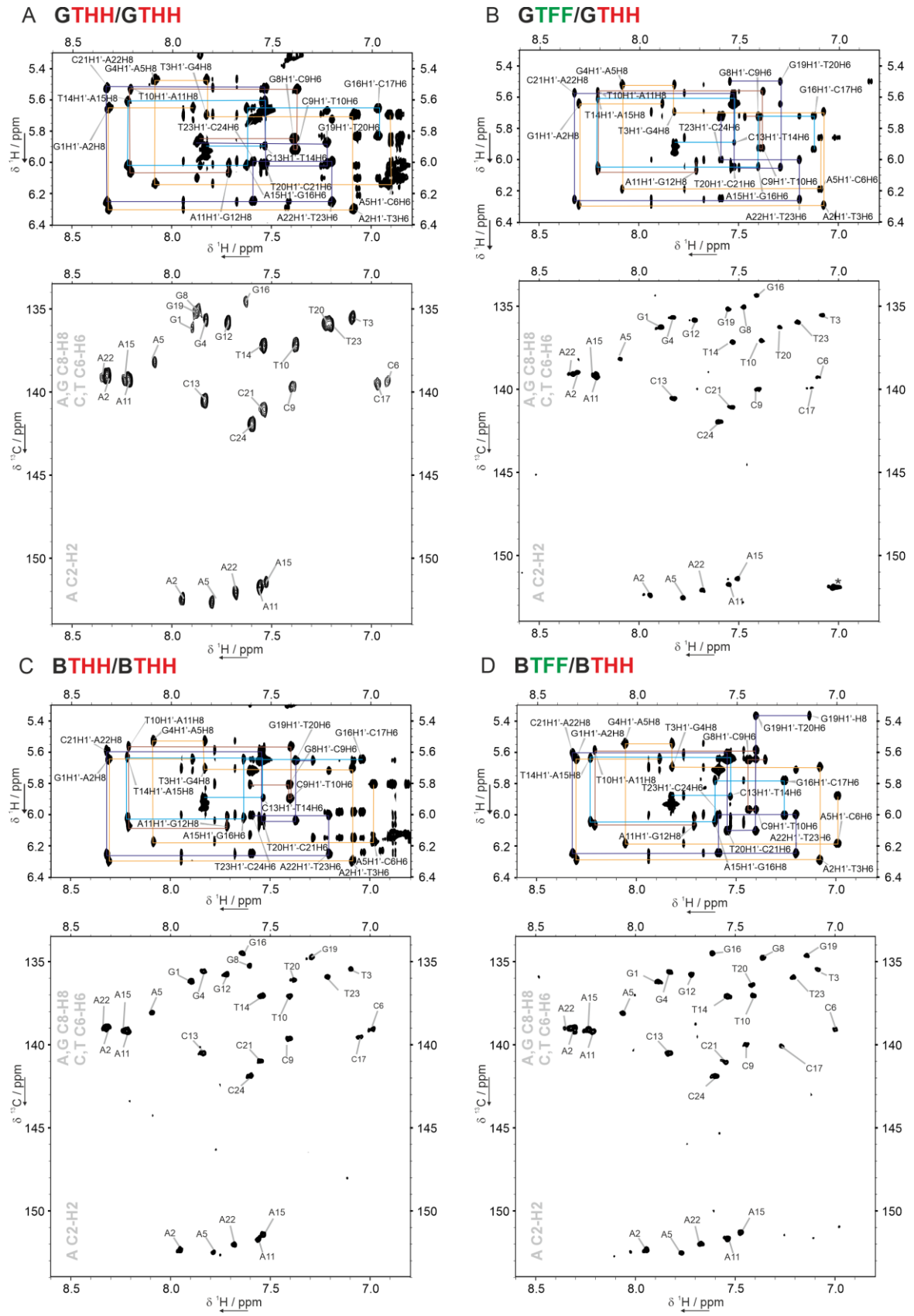


Figure S3.9 Aromatic-anomeric region of the 2D ^{1}H,^{1}H NOESY (top) spectrum and aromatic region of the ^{1}H,^{13}C HSQC (bottom) spectrum of the duplexes (A) GTHH/GTHH, (B) GTFF/GTHH, (C) BTHH/BTHH, (D) BTFF/BTHH. Assignment on the NOESY spectra is reported for clarity only for the sequential step 5'-H1' (n)-H6/H8(n+1)-3'.

Sequential walk is indicated and color coded as follows: G1-C6 orange, G8-G12 brown, C13-C17 cyan, G19-C24 blue. Spectra recorded on 1.0 mM DNA duplex in NMR buffer, 100%D_2O, 600 MHz, 25 °C.

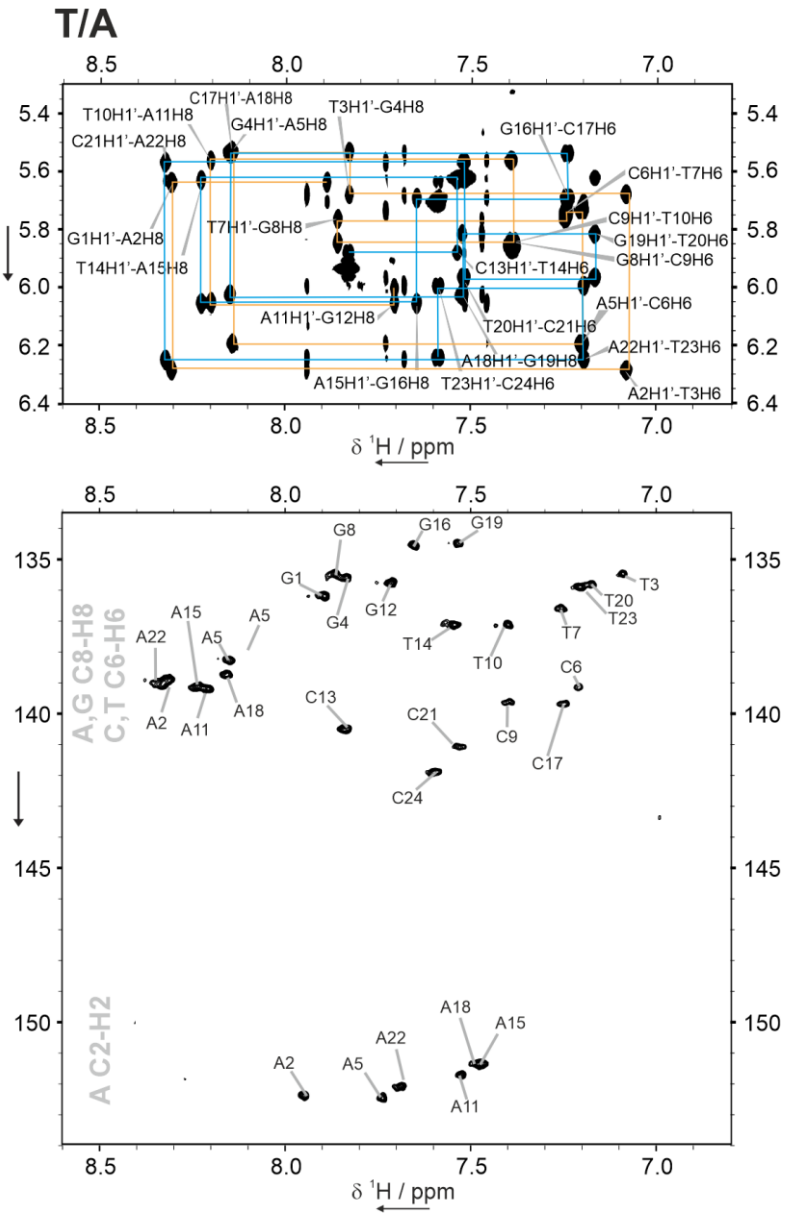


Figure S3.10 Aromatic-anomeric region of the 2D ^{1}H,^{1}H NOESY (top) spectrum and aromatic region of the ^{1}H,^{13}C HSQC (bottom) spectrum of the T/A duplex. Assignment on the NOESY spectra is reported for clarity only for the sequential step 5'-H1' (n)-H6/H8(n+1)-3'. Sequential walk is indicated and color coded as follows: G1-G12 orange, C13-C24 cyan. Spectra recorded on 1.0 mM DNA duplex in NMR buffer, 100%D_2O, 600 MHz, 25 °C.

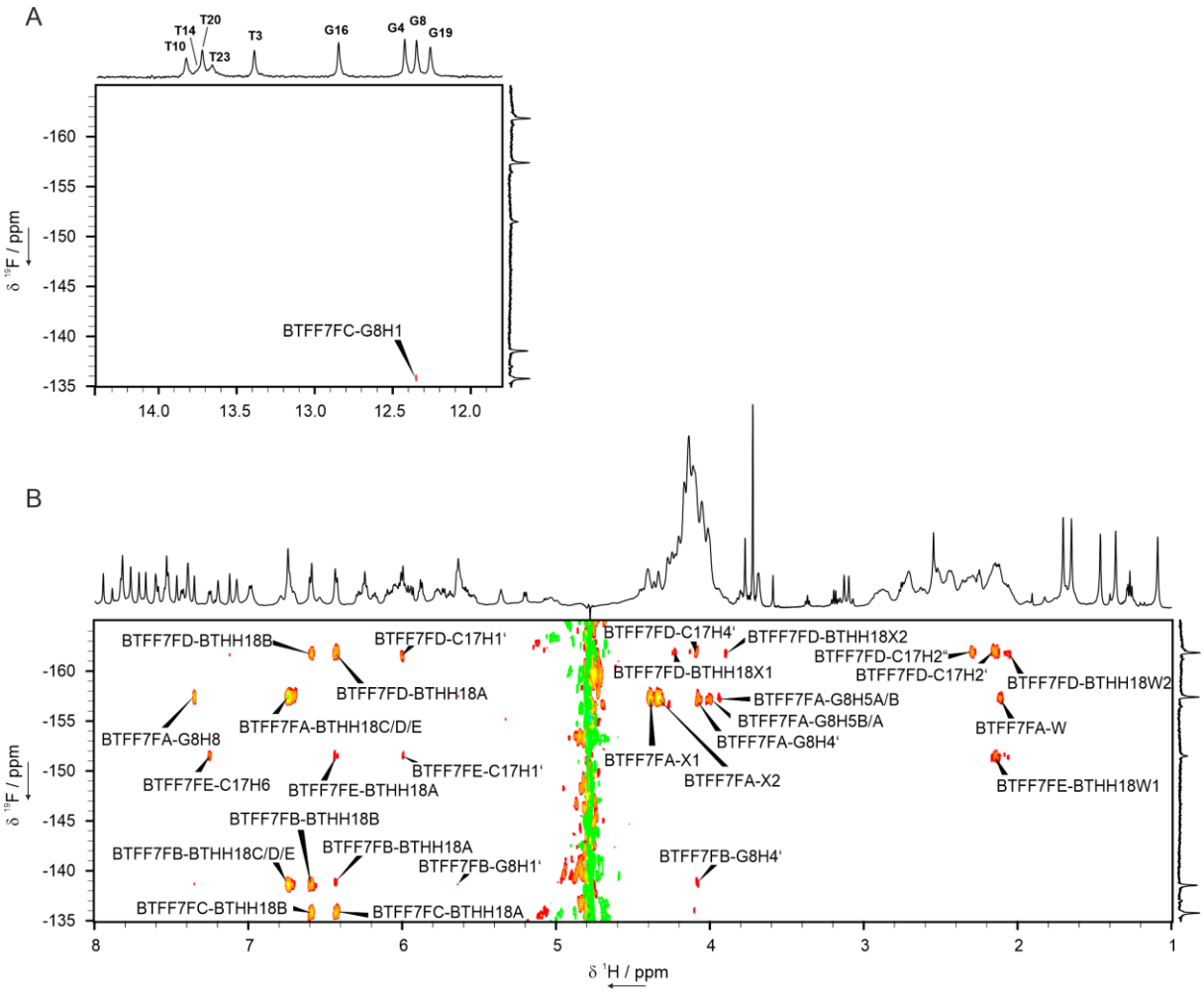


Figure S3.11 Imino (A) and aromatic/sugar region (B) of the 2D $^{1}H,^{19}F$ HOESY spectrum of BTFF/BTHH duplex with assignment. Spectra recorded on 1.0 mM DNA duplex in NMR buffer, 90%H_2O/10%D_2O, 600 MHz, 25 °C.

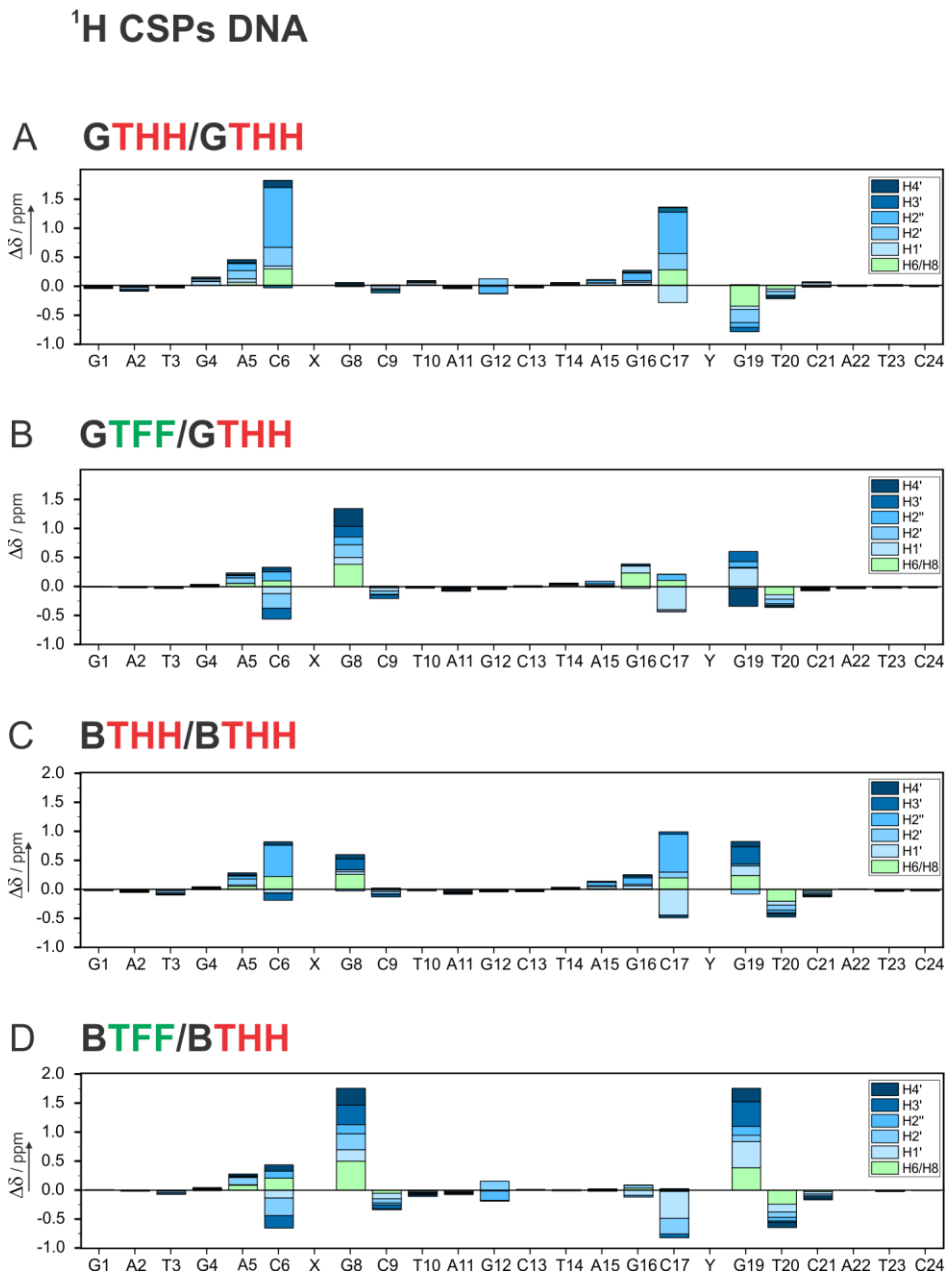


Figure S3.12 Chemical shift perturbation (CSP, Δδ) analysis of base and sugar ^{1}H for the duplexes (A) GTHH/GTHH, (B) GTFF/GTHH, (C) BTHH/BTHH, (D) BTFF/BTHH. $\Delta\delta = (^1\text{H c.s.}_{X/Y}) - (^1\text{H c.s.}_{T/A})$.

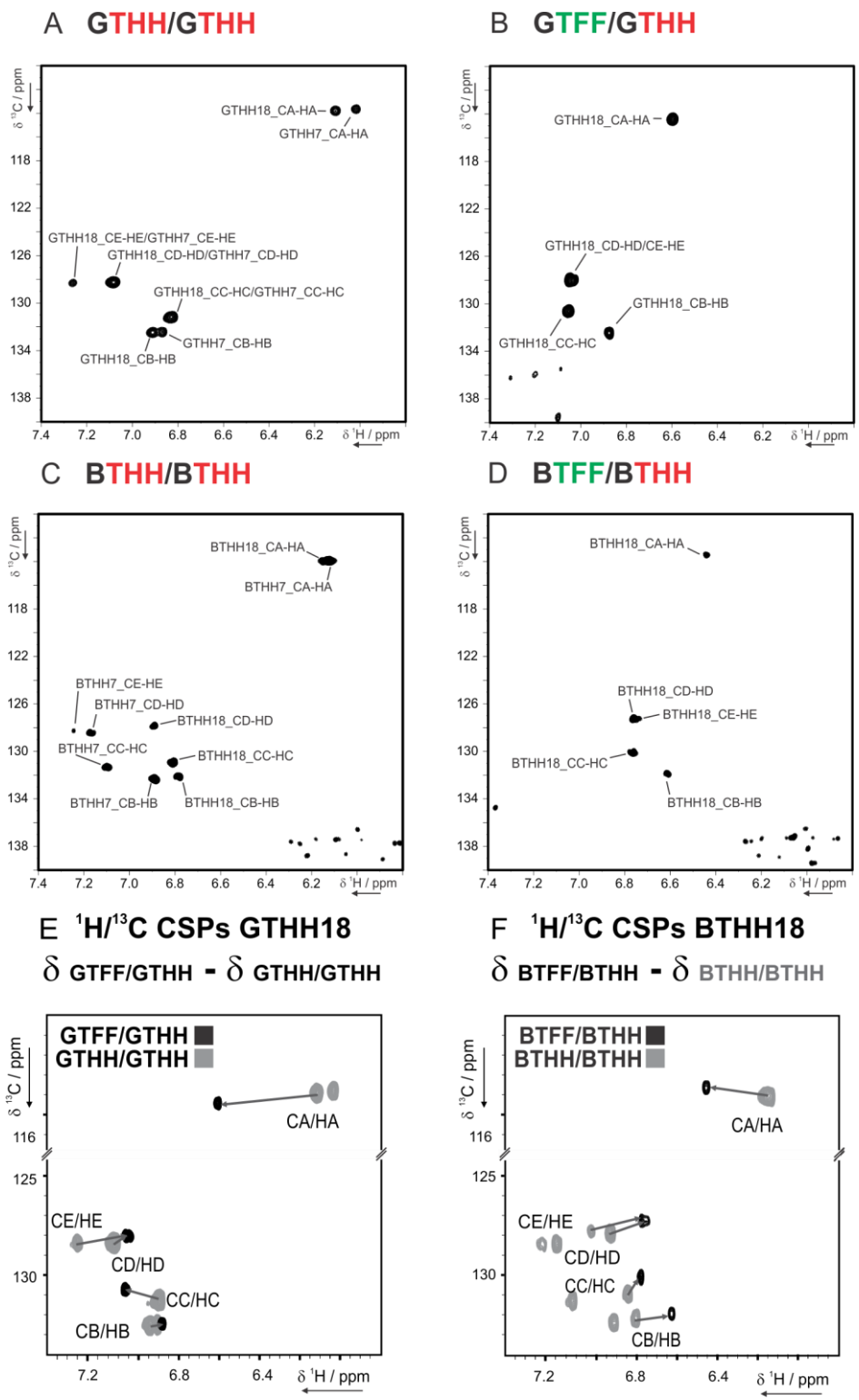


Figure S3.13 Tolane region of the $^{1}H,^{13}C$ HSQC spectrum recorded on the duplexes (A) GTHH/GTHH, (B) GTFF/GTHH, (C) BTHH/BTHH, (D) BTFF/BTHH with assignment. Spectra recorded on 1.0 mM DNA duplex in NMR buffer, 90%H_2O/10%D_2O (A,B) or 100%D_2O (C,D), 600 MHz, 10 °C. Spectra A,B were recorded with a F1 FID resolution of 178 Hz without folding of other signals; spectra C,D were recorded with a F1 FID resolution of 115 Hz and contain sugar C1'-H1' folded signals. (E) Combined $^{1}H/^{13}C$ CSP of the tolane unit THH18 in GTFF/GTHH compared to GTHH/GTHH, represented as dark gray arrows on the overlay of the aromatic region of the $^{1}H,^{13}C$ HSQC of GTHH/GTHH (gray spectrum) and GTFF/GTHH (black spectrum). (F) Same as in (E) but with BuNA backbone.

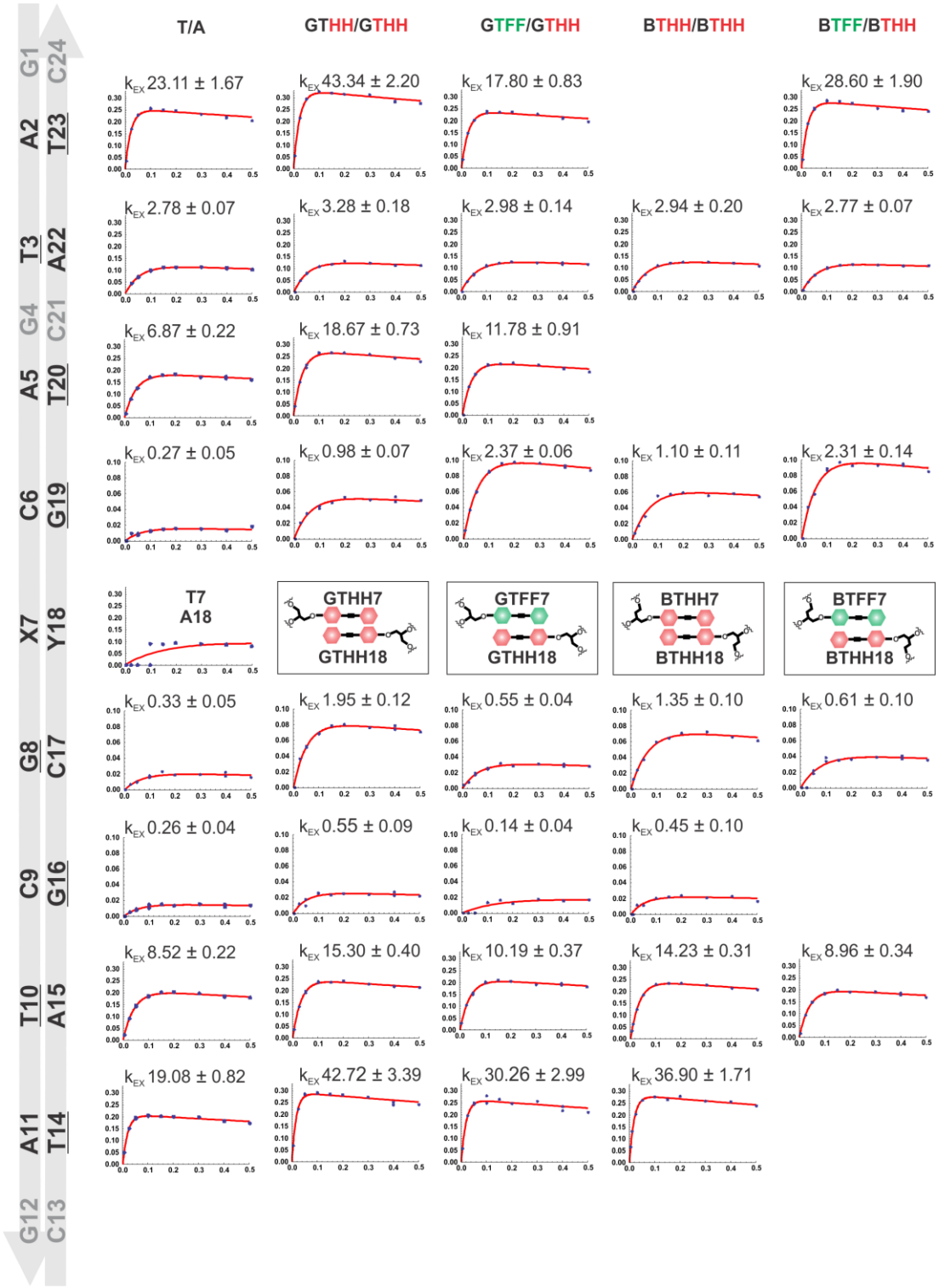


Figure S3.14 Relative peak intensities of the imino proton NMR signals as a function of the mixing time after water inversion in the CLEANEX-PM experiments. Plots are arranged in rows, as a function of the imino-containing residue position along the duplex, and in columns, according to the different duplex (label on top). The sequence is reported on the left. The residues whose imino proton-water k_{EX} could not be determined due to too slow exchange (G4 in all the duplexes; G16 in BTFF7/BTHH18) or to the peak broadening beyond detection at 25 °C (G1 and G12 in all the duplexes) are gray. The imino proton-water k_{EX} of residue T20 and T23 in BTHH7/BTHH18 as well as T20 and T14 in BTFF7/BTHH18 could not be determined due to peak overlap. The best fit to Eq. 3.12 is depicted as a solid red line. Results of the fitting are reported on top of each graph (error from the fitting).

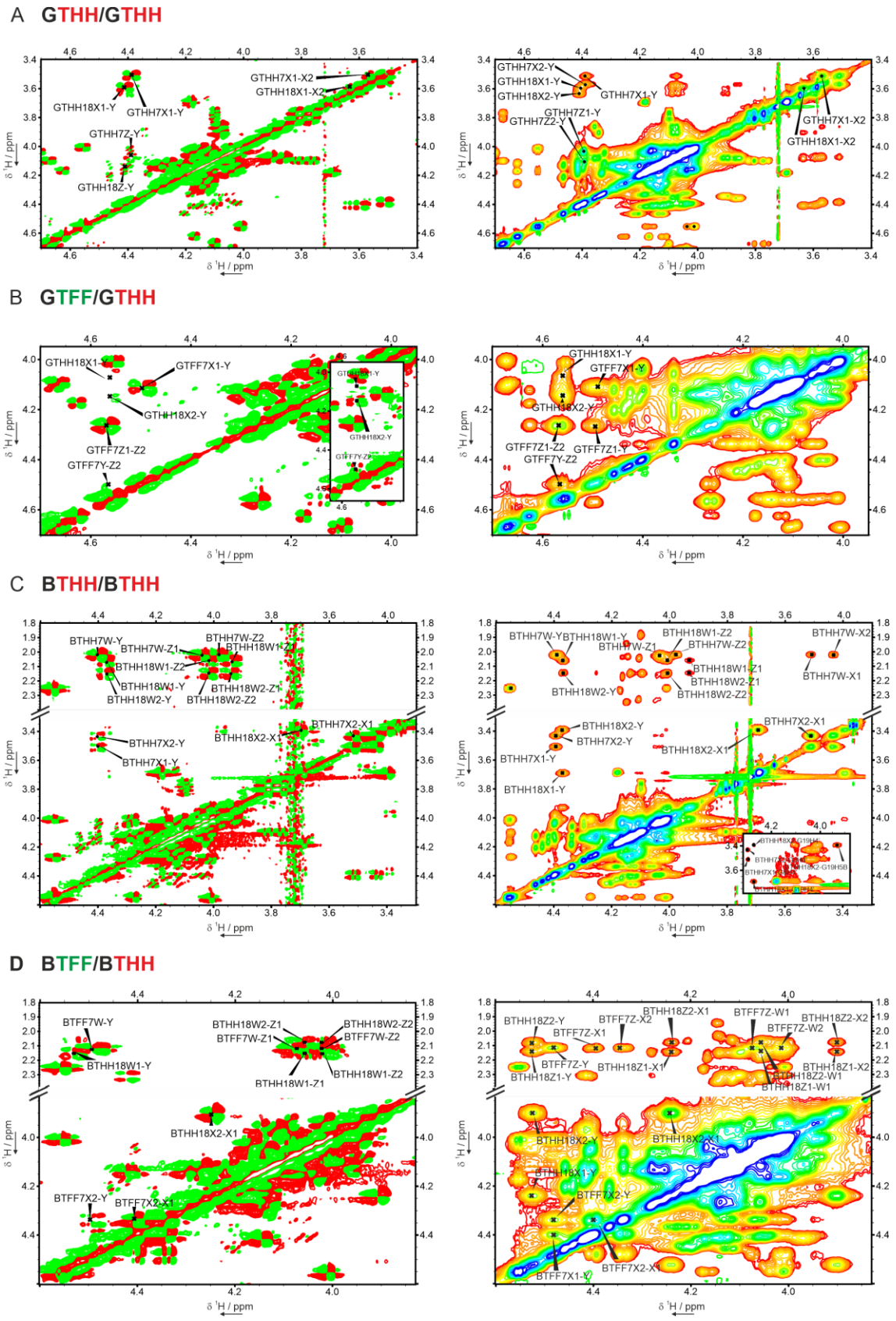


Figure S3.15 Acyclic backbone region of ^{1}H,^{1}H DQF-COSY (left) and ^{1}H,^{1}H NOESY (right) spectra recorded on the duplexes (A) GTHH/GTHH, (B) GTFF/GTHH, (C) BTHH/BTHH, (D) BTFF/BTHH. Spectra recorded on 1.0 mM DNA duplex in NMR buffer, 100%D_2O, 600 MHz, 25 °C.

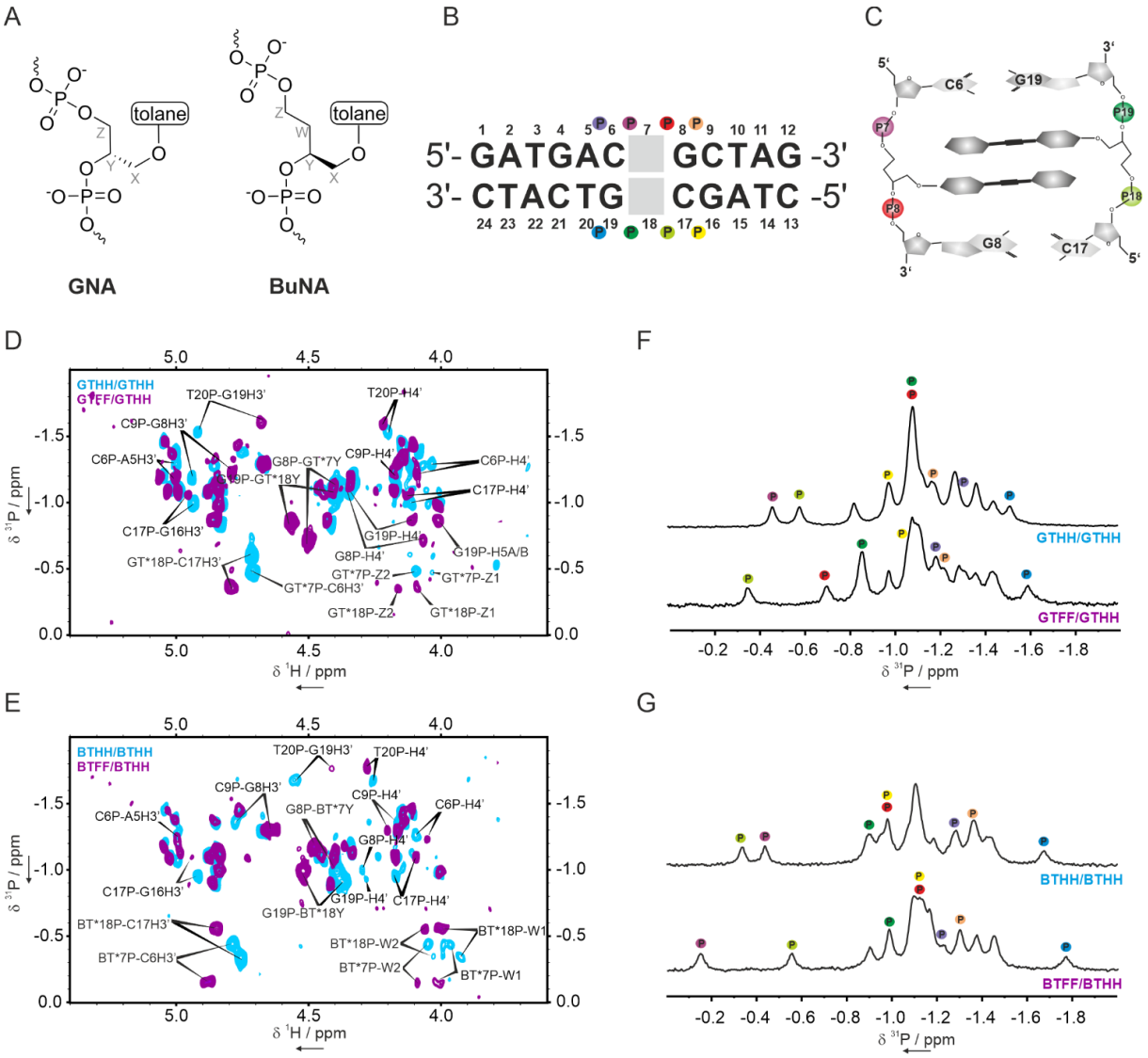


Figure S3.16 Assignment of the ^{31}P NMR spectra. Letter code, color code and numbering used for labeling acyclic backbone (A) and phosphate groups (B-C). Overlay of the ^{1}H,^{31}P HSQC of (D) GTHH/GTHH (cyan) and GTFF/GTHH (purple) duplex and (E) BTHH/BTHH (cyan) and BTFF/BTHH (purple) duplex, with H3' (n-1)-P(n) and H4' (n)-P(n) assignment indicated only for phosphate residues 6-9 and 17-20. Overlay of the 1D ^{31}P NMR spectrum of (F) GTHH/GTHH (top) and GTFF/GTHH (bottom), and F) BTHH/BTHH (top) and BTFF/BTHH (bottom), with assignment of the phosphate residues 6-9 and 17-20. Spectra recorded on 1.0 mM DNA duplex in NMR buffer, 100%D_2O, 600 MHz, 25 °C. We note that P7 in GTFF/GTHH could not be assigned because the correlations to the glycerol backbone Z protons and to C6 H3' in the 1H,31P HSQC are missing. This can be due either to a particularly small ^{3}JHP or to a highly dynamic backbone. In either case, GTFF/GTHH P7 has a peculiar behavior. Overall, the GTFF/GTHH combination produced the most significant perturbation on the phosphate backbone compared to the reference duplex.

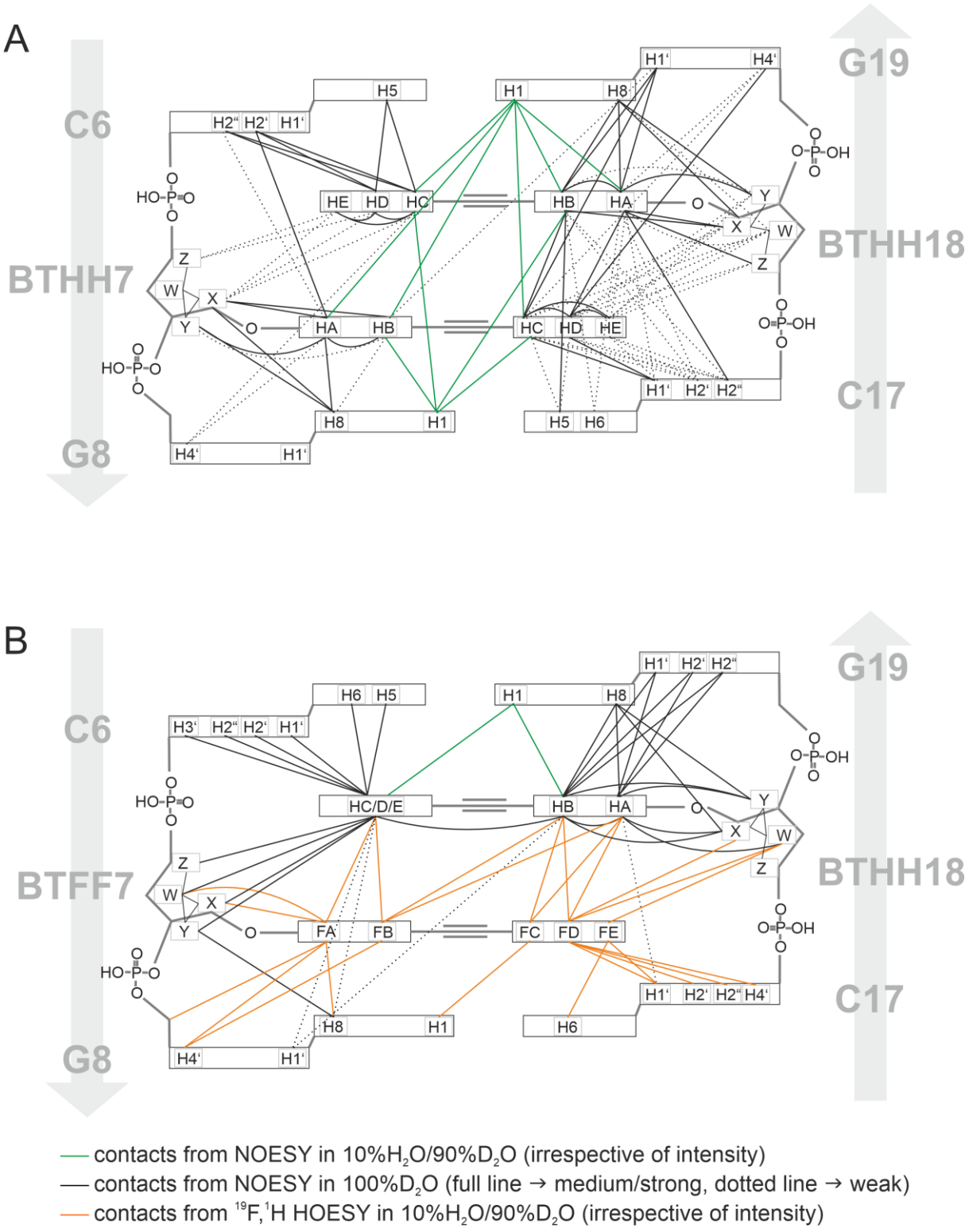


Figure S3.17 (A) Schematics of the tolane NOE contacts detected for BTHH/BTHH. (B) Schematics of the tolane NOE and hetero-NOE contacts detected for BTFF/BTHH. Color code explained in the legend on the bottom. C6 H1' and G8 H1' resonances in BTHH/BTHH are overlapped and the ambiguous cross peaks observed to BTHH7 HA and BTHH18 HC/D are not indicated. NOESY spectra recorded with 200 ms mixing time on 1.0 mM DNA duplex in NMR buffer, 100%D_2O, 600 MHz, 25 °C.

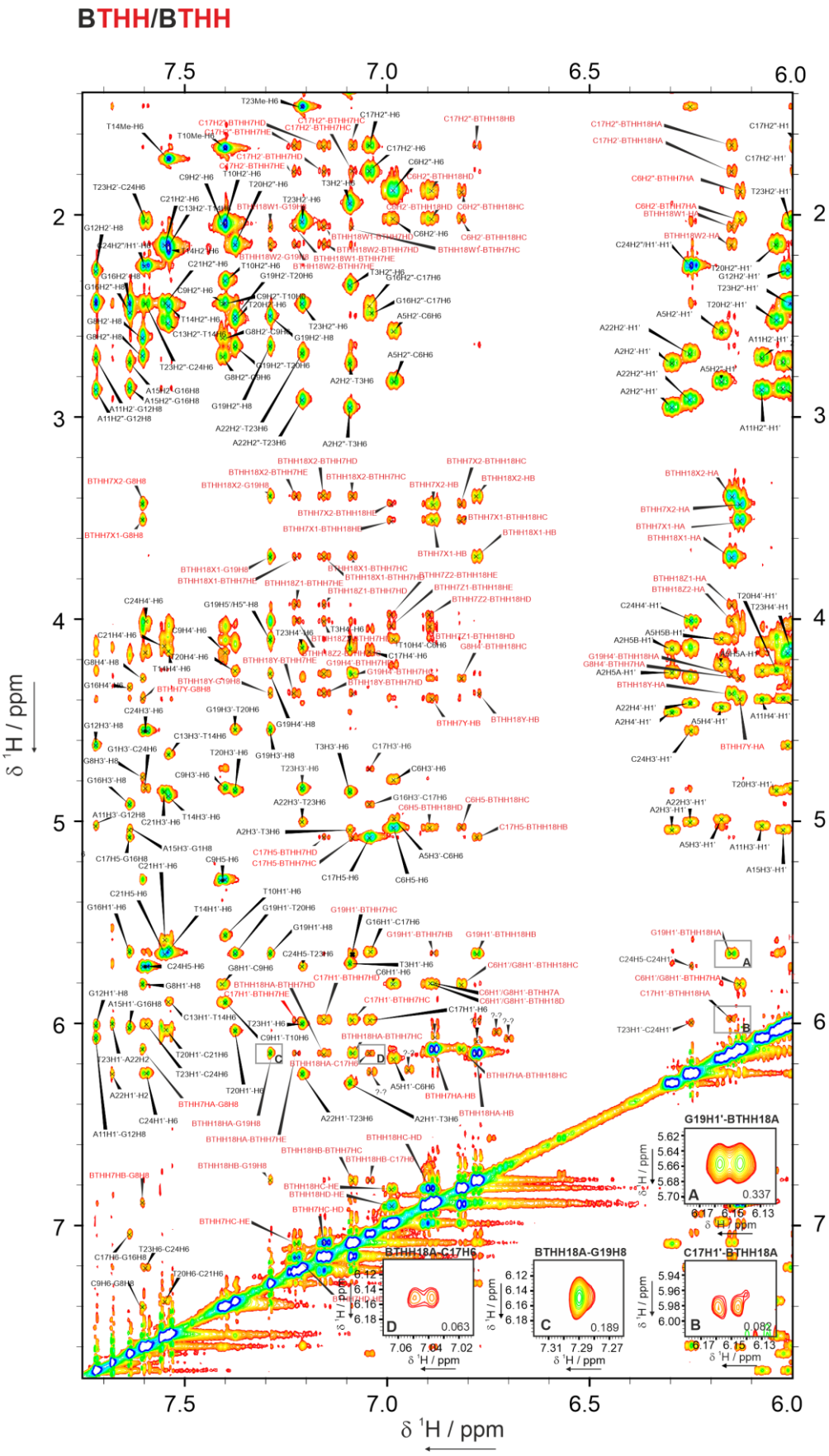


Figure S3.18 Aromatic region of the NOESY spectrum of BTHH/BTHH 1 mM in 100% D_2O (600 MHz, 25 °C) with assignment. Cross peaks with the tolane aromatic protons are highlighted in red. Spectrum shown at lowest positive contour level adjusted to be comparable with spectrum shown in Figure S3.17. The cross peaks boxed in gray, corresponding to NOE interaction of BTHH18 HA with aromatic and anomeric protons of its adjacent residues (*i.e.*, C17 and G19), did not suffer from overlap with other peaks and were suitable for integration. A zoom on the selected cross peaks is shown on the bottom right corner of the spectrum. The integral of each peak, normalized to the integral of peak C9_H5-H6, is shown in the bottom right corner of the corresponding spectrum excerpt. See for comparison the corresponding cross peaks in BTFF/BTHH (Figure S3.19).

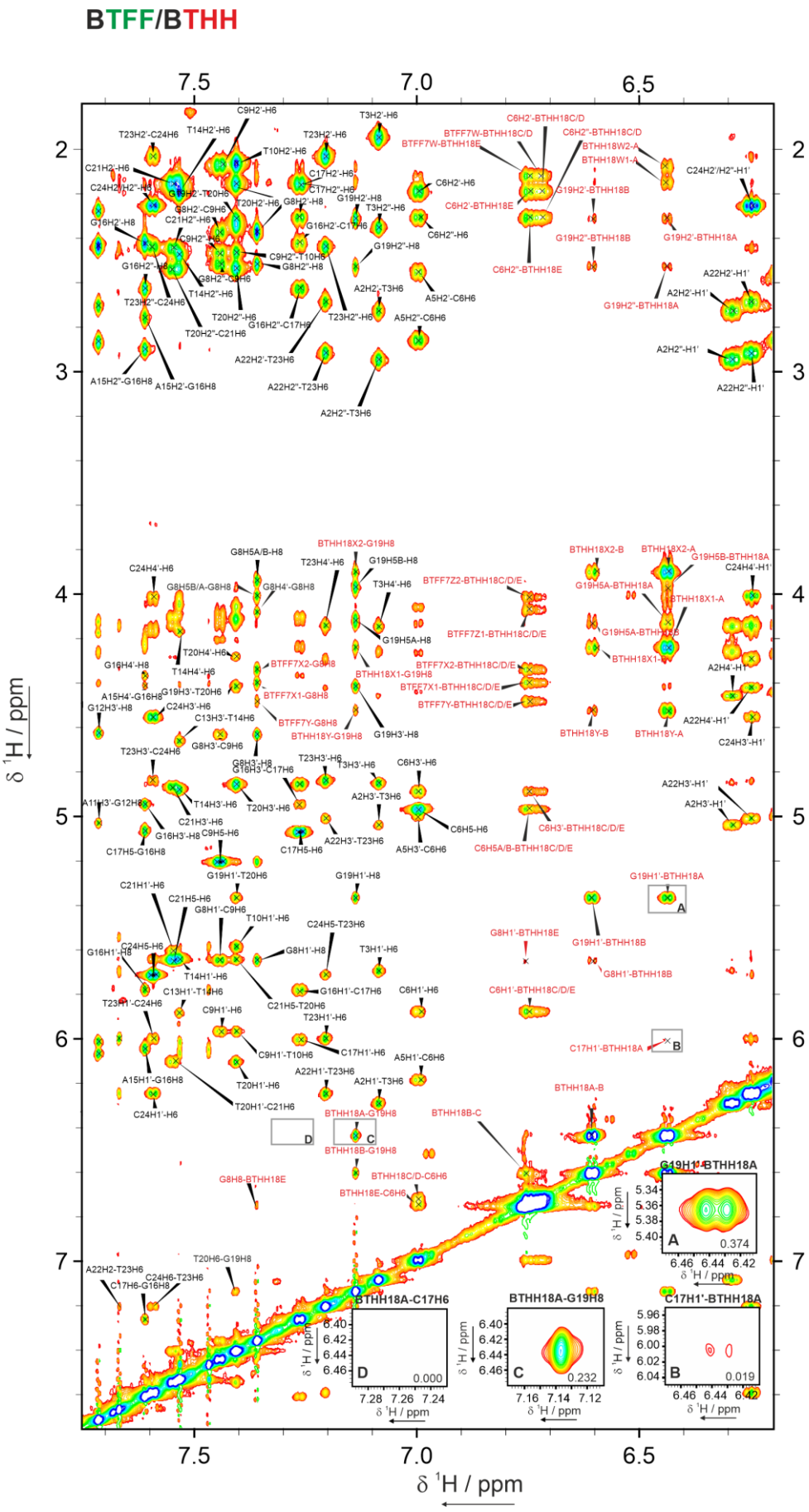


Figure S3.19 Aromatic region of the NOESY spectrum of BTFF/BTHH 1 mM in 100% D_2O (600 MHz, 25 °C) with assignment. Cross peaks with the tolane aromatic protons are highlighted in red. Spectrum shown at lowest positive contour level adjusted to be comparable with spectrum shown in Figure S3.17. The cross peaks boxed in gray, corresponding to NOE interaction of BTHH18 HA with aromatic and anomeric protons of its adjacent residues (*i.e.*, C17 and G19), did not suffer from overlap with other peaks and were suitable for integration. A zoom on the selected cross peaks is shown on the bottom right corner of the spectrum. The integral of each peak, normalized to the integral of peak C9_H5-H6, is shown in the bottom right corner of the corresponding spectrum excerpt. See for comparison the corresponding cross peaks in BTHH/BTHH (Figure **S3.18**).

3.4.2 Experimental procedures

3.4.2.1 General material and methods

All standard chemicals were purchased from commercial suppliers. D-(+)-solketal = (*S*)-(+)-2,2-dimethyl-1,3-dioxolane-4-methanol, trimethylsilylacetylene, (*R*)-(+)-glycidol were obtained from Tokyo Chemical Industry. Iodopentafluorobenzene, 4-iodophenol, L-malic acid, borane dimethyl sulfide complex solution (2 M in THF), trimethyl borate, benzaldehyde dimethyl acetal and tetrakis(triphenylphosphine)palladium(0) were purchased from Sigma Aldrich. 4,4'-Dimethoxytrityl chloride and 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite were obtained from Chem Genes Corporation. Triphenylphosphine was obtained from Acros Organics and diisopropyl azodicarboxylate was purchased from Alfa Aesar.

4-(2-Phenylethynyl)phenol (**3.1**)[413] as well as the backbone precursors (*S*)-DMT-*O*-glycidol (**3.2**)[414] and 4-(*S*)-hydroxymethyl-2-phenyl-1,3-dioxan (**3.3**)[415] were synthesized according to previously published protocols.

Dry solvents DCM and THF were obtained via a solvent purification system (SPS) from Inert Corporation. Pyridine for DMT-protection was obtained from Acros Organics and dried over activated molecular sieves. Solvents for column chromatography were used in technical quality and distilled prior to use. All other organic solvents were used in pro analysis or for synthesis quality without further purification.

Column chromatography was carried out on silica gel (Kieselgel 60, Merck, 0.063 – 0.200 mm). Thin layer chromatography (TLC) was performed on aluminum-backed plates coated with silica gel and a fluorescent indicator (Alugram SIL G/UV254, Macherey-Nagel, UV visualization, 254 nm). For UV-inactive substances a cerium molybdate stain was used.

3.4.2.2 NMR-Spectroscopy and mass spectrometry

^{1}H-, ^{13}C-, ^{19}F- and ^{31}P-NMR spectra were recorded on a Bruker Avance HD III spectrometer at 400 MHz. Spectra were calibrated to the residual solvent peak from $CDCl_3$ (δ = 7.26 (^{1}H) and δ = 77.16 (^{13}C)) and DMSO-d_6 (δ = 2.50 (^{1}H) and δ = 39.52 (^{13}C)). Other nuclei were referenced on the unified scale. Chemical shifts δ are reported in ppm and coupling constants *J* are given in Hz. Multiplicities are denoted as follows: s (singlet), d (doublet), t (triplet), q

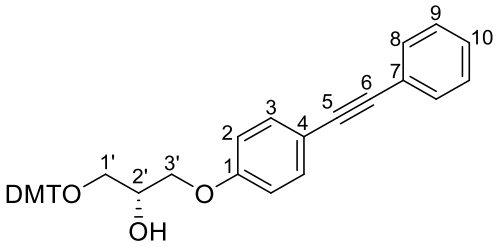
(quartet), p (pentet), dd (doublet of doublet), td (triplet of doublet), m (multiplet), br (broad).

NMR spectra of building blocks were evaluated with MestReNova v12.0.4.

High resolution ESI mass spectra were measured on a Bruker micrOTOF-Q III spectrometer.

3.4.2.3 Synthetic procedures

Compound **S3.1**



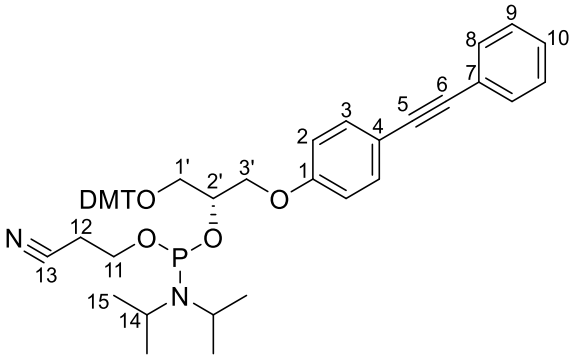
Under nitrogen atmosphere, 4-(2-phenylethynyl)phenol (**3.1**) (500 mg, 2.57 mmol, 1.12 eq.) and sodium hydride (60% oil dispersion, 20.6 mg, 515 µmol, 0.20 eq.) were dissolved in anhydrous DMF (6 mL) and stirred at ambient temperature for 2 h. A solution of (*S*)-DMT-*O*-glycidol (**3.2**) (864 mg, 2.29 mmol. 1.00 eq.) in anhydrous DMF (8 mL) was added and the mixture was stirred for 1 d at 110°C. The solvent was removed under reduced pressure and the residue was purified by column chromatography (hexane/EtOAc 2:1 + 3% Et_3N –1:1 + 1% Et_3N) to afford compound **S3.1** (863 mg, 1.37 mmol, 60%) as a yellow foam.

^{1}H NMR (400 MHz, $CDCl_3$): δ (ppm) = 7.54 – 7.49 (m, 2H, 8-H), 7.48 – 7.44 (m, 2H, 3-H), 7.44 – 7.40 (m, 2H, DMT-H), 7.37 – 7.26 (m, 9H, 9-H, 10-H, DMT-H), 7.25 – 7.18 (m, 1H, DMT-H), 6.88 – 6.84 (m, 2H, 2-H), 6.84 – 6.79 (m, 4H, DMT-H), 4.18 – 4.10 (m, 1H, 2'-H), 4.11 – 4.01 (m, 2H, 3'-H), 3.79 (s, 6H, DMT-OMe-H), 3.40 – 3.30 (m, 2H, 1'-H), 2.44 (d, J = 5.3 Hz, 1H, 2'-OH);

^{13}C{^{1}H} NMR (100 MHz, $CDCl_3$): δ (ppm) = 158.69 (1-C), 158.68 (DMT-C), 144.81 (DMT-C), 135.93 (DMT-C), 135.91 (DMT-C), 133.19 (3-C), 131.60 (8-C), 130.17 (DMT-C), 128.46 (DMT-C), 128.20 (DMT-C), 128.13 (10-C), 128.04 (9-C), 127.02 (DMT-C), 123.65 (7-C), 115.91 (4-C), 114.74 (2-C), 113.31 (DMT-C), 89.38 (5-C), 88.35 (6-C), 86.45 (DMT-C), 69.58 (2'-C), 69.21 (3'-C), 64.04 (1'-C), 55.37 (DMT-C);

HR-MS (ESI+): *m/z* calc. ($C_{38}H_{34}O_5Na$, $[M+Na]^+$): 593.22984, found: 593.23111

Compound **S3.2** = GTHH-PA



Compound **S3.1** (340 mg, 596 µmol, 1.00 eq.) was dissolved with DIPEA (596 µL, 442 mg, 3.42 mmol, 5.74 eq.) in anhydrous DCM (15 mL). After 10 min CEP-Cl (173 µL, 773 µmol, 1.30 eq.) was added. Additional CEP-Cl (17.0 µL, 18.3 mg, 77.3 µmol, 0.13 eq.) was added after 2 h stirring at room temperature. The reaction mixture was stirred at ambient temperature for additionally 4 h. The solvent was removed under reduced pressure and the residue was purified by column chromatography (DCM/hexane 3:1 + 1% Et_3N) to afford compound **S3.2** as a colorless foam (358 mg, 464 µmol, 78%).

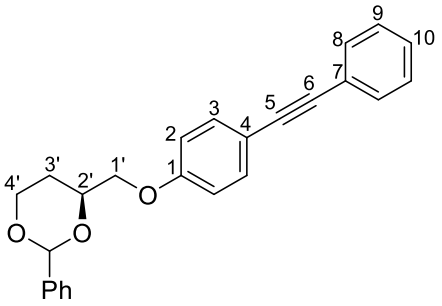
^{1}H NMR (400 MHz, $CDCl_3$): δ (ppm) = 7.54 – 7.49 (m, 4H, 8-H), 7.49 – 7.38 (m, 8H, 3-H, DMT-H), 7.39 – 7.10 (m, 20H, 9-H, 10-H, DMT-H), 6.92 – 6.74 (m, 12H, 2-H, DMT-H), 4.39 – 4.23 (m, 3H, 2'-H, 3'-H), 4.19 – 4.03 (m, 3H, 3'-H), 3.91 – 3.67 (m, 16H, 11-H; DMT-H), 3.66 – 3.47 (m, 4H, 14-H), 3.42 – 3.20 (m, 4H, 1'-H), 2.72 – 2.37 (m, 4H, 12-H), 1.25 – 0.98 (m, 24H, 15-H);

^{13}C{^{1}H} NMR (100 MHz, $CDCl_3$): δ (ppm) = 158.89 (1-C), 158.85 (1-C), 158.60 (DMT-C), 158.59 (DMT-C), 144.92 (DMT-C), 144.90 (DMT-C), 136.12 (DMT-C), 136.09 (DMT-C), 136.07 (DMT-C), 133.19 (3-C), 133.13 (3-C), 131.59 (8-C), 130.25 (DMT-C), 130.20 (DMT-C), 130.17 (DMT-C), 128.45, 128.33, 128.26, 128.08, 127.93, 126.93 (DMT-C), 126.89 (DMT-C), 123.69 (7-C), 117.86 (13-C), 117.75 (13-C), 115.65 (4-C), 115.59 (4-C), 114.79 (2-C), 114.72 (2-C), 113.20 (DMT-C), 89.47 (5-C), 88.27 (6-C), 88.24 (6-C), 86.29 (DMT-C), 86.25 (DMT-C), 72.21 (3'-C), 72.03 (3'-C), 71.82 (3'-C), 71.65 (3'-C), 68.98 (2'-C), 68.95 (2'-C), 68.74 (2'-C), 68.70 (2'-C), 63.76 (1'-C), 63.72 (1'-C), 63.67 (1'-C), 63.64 (1'-C), 58.67 (11-C), 58.59 (11-C), 58.49 (11-C), 58.41 (11-C), 55.36 (DMT-C), 55.35 (DMT-C), 43.41 (14-C), 43.37 (14-C), 43.29 (14-C), 43.24 (14-C), 24.85 (15-C), 24.83 (15-C), 24.78 (15-C), 24.75 (15-C), 24.71 (15-C), 24.64 (15-C), 24.57 (15-C), 20.50 (12-C), 20.44 (12-C), 20.37 (12-C), 20.30 (12-C);

^{31}P NMR (162 MHz, $CDCl_3$): δ (ppm) = 149.87, 149.61;

HR-MS (ESI+): *m/z* calc. ($C_{47}H_{51}O_6N_2PNa$, $[M+Na]^+$): 793.33769, found: 793.33746.

Compound **S3.3**



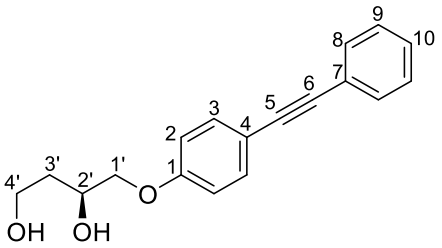
General procedure was adapted from He et. al.[416] 4-(2-Phenylethynyl)phenol (**3.1**) (500 mg, 2.57 mmol, 1.00 eq.), 4-(*S*)-hydroxmethyl-2-phenyl-1,3-dioxan (**3.3**) (649 mg, 3.34 mmol, 1.30 eq.) and triphenylphosphine (709 mg, 2.70 mmol, 1.05 eq.) were dissolved in THF (0.9 mL). Diisopropyl azodicarboxylate (530 µL, 2.70 mmol, 1.05 eq.) was added dropwise and the reaction mixture was sonicated for 30 min. Purification of the residue by column chromatography (hexane/DCM 3:2) afforded compound **S3.3** as a yellow solid (722 mg, 1.95 mmol, 76%).

^{1}H NMR (400 MHz, $CDCl_3$): δ (ppm) = 7.54 – 7.49 (m, 4H, 8-H, Ph-H), 7.49 – 7.44 (m, 2H, 3-H), 7.42 – 7.28 (m, 6H, 9-H, 10-H, Ph-H), 6.93 – 6.88 (m, 2H, 2-H), 5.60 (s, 1H, C*H*-Ph), 4.41 – 4.32 (m, 1H), 4.37 – 4.26 (m, 1H), 4.23 – 4.17 (m, 1H), 4.11 – 3.99 (m, 1H), 4.06 – 3.98 (m, 1H), 2.09 – 1.92 (m, 1H), 1.77 – 1.68 (m, 1H);

^{13}C{^{1}H} NMR (100 MHz, $CDCl_3$): δ (ppm) = 158.86 (C-1), 138.38 (C-12), 133.19 (C-3), 131.59 (C-8), 129.07, 128.45, 128.44, 128.11, 126.26 (13-C), 123.67 (C-7), 115.85 (C-4), 114.83 (C-2), 101.48(*C*H-Ph), 89.41 (C-5), 88.32 (C-6), 75.44 (C-2'), 70.77 (C-1'), 66.88 (C-4'), 28.31 (C-3');

HR-MS (ESI+): *m/z* calc. ($C_{25}H_{22}NaO_3$, $[M+Na]^+$): 393.14666, found: 393.14680.

Compound **S3.4**



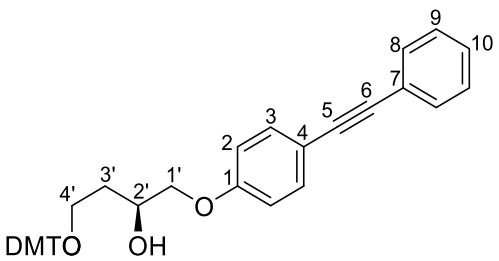
Compound **S3.3** (650 mg, 1.75 mmol, 1.00 eq.) was dissolved in methanol (24 mL). *p*-Toluenesulfonic acid monohydrate (501 mg, 2.63 mmol, 1.50 eq.) was added and the reaction mixture stirred at ambient temperature for 40 h. Afterwards, the reaction was diluted with a saturated potassium carbonate solution (2.5 mL) and extracted with DCM (4x50 mL). The combined organic layers were dried over Na_2SO_4, evaporated and the solid residue was purified by column chromatography (hexane/EtOAc 1:1–1:2) to afford compound **S3.4** as a colorless solid (416 mg, 1.47 mmol, 84%).

^{1}H NMR (400 MHz, DMSO-d_6): δ (ppm) = 7.55 – 7.50 (m, 2H, 8-H), 7.50 – 7.45 (m, 2H, 3-H), 7.44 – 7.36 (m, 3H, 9-H, 10-H), 7.01 – 6.95 (m, 2H, 2-H), 4.87 (d, *J* = 4.8 Hz, 1H, 2'-OH), 4.43 (t, *J* = 5.1 Hz, 1H, 4'-OH), 3.98 – 3.84 (m, 3H, 1'-H, 2'-H), 3.60 – 3.51 (m, 2H, 4'-H), 1.74 – 1.64 (m, 1H, 3'-H), 1.62 – 1.51 (m, 1H, 3'-H);

^{13}C{^{1}H} NMR (100 MHz, DMSO-d_6): δ (ppm) = 159.10 (1-C), 132.95 (3-C), 131.19 (8-C), 128.74 (9-C), 128.44 (10-C), 122.70 (7-C), 114.96 (2-C), 114.07 (4-C), 89.53 (5-C), 87.95 (6-C), 72.54 (1'-C), 65.85 (2'-C), 57.58 (4'-C), 36.74 (3'-C);

HR-MS (ESI+): *m/z* calc. ($C_{18}H_{18}NaO_3$, $[M+Na]^+$): 305.11482, found: 305.11540.

Compound **S3.5**



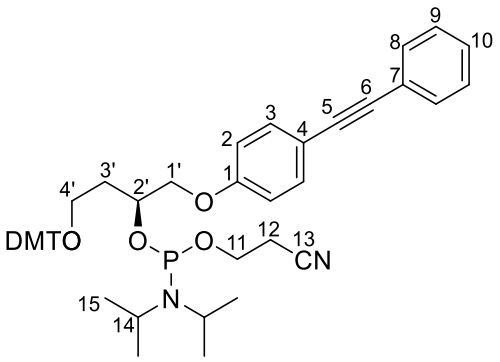
Compound **S3.4** (350 mg, 1.24 mmol, 1.00 eq.) was dissolved in anhydrous pyridine (19 mL). 4,4'-Dimethoxytrityl chloride (504 mg, 1.49 mmol, 1.20 eq.) was added in small portions over a time period of 25 min and the reaction mixture stirred at ambient temperature for 3 h. The solvent was removed under reduced pressure and the residue was purified by column chromatography (hexane/EtOAc 6:1–3:1 +1% Et_3N) to afford compound **S3.5** as a colorless foam (563 mg, 963 µmol, 78%).

^{1}H NMR (400 MHz, $CDCl_3$): δ (ppm) = 7.54 – 7.49 (m, 2H, 8-H), 7.47 – 7.42 (m, 4H, H-3, DMT-H), 7.38 – 7.27 (m, 9H, 9-H, 10-H, DMT-H), 7.24 – 7.19 (m, 1H, DMT-H), 6.88 – 6.79 (m, 6H, 2-H, DMT-H), 4.27 – 4.16 (m, 1H, 2'-H), 3.98 – 3.85 (m, 2H, 1'-H), 3.79 (s, 6H, DMT-H), 3.46 – 3.37 (m, 1H, 4'-H), 3.35 – 3.26 (m, 1H, 4'-H), 3.09 (d, *J* = 3.3 Hz, 1H, 2'-OH), 1.98 – 1.84 (m, 2H, 3'-H);

^{13}C{^{1}H} NMR (100 MHz, $CDCl_3$): δ (ppm) = 158.82 (C-1), 158.64 (DMT-C), 144.91 (DMT-C), 136.13 (DMT-C), 136.04 (DMT-C), 133.19 (8-C), 131.59 (3-C), 130.10, 128.15, 128.10, 128.06, 126.99, 123.67 (7-C), 115.79 (4-C), 114.72 (2-C), 113.32 (DMT-C), 89.42 (5-C), 88.30 (6-C), 86.80 (DMT-C), 71.89 (1'-C), 69.37 (2'-C), 61.38 (4'-C), 55.36 (DMT-C), 33.36 (3'-C);

HR-MS (ESI+): *m/z* calc. ($C_{39}H_{36}NaO_5$, $[M+Na]^+$): 607.24549, found: 607.24702

Compound **S3.6** = BTHH-PA



Compound **S3.5** (300 mg, 513 µmol, 1.00 eq.) was dissolved under nitrogen atmosphere with DIPEA (536 µL, 3.08 mmol, 6.00 eq.) in anhydrous DCM. After 10 min CEP-Cl (149 µL, 667 µmol, 1.30 eq.) was added. Additional CEP-Cl (11.5 µL, 51.3 µmol, 0.10 eq.) was added after 2 h stirring at room temperature. The reaction mixture was stirred additionally at ambient temperature for 1 h. The solvent was removed under reduced pressure and the residue

was purified by column chromatography (hexane/EtOAc 6:1–5:1 + 1% Et_3N) to afford compound **S3.6** as a colorless foam (349 mg, 445 µmol, 87%).

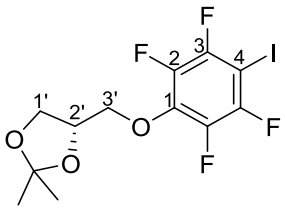
^{1}H NMR (400 MHz, $CDCl_3$): δ (ppm) = δ 7.55 – 7.47 (m, 4H, 8-H), 7.48 – 7.38 (m, 8H, 3-H, DMT-H), 7.40 – 7.21 (m, 17H, 9-H, 10-H, DMT-H), 7.26 – 7.15 (m, 2H, DMT-H), 6.87 – 6.75 (m, 12H, 2-H, DMT-H), 4.41 – 4.30 (m, 2H, 2'-H), 4.06 – 3.95 (m, 4H, 1'-H), 3.88 – 3.67 (m, 14H, 11-H, DMT-H), 3.68 – 3.46 (m, 6H, 11-H, 14-H), 3.32 – 3.18 (m, 4H, 4'-H), 2.61 – 2.36 (m, 4H, 12-H), 2.11 – 1.94 (m, 4H, 3'-H), 1.20 – 1.00 (m, 24H, 15-H).;

^{13}C{^{1}H} NMR (100 MHz, $CDCl_3$): δ (ppm) = 158.98 (C-1), 158.92 (C-1), 158.51 (DMT-C), 145.23 (DMT-C), 145.22 (DMT-C), 136.53 (DMT-C), 136.50 (DMT-C), 136.45 (DMT-C), 136.40 (DMT-C), 133.19 (3-C), 133.14 (3-C), 131.58 (8-C), 130.18 (DMT-C), 130.12 (DMT-C), 130.11 (DMT-C), 128.45, 128.33, 128.23, 128.08, 127.92, 126.84 (DMT-C), 126.83 (DMT-C), 123.71 (7-C), 117.88 (13-C), 117.78 (13-C), 115.56 (4-C), 115.51 (4-C), 114.77 (2-C), 114.75 (2-C), 113.18 (DMT-C), 113.16 (DMT-C), 89.51 (6-C), 89.49 (6-C), 88.24 (5-C), 88.22 (5-C), 86.26 (DMT-C), 86.23 (DMT-C), 71.00 (1'-C), 70.98 (1'-C), 70.77 (2'-C), 70.59(2'-C), 70.52 (2'-C), 70.35 (2'-C), 59.90 (4'-C), 59.82 (4'-C), 58.51(11-C), 58.48 (11-C), 58.33 (11-C), 58.29 (11-C), 55.36 (DMT-C), 55.34 (DMT-C), 43.32 (14-C), 43.19 (14-C), 34.00 (3'-C), 33.96 (3'-C), 33.93 (3'-C), 33.89 (3'-C), 24.86 (15-C), 24.78 (15-C), 24.71 (15-C), 24.67 (15-C), 24.62 (15-C), 24.59 (15-C), 24.54 (15-C), 20.47 (12-C), 20.40 (12-C), 20.37 (12-C), 20.30 (12-C);

^{31}P{^{1}H} NMR (162 MHz, $CDCl_3$): δ (ppm) = 148.94, 148.61;

HR-MS (ESI+): *m/z* calc. ($C_{48}H_{53}N_2NaO_6P$, $[M+Na]^+$): 807.3533, found: 807.3523.

Compound **S3.7**



(*S*)-(+)-2,2-Dimethyl-1,3-dioxolane-4-methanol (**3.4**, also known as (*S*)-(+)-1,2-isopropylidenglycerin or D-(+)-Solketal) (1.87 mL, 2.00 g, 15.1 mmol, 1.00 eq.) and sodium hydride (60% oil dispersion, 605 mg, 15.1 mmol, 1.00 eq.) was dissolved in anhydrous THF (75 mL) and stirred under ambient temperature for 30 min. Iodopentafluorobenzene (2.02 mL, 4.45 g,

15.1 mmol, 1.00 eq.) was added dropwise to the mixture. After 3 h water (50 mL) was added. The mixture was extracted with Et_2O (3x100 mL). The combined organic phases were washed with brine (50 mL) and dried over Na_2SO_4, evaporated and the residue was purified by column chromatography (DCM/hexane 1:1) to afford compound **S3.7** (3.66 g, 9.01 mmol, 60%) as a colorless oil.

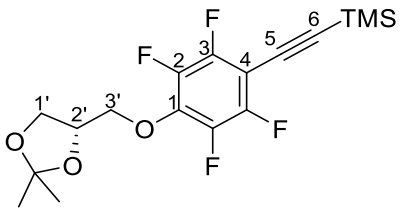
^{1}H NMR (400 MHz, $CDCl_3$): δ (ppm) = 4.44 (p, *J* = 5.6 Hz, 1H, 2'-H), 4.32 – 4.27 (m, 1H, 3'-H), 4.22 – 4.17 (m, 1H, 3'-H), 4.17 – 4.12 (m, 1H, 1'-H), 3.97 – 3.92 (m, 1H, 1'-H), 1.42 (q, *J* = 0.7 Hz, 3H, CH_3), 1.38 (q, *J* = 0.7 Hz, 3H, CH_3);

^{13}C{^{1}H} NMR (100 MHz, $CDCl_3$): δ (ppm) = 148.84 – 148.42 (m), 146.48 – 145.80 (m), 142.32 – 141.39 (m), 139.82 – 139.05 (m), 138.46 – 137.80 (m), 110.10 ($C(CH_3)_2$), 74.97 (t, *J* = 3.2 Hz, 3'-C), 74.13 (2'-C), 66.28 (1'-C), 64.53 (t, J = 28.2 Hz, 4-C), 26.78 (CH_3), 25.39 (CH_3);

^{19}F{^{1}H} NMR (376 MHz, $CDCl_3$): δ (ppm) = -120.74 – -121.37 (m, 2F, 3-F), -153.82 – -154.27 (m, 2F, 2-F);

HR-MS (ESI+): *m/z* calc. ($C_{12}H_{11}F_4IO_3Na$, $[M+Na]^+$): 428.95813, found: 428.95787.

Compound **S3.8**



Compound **S3.8** was synthesized according to Tschierske *et al.*[417] Under nitrogen atmosphere, CuI (79.7 mg, 419 µmol, 10 mol%) and $Pd(PPh_3)_4$ (242 mg, 209 µmol, 5.0 mol%) was dissolved in argon-purged Et_3N (20 mL). A solution of compound **S3.7** (1.70 g, 4.19 mmol, 1.00 eq.) in argon-purged Et_3N (7 mL) was added and stirred at ambient temperature for 10 min. Then, trimethylsilylacetylene (1.16 mL, 822 mg, 8.37 mmol, 2.00 eq) was added and stirring was continued for 20 h at 80 °C. The precipitate mixture was filtered off and washed with Et_2O. The filtrate was washed with a saturated solution of ammonium chloride (2x50 mL). The aqueous phase was extracted with Et_2O (2x50 mL). The combined organic phases were washed

with brine (50mL), dried over Na_2SO_4, evaporated and the residue was purified by column chromatography (DCM/hexane 1:1) to afford compound **S3.8** as a yellow oil (1.38 g, 3.67 mmol, 88%).

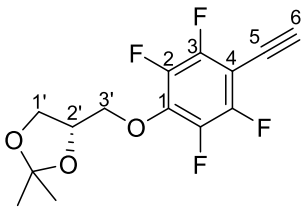
^{1}H NMR (400 MHz, $CDCl_3$): δ (ppm) = 4.48 – 4.39 (m, 1H, 2'-H), 4.33 – 4.27 (m, 1H, 3'-H), 4.24 – 4.18 (m, 1H, 3'-H), 4.17 – 4.11 (m, 1H, 2'-H), 3.97 – 3.91 (m, 1H, 2'-H), 1.42 (q, *J* = 0.8 Hz, 3H, CH_3), 1.37 (q, *J* = 0.7 Hz, 3H, CH_3), 0.27 (s, 9H, TMS);

^{13}C{^{1}H} NMR (100 MHz, $CDCl_3$): δ (ppm) = 149.24 – 148.35 (m), 146.92 – 145.96 (m), 142.82 – 141.39 (m), 140.15 – 139.37 (m), 138.50 – 137.67 (m, 1-C), 110.11 ($C(CH_3)_2$), 107.93 (t, *J* = 3.6 Hz, C-6), 98.65 (C-4), 88.49 (t, *J* = 3.7 Hz, C-5), 74.93 (t, *J* = 3.4 Hz, C-3'), 74.17 (C-2'), 66.24 (C-1'), 26.76 (CH_3), 25.40 (CH_3), -0.24 (TMS).;

^{19}F{^{1}H} NMR (376 MHz, $CDCl_3$): δ (ppm) = -134.68 – -140.68 (m, 2F, 2-F), -154.78 – -159.33 (m, 2F, 3-F);

HR-MS (ESI+): *m/z* calc. ($C_{17}H_{20}F_4O_3SiNa$, $[M+Na]^+$): 399.10100, found: 399.10083.

Compound **S3.9**



Compound **S3.8** (442 mg, 1.06 mmol, 1.00 eq.) was dissolved in THF/MeOH (1:1, 10 mL). K_2CO_3 (147 mg, 1.06 mmol, 1.00 eq.) was added and the mixture was stirred for 2 h at room temperature. The suspension was diluted with water (30 mL), extracted with Et_2O (3x50 mL) and washed with brine (2x30 mL). The organic phase was dried over Na_2SO_4, evaporated and the residue was purified by column chromatography (DCM/hexane 1:1) to afford compound **S3.9** as a light brown oil (312 mg, 1.03 mmol, 96%).

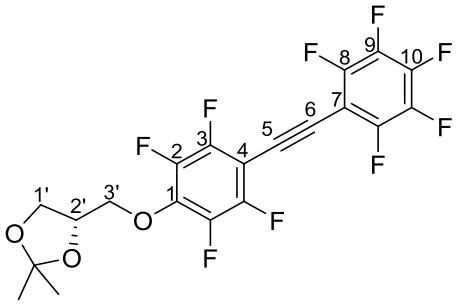
^{1}H NMR (400 MHz, $CDCl_3$): δ (ppm) = 4.48 – 4.40 (m, 1H, H-2'), 4.36 – 4.28 (m, 1H, H-3'), 4.27 – 4.19 (m, 1H, H-3'), 4.18 – 4.12 (m, 1H, H-3'), 3.99 – 3.91 (m, 1H, H-3'), 3.56 (q, *J* = 0.8 Hz, 1H, H-6), 1.42 (q, *J* = 0.8 Hz, 3H, CH_3), 1.38 (q, *J* = 0.7 Hz, 3H, CH_3).;

$^{13}C\{^{1}H\}$ NMR (100 MHz, $CDCl_3$): δ (ppm) = 149.96 – 148.75 (m), 147.36 – 146.35 (m), 142.65 – 141.50 (m), 140.05 – 139.24 (m), 139.04 – 138.20 (m), 110.13 ($C(CH_3)_2$), 97.33 (t, *J* = 18.1 Hz, C-4), 88.85 (t, *J* = 3.8 Hz, C-5), 74.94 (t, *J* = 3.4 Hz, C-3‘), 74.15 (C-2‘), 68.69 (t, *J* = 4.0 Hz, C-6), 66.20, 26.75, 25.38;

$^{19}F\{^{1}H\}$ NMR (376 MHz, $CDCl_3$): δ (ppm) = -136.42 – -138.69 (m, 2F, 2-F), -155.34 – -158.80 (m, 2F, 3-F);

HR-MS (ESI+): *m/z* calc. ($C_{14}H_{12}F_4O_3Na$, $[M+Na]^+$): 327.06148, found: 327.06102.

Compound **S3.10**



Under nitrogen atmosphere, CuI (17.2 mg, 90.3 µmol, 10 mol%), $Pd(PPh_3)_4$ (52.0 mg, 45.0 µmol, 5.0 mol%) and iodopentafluorobenzene (180 µL, 397 mg, 1.35 mmol, 1.50 eq.) were dissolved in argon-purged Et_3N (8.7 mL). A solution of compound **S3.9** (274 mg, 900 µmol, 1.00 eq.) in argon-purged Et_3N (2.4 mL) was added and stirring was continued for 20 h at 60°C. The precipitate mixture was filtered off and washed with Et_2O. The filtrate was washed with a saturated solution of ammonium chloride (2x30 mL). The aqueous phase was extracted with Et_2O (2 x 30 mL). The combined organic phases were washed with brine (30mL), dried over Na_2SO_4, evaporated and the residue was purified by column chromatography (DCM/hexane 1:1) to afford compound **S3.10** as a pale yellow solid (383 mg, 815 µmol, 91%).

^{1}H NMR (400 MHz, $CDCl_3$): δ (ppm) = 4.53 – 4.41 (m, 1H, 2’-H), 4.40 – 4.32 (m, 1H, 3’-H), 4.32 – 4.24 (m, 1H, 3’-H), 4.21 – 4.12 (m, 1H, 1’-H), 4.01 – 3.90 (m, 1H, 1’-H), 1.43 (q, *J* = 0.7 Hz, 3H, CH_3), 1.38 (q, *J* = 0.7 Hz, 3H, CH_3);

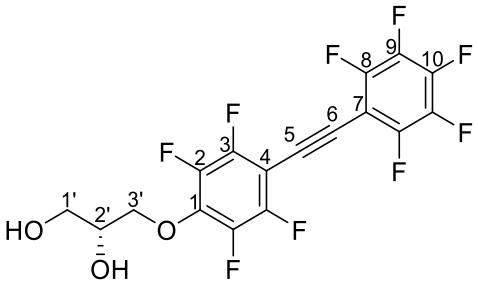
$^{13}C\{^{1}H\}$ NMR (100 MHz, $CDCl_3$): δ (ppm) = 148.99 – 148.16 (m), 146.70 – 145.67 (m), 142.53 – 141.72 (m), 139.95 – 138.69 (m), 136.77 – 136.14 (m), 110.19 ($C(CH_3)_2$), 99.29, 96.99, 87.74 –

85.67 (m), 85.00 – 81.68 (m), 74.97 (t, *J* = 3.5 Hz, 3'-C), 74.17 (2'-C), 66.17 (1'-C), 26.75 (CH_3), 25.38 (CH_3);

$^{19}F\{^1H\}$ NMR (376 MHz, $CDCl_3$): δ (ppm) = -133.73 – -135.61 (m, 2F), -135.92 – -136.54 (m, 2F), -150.24 (tt, *J* = 20.7, 2.4 Hz, 1F), -156.04 – -156.61 (m, 2F), -160.85 – -161.32 (m, 2F);

HR-MS (ESI+): *m/z* calc. ($C_{20}H_{11}F_9O_3Na$, $[M+Na]^+$): 493.04567 , found:. 493.04563

Compound **S3.11**



Compound **S3.10** (395 mg, 841 µmol, 1.00 eq.) was dissolved in THF (7 mL). 2 M Hydrochloric acid (7 mL) was added and the reaction mixture was stirred for 4 d at room temperature. The mixture was diluted with water (20 mL) and extracted with EtOAc (3x30 mL). The combined organic phases were washed with brine (20 mL) and dried over Na_2SO_4. The solvent was removed under reduced pressure and the residue was purified by column chromatography (hexane/EtOAc 3:2) to afford compound **S3.11** as a colorless solid (268 mg, 623 µmol, 74%).

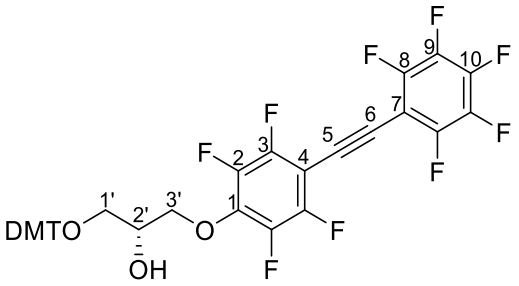
^1H NMR (400 MHz, $CDCl_3$): δ (ppm) = 4.49 – 4.27 (m, 2H, 3'-H), 4.12 (p, *J* = 4.4 Hz, 1H, 2'-H), 3.93 – 3.69 (m, 2H, 1'-H), 2.58 (s_{br}, 1H, OH), 1.91 (s_{br}, 1H, OH);

$^{13}C\{^1H\}$ NMR (100 MHz, $CDCl_3$): δ (ppm) = 86.11 – 85.84 (m), 84.12 – 83.71 (m), 76.35 (t, *J* = 3.4 Hz, 3'-C), 70.69 (2'-C), 63.10 (1'-C);

$^{19}F\{^1H\}$ NMR (376 MHz, $CDCl_3$): δ (ppm) = -134.44 – -134.94 (m, 2F), -135.73 – -136.17 (m, 2F), -150.14 (tt, *J* = 20.8, 2.5 Hz, 1F), -156.27 – -156.68 (m, 2F), -160.85 – -161.20 (m, 2F);

HR-MS (ESI+): *m/z* calc. ($C_{17}H_7F_9O_3Na$, $[M+Na]^+$): 453.01437 , found: 453.01494.

Compound **S3.12**



Under nitrogen atmosphere, compound **S3.11** (200 mg, 465 µmol, 1.00 eq.) was dissolved in anhydrous pyridine (6.5 mL). 4,4'-dimethoxytrityl chloride (205 mg, 604 µmol, 1.30 eq.) was added in small portions over a time period of 30 min and the reaction mixture was stirred at ambient temperature for 4 h. The mixture was diluted with DCM (50 mL), washed with sodium hydrogen carbonate solution (5%, 3x30 mL) and dried over Na_2SO_4.The solvent was removed under reduced pressure and the residue was purified twice by column chromatography (hexane/EtOAc 4:1+1% Et_3N and DCM/hexane 1:1+1% Et_3N) to afford compound **S3.12** as a colorless foam (277 mg, 378 µmol, 81%).

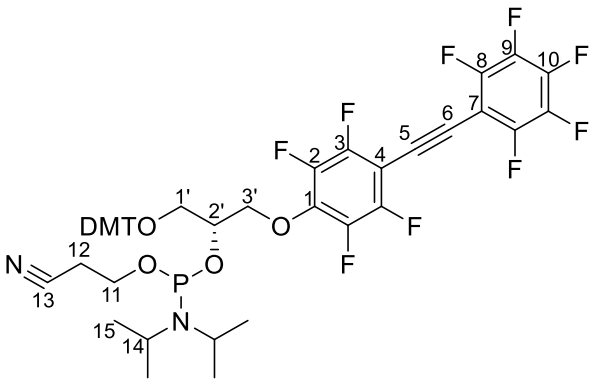
^{1}H NMR (400 MHz, $CDCl_3$): δ (ppm) = 7.43 – 7.38 (m, 2H, DMT-H), 7.34 – 7.25 (m, 6H, DMT-H), 7.25 – 7.20 (m, 1H, DMT-H), 6.86 – 6.80 (m, 4H, DMT-H), 4.45 – 4.30 (m, 2H, 3'-H), 4.15 – 4.05 (m, 1H, 2'H), 3.79 (s, 6H, DMT-H), 3.40 – 3.28 (m, 2H, 1'-H), 2.42 (d, *J* = 5.4 Hz, 1H, 2'-OH);

^{13}C{^{1}H} NMR (100 MHz, $CDCl_3$): δ (ppm) = 158.74 (DMT-C), 144.65 (DMT-C), 135.75 (DMT-C), 130.12 (DMT-C), 128.14 (DMT-C), 128.06 (DMT-C), 127.09 (DMT-C), 113.33 (DMT-C), 86.57 (DMT-C), 86.20 – 86.06 (m), 83.80 – 83.73 (m), 76.28 (t, *J* = 3.5 Hz, 3'-C), 70.01 (2'-C), 63.63 (1'-C), 55.36 (DMT-C).;

^{19}F{^{1}H} NMR (376 MHz, $CDCl_3$): δ (ppm) = -134.53 – -134.96 (m, 2F), -135.98 – -136.53 (m, 2F), -150.27 (tt, *J* = 20.8, 2.4 Hz, 1F), -155.97 – -156.83 (m, 2F), -160.68 – -161.51 (m, 2F);

HR-MS (ESI+): *m/z* calc. ($C_{38}H_{25}F_9O_5Na$, $[M+Na]^+$): 755.14505, found: 755.14427.

Compound **S3.13** = GTFF-PA



Compound **S3.12** (254 mg, 347 µmol, 1.00 eq.) was dissolved with DIPEA (354 µL, 269 mg, 2.08 mmol, 6.00 eq.) in anhydrous DCM (7 mL). After 20 min CEP-Cl (101 µL, 106 mg, 451 µmol, 1.30 eq.) was added. The reaction mixture was stirred additionally at ambient temperature for 3 h. The solvent was removed under reduced pressure and the residue was purified by column chromatography (hexane/EtOAc 5:1 + 1% Et_3N) to afford compound **S3.13** as a colorless foam (304 mg, 326 µmol, 94%).

^{1}H NMR (400 MHz, $CDCl_3$): δ (ppm) = 7.47 – 7.38 (m, 4H, DMT-H), 7.36 – 7.22 (m, 12H, DMT-H), 7.26 – 7.16 (m, 2H, DMT-H), 6.88 – 6.75 (m, 8H, DMT-H), 4.71 – 4.57 (m, 1H, 3'-H), 4.56 – 4.40 (m, 3H, 3'-H), 4.32 – 4.14 (m, 2H, 2'-H), 3.90 – 3.65 (m, 16H, DMT-H, 11-H), 3.64 – 3.52 (m, 2H, 14-H), 3.48 – 3.32 (m, 5H, 1'-H, 14-H), 3.25 (dd, J = 9.5, 6.6 Hz, 1H, 1'-H), 2.61 (t, J = 6.4 Hz, 2H, 12-H), 2.53 – 2.40 (m, 2H, 12-H), 1.28 – 0.99 (m, 24H, 15-H).;

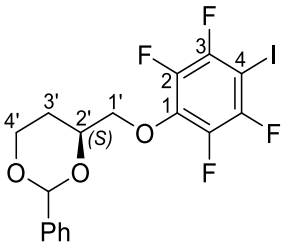
^{13}C{^{1}H} NMR (100 MHz, $CDCl_3$): δ (ppm) = 158.65 (DMT-C), 158.63 (DMT-C), 144.83 (DMT-C), 136.06 (DMT-C), 136.00 (DMT-C), 135.97 (DMT-C), 135.95 (DMT-C), 130.19 (DMT-C), 130.16 (DMT-C), 130.13 (DMT-C), 128.26 (DMT-C), 128.21 (DMT-C), 127.95 (DMT-C), 126.98 (DMT-C), 126.93 (DMT-C), 117.75 (13-C), 117.67(13-C), 113.22 (DMT-C), 86.39 (DMT-C), 86.34 (DMT-C), 75.49 (3'-C), 72.77 (2'-C), 72.60 (2'-C), 72.41 (2'-C), 72.25 (2'-C), 63.14 (1'-C), 63.13 (1'-C), 63.10 (1'-C), 58.52 (11-C), 58.39 (11-C), 58.33 (11-C), 58.21 (11-C), 55.36 (DMT-C), 55.35 (DMT-C), 43.41 (14-C), 43.32 (14-C), 43.29 (14-C), 43.19 (14-C), 24.82 (15-C), 24.80 (15-C), 24.75 (15-C), 24.73 (15-C), 24.56 (15-C), 24.54 (15-C), 24.49 (15-C), 24.47 (15-C), 20.45 (12-C), 20.39 (12-C), 20.34 (12-C), 20.27 (12-C);

^{19}F{^{1}H} NMR (376 MHz, $CDCl_3$): δ (ppm) = -134.66 – -134.90 (m, 4F), -136.45 – -136.83 (m, 4F), -150.07 – -150.61 (m, 2F), -155.59 – -156.33 (m, 4F), -160.87 – -161.25 (m, 4F);

$^{31}P\{^{1}H\}$ NMR (162 MHz, $CDCl_3$): δ (ppm) = 150.04 (t, *J* = 2.2 Hz, 1P), 149.72 (t, *J* = 2.1 Hz, 1P);

HR-MS (ESI+): *m/z* calc. ($C_{47}H_{42}F_9O_6N_2PNa$, $[M+Na]^+$): 955.25290, found: 955.25188.

Compound **S3.14**



4-(*S*)- Hydroxmethyl-2-phenyl-1,3-dioxan (**3.3**) (1.00 g, 5.15 mmol, 1.00 eq.) and sodium hydride (60% oil dispersion, 206 mg, 5.15 mmol, 1.00 eq.) were dissolved in anhydrous THF (25 mL) and stirred at ambient temperature. Iodopentafluorobenzene (687 µL, 1.51 g, 5.15 mmol, 1.00 eq.) was added dropwise to the mixture. After 3 h additional iodopentafluorobenzene (138 µL, 303 mg, 1.03 mmol, 0.20 eq.) was added. The solvent was evaporated after 30 min and the residue was purified by column chromatography (hexane/EtOAc 9:1) to afford compound **S3.14** (1.61 g, 3.43 mmol, 67%) as a colorless solid.

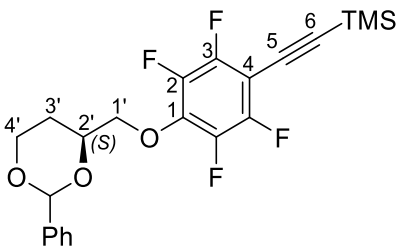
^{1}H NMR (400 MHz, $CDCl_3$): δ (ppm) = 7.42 – 7.29 (m, 5H, Ph-H), 5.54 (s, 1H, C*H*-Ph), 4.44 – 4.23 (m, 4H, 1'-H, 2'-H, 4'-H), 4.17 – 3.92 (m, 1H, 4'-H), 2.12 – 1.91 (m, 1H, 2'-H), 1.66 – 1.59 (m, 1H, 2'-H);

$^{13}C\{^{1}H\}$ NMR (100 MHz, $CDCl_3$): δ (ppm) = 138.12 (Ph-C), 129.03 (Ph-C), 128.36 (Ph-C), 125.97 (Ph-C), 101.18 (*C*H-Ph), 76.66 (t, *J* = 3.1 Hz, 1'-C), 76.03 (2'-C), 66.65 (4'-C), 64.23 (t, *J* = 28.1 Hz, 4-C), 27.15 (3'-C);

$^{19}F\{^{1}H\}$ NMR (376 MHz, $CDCl_3$): δ (ppm) = -121.33 – -121.66 (m, 2F, 3-F), -153.54 – -154.31 (m, 2F, 2-F);

HR-MS (ESI+): *m/z* calc. ($C_{17}H_{13}F_4O_3NaI$, $[M+Na]^+$): 490.97378, found: 490.97517.

Compound **S3.15**



Under nitrogen atmosphere, CuI (51.5 mg, 271 µmol, 10 mol%) and $Pd(PPh_3)_4$ (93.8 mg, 81.1 µmol, 3.0 mol%) were dissolved in argon-purged Et_3N (14 mL). A solution of compound **S3.14** (1.27 g, 2.71 mmol, 1 eq.) in argon-purged Et_3N (11 mL) was added and stirred at ambient temperature for 10 min. Then, trimethylsilylacetylene (487 µL, 345 mg, 3.52 mmol, 1.3 eq) was added and stirring was continued for 3 h at 60°C. The precipitate was filtered off and washed with Et_2O. The filtrate was washed with a saturated solution of ammonium chloride (2x50 mL) and brine (2x50 mL). The organic phases were dried over Na_2SO_4, evaporated and the residue was purified by column chromatography (hexane/EtOAc 20:1–15:1) to afford compound **S3.15** as a brown oil (1.07 g, 2.44 mmol, 90%).

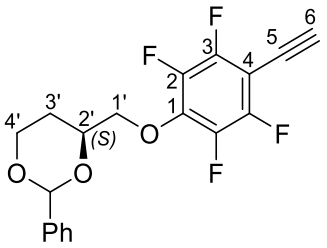
^{1}H NMR (400 MHz, $CDCl_3$): δ (ppm) = 7.46 – 7.28 (m, 5H, Ph-H), 5.54 (s, 1H, C*H*-Ph), 4.43 – 4.24 (m, 4H, 1'-H, 2'-H, 4'H), 4.10 – 3.95 (m, 1H, 4'-H), 2.07 – 1.91 (m, 1H, 3'-H), 1.67 – 1.58 (m, 1H, 3'-H), 0.28 (s, 9H, TMS);

^{13}C{^{1}H} NMR (100 MHz, $CDCl_3$): δ (ppm) = 149.21 – 148.81 (m), 146.71 – 146.31 (m), 142.44 – 142.02 (m), 140.01 – 139.56 (m), 138.97 – 138.42 (m), 138.15 (Ph-C), 129.02 (Ph-C), 128.36 (Ph-C), 126.01 (Ph-C), 107.73 (t, *J* = 3.7 Hz, 6-C), 101.21 (*C*H-Ph), 98.47 (t, *J* = 18.6 Hz), 88.58 (t, *J* = 3.7 Hz, 5-C), 76.87 (1'-C), 75.89 (2'-C), 66.66 (4'-C), 27.22 (3'-C), -0.23 (TMS).

^{19}F{^{1}H} NMR (376 MHz, $CDCl_3$): δ (ppm) = -134.91 – -140.25 (m, 2F, 3-F), -155.45 – -158.93 (m, 2F, 2-F);

HR-MS (ESI+): *m/z* calc. ($C_{22}H_{22}F_4O_3NaSi$, $[M+Na]^+$): 461.11665, found: 461.11709.

Compound **S3.16**



Compound **S3.15** (371 mg, 846 µmol, 1 eq.) was dissolved in THF/MeOH (1:1, 8 mL). K_2CO_3 (234 mg, 1.69 mmol, 2.00 eq.) was added and the mixture was stirred for 2.5 h at room temperature. The suspension was diluted with water (30 mL), extracted with Et_2O (3x50 mL) and washed with brine (2x30 mL). The organic phases were dried over $MgSO_4$, evaporated and the residue was purified by column chromatography (hexane/EtOAc 20:1–15:1) to afford compound **S3.16** as a brown solid (267 mg, 729 µmol, 86%).

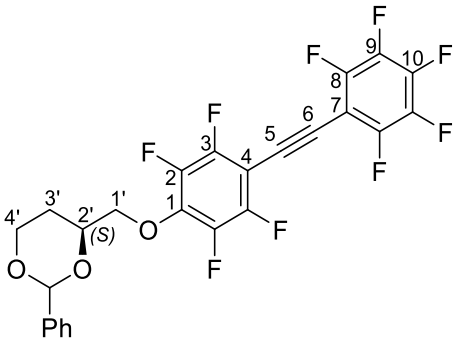
^{1}H NMR (400 MHz, $CDCl_3$): δ (ppm) = 7.41 – 7.29 (m, 5H, Ph-H), 5.54 (s, 1H, C*H*-Ph), 4.46 – 4.24 (m, 4H, 1'-H, 2'-H, 4'-H), 4.02 (ddd, J = 12.3, 11.5, 2.6 Hz, 1H, 4'-H), 3.55 (t, *J* = 0.8 Hz, 1H, 6-H), 2.05 – 1.91 (m, 1H, 3'-H), 1.66 – 1.58 (m, 1H, 3'-H);

^{13}C{^{1}H} NMR (100 MHz, $CDCl_3$): δ (ppm) = 149.60 – 149.10 (m), 147.08 – 146.65 (m), 142.55 – 141.88 (m), 140.00 – 139.42 (m), 139.48 – 138.87 (m), 138.13 (Ph-C), 129.06 (Ph-C), 128.37 (Ph-C), 125.99 (Ph-C), 101.21 (*C*H-Ph), 97.40 – 96.84 (m, 4-C), 88.68 (t, *J* = 3.8 Hz, 6-C), 76.89 (1'-C), 75.96 (2'-C), 68.77 (t, *J* = 4.0 Hz, 5-C), 66.65 (4'-C), 27.16 (3'-C);

^{19}F{^{1}H} NMR (376 MHz, $CDCl_3$): δ (ppm) = -134.91 – -140.00 (m, 2F, 3-F), -154.94 – -160.80 (m, 2F, 2-F);

HR-MS (ESI+): *m/z* calc. ($C_{19}H_{14}F_4O_3Na$, $[M+Na]^+$):389.07713, found:389.07619.

Compound **S3.17**



Under nitrogen atmosphere, CuI (34.0 mg, 179 µmol, 10 mol%) and $Pd(PPh_3)_4$ (62.2 mg, 53.6 µmol, 3.0 mol%) were dissolved in argon-purged Et_3N (13 mL). A solution of compound **S3.16** (654 mg, 1.79 mmol, 1.00 eq.) in argon-purged Et_3N (7.5 mL) was added and stirred at ambient temperature for 10 min. Then, iodopentafluorobenzene (262 µL, 578 mg, 1.96 mmol, 1.10 eq) was added and stirring was continued for 3 h at 80°C. The mixture was diluted with Et_2O (50 mL) and washed with a saturated solution of ammonium chloride (2x30 mL). The aqueous phase was extracted again with Et_2O (2x30 mL). The organic phases were dried over Na_2SO_4, evaporated and the residue was purified by column chromatography (hexane/EtOAc 20:1) to afford compound **S3.17** as a yellow solid (744 mg, 1.40 mmol, 78%).

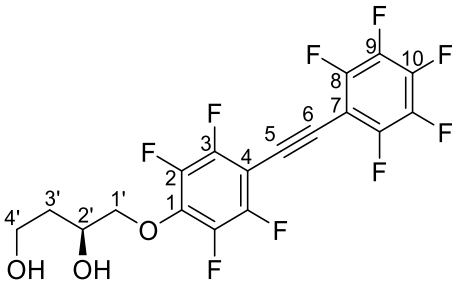
^{1}H NMR (400 MHz, $CDCl_3$): δ (ppm) = 7.40 – 7.31 (m, 5H, Ph-H), 5.54 (s, 1H, C*H*-Ph), 4.50 – 4.25 (m, 4H, 1'-H, 2'-H, 4'-H), 4.09 – 3.96 (m, 1H, 4'-H), 2.08 – 1.92 (m, 1H, 3'H), 1.70 – 1.57 (m, 1H, 3'-H);

^{13}C{^{1}H} NMR (100 MHz, $CDCl_3$): δ (ppm) = 138.09 (12-C), 129.08 (15-C), 128.36 (14-C), 125.97 (13-C), 101.22 (11-C), 99.28, 96.75, 86.17, 83.63, 76.88 (1'-C), 75.98 (2'-C), 66.63 (4'-C), 27.08 (3'-C);

^{19}F{^{1}H} NMR (376 MHz, $CDCl_3$): δ (ppm) = -134.10 – -135.15 (m, 2F), -136.31 – -136.87 (m, 2F), -150.36 (tt, *J* = 20.9, 2.4 Hz, 1F), -155.66 – -156.33 (m, 2F), -160.76 – -161.37 (m, 2F);

HR-MS (ESI+): *m/z* calc. ($C_{25}H_{13}F_9O_3Na$, $[M+Na]^+$): 555.06187, found: 555.06280.

Compound **S3.18**



Compound **S3.17** (455 mg, 854 µmol, 1.00 eq.) was dissolved in THF (8.6 mL). Then 2 M hydrochloric acid (8.6 mL) was added and the reaction mixture was stirred for 5 d at room temperature. The mixture was diluted with water (20 mL) and extracted with EtOAc (3x30 mL). The combined organic phases were washed with brine (20 mL). The solvent was removed under reduced pressure and the residue was purified by column chromatography (hexane/EtOAc 1:1 + 1% Et_3N) to afford compound **S3.18** as a colorless solid (164 mg, 369 µmol, 43%). Compound **S3.17** (208 mg. 391 µmol, 46%) could be recovered for further reactions.

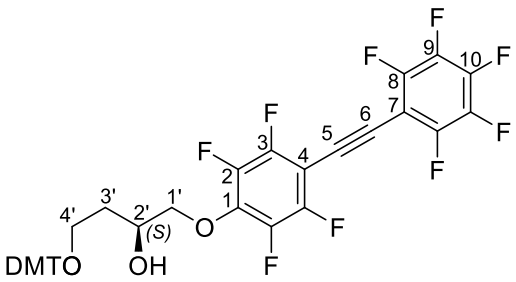
^{1}H NMR (400 MHz, $CDCl_3$): δ (ppm) = 4.37 – 4.18 (m, 3H, 1'-H, 2'-H), 3.99 – 3.87 (m, 2H, 4'-H), 2.97 (s_{br}, 1H, OH), 2.08 (s_{br}, 1H, OH), 1.90 – 1.75 (m, 2H, 3'-H);

^{13}C{^{1}H} NMR (100 MHz, $CDCl_3$): δ (ppm) = 79.11 (t, J = 3.3 Hz, 1'C), 70.25 (2'-C), 60.95 (4'-C), 34.21 (3'-C);

^{19}F{^{1}H} NMR (376 MHz, $CDCl_3$): (ppm) = -134.45 – -135.10 (m, 2F), -135.77 – -136.55 (m, 2F), -150.19 (tt, J = 20.8, 2.5 Hz, 1F), -156.11 – -156.67 (m, 2F), -160.73 – -161.59 (m, 2F);

HR-MS (ESI+): m/z calc. ($C_{18}H_9F_9O_3Na$, $[M+Na]^+$): 467.03002, found: 467.03038.

Compound **S3.19**



Under nitrogen atmosphere compound **S3.18** (160 mg, 361 µmol, 1.00 eq.) was dissolved in anhydrous pyridine (15 mL). 4,4'-dimethoxytrityl chloride (147 mg, 433 µmol, 1.20 eq.) was added in small portions over a time period of 10 min and the reaction mixture was stirred at ambient temperature for 23 h. The solvent was removed under reduced pressure and the residue was purified by column chromatography (hexane/EtOAc 7:1–6:1 +1% Et_3N) to afford compound **S3.19** as a colorless foam (203 mg, 272 µmol, 75%).

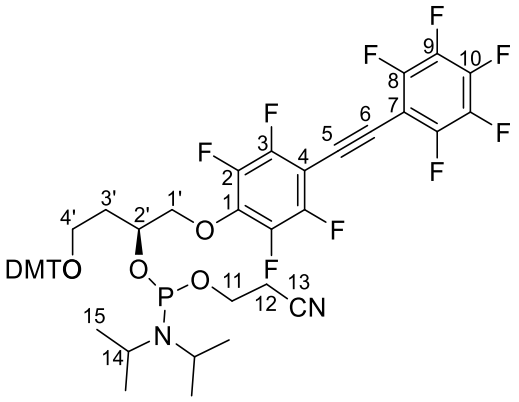
^{1}H NMR (400 MHz, $CDCl_3$): δ (ppm) = 7.43 – 7.39 (m, 2H, DMT-H), 7.34 – 7.27 (m, 6H, DMT-H), 7.24 – 7.19 (m, 1H, DMT-H), 6.86 – 6.81 (m, 4H, DMT-H), 4.31 – 4.12 (m, 3H, 1'-H, 2'-H), 3.79 (s, 6H, DMT-H), 3.45 – 3.25 (m, 2H, 4'-H), 3.09 (d, J = 3.2 Hz, 1H, 2'-OH), 1.98 – 1.81 (m, 2H, 3'-OH);

^{13}C{^{1}H} NMR (100 MHz, $CDCl_3$): δ (ppm) = 158.67 (DMT-C), 144.78 (DMT-C), 136.00 (DMT-C), 135.92 (DMT-C), 130.06 (DMT-C), 128.10 (DMT-C), 128.08 (DMT-C), 127.02 (DMT-C), 113.32 (DMT-C), 86.90 (DMT-C), 78.69 (t, J = 3.4 Hz, 1'-C), 69.83 (2'-C), 61.22 (4'-C), 55.35 (DMT-C), 32.74 (3'-C);

^{19}F{^{1}H} NMR (376 MHz, $CDCl_3$): δ (ppm) = -134.69 – -134.83 (m, 2F), -135.86 – -136.94 (m, 2F), -150.30 (tt, J = 20.7, 2.3 Hz, 1F), -155.72 – -157.26 (m, 2F), -160.63 – -161.59 (m, 2F);

HR-MS (ESI+): m/z calc. ($C_{39}H_{27}F_9NaO_5$, $[M+Na]^+$): 769.16070, found: 769.16267.

Compound **S3.20** = BTFF-PA



Under nitrogen atmosphere, compound **S3.19** (148 mg, 198 µmol, 1.00 eq.) was dissolved with DIPEA (202 µL, 153 mg 1.19 mmol, 6.00 eq.) in anhydrous DCM (6 mL). After 10 min CEP-

Cl (60.8 mg, 257 µmol, 1.30 eq.) was added. The reaction mixture was stirred at ambient temperature for 2 h. The solvent was removed under reduced pressure and the residue was purified by column chromatography (hexane/EtOAc 7:1 + 1% Et_3N) to afford compound **S3.20** as a colorless foam (156 mg, 165 µmol, 83%).

^{1}H NMR (400 MHz, $CDCl_3$): δ (ppm) = 7.52 – 7.37 (m, 4H, DMT-H), 7.34 – 7.22 (m, 12H, DMT-H), 7.24 – 7.17 (m, 2H, DMT-H), 6.84 – 6.78 (m, 8H, DMT-H), 4.53 – 4.42 (m, 1H, 1'-H), 4.39 – 4.24 (m, 5H, 1'-H, 2'-H), 3.87 – 3.70 (m, 13H, 11-H, DMT-H), 3.70 – 3.53 (m, 3H, 11-H), 3.56 – 3.28 (m, 4H, 14-H), 3.31 – 3.17 (m, 4H, 4'-H), 2.63 – 2.55 (m, 2H, 12-H), 2.51 – 2.42 (m, 2H, 12-H), 2.16 – 1.96 (m, 4H, 3'-H), 1.19 – 0.98 (m, 24H, 15-H);

^{13}C{^{1}H} NMR (100 MHz, $CDCl_3$): δ (ppm) = 158.54 (DMT-C), 158.53 (DMT-C), 145.19 (DMT-C), 145.14 (DMT-C), 136.45 (DMT-C), 136.36 (DMT-C), 136.34 (DMT-C), 130.16 (DMT-C), 130.13 (DMT-C), 130.11 (DMT-C), 130.09 (DMT-C), 128.27 (DMT-C), 128.19 (DMT-C), 127.91 (DMT-C), 127.90 (DMT-C), 126.86 (DMT-C), 126.83 (DMT-C), 117.73 (13-C), 113.16 (DMT-C), 113.13 (DMT-C), 86.30 (DMT-C), 86.28 (DMT-C), 71.80 (2'-C), 71.63 (2'-C), 70.75 (2'-C), 70.60 (2'-C), 59.83 (4'-C), 59.67 (4'-C), 58.46 (11-C), 58.27 (11-C), 58.15 (11-C), 57.96 (11-C), 55.34 (DMT-C), 55.32 (DMT-C), 43.31 (14-C), 43.21 (14-C), 43.19 (14-C), 43.09 (14-C), 33.27(3'-C), 33.22 (3'-C), 33.18 (3'-C), 24.82 (15-C), 24.75 (15-C), 24.68 (15-C), 24.62 (15-C), 24.55 (15-C), 24.47 (15-C), 24.46 (15-C), 24.38 (15-C), 20.38 (12-C), 20.31 (12-C);

^{19}F{^{1}H} NMR (376 MHz, $CDCl_3$): δ (ppm) = -134.64 – -135.02 (m, 4F), -136.41 – -137.10 (m, 4F), -150.17 – -150.76 (m, 2F), -155.51 – -156.18 (m, 4F), -160.87 – -161.52 (m, 4F);

^{31}P{^{1}H} NMR (162 MHz, $CDCl_3$): δ (ppm) = 149.16 (t, J = 2.0 Hz, 1P), 148.63 (t, J = 2.7 Hz, 1P);

HR-MS (ESI+): m/z calc. ($C_{48}H_{44}F_9N_2NaO_6P$, $[M+Na]^+$): 969.26855, found: 969.26957.

3.4.2.4 Oligonucleotide synthesis

Oligonucleotides were prepared on an Applied Biosystems ABI 392 DNA/RNA synthesizer on a 0.6 µmol scale using standard phosphoramidite chemistry. DMT-dC(ac)-CPG and DMT-dG(dmf)-CPG with a pore size of 1000 Å and loading density of 25–35 µmol/g were used as solid supports and were obtained from Sigma Aldrich. DMT-dA(bz), DMT-dC(ac), DMT-dG(dmf), DMT-dT were purchased from Chem Genes Corporation. Hex-5-yn-1-*O*-(2-

cyanoethyl-*N*,*N*-diisopropyl)phosphoramidite[418] and 3-(4,4'-dimethoxytrityloxy)propyl-1-(2-cyanoethyl-*N*,*N*-diisopropyl)phosphoramidite[419] were synthesized as previously described. Solutions of standard phosphoramidites (70 mM) and synthesized phosphoramidites (80 mM) were prepared in anhydrous MeCN. Following solutions were used during synthesis:

Activator: 0.25 M ethylthiotetrazole (ETT) in anhydrous acetonitrile,
Oxidation: 20 mM iodine in THF/water/pyridine (66:12:22, *v*/*v*/*v*)
Cap A: pyridine/acetic anhydride/THF (10/10/80, *v*/*v*/*v*)
Cap B: NMI in THF (84/16, *v*/*v*).
Deprotection: 3% trichloro acetic acid in 1,2-dichloroethane

Cleavage from the solid support and removal of the base labile protecting groups were performed by treatment with concentrated ammonium hydroxide (33% NH_3) at 25°C overnight in a 1.5 mL screw-capped tube. The solid support was filtered off and the solvent was removed under reduced pressure. The residue was dissolved in water. The crude product was purified by denaturing PAGE. Gels (0.7x200x300 mm) were prepared using a 20% acrylamide solution containing 7 M urea in 1x TBE. After polymerization gels were run with 1x TBE buffer (89 mM Tris, 89 mM boric acid, 2 mM EDTA, pH 8.3) for 2.5 h at a constant power (35 W). Product bands were visualized with a TLC plate under UV irradiation. The oligonucleotides were extracted with TEN buffer (10 mM Tris-HCl, 0.1 mM EDTA, 300 mM NaCl, pH 8.0) and recovered after precipitation with ethanol. DNA strands containing the TFF moiety needed an additional RP-HPLC purification step. Purification was performed on an ÄKTAmicro from GE Healthcare using a Nucleodur reversed-phase column (C18, 250x8 mm, 100 Å, 7 µm) from Machery Nagel at a flow rate of 2 mL min^{-1}. Linear gradients from 10–40% B over 50 min of buffer A (100 mM TEAA in water, pH 7.0) and buffer B (100 mM TEAA in MeCN/water 9:1, *v*/*v*) were used for purification. Chromatograms were monitored at 260 nm and 320 nm. All runs were performed at 40°C. HPLC purified oligos were lyophilized, dissolved in TEN buffer and precipitated with ethanol.

Purified oligonucleotides were analyzed by anion exchange HPLC using an ÄKTApurifier from GE Healthcare with a DNAPac PA 200 column (2x250 mm) from Thermo Scientific at a flow rate of 0.5 mL min^{-1}. Linear gradients from 0–48% B over 24 min of buffer A (25 mM Tris-HCl,

6 M urea, pH 8.0) and buffer B (25 mM Tris-HCl, 0.5 M $NaClO_4$, 6 M urea, pH 8.0) were used for analysis. Chromatograms were monitored at 260 nm. All analyses were performed at 80°C. High resolution ESI mass spectra were recorded on a Bruker micrOTOF-Q III spectrometer.

3.4.2.5 Labeling of 5'-alkyne functionalized oligonucleotides

5'-Alkyne functionalized oligonucleotides were fluorescently labeled using copper(I)-catalyzed alkyne-azide cycloaddition (CuAAC). Sulfo-Cy3-azide was obtained by Lumiprobe GmbH or Sulfo-Cy5-azide was purchased from Jena Bioscience GmbH. The freeze-dried DNA oligonucleotide (5 nmol) was dissolved in water (5 µL) and mixed with a DMSO/^{t}BuOH mixture (3:1, 3 µL). A solution of azide (0.63 µL, 50 mM) in DMSO/^{t}BuOH (3:1, v/v) was added. A freshly prepared solution of CuBr (0.63 µL, 100 mM) in DMSO/ ^{t}BuOH (3:1) was combined with a solution of tris(benzyltriazolylmethyl)amine (1.26 µL, 100 mM) in DMSO/ ^{t}BuOH (3:1) and then added to the reaction mixture. After 3 h of incubation in the dark at 37 °C, the reaction mixture was purified by PAGE (20% polyacrylamide).

3.4.2.6 UV/Vis spectroscopy / Thermal denaturing experiments

UV denaturing melting experiments were recorded on Varian Cary100 equipped with a 6x6 Multicell Block Peltier Series II cell changer and a VARIAN CARY Temperature Controller. Absorption was measured at 250, 260 and 280 nm. The absorption was recorded with a spectral bandwidth of 1 nm and the averaging time was set to 2s. The temperature cycle was programmed as follows:

20 °C to 90 °C
90 °C to 10 °C
10 °C to 90 °C
90 °C to 10 °C
10 °C to 90 °C

The first ramp was performed for annealing and was not considered for further melting temperature analysis. The heating rate was set to 0.5 °C min^{-1}. 500 µL sample in phosphate buffer (100 mM NaCl, 10 mM sodium phosphate, pH 7.0) with a duplex concentration of 1 µM, 2 µM and 5 µM was measured in 10 mm quartz cuvettes from VARIAN and 300µL sample in phosphate buffer (100 mM NaCl, 10 mM sodium phosphate, pH 7.0) with a duplex concentration of 10 µM and 20 µM were measured in 1 mm quartz cuvettes from HELLMA. The samples in the cuvettes were overlaid with silicone oil.

The obtained curves were fitted in a two-state transition model with upper and lower limit to obtain the melting temperature T_m. Absorption of the melted duplexes was used to recalculate $c_{total.}$ For the estimation of the thermodynamic parameters melting curves were analyzed according to Breslauer et al[420]:

$$\frac{1}{T_m} = \frac{(n-1)R}{\Delta H^0} \ln c_{total} + \frac{\Delta S^0 - (n-1) \cdot R \cdot \ln 2n}{\Delta H^0} \tag{3.1}$$

Assuming a bimolecular association of two non-self-complementary strands leads to a molecularity $n = 2$:

$$\frac{1}{T_m} = \frac{\mathrm{R} \ln c_{total}}{\Delta H^0} + \frac{\Delta S^0 - R \cdot \ln 4}{\Delta H^0} \tag{3.2}$$

Using van't Hoff analysis enthalpy ΔH^0 could be obtained from the slope and entropy ΔS^0 from the intercept of a linear fit of ln c_{total} vs $1/T_m$. Afterwards the free energy ΔG^{298} was calculated using the Gibbs-Helmholtz equation with T = 298 K:

$$\Delta G^{298} = \Delta H^0 - T\Delta S^0 \tag{3.3}$$

3.4.2.7 Calculation of double mutant cycle

The replacement with fluorinated tolanes leads to additional interactions with the surrounding DNA and the other tolane unit. The changes in free energy contributed to DNA/tolane interactions are defined as $\Delta\Delta G_{X7}$ for a substitution with a TFF at position X7, and $\Delta\Delta G_{Y18}$ for a substitution with a TFF at position Y18, respectively. $\Delta\Delta G_{het}$ is defined as the additionally energy gained by the introduction of a tolane heteropair. Starting from the THH homodimer the free energies of the other systems can be described as followed:

$$\Delta G_{\mathrm{FH}} = \Delta G_{\mathrm{HH}} + \Delta\Delta G_{\mathrm{X7}} + \Delta\Delta G_{\mathrm{het}} \tag{3.4}$$

$$\Delta G_{\mathrm{HF}} = \Delta G_{\mathrm{HH}} + \Delta\Delta G_{\mathrm{Y18}} + \Delta\Delta G_{\mathrm{het}} \tag{3.5}$$

$$\Delta G_{\mathrm{FF}} = \Delta G_{\mathrm{HH}} + \Delta\Delta G_{\mathrm{X7}} + \Delta\Delta G_{Y18} \tag{3.6}$$

Solving the system of linear equations leads to following terms for $\Delta\Delta G_{X7}$, $\Delta\Delta G_{Y18}$, $\Delta\Delta G_{het}$:

$$\Delta\Delta G_{\mathrm{X7}} = \frac{-\Delta G_{\mathrm{HH}} + \Delta G_{\mathrm{FH}} - \Delta G_{\mathrm{HF}} + \Delta G_{\mathrm{FF}}}{2} \quad (3.7)$$

$$\Delta\Delta G_{\mathrm{Y18}} = \frac{-\Delta G_{\mathrm{HH}} - \Delta G_{\mathrm{FH}} + \Delta G_{\mathrm{HF}} + \Delta G_{\mathrm{FF}}}{2} \quad (3.8)$$

$$\Delta\Delta G_{\mathrm{het}} = \frac{-\Delta G_{\mathrm{HH}} + \Delta G_{\mathrm{FH}} + \Delta G_{\mathrm{HF}} - \Delta G_{\mathrm{FF}}}{2} \quad (3.9)$$

3.4.2.8 Fluorescence spectroscopy

Fluorescence spectra and time-course measurements were recorded on a JASCO FP-8300 spectrofluorometer equipped with a FCT-817S cell changer and a F12 temperature control device from Julabo. All measurements were performed in a FMM-200 5 mm quartz fluorescence microcell from JASCO with a magnetic stir bar using a FMH-802 5 mm microcell jacket from JASCO as cell adapter. Emission spectra were recorded for Cy3 (excitation: 546 nm; emission: 561–750 nm) and Cy5 (excitation: 646 nm; emission: 661–750 nm) with following settings:

- Ex bandwidth: 5 nm
- Em bandwidth: 5 nm
- Response: 0.2 s
- PMT:voltage 480 V
- Data interval: 0.5 nm
- Scan speed: 500 nm/min

Time-course measurements wererecorded with following settings:

- Ex wavelength: 546 nm
- Em wavelength: 565 nm
- Ex bandwidth: 5 nm
- Em bandwidth: 5 nm
- Response: 0.2 s

- PMT:voltage 480 V
- Data interval: 0.2 s

3.4.2.9 FRET exchange experiments

A solution of a Cy3 and Cy5 labelled DNA duplex (0.104 µM, 270 µL) strands was prepared in phosphate buffer (104 mM NaCl, 10.4 mM phosphate, pH7.0). The duplex was annealed by heating to 95°C for 1 min and cooling down to room temperature over 30 min. The sample was transferred into a cuvette and incubated at 33°C for 5 min. Emission spectra of Cy3 and Cy5 was measured. The starting Cy3 (F_0) emission was recorded for 1 min and the FRET exchange experiment was started by adding an unlabeled single strand (2.8 µM, 10 µL), which sequence is comparable to the Cy5 labeled single strand. The increase of Cy3 emission (F) was recorded for 30 min if the labeled duplex contains a THH homopair or 60 min for a TFF homopair. An oligo with a C_3-spacer instead of the tolane moiety was added in excess (240 µM, 10 µL) to obtain the maximal Cy3 fluorescence (F_{total}) after 60 min. Emission spectra of Cy3 was measured to check if the Cy5 emission completely vanished.

The displacement curves were normalized with following equation:

$$F_{\mathrm{norm}} = \frac{F - F_0}{F_{\mathrm{total}} - F_0} \tag{3.10}$$

The resulted curve was fitted with a reaction kinetic first order to obtain the total strand displacement (F_{max}) and the displacement rate (k_{obs}):

$$F_{\mathrm{norm}} = F_{\mathrm{max}} \cdot (1 - e^{-x \cdot k_{\mathrm{obs}}}) + y_0 \tag{3.11}$$

To obtain the mean curve the first measurement point after the addition of the unlabeled single strand was set to zero.

3.4.2.10 NMR spectroscopy of DNA duplexes

After HPLC purification, all the oligonucleotides were precipitated overnight at -20°C with five volumes of $LiClO_4$ 2% w/v in acetone. The pellet was then solved in ddH_2O and lyophilized to remove residual acetone. The complementary oligonucleotides were annealed in the final

volume of the NMR sample (180 µL) by heating to 95°C 5 minutes and slowly cooling down to room temperature overnight. The samples used for NMR assignments contained 0.2-1 mM DNA duplex and were dissolved in either 90% H_2O/10% D_2O or 99.95% D_2O containing NMR buffer (20 mM NaPi buffer at pH 7.0, 100 mM NaCl).

The samples used for CLEANEX-PM experiments contained 0.2 mM DNA duplex and were dissolved in 90% H_2O/10% D_2O containing 10 mM NaPi buffer (pH 7.0) and 100 mM NaCl.

The same buffer stock solution was used for all the samples.

3-(Trimethylsilyl)-1-propanesulfonic acid (DSS) was added to all the samples as internal reference for ^{1}H. ^{13}C, ^{19}F and ^{31}P were indirectly referenced from the ^{1}H chemical shift.[421-422]

Nearly complete assignment of ^{1}H, ^{13}C, ^{19}F and ^{31}P resonances was achieved by means of 2D homo- and heteronuclear NMR experiments and following established assignment strategies (*i.e.*, imino-imino NOESY walk and sequential intra-strand H1' (n)-H8/6(n+1)-H1' (n+1) NOESY walk).

All the NMR experiments were performed on a Bruker Avance III 600 NMR spectrometer equipped with a DCH ^{13}C / ^{1}H cryoprobe, a Bruker Avance III 600 NMR spectrometer equipped with a BBFO room temperature probe or a Bruker Avance III HD 400 equipped with a BBFO room temperature probe.

The NMR spectra were acquired and processed using the software Topspin 3.2 (Bruker BioSpin, Germany). The spectra analysis was performed using Sparky 3.114[423] or NMRFAM-SPARKY.[424] Integrals of the NOESY cross peaks shown in Figure **S3.18**-S3.19 were calculated using the integration module of Sparky 3.114 (integration method: Gaussian fit).

2D ^{1}H,^{1}H NOESY spectra recorded on the samples in 90% H_2O/10% D_2O used jump-return-echo (excitation maximum on the middle of the imino region)[425] or excitation sculpting with gradients[426] for water suppression and contained a 200 ms mixing time. 1D ^{19}F NMR and 2D ^{1}H,^{19}F HOESY spectra were recorded on the samples containing fluorinated tolane units in 90% H_2O/10% D_2O. The 2D ^{1}H,^{19}F HOESY used excitation sculpting for water suppression[426] and contained a mixing time of 200 ms. 1D ^{31}P spectra were recorded with power-gated decoupling.

^{1}H,^{13}C HSQC for aromatic region ($^{1}J_{CH}$ 200Hz) and for sugar region ($^{1}J_{CH}$ 160Hz), DQF-COSY[427], TOCSY[426, 428], ^{1}H,^{1}H NOESY[429-430], ^{1}H,^{31}P HSQC spectra ($^{1}J_{PH}$ 20 Hz) were recorded on the samples in 99.95% D_2O for assignment. NOESY and COSY data show that the sugar pucker of most deoxyribose residues is C2'-endo, except for some residues adjacent to the modification site and closing base pairs.

Water hydrogen exchange rates of imino protons (k_{EX}) were measured using a 1D version of the CLEANEX-PM pulse sequence[410, 431-432] according to the protocol described before.[409, 433] A series of exchange experiments with different mixing times τ_m (5, 25, 50, 50, 100, 100, 150, 200, 300, 400, 400, 500 ms) were performed using the standard Bruker pulse program zgcxesgp, employing excitation sculpting with gradients as water suppression scheme.[426] The Bruker standard program zgesgp, employing as well excitation sculpting with gradients as water suppression scheme,[426] was used as reference experiment, without CLEANEX-PM element and water saturation pulse. 2560 scans were recorded for each experiment using an interscan delay of 2.0 s and 16 dummy scans. The spectra were processed using a line broadening factor of 10 Hz and the peak intensities were used for the data analysis.

The water ^{1}H longitudinal relaxation rate R_{1w} (s^{-1}) was determined for each sample using a saturation recovery experiment as described by Szulik *et al.*[434]

For each imino signal, the ratio between the intensity of peak in the CLEANEX-PM experiment (I) and the intensity of the peak in the reference experiment (I_0) was plotted as a function of the mixing time (τ_m). The imino proton water exchange rate k_{EX} (s^{-1}) and the apparent imino proton relaxation rate R_{1A} were obtained by fitting the resulting curve to the following equation:

$$\frac{I}{I_0} = \frac{k_{\mathrm{EX}}}{(R_{1\mathrm{A}} + k_{\mathrm{EX}} - R_{1\mathrm{W}})} \cdot (e^{-R_{1\mathrm{W}} \cdot \tau_{\mathrm{m}}} - e^{(-R_{1\mathrm{A}} + k_{\mathrm{EX}})\tau_{\mathrm{m}}}) \qquad (3.12)$$

With:

I = intensity of the imino peak at the mixing time τ_m

I_0 = intensity of the imino peak in the reference experiment

k_{EX} = imino proton to water exchange rate (s^{-1})

τ_m = mixing time (s)

R_{1w} = water ^{1}H longitudinal relaxation rate (s^{-1})

R_{1A} = imino ^{1}H apparent relaxation rate (combination of longitudinal and transverse relaxation rate, s^{-1})

The fitting was performed with Mathematica 8.0. The error for the fitted parameter results from the fitting.

3.4.2.11 Computational methods

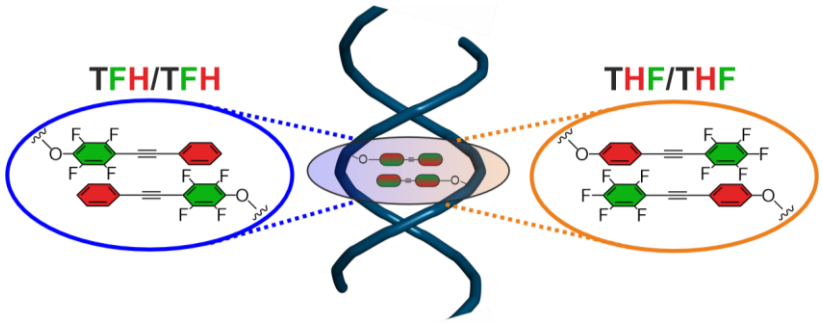
ORCA version 4.2.1[435-436] was used for the calculation of the DFT-optimized geometries and electrostatic potential maps. The acyclic backbone was replaced with a methyl group. B3LYP functional with D3BJ dispersion correction[437-438], a def2-TZVP basis set[439-440] and the corresponding auxiliary basis set for the RIJCOSX approximation[441] on all atoms was used for the calculation.

Electrostatic potential maps were generated from a cube file of the electron density with a resolution of 80×80×80 grid points using a python script that is available online.[442]

3.4.3 Appendix

The characterization of the synthesized products (NMR, HPLC), the melting curves and the FRET exchange experiments can be found in the published supporting information.[443]

4 Expansion of a tolane-fluorotolane base pair with self-recognition moieties



This chapter and the corresponding supporting information are unpublished:

H. Neitz, I. Bessi, J. Müller, M. Michel, C. Höbartner, **2023**, *Manuscript in preparation.*

Abstract: Noncovalent interactions provide an attractive way to control higher order assembly structures. The possibility to introduce polar π-stacking in a biological system by arene-fluoroarene brings the opportunity to expand the recognition motifs in DNA. This differs from most known unnatural base pairs, which use hydrogen bonding or hydrophobic shape complementarity. Here we present an expansion of supramolecular base pair replacement based on arene-fluoroarene interactions. The tolane moiety attached to the acyclic backbone of the glycol nucleic acid (GNA) and butyl nucleic acid (BuNA) was used as basis for a base pair replacement. The two nucleobase surrogates TFH and THF, each of which can form a self-complementary homopair, were generated by fluorination of one of the aromatic rings of the tolane moiety. It was found that the thermodynamic stability of the two base pairs is influenced by the structure of the backbone. A remarkable stabilization was observed for the THF homodimer attached to a BuNA backbone, which showed a 5 °C higher melting temperature compared to a T/A base pair. Based on this stabilization of the DNA, first attempts were made to introduce a recognition motif using three consecutive THF moieties, which resulted in a duplex-like (dimerized) assembly of a non-self-complementary DNA sequence.

4.1 Introduction

The combination of hydrogen bonding and stacking interactions guides the formation of the DNA duplex. Supramolecular synthons have been used as nucleobase replacements and led to the development of new artificial base pairs. Unnatural base pairs (UBPs)[10, 217, 272] exploit alternative hydrogen bonding patterns[265], metal coordination[338, 393] or hydrophobic interactions[282] and have been employed for a variety of applications such as structure stabilization[444-445], functionalization[268, 390, 446], site-specific labeling of nucleic acids[447-448] or even for the design of novel semi-synthetic organisms.[284, 449] Although π-stacking interactions contribute significantly to helix stabilization, the programmability of the DNA double helix is limited to hydrogen bonding of nucleobases or hydrophobic shape complementarity of suitable nucleobase analogs.[280, 301, 450] In contrast, the use of artificial aromatic surrogates allows pairing based on π-π stacking or dipole moments in the ring.[305, 308-310] Base pair replacements by aromatic rings could be developed in a homo- or heteropair design. Dispersion, electrostatic interactions, and solvophobic effects are the primary forces driving the assembly of aromatic hydrocarbons. Quadrupole electrostatic interactions play an important role in determining the stacking geometry of the aromatic rings. Same-sign quadrupole moments, like in arene-arene interactions, induce a T-shape or a slip stack arrangement.[293, 295, 315] In supramolecular chemistry, arene-fluoroarene interactions offer additional ways to exploit electrostatic interactions to control the stacking arrangement of aromatic rings. This interaction has been harnessed for the controlled assembly of organic materials[317, 396-397], polymers[318-319, 398] and crystals[320-321, 400]. Quadrupole moments of opposite sign, as found in arene-fluoroarene systems, promote a stacked alternate arrangement. The positive quadrupole moment of the fluoroarene is due to the electron-withdrawing nature of the fluorine atoms and leads to an enhanced electrostatic interaction with the negative quadrupole moment of the arene.[314] The use of arene-fluoroarene interactions in aqueous buffered systems is more challenging due to the rather weak binding affinity ($\Delta G \approx -1$ kcal mol^{-1}).[325] Arene moieties often stabilize the assembly due to the hydrophobic effect, but this can lead to uncontrolled aggregation.[296, 401] Nevertheless, several systems using arene-fluoroarene interactions in water have been reported to control the assembly of supramolecular structures and peptides.[327-328, 402] In DNA, fluorinated aromatic building blocks have also been used. For example, the shape complementarity of 2,4-difluorotoluene-2'-deoxyriboside with adenine allowed its sequence-specific

enzymatic incorporation.[404] Other attempts of selective base pairing with fluorinated base analogs have been described by Kool and Hunziker. In these systems, the higher stability of the modified duplexes was attributed to the hydrophobic character of the fluorinated structures and not to improved stacking capabilities.[330] Hunziker showed that electrostatic interactions of aromatic rings can control the DNA duplex formation if multiple arenes and fluoroarenes are arranged intramolecularly in an alternating fashion, but no additional stabilization was observed for an arene-fluoroarene pair compared to an arene or fluoroarene homopair replacing a canonical base pair.[329] Leumann and coworkers used the biphenyl moiety to implement arene-fluoroarene interactions in a zipper-like arrangement when incorporated at opposite positions in the DNA strands. The observed stabilization was of entropic origin and was explained in terms of the hydrophobic effect, as a result of the increased hydrophobicity of the fluorinated compound.[331] Alternative structural scaffolds are therefore needed to fully exploit the potential of quadrupole moments of opposite sign and enhanced electrostatic interactions.

Previously, we reported an artificial base pair based on the tolane moiety and its fluorinated analog. Using the acyclic backbone analog butyl nucleic acid (BuNA), a stabilization energy of −1.4 kcal mol^{-1} between the tolane heteropair was observed.[443] Here we present novel derivatives of this system, in with the tolane units are only partially fluorinated. Depending on the fluorination and the acyclic backbone used, two new homopairs are generated.

4.2 Results and discussion

4.2.1 Design of the study

The tolane moiety has an extended π-conjugated structure with two phenyl rings linked by a triple bond. Fluorination of one of the two rings as in the pentafluorinated tolane 1-pentafluorophenyl-2-phenylacetylene results in an unsymmetrically distributed electron density. Anti-parallel assembly of two such units allows for favorable arene-fluoroarene interactions within a homodimer.[451] In addition, this pentafluorinated tolane contains a strong dipol moment, which is caused by the distribution of electron density along the molecule. These considerations make the partially fluorinated tolane a good candidate as a supramolecular synthon.[452] A linker is necessary for the incorporation of the tolane moiety as a base pair substitute in a DNA strand. Therefore, it was attached as a phenyl ether to an acyclic linker in the phosphodiester backbone. The asymmetric composition of the partially fluorinated tolane

allows for two different attachment points for the phenyl ether, resulting in two different moieties. The introduction of the ether in the fluorinated aromatic ring results into the building block TFH, while the other building block THF is formed by the introduction of the phenyl ether on the protonated ring. Several types of acyclic backbone structures are compatible with artificial DNA structures and functions.[240] The higher flexibility of an acyclic backbone may facilitate the packing of the aromatic moieties within the DNA structure. Therefore, we chose to compare glycol nucleic acid (GNA)[249, 251] with butyl nucleic acid (BuNA)[258] units connected to the tolane derivatives. Depending on its artificial backbone, we named GTFH and GTHF the artificial base containing GNA attached to TFH or THF, respectively. BTFH and BTHF are the corresponding bases embedded in the BuNA backbone. In a head-to-tail arrangement of a homopair the aromatic rings with opposite quadrupole moments will be in a stacked arrangement (see Figure 4.1).

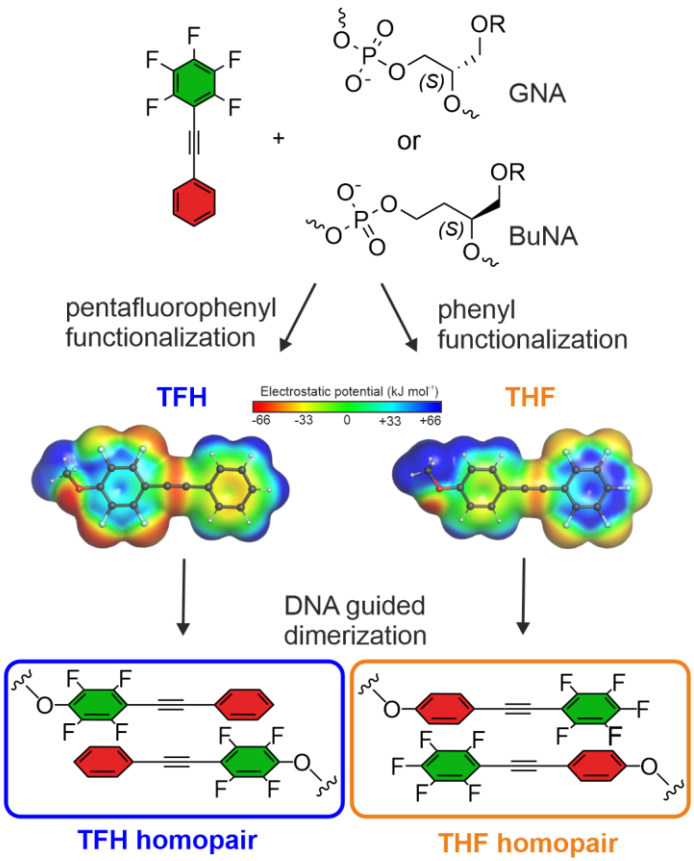


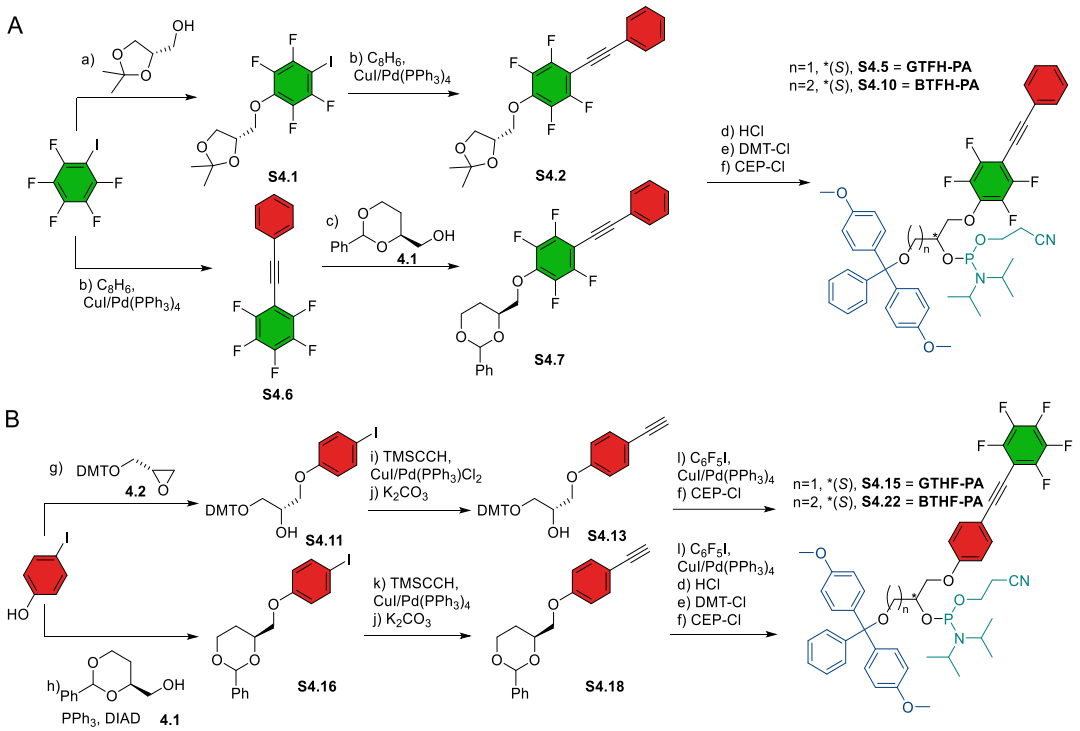
Figure 4.1 Concept and design of artificial base pairs TFH and THF using semifluorinated tolane and acyclic backbones GNA and BuNA to form self-recognizing homopairs in the DNA context.

The stacking arrangement has also been shown in the crystal structures of the corresponding tolane analogs, which contain an alkyl chain as phenyl ether residue.[323, 451] In addition, the

dipole moments of the monomers stabilize the homopair. Theoretical calculations of the THF building block showed a longitudinal molecular dipole moment of 4.84 D.[323] DFT calculations using a def2-TZVP basis provide a dipole moment of 1.32 D for the TFH and 4.53 D for the THF analog containing a methyl group as a residue. To test their base pair abilities depending on their electrostatic interactions, the building block was incorporated in a DNA dodecamer via phosphoramidite chemistry.

4.2.2 Synthesis

To incorporate the TFH or THF building blocks into a DNA strand, the corresponding phosphoramidites (PA) have been synthesized (see Scheme 4.1). For the synthesis of the TFH building block with the GNA backbone GTFH, a nucleophilic aromatic substitution was performed with D-(+)-solketal and iodopentafluorobenzene. The TFH tolane moiety was then completed via a Sonogashira reaction with phenylacetylene. For the BuNA analog BTFH, 1-pentafluorophenyl-2-phenylacetylene (**S4.6**) was prepared. Nucleophilic substitution with 4-(*S*)-hydroxmethyl-2-phenyl-1,3-dioxan (**4.1**) was then performed. After acidic acetal deprotection, the primary hydroxy group was protected with 4,4'-dimethoxytritylchloride (DMT-Cl) and converted to the corresponding 2-cyanoethyl *N*,*N*-diisopropyl phosphoramidites.



Scheme 4.1 Synthesis of phosphoramidite buildingblocks with the acyclic backbones GNA and BuNA containing the (A) TFH and (B) THF moiety. DIAD=diisopropyl azodicarboxylate, DMT-Cl=4,4'-dimethoxytritylchloride, CEP-Cl=2-cyanoethy *N,N*-diisopropylchlorophosphoramidite. For details see Schemes S4.1-S4.4.

The introduction of the acyclic backbone for the THF building blocks was achieved using two different strategies. For both phosphoramidites, 4-iodophenol was used as the starting material. For the GNA analog GTHF, nucleophilic ring opening of DMT-protected (*R*)-glycidol (**4.2**) under basic conditions was used, and for the BuNA analog BTHF, the backbone precursor 4-(*S*)-hydroxymethyl-2-phenyl-1,3-dioxan (**4.1**) was introduced via a Mitsunobu reaction. The tolane moiety for THF was completed by two Sonogashira reactions using TMS acetylene and iodopentafluorobenzene with a TMS deprotection in between. Additionally, the acetal protecting group of the BTHF precursor was removed and the primary alcohol was protected with DMT-Cl. The DMT-protected precursors were then converted to the corresponding phosphoramidites GTHF-PA and BTHF-PA. The modified oligonucleotides were obtained by solid-phase DNA synthesis, cleavage from controlled pore glass and deprotection with ammonium hydroxide, followed by PAGE purification. The purity and identity of the synthetic oligonucleotides were confirmed by analytical anion exchange HPLC and HR-ESI mass spectrometry (see Supporting information).

4.2.3 Thermodynamic analysis

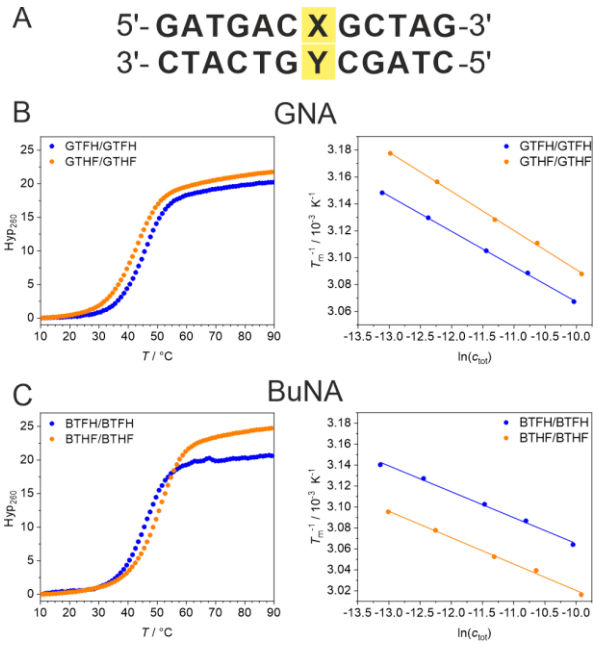


Figure 4.2 (A) Sequence of the DNA duplex containing the base surrogates at position X and Y. Melting curves and the corresponding van't Hoff analysis of the homopair containing DNA-duplexes for the (B) GNA and (C) BuNA backbone (1 µM DNA duplex in 10 mM phosphate buffer, 100 mM NaCl, pH 7.0).

Melting curve analyses and corresponding van't Hoff analyses were performed on the synthesized oligos. The two homopair arrangements for TFH and THF were tested with GNA or BuNA,

located in the center of an otherwise unaltered DNA duplex (see Figure 4.2A). Both homopairs showed a higher melting temperature with a BuNA backbone, suggesting a better tolerance for the base substitutions in a DNA duplex when attached to a BuNA backbone (see Table 4.1). Interestingly, the thermal stabilities of TFH and THF containing duplexes differ, depending on the backbone nature. In GNA, the TFH homopair (T_m = 44.5 °C) is about 3 °C higher than the THF homopair (T_m = 41.6 °C, see Figure 4.2B). The stabilization of the GTHF homopair appears to be entropically driven with a ΔS^0 of −189 cal mol^{-1} K^{-1}, compared to the GTFH/GTFH with a ΔS^0 of −210 cal mol^{-1} K^{-1}. This suggests that the fluorination of the outer ring leads to a stronger demand for the exclusion of water in the GNA context, resulting in a stabilization driven by the hydrophobic effect, instead of the expected polar π interactions. Using BuNA as a backbone replacement changes the trend of thermal stabilization (see Figure 4.2C). While the BTFH homodimer shows only a slight stabilization of 0.8 °C compared to the GNA analog, the BTHF homodimer shows a strong stabilization of 8.3 °C, which is also 5.1°C higher than a natural T/A base pair. The thermodynamic parameters for BTFH and BTHF homodimer show similar trends, with a ΔS^0 of about −220 cal mol^{-1} K^{-1}, which is less stabilizing than the GNA series, suggesting a stabilization that is not due to the hydrophobic effect. Given this interesting thermodynamic stabilization, the BTHF/BTHF interaction was studied in more detail using NMR spectroscopy to understand the structural arrangement of the homodimer within a DNA duplex.

Table 4.1 Melting temperatures and thermodynamic parameters of DNA duplexes containing a combination of TFH and THF homopairs at the center of the duplex.

Name	back-bone[a]	X	Y	T_m[b] [°C]	ΔT_m[c] [°C]	ΔH^{0}[d] [kcal mol^{-1}]	ΔS^{0}[d] [cal mol^{-1} K^{-1}]	ΔG^{298}[e] [kcal mol^{-1}]
GTFH/GTFH	G	TFH	TFH	44.5		−75.8 ± 0.5	−210 ± 1.3	−13.3 ± 0.6
BTFH/BTFH	B	TFH	TFH	45.3	+0.8	−80.4 ± 1.5	−224 ± 4.2	−13.6 ± 1.9
GTHF/GTHF	G	THF	THF	41.6		−68.3 ± 0.5	−189 ± 1.3	−12.1 ± 0.6
BTHF/BTHF	B	THF	THF	49.9	+8.3	−78.8 ± 1.2	−215 ± 3.4	−14.5 ± 1.6

[a] G (GNA), B (BuNA), [b] T_m at 1 µM DNA, in 10 mM phosphate buffer, 100 mM NaCl, pH 7.0. [c] Difference to T_m of BuNA and GNA derivative. [d] Derived from van't Hoff analyses with five concentrations (1, 2, 5, 10, 20 µM; Table S4.2). [e] $\Delta G^{298}=\Delta H^{0}-T\Delta S^{0}$. Calculated at 25 °C.

4.2.4 Structural analysis

NMR was used to gain insight into the architecture of the BTHF tolane units at the modification site. Nearly complete assignment of ^{1}H, ^{13}C, ^{19}F and ^{31}P resonances was achieved by means of

2D homo- and heteronuclear NMR experiments and following established assignment strategies (*i.e.*, imino-imino NOESY walk and sequential intrastrand H1'(n)-H8/6(n+1)-H1'(n+1) NOESY walk). Detailed assignment is reported in Figures S4.1-4.5. All the non-terminal imino protons were observed in the ^{1}H 1D NMR spectrum of the modified duplexes at 25°C, and assigned as indicated in Figure 4.3A. This indicates that no major distortion of the base pairs occurs due to the incorporation of the BTHF homopair. To further characterize the effect of the BTHF homopair on the base pair stability, CLEANEX-PM experiments were performed to obtain imino-water exchange rates (see Figure S4.6). Surprisingly, the exchange rates (k_{EX}) observed for the imino protons G8H1 and G19H1 flanking the BTHF homopair are comparable to those observed for the corresponding imino protons neighboring a Watson Crick T/A base pair.[409-410] So the BTHF homopair does not destabilize the adjacent base pairs and contributes to the stabilization of the duplex, in agreement with the high melting temperature observed for the BTHF/BTHF duplex.

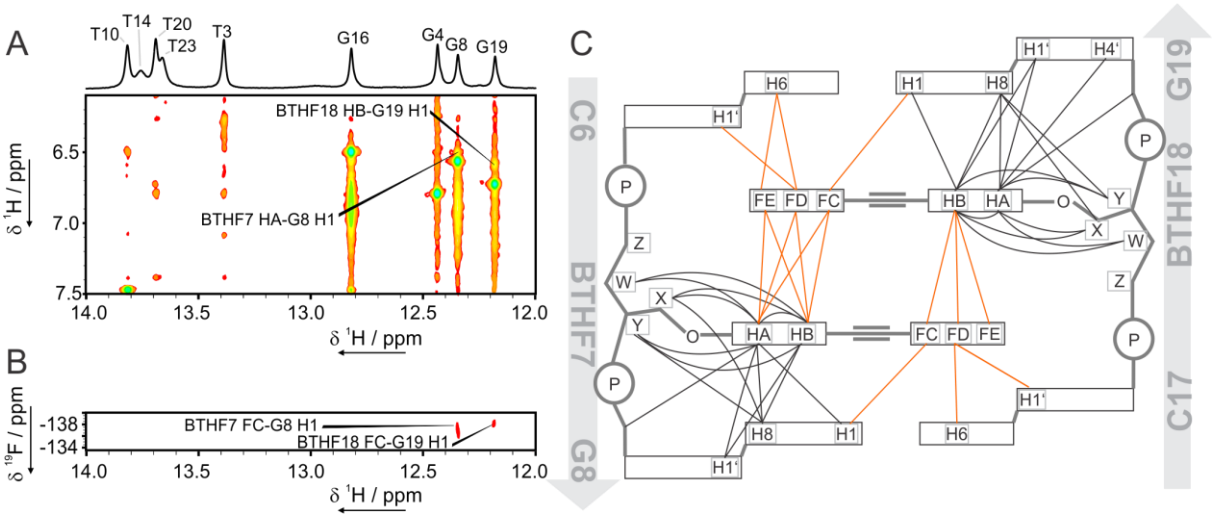


Figure 4.3 (A) Imino/aromatic region of ^{1}H,^{1}H NOESY spectrum of BTHF/BTHF and (B) ^{1}H,^{19}F HOESY with cross peaks between the tolane units and the imino protons of G8 and G19. (C) Schematic representation of all the homo- and heteronuclear NOE correlations with the tolane groups observed in ^{1}H,^{19}F HOESY (orange lines) and ^{1}H,^{1}H NOESY (black lines) spectra.

To gain further insight into the arrangement of the BTHF homopair within the DNA duplex, the NOESY and HOESY spectra were analyzed. For the imino protons G8H1 and G19H1 NOE and HOE intra-strand cross peaks to the neighboring tolane were observed. The NOE contact analysis was extended to all the residues surrounding the tolane. The results are summarized in the schematic drawing in Figure 4.3C and Figure S4.7. The observed contacts are consistent

with a head-to-tail arrangement where the tolane units are stacked on top of each other. The tolane units only show contacts with the flanking base pair and the other tolane. Importantly, inter-tolane contacts were observed only between the outer and inner ring, supporting the proposed arene-fluoroarene stacked arrangement of the aromatic rings.

4.2.5 Double mutant cycle

The fluorination of a tolane influences the interaction to the DNA and the opposite tolane. To understand these effects for both tolane analogs and to disentangle the changes in the interactions between DNA and the tolane pair, double mutant cycles were performed (see Figure 4.4). The double mutant cycle starts with a completely non-fluorinated tolane homopair. Introduction of TFH at position X results in a free energy change caused by the interaction of the DNA with the partially fluorinated tolane ($\Delta\Delta G_{X,TFH}$) and the electrostatic interaction with the other tolane ($\Delta\Delta G_{Tol}$). Accordingly, the introduction of THF at position Y results in a change of free energy caused by DNA ($\Delta\Delta G_{Y,THF}$) and tolane ($\Delta\Delta G_{Tol}$) interactions. Finally, the free energy of a duplex containing heteropair of THF and TFH was used, which will cancel the polar π stacking interaction $\Delta\Delta G_{Tol}$.

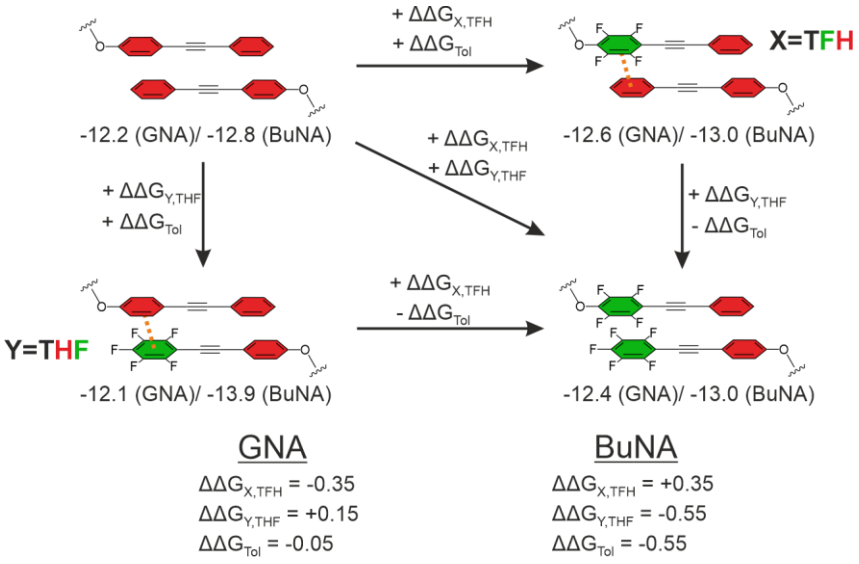


Figure 4.4 Double mutant cycle for introduction of TFH at position X and THF at position Y. Values are given in kcal mol^{-1}.

Thermodynamic parameters of all mixed tolane combinations were measured using van't Hoff analysis to perform the double mutant cycle (see Figure S4.8 and S4.9). The double mutant cycle shows an attractive interaction between the DNA and TFH ($\Delta\Delta G_{X,TFH}$ = −0.35 kcal mol^{-1}) and a destabilizing interaction with THF ($\Delta\Delta G_{Y,THF}$ = +0.15 kcal mol^{-1}) in the case of GNA. Interestingly, the interaction behavior of the tolane derivatives with DNA was reversed using the

BuNA backbone. Here, the interaction with TFH is unattractive ($\Delta\Delta G_{X,TFH}$ = +0.35 kcal mol^{-1}) and the interaction with THF is favorable ($\Delta\Delta G_{Y,THF}$ = −0.55 kcal mol^{-1}). When the position X and Y were switched for TFH and THF, the same trends were observed for GNA and BuNA (see Figure S4.10). The proposed polar π interaction between stacked arene fluoroarene is stronger in the case of BuNA with $\Delta\Delta G_{Tol}$ = −0.55 and −0.35 kcal mol^{-1} compared to GNA with $\Delta\Delta G_{Tol}$ = −0.05 and −0.15 kcal mol^{-1} confirming that the BuNA backbone is a privileged scaffold to establish arene-fluoroarene interactions for the tolane base pair surrogate.

4.2.6 Kinetic analysis

The distinct preference of GNA and BuNA for the TFH and THF building blocks, respectively, was also observed using a FRET-supported strand displacement experiment (see Figure 4.5). A duplex containing a THH homodimer was labeled at its 5' ends with the cyanine dyes Cy3 and Cy5, resulting in the formation of a FRET pair and quenching of the Cy3 fluorescence. A single strand containing the TFH or THF modifications and complementary to the Cy3-labeled strand was added to the duplex, resulting in strand exchange with the Cy5-labeled strand until equilibrium was reached. The exchange was monitored by the increase in Cy3 fluorescence. After 30 min, the maximal Cy3 fluorescence was determined by adding an excess of a complementary strand containing an alkyl linker instead of a tolane moiety. In the case of GNA, the exchange with a TFH containing DNA strand is higher (42.5±1.0 %), than with the THF containing DNA strand (36.0±1.0 %). For the BuNA backbone, the opposite behavior was again observed. The extent of strand displacement is higher upon addition of the THF containing single strand (59.5±3.3 %) than with the TFH single strand (50.8±1.1 %). The observed exchange rates were similar for both tolane derivatives in the same acyclic backbone substitution. The strand exchange is consistent with the higher preference of the TFH homopair for the GNA backbone and of the THF homopair for the BuNA backbone. The destabilization of THF in the GNA backbone could be due to the reduced space for the tolane in presence of a shorter acyclic backbone, which does not allow for efficient stacking. To avoid exposure of the outer fluorinated ring to water, the tolane pair disturbs the DNA, resulting in lower thermal stability. On the other side, the BuNA backbone has the optimal size for allowing the stacking of an homopair in the DNA context, resulting in a high melting temperature of the BTHF homodimer.

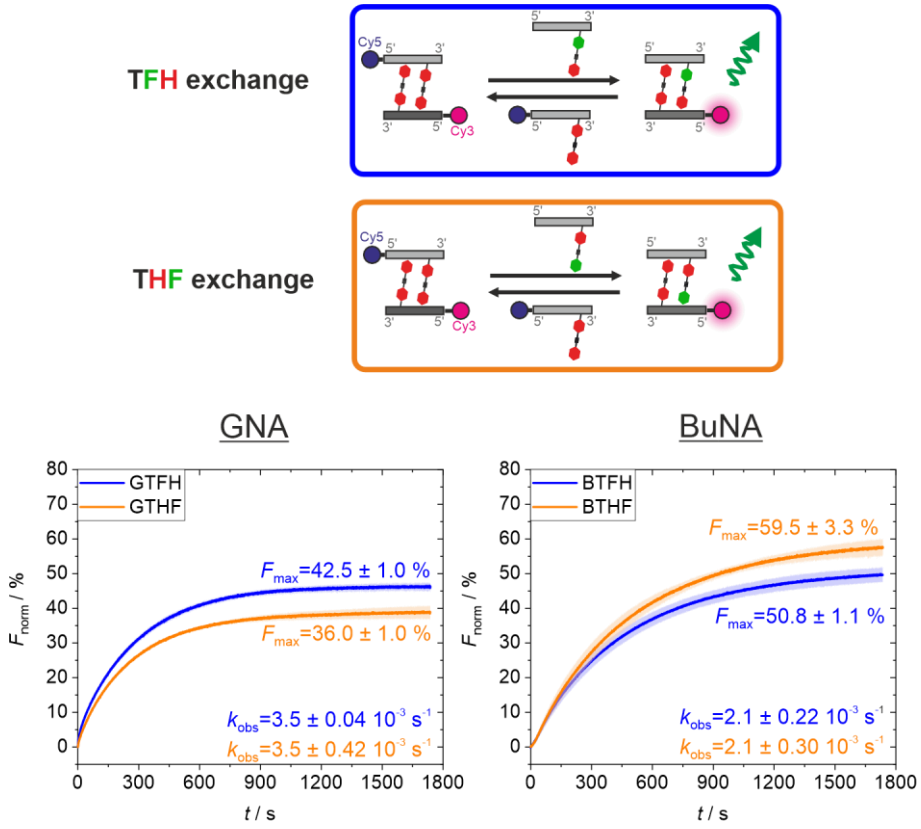


Figure 4.5 FRET supported strand displacement experiments using a double-labeled THH homopair duplex and a TFH (blue) or THF (orange) single strand as competitor. Experiments were performed in the context of GNA and BuNA.

4.2.7 Successive tolane incorporation

To take advantage of the DNA duplex stabilization provided the BTHF homodimer, we incorporated two and three BTHF units consecutively into a DNA strand to generate a self-recognizing motif. As a reference, the same series has been synthesized for the BTHH building block (see Table 4.2). Interestingly, an increase of 8.5 °C was observed for two BTHH homopairs compared to a single pair. With a difference of about −10 °kcal mol^{-1}, the higher stability seems to be enthalpy driven, suggesting that four THH moieties with a BuNA backbone allow a stacked arrangement, which is not possible for a single pair in the DNA context. The introduction of a third pair of BTHHs resulted in an increased melting temperature (55.1 °C) and a higher enthalpy (93.1 kcal mol^{-1}). For BTHF, the melting temperature increased by an average of 5 °C with the introduction of an additional BTHF homopair. The stabilization is more entropically driven by increasing number of tolanes. In all cases, the hyperchromicity decreases with increasing number of tolane homopairs (see Figure S4.11).

Table 4.2 Melting temperatures and thermodynamic parameters of DNA duplex 5'-d(GATGAC(**X**)$_n$GCTAG)3' / 3'-d(CTACTG(**Y**)$_n$CGATC)-5' containing 1-3 homopairs of BTHH and BTHF.

Homopair	n	T_m[a] [°C]	ΔT_m[b] [°C]	ΔH^0[c] [kcal mol^{-1}]	ΔS^0[c] [cal mol^{-1} K^{-1}]	ΔG^{298}[d] [kcal mol^{-1}]
(BTHH/BTHH)$_n$	1	42.0		−76.7 ± 0.7	−214 ± 2.1	−12.8 ± 1.0
	2	50.5	+8.5	−86.8 ± 1.1	−240 ± 3.0	−15.4 ± 1.4
	3	55.1	+4.6	−93.1 ± 1.6	−255 ± 4.3	−17.0 ± 2.0
(BTHF/BTHF)$_n$	1	49.9		−78.8 ± 1.2	−215 ± 3.4	−14.5 ± 1.6
	2	54.8	+4.9	−77.1 ± 2.3	−207 ± 6.3	−15.4 ± 3.0
	3	60.0	+5.2	−75.8 ± 2.4	−200 ± 6.4	−16.3 ± 3.0

[a] T_m at 1 µM DNA, in 10 mM phosphate buffer, 100 mM NaCl, pH 7.0. [b] $\Delta T_m = T_{m,n+1} - T_{m,n}$ [c] Derived from van't Hoff analyses with five concentrations (1, 2, 5, 10, 20 µM). [d] $\Delta G^{298} = \Delta H^0 - T\Delta S^0$. Calculated at 25 °C.

Thermal melting experiments of the single strands containing multiple tolane units were performed to test if they could form a recognition motif driving the formation of a duplex, with only partial base pairing of the surrounding DNA (see Figure 4.6). Only the DNA strand (BTHF)$_3$1 with the three incorporated BTHF tolanes showed a cooperative melting behavior with a melting temperature of 42.7 °C for a 5 µM sample. The formation of a duplex was confirmed by melting experiments performed at different duplex concentrations. The van't Hoff plot in Figure 4.6B shows a concentration dependent melting behavior, proving duplex formation. The melting temperature of 36.3 °C for a 1 µM sample is much lower compared to the fully complementary DNA duplex (BTHF/BTHF)$_3$, also with a ΔS^0 of −165 cal mol^{-1} K^{-1} the duplex formation seems to be entropically driven by the hydrophobic tolane units. However, base pairing of some bases is possible (see Figure 4.6C) if the oligonucleotide is arranged in an anti-parallel fashion. To show that the recognition is independent on the possible base pairing, the recognition motif consisting of three BTHF units was introduced at the 5' end of a DNA 8mer. Asanuma and coworkers introduced a triplet of *p*-methylstilbazoles attached to D-threoninols, which led to self-recognition due to stacking interactions of the aromatic rings and electrostatic interactions between the positively charged *p*-methylstilbazoles and the negatively charged phosphate backbone. Using it as a 5' overhang of a DNA, the triplet can recognize itself and form polymeric bonds between the two duplexes.[193] The same design was used for the BTHF trimer. Temperature dependent UV/Vis spectra were recorded for the duplex (BTHF/BTHF)$_3$ and the 5' modified BTHF trimer DNA duplex (see Figure S4.12). For the

(BTHF/BTHF)$_3$ duplex, the absorbance at 308 nm increases at higher temperatures, indicating a melting of the BTHF units. This melting behavior was not observed for the 5' modified BTFF trimer DNA duplex. The absence of absorption changes at 308 nm indicates that an interaction between two BTHF triplets do not occur. This suggests that the interaction of the BTHF triplet is not sufficient for self-recognition and that additional canonical base pairs must support the recognition.

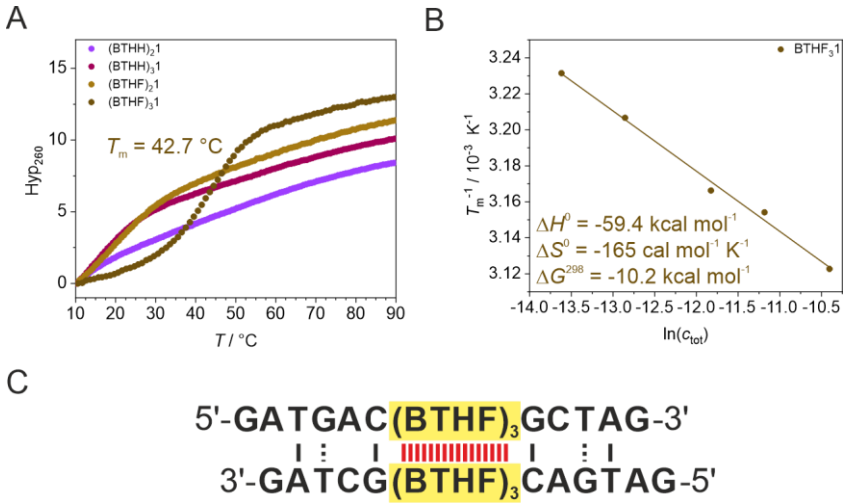


Figure 4.6 (A) Melting curves of the single strand 5'-d(GATGAC(**X**)$_n$GCTAG)3' for two or three incorporated BTHH or BTHF units (5 µM single strand concentration). (B) Van't Hoff analysis of the oligo BTHF$_3$1 containing three BTHF homopairs (10 mM phosphate buffer, 100 mM NaCl, pH 7.0). (C) Possible base pairing of the BTHF$_3$1 homodimer with recognition of the BTHF trimer (red lines), Watson-Crick base pairs (solid lines) and wobble base pairs (dashed lines).

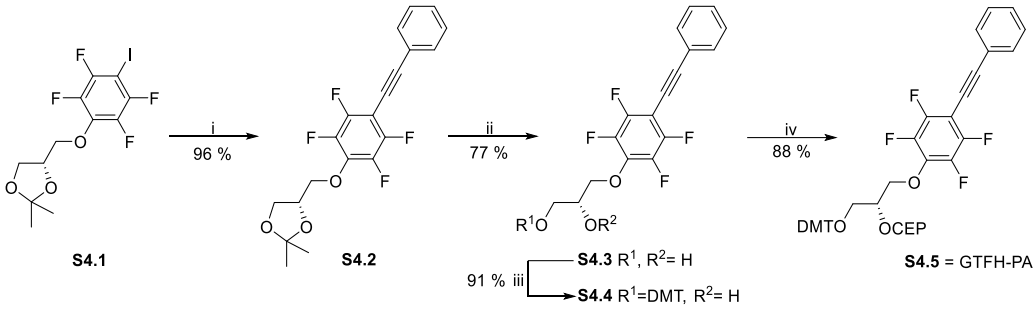
4.3 Conclusion

In conclusion, we introduced two new homopairs based on a semifluorinated tolane as artificial base pair analogs in DNA. The THF moiety shows a remarkable DNA duplex stabilization using BuNA as an acyclic backbone replacement. The high level of stabilization could be achieved by polar π interactions and additionally by dipole-dipole interactions resulting from the fluorination of the outer ring of the tolane moiety. NMR characterization supports a stacked conformation of the BTHF dimer in the DNA duplex. Analysis using a double mutation cycle and FRET-supported strand displacement revealed a preference in terms of duplex stabilization for the THF building block in the BuNA backbone. Conversely, the TFH building block is better accepted with a GNA backbone. To establish a recognition unit based on the backbone BTHF was incorporated consecutively multiple times inside the DNA. Incorporation of a BTHF triplet in the DNA context allows the formation of a homodimer formed by two identical strands. This is a promising result to establish the tolane moiety as self-recognizing motif

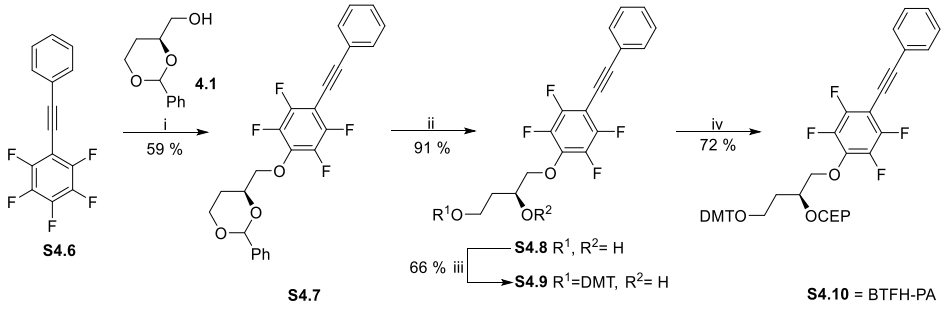
leading to amphiphilic supramolecular structures. Combined with the heterodimer pair THH/TFF the TFH and THF homodimers complete a supramolecular language in the DNA context based on the electrostatic optimization of the tolane moiety. This can be applied to the extension of DNA nanostructures to obtain an additional level of control by using arene-fluoroarene interaction orthogonal to DNA.

4.4 Supporting information

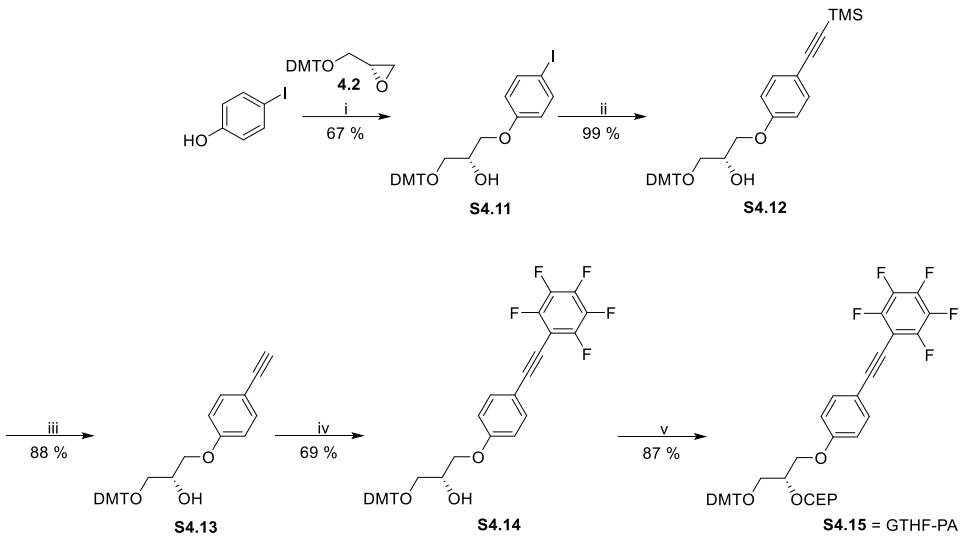
4.4.1 Additional data



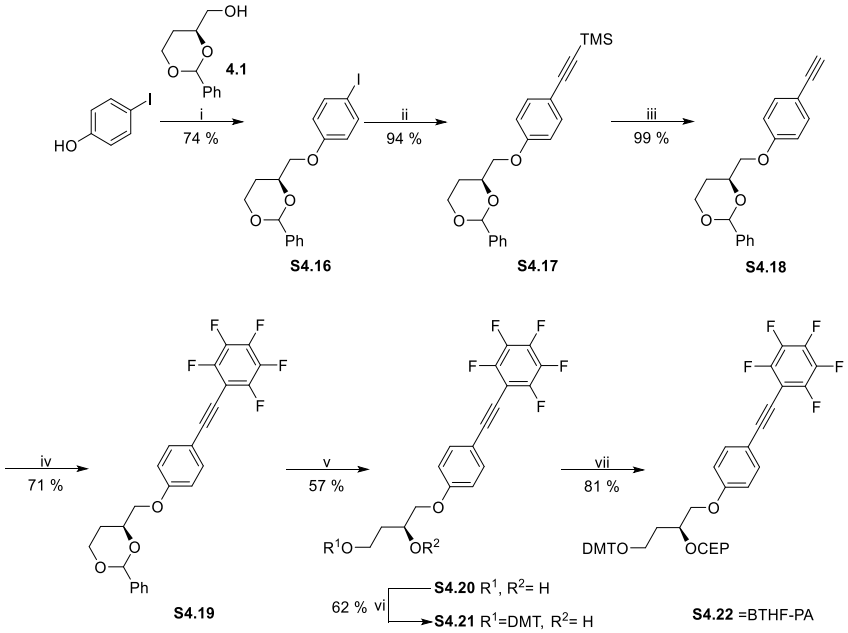
Scheme S4.1 Synthesis of GTFH phosphoramidite i) phenylacetylene, CuI, $Pd(PPh_3)_4$, 80 °C, 18 h; ii) 2 M HCl, THF, r.t., 4 d; iii) DMT-Cl, pyridine, r.t.; 4 h; iv) CEP-Cl, DIPEA, DCM, r.t., 3 h.



Scheme S4.2 Synthesis of BTFH phosphoramidite i) NaH, THF, r.t., 3.5 h; ii) *p*-toluenesulfonic acid, MeOH, r.t., 20 h; iii) DMT-Cl, pyridine, r.t.; 24 h; iv) CEP-Cl, DIPEA, DCM, r.t., 2 h.



Scheme S4.3 Synthesis of GTHF phosphoramidite i) NaH, DMF, 90 C, 20h; ii) trimethylsilylacetylene, CuI, $Pd(PPh_3)_2Cl_2$, 60 °C, 16 h; iii) K_2CO_3, THF/MeOH, r.t., 2 h; iv) C_6F_5I, CuI, $Pd(PPh_3)_2Cl_2$, 60 °C, 18 h; v) CEP-Cl, DIPEA, DCM, r.t., 3 h.



Scheme S4.4 Synthesis of BTHF phosphoramidite i) PPh_3, DIAD, THF, r.t., 30 min; ii) trimethylsilylacetylene, CuI, $Pd(PPh_3)_4$, 80 °C, 15 h; iii) K_2CO_3, THF/MeOH, r.t., 2 h; iv) C_6F_5I, CuI, $Pd(PPh_3)_2Cl_2$, 80 °C, 17 h; v) 4 M HCl, THF, r.t., 3 d; vi) DMT-Cl, pyridine, r.t.; 18 h; vii) CEP-Cl, DIPEA, DCM, r.t., 3 h.

Table S4.1. Sequences and ESI-MS results of the used DNA oligonucleotides.

Oligo	5'-Sequence-3'	Chemical formula	Calc. Mass	Measured Mass
GTFH1	GATGAC(GTFH)GCTAG	$C_{125}H_{146}F_4N_{45}O_{68}P_{11}$	3781.64052	3781.68775
GTFH2	CTAGC(GTFH)GTCATC	$C_{123}H_{147}F_4N_{38}O_{70}P_{11}$	3692.61665	3692.63333
GTHF1	GATGAC(GTHF)GCTAG	$C_{125}H_{145}F_5N_{45}O_{68}P_{11}$	3799.63110	3799.67232
GTHF2	CTAGC(GTHF)GTCATC	$C_{123}H_{146}F_5N_{38}O_{70}P_{11}$	3710.60723	3710.64565
BTFH1	GATGAC(BTFH)GCTAG	$C_{126}H_{148}F_4N_{45}O_{68}P_{11}$	3795.65617	3795.64847
BTFH2	CTAGC(BTFH)GTCATC	$C_{124}H_{149}F_4N_{38}O_{70}P_{11}$	3706.63230	3706.62075
BTHF1	GATGAC(BTHF)GCTAG	$C_{126}H_{147}F_5N_{45}O_{68}P_{11}$	3813.64675	3813.62694
BTHF2	CTAGC(BTHF)GTCATC	$C_{124}H_{148}F_5N_{38}O_{70}P_{11}$	3724.62233	3724.64650
$(BTHH)_2$1	GATGAC$(BTHH)_2$GCTAG	$C_{144}H_{169}N_{45}O_{73}P_{12}$	4067.77467	4067.77160
$(BTHH)_2$2	CTAGC$(BTHH)_2$GTCATC	$C_{142}H_{170}N_{38}O_{75}P_{12}$	3978.75080	3978.75321
$(BTHH)_3$1	GATGAC$(BTHH)_3$GCTAG	$C_{162}H_{186}N_{45}O_{78}P_{13}$	4411.85603	4411.84023
$(BTHH)_3$2	CTAGC$(BTHH)_3$GTCATC	$C_{160}H_{187}N_{38}O_{80}P_{13}$	4322.83216	4322.84937
$(BTHF)_2$1	GATGAC$(BTHF)_2$GCTAG	$C_{144}H_{159}F_{10}N_{45}O_{73}P_{12}$	4247.68045	4247.69724
$(BTHF)_2$2	CTAGC$(BTHF)_2$GTCATC	$C_{142}H_{160}F_{10}N_{38}O_{75}P_{12}$	4158.65659	4158.63710
$(BTHF)_3$1	GATGAC$(BTHF)_3$GCTAG	$C_{162}H_{171}F_{15}N_{45}O_{78}P_{13}$	4681.71470	4681.69489
$(BTHF)_3$2	CTAGC$(BTHF)_3$GTCATC	$C_{160}H_{172}F_{15}N_{38}O_{80}P_{13}$	4592.69084	4592.69091
$(BTFF)_2$1	GATGAC$(BTFF)_2$GCTAG	$C_{144}H_{151}F_{18}N_{45}O_{73}P_{12}$	4391.60507	4391.60071
$(BTFF)_2$2	CTAGC$(BTFF)_2$GTCATC	$C_{142}H_{152}F_{18}N_{38}O_{75}P_{12}$	4302.58121	4302.58433
$(BTHF)_3$_DNA8mer1	$(BTHF)_3$GCATCAGT	$C_{132}H_{135}F_{15}N_{30}O_{61}P_{10}$	3710.55206	3710.53845
DNA8mer2	ACTGATGC	$C_{78}H_{99}N_{30}O_{46}P_7$	2408.44930	2408.44124

Table S4.2 Thermodynamic data for DNA duplexes in phosphate buffer (100 mM NaCl, 10 mM sodium phosphate, pH 7.0).

Duplex	Oligo	Sequence	c_{total}[a] [µM]	T_m [°C]	ΔH^0 [kcal mol^{-1}]	ΔS^0 [cal mol^{-1} K^{-1}]	ΔG^{298}[b] [kcal mol^{-1}]
GTHH/GTFH	GTHH1 GTFH2	5'…C**(GTHH)**G…3' 3'…G**(GTFH)**C…5'	2.1 4.3 10.9 21.0 43.9	41.1 43.1 45.6 47.3 49.5	–74.4±0.3	–208±0.9	–12.5±0.4
GTHH/GTHF	GTHH1 GTHF2	5'…C**(GTHH)**G…3' 3'…G**(GTHF)**C…5'	2.2 4.4 11.4 22.6 45.6	40.3 42.2 45.0 46.8 48.9	–70.8±0.4	–197±1.2	–12.1±0.5
GTFH/GTHH	GTFH1 GTHH2	5'…C**(GTFH)**G…3' 3'…G**(GTHH)**C…5'	2.0 4.3 10.7 20.5 42.4	41.7 43.6 46.2 47.8 49.8	-75.1±0.5	-210±1.4	-12.6±0.6
GTFH/GTFH	GTFH1 GTFH2	5'…C**(GTFH)**G…3' 3'…G**(GTFH)**C…5'	2.0 4.2 10.7	44.5 46.4 48.9	–75.8±0.5	–210±1.3	–13.3±0.6

			20.8 43.6	50.6 52.9			
GTFH/GTHF	GTFH1 GTHF2	5'…C(GTFH)G…3' 3'…G(GTHF)C…5'	2.2 4.6 11.5 23.0 47.0	41.5 43.4 46.1 47.8 50.0	−72.7±0.7	−202±1.9	−12.4±0.9
GTFH/GTFF	GTFH1 GTFF2	5'…C(GTFH)G…3' 3'…G(GTFF)C…5'	2.1 4.3 10.8 21.2 43.7	44.2 46.2 49.2 50.8 53.2	−70.1±0.7	−192±1.9	−12.8±0.9
GTHF/GTHH	GTHF1 GTHH2	5'…C(GTHF)G…3' 3'…G(GTHH)C…5'	2.2 4.6 11.6 22.6 46.2	40.4 42.3 45.1 46.8 49.1	−70.6±0.4	−196±1.2	−12.1±0.5
GTHF/GTFH	GTHF1 GTFH2	5'…C(GTHF)G…3' 3'…G(GTFH)C…5'	2.2 4.6 11.3 21.9 45.5	41.7 43.8 46.4 48.2 50.5	−69.8±0.2	−193±0.7	−12.1±0.3
GTHF/GTHF	GTHF1 GTHF2	5'…C(GTHF)G…3' 3'…G(GTHF)C…5'	2.3 4.9 12.3 24.2 49.3	41.6 43.7 46.5 48.3 50.7	−68.3±0.5	−189±1.3	−12.1±0.6
GTHF/GTFF	GTHF1 GTFF2	5'…C(GTHF)G…3' 3'…G(GTFF)C…5'	2.4 4.7 12.0 24.1 48.2	42.4 44.6 47.7 49.6 52.0	−64.6±0.5	−176±1.5	−12.1±0.7
BTHH/BTFH	BTHH1 BTFH2	5'…C(BTHH)G…3' 3'…G(BTFH)C…5'	2.1 4.3 10.9 22.5 43.4	42.5 44.4 46.8 48.5 50.8	−76.2±0.9	−213±2.7	−12.7±1.2
BTHH/BTHF	BTHH1 BTHF2	5'…C(BTHH)G…3' 3'…G(BTHF)C…5'	2.4 4.8 11.9 23.0 47.6	45.9 47.4 50.0 51.4 53.4	−81.9±0.8	−228±2.3	−13.9±1.0
BTFH/BTHH	BTFH1 BTHH2	5'…C(BTFH)G…3' 3'…G(BTHH)C…5'	2.0 4.0 10.7 20.9 43.8	42.5 44.5 47.0 48.5 50.8	−77.9±0.8	−218±2.2	−13.0±1.0
BTFH/BTFH	BTFH1 BTFH2	5'…C(BTFH)G…3' 3'…G(BTFH)C…5'	2.0 3.9 10.5 20.3 43.2	45.3 46.6 49.2 50.8 53.2	−80.4±1.5	−224±4.2	−13.6±1.9
BTFH/BTHF	BTFH1 BTHF2	5'…C(BTFH)G…3' 3'…G(BTHF)C…5'	2.2 4.7 11.7 23.1 47.7	44.1 46.2 48.6 50.3 52.7	−75.0±0.7	−208±2.0	−13.0±0.9
BTFH/BTFF	BTFH1 BTFF2	5'…C(BTFH)G…3' 3'…G(BTFF)C…5'	2.1 4.4 10.9 20.9 43.6	46.8 48.8 51.4 53.1 56.0	−70.4±1.3	−192±3.7	−13.3±1.7
BTHF/BTHH	BTHF1 BTHH2	5'…C(BTHF)G…3' 3'…G(BTHH)C…5'	2.3 4.6 11.7 23.8	45.1 47.0 49.5 51.2	−75.8±1.1	−210±3.2	−13.3±1.5

			46.5	53.6			
BTHF/BTFH	BTHF1 BTFH2	5'…C(**BTHF**)G…3' 3'…G(**BTFH**)C…5'	2.3 4.7 11.7 24.6 46.9	43.5 45.4 47.9 49.6 52.1	–73.3±1.1	–203±3.4	–12.7±1.6
BTHF/BTHF	BTHF1 BTHF2	5'…C(**BTHF**)G…3' 3'…G(**BTHF**)C…5'	2.2 4.8 12.2 23.9 49.3	49.9 51.8 54.4 55.9 58.4	–78.8±1.2	–215±3.4	–14.5±1.6
BTHF/BTFF	BTHF1 BTFF2	5'…C(**BTHF**)G…3' 3'…G(**BTFF**)C…5'	2.3 4.8 12.3 24.3 49.9	47.8 49.8 52.3 54.1 56.5	–74.6±0.8	–204±2.4	–13.8±1.1
$(BTHH/BTHH)_2$	$(BTHH)_2$1 $(BTHH)_2$2	5'…C(BT**HH**)$_2$G…3' 3'…G(BT**HH**)$_2$C…5'	2.1 4.3 11.0 21.1 42.8	50.5 52.0 54.5 55.8 57.9	–86.8±1.1	–240±3.0	–15.4±1.4
$(BTHH/BTHH)_3$	$(BTHH)_3$1 $(BTHH)_3$2	5'…C(BT**HH**)$_3$G…3' 3'…G(BT**HH**)$_3$C…5'	2.0 4.2 10.8 20.5 43.0	55.1 56.5 58.9 60.1 62.3	–93.1±1.6	–255±4.3	–17.0±2.0
$(BTHF/BTHF)_2$	$(BTHF)_2$1 $(BTHF)_2$2	5'…C(BT**HF**)$_2$G…3' 3'…G(BT**HF**)$_2$C…5'	2.5 5.3 11.4 26.5 55.0	54.8 56.6 59.5 60.9 63.8	–77.1±2.3	–207±6.3	–15.4±3.0
$(BTHF/BTHF)_3$	$(BTHF)_3$1 $(BTHF)_3$2	5'…C(BT**HF**)$_3$G…3' 3'…G(BT**HF**)$_3$C…5'	2.5 5.5 15.1 29.6 62.3	60.0 61.8 65.0 66.4 69.9	–75.8±2.4	–200±6.4	–16.3±3.0
$(BTFF/BTHH)_2$	$(BTFF)_2$1 $(BTHH)_2$2	5'…C(BT**FF**)$_2$G…3' 3'…G(BT**HH**)$_2$C…5'	2.0 4.1 10.5 20.4 42.2	52.2 53.9 56.6 58.4 59.9	–83.4±1.2	–227±3.4	–15.6±1.6
$(BTFF/BTFF)_2$	$(BTFF)_2$1 $(BTFF)_2$2	5'…C(BT**FF**)$_2$G…3' 3'…G(BT**FF**)$_2$C…5'	1.8 4.2 11.1 21.8 47.6	50.6 53.5 56.0 57.5 60.3	–75.1±1.4	–203±3.8	–14.6±1.8
$(BTHF)_3$_Dimer	$(BTHF)_3$1	5'…C(BT**HF**)$_3$G…3'	1.2 2.6 7.3 13.9 30.3	36.3 38.7 42.7 43.9 47.1	–59.4±1.2	–165±3.4	–10.2±1.6

[a] Total concentration of DNA strands. Values were calculated with the absorption at 260 nm of the melted duplex [b] Calculated for T = 25 °C.

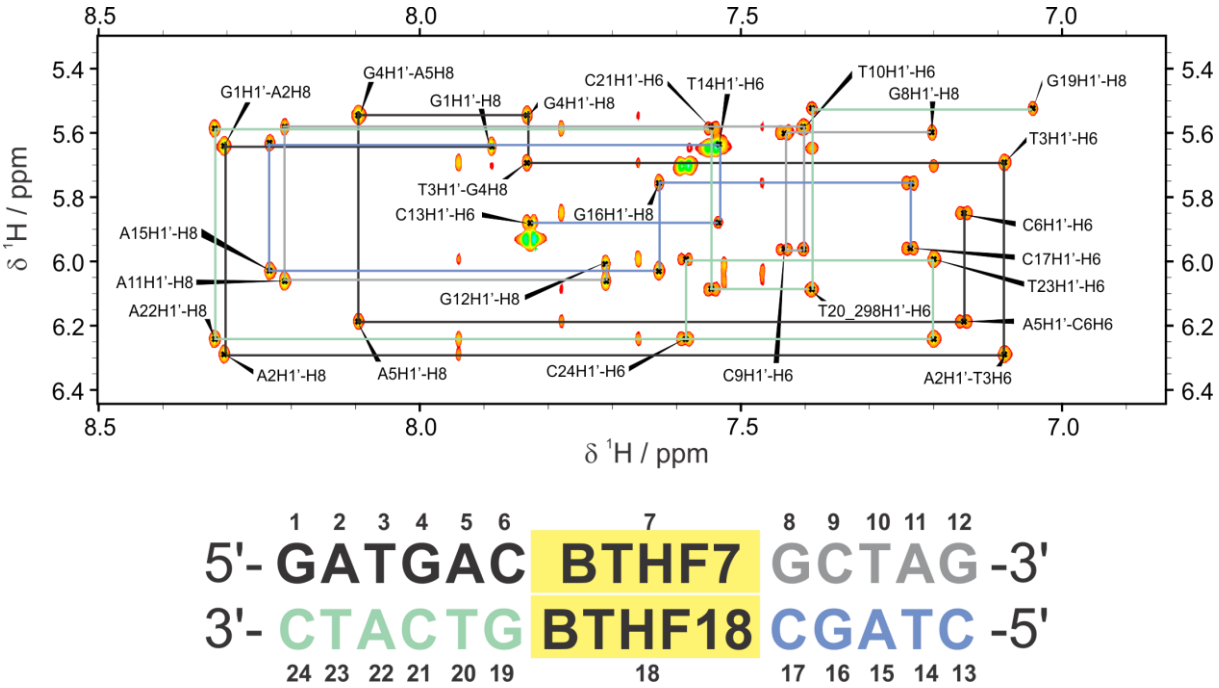


Figure S4.1 Aromatic-anomeric region of the 2D ^{1}H,^{1}H NOESY NMR spectrum of the duplex BTHF/BTHF (sequence and numbering reported on the bottom). Assignment on the NOESY spectra is reported for clarity only for the sequential step 5'-H1'(n)-H6/H8(n+1)-3'. Sequential walk is indicated and color coded as follows: G1-C6 black, G8-G12 gray, C13-C17 blue, G19-C24 aquamarine. Spectrum recorded on 1.0 mM DNA duplex in NMR buffer, 100%D_2O, 600 MHz, 25 °C.

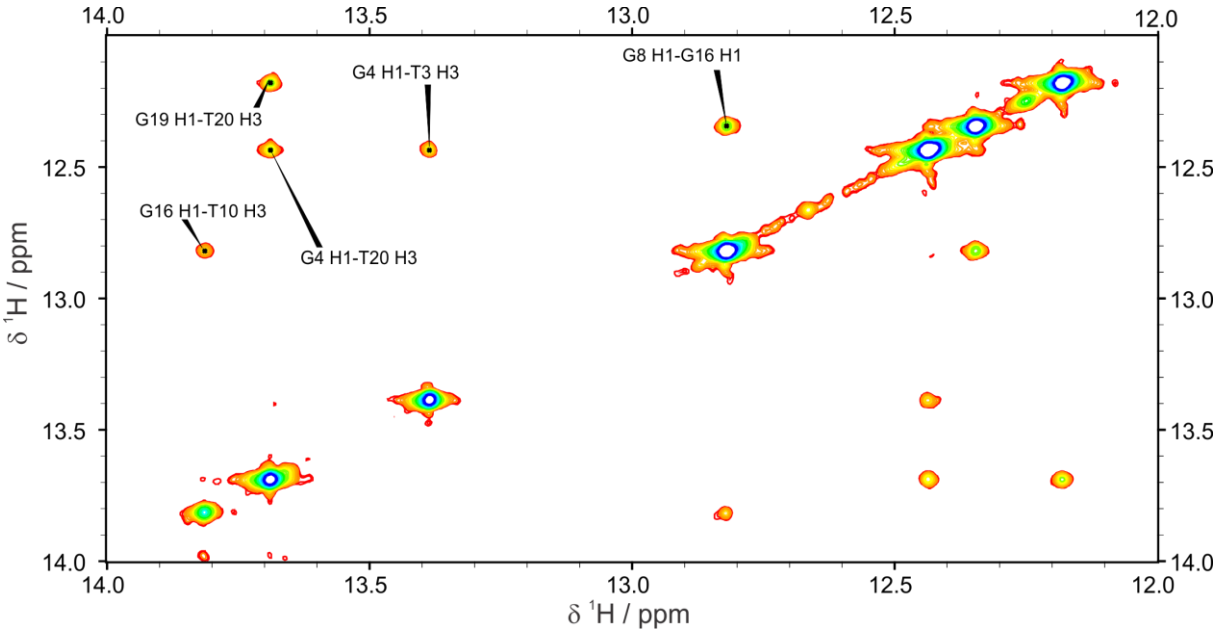


Figure S4.2 Imino region of the 2D ^{1}H,^{1}H NOESY NMR spectrum of the duplex BTHF/BTHF with assignment of the imino-imino connectivities. Spectrum recorded on 1.0 mM DNA duplex in NMR buffer, 90% H_2O / 10%D_2O, 600 MHz, 25 °C.

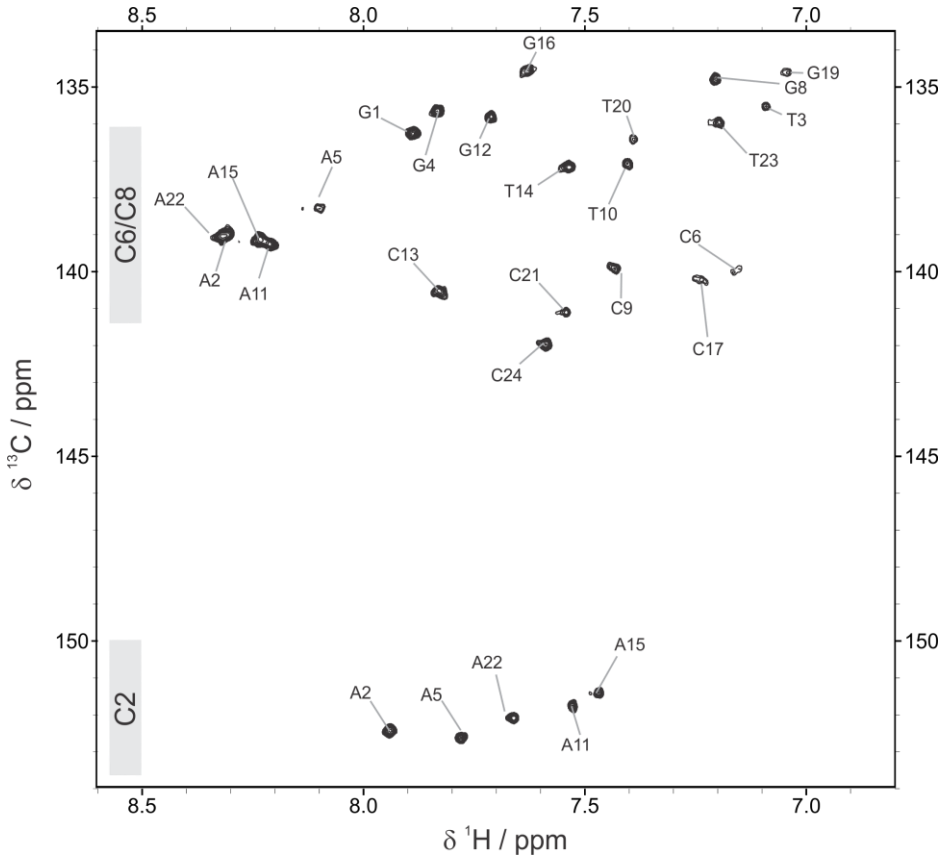


Figure S4.3 Aromatic region of the 2D $^{1}H,^{13}C$ HSQC NMR spectrum of the duplex BTHF/BTHF (sequence and numbering reported on the bottom). Spectrum recorded on 1.0 mM DNA duplex in NMR buffer, 100%D_2O, 600 MHz.

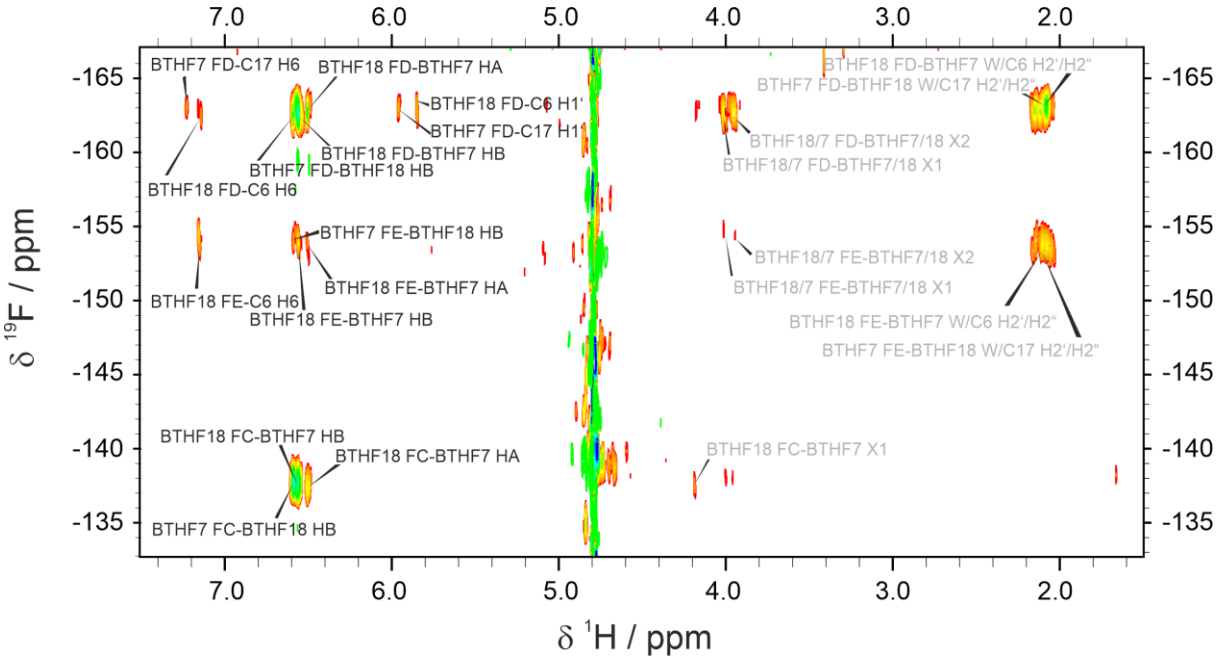


Figure S4.4 Aromatic/sugar region of the 2D $^{1}H,^{19}F$ HOESY spectrum of BTHF/BTHF duplex with assignment. Ambiguous assignment due to overlap in the ^{1}H dimension is reported in gray. Spectrum recorded on 1.0 mM DNA duplex in NMR buffer, 90%H_2O / 10%D_2O, 600 MHz, 25 °C.

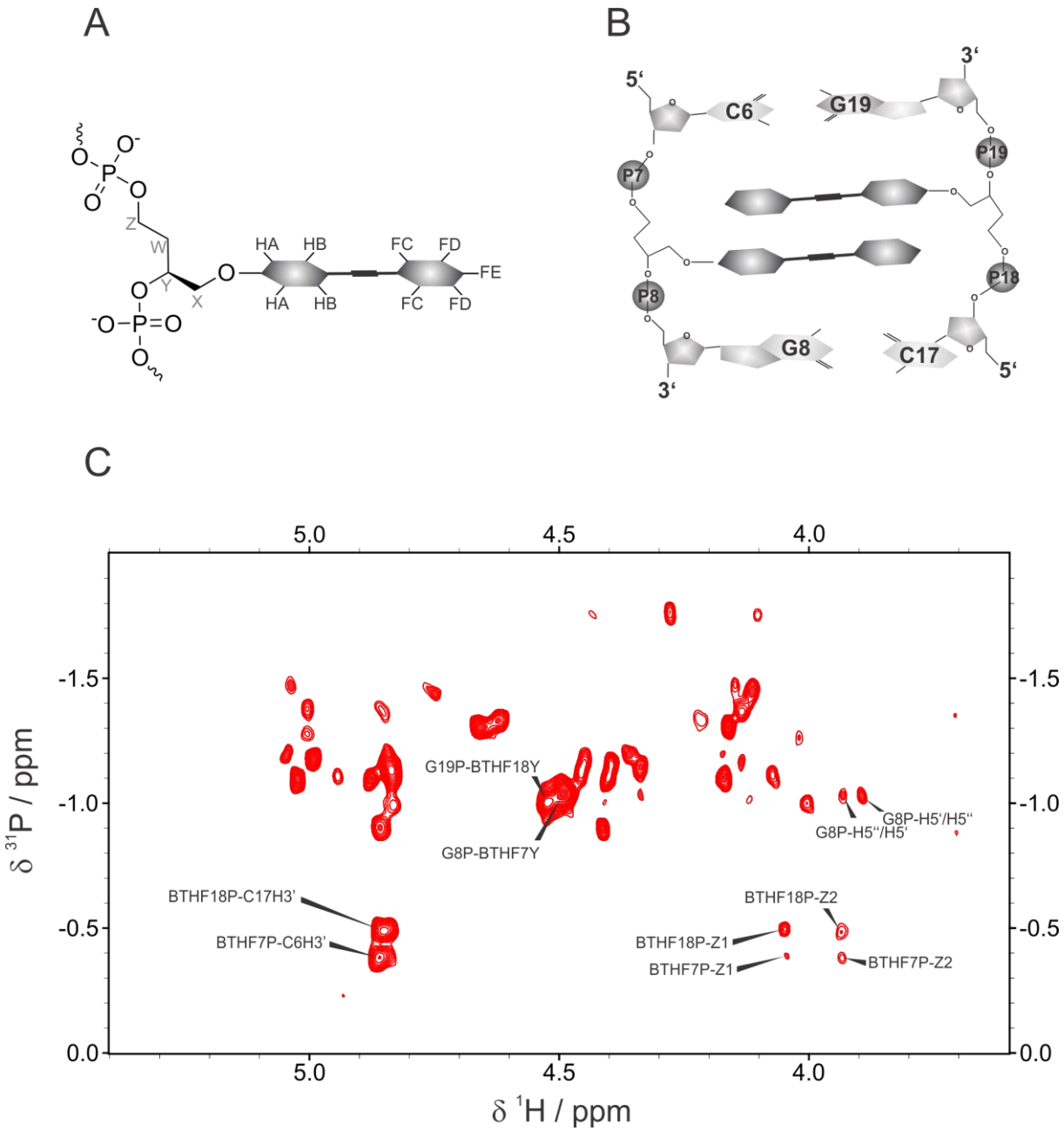


Figure S4.5 Letter code and numbering used for labeling acyclic backbone (A) and phosphate groups (B). $^{1}H,^{31}P$ HSQC of BTHF/BTHF with H3'(n-1)-P(n) and H4'(n)-P(n) assignment indicated only for phosphate residues at the modification site. Spectra recorded on 1.0 mM DNA duplex in NMR buffer, 100%D_2O, 600 MHz, 25 °C.

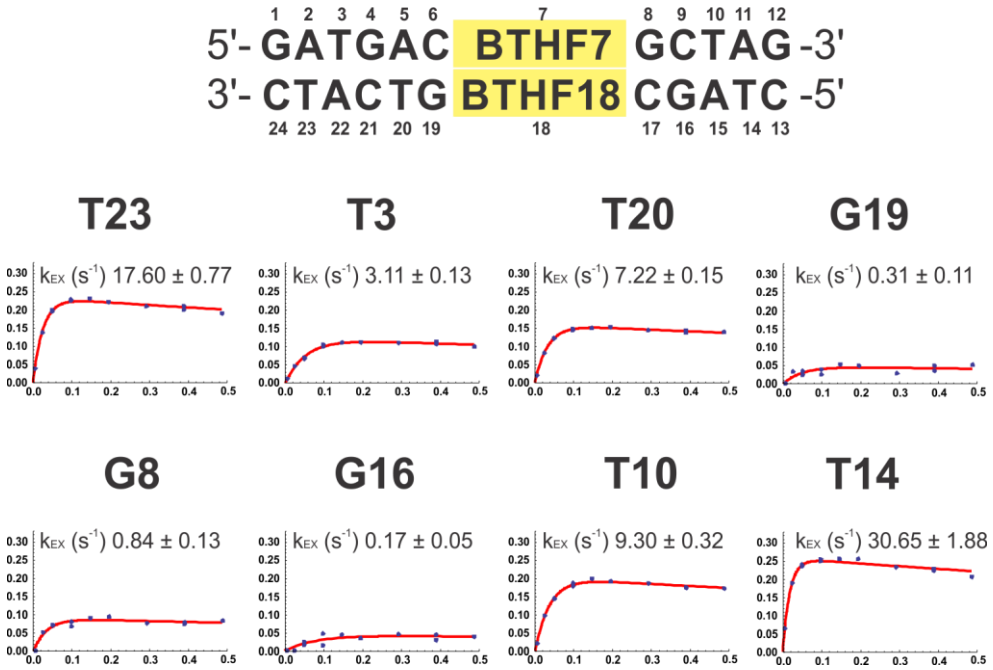


Figure S4.6 Relative peak intensities of the imino proton NMR signals as a function of the mixing time (in seconds) after water inversion in the CLEANEX-PM experiments. The sequence is reported on the top. G4 imino proton water k_{EX} could not be determined due to too slow exchange. G1 and G12 imino proton k_{EX} could not be determined due to the peak broadening beyond detection at 25 °C. The best fit to Eq. 3.12 is depicted as a solid red line. Results of the fitting are reported on top of each graph (error from the fitting).

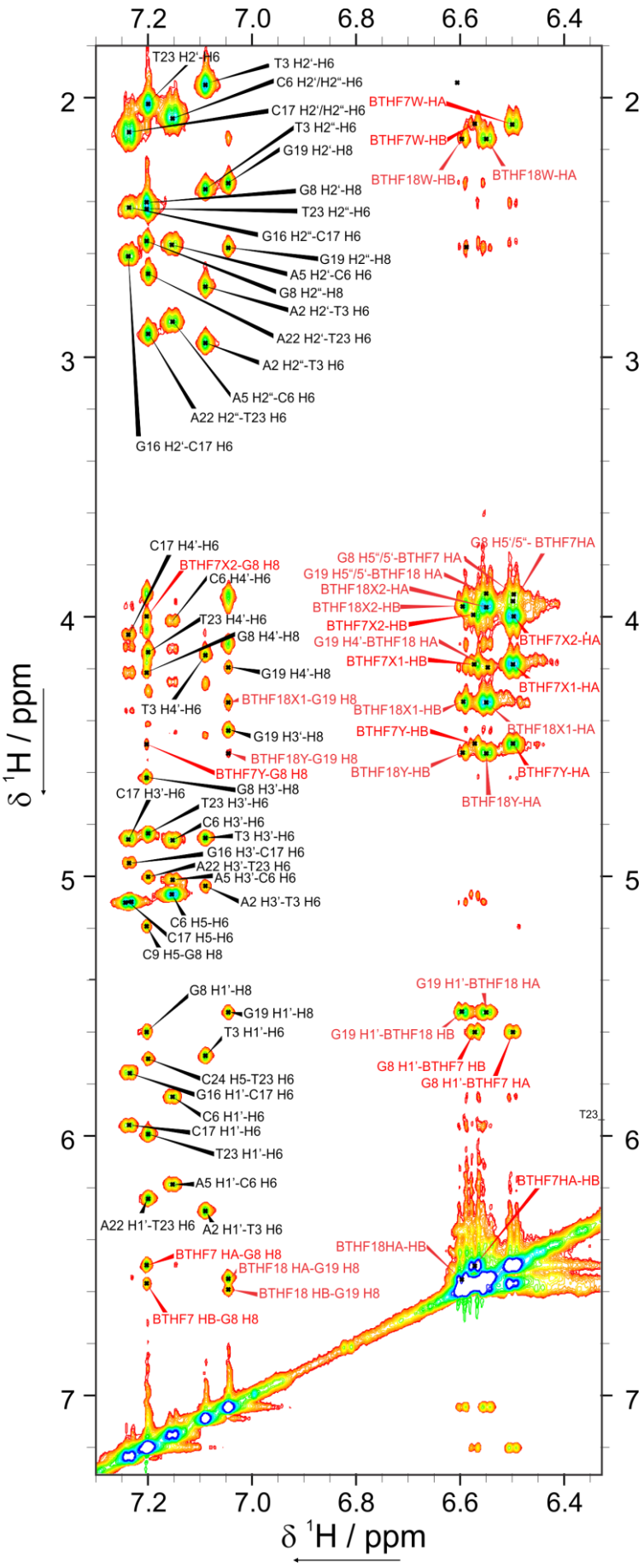


Figure S4.7 Aromatic region of the NOESY spectrum of BTHF/BTHF 1 mM in NMR buffer, 100% D_2O (600 MHz, 25 °C) with assignment. Cross peaks with the tolane aromatic protons are highlighted in red.

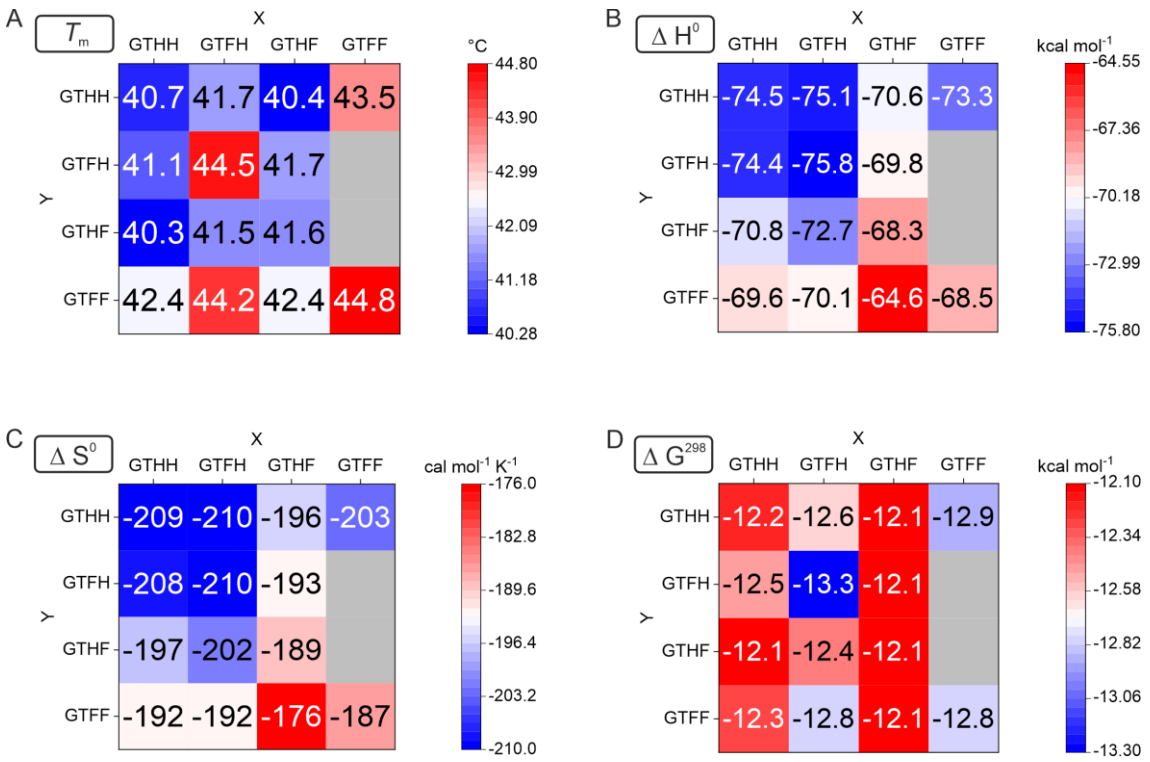


Figure S4.8 Summary of (A) melting points at 1 µM DNA duplex concentration, (B) enthalpy, (C) entropy and (D) free energy at 25 °C containing a tolane pair connected on an artificial GNA backbone. For detail information see Table S4.2. Values for GTHH/GTHH, GTHH/GTFF, GTFF/GTHH, GTFF/GTFF are taken from literature.[443]

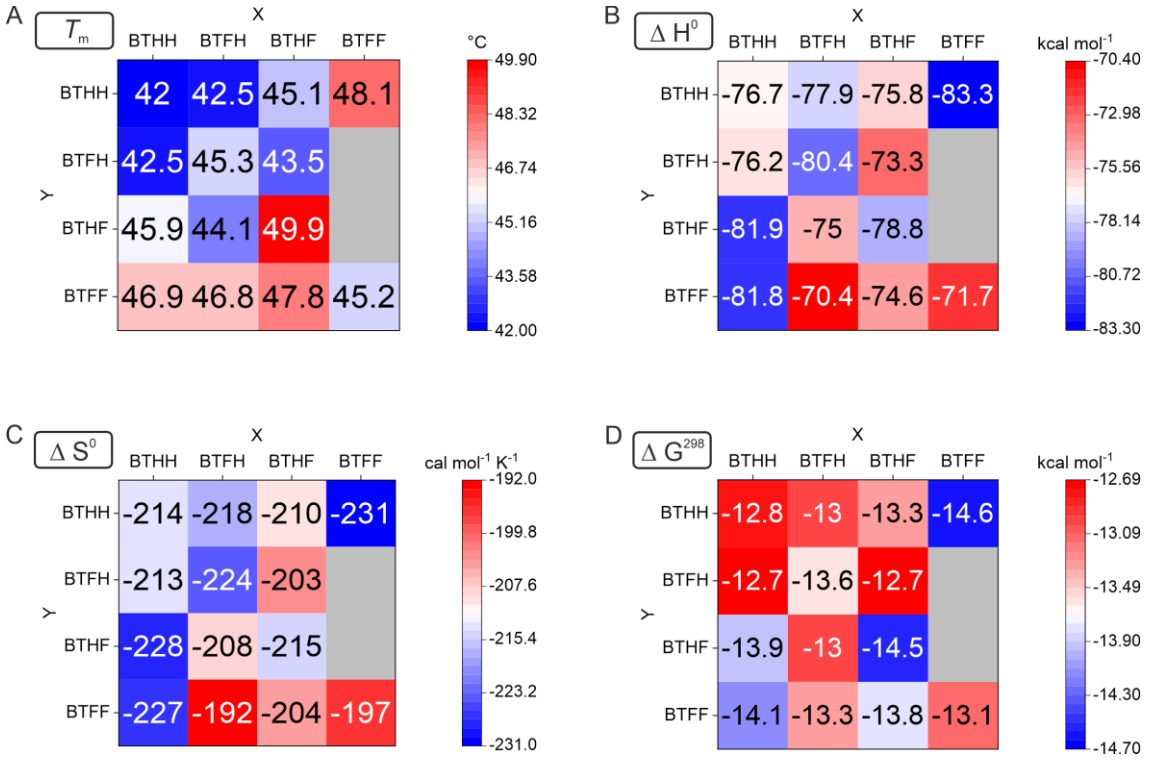


Figure S4.9 Summary of (A) melting points at 1 µM DNA duplex concentration, (B) enthalpy, (C) entropy and (D) free energy at 25 °C containing a tolane pair connected on an artificial BuNA backbone. For detail information see Table S4.2. Values for BTHH/BTHH, BTHH/BTFF, BTFF/BTHH, BTFF/BTFF are taken from literature.[443]

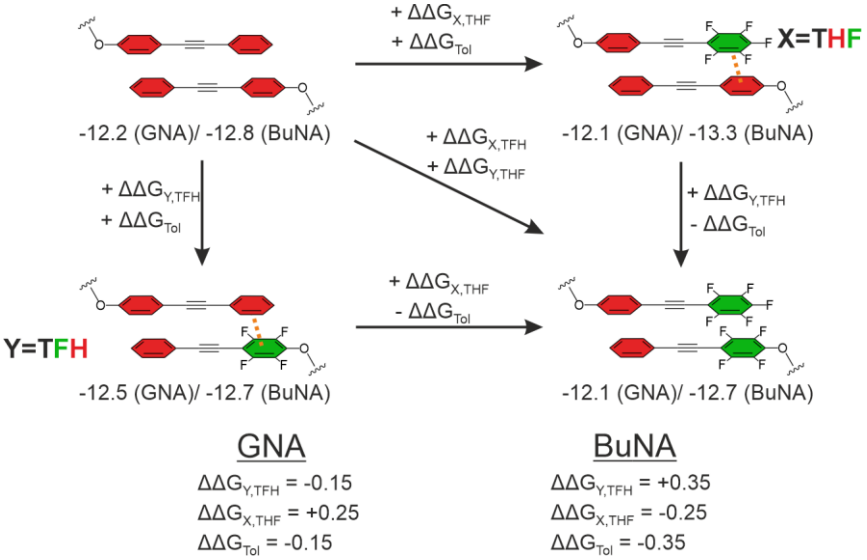


Figure S4.10 Double mutant cycle for the replacement THF at position X and TFH at position Y. Values are given in kcal mol^{-1}.

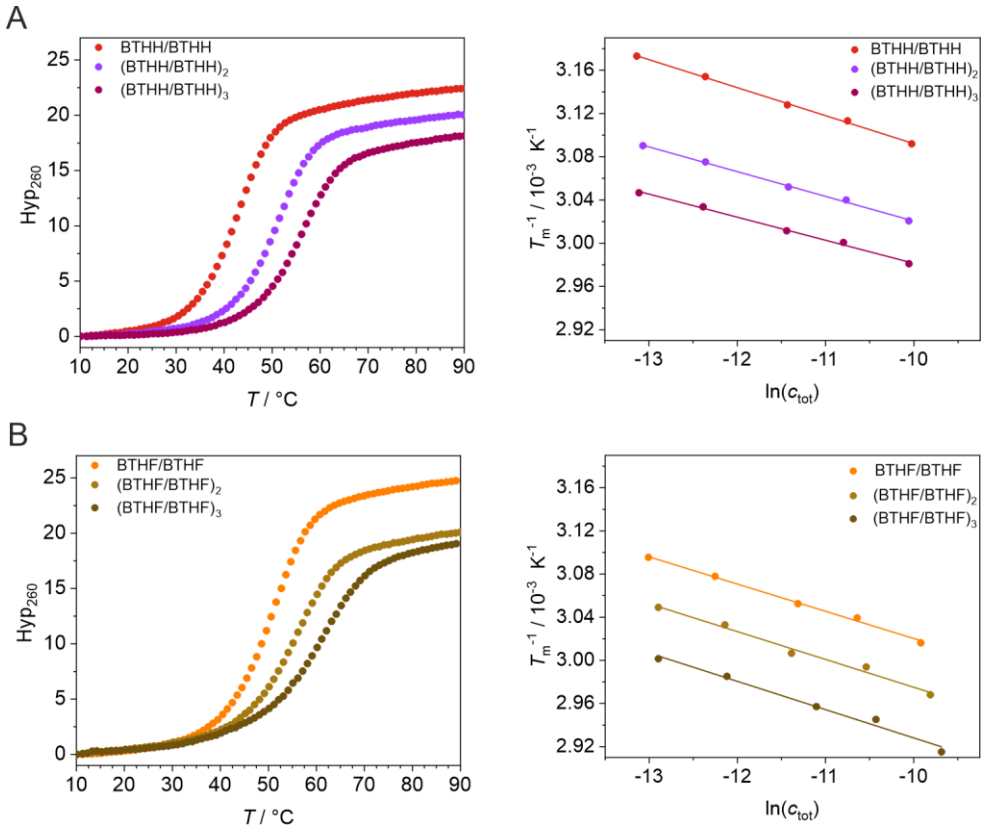


Figure S4.11 Melting curves and the corresponding van't Hoff analysis of the multiple homopairs containing DNA-duplexes for the (A) BTHH and (B) BTHF (1 µM DNA duplex in 10 mM phosphate buffer, 100 mM NaCl, pH 7.0).

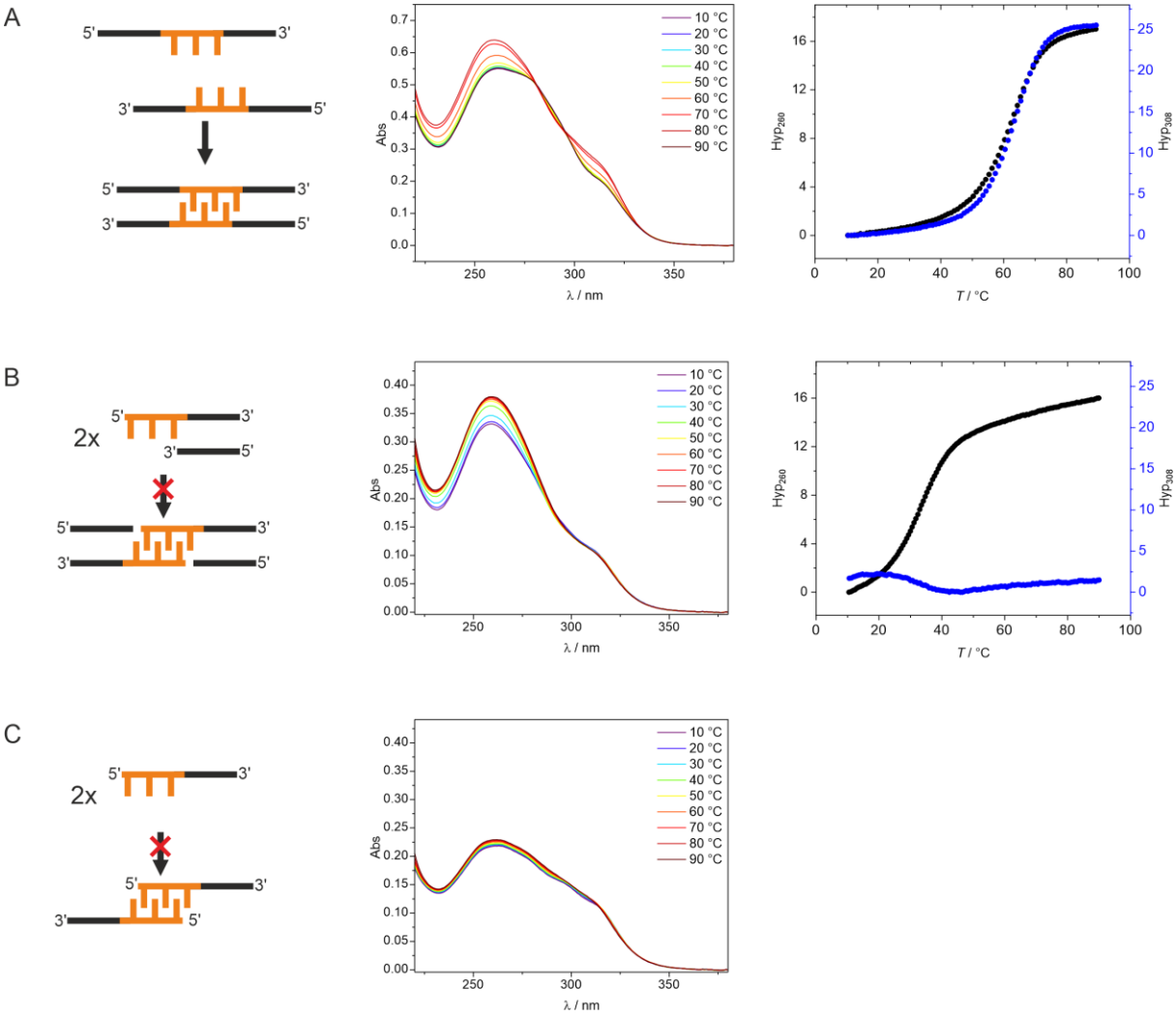


Figure S4.12 Thermal UV/Vis analysis of DNA strands containing the BTHF triplet motif (orange) for (A) the DNA duplex (BTHF/BTHF)$_3$, (B) 5'-(BTHF)$_3$-8merDNA with complementary DNA strand and (C) only 5'-(BTHF)$_3$-8merDNA (2 µM DNA duplex in 10 mM phosphate buffer, 100 mM NaCl, pH 7.0).

4.4.2 Experimental procedures

4.4.2.1 General material and methods

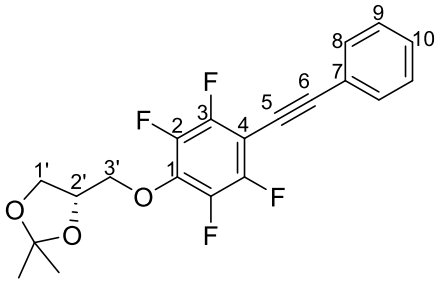
See 3.4.2.1. General material and methods for further information. Synthesis of 1-pentafluorophenyl-2-phenylacetylene (**S4.6**) was adapted from literature.[453] Column chromatography (hexane) was used instead of sublimation purification.

4.4.2.2 NMR-Spectroscopy and mass spectrometry

See 3.4.2.2 NMR-Spectroscopy and mass spectrometry for further information.

4.4.2.3 Synthetic procedures

Compound **S4.2**



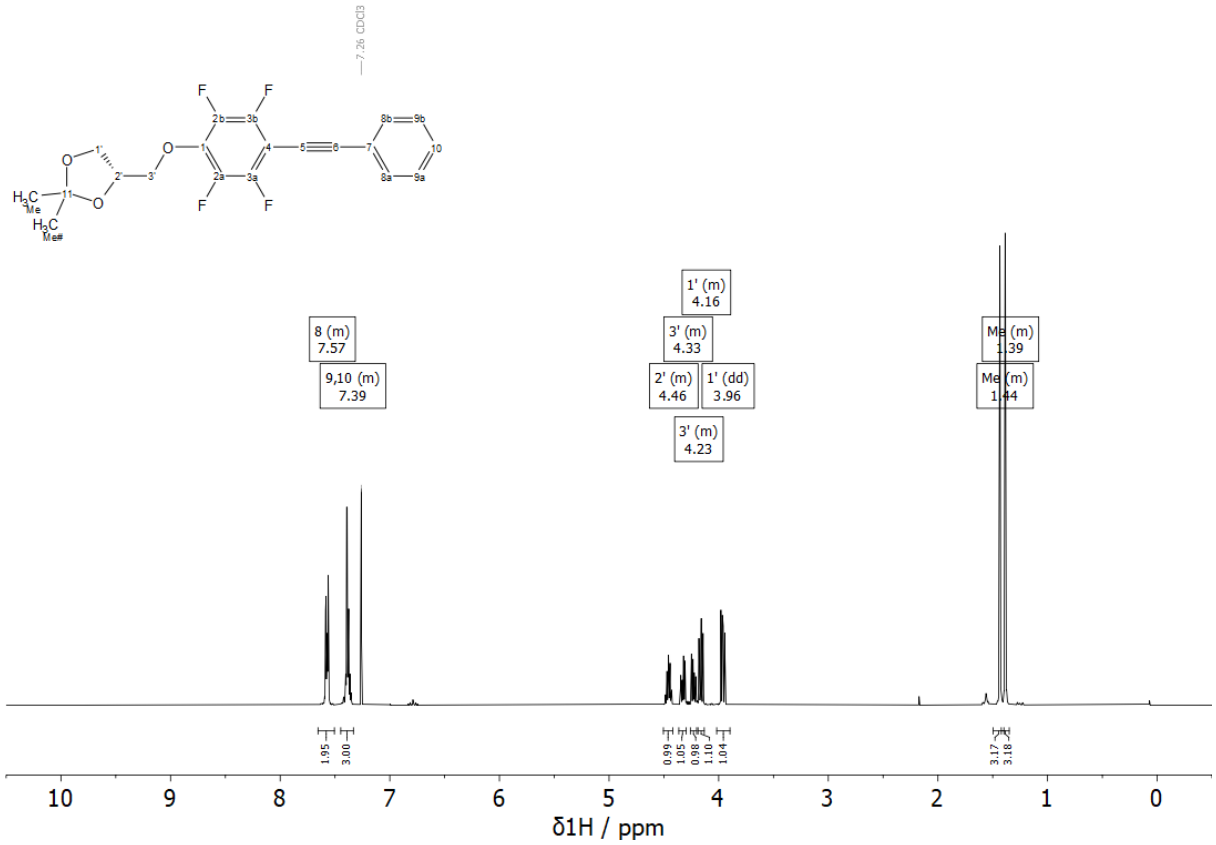
Under nitrogen atmosphere, CuI (36.5 mg, 192 µmol, 5.4 mol%), $Pd(PPh_3)_4$ (122 mg, 106 µmol, 3.0 mol%) and (*S*)-2,2-dimethyl-4-((2,3,5,6-tetrafluoro-4-iodophenoxy)methyl)-1,3-dioxolane (**S4.1**) (1.43 g, 3.52 mmol, 1.00 eq.) were dissolved in argon-purged toluene (14 mL) and diisopropylamine (4.2 mL). Phenylacetylene (387 µL, 360 mg, 3.52 mmol, 1.00 eq.) was added and stirring was continued for 18 h at 80°C. The precipitate was filtered off and washed with Et_2O. The filtrate was washed with a saturated solution of ammonium chloride. The aqueous phase was extracted with Et_2O. The combined organic phases were washed with brine, dried over Na_2SO_4, evaporated and the residue was purified by column chromatography (DCM/hexane 1:1) to afford compound **S4.2** as a brown solid (1.28 g, 3.37 mmol, 96%).

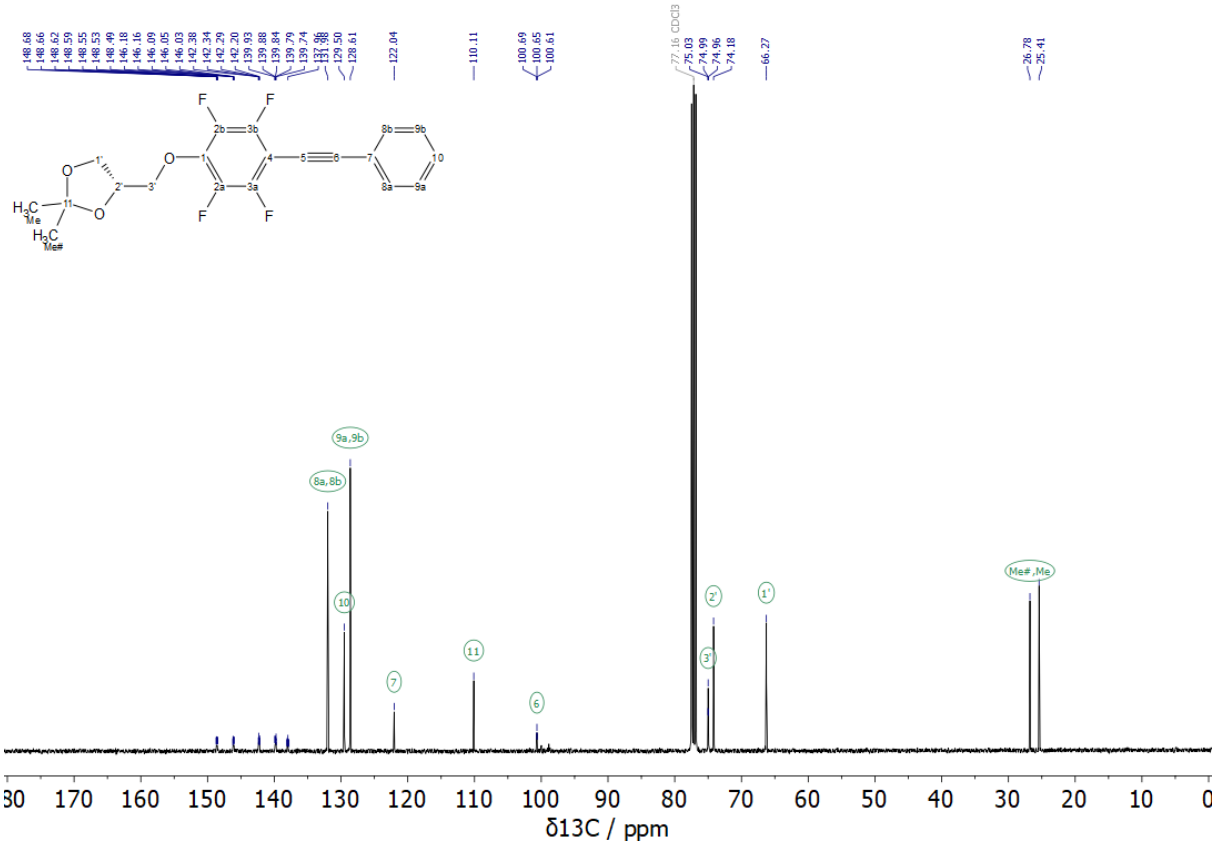
^{1}H NMR (400 MHz, $CDCl_3$): δ (ppm) = 7.65 – 7.51 (m, 2H, 8-H), 7.45 – 7.33 (m, 3H, 9-H, 10-H), 4.50 – 4.42 (m, 1H, 2',H), 4.36 – 4.30 (m, 1H, 3'-H), 4.26 – 4.20 (m, 1H, 3'-H), 4.19 – 4.13 (m, 1H, 1'-H), 3.96 (dd, *J* = 8.6, 5.5 Hz, 1H, 1'-H), 1.50 – 1.40 (m, 3H, CH_3), 1.42 – 1.35 (m, 3H, CH_3);

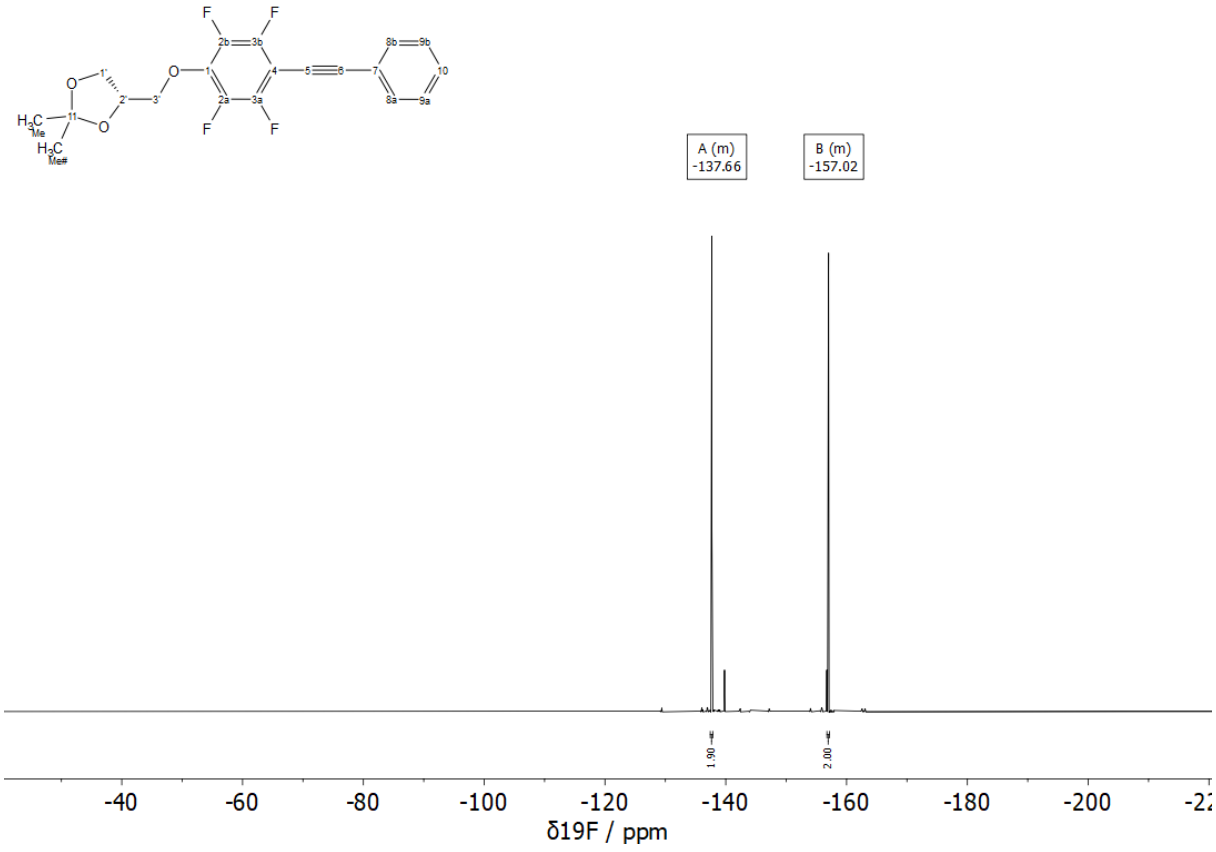
^{13}C{^{1}H} NMR (100 MHz, $CDCl_3$): δ (ppm) = 148.73 – 148.39 (m), 146.27 – 145.97 (m), 142.59 – 142.04 (m), 140.02 – 139.49 (m), 138.24 – 137.66 (m), 131.98 (8-C), 129.50 (10-C); 128.61 (9-C), 122.04 (7-C), 110.11 ($C(CH_3)_2$), 100.65 (t, *J* = 3.7 Hz, 6-C), 74.99 (t, *J* = 3.3 Hz, 3'-C), 74.18 (2'-C), 66.27 (1'-C), 26.78 (CH_3), 25.41 (CH_3);

^{19}F{^{1}H} NMR (376 MHz, $CDCl_3$): δ (ppm) = -137.43 – -137.85 (m, 2F), -156.75 – -157.16 (m, 2F);

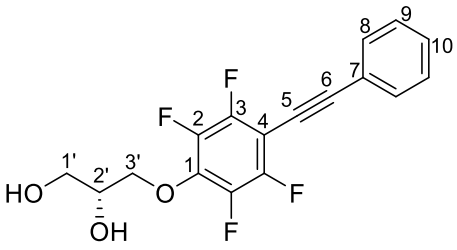
HR-MS (ESI+): *m/z* calc. ($C_{20}H_{16}F_4O_3Na$, $[M+Na]^+$): 403.09278, found: 403.09388.


7.26 CDCl3
8 (m) 7.57
9,10 (m) 7.39
1' (m) 4.16
3' (m) 4.33
2' (m) 4.46
1' (dd) 3.96
3' (m) 4.23
Me (m) 1.39
Me (m) 1.44
δ1H / ppm


77.16 CDCl3
9a,9b
8a,8b
10
7
11
6
2'
1'
3'
Me#,Me
δ13C / ppm



Compound **S4.3**



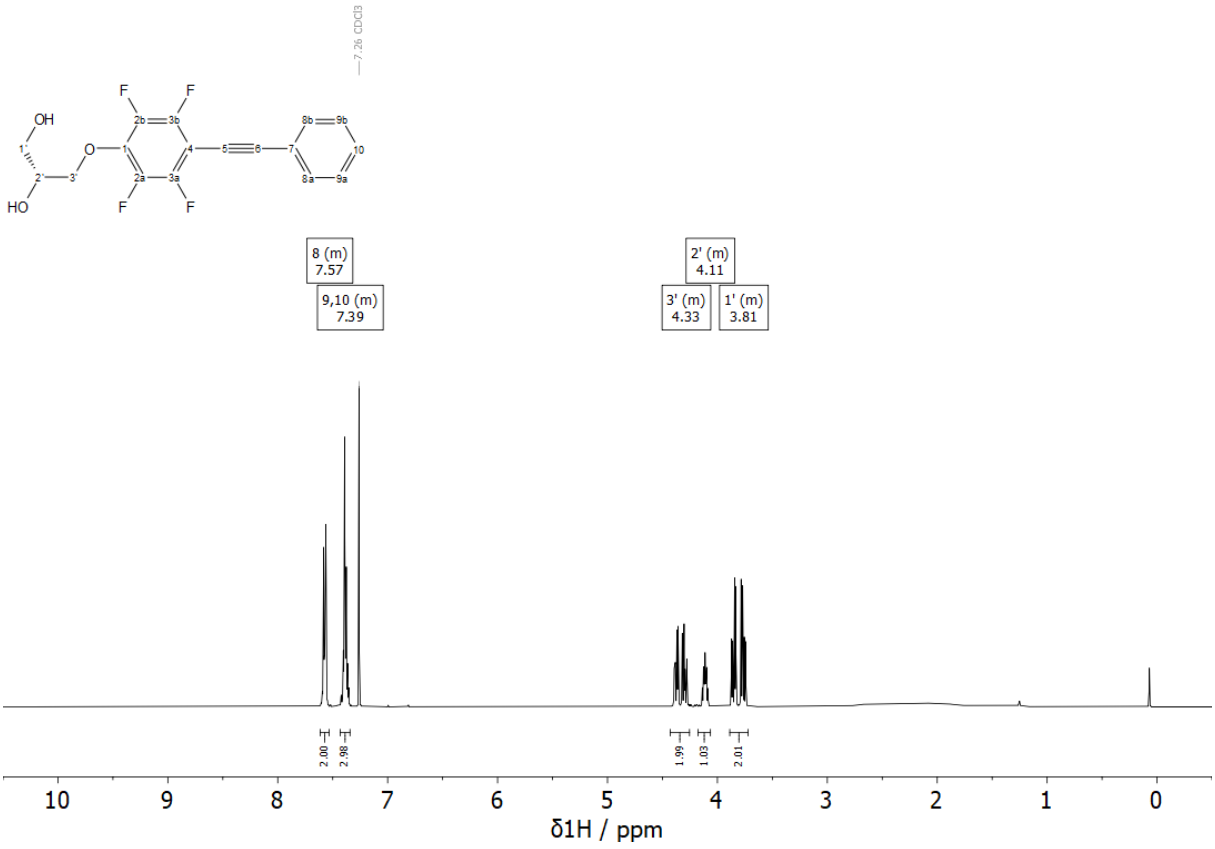
Compound **S4.2** (603 mg, 1.59 mmol, 1.00 eq.) was dissolved in THF (10 mL) and 2 M hydrochloric acid (10 mL) was added. The reaction mixture was stirred for 4 d at room temperature. The mixture was diluted with water and extracted with EtOAc. The combined organic phases were washed with brine and dried over Na_2SO_4. The solvent was removed under reduced pressure and the residue was purified by column chromatography (hexane/EtOAc 3:2 → 1:1) to afford compound **S4.3** as a colorless solid (414 mg, 1.22 mmol, 77%).

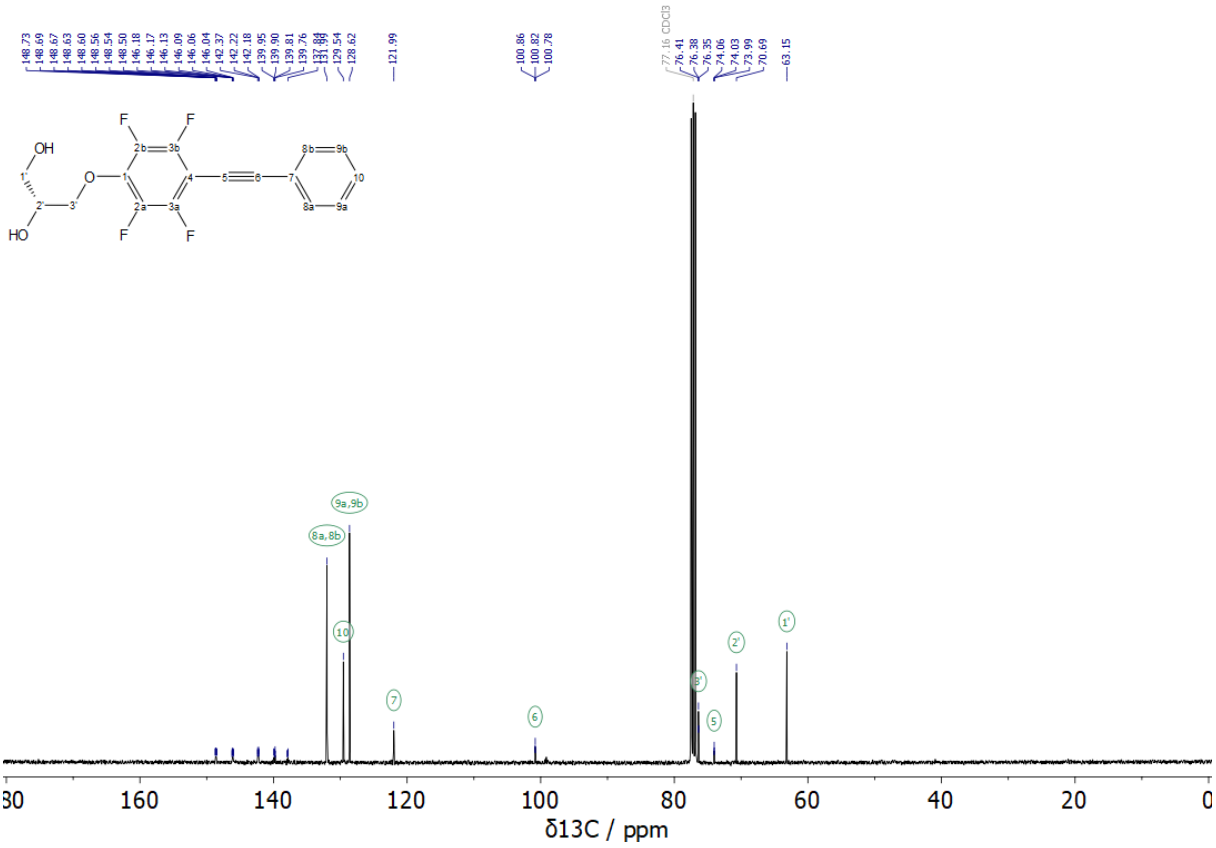
^{1}H NMR (400 MHz, $CDCl_3$): δ (ppm) = 7.61 – 7.53 (m, 2H, 8-H), 7.43 – 7.34 (m, 3H, 9-H, 10-H), 4.43 – 4.25 (m, 2H, 3'-H), 4.18 – 4.06 (m, 1H, 2'-H), 3.89 – 3.72 (m, 2H, 1'-H);

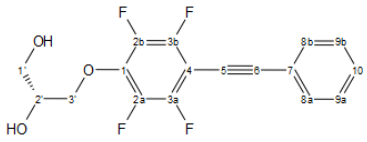
$^{13}C\{^{1}H\}$ NMR (100 MHz, $CDCl_3$): δ (ppm) = 148.77 – 148.42 (m), 146.41 – 145.83 (m), 142.51 – 142.02 (m), 139.98 – 139.54 (m), 138.03 – 137.74 (m), 131.99 (8-C), 129.54 (10-C), 128.62 (9-C), 121.99 (7-C), 100.82 (t, *J* = 3.8 Hz, 6-C), 76.38 (t, *J* = 3.2 Hz, 3‘-C), 74.03 (t, *J* = 3.9 Hz, 5-C), 70.69 (2’-C), 63.15 (1’-C);

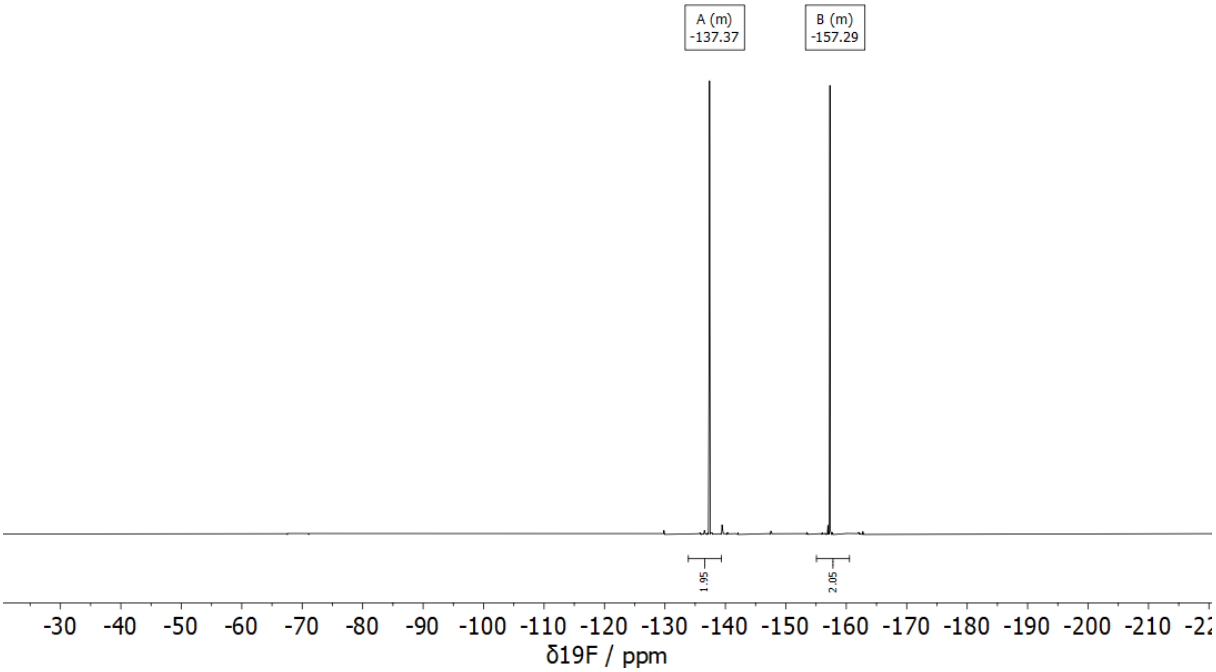
$^{19}F\{^{1}H\}$ NMR (376 MHz, $CDCl_3$): δ (ppm) = -133.87 – -139.33 (m, 2F), -155.07 – -160.53 (m, 2F);

HR-MS (ESI+): *m/z* calc. ($C_{17}H_{12}F_4O_3Na$, $[M+Na]^+$): 363.06148, found: 363.06226.

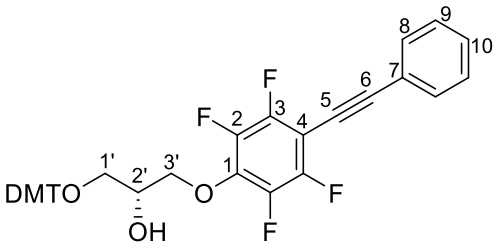



148.73
148.69
148.67
148.63
148.60
148.56
148.54
148.50
146.18
146.17
146.13
146.09
146.06
146.04
142.37
142.22
142.18
139.95
139.90
139.81
139.76
131.99
129.54
128.62
121.99
100.86
100.82
100.78
77.16 CDCl3
76.41
76.38
76.35
74.06
74.03
73.99
70.69
63.15
OH
HO
F
O
9a,9b
8a,8b
10
7
6
5
2'
1'
δ13C / ppm


OH
HO
F
O


A (m)
-137.37
B (m)
-157.29
1.95
2.05
δ19F / ppm

Compound **S3.4**



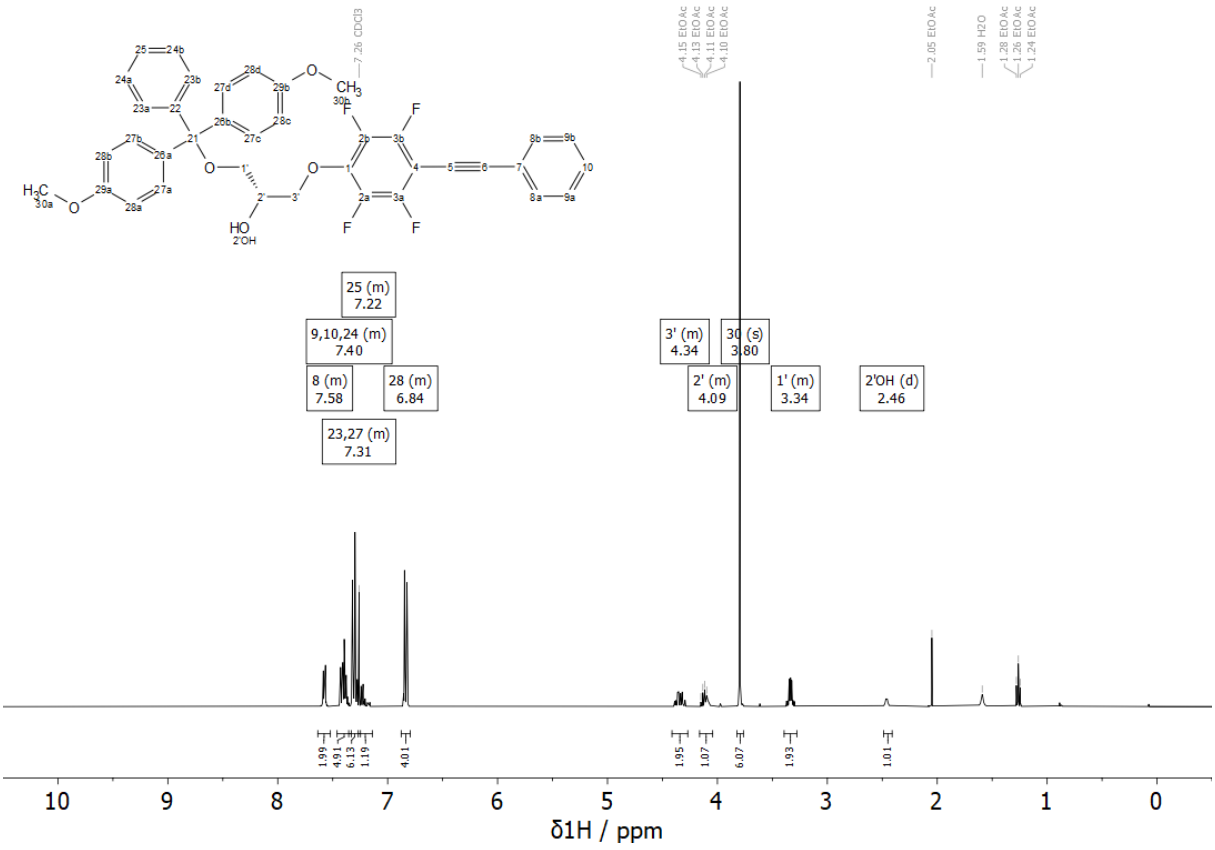
Under nitrogen atmosphere, compound **S4.3** (350 mg, 1.03 mmol, 1.00 eq.) was dissolved in anhydrous pyridine (11 mL). 4,4'-dimethoxytrityl chloride (453 mg, 1.34 mmol, 1.30 eq.) was added in small portions over a time period of 30 min and the reaction mixture was stirred at ambient temperature for 4 h. The mixture was diluted with DCM, washed three times with sodium hydrogen carbonate solution (5%) and dried over Na_2SO_4.The solvent was removed under reduced pressure and the residue was purified by column chromatography (hexane/EtOAc 4:1+1%) to afford compound **S4.4** as a colorless foam (600 mg, 934 µmol, 91%).

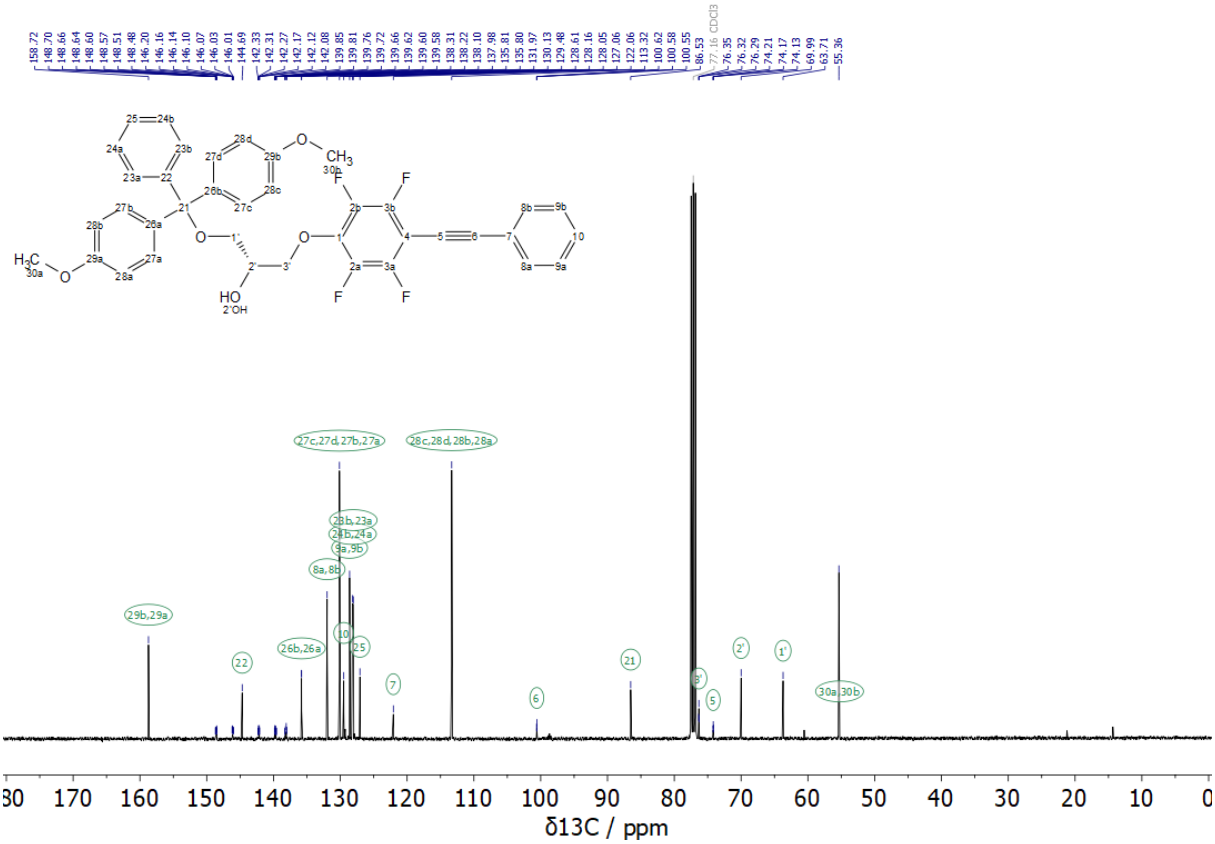
^{1}H NMR (400 MHz, $CDCl_3$): δ (ppm) = 7.64 – 7.52 (m, 2H, 8-H), 7.46 – 7.33 (m, 5H, 9-H, 10-H, DMT-H), 7.35 – 7.25 (m, 6H, DMT-H), 7.27 – 7.14 (m, 1H, DMT-H), 6.88 – 6.79 (m, 4H, DMT-H), 4.41 – 4.27 (m, 2H, 3'-H), 4.16 – 4.04 (m, 1H, 2'-H), 3.80 (s, 6H, DMT-H), 3.39 – 3.28 (m, 2H, 1'-H), 2.46 (d, *J* = 5.2 Hz, 1H, 2'-OH);

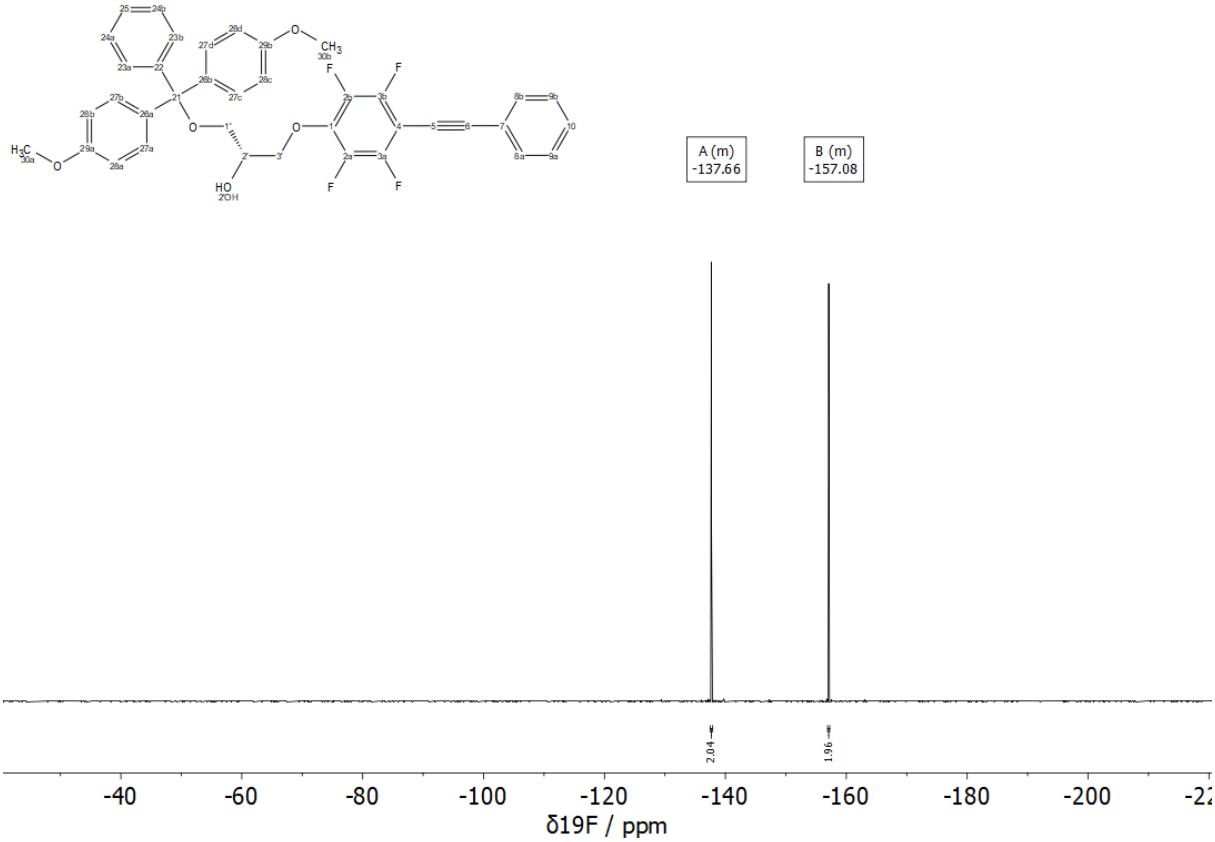
^{13}C{^{1}H} NMR (100 MHz, $CDCl_3$): δ (ppm) = 158.72 (29-C), 148.76 – 148.41 (m), 146.33 – 145.84 (m), 144.69 (DMT-C), 142.35 – 141.98 (m), 139.97 – 139.43 (m), 138.50 – 137.79 (m, 1-C), 135.81 (DMT-C), 135.80 (DMT-C), 131.97 (8-C), 130.13 (DMT-C), 129.48 (10-C), 128.61 (9-C), 128.16 (DMT-C), 128.05 (DMT-C), 127.06 (DMT-C), 122.06 (7-C), 113.32 (DMT-C), 100.58 (t, *J* = 3.5 Hz, 6-C), 86.53 (DMT-C), 76.32 (t, *J* = 3.2 Hz, 3'-C), 74.17 (t, *J* = 4.0 Hz, 5-C), 69.99 (2'-C), 63.71 (1'-C), 55.36 (DMT-C);

^{19}F{^{1}H} NMR (376 MHz, $CDCl_3$): δ (ppm) = -137.47 – -137.90 (m, 2F), -156.92 – -157.26 (m, 2F);

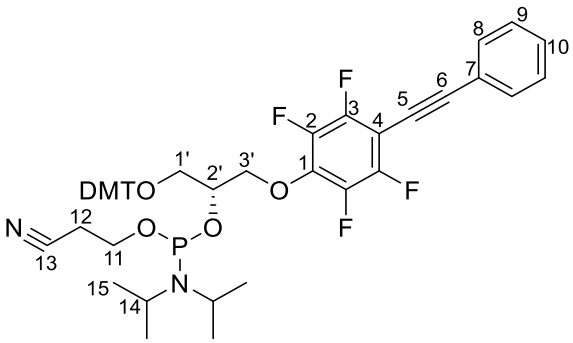
HR-MS (ESI+): *m/z* calc. ($C_{38}H_{30}F_4O_5Na$, $[M+Na]^+$): 665.19216, found: 665.19391.


25 (m)
7.22
9,10,24 (m)
7.40
8 (m)
7.58
28 (m)
6.84
23,27 (m)
7.31
3' (m)
4.34
30 (s)
3.80
2' (m)
4.09
1' (m)
3.34
2'OH (d)
2.46
δ1H / ppm


27c,27d,27b,27a
28c,28d,28b,28a
23b,23a
24b,24a
9a,9b
8a,8b
29b,29a
26b,26a
δ13C / ppm



Compound **S4.5** = GTFH-PA



Compound **S4.4** (160 mg, 249 µmol, 1.00 eq.) was dissolved under nitrogen atmosphere with DIPEA (260 µL, 193 mg 1.49 mmol, 6.00 eq.) in anhydrous DCM (4 mL). After 10 min CEP-Cl (72.2 µL, 76.6 mg, 324 µmol, 1.30 eq.) was added. Additional CEP-Cl (7.2 µL, 7.7 mg, 32.5 µmol, 0.13 eq.) was added after 2 h stirring at room temperature. The reaction mixture was stirred additionally at ambient temperature for 1 h. The solvent was removed under reduced pressure and the residue was purified by column chromatography (hexane/EtOAc 4:1 + 1% Et_3N) to afford compound **S4.5** as a colorless foam (185 mg, 218 µmol, 88%).

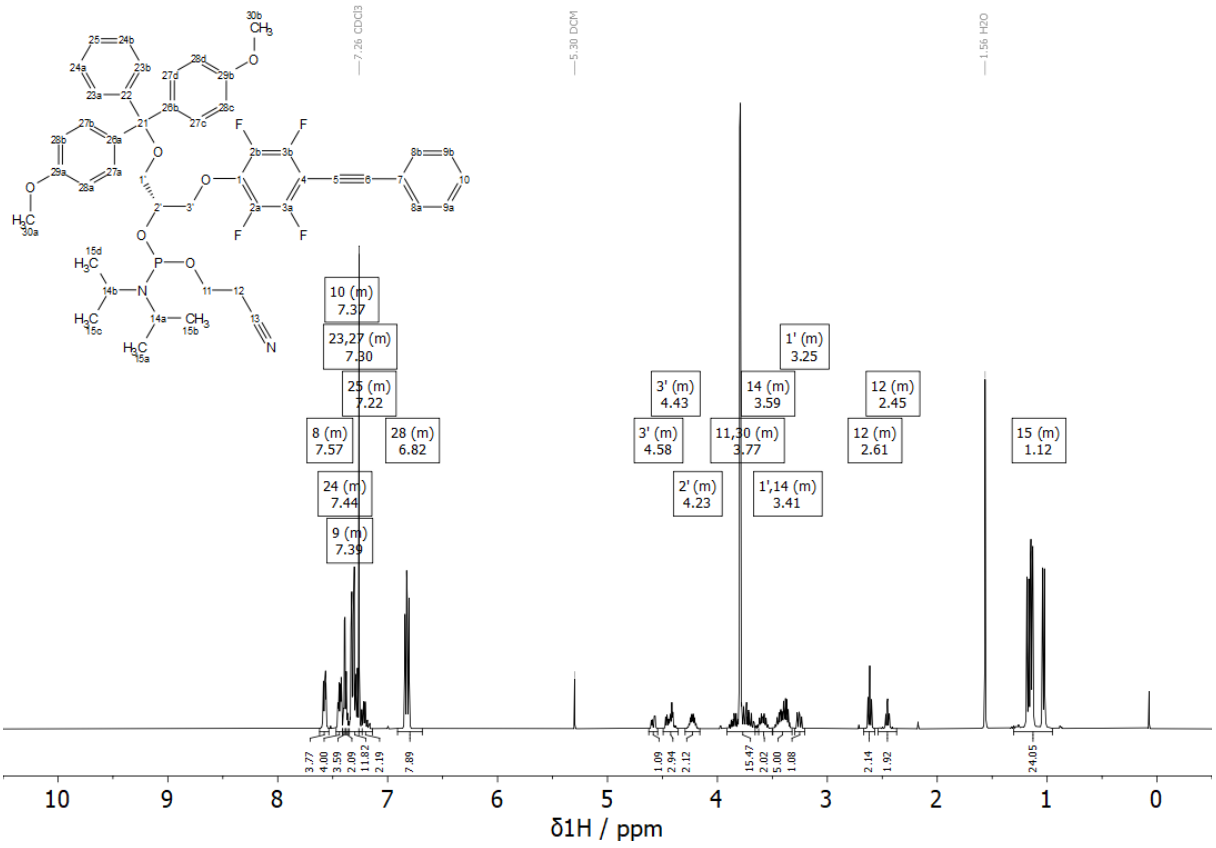
^{1}H NMR (400 MHz, $CDCl_3$): δ (ppm) = 7.62 – 7.54 (m, 4H, 8-H), 7.47 – 7.41 (m, 4H, DMT-H), 7.41 – 7.38 (m, 4H, 9-H), 7.38 – 7.35 (m, 2H, 10-H), 7.36 – 7.23 (m, 12H, DMT-H), 7.26 – 7.14 (m, 2H, DMT-H), 6.91 – 6.68 (m, 8H, DMT-H), 4.62 – 4.54 (m, 1H, 3'-H), 4.49 – 4.36 (m, 3H, 3'-H), 4.29 – 4.16 (m, 2H, 2'-H), 3.91 – 3.62 (m, 16H, 11-H, DMT-H), 3.66 – 3.50 (m, 2H, 14-H), 3.50 – 3.32 (m, 5H, 1'-H, 14-H), 3.30 – 3.21 (m, 1H, 1'-H), 2.67 – 2.57 (m, 2H, 12-H), 2.54 – 2.37 (m, 2H, 12-H), 1.30 – 0.95 (m, 24H, 15-H);

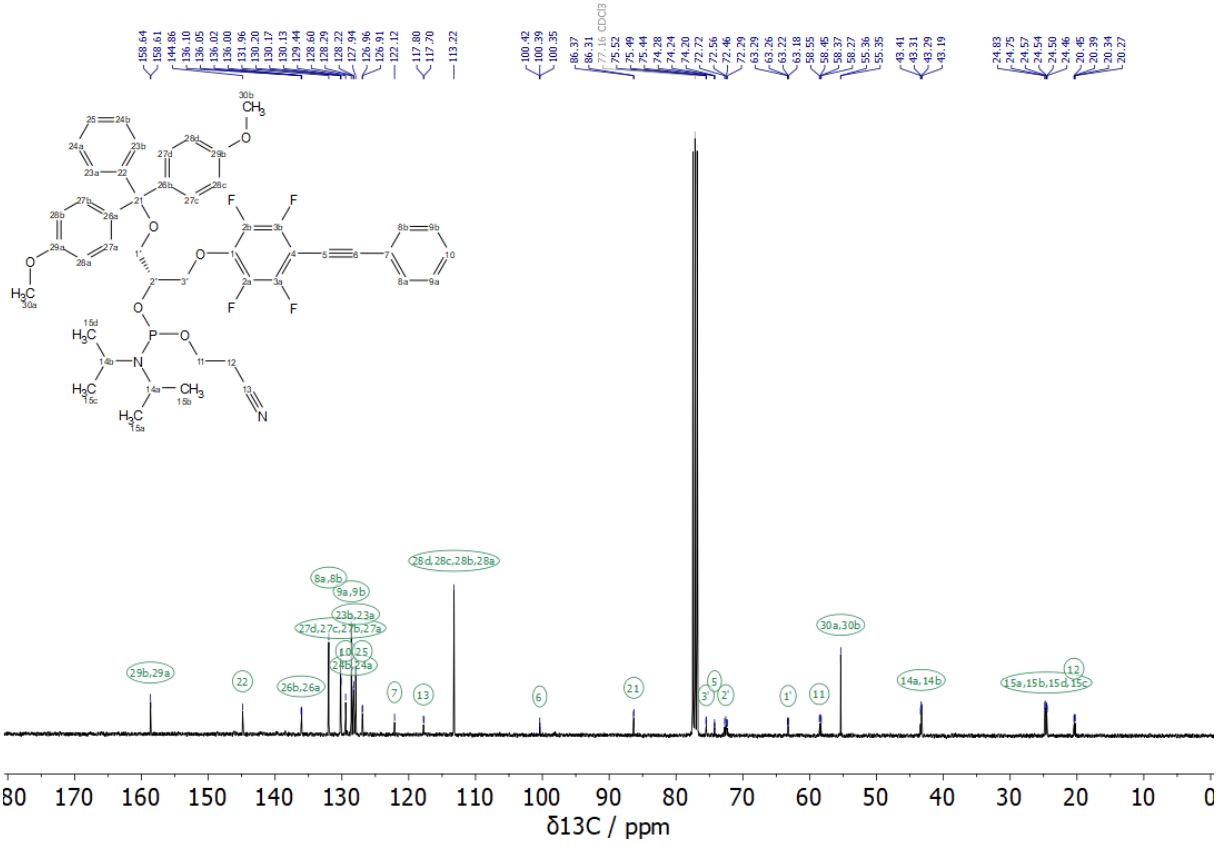
^{13}C{^{1}H} NMR (100 MHz, $CDCl_3$): δ (ppm) = 158.64 (DMT-C), 158.61 (DMT-C), 144.86 (DMT-C), 136.10 (DMT-C), 136.05 (DMT-C), 136.02 (DMT-C), 136.00 (DMT-C), 131.96 (8-C), 130.20 (DMT-C), 130.17 (DMT-C), 130.13 (DMT-C), 129.44 (10-C), 128.60 (9-C), 128.29 (DMT-C), 128.22 (DMT-C), 127.94 (DMT-C), 126.96 (DMT-C), 126.91 (DMT-C), 122.12 (7-C), 117.80 (13-C), 117.70 (13-C), 113.22 (DMT-C), 100.42 (6-C), 100.39 (6-C), 100.35 (6-C), 86.37 (DMT-C), 86.31 (DMT-C), 75.52 (3'-C), 75.49 (3'-C), 75.44 (3'-C), 74.28 (5-C), 74.24 (5-C), 74.20 (5-C), 72.72 (2'-C), 72.56 (2'-C), 72.46 (2'-C), 72.29 (2'-C), 63.29 (1'-C), 63.26 (1'-C), 63.22 (1'-C), 63.18 (1'-C), 58.55 (11-C), 58.45 (11-C), 58.37 (11-C), 58.27 (11-C), 55.36 (DMT-C), 55.35 (DMT-C), 43.41 (14-C), 43.31 (14-C), 43.29 (14-C), 43.19 (14-C), 24.83 (15-C), 24.75 (15-C), 24.57 (15-C), 24.54 (15-C), 24.50 (15-C), 24.46 (15-C), 20.45 (12-C), 20.39 (12-C), 20.34 (12-C), 20.27 (12-C);

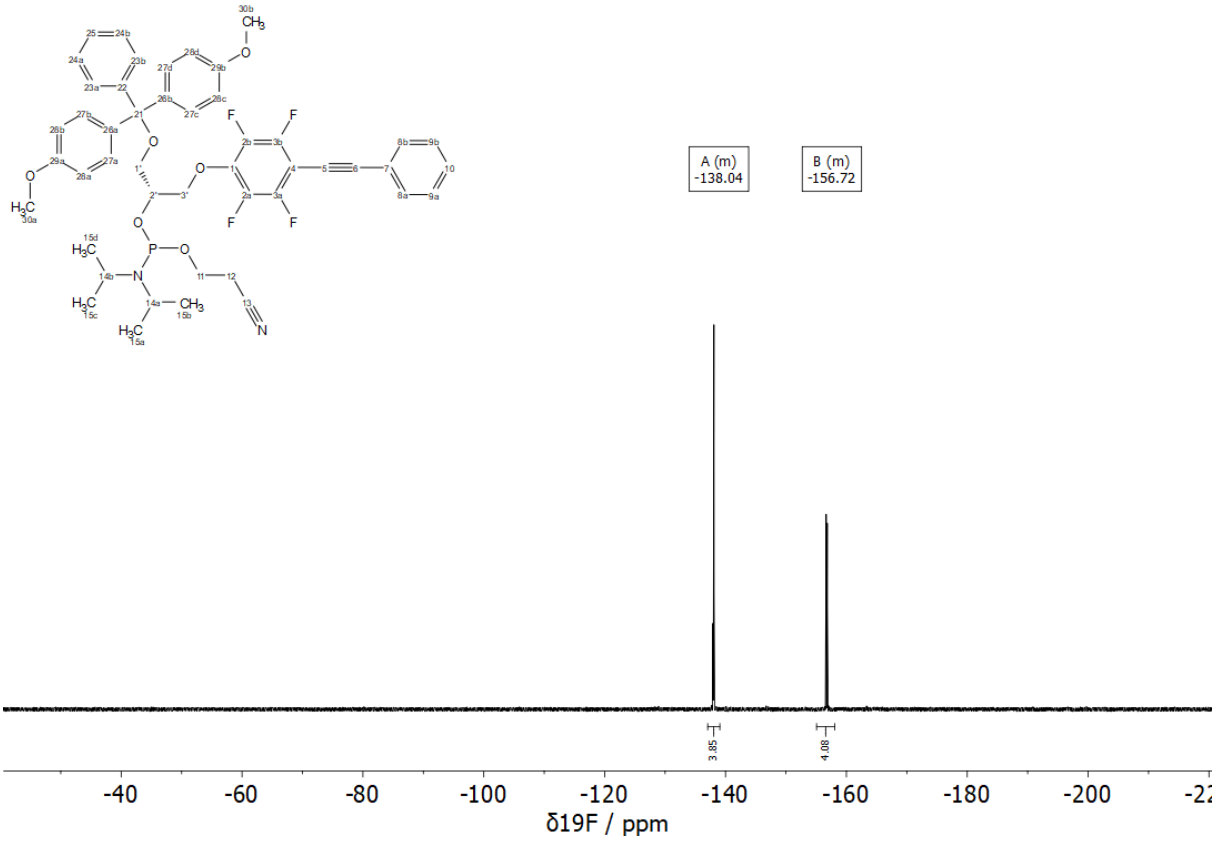
^{19}F{^{1}H} NMR (376 MHz, $CDCl_3$): δ (ppm) = -137.06 – -139.06 (m, 4F), -155.07 – -158.07 (m, 4F);

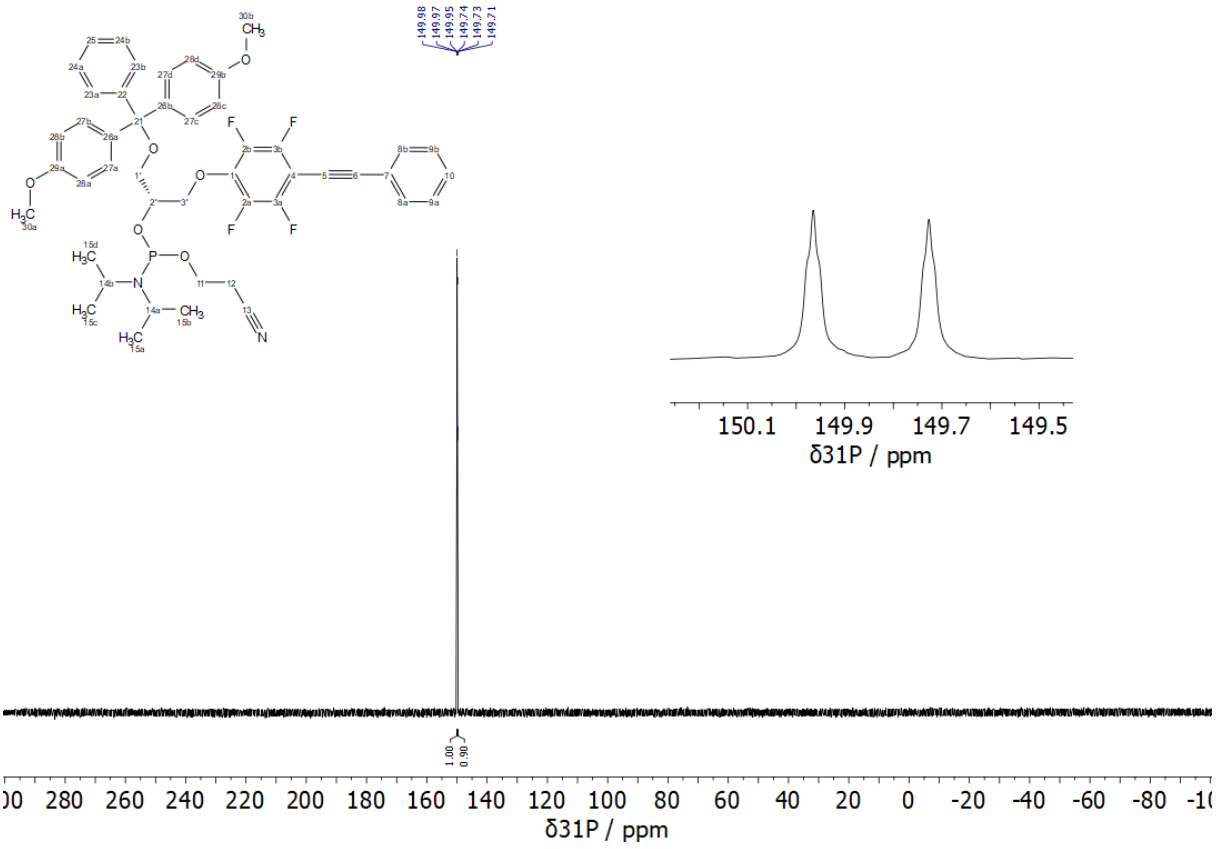
^{31}P{^{1}H} NMR (162 MHz, $CDCl_3$): δ (ppm) = 149.97 (t, J = 2.2 Hz), 149.73 (t, J = 2.0 Hz);

HR-MS (ESI+): m/z calc. ($C_{47}H_{48}F_4N_2O_6P$, $[M+H]^+$): 843.31806, found: 843.31963.

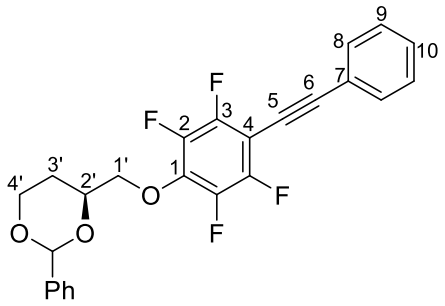

7.26 CDCl3
5.30 DCM
1.56 H2O
10 (m) 7.37
23,27 (m) 7.30
25 (m) 7.22
8 (m) 7.57
28 (m) 6.82
24 (m) 7.44
9 (m) 7.39
3' (m) 4.43
3' (m) 4.58
2' (m) 4.23
14 (m) 3.59
11,30 (m) 3.77
1',14 (m) 3.41
1' (m) 3.25
12 (m) 2.45
12 (m) 2.61
15 (m) 1.12
3.77
4.00
3.59
2.09
11.82
2.19
7.89
1.09
2.94
2.12
15.47
2.02
5.00
1.08
2.14
1.92
24.05
δ1H / ppm


77.16 CDCl3
28d,28c,28b,28a
8a,8b
9a,9b
23b,23a
27d,27c,27b,27a
10,25
24b,24a
29b,29a
22
26b,26a
7
13
6
21
5
3'
2'
1'
11
30a,30b
14a,14b
15a,15b,15d,15c
12
δ13C / ppm


A (m)
-138.04
B (m)
-156.72
3.85
4.08
-40
-60
-80
-100
-120
-140
-160
-180
-200
δ19F / ppm


149.98
149.97
149.95
149.74
149.73
149.71
150.1
149.9
149.7
149.5
δ31P / ppm
1.00
0.90
δ31P / ppm

Compound **S4.7**



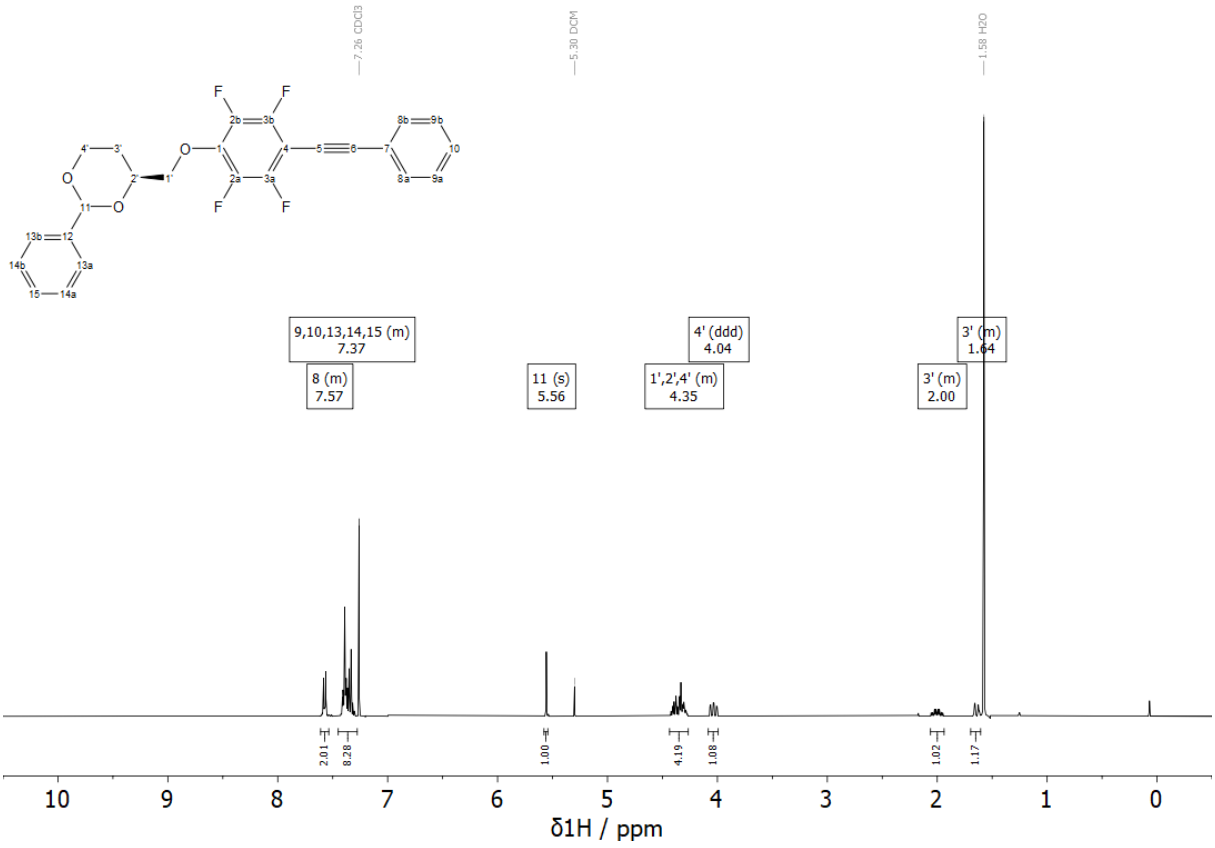
4-(*S*)-Hydroxmethyl-2-phenyl-1,3-dioxan (**4.1**) (500 mg, 2.57 mmol, 1.20 eq.) and sodium hydride (60% oil dispersion, 51.5 mg, 1.29 mmol, 0.60 eq.) were dissolved in anhydrous THF (25 mL) and stirred at ambient temperature for 10 min. 1,2,3,4,5-Pentafluoro-6-(2-phenylethynyl)benzene (**S4.6**) (575 mg, 2.15 mmol, 1.00 eq.) was added slowly to the mixture. After 3.5 h the solvent was evaporated, and the residue was purified by column chromatography (DCM/hexane 1:1) to afford compound **S4.7** as a colorless oil (562 mg, 1.27 mmol, 59%).

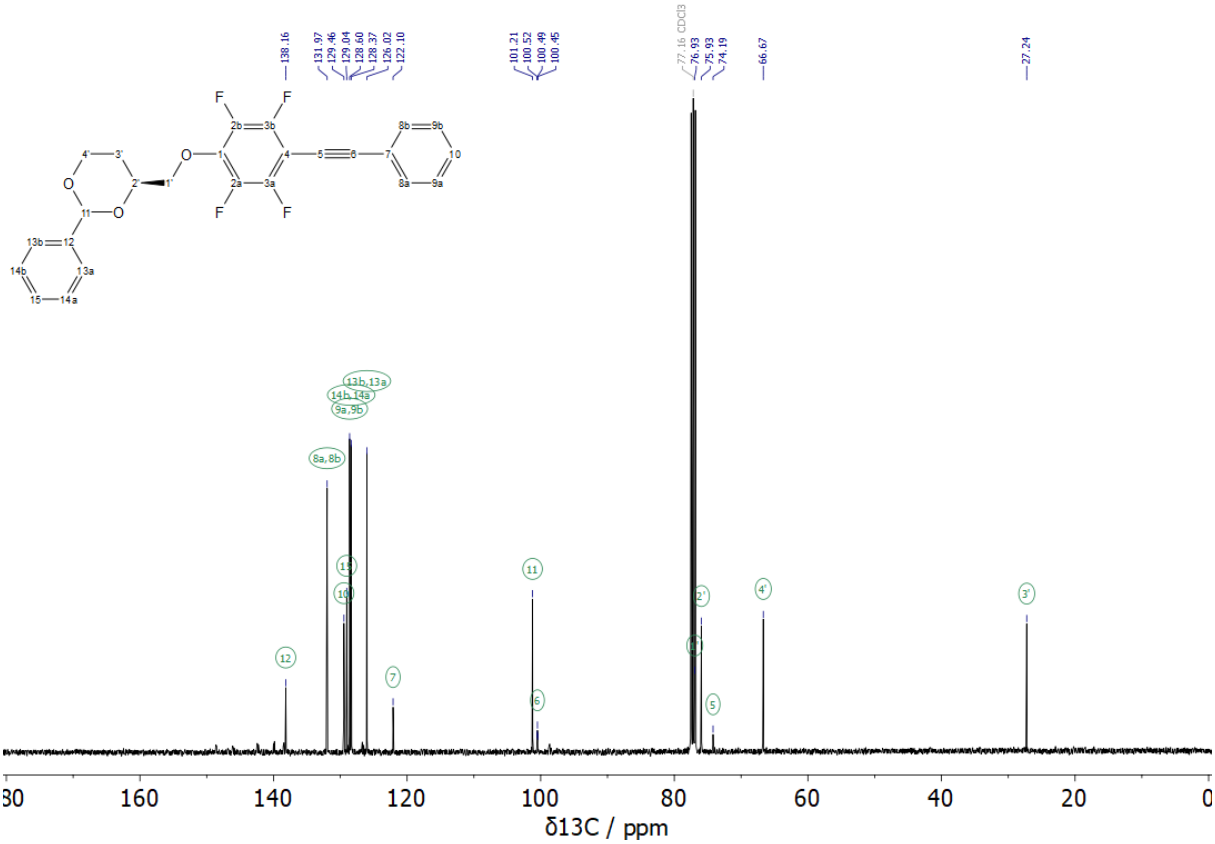
^{1}H NMR (400 MHz, $CDCl_3$): δ (ppm) = 7.61 – 7.54 (m, 2H, 8-H), 7.45 – 7.28 (m, 8H, 9-H, 10-H, Ph-H), 5.56 (s, 1H, C*H*-Ph), 4.44 – 4.27 (m, 4H, 1'-H, 2'-H, 3'-H), 4.04 (ddd, *J* = 12.3, 11.5, 2.6 Hz, 1H, 4'-H), 2.06 – 1.94 (m, 1H, 3'-H), 1.70 – 1.60 (m, 1H, 3'-H);

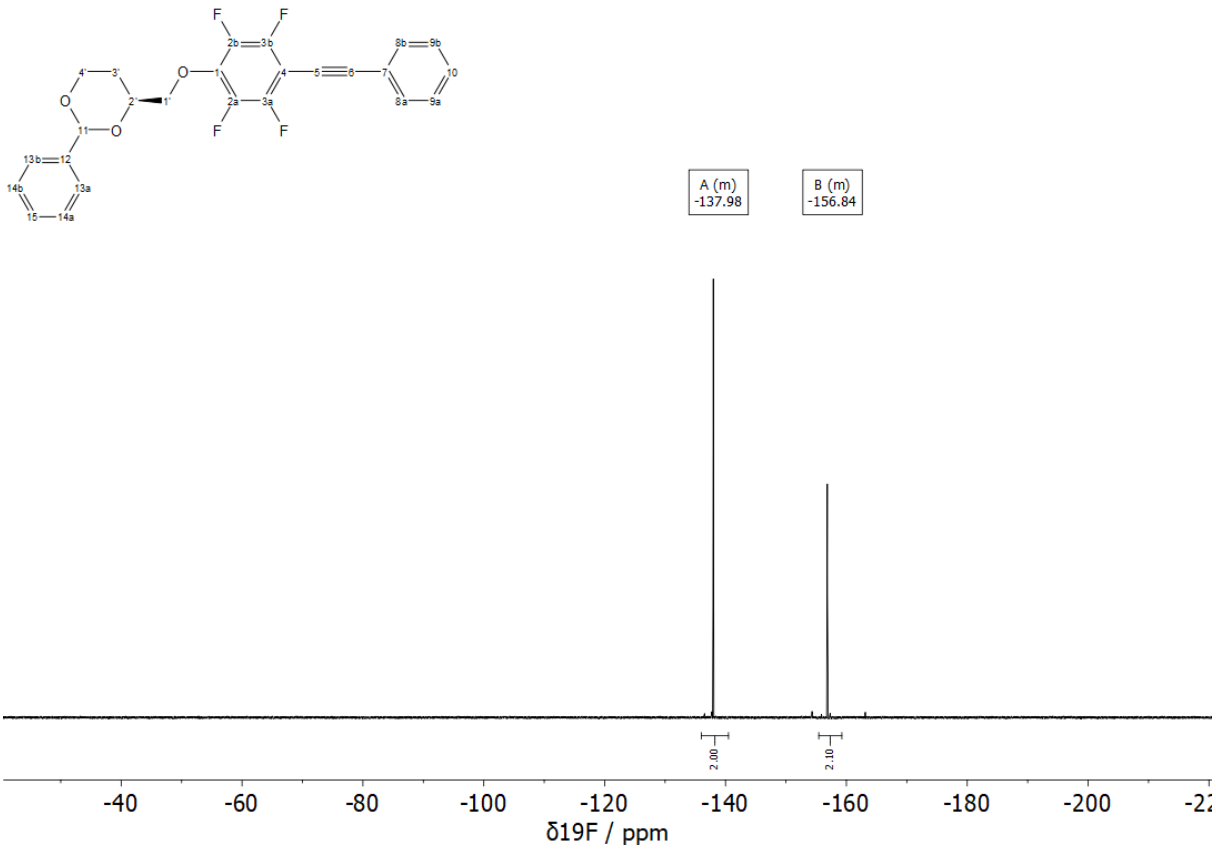
^{13}C{^{1}H} NMR (100 MHz, $CDCl_3$): δ (ppm) = 138.16 (12-C), 131.97 (8-C), 129.46 (10-C), 129.04 (15-C), 128.60 (9-C), 128.37 (14-C), 126.02 (13-C), 122.10 (7-C), 101.21 (*C*H-Ph), 100.49 (t, *J* = 3.6 Hz, 6-C), 76.93 (1'-C), 75.93 (2'-C), 74.19 (5-C), 66.67 (4'-C), 27.24 (3'-C);

^{19}F{^{1}H} NMR (376 MHz, $CDCl_3$): δ (ppm) = -136.00 – -140.49 (m), -155.43 – -159.25 (m);

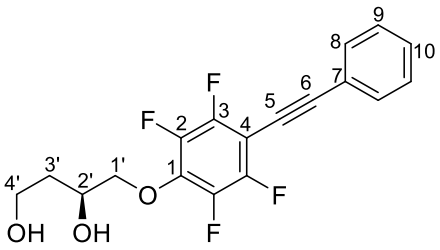
HR-MS (ESI+): *m/z* calc. ($C_{25}H_{18}F_4O_3Na$, $[M+Na]^+$): 465.10843, found: 465.11148.


7.26 CDCl3
5.30 DCM
1.58 H2O
9,10,13,14,15 (m)
7.37
8 (m)
7.57
11 (s)
5.56
1',2',4' (m)
4.35
4' (ddd)
4.04
3' (m)
2.00
3' (m)
1.64
2.01
8.28
1.00
4.19
1.08
1.02
1.17
δ1H / ppm


138.16
131.97
129.46
129.04
128.60
128.37
126.02
122.10
101.21
100.52
100.49
100.45
77.16 CDCl3
76.93
75.93
74.19
66.67
27.24
13b,13a
14b,14a
9a,9b
8a,8b
δ13C / ppm



Compound **S4.8**



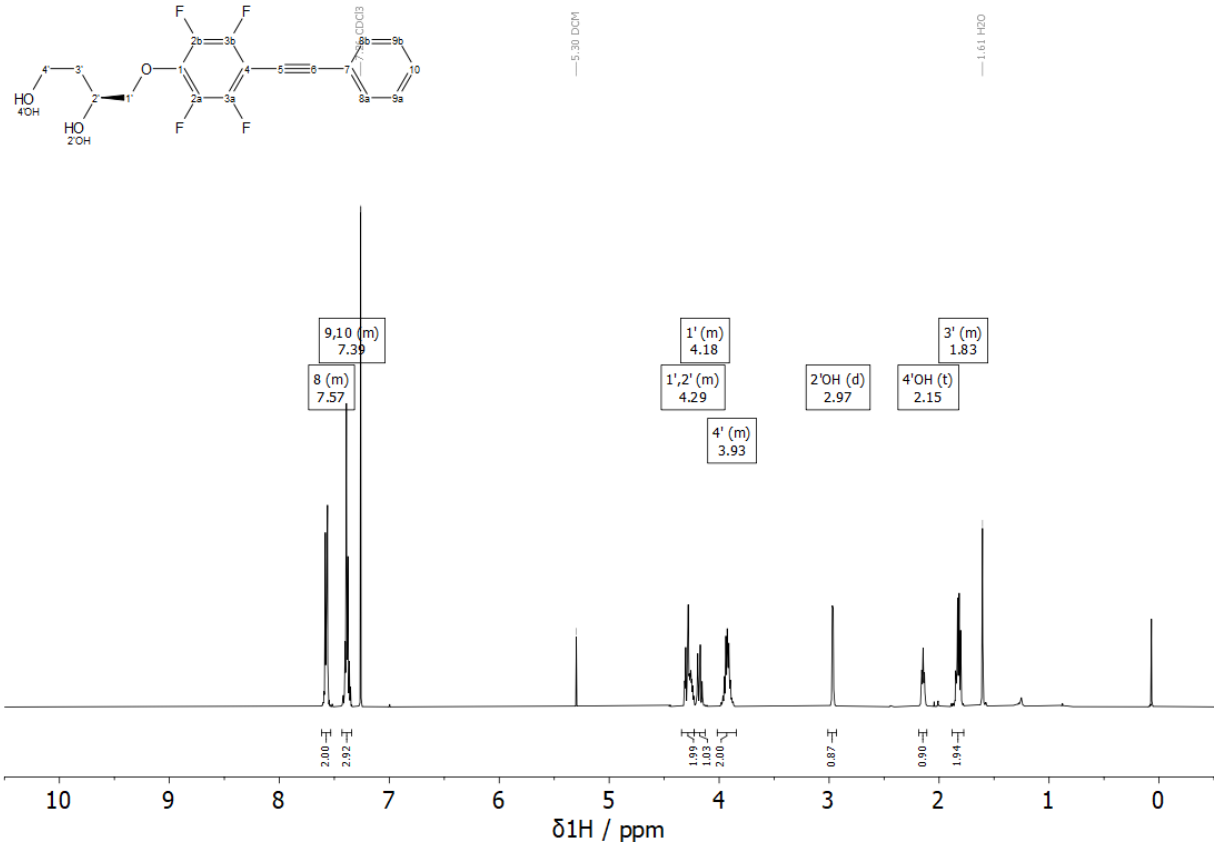
Compound **S4.7** (411 mg, 929 µmol, 1.00 eq.) was dissolved in methanol (7.5 mL). *p*-Toluenesulfonic acid monohydrate (283 mg, 1.49 mmol, 1.60 eq.) was added and the reaction mixture stirred at ambient temperature for 22 h. Afterwards, the reaction was diluted with a potassium carbonate solution (10%, 2 mL) and extracted with DCM. The combined organic layers were dried over Na_2SO_4, evaporated and the solid residue was purified by column chromatography (hexane/EtOAc 1:1→1:2) to afford compound **S4.8** as a colorless solid (300 mg, 846 µmol, 91%).

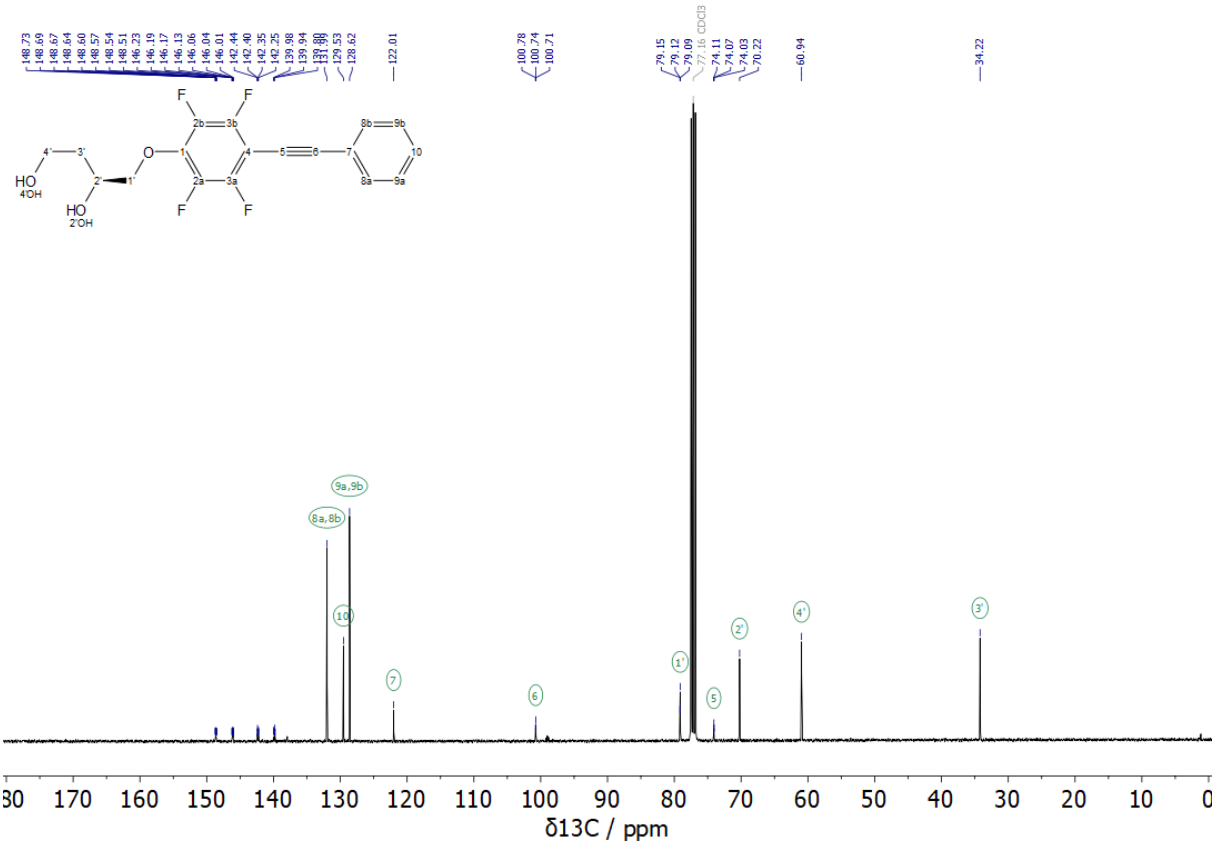
^{1}H NMR (400 MHz, $CDCl_3$): δ (ppm) = 7.61 – 7.53 (m, 2H, 8-H), 7.43 – 7.34 (m, 3H, 9-H, 10-H), 4.34 – 4.23 (m, 2H, 1'-H. 2'-H), 4.23 – 4.13 (m, 1H, 1'-H), 4.02 – 3.84 (m, 2H, 4'-H), 2.97 (d, *J* = 3.2 Hz, 1H, 2'-OH), 2.15 (t, *J* = 5.0 Hz, 1H, 4'-OH), 1.88 – 1.78 (m, 2H, 3'-H);

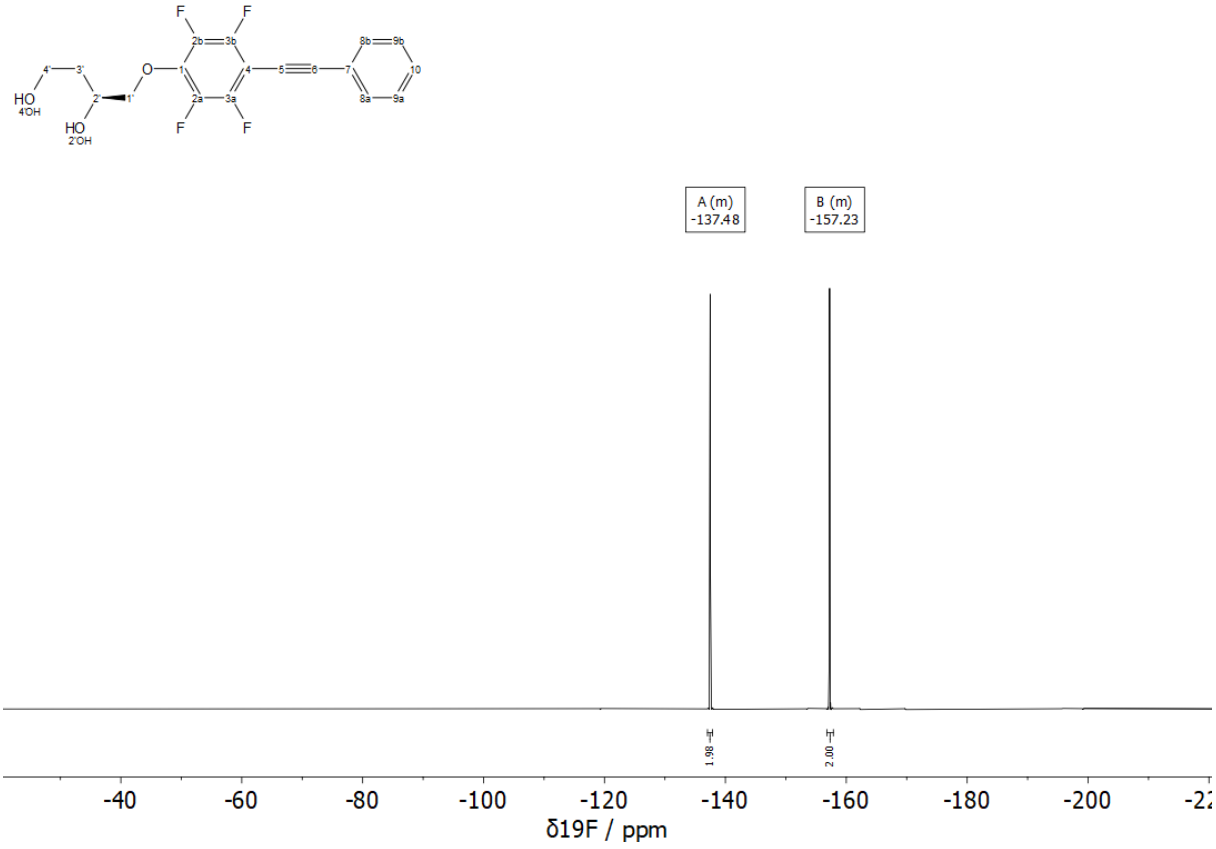
^{13}C{^{1}H} NMR (100 MHz, $CDCl_3$): δ (ppm) = 148.87 – 148.33 (m), 146.34 – 145.91 (m), 142.63 – 142.02 (m), 140.22 – 139.63 (m), 131.99 (8-C), 129.53 (10-C), 128.62 (9-C), 122.01 (7-C), 100.74 (t, *J* = 3.6 Hz, 6-C), 79.12 (t, *J* = 3.1 Hz, 1'-C), 74.07 (t, *J* = 4.1 Hz, 5-C), 70.22 (2'-C), 60.94 (4'-C), 34.22 (3'-C);

^{19}F{^{1}H} NMR (376 MHz, $CDCl_3$): δ (ppm) = -137.04 – -137.85 (m), -156.78 – -157.86 (m);

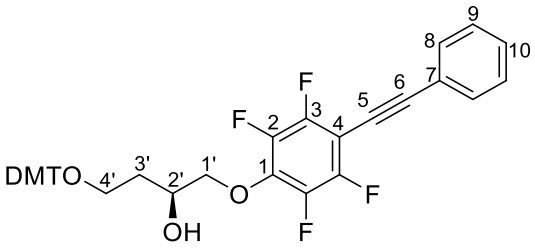
HR-MS (ESI+): *m/z* calc. ($C_{18}H_{14}F_4O_3Na$, $[M+Na]^+$): 377.07713, found: 377.07868.




148.73
148.69
148.67
148.64
148.60
148.57
148.54
148.51
146.23
146.19
146.17
146.13
146.06
146.04
146.01
142.44
142.40
142.35
142.25
139.98
139.94
131.99
129.53
128.62
122.01
100.78
100.74
100.71
79.15
79.12
79.09
77.16 CDCl3
74.11
74.07
74.03
70.22
60.94
34.22
F
F
2b
3b
8b
9b
4'
3'
O
1
4
5
6
7
10
1'
2'
2a
3a
8a
9a
HO
4'OH
HO
2'OH
9a,9b
8a,8b
10
7
6
1'
5
2'
4'
3'
δ13C / ppm


A (m)
-137.48
B (m)
-157.23
1.98
2.00
δ19F / ppm

Compound **S4.9**



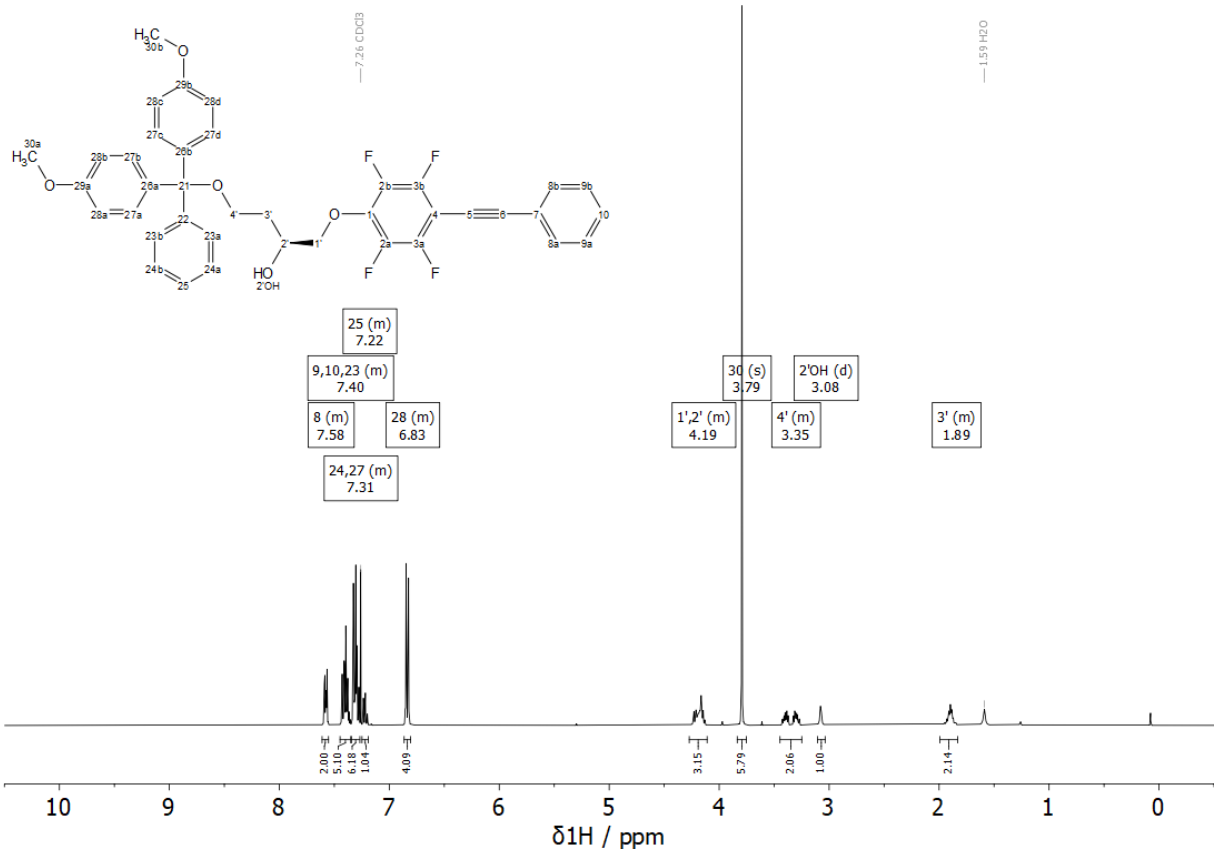
Compound **S4.8** (253 mg, 715 µmol, 1.00 eq.) was dissolved in anhydrous pyridine (5.5 mL). 4,4'-Dimethoxytrityl chloride (291 mg, 859 µmol, 1.20 eq.) was added in small portions over a time period of 30 min and the reaction mixture stirred at ambient temperature for 24 h. The solvent was removed under reduced pressure and the residue was purified by column chromatography (hexane/EtOAc 7:1 +1% Et_3N→4:1 +1% Et_3N) to afford compound **S4.9** as a colorless foam (312 mg, 475 µmol, 66%).

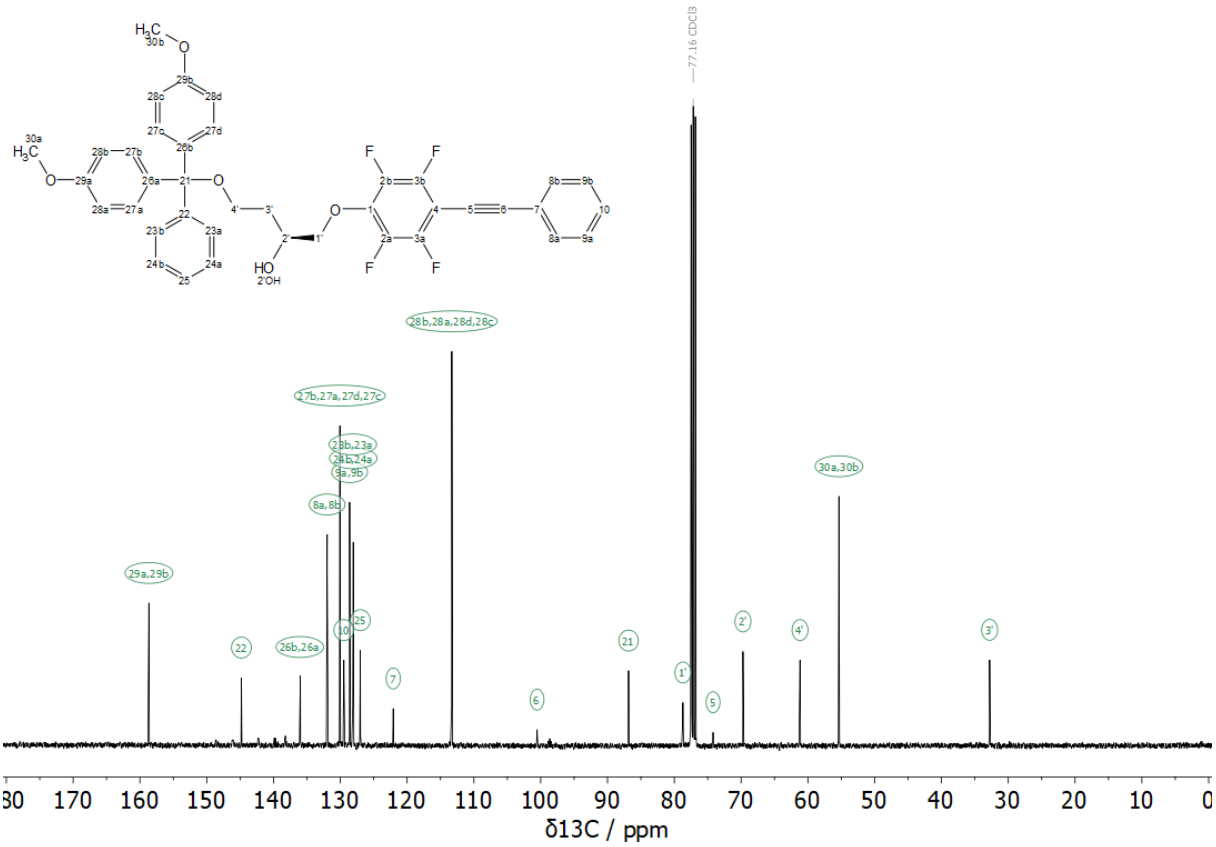
^{1}H NMR (400 MHz, $CDCl_3$): δ (ppm) = 7.61 – 7.55 (m, 2H, 8-H), 7.45 – 7.35 (m, 5H, 9-H, 10-H, DMT-H), 7.34 – 7.27 (m, 6H, DMT-H), 7.25 – 7.19 (m, 1H, DMT-H), 6.86 – 6.81 (m, 4H, DMT-H), 4.27 – 4.11 (m, 3H, 1'-H, 2'-H), 3.79 (s, 6H, DMT-H), 3.45 – 3.25 (m, 2H, 4'-H), 3.08 (d, *J* = 2.9 Hz, 1H, 2'-OH), 1.99 – 1.83 (m, 2H, 3'-H);

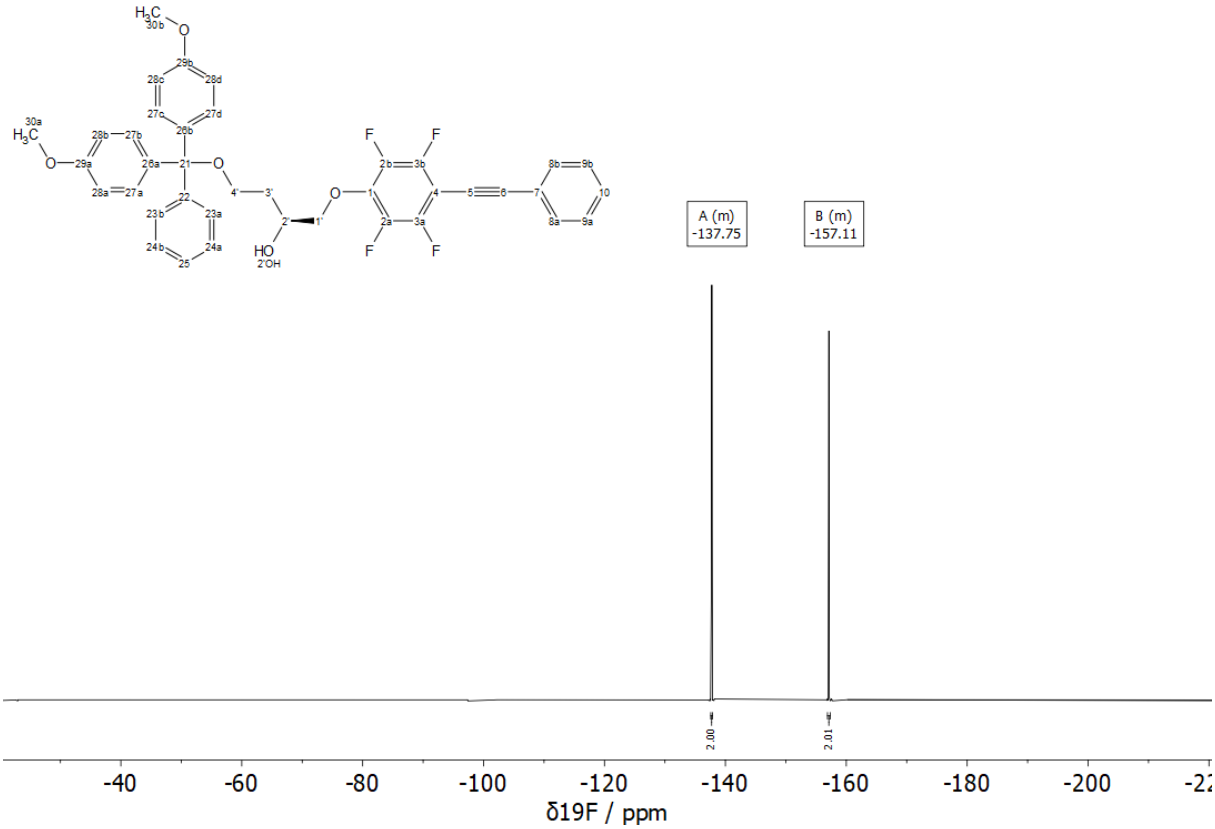
^{13}C{^{1}H} NMR (100 MHz, $CDCl_3$): δ (ppm) = 158.65 (DMT-C), 148.80 – 148.45 (m), 146.28 – 145.93 (m), 144.81 (DMT-C), 142.62 – 142.04 (m), 140.07 – 139.62 (m), 138.49 – 137.91 (m), 136.06 (DMT-C), 135.98 (DMT-C), 131.97 (8-C), 130.06 (DMT-C), 129.47 (10-C), 128.60 (9-C), 128.12 (DMT-C), 128.06 (DMT-C), 127.00 (DMT-C), 122.07 (7-C), 113.31 (DMT-C), 100.54 (t, *J* = 3.6 Hz, 6-C), 86.84 (DMT-C), 78.71 (t, *J* = 3.0 Hz, 1'-C), 74.19 (t, *J* = 4.3 Hz, 5-C), 69.72 (2'-C), 61.18 (4'-C), 55.35 (DMT-C), 32.79 (3'-C);

^{19}F{^{1}H} NMR (376 MHz, $CDCl_3$): δ (ppm) = -137.58 – -137.88 (m), -156.89 – -157.34 (m);

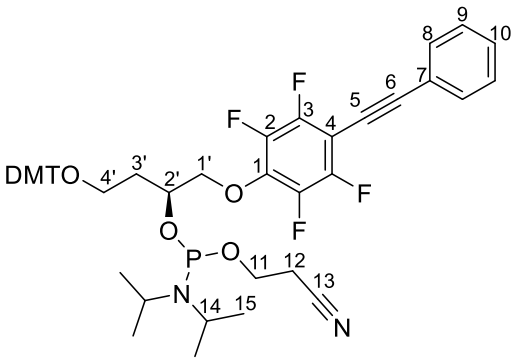
HR-MS (ESI+): *m/z* calc. ($C_{39}H_{32}F_4O_5Na$, $[M+Na]^+$): 679.20781, found: 679.20858.


7.26 CDCl3
1.59 H2O
25 (m) 7.22
9,10,23 (m) 7.40
8 (m) 7.58
28 (m) 6.83
24,27 (m) 7.31
30 (s) 3.79
2'OH (d) 3.08
1',2' (m) 4.19
4' (m) 3.35
3' (m) 1.89
2.00
5.10
6.18
1.04
4.09
3.15
5.79
2.06
1.00
2.14
δ1H / ppm


77.16 CDCl3
28b,28a,28d,28c
27b,27a,27d,27c
23b,23a
26b,24a
5a,9b
8a,8b
30a,30b
29a,29b
22
26b,26a
10
25
7
6
21
1
5
2'
4'
3'
δ13C / ppm



Compound **S4.10** = BTFH-PA



Compound **S4.9** (246 mg, 375 µmol, 1.00 eq.) was dissolved under nitrogen atmosphere with DIPEA (399 µL, 296 mg 2.29 mmol, 6.1 eq.) in anhydrous DCM (10 mL). After 10 min CEP-Cl (120 mg, 507 µmol, 1.35 eq.) was added. Additional CEP-Cl (12.0 mg, 50.7 µmol, 0.14 eq.) was added after 2 h stirring at room temperature. The reaction mixture was stirred additionally at ambient temperature for 30 min. The solvent was removed under reduced pressure and the residue was purified by column chromatography (hexane/EtOAc 7:1 + 1% Et_3N –6:1 + 1% Et_3N) to afford compound **S4.10** as a colorless foam (231 mg, 270 µmol, 72%).

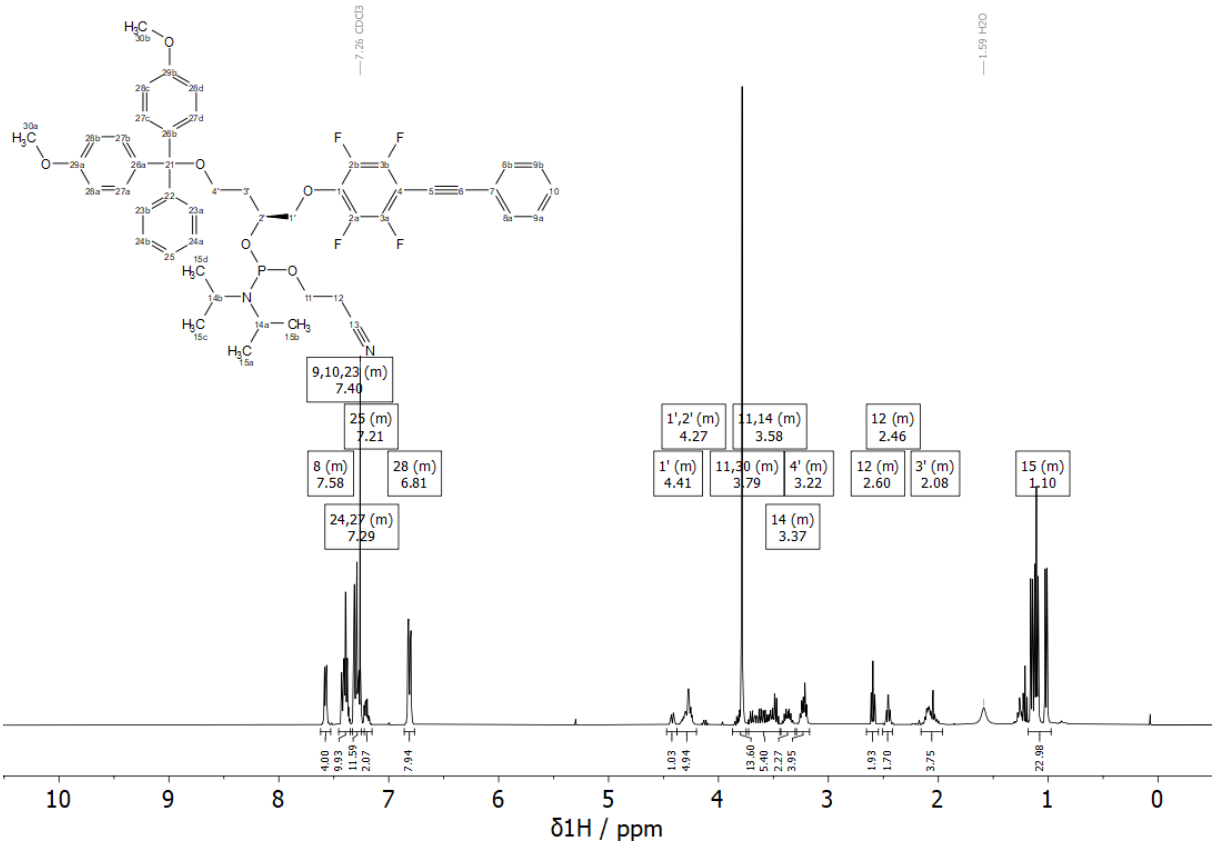
^{1}H NMR (400 MHz, $CDCl_3$): δ (ppm) = 7.62 – 7.53 (m, 4H, 8-H), 7.46 – 7.33 (m, 10H, 9-H, 10-H, DMT-H), 7.35 – 7.22 (m, 12H, DMT-H), 7.25 – 7.15 (m, 2H, DMT-H), 6.86 – 6.77 (m, 8H, DMT-H), 4.47 – 4.38 (m, 1H, 1'-H), 4.38 – 4.20 (m, 5H, 1'-H, 2'-H), 3.87 – 3.73 (m, 13H, 11-H, DMT-H), 3.75 – 3.43 (m, 5H, 11-H, 14-H), 3.44 – 3.30 (m, 2H, 14-H), 3.29 – 3.17 (m, 4H, 4'-H), 2.65 – 2.54 (m, 2H, 12-H), 2.51 – 2.42 (m, 2H, 12-H), 2.16 – 1.96 (m, 4H, 3'-H), 1.18 – 0.97 (m, 24H, 15-H);

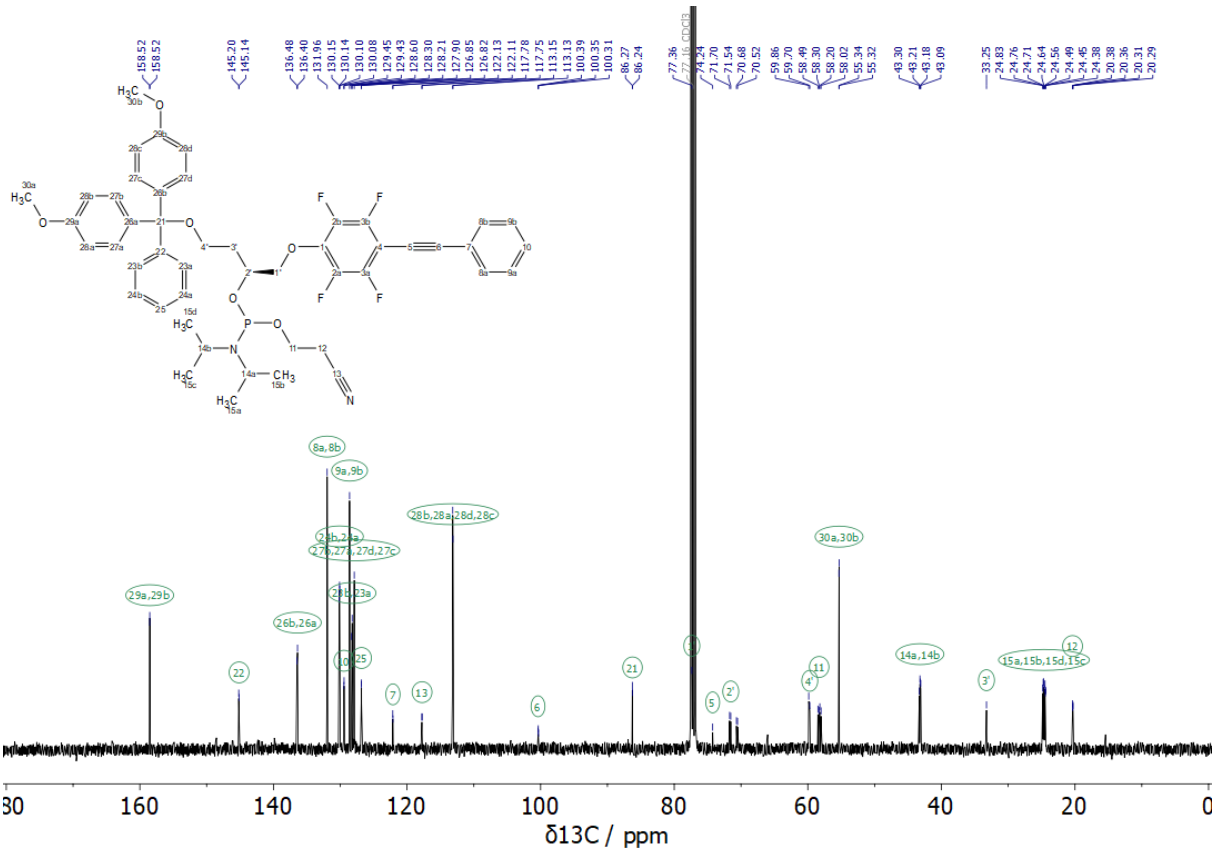
^{13}C{^{1}H} NMR (100 MHz, $CDCl_3$): δ (ppm) = 158.52 (DMT-C), 158.52 (DMT-C), 145.20 (DMT-C), 145.14 (DMT-C), 136.48 (DMT-C), 136.40 (DMT-C), 131.96 (8-C), 130.15 (DMT-C), 130.14 (DMT-C), 130.10 (DMT-C), 130.08 (DMT-C), 129.45 (10-C), 129.43 (10-C), 128.60 (9-C), 128.30 (DMT-C), 128.21 (DMT-C), 127.90 (DMT-C), 126.85 (DMT-C), 126.82 (DMT-C), 122.13 (7-C), 122.11 (7-C), 117.78 (13-C), 117.75 (13-C), 113.15 (DMT-C), 113.13 (DMT-C), 100.39 (6-C), 100.35 (6-C), 100.31 (6-C), 86.27 (DMT-C), 86.24 (DMT-C), 77.36 (1'-C), 74.24 (5-C), 71.70 (2'-C), 71.54 (2'-C), 70.68 (2'-C), 70.52 (2'-C), 59.86 (4'-C), 59.70 (4'-C), 58.49 (11-C), 58.30 (11-C), 58.20 (11-C), 58.02 (11-C), 55.34 (DMT-C), 55.32 (DMT-C), 43.30 (14-C), 43.21 (14-C), 43.18 (14-C), 43.09 (14-C), 33.25 (3'-C), 24.83 (15-C), 24.76 (15-C), 24.71 (15-C), 24.64 (15-C), 24.56 (15-C), 24.49 (15-C), 24.45 (15-C), 24.38 (15-C), 20.38 (12-C), 20.36 (12-C), 20.31 (12-C), 20.29 (12-C);

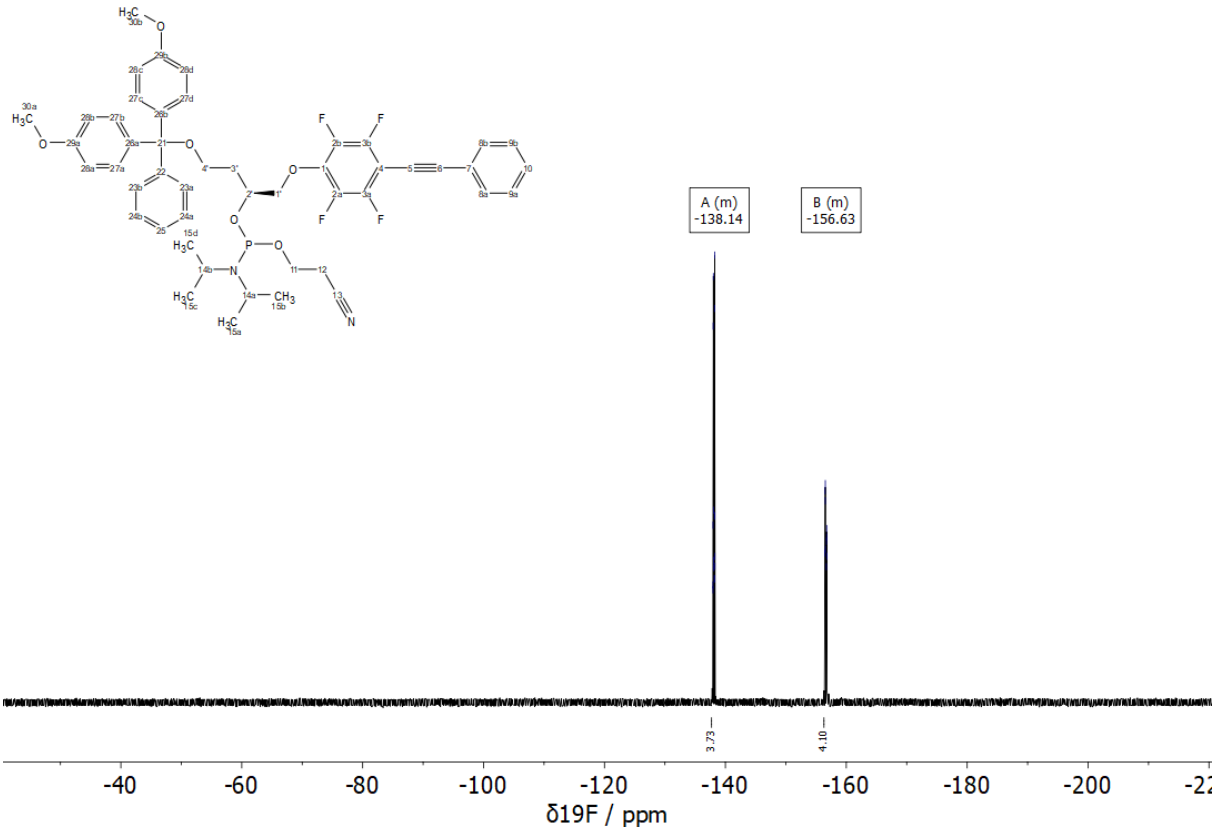
^{19}F{^{1}H} NMR (376 MHz, $CDCl_3$): δ (ppm) = -136.76 – -139.73 (m, 4F), -153.91 – -161.12 (m, 4F);

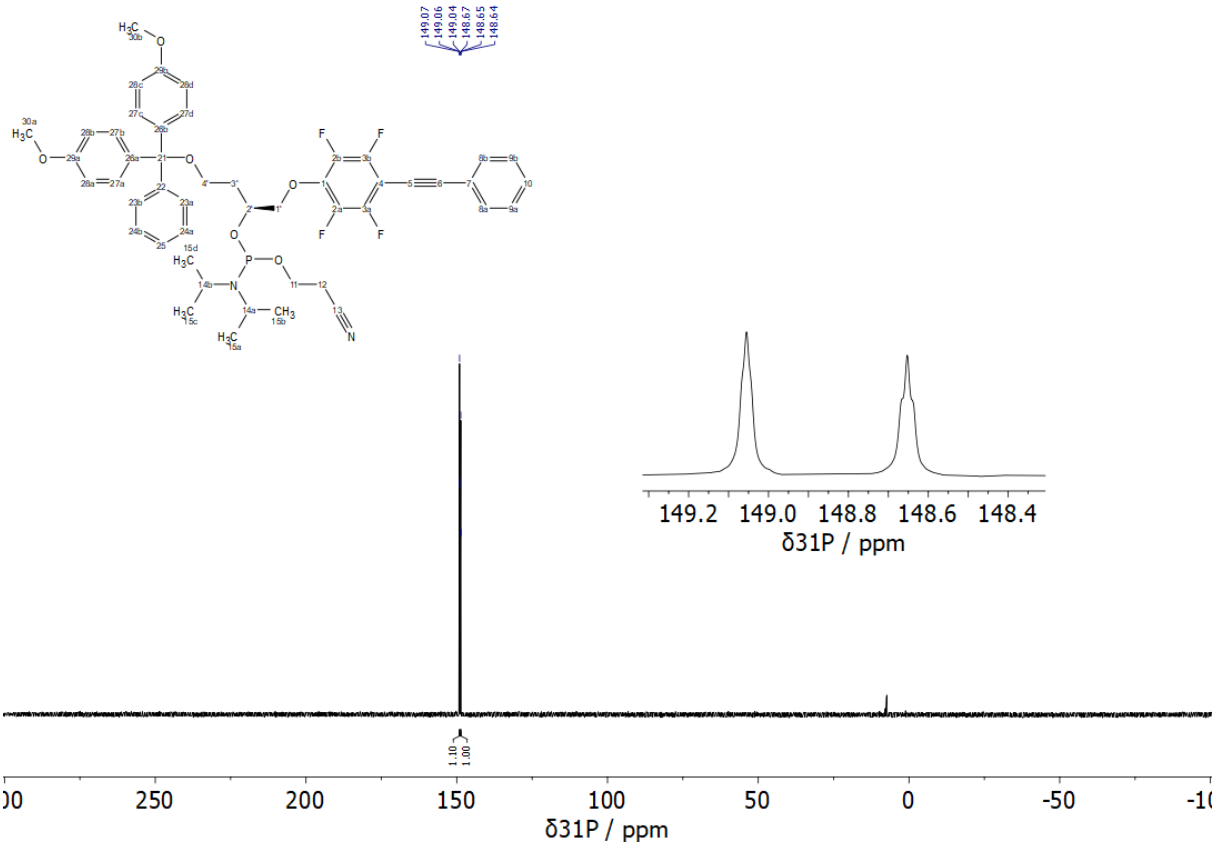
^{31}P{^{1}H} NMR (162 MHz, $CDCl_3$): δ (ppm) = 149.06 (t, J = 2.0 Hz), 148.65 (t, J = 2.6 Hz);

HR-MS (ESI+): m/z calc. ($C_{48}H_{49}F_4N_2O_6PNa$, $[M+Na]^+$): 879.31566, found: 879.3158.

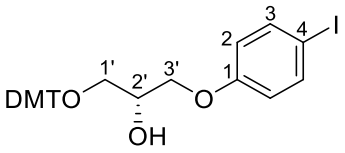

7.26 CDCl3
1.59 H2O
9,10,23 (m)
7.40
25 (m)
7.21
8 (m)
7.58
28 (m)
6.81
24,27 (m)
7.29
1',2' (m)
4.27
11,14 (m)
3.58
1' (m)
4.41
11,30 (m)
3.79
4' (m)
3.22
14 (m)
3.37
12 (m)
2.46
12 (m)
2.60
3' (m)
2.08
15 (m)
1.10
4.00
9.93
11.59
2.07
7.94
1.03
4.94
13.60
5.40
2.27
3.95
1.93
1.70
3.75
22.98
10
9
8
7
6
5
4
3
2
1
0
δ1H / ppm


158.52
158.52
145.20
145.14
136.48
136.40
131.96
130.15
130.14
130.10
130.08
129.45
129.43
128.60
128.30
128.21
127.90
126.85
126.82
122.13
122.11
117.78
117.75
113.15
113.13
100.39
100.35
100.31
86.27
86.24
77.36
77.16 CDCl3
74.24
71.70
71.54
70.68
70.52
59.86
59.70
58.49
58.30
58.20
58.02
55.34
55.32
43.30
43.21
43.18
43.09
33.25
24.83
24.76
24.71
24.64
24.56
24.49
24.45
24.38
20.38
20.36
20.31
20.29
8a,8b
9a,9b
28b,28a,28d,28c
24b,24a
27b,27a,27d,27c
26b,26a
29a,29b
30a,30b
14a,14b
15a,15b,15d,15c
22
25
7
13
6
21
5
2'
4
11
3'
12
180
160
140
120
100
80
60
40
20
0
δ13C / ppm


A (m)
-138.14
B (m)
-156.63
3.73
4.10
-40
-60
-80
-100
-120
-140
-160
-180
-200
δ19F / ppm


149.07
149.06
149.04
148.67
148.65
148.64
149.2
149.0
148.8
148.6
148.4
δ31P / ppm
1.10
1.00
250
200
150
100
50
0
-50
δ31P / ppm

Compound **S4.11**

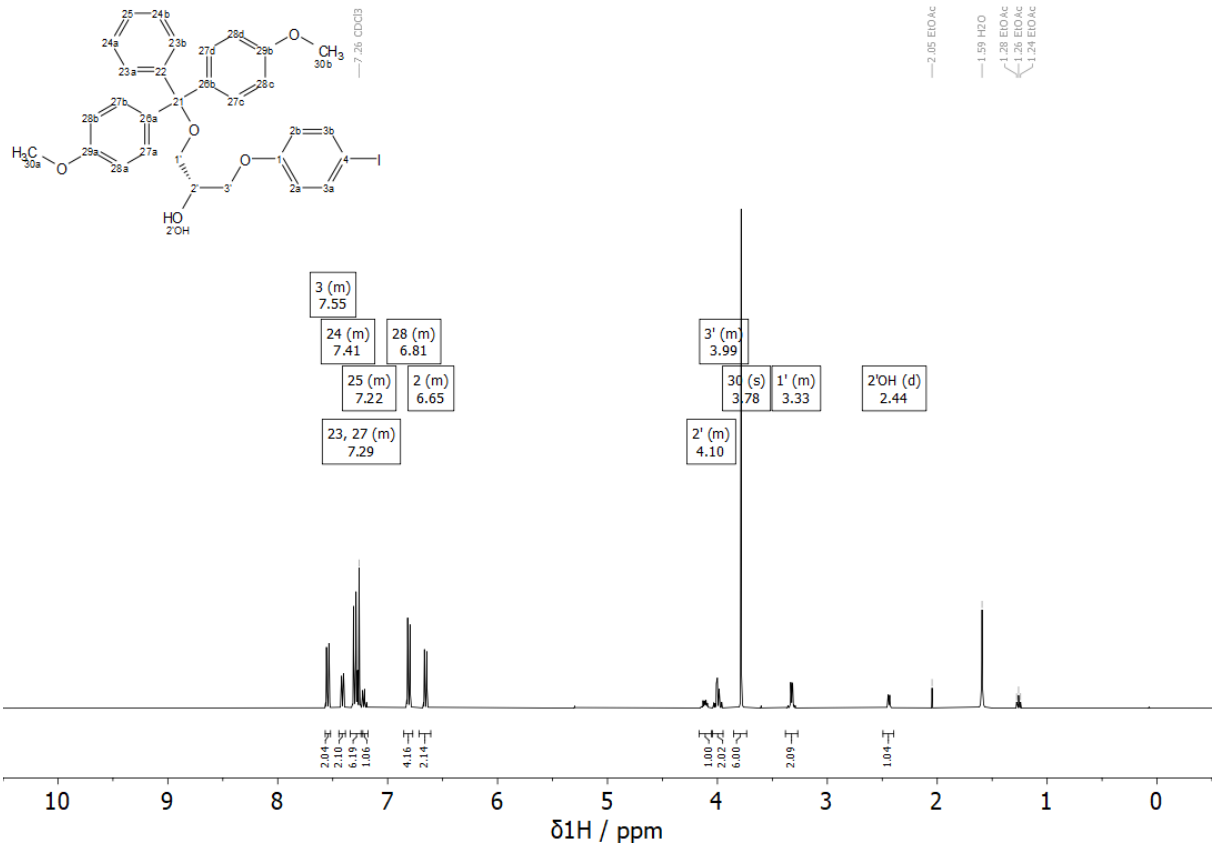


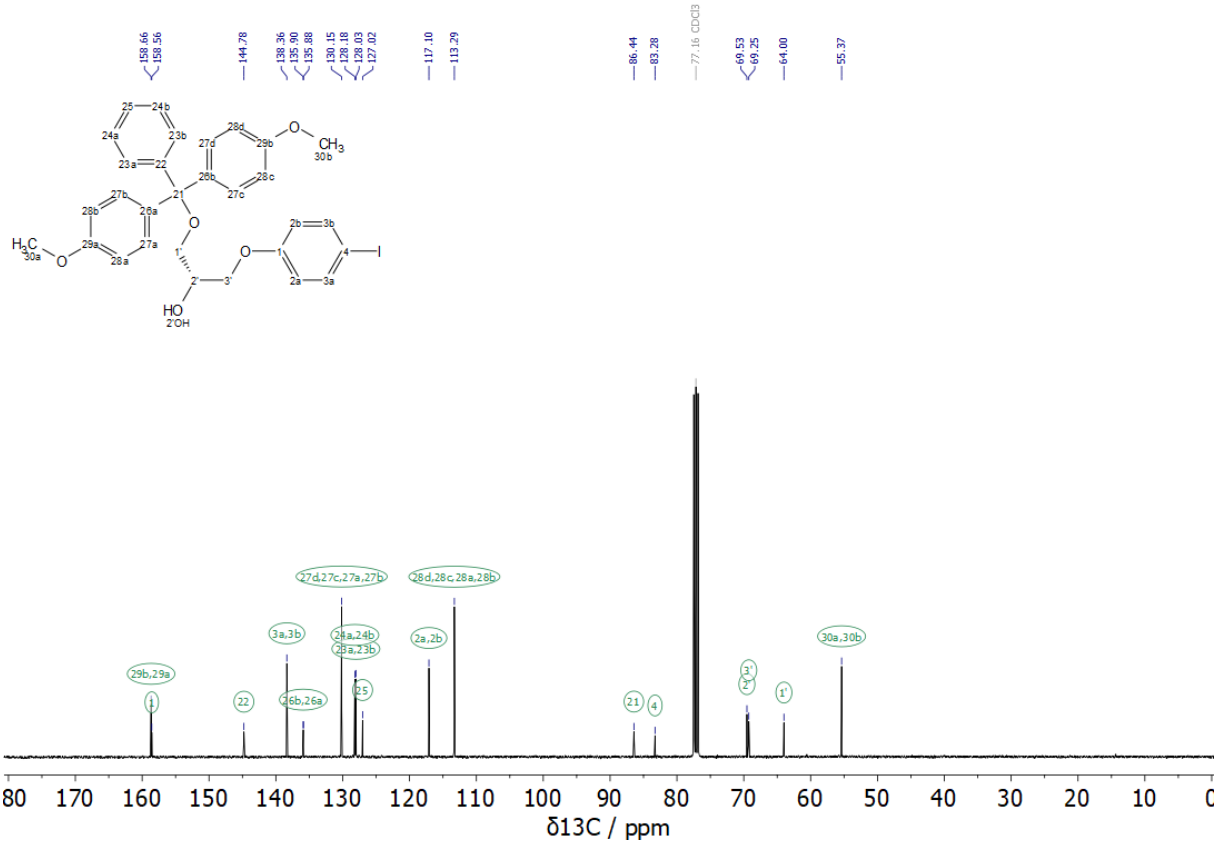
Under nitrogen atmosphere 4-iodophenol (961 mg, 4.37 mmol, 1.12 eq.) and sodium hydride (60% oil dispersion, 77.7 mg, 1.94 mmol, 0.50 eq.) were suspended in DMF (12 mL) and stirred for 1 h. A solution of (*S*)-DMT-*O*-glycidol (**4.2**) (1.46 g, 3.88 mmol. 1.00 eq.) in anhydrous DMF (12 mL) was added to a mixture and stirred additional 20 h at 90°C. The solvent was removed under reduced pressure and the residue was purified by column chromatography (hexane/EtOAc 4:1 + 1% Et_3N → 3:1 + 1% Et_3N) to afford compound **S4.11** (1.56 g, 2.62 mmol, 67%) as a colorless foam.

^{1}H NMR (400 MHz, $CDCl_3$): δ (ppm) = 7.57 – 7.52 (m, 2H, 3-H), 7.44 – 7.38 (m, 2H, DMT-H), 7.34 – 7.23 (m, 6H, DMT-H), 7.24 – 7.18 (m, 1H, DMT-H), 6.85 – 6.77 (m, 4H, DMT-H), 6.71 – 6.61 (m, 2H, 2-H), 4.17 – 4.06 (m, 1H, 2'-H), 4.05 – 3.95 (m, 2H, 3'-H), 3.78 (s, 6H, DMT-H), 3.38 – 3.27 (m, 2H, 1'-H), 2.44 (d, J = 5.4 Hz, 1H, 2'-OH);

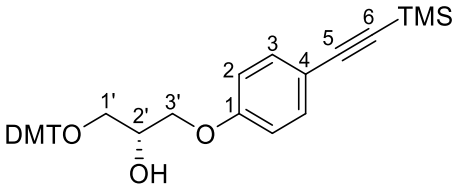
^{13}C{^{1}H} NMR (100 MHz, $CDCl_3$): δ (ppm) = 158.66 (DMT-C), 158.56 (1-C), 144.78 (DMT-C), 138.36 (3-C), 135.90 (DMT-C), 135.88 (DMT-C), 130.15 (DMT-C), 128.18 (DMT-C), 128.03 (DMT-C), 127.02 (DMT-C), 117.10 (2-C), 113.29 (DMT-C), 86.44 (DMT-C), 83.28 (4-C), 69.53 (2'-C), 69.25 (3'-C), 64.00 (1'-C), 55.37 (DMT-C).

HR-MS (ESI+): *m/z* calc. ($C_{30}H_{29}IO_5Na$, $[M+Na]^+$): 619.09519, found: 619.09632.


3 (m) 7.55
24 (m) 7.41
28 (m) 6.81
25 (m) 7.22
2 (m) 6.65
23, 27 (m) 7.29
3' (m) 3.99
30 (s) 3.78
1' (m) 3.33
2'OH (d) 2.44
2' (m) 4.10
δ1H / ppm


δ13C / ppm

Compound **S4.12**

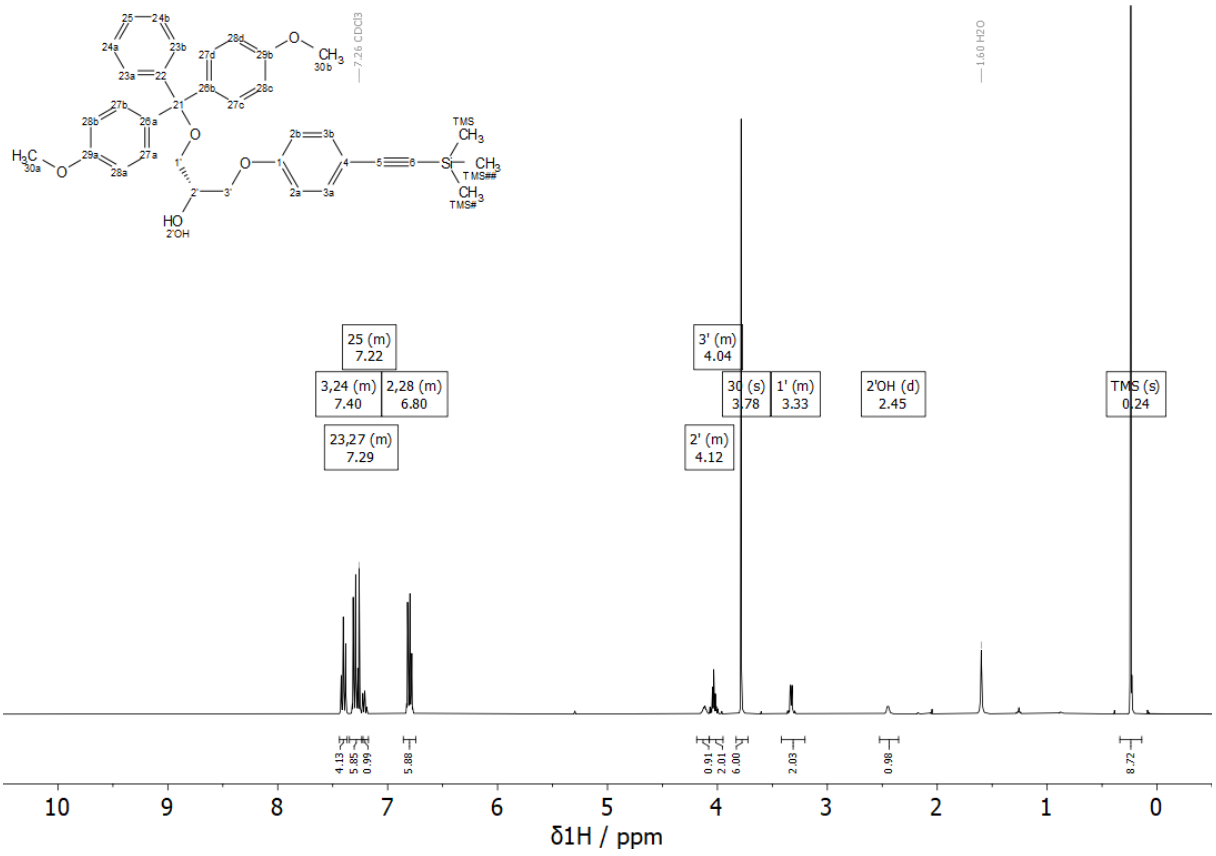


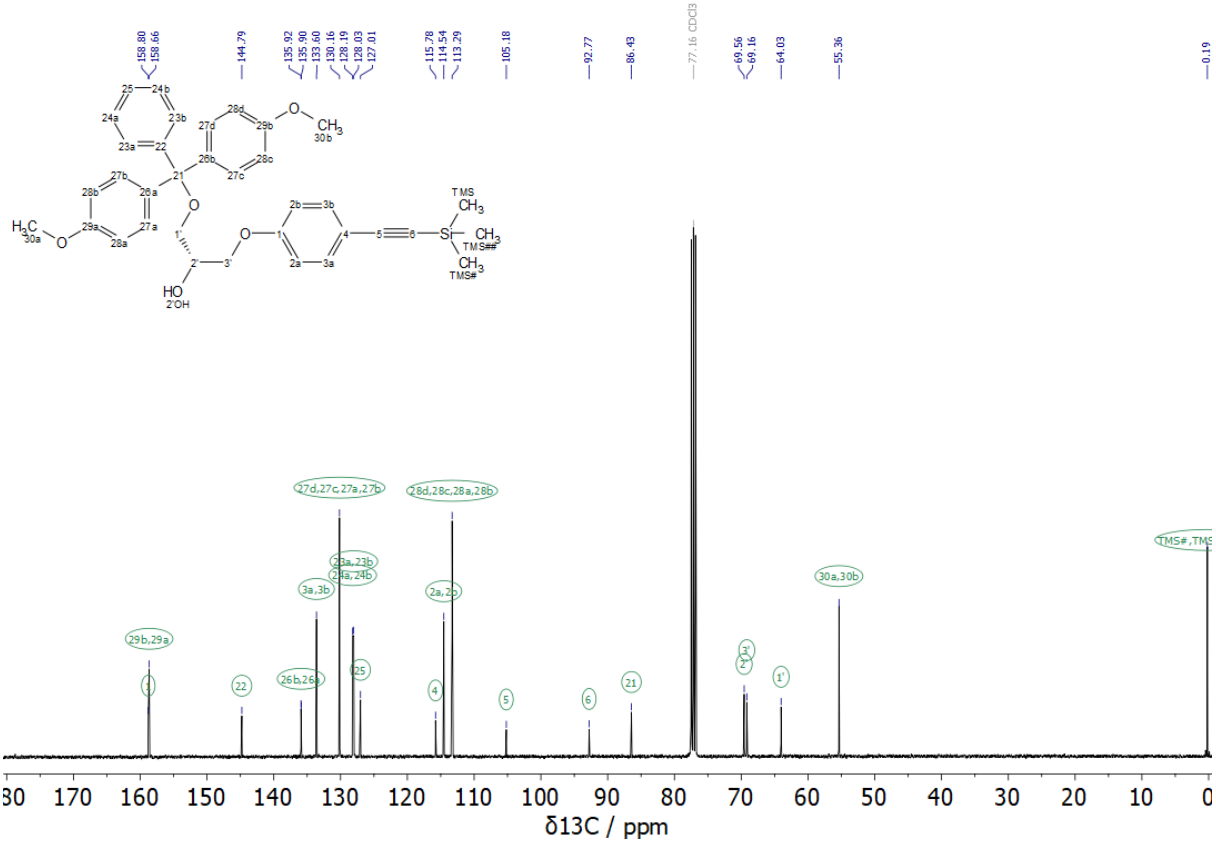
Under nitrogen atmosphere, CuI (31.9 mg, 168 µmol, 10 mol%), $Pd(PPh_3)_2Cl_2$ (58.8 mg, 83.8 µmol, 5.0 mol%) and compound **S4.11** (1.00 g, 1.68 mmol, 1.00 eq.) were dissolved in argon-purged Et_3N (27 mL). Then, trimethylsilylacetylene (348 µL, 247 mg, 2.51 mmol, 1.50 eq) was added and stirred for 16 h at 60 °C. The precipitate was filtered off and washed with EtOAc. The filtrate was washed with a saturated solution of ammonium chloride (30 mL). The aqueous phase was extracted with EtOAc (3x30 mL). The combined organic phases were washed with brine, dried over Na_2SO_4, evaporated and the residue was purified by column chromatography (hexane/EtOAc 4:1 + 1% Et_3N → 2:1 + 1% Et_3N) to afford compound **S4.12** as a colorless foam (943 mg, 1.66 mmol, 99%).

^{1}H NMR (400 MHz, $CDCl_3$): δ (ppm) = 7.44 – 7.37 (m, 4H, 3-H, DMT-H), 7.35 – 7.22 (m, 6H, DMT-H), 7.24 – 7.18 (m, 1H, DMT-H), 6.86 – 6.74 (m, 6H, 2-H, DMT-H), 4.19 – 4.07 (m, 1H, 2'-H), 4.08 – 3.95 (m, 2H, 3'-H), 3.78 (s, 6H, DMT-H), 3.42 – 3.20 (m, 2H, 1'-H), 2.45 (d, J = 5.1 Hz, 1H, 2'-OH), 0.24 (s, 9H, TMS);

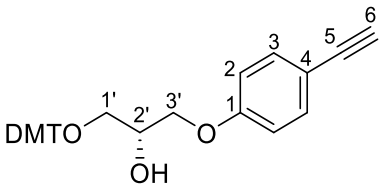
^{13}C{^{1}H} NMR (100 MHz, $CDCl_3$): δ (ppm) = 158.80 (1-C), 158.66 (DMT-C), 144.79 (DMT-C), 135.92 (3-C), 135.90 (DMT-C), 133.60 (DMT-C), 130.16 (DMT-C), 128.19 (DMT-C), 128.03 (DMT-C), 127.01 (DMT-C), 115.78 (4-C), 114.54 (2-C), 113.29 (DMT-C), 105.18 (5-C), 92.77 (6-C), 86.43 (DMT-C), 69.56 (2'-C), 69.16 (3'-C), 64.03 (1'-C), 55.36 (DMT-C), 0.19 (TMS);

HR-MS (ESI+): m/z calc. ($C_{35}H_{38}O_5SiNa$, $[M+Na]^+$): 589.23807, found: 589.23869.


25 (m) 7.22
3,24 (m) 7.40
2,28 (m) 6.80
23,27 (m) 7.29
3' (m) 4.04
30 (s) 3.78
1' (m) 3.33
2'OH (d) 2.45
TMS (s) 0.24
2' (m) 4.12
δ1H / ppm


27d,27c,27a,27b
28d,28c,28a,28b
30a,30b
δ13C / ppm

Compound **S4.13**

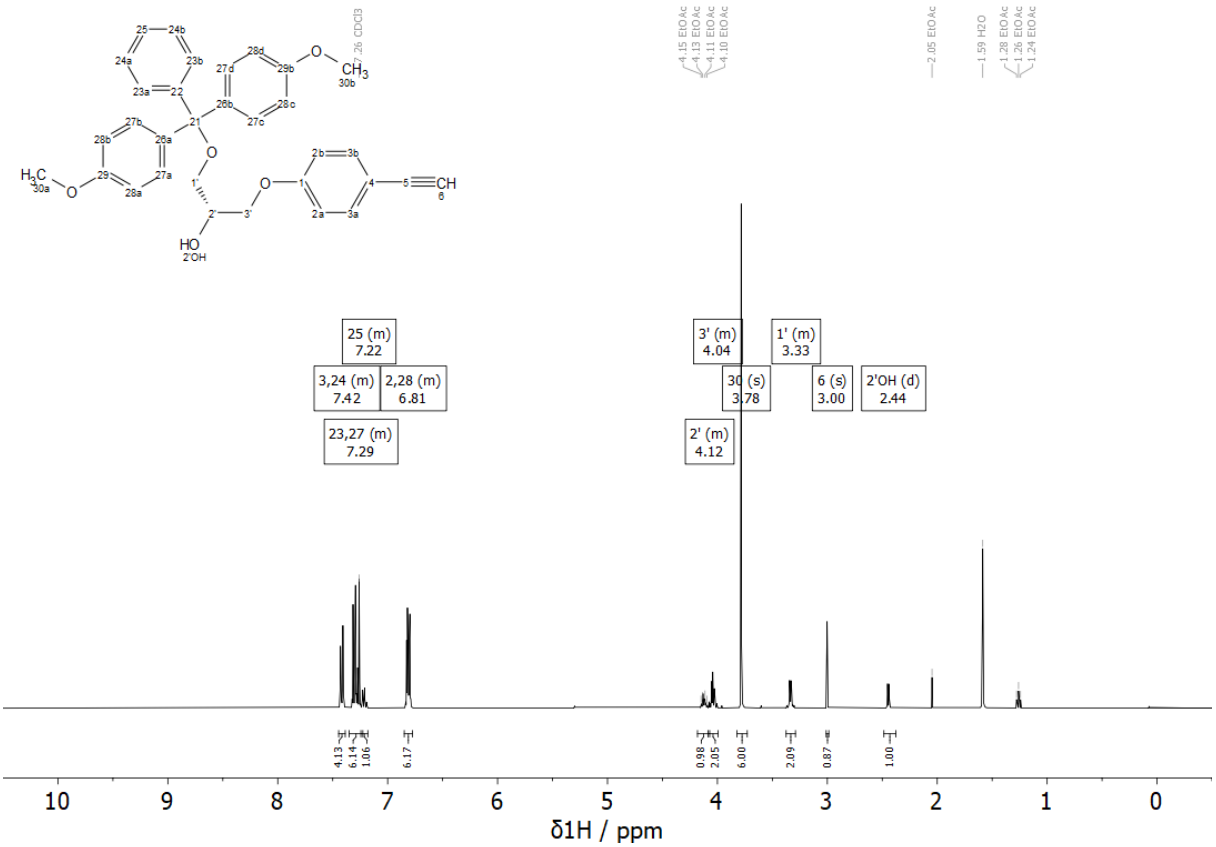


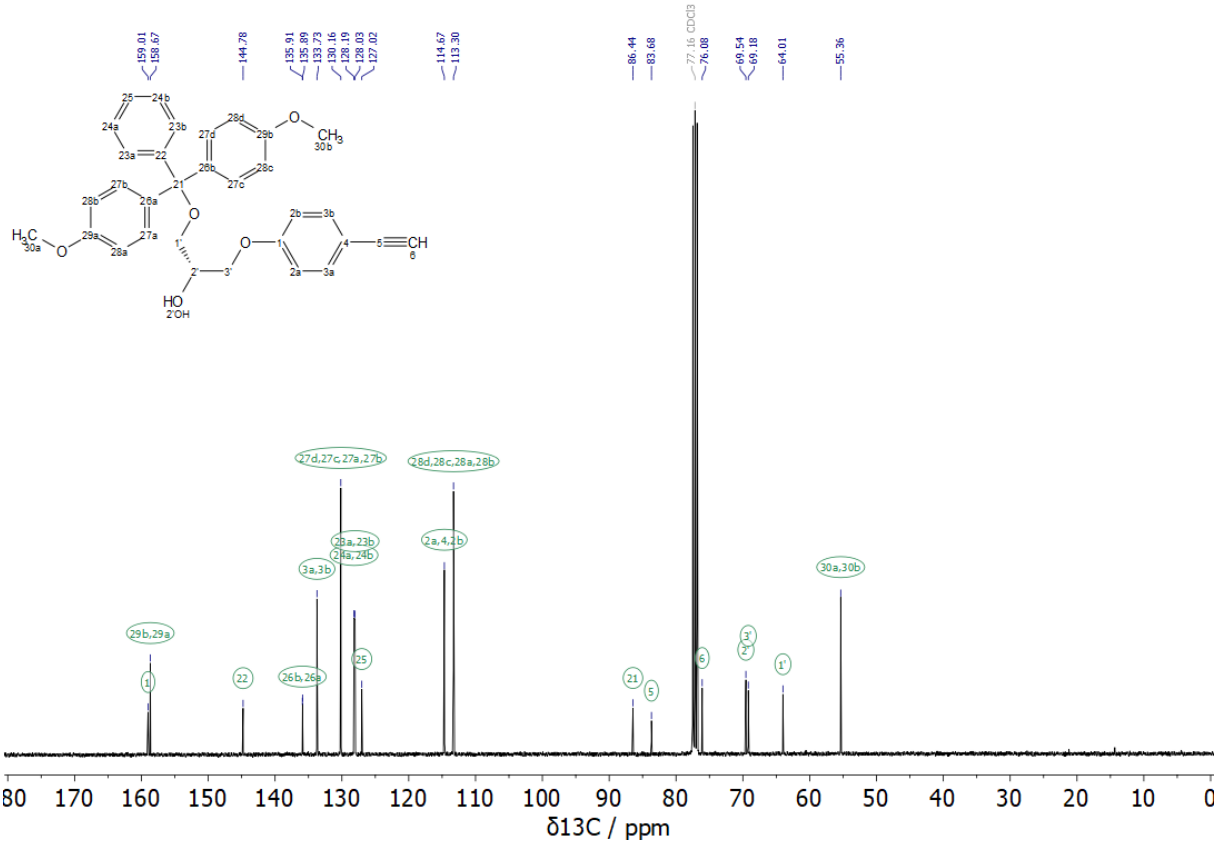
Compound **S4.12** (756 mg, 1.33 mmol, 1.00 eq.) was dissolved in THF/MeOH (1:1, 22 mL). K_2CO_3 (184 mg, 1.33 mmol, 1.00 eq.) was added and the mixture was stirred for 2 h at room temperature. The suspension was diluted with water (40 mL), extracted with EtOAc and washed with brine. The organic phase was dried over Na_2SO_4, evaporated and the residue was purified by column chromatography (hexane/EtOAc 3:1 + 1% Et_3N → 2:1 + 1% Et_3N) to afford compound **S4.13** as a colorless foam (580 mg, 1.17 mmol, 88%).

^{1}H NMR (400 MHz, $CDCl_3$): δ (ppm) = 7.45 – 7.39 (m, 4H), 7.35 – 7.23 (m, 6H, 3-H, DMT-H), 7.24 – 7.18 (m, 1H, DMT-H), 6.85 – 6.77 (m, 6H, 2-H, DMT-H), 4.18 – 4.07 (m, 1H, 2'-H), 4.08 – 3.99 (m, 2H, 3'-H), 3.78 (s, 6H, DMT-H), 3.37 – 3.29 (m, 2H, 1'-H), 3.00 (s, 1H, 6-H), 2.44 (d, *J* = 5.4 Hz, 1H, 2'-OH);

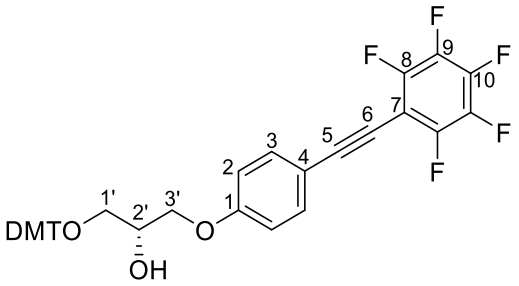
^{13}C{^{1}H} NMR (100 MHz, $CDCl_3$): δ (ppm) = 159.01 (1-C), 158.67 (DMT-C), 144.78 (DMT-C), 135.91 (DMT-C), 135.89 (DMT-C), 133.73 (3-C), 130.16 (DMT-C), 128.19 (DMT-C), 128.03 (DMT-C), 127.02 (DMT-C), 114.67 (2-C, 4-C), 113.30 (DMT-C), 86.44 (DMT-C), 83.68 (5-C), 76.08 (6-C), 69.54 (2'-C), 69.18 (3'-C), 64.01 (1'-C), 55.36 (DMT-C);

HR-MS (ESI+): *m/z* calc. ($C_{32}H_{30}O_5Na$, $[M+Na]^+$): 517.19854, found: 517.19917.


25 (m) 7.22
3,24 (m) 7.42
2,28 (m) 6.81
23,27 (m) 7.29
3' (m) 4.04
1' (m) 3.33
30 (s) 3.78
6 (s) 3.00
2'OH (d) 2.44
2' (m) 4.12
δ1H / ppm


δ13C / ppm

Compound **S4.14**



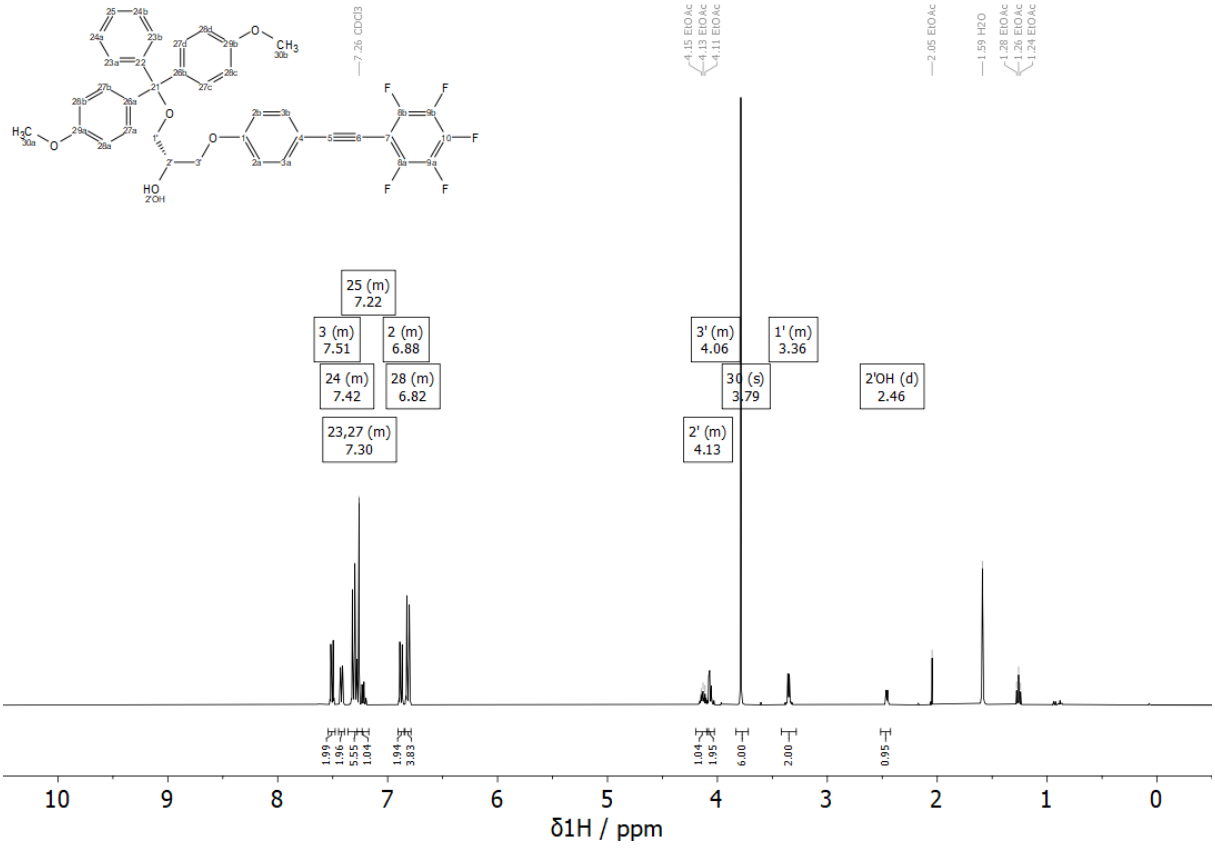
Under nitrogen atmosphere, CuI (17.6 mg, 92.4 µmol, 10 mol%), $Pd(PPh_3)_2Cl_2$ (32.0 mg, 45.6 µmol, 5.0 mol%), PPh_3 (12.1 mg, 46.1 µmol, 5.0 mol%) and iodopentafluorobenzene (185 µL, 407 mg, 1.38 mmol, 1.50 eq.) were dissolved in argon-purged Et_3N (8 mL). A solution of compound **S4.13** (457 mg, 924 µmol, 1.00 eq.) in argon-purged Et_3N (5 mL) was added and stirring was continued for 18 h at 60°C. The precipitate was filtered off and washed with EtOAc. The filtrate was washed with a saturated solution of ammonium chloride (30 mL). The aqueous phase was extracted with EtOAc. The combined organic phases were washed with brine (30mL), dried over Na_2SO_4, evaporated and the residue was purified by column chromatography (hexane/EtOAc 3:1 + 1% Et_3N) to afford compound **S4.14** as a colorless foam (423 mg, 640 µmol, 69%).

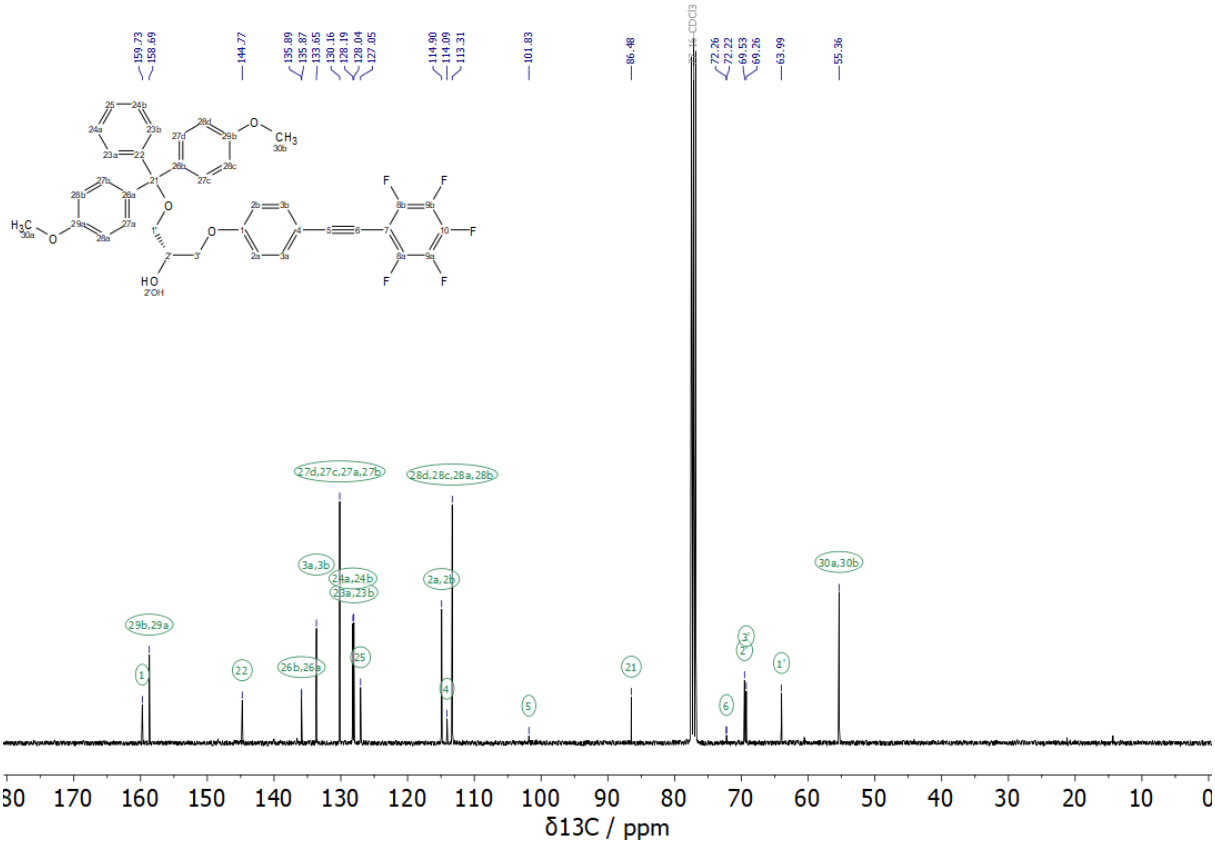
^{1}H NMR (400 MHz, $CDCl_3$): δ (ppm) = 7.54 – 7.48 (m, 2H, 3-H), 7.45 – 7.39 (m, 2H, DMT-H), 7.36 – 7.23 (m, 6H, DMT-H), 7.28 – 7.17 (m, 1H, DMT-H), 6.91 – 6.85 (m, 2H, 2-H), 6.84 – 6.79 (m, 4H, DMT-H), 4.20 – 4.08 (m, 1H, 2'-H), 4.10 – 4.03 (m, 2H, 3'-H), 3.79 (s, 6H, DMT-H), 3.42 – 3.28 (m, 2H, 1'-H), 2.46 (d, J = 5.3 Hz, 1H, 2'-OH);

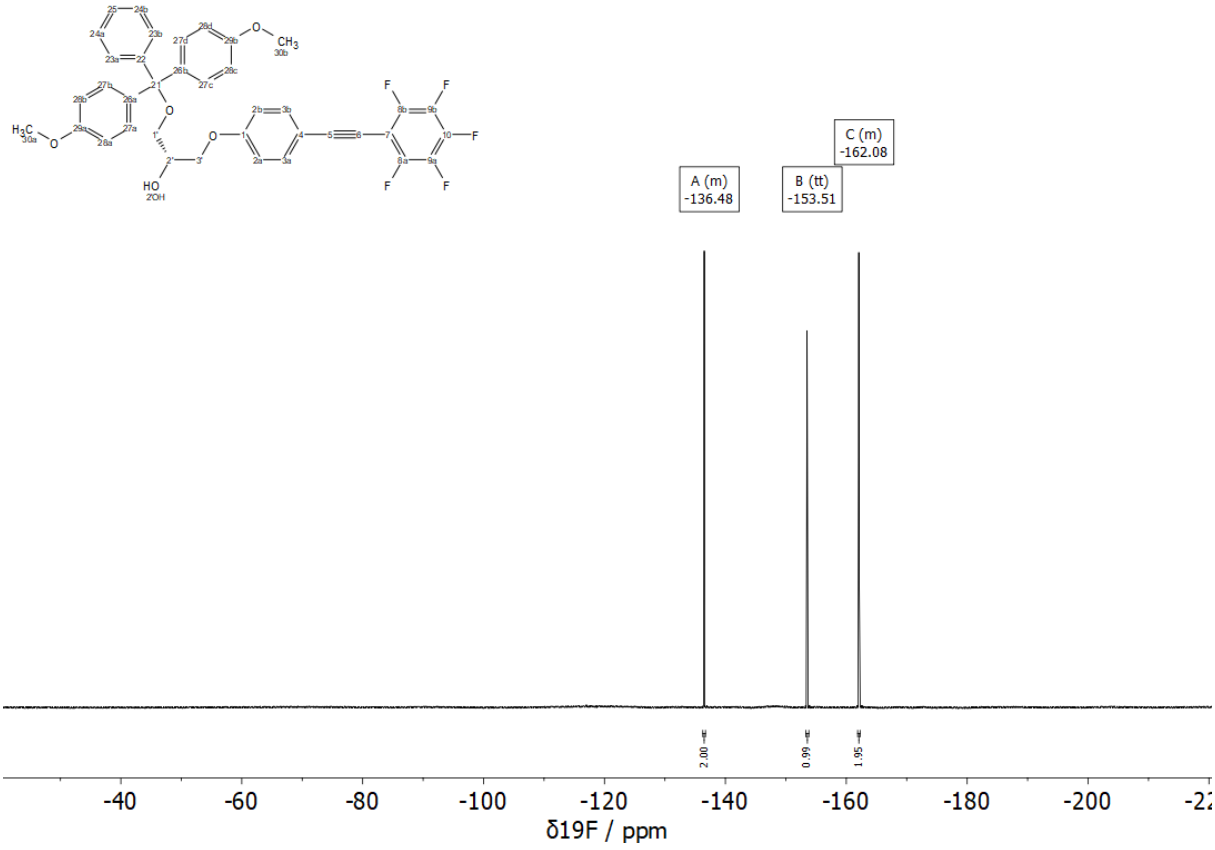
^{13}C{^{1}H} NMR (100 MHz, $CDCl_3$): δ (ppm) = 159.73 (1-C), 158.69 (DMT-C), 144.77 (DMT-C), 135.89 (DMT-C), 135.87 (DMT-C), 133.65 (3-C), 130.16 (DMT-C), 128.19 (DMT-C), 128.04 (DMT-C), 127.05 (DMT-C), 114.90 (2-C), 114.09 (4-C), 113.31 (DMT-C), 101.83 (5-C), 86.48 (DMT-C), 72.26 (6-C), 72.22 (6-C), 69.53 (2'-C), 69.26 (3'-C), 63.99 (1'-C), 55.36 (DMT-C);

^{19}F{^{1}H} NMR (376 MHz, $CDCl_3$): δ (ppm) = -136.27 – -136.71 (m, 2F), -153.51 (tt, J = 20.7, 1.2 Hz, 1F), -161.88 – -162.29 (m, 2F);

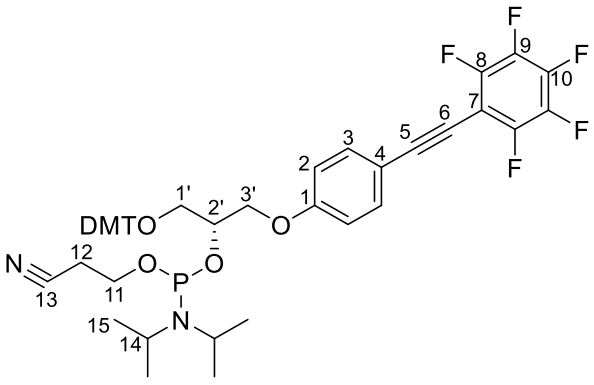
HR-MS (ESI+): m/z calc. ($C_{38}H_{29}F_5O_5Na$, $[M+Na]^+$): 683.18328, found: 683.18494.


25 (m)
7.22
3 (m)
7.51
2 (m)
6.88
24 (m)
7.42
28 (m)
6.82
23,27 (m)
7.30
3' (m)
4.06
1' (m)
3.36
30 (s)
3.79
2'OH (d)
2.46
2' (m)
4.13
δ1H / ppm


27d,27c,27a,27b
28d,28c,28a,28b
30a,30b
δ13C / ppm



Compound **S4.15** = GTHF-PA



Compound **S4.14** (120 mg, 182 µmol, 1.00 eq.) was dissolved under nitrogen atmosphere with DIPEA (190 µL, 141 mg 1.10 mmol, 6.00 eq.) in anhydrous DCM (3 mL). After 10 min CEP-Cl (52.7 µL, 55.9 mg, 236 µmol, 1.30 eq.) was added. Additional CEP-Cl (5.3 µL, 5.6 mg, 22.4 µmol, 0.12 eq.) was added after 2 h stirring at room temperature. The reaction mixture was stirred additionally at ambient temperature for 1 h. The solvent was removed under reduced pressure and the residue was purified by column chromatography (hexane/EtOAc 4:1 + 1% Et_3N) to afford compound **S4.15** as a colorless foam (136 mg, 158 µmol, 87%).

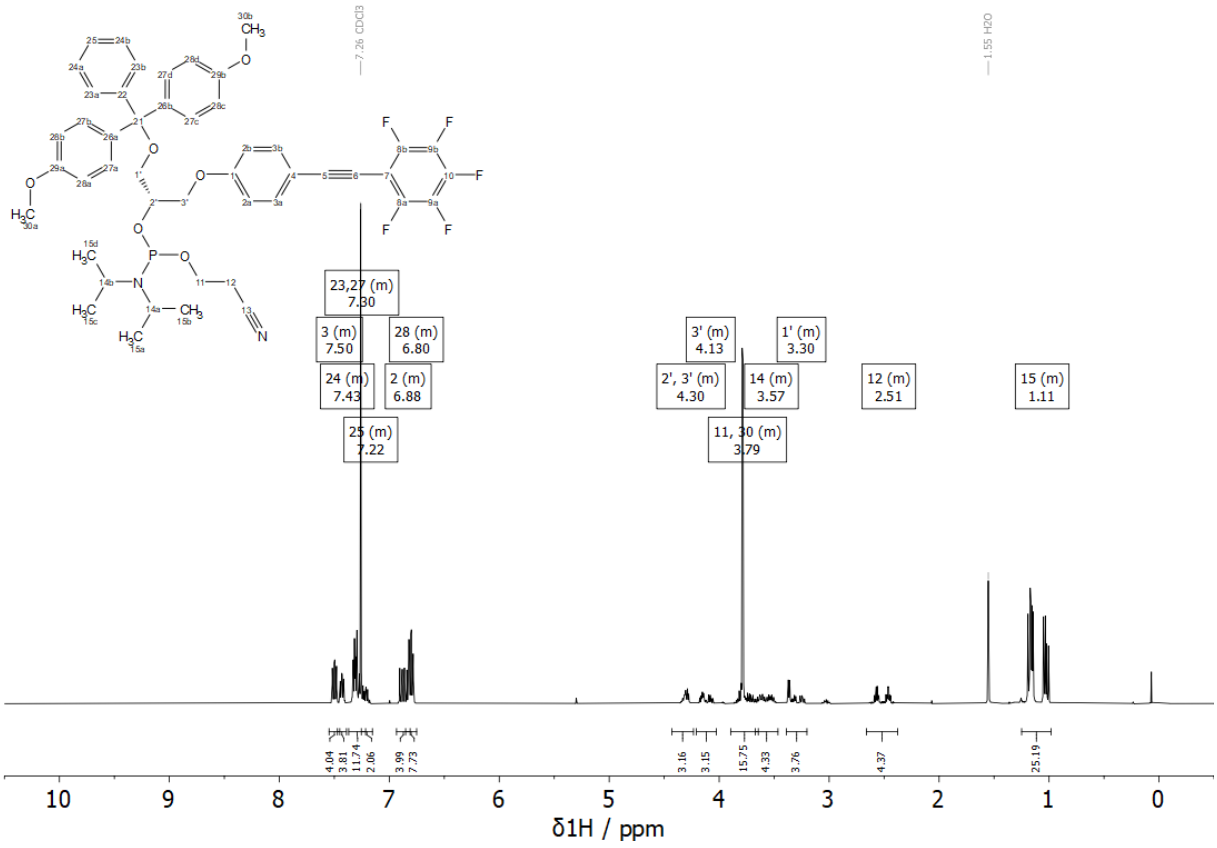
^{1}H NMR (400 MHz, $CDCl_3$): δ (ppm) = 7.55 – 7.45 (m, 4H, 3-H), 7.47 – 7.39 (m, 4H, DMT-H), 7.37 – 7.22 (m, 12H, DMT-H), 7.25 – 7.15 (m, 2H, DMT-H), 6.93 – 6.81 (m, 4H, 2-H), 6.85 – 6.75 (m, 8H, DMT-H), 4.43 – 4.23 (m, 3H, 2'-H, 3'-H), 4.21 – 4.03 (m, 3H, 3'-H), 3.90 – 3.64 (m, 16H, 11-H), 3.67 – 3.47 (m, 4H, 14-H), 3.39 – 3.20 (m, 4H, 1'-H), 2.66 – 2.38 (m, 4H, 12-H), 1.25 – 0.98 (m, 25H, 15-H);

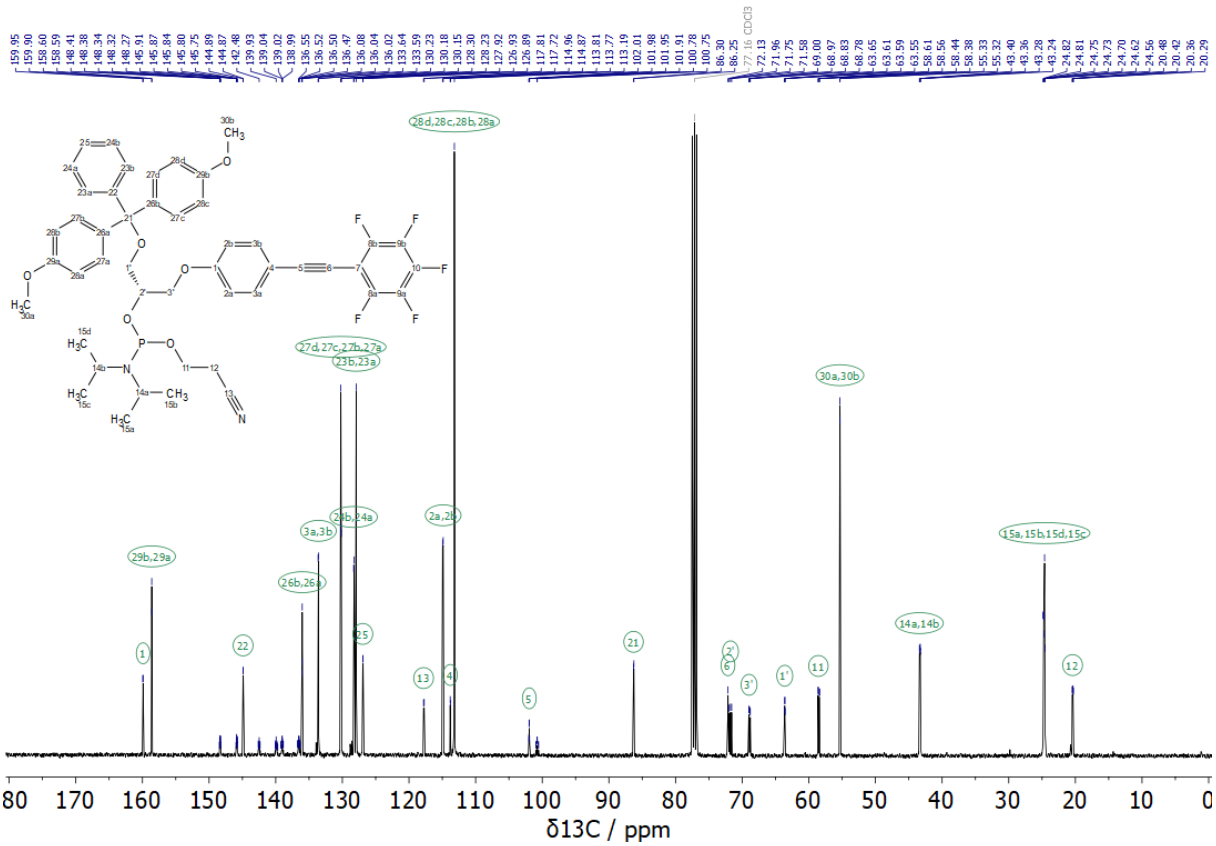
^{13}C{^{1}H} NMR (100 MHz, $CDCl_3$): δ (ppm) = 159.95 (1-C), 159.90 (1-C), 158.60 (DMT-C), 158.59 (DMT-C), 148.58 – 148.02 (m), 146.06 – 145.56 (m), 144.89 (DMT-C), 144.87 (DMT-C), 142.70 – 142.18 (m), 140.22 – 139.65 (m), 139.24 – 138.68 (m), 136.92 – 136.21 (m), 136.08 (DMT-C), 136.04 (DMT-C), 136.02 (DMT-C), 133.64 (3-C), 133.59 (3-C), 130.23 (DMT-C), 130.18 (DMT-C), 130.15 (DMT-C), 128.30 (DMT-C), 128.23 (DMT-C), 127.92 (DMT-C), 126.93 (DMT-C), 126.89 (DMT-C), 117.81 (13-C), 117.72 (13-C), 114.96 (2-C), 114.87 (2-C), 113.81 (4-C), 113.77 (4-C), 113.19 (DMT-C), 102.01 (5-C), 101.98 (5-C), 101.95 (5-C), 101.91 (5-C), 100.96, 100.93, 100.78, 100.75, 100.60, 100.56, 86.30 (DMT-C), 86.25 (DMT-C), 72.13 (6-C), 71.96 (2'-C), 71.75 (2'-C), 71.58 (2'-C), 69.00 (3'-C), 68.97 (3'-C), 68.83 (3'-C), 68.78 (3'-C), 63.65 (1'-C), 63.61 (1'-C), 63.59 (1'-C), 63.55 (1'-C), 58.61 (11-C), 58.56 (11-C), 58.44 (11-C), 58.38 (11-C), 55.33 (DMT-C), 55.32 (DMT-C), 43.40 (14-C), 43.36 (14-C), 43.28 (14-C), 43.24 (14-C), 24.82 (15-C), 24.81 (15-C), 24.75 (15-C), 24.73 (15-C), 24.70 (15-C), 24.62 (15-C), 24.56 (15-C), 20.48 (12-C), 20.42 (12-C), 20.36 (12-C), 20.29 (12-C);

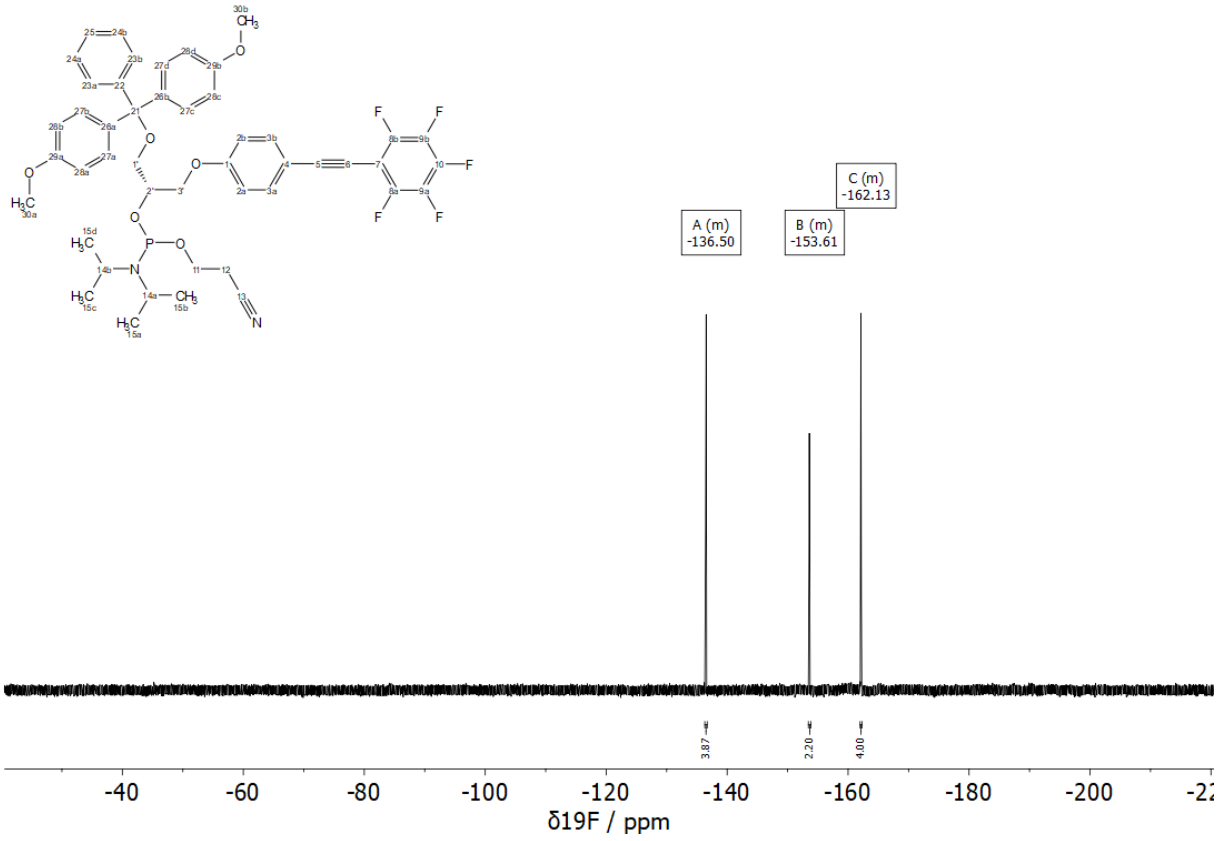
^{19}F{^{1}H} NMR (376 MHz, $CDCl_3$): δ (ppm) = -136.31 – -136.70 (m, 4F), -153.43 – -153.86 (m, 2F), -161.99 – -162.30 (m, 4F);

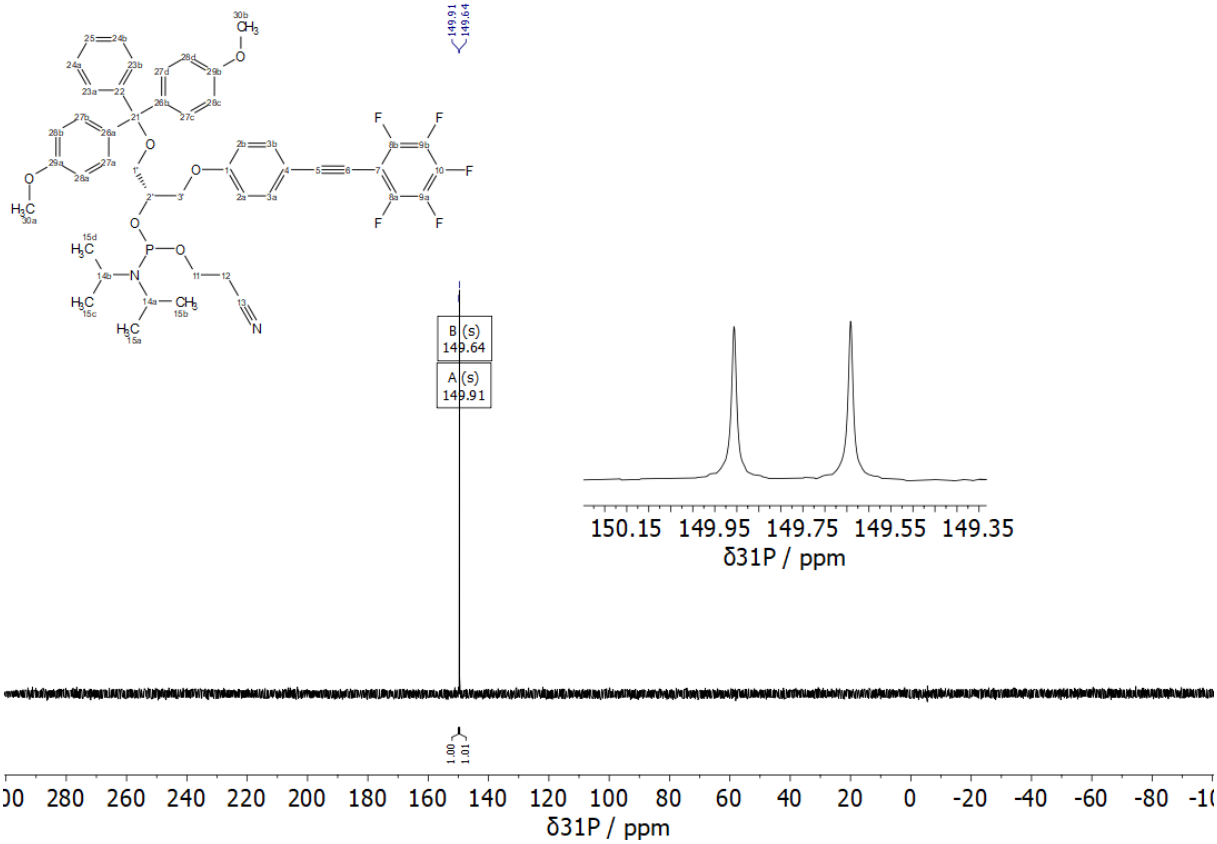
^{31}P{^{1}H} NMR (162 MHz, $CDCl_3$): δ (ppm) = 149.91, 149.64;

HR-MS (ESI+): *m/z* calc. ($C_{47}H_{46}F_5N_2O_6PNa$, $[M+Na]^+$): 883.29059, found: 883.29040.

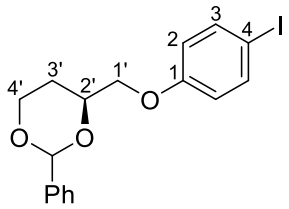

7.26 CDCl3
1.55 H2O
23,27 (m) 7.30
3 (m) 7.50
28 (m) 6.80
24 (m) 7.43
2 (m) 6.88
25 (m) 7.22
3' (m) 4.13
1' (m) 3.30
2', 3' (m) 4.30
14 (m) 3.57
12 (m) 2.51
15 (m) 1.11
11, 30 (m) 3.79
4.04
3.81
11.74
2.06
3.99
7.73
3.16
3.15
15.75
4.33
3.76
4.37
25.19
10
9
8
7
6
5
4
3
2
1
0
δ1H / ppm


77.16 CDCl3
28d,28c,28b,28a
27d,27c,27b,27a
23b,23a
30a,30b
26b,26a
3a,3b
2a,2b
29b,29a
15a,15b,15d,15c
14a,14b
80
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
δ13C / ppm


A (m)
-136.50
B (m)
-153.61
C (m)
-162.13
3.87
2.20
4.00
-40
-60
-80
-100
-120
-140
-160
-180
-200
δ19F / ppm


149.91
149.64
B (s)
149.64
A (s)
149.91
150.15
149.95
149.75
149.55
149.35
δ31P / ppm
1.00
1.01
δ31P / ppm

Compound **S4.16**

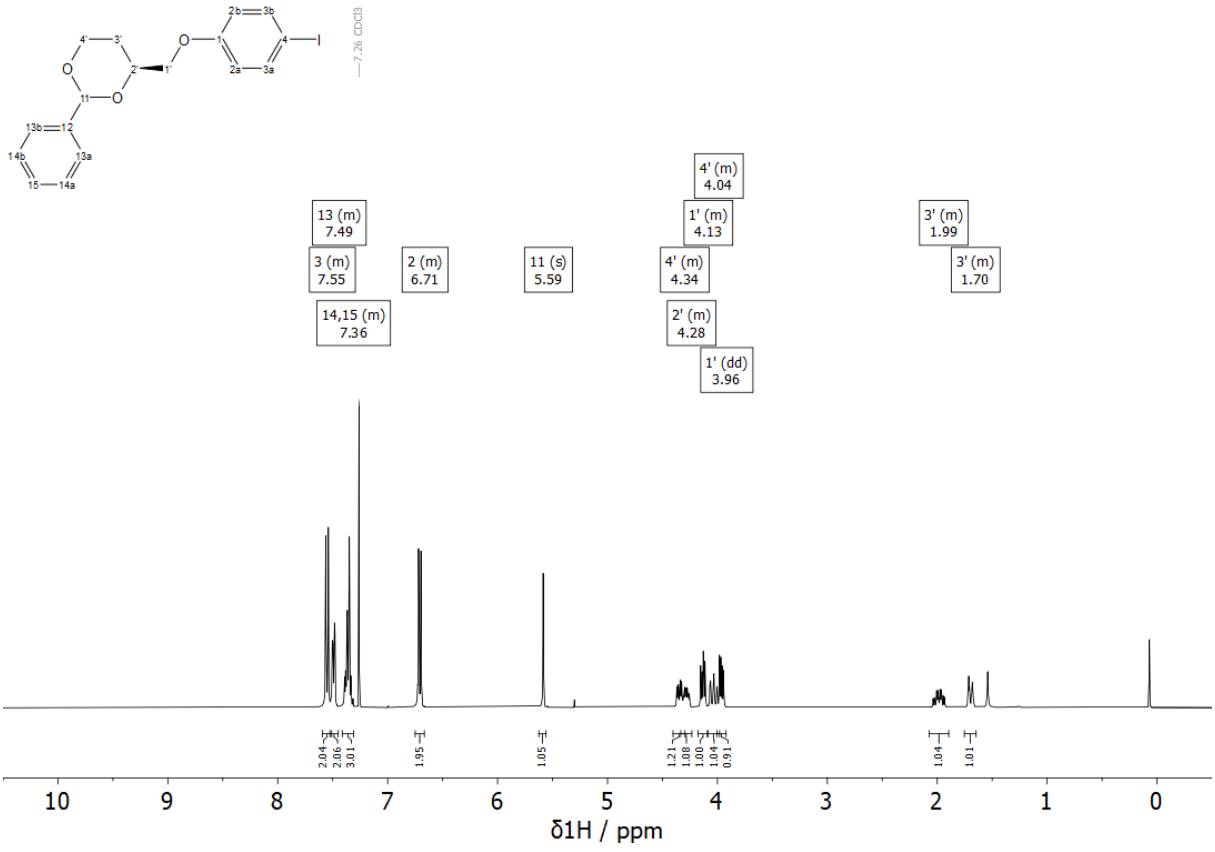


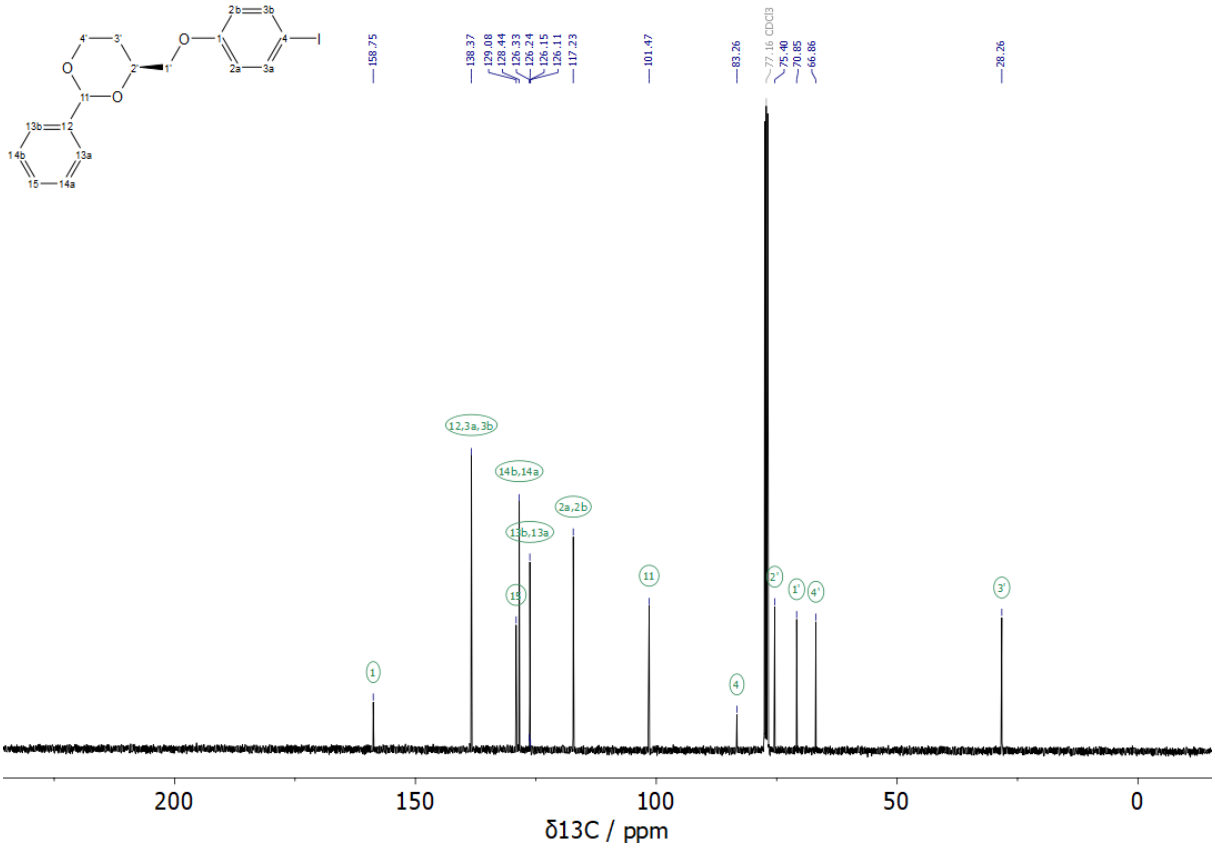
General procedure was adapted from He *et al.*[416] 4-Iodophenol (500 mg, 2.27 mmol, 1.00 eq.), 4-(*S*)-hydroxmethyl-2-phenyl-1,3-dioxan (**4.1**) (573 mg, 2.95 mmol, 1.30 eq.) and triphenylphosphine (626 mg, 2.39 mmol, 1.05 eq.) were dissolved in THF (0.8 mL). Diisopropyl azodicarboxylate (468 µL, 483 mg, 2.39 mmol, 1.05 eq.) was added dropwise and the reaction mixture was sonicated for 30 min. Purification of the residue by column chromatography (hexane/DCM 3:2→1:1) afforded compound **S4.16** as a colorless solid (669 mg, 1.69 mmol, 74%).

^{1}H NMR (400 MHz, $CDCl_3$): δ (ppm) = 7.59 – 7.51 (m, 2H, 3-H), 7.53 – 7.45 (m, 2H, Ph-H), 7.41 – 7.31 (m, 3H, Ph-H), 6.75 – 6.67 (m, 2H, 2-H), 5.59 (s, 1H, C*H*-Ph), 4.40 – 4.30 (m, 1H, 4'-H), 4.33 – 4.23 (m, 1H, 2'-H), 4.18 – 4.09 (m, 1H, 1'H), 4.09 – 3.98 (m, 1H, 4'-H), 3.96 (dd, *J* = 9.8, 5.1 Hz, 1H, 1'-H), 2.07 – 1.90 (m, 1H, 3'-H), 1.75 – 1.65 (m, 1H, 3'-H);

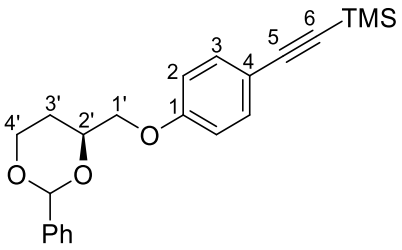
^{13}C{^{1}H} NMR (100 MHz, $CDCl_3$): δ (ppm) = 158.75 (1-C), 138.37 (3-C, 12-C), 129.08 (15-C), 128.44 (14-C), 126.24 (13-C), 117.23 (2-C), 101.47 (*C*H-Ph), 83.26 (4-C), 75.40 (2'-C), 70.85 (1'-C), 66.86 (4'-C), 28.26 (3'-C);

HR-MS (ESI+): *m/z* calc. ($C_{17}H_{17}IO_3Na$, $[M+Na]^+$): 419.01146, found: 419.01094.


7.26 CDCl3
13 (m)
7.49
3 (m)
7.55
14,15 (m)
7.36
2 (m)
6.71
11 (s)
5.59
4' (m)
4.04
1' (m)
4.13
4' (m)
4.34
2' (m)
4.28
1' (dd)
3.96
3' (m)
1.99
3' (m)
1.70
2.04
2.06
3.01
1.95
1.05
1.21
1.08
1.00
1.04
0.91
1.04
1.01
10
9
8
7
6
5
4
3
2
1
0
δ1H / ppm


158.75
138.37
129.08
128.44
126.33
126.24
126.15
126.11
117.23
101.47
83.26
77.16 CDCl3
75.40
70.85
66.86
28.26
12,3a,3b
14b,14a
2a,2b
13b,13a
15
11
1
4
2'
1'
4'
3'
200
150
100
50
0
δ13C / ppm

Compound **S4.17**

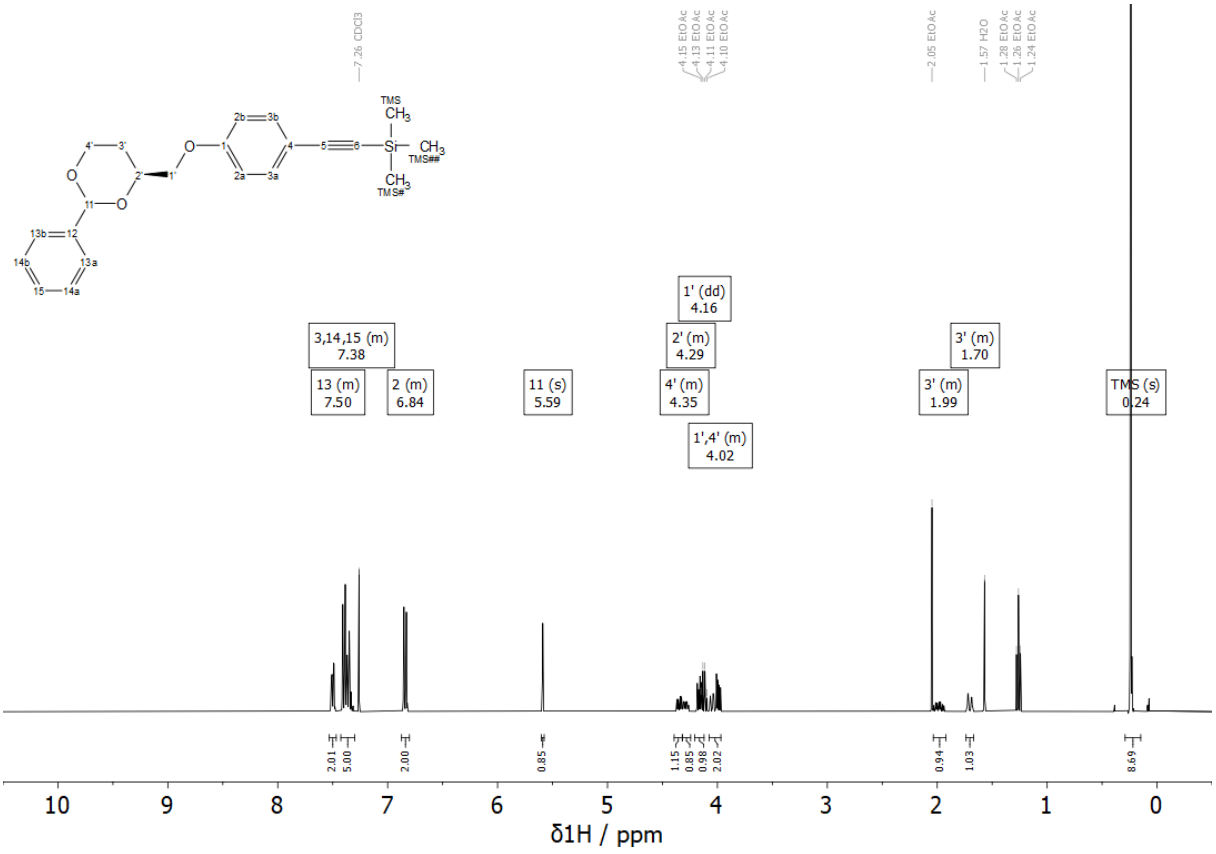


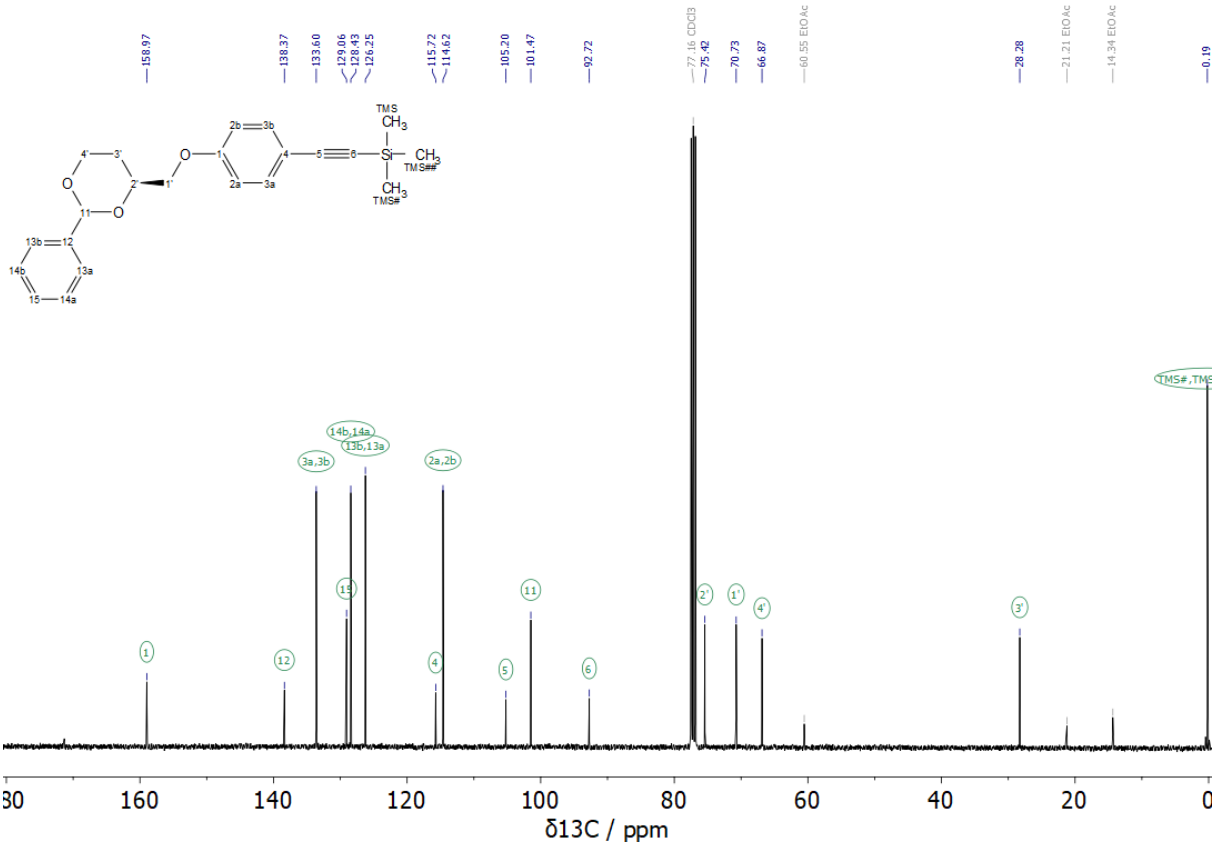
Under nitrogen atmosphere, CuI (28.8 mg, 151 µmol, 10 mol%), $Pd(PPh_3)_4$ (52.5 mg, 45.4 µmol, 3.0 mol%) and compound **S4.16** (600 mg, 1.51 mmol, 1.00 eq.) were dissolved in argon-purged Et_3N (20 mL). Then, trimethylsilylacetylene (293 µL, 208 mg, 2.12 mmol, 1.40 eq) was added and stirred for 15 h at 80 °C. The precipitate was filtered off and washed with Et_2O. The filtrate was washed with a saturated solution of ammonium chloride (2x25 mL) and brine (25 mL). The aqueous phase was extracted with Et_2O (2x25 mL). The combined organic phases were dried over Na_2SO_4, evaporated and the residue was purified by column chromatography (hexane/EtOAc 20:1) to afford compound **S4.17** as a colorless oil (520 mg, 1.42 mmol, 94%).

^{1}H NMR (400 MHz, $CDCl_3$): δ (ppm) = 7.53 – 7.47 (m, 2H, Ph-H), 7.43 – 7.30 (m, 5H, 3-H, Ph-H), 6.87 – 6.80 (m, 2H, 2-H), 5.59 (s, 1H, C*H*-Ph), 4.39 – 4.32 (m, 1H, 4'-H), 4.32 – 4.24 (m, 1H, 2'-H), 4.16 (dd, *J* = 9.8, 5.7 Hz, 1H, 1'-H), 4.07 – 3.97 (m, 2H, 4'-H, 1'-H), 2.04 – 1.92 (m, 1H, 3'-H), 1.74 – 1.67 (m, 1H, 3'-H), 0.24 (s, 9H, TMS);

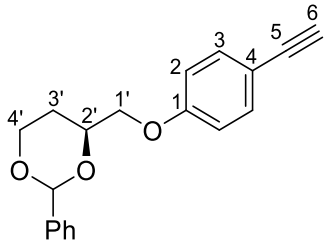
^{13}C{^{1}H} NMR (100 MHz, $CDCl_3$): δ (ppm) = 158.97 (1-C), 138.37 (12-C), 133.60 (3-C), 129.06 (15-C), 128.43 (14-C), 126.25 (13-C), 115.72 (4-C), 114.62 (2-C), 105.20 (5-C), 101.47 (*C*H-Ph), 92.72 (6-C), 75.42 (2'-C), 70.73 (1'-C), 66.87 (4'-C), 28.28 (3'-C), 0.19 (TMS);

HR-MS (ESI+): *m/z* calc. ($C_{22}H_{26}O_3SiNa$, $[M+Na]^+$): 389.15434, found: 389.15490.


7.26 CDCl3
4.15 EtOAc
4.13 EtOAc
4.11 EtOAc
4.10 EtOAc
2.05 EtOAc
1.57 H2O
1.28 EtOAc
1.26 EtOAc
1.24 EtOAc
TMS
CH3
Si
CH3
TMS##
CH3
TMS#
3,14,15 (m)
7.38
13 (m)
7.50
2 (m)
6.84
11 (s)
5.59
1' (dd)
4.16
2' (m)
4.29
4' (m)
4.35
1',4' (m)
4.02
3' (m)
1.70
3' (m)
1.99
TMS (s)
0.24
2.01
5.00
2.00
0.85
1.15
0.85
0.98
2.02
0.94
1.03
8.69
δ1H / ppm


158.97
138.37
133.60
129.06
128.43
126.25
115.72
114.62
105.20
101.47
92.72
77.16 CDCl3
75.42
70.73
66.87
60.55 EtOAc
28.28
21.21 EtOAc
14.34 EtOAc
0.19
TMS
CH3
Si
CH3
TMS##
CH3
TMS#
14b,14a
13b,13a
3a,3b
2a,2b
TMS#,TMS
δ13C / ppm

Compound **S4.18**

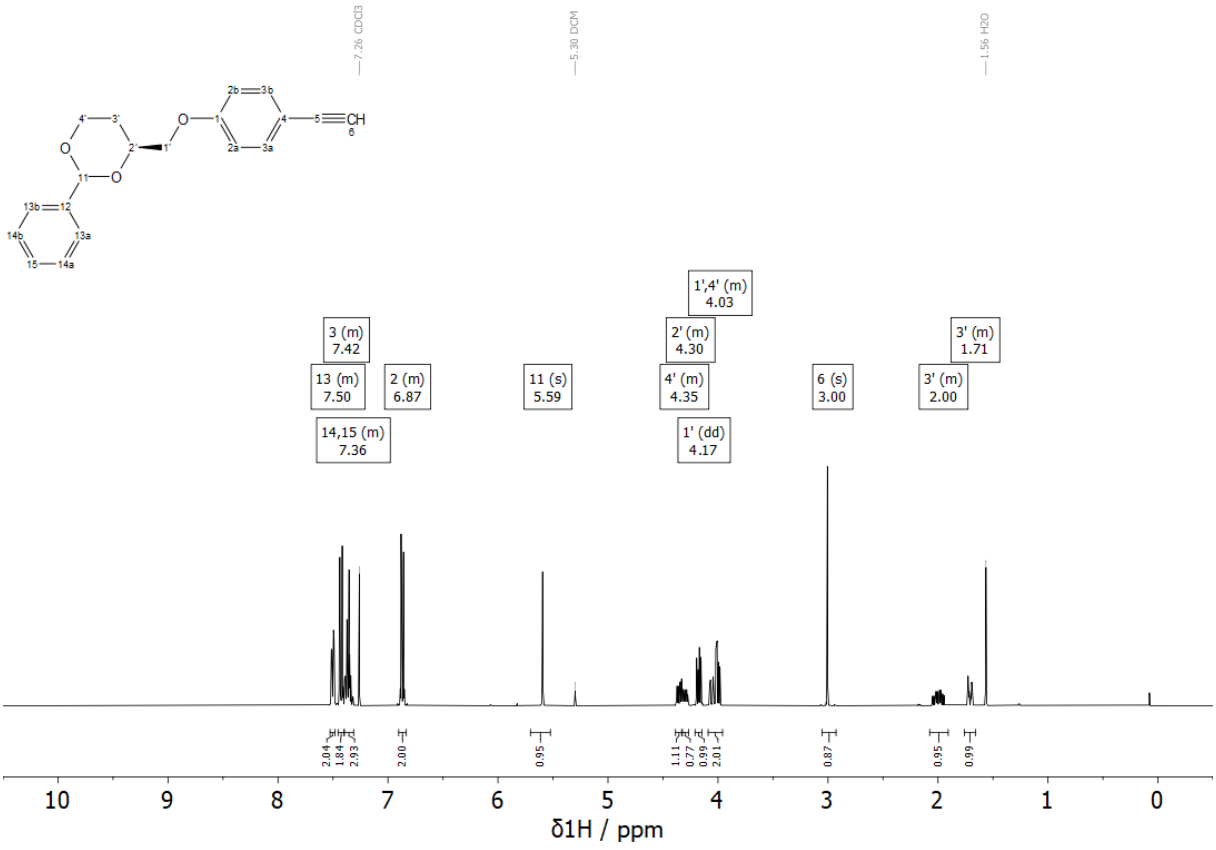


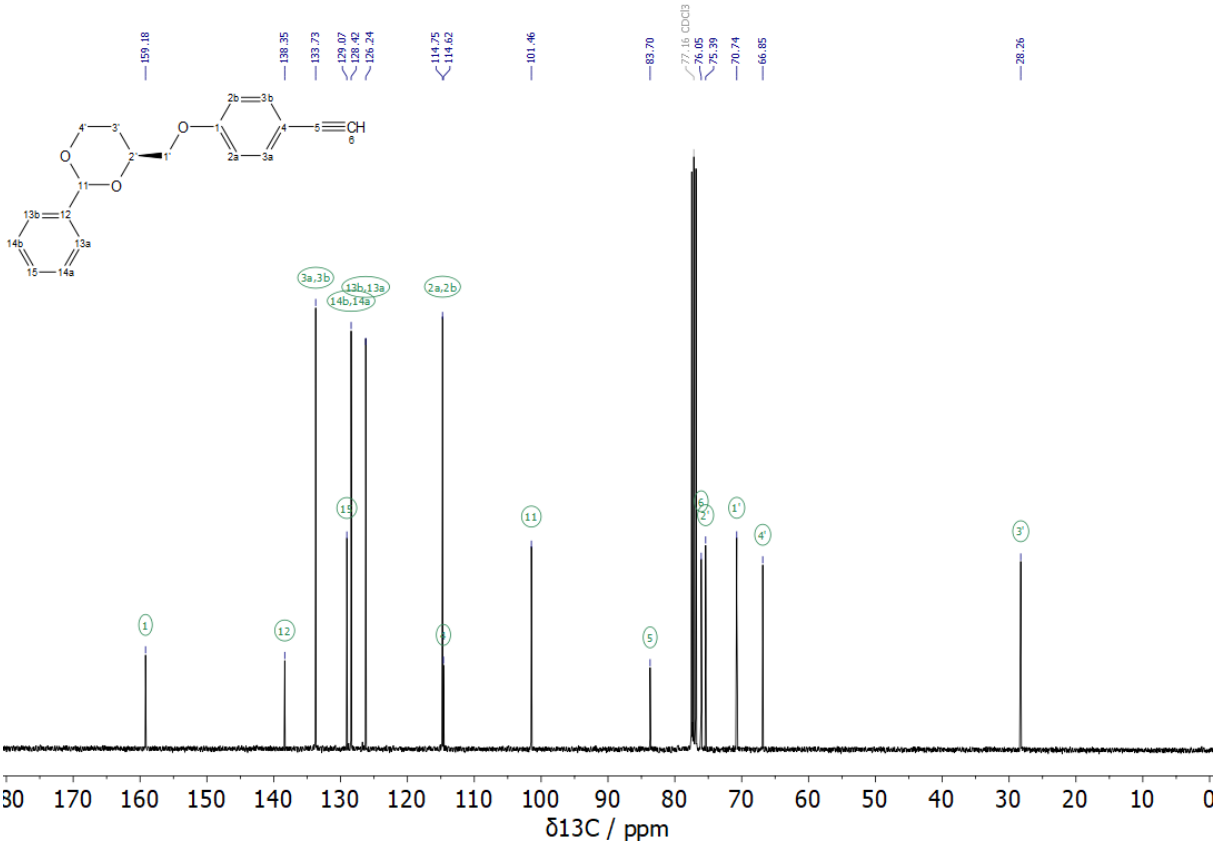
Compound **S4.17** (500 mg, 1.36 mmol, 1.00 eq.) was dissolved in THF/MeOH (1:1, 10 mL). K_2CO_3 (377 mg, 2.73 mmol, 2.00 eq.) was added and the mixture was stirred for 2 h at room temperature. The suspension was diluted with water (20 mL), extracted with DCM (4x25 mL). The organic phase was dried over Na_2SO_4 and evaporated to afford compound **S4.18** as a colorless solid (397 mg, 1.35 mmol, 99%).

^{1}H NMR (400 MHz, $CDCl_3$): δ (ppm) = 7.53 – 7.48 (m, 2H, Ph-H), 7.45 – 7.40 (m, 2H, 3-H), 7.40 – 7.31 (m, 3H, Ph-H), 6.90 – 6.83 (m, 2H, 2-H), 5.59 (s, 1H, C*H*-Ph), 4.39 – 4.32 (m, 1H, 4'-H), 4.33 – 4.27 (m, 1H, 2'-H), 4.17 (dd, J = 9.9, 5.7 Hz, 1H, 1'-H), 4.09 – 3.96 (m, 2H, 1'-H, 4'-H), 3.00 (s, 1H, 6-H), 2.08 – 1.91 (m, 1H, 3'-H), 1.76 – 1.66 (m, 1H, 3'-H);

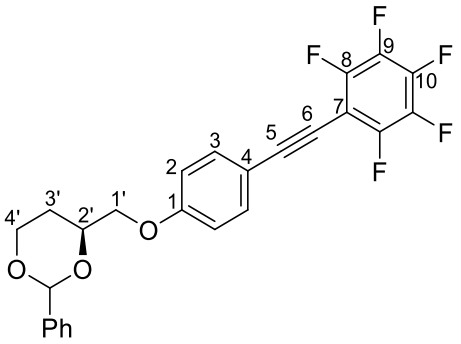
^{13}C{^{1}H} NMR (100 MHz, $CDCl_3$): δ (ppm) = 159.18 (1-C), 138.35 (12-C), 133.73 (3-C), 129.07 (15-C), 128.42 (14-C), 126.24 (13-C), 114.75 (2-C), 114.62 (4-C), 101.46 (*C*H-Ph), 83.70 (5-C), 76.05 (6-C), 75.39 (2'-C), 70.74 (1'-C), 66.85 (4'-C), 28.26 (3'-C);

HR-MS (ESI+): *m/z* calc. ($C_{19}H_{18}O_3Na$, $[M+Na]^+$): 317.11482, found: 317.11624.


7.26 CDCl3
5.30 DCM
1.56 H2O
3 (m) 7.42
13 (m) 7.50
14,15 (m) 7.36
2 (m) 6.87
11 (s) 5.59
1',4' (m) 4.03
2' (m) 4.30
4' (m) 4.35
1' (dd) 4.17
6 (s) 3.00
3' (m) 1.71
3' (m) 2.00
δ1H / ppm


159.18
138.35
133.73
129.07
128.42
126.24
114.75
114.62
101.46
83.70
77.16 CDCl3
76.05
75.39
70.74
66.85
28.26
δ13C / ppm

Compound **S4.19**



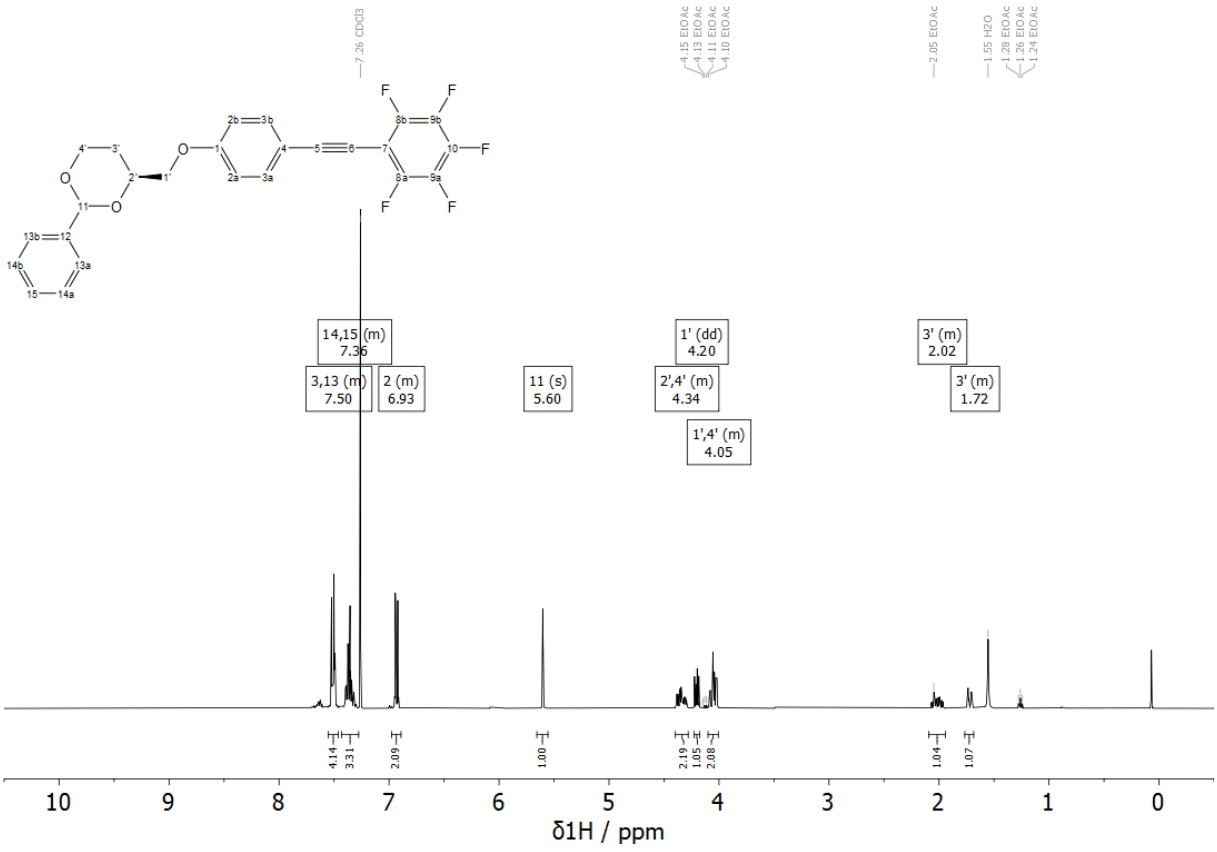
Under nitrogen atmosphere, CuI (38.8 mg, 204 µmol, 10 mol%), $Pd(PPh_3)_4$ (70.7 mg, 61.2 µmol, 3.0 mol%) and iodopentafluorobenzene (299 µL, 659 mg, 2.24 mmol, 1.10 eq.) were dissolved in argon-purged Et_3N (40 mL) and stirred for 10 min. Compound **S4.18** (600 mg, 2.04 mmol, 1.00 eq.) was added in small portions over a time period of 25 min and stirring was continued for 17 h at 80°C. The precipitate was filtered off and washed with Et_2O. The filtrate was washed with a saturated solution of ammonium chloride (40 mL). The aqueous phase was extracted with Et_2O (3 x 40 mL). The combined organic phases were washed with water and brine, dried over Na_2SO_4, evaporated and the residue was purified by column chromatography (hexane/EtOAc 20:1) to afford compound **S4.19** as a yellow solid (664 mg, 1.44 mmol, 71%).

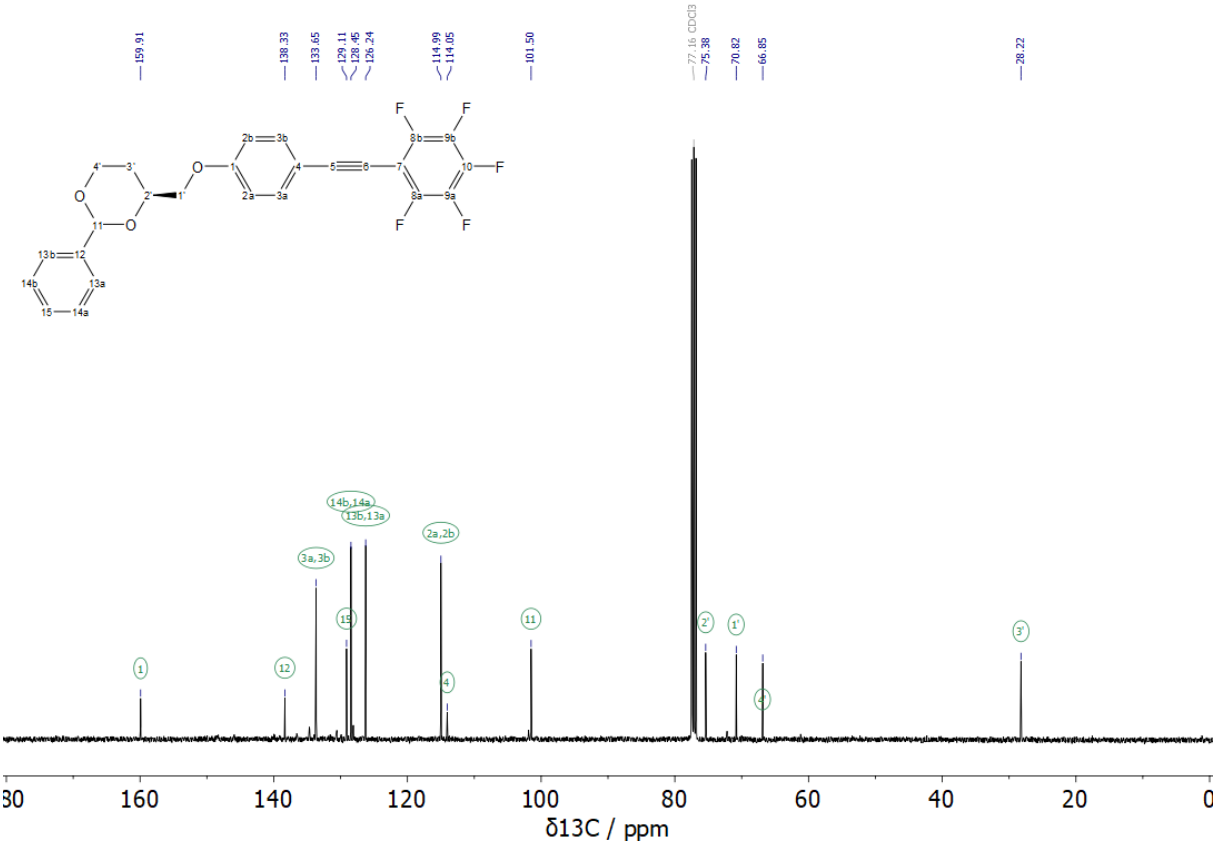
^{1}H NMR (400 MHz, $CDCl_3$): δ (ppm) = 7.55 – 7.46 (m, 4H, 3-H, Ph-H), 7.43 – 7.28 (m, 3H, Ph-H), 6.98 – 6.89 (m, 2H, 2-H), 5.60 (s, 1H, C*H*-Ph), 4.40 – 4.28 (m, 2H, 2'-H, 4'-H), 4.20 (dd, *J* = 9.9, 5.7 Hz, 1H, 1'-H), 4.10 – 4.01 (m, 2H, 1'-H, 4'-H), 2.09 – 1.94 (m, 1H, 3'-H), 1.77 – 1.68 (m, 1H, 3'-H);

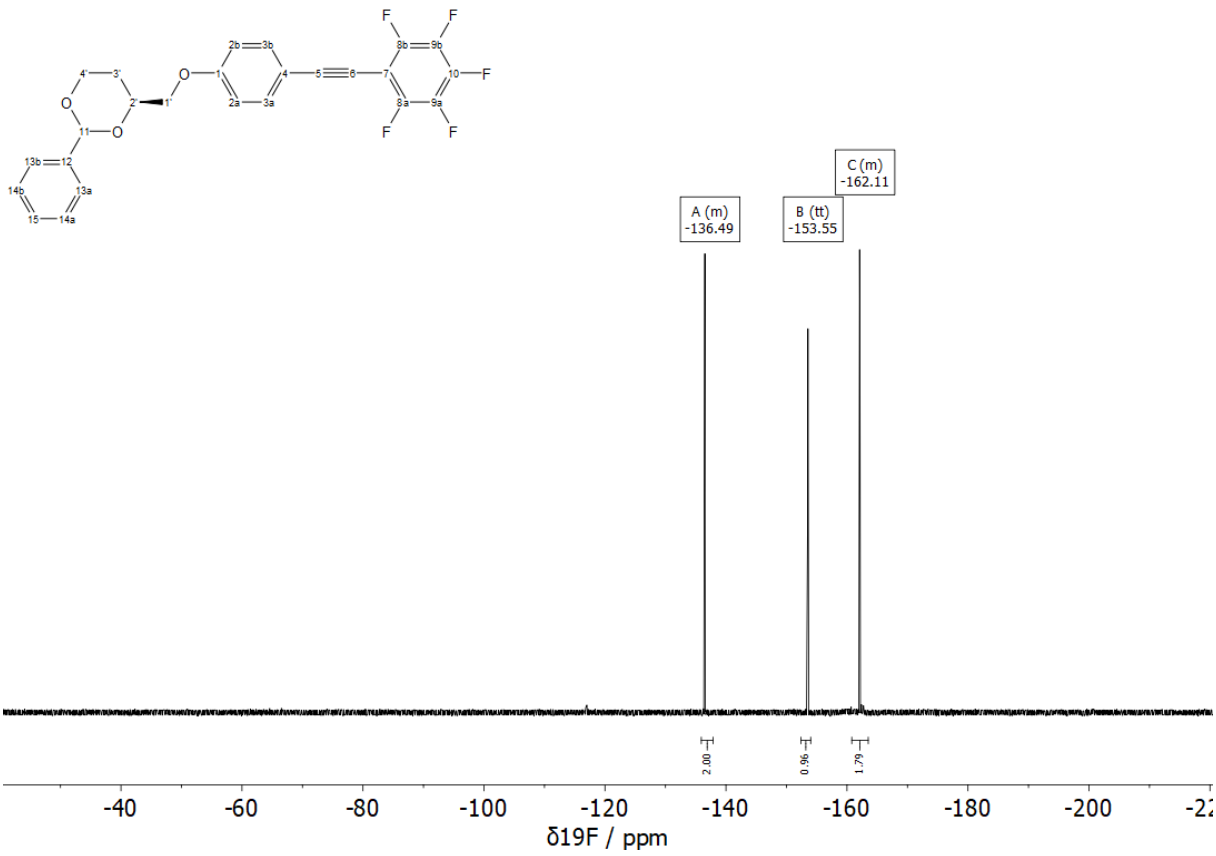
^{13}C{^{1}H} NMR (100 MHz, $CDCl_3$): δ (ppm) = 159.91 (1-C), 138.33 (12-C), 133.65 (3-C), 129.11 (15-C), 128.45 (14-C), 126.24 (12-C), 114.99 (2-C), 114.05 (4-C), 101.50 (*C*H-Ph), 75.38 (2'-C), 70.82 (1'-C), 66.85 (4'-C), 28.22 (3'-C);

^{19}F{^{1}H} NMR (376 MHz, $CDCl_3$): δ (ppm) = -135.92 – -137.88 (m,2F), -153.55 (tt, J = 20.7, 1.2 Hz, 1F), -160.80 – -163.51 (m, 2F);

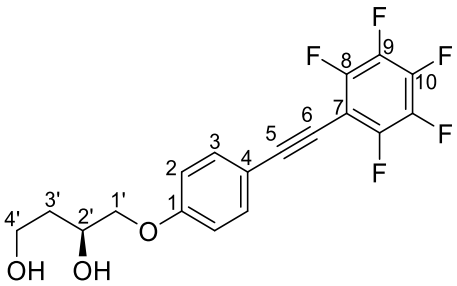
HR-MS (ESI+): *m/z* calc. ($C_{25}H_{17}F_5O_3Na$, $[M+Na]^+$): 483.09901, found: 483.10099.


7.26 CDCl3
4.15 EtOAc
4.13 EtOAc
4.11 EtOAc
4.10 EtOAc
2.05 EtOAc
1.55 H2O
1.28 EtOAc
1.26 EtOAc
1.24 EtOAc
14,15 (m) 7.36
3,13 (m) 7.50
2 (m) 6.93
11 (s) 5.60
1' (dd) 4.20
2',4' (m) 4.34
1',4' (m) 4.05
3' (m) 2.02
3' (m) 1.72
4.14
3.31
2.09
1.00
2.19
1.05
2.08
1.04
1.07
δ1H / ppm


159.91
138.33
133.65
129.11
128.45
126.24
114.99
114.05
101.50
77.16 CDCl3
75.38
70.82
66.85
28.22
14b,14a
13b,13a
2a,2b
3a,3b
δ13C / ppm



Compound **S4.20**



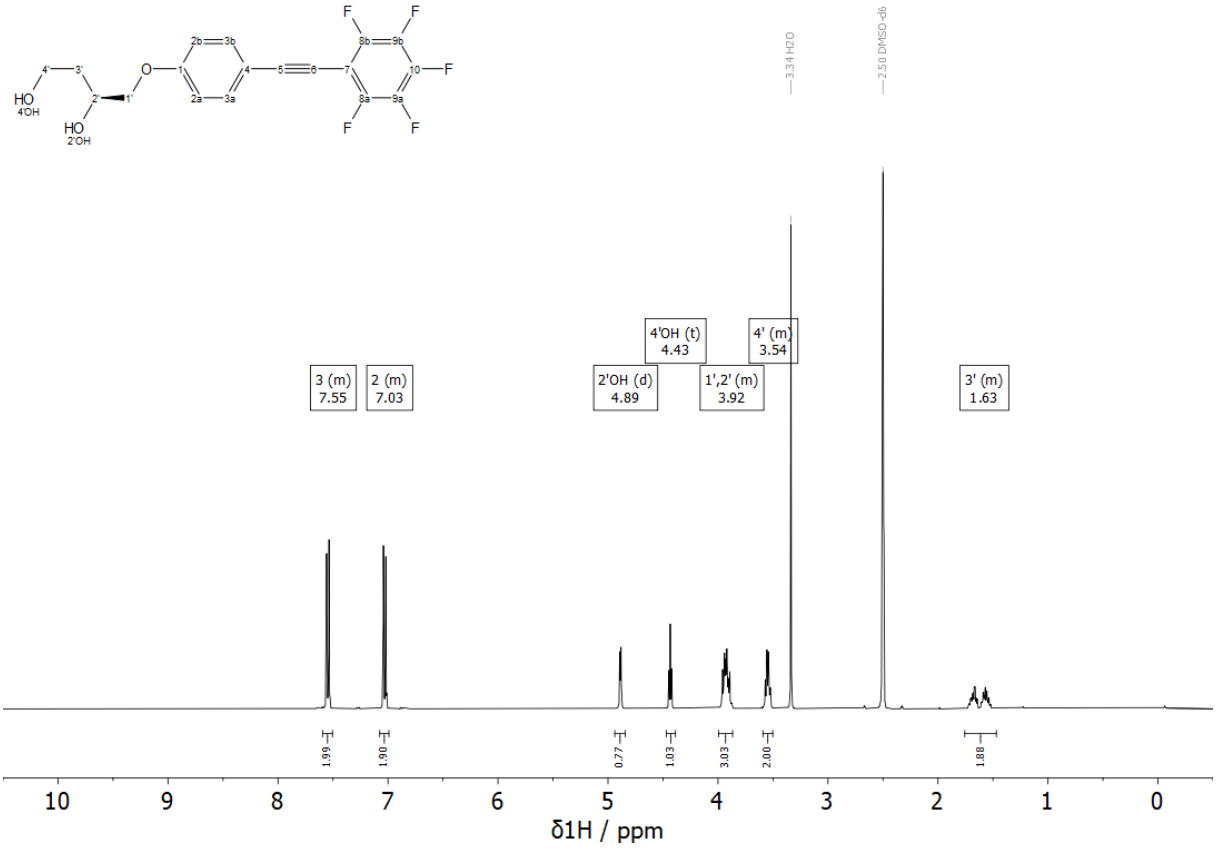
Compound **S4.19** (442 mg, 960 µmol, 1.00 eq.) was dissolved in THF (7.5 mL) and 4 M hydrochloric acid (7.5 mL) was added. The reaction mixture was stirred for 3 d at room temperature. The mixture was diluted with a saturated sodium hydrogen carbonate solution (25 mL) and extracted with EtOAc (3x30 mL). The combined organic phases were washed with brine and dried over Na_2SO_4. The solvent was removed under reduced pressure and the residue was purified by column chromatography (hexane/EtOAc 1:1) to afford compound **S4.20** as a colorless solid (205 mg, 551 µmol, 57%).

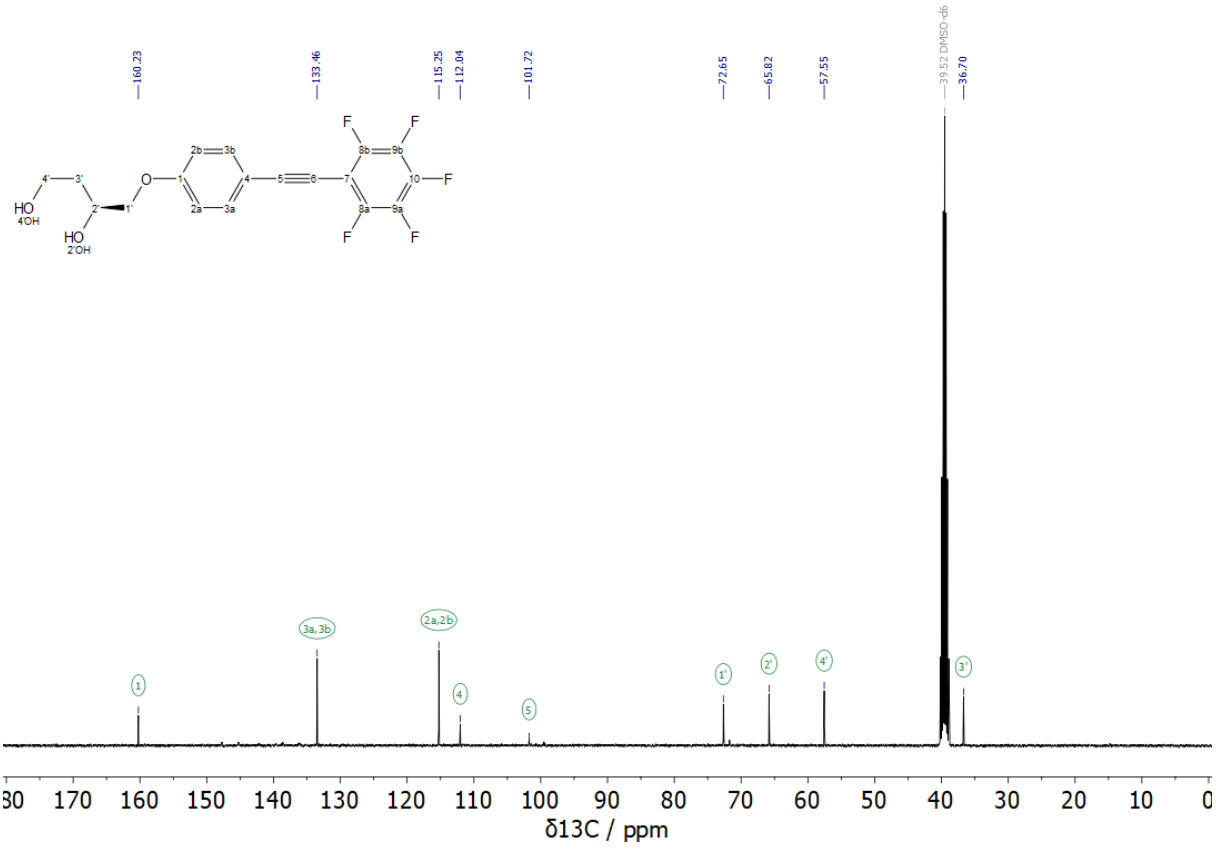
^{1}H NMR (400 MHz, DMSO-d_6): δ (ppm) = 7.59 – 7.51 (m, 2H, 3-H), 7.08 – 6.99 (m, 2H, 2-H), 4.89 (d, J = 4.7 Hz, 1H, 2'-OH), 4.43 (t, J = 5.1 Hz, 1H, 4'-OH), 4.00 – 3.87 (m, 3H, 1'-H, 2'-H), 3.59 – 3.50 (m, 2H, 4'-H), 1.76 – 1.47 (m, 2H, 3'-H);

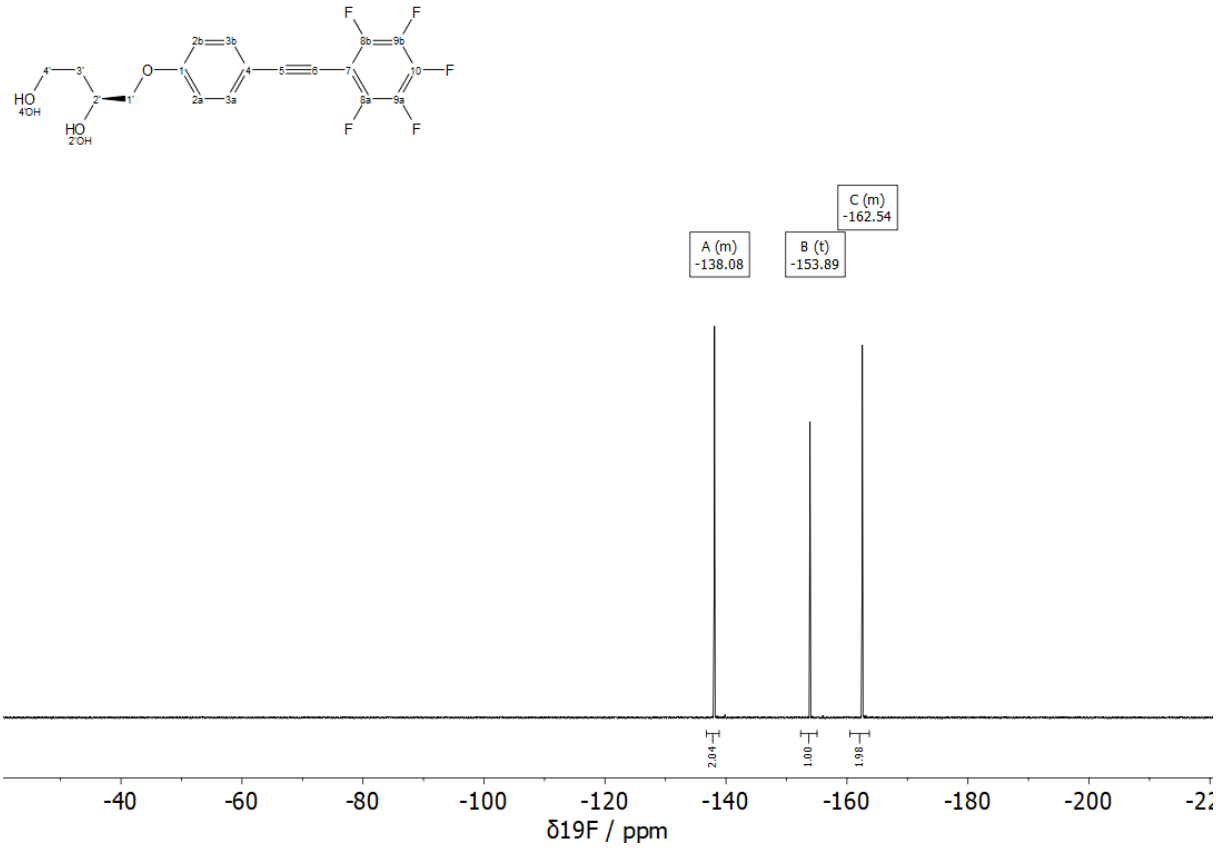
^{13}C{^{1}H} NMR (100 MHz, DMSO-d_6): δ (ppm) = 160.23 (1-C), 133.46 (3-C), 115.25 (2-C), 112.04 (4-C), 101.72 (5-C), 72.65 (1'-C), 65.82 (2'-C), 57.55 (4'-C), 36.70 (3'-C);

^{19}F{^{1}H} NMR (376 MHz, DMSO-d_6): δ (ppm) = -136.76 – -138.87 (m, 2F), -153.89 (t, J = 22.3 Hz, 1F), -160.48 – -163.70 (m, 2F);

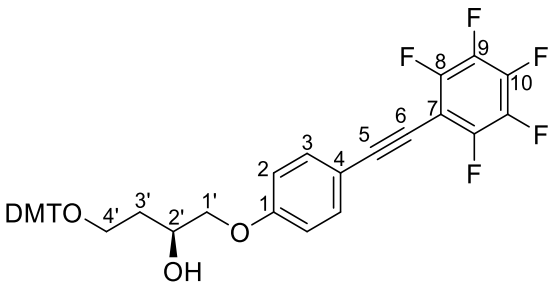
HR-MS (ESI+): m/z calc. ($C_{18}H_{13}F_5O_3Na$, $[M+Na]^+$): 395.06771, found: 395.06603.







Compound **S4.21**



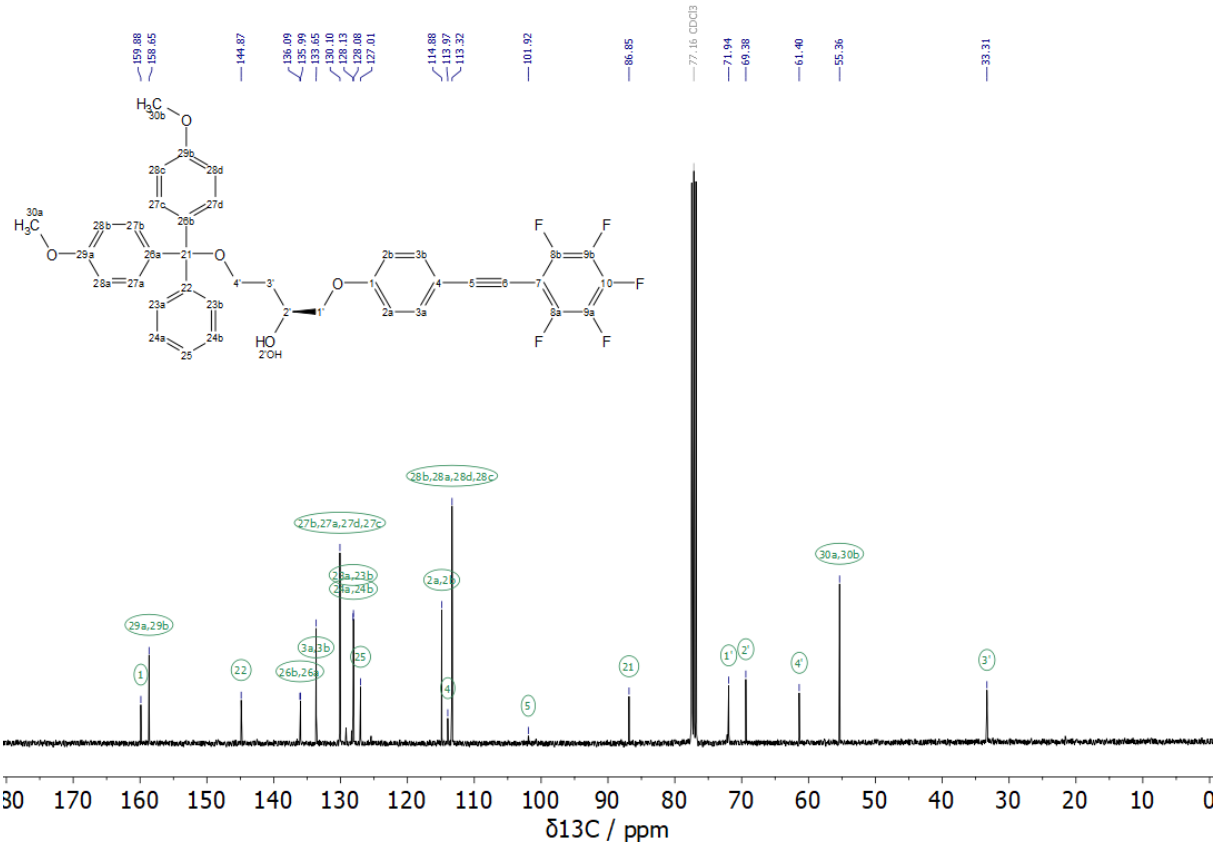
Compound **S4.20** (145 mg, 389 µmol, 1.00 eq.) was dissolved in anhydrous pyridine (14.5 mL). 4,4'-Dimethoxytrityl chloride (158 mg, 466 µmol, 1.20 eq.) was added in small portions over a time period of 30 min and the reaction mixture stirred at ambient temperature for 18 h. The solvent was removed under reduced pressure and the residue was purified by column chromatography (hexane/EtOAc 7:1 +1% Et_3N →4:1 +1% Et_3N) to afford compound **S4.21** as a colorless foam (162 mg, 240 µmol, 62%).

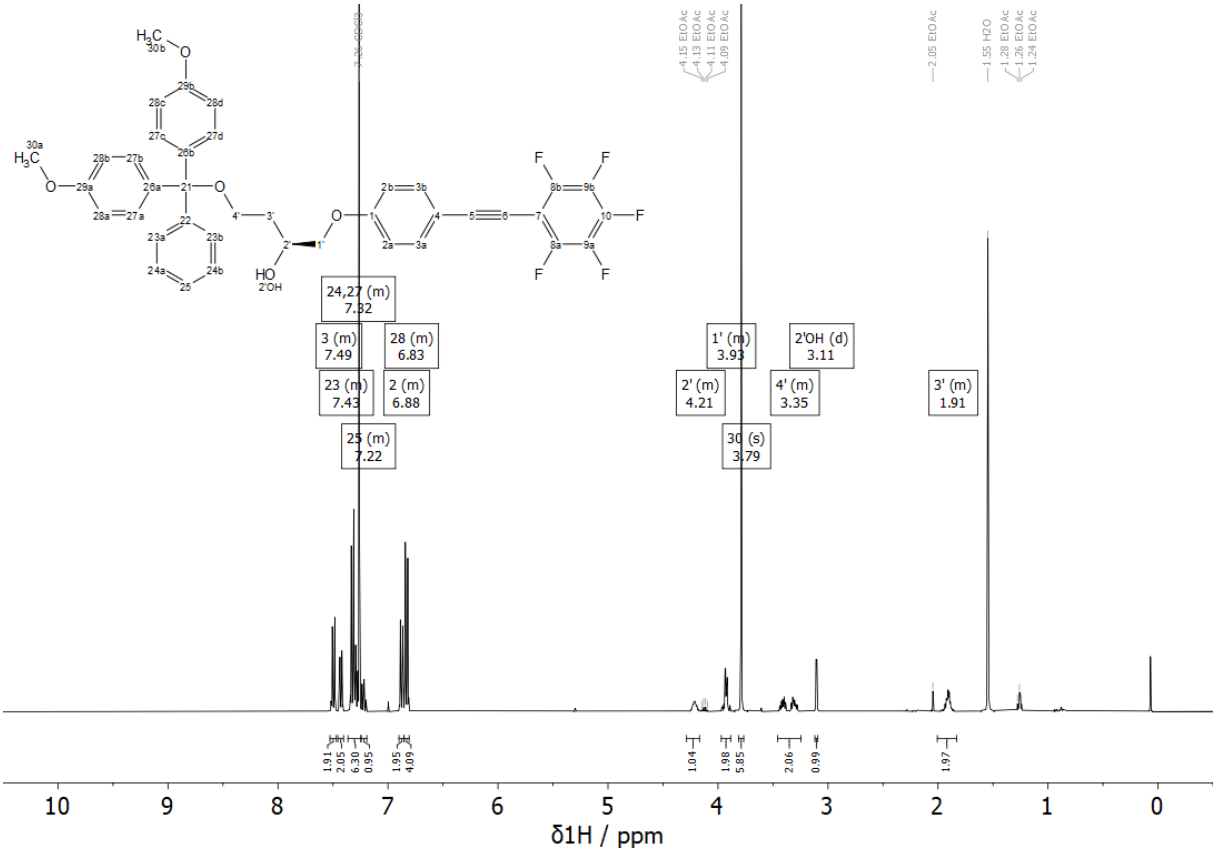
^{1}H NMR (400 MHz, $CDCl_3$): δ (ppm) = 7.53 – 7.47 (m, 2H, 3-H), 7.46 – 7.40 (m, 2H, DMT-H), 7.36 – 7.24 (m, 6H, DMT-H), 7.25 – 7.19 (m, 1H, DMT-H), 6.90 – 6.85 (m, 2H, 2-H), 6.85 – 6.81 (m, 4H, DMT-H), 4.29 – 4.17 (m, 1H, 2'-H), 3.97 – 3.89 (m, 2H, 1'-H), 3.79 (s, 6H, DMT-H), 3.46 – 3.25 (m, 2H, 4-H), 3.11 (d, J = 3.3 Hz, 1H, 2'-OH), 2.01 – 1.83 (m, 2H, 3'-H);

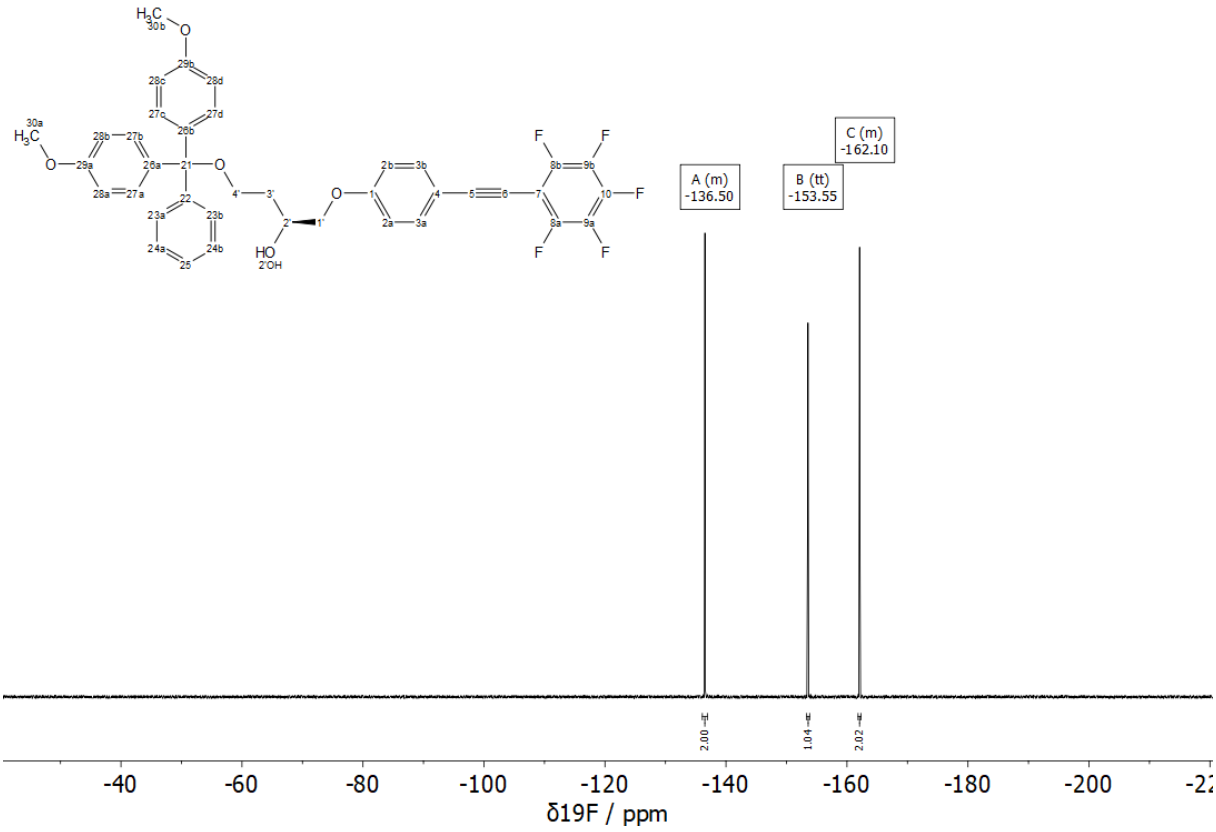
^{13}C{^{1}H} NMR (100 MHz, $CDCl_3$): δ (ppm) = 159.88 (1-C), 158.65 (DMT-C), 144.87 (DMT-C), 136.09 (DMT-C), 135.99 (DMT-C), 133.65 (3-C), 130.10 (DMT-C), 128.13 (DMT-C), 128.08 (DMT-C), 127.01 (DMT-C), 114.88 (2-C), 113.97 (4-C), 113.32 (DMT-C), 101.92 (5-C), 86.85 (DMT-C), 71.94 (1'-C), 69.38 (2'-C), 61.40 (4'-C), 55.36 (DMT-C), 33.31 (3'-C);

^{19}F{^{1}H} NMR (376 MHz, $CDCl_3$): δ (ppm) = -136.09 – -136.93 (m, 2F), -153.55 (tt, J = 21.1, 1.2 Hz, 1F), -161.84 – -162.32 (m, 2F);

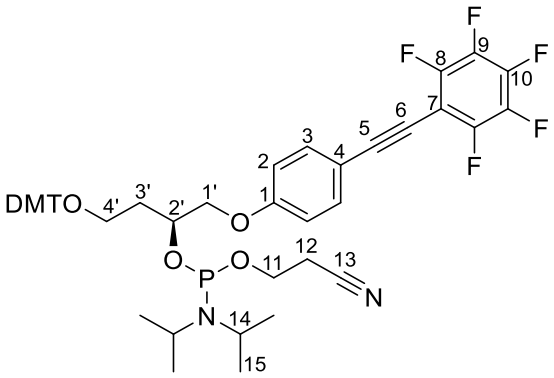
HR-MS (ESI+): m/z calc. ($C_{39}H_{31}F_5O_5Na$, $[M+Na]^+$): 697.19839, found: 697.19899.


159.88
158.65
144.87
136.09
135.99
133.65
130.10
128.13
128.08
127.01
114.88
113.97
113.32
101.92
86.85
77.16 CDCl3
71.94
69.38
61.40
55.36
33.31
δ13C / ppm


24,27 (m) 7.32
3 (m) 7.49
28 (m) 6.83
23 (m) 7.43
2 (m) 6.88
25 (m) 7.22
1' (m) 3.93
2'OH (d) 3.11
2' (m) 4.21
4' (m) 3.35
30 (s) 3.79
3' (m) 1.91
δ1H / ppm



Compound **S4.22**



Compound **S4.21** (140 mg, 208 µmol, 1.00 eq.) was dissolved under nitrogen atmosphere with DIPEA (217 µL, 161 mg 1.25 mmol, 6.00 eq.) in anhydrous DCM (6.5 mL). After 10 min CEP-Cl (63.9 mg, 270 µmol, 1.30 eq.) was added. Additional CEP-Cl (4.9 mg, 20.8 µmol, 0.10 eq.) was added after 2 h stirring at room temperature. The reaction mixture was stirred additionally at ambient temperature for 1 h. The solvent was removed under reduced pressure and the residue was purified by column chromatography (hexane/EtOAc 5:1 + 1% Et_3N) to afford compound **S4.22** as a colorless foam (148 mg, 169 µmol, 81%).

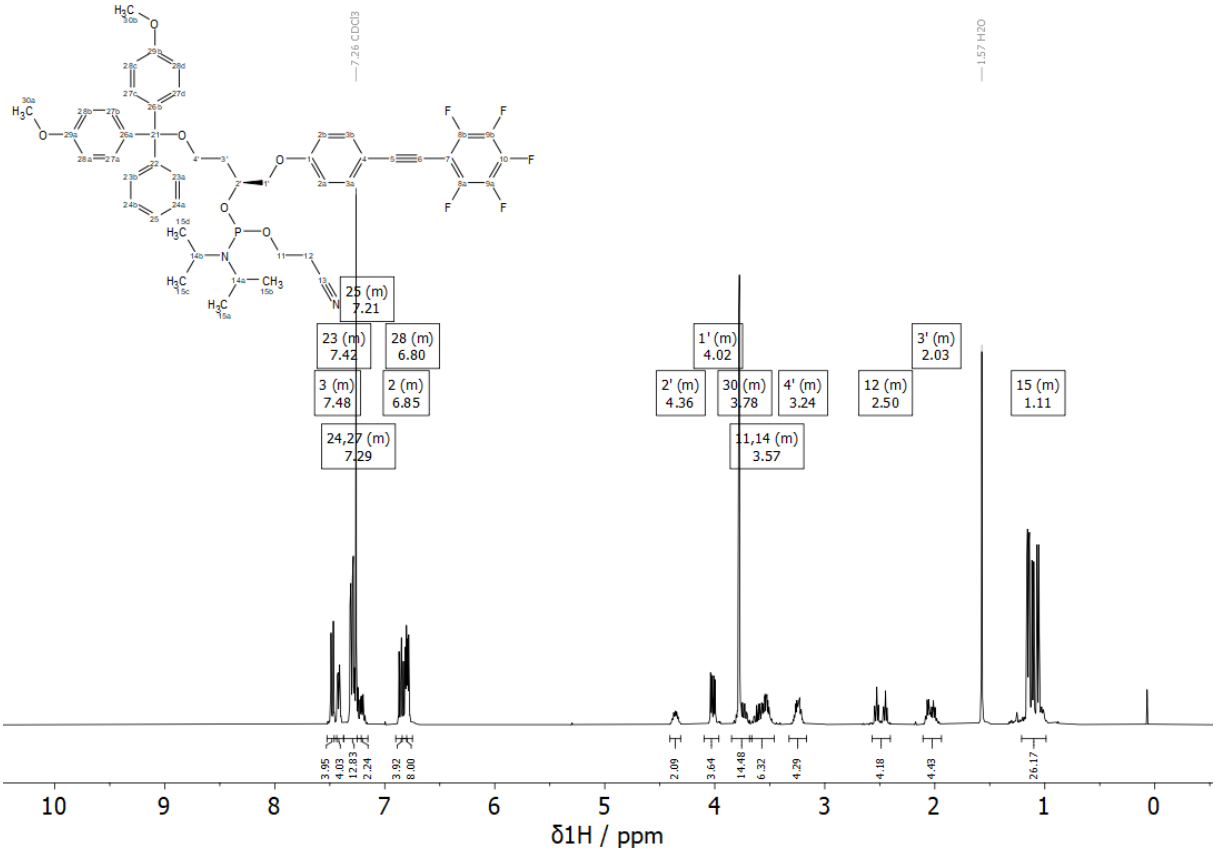
^{1}H NMR (400 MHz, $CDCl_3$): δ (ppm) = 7.53 – 7.44 (m, 4H, 3-H), 7.46 – 7.38 (m, 4H, DMT-H), 7.37 – 7.21 (m, 12H, DMT-H), 7.25 – 7.15 (m, 2H, DMT-H), 6.90 – 6.80 (m, 4H, 2-H), 6.84 – 6.75 (m, 8H, DMT-H), 4.41 – 4.31 (m, 2H, 2'H), 4.10 – 3.96 (m, 4H, 1'-H), 3.85 – 3.66 (m, 14H, 11-H, DMT-H), 3.68 – 3.46 (m, 6H, 11-H, 14-H), 3.32 – 3.17 (m, 4H, 4'-H), 2.57 – 2.40 (m, 4H, 12-H), 2.11 – 1.94 (m, 4H, 3'-H), 1.21 – 0.99 (m, 26H, 15-H);

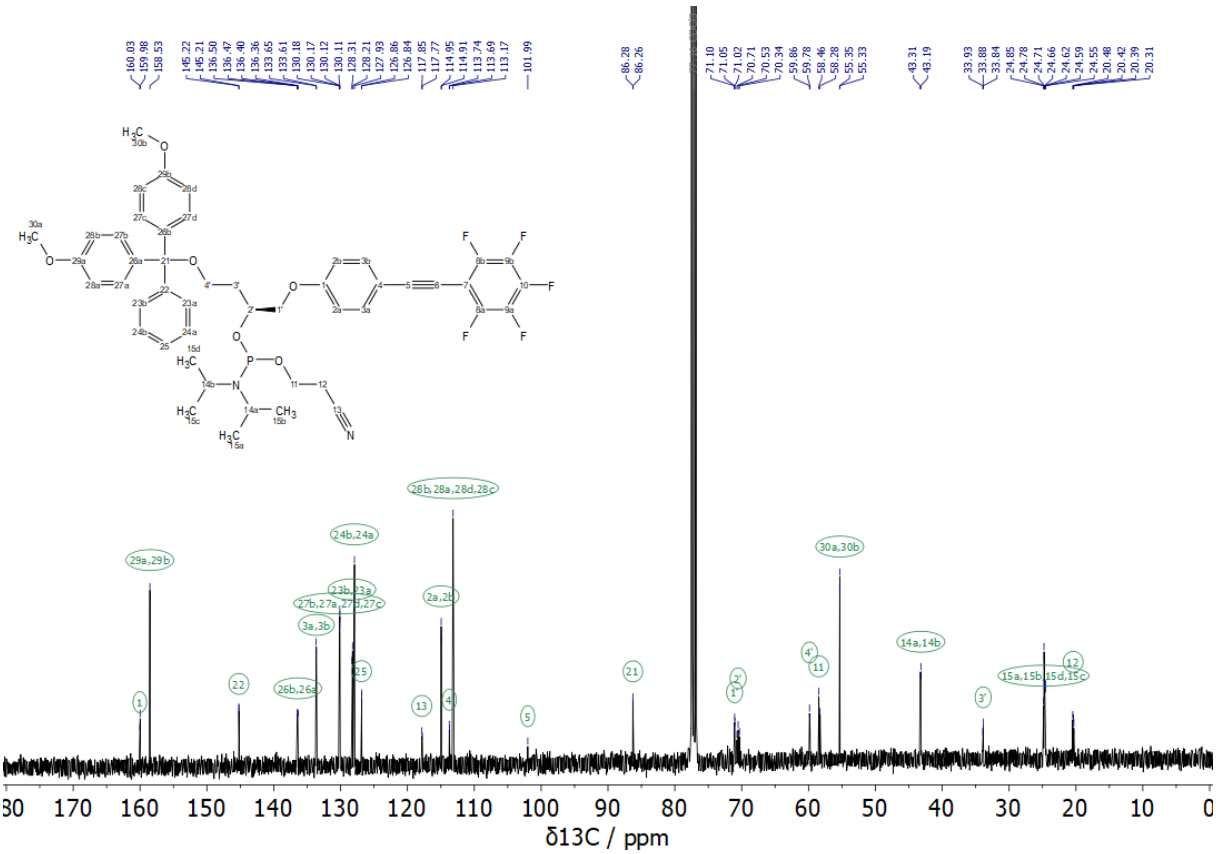
^{13}C{^{1}H} NMR (100 MHz, $CDCl_3$): δ (ppm) = 160.03 (1-C), 159.98 (1-C), 158.53 (DMT-C), 145.22 (DMT-C), 145.21 (DMT-C), 136.50 (DMT-C), 136.47 (DMT-C), 136.40 (DMT-C), 136.36 (DMT-C), 133.65 (3-C), 133.61 (3-C), 130.18 (DMT-C), 130.17 (DMT-C), 130.12 (DMT-C), 130.11 (DMT-C), 128.31 (DMT-C), 128.21 (DMT-C), 127.93 (DMT-C), 126.86 (DMT-C), 126.84 (DMT-C), 117.85 (13-C), 117.77 (13-C), 114.95 (2-C), 114.91 (2-C), 113.74 (4-C), 113.69 (4-C), 113.17 (DMT-C), 101.99 (5-C), 86.28 (DMT-C), 86.26 , 71.10 (1'-C), 71.05 (1'-C), 71.02 (1'-C), 70.71 (2'-C), 70.53 (2'-C), 70.34 (2'-C), 59.86 (4'-C), 59.78 (4'-C), 58.46 (11-C), 58.28 (11-C), 55.35 (DMT-C), 55.33 (DMT-C), 43.31 (14-C), 43.19 (14-C), 33.93 (3'-C), 33.88 (3'-C), 33.84 (3'-C), 24.85 (15-C), 24.78 (15-C), 24.71 (15-C), 24.66 (15-C), 24.62 (15-C), 24.59 (15-C), 24.55 (15-C), 20.48 (12-C), 20.42 (20-C), 20.39 (20-C), 20.31 (20-C);

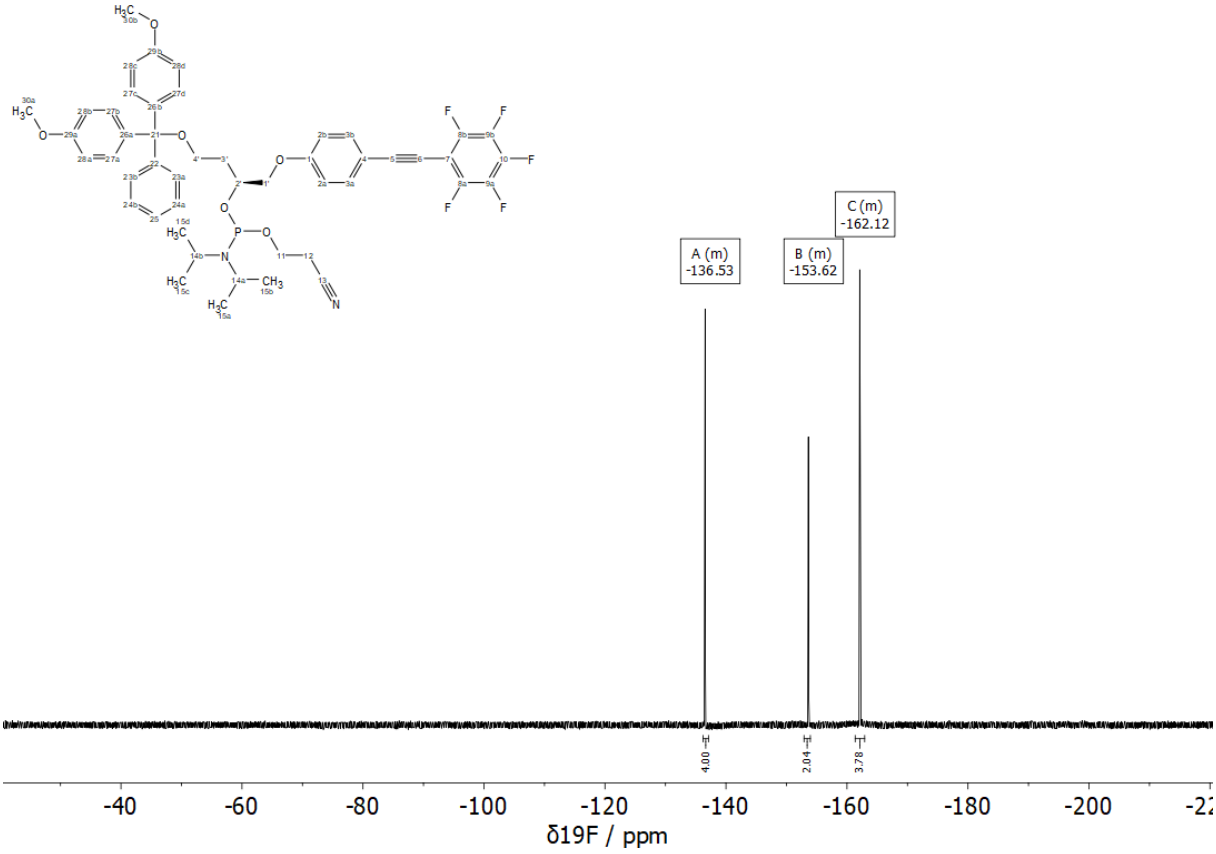
^{19}F{^{1}H} NMR (376 MHz, $CDCl_3$): δ (ppm) = -136.32 – -136.89 (m, 4F), -153.15 – -153.80 (m, 2F), -161.91 – -162.83 (m, 4F);

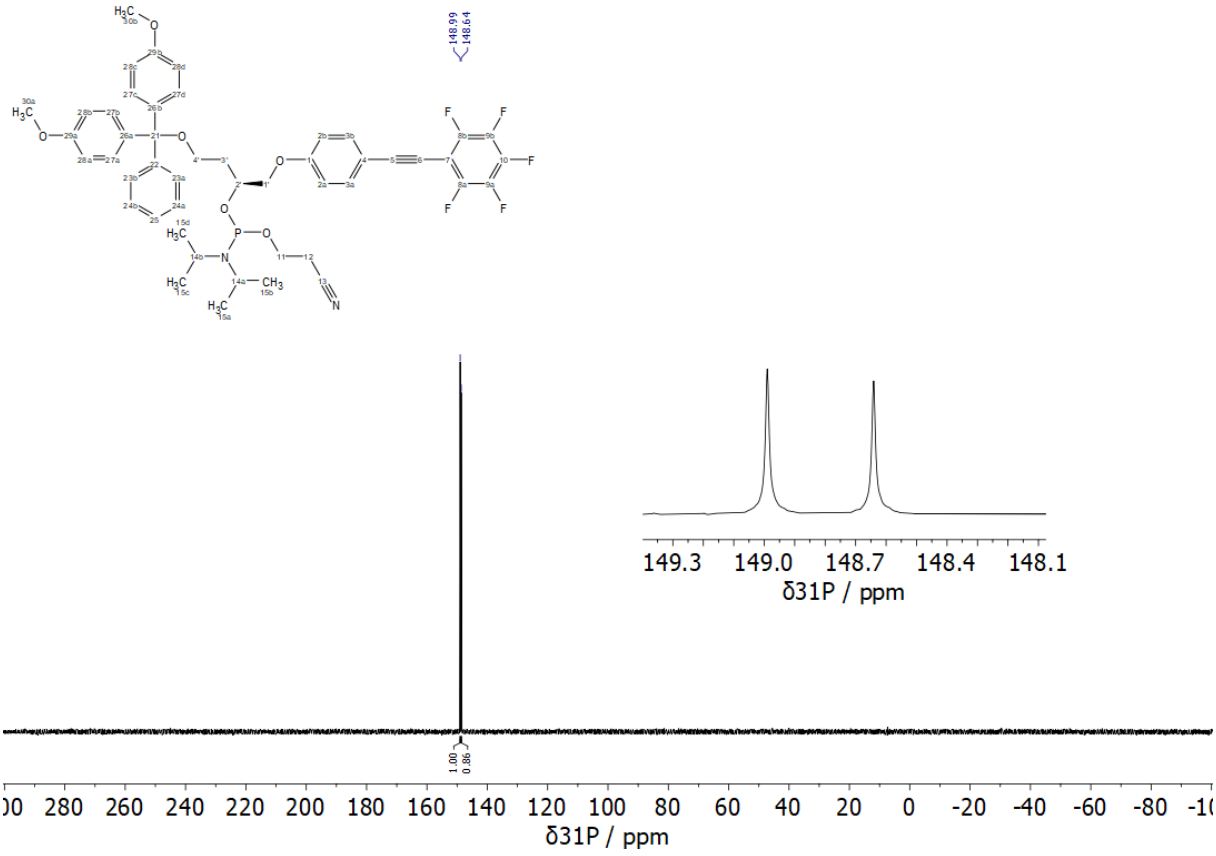
^{31}P{^{1}H} NMR (162 MHz, $CDCl_3$): δ (ppm) = 148.99, 148.64;

HR-MS (ESI+): *m/z* calc. ($C_{48}H_{48}F_5N_2O_6PNa$, $[M+Na]^+$): 897.30624, found: 897.30683.


7.26 CDCl3
1.57 H2O
25 (m) 7.21
23 (m) 7.42
28 (m) 6.80
3 (m) 7.48
2 (m) 6.85
24,27 (m) 7.29
1' (m) 4.02
2' (m) 4.36
30 (m) 3.78
4' (m) 3.24
11,14 (m) 3.57
12 (m) 2.50
3' (m) 2.03
15 (m) 1.11
δ1H / ppm


28b,28a,28d,28c
24b,24a
29a,29b
30a,30b
23b,23a
27b,27a,27d,27c
2a,2b
3a,3b
14a,14b
15a,15b,15d,15c
26b,26a
δ13C / ppm


A (m)
-136.53
B (m)
-153.62
C (m)
-162.12
4.00
2.04
3.78
-40
-60
-80
-100
-120
-140
-160
-180
-200
δ19F / ppm


148.99
148.64
149.3
149.0
148.7
148.4
148.1
δ31P / ppm
1.00
0.86
δ31P / ppm

4.4.2.4 Oligonucleotide synthesis

For further details see 3.4.2.4 Oligonucleotide synthesis.

The DNA strand (BTHF)3_DNA8mer1 was purified by RP-HPLC. Purification was performed on an ÄKTAmicro from GE Healthcare using a Nucleodur reversed-phase column (C18, 250x8 mm, 100 Å, 7 µm) from Machery Nagel at a flow rate of 2 mL min^{-1}. Linear gradients from 20–50% B over 35 min of buffer A (10 mM NH_4OAc, pH 6.6) and buffer B (MeCN) were used for purification. Chromatograms were monitored at 260 nm and 320 nm. All runs were performed at 40°C. HPLC purified DNA strands were lyophilized and dissolved in water.

4.4.2.5 Labeling of 5'-alkyne functionalized oligonucleotides

For further details see 3.4.2.5 Labeling of 5'-alkyne functionalized oligonucleotides.

4.4.2.6 UV/VIS spectroscopy / Thermal denaturing experiments

For further details see 3.4.2.6 UV/Vis spectroscopy / Thermal denaturing experiments.

UV/Vis spectra were measured in 10 mm quartz cuvettes from Varian using an Varian Cary100 with following settings:

- Wavelength: 200-750 nm
- Averaging time: 0.1 s
- Data interval: 1.0 nm
- Scan rate: 600 nm min^{-1}

Temperature dependent UV/Vis spectra were recorded every 10 °C in a range of 10-90 °C. 2 µM DNA samples in phosphate buffer (100 mM NaCl, 10 mM sodium phosphate, pH 7.0) were heated for 1 min at 95 °C and stored at least 30 min at room temperature prior to measurement. The samples in the cuvettes were overlaid with silicone oil.

For the estimation of the thermodynamic parameters for self-complementary dimers melting curves were analyzed according to Breslauer *et al*:[420]

$$\frac{1}{T_m} = \frac{(n-1)\cdot R}{\Delta H^0}\ln c_{total} + \frac{\Delta S^0 - (n-1)\cdot R\cdot \ln 2 + R\cdot \ln n}{\Delta H^0} \qquad (4.1)$$

4.4.2.7 Fluorescence spectroscopy

For further details see 3.4.2.8 Fluorescence spectroscopy.

4.4.2.8 FRET exchange experiments

For further details see 3.4.2.9 FRET exchange experiments.

4.4.2.9 NMR spectroscopy of DNA duplexes

For further details see 3.4.2.10 NMR spectroscopy of DNA duplexes.

4.4.2.10 Computational methods

For further details see 3.4.2.11 Computational methods.

4.4.3 Appendix

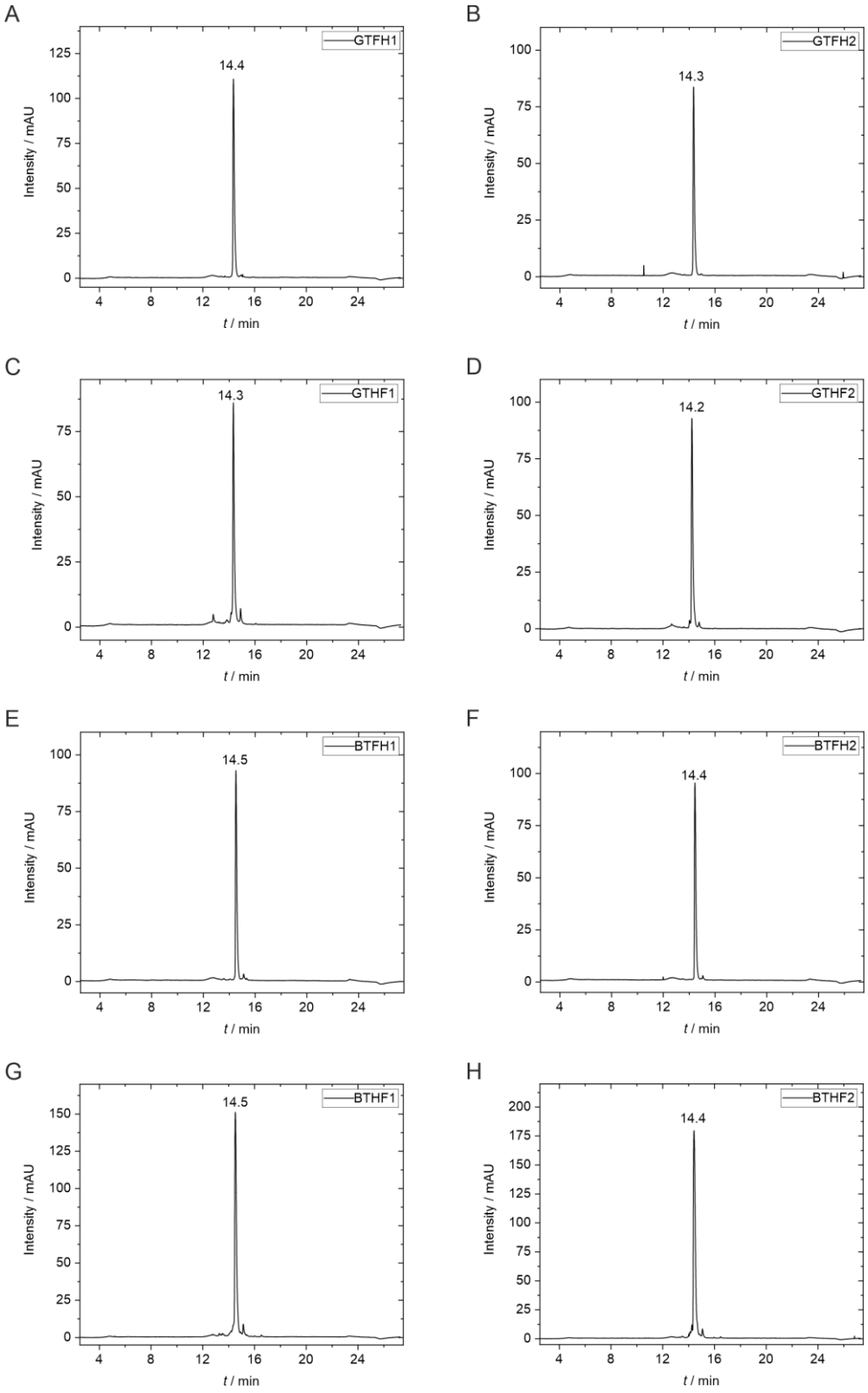


Figure S4.13 Analytical anion exchange HPLC chromatograms of the DNA oligonucleotides GTFH1 (A), GTFH2 (B), GTHF1 (C), GTHF2 (D), BTFH1 (E), BTFH2(F), BTHF1 (G) and BTHF2(H).

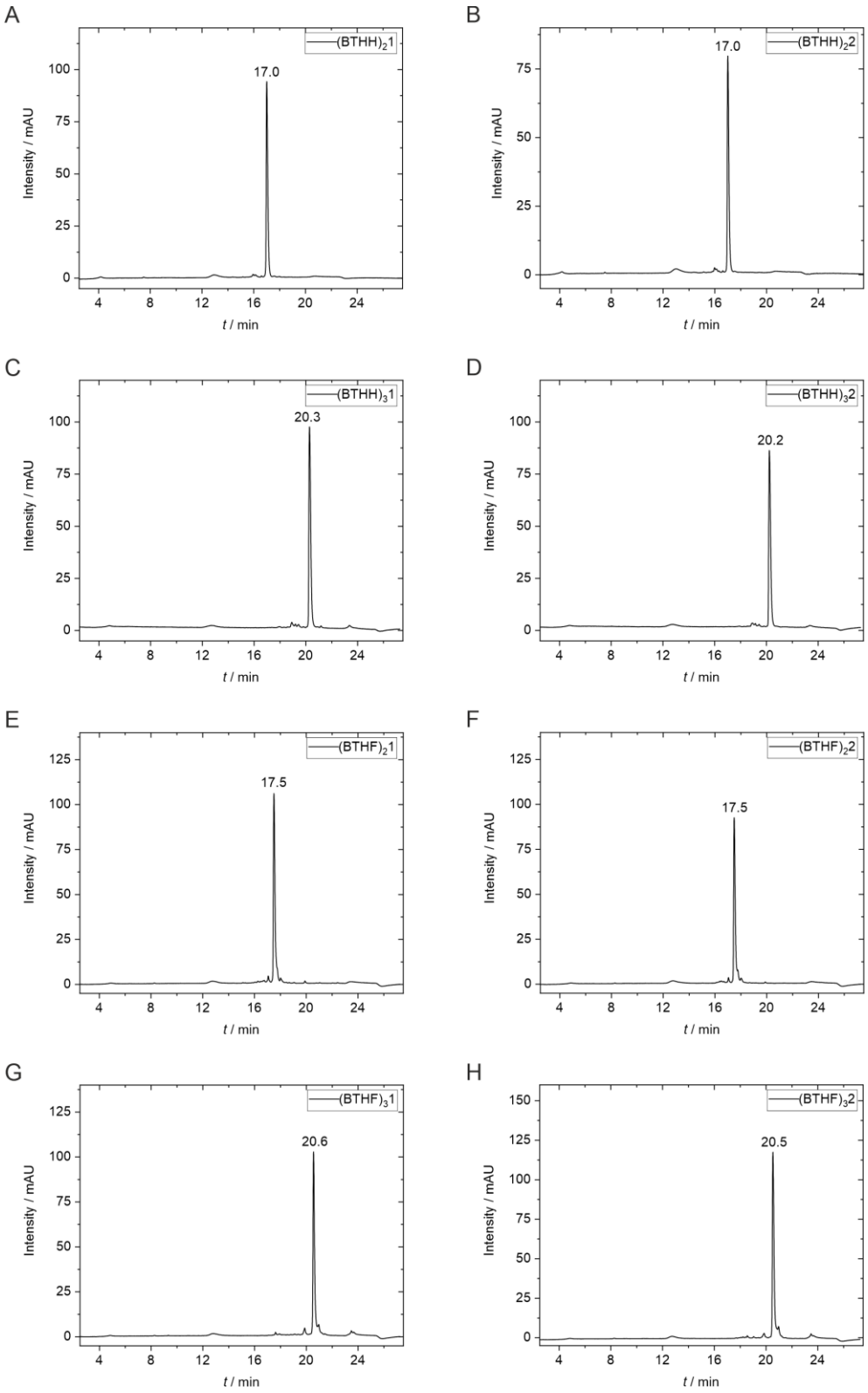


Figure S4.14 Analytical anion exchange HPLC chromatograms of the DNA oligonucleotides $(BTHH)_21$ (A), $(BTHH)_22$ (B), $(BTHH)_31$ (C), $(BTHH)_32$ (D), $(BTHF)_21$ (E), $(BTHF)_22$ (F), $(BTHF)_31$ (G) and $(BTHF)_32$ (H).

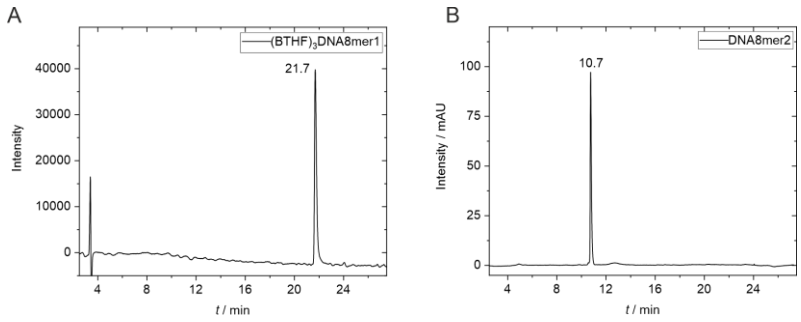


Figure S4.15 Analytical RP HPLC chromatogram of the DNA oligonucleotide $(BTHF)_3$DNA8mer1 (A) and anion exchange HPLC chromatogram of the DNA oligonucleotide DNA8mer1 (B).

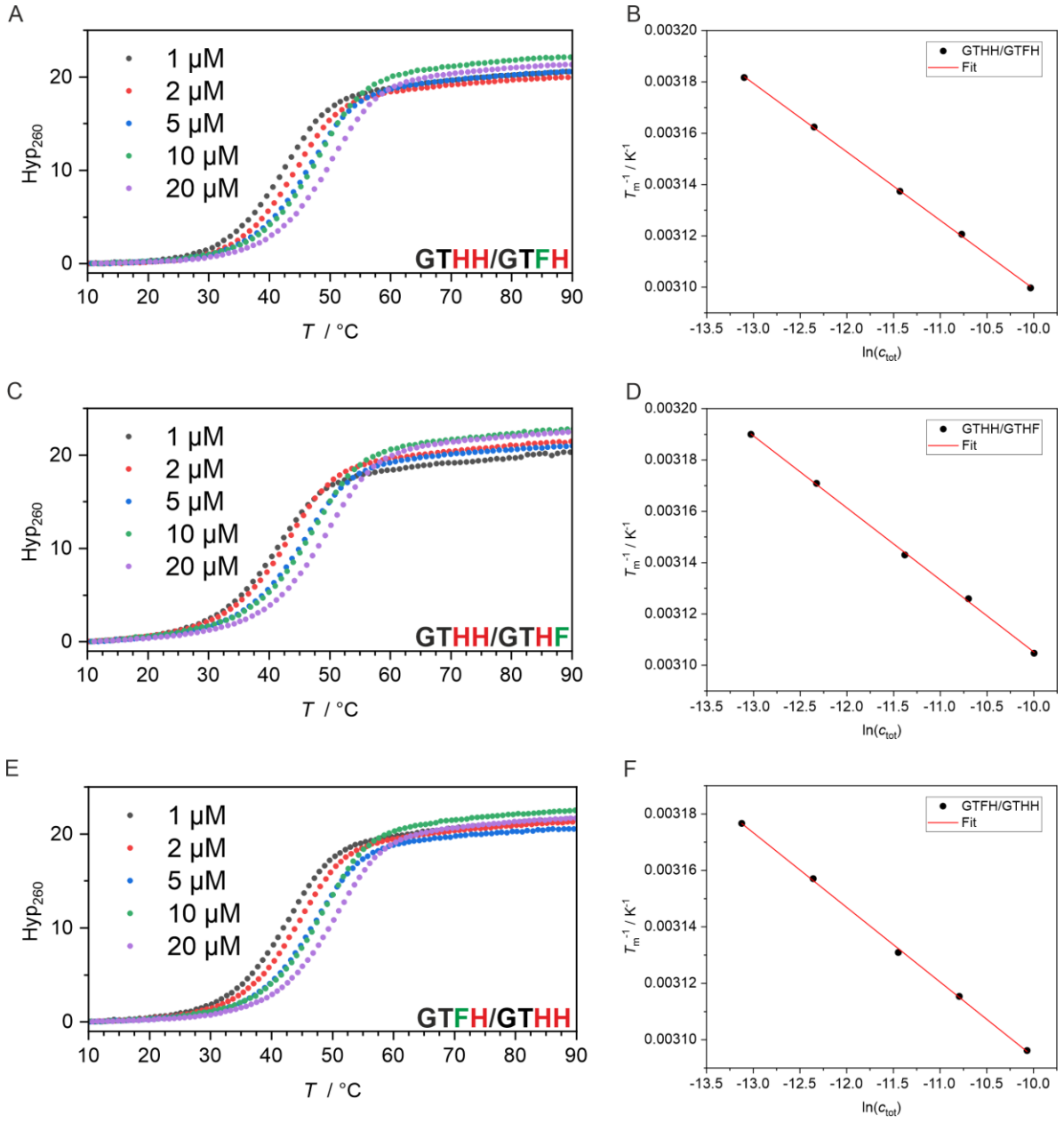


Figure S4.16 UV-thermal denaturation curves at different duplex concentrations in phosphate buffer (100 mM NaCl, 10 mM phosphate, pH 7.0) and the corresponding van't Hoff plot for GTHH/GTFH (A and B), GTHH/GTHF (C and D) and GTFH/GTHH (E and F).

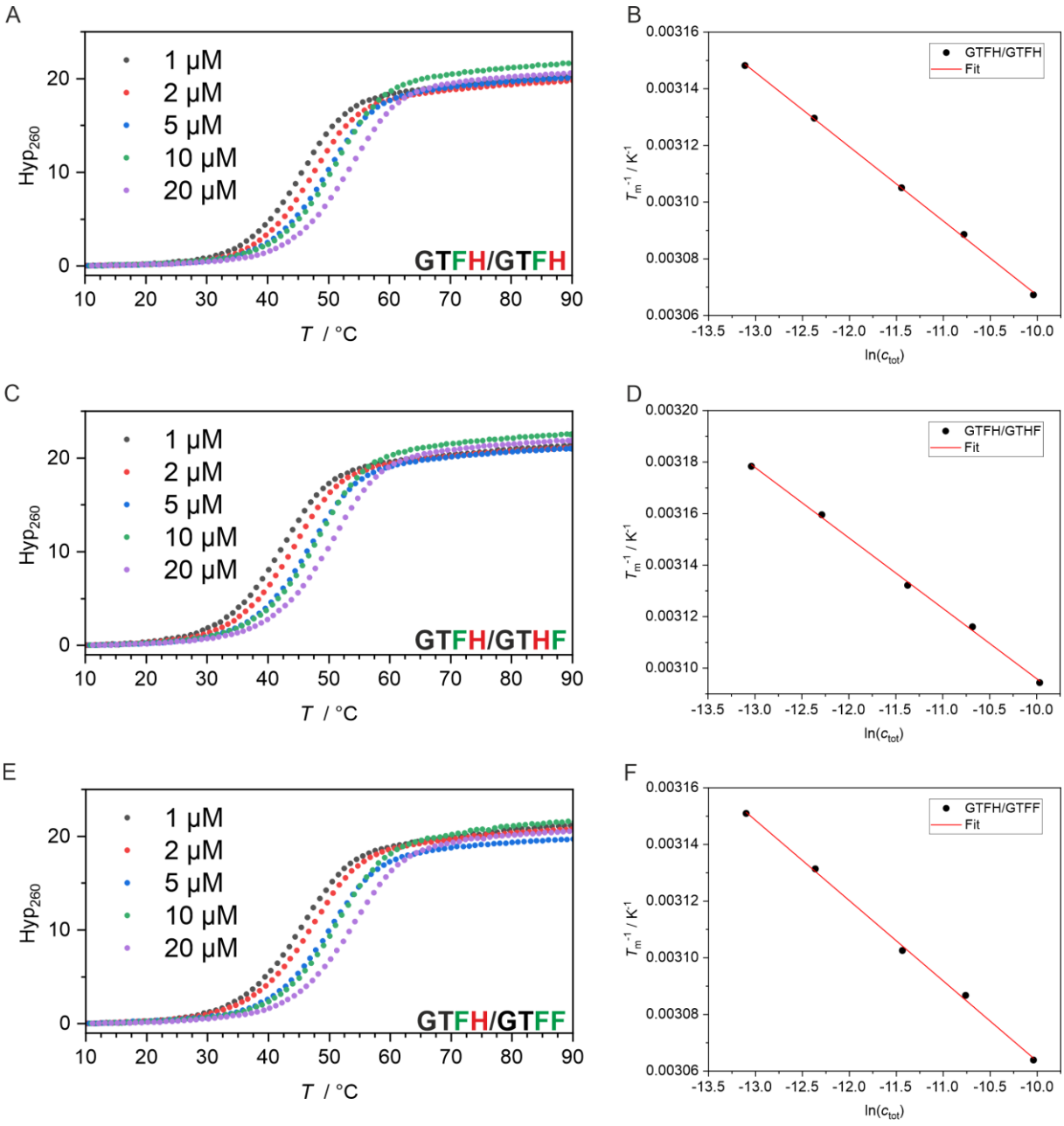


Figure S4.17 UV-thermal denaturation curves at different duplex concentrations in phosphate buffer (100 mM NaCl, 10 mM phosphate, pH 7.0) and the corresponding van't Hoff plot for GTFH/GTFH (A and B), GTFH/GTHF (C and D) and GTFH/GTFF (E and F).

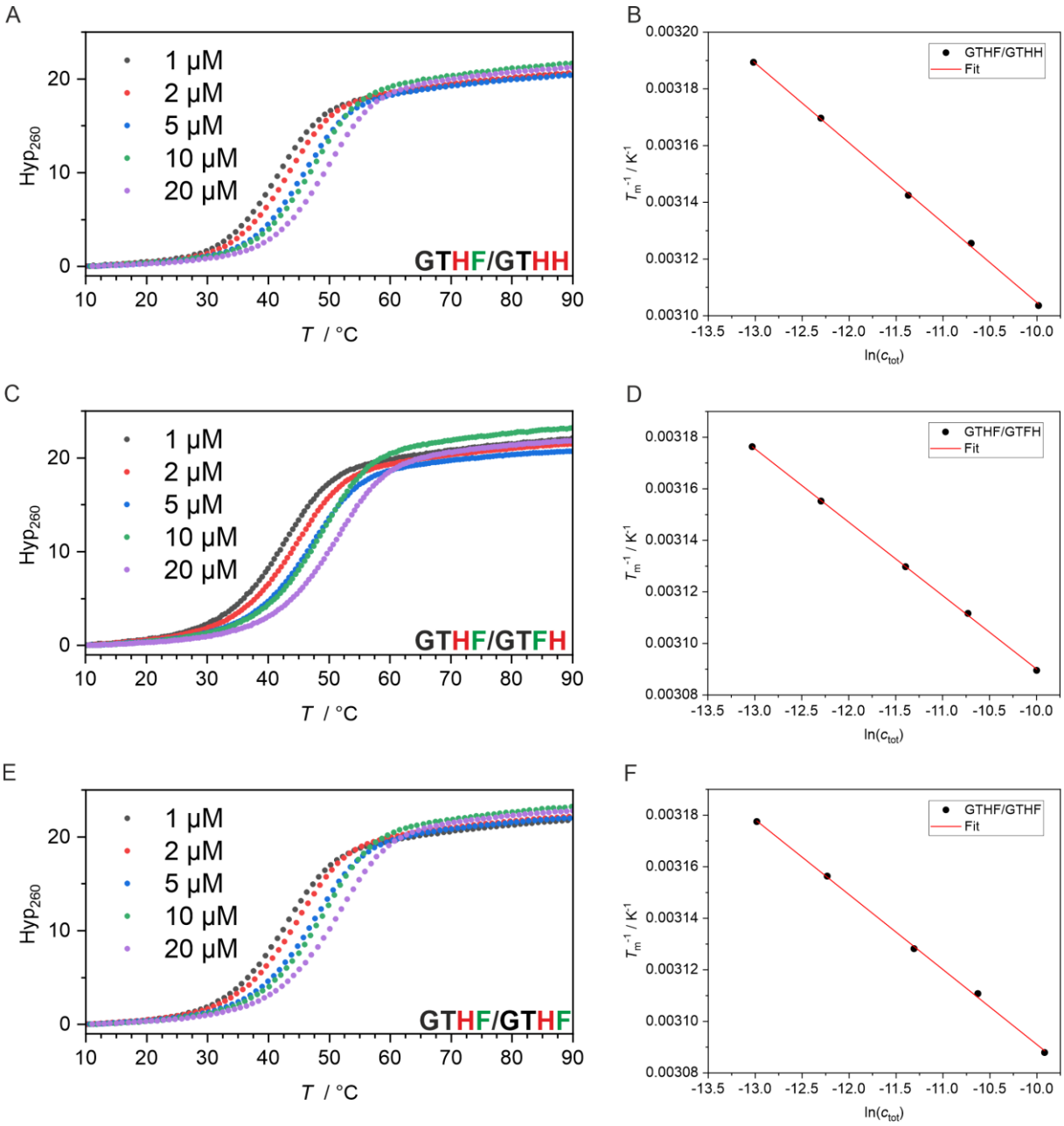


Figure S4.18 UV-thermal denaturation curves at different duplex concentrations in phosphate buffer (100 mM NaCl, 10 mM phosphate, pH 7.0) and the corresponding van't Hoff plot for GTHF/GTHH (A and B), GTHF/GTFH (C and D) and GTHF/GTHF (E and F).

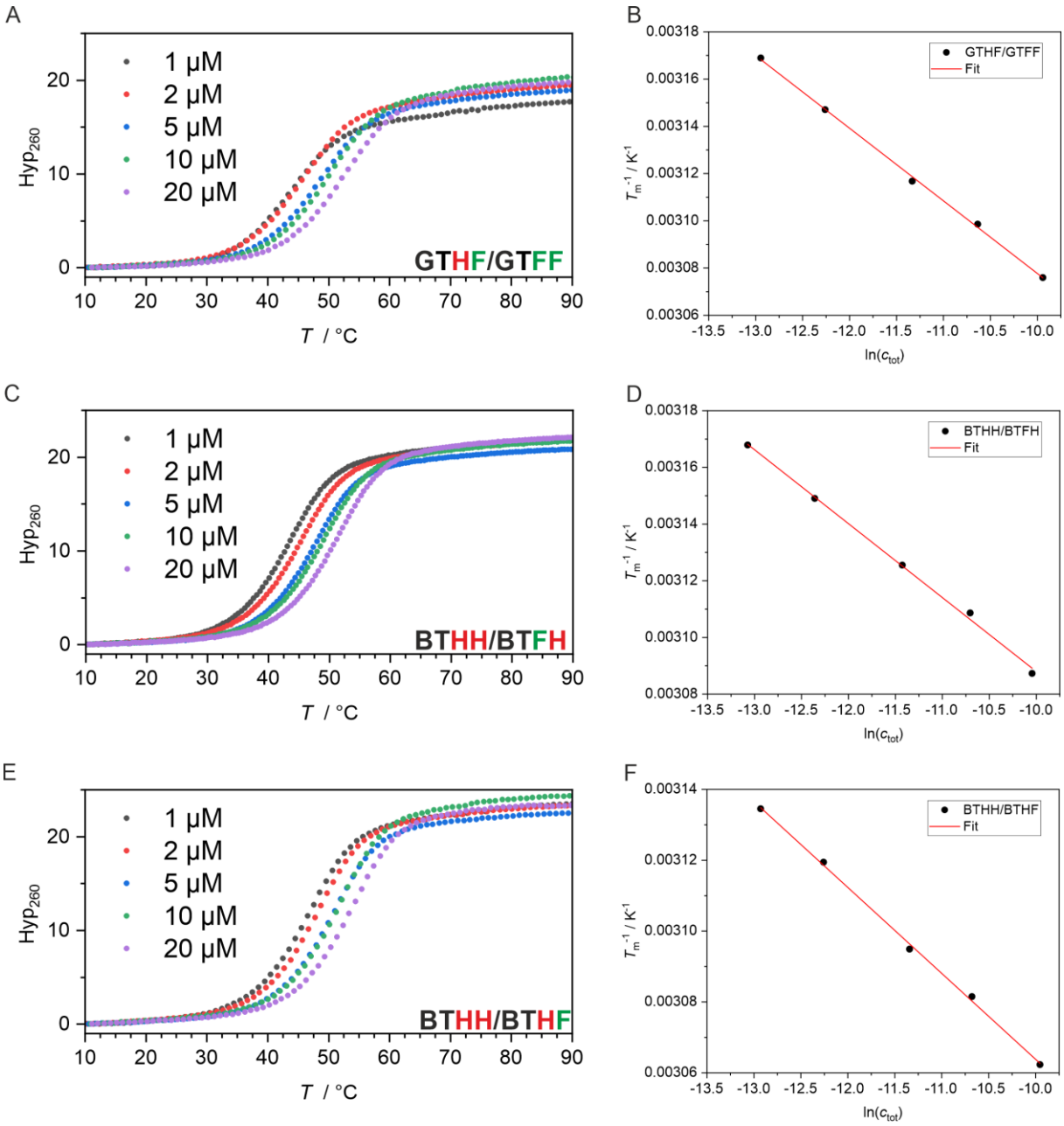


Figure S4.19 UV-thermal denaturation curves at different duplex concentrations in phosphate buffer (100 mM NaCl, 10 mM phosphate, pH 7.0) and the corresponding van't Hoff plot for GTHF/GTFF (A and B), BTHH/BTFH (C and D) and BTHH/BTHF (E and F).

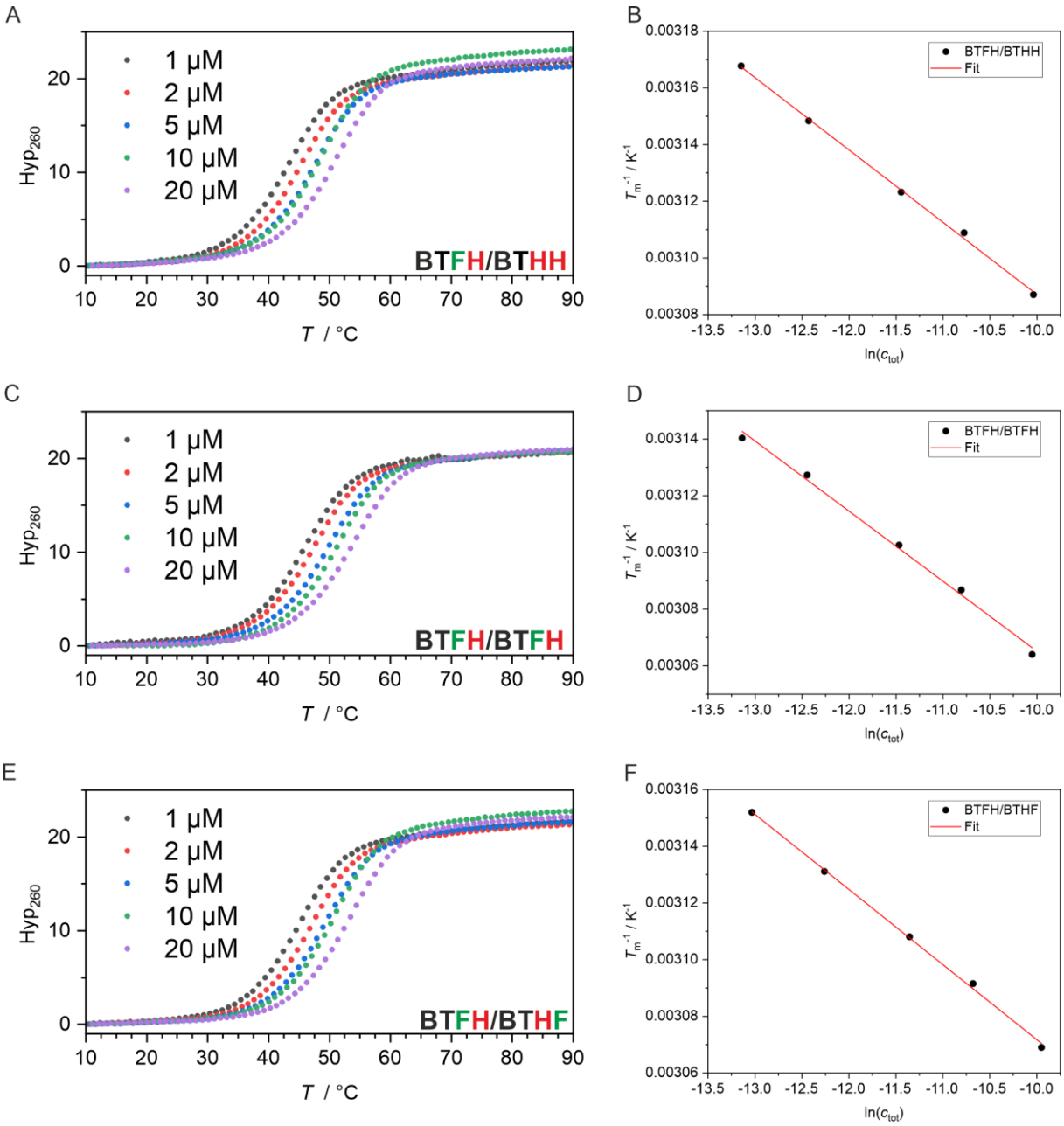


Figure S4.20 UV-thermal denaturation curves at different duplex concentrations in phosphate buffer (100 mM NaCl, 10 mM phosphate, pH 7.0) and the corresponding van't Hoff plot for BTFH/BTHH (A and B), BTFH/BTFH (C and D) and BTFH/BTHF (E and F).

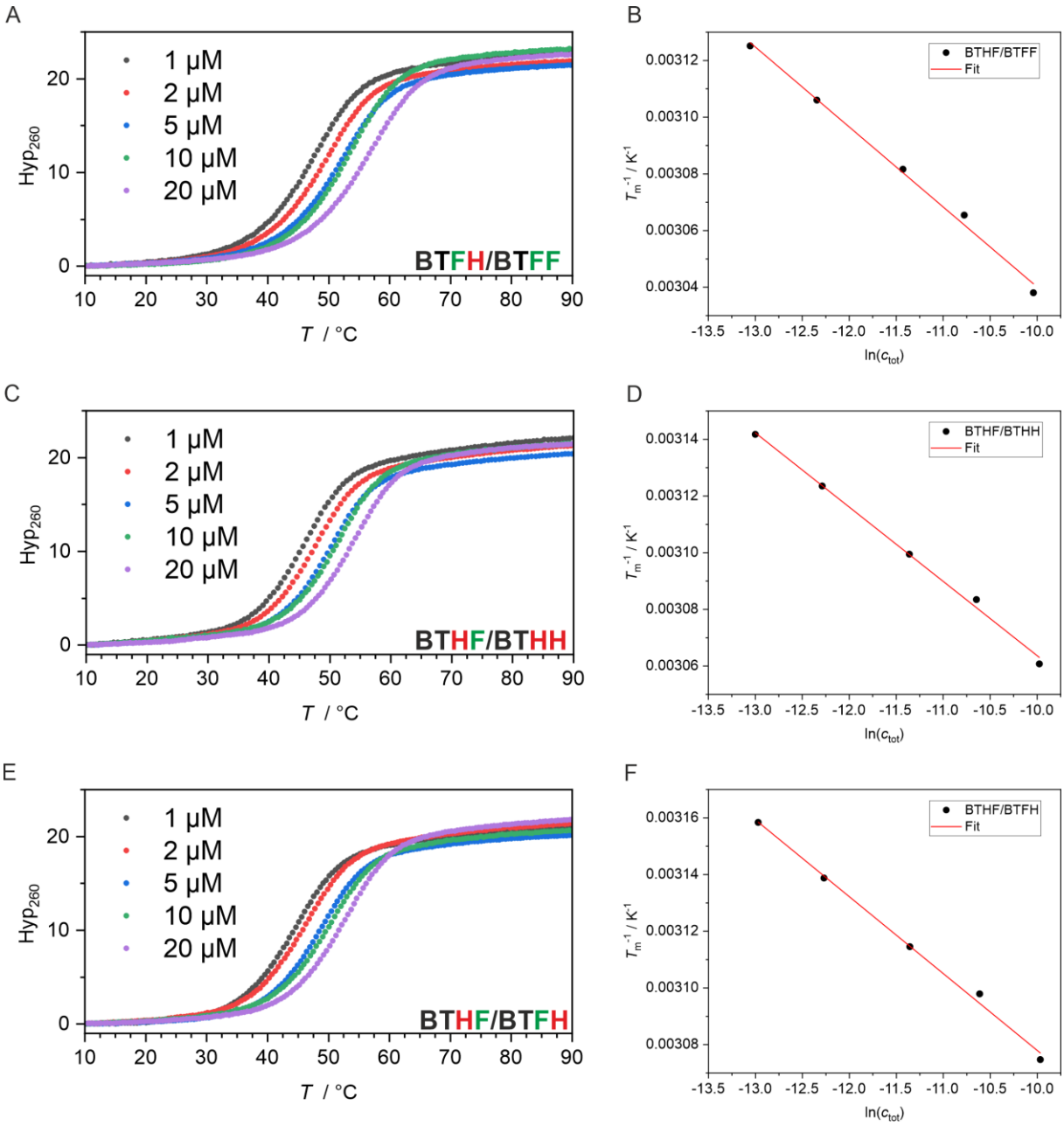


Figure S4.21 UV-thermal denaturation curves at different duplex concentrations in phosphate buffer (100 mM NaCl, 10 mM phosphate, pH 7.0) and the corresponding van't Hoff plot for BTFH/BTFF (A and B), BTHF/BTHH (C and D) and BTHF/BTFH (E and F).

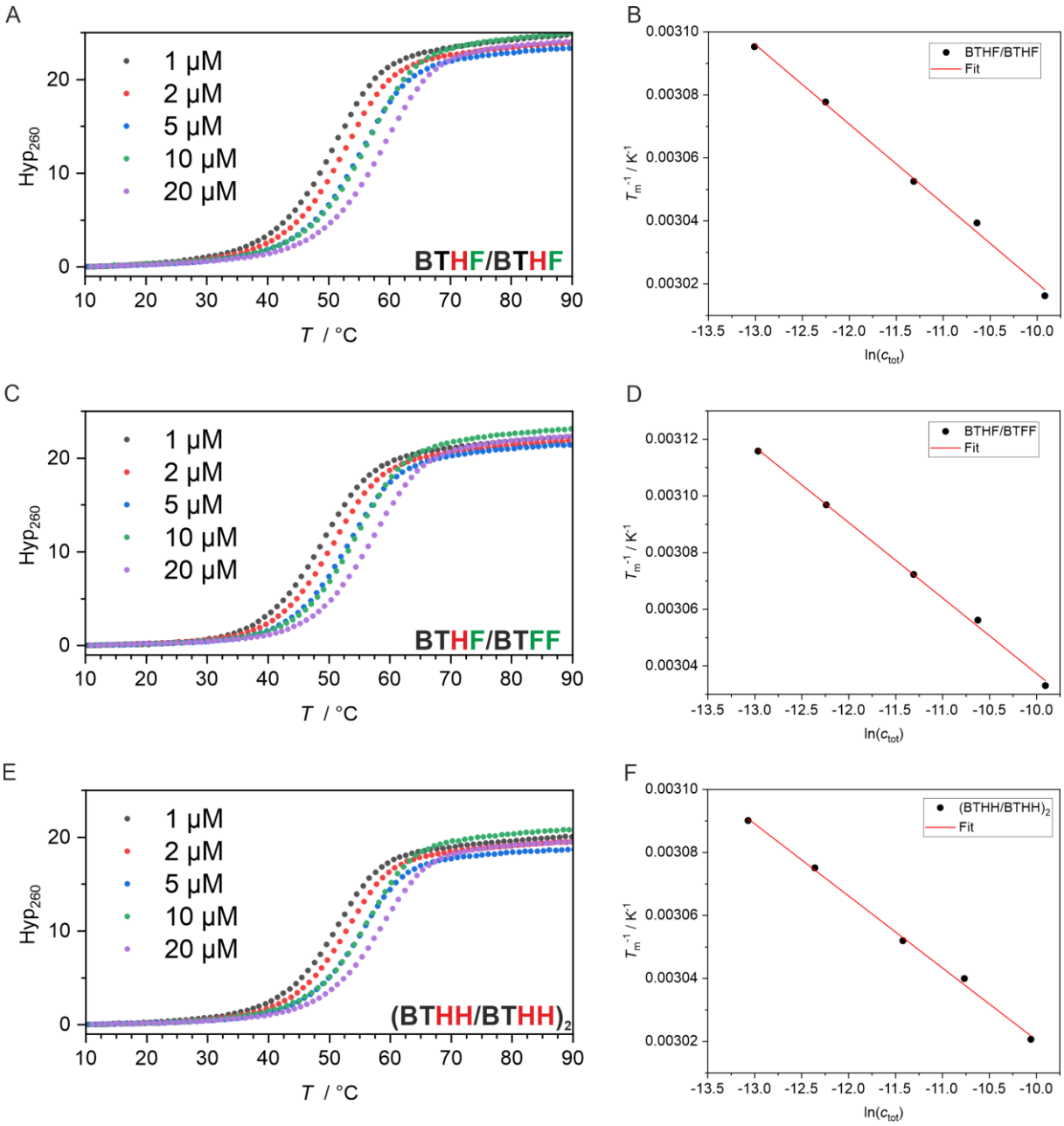


Figure S4.22 UV-thermal denaturation curves at different duplex concentrations in phosphate buffer (100 mM NaCl, 10 mM phosphate, pH7.0) and the corresponding van't Hoff plot for BTHF/BTHF (A and B), BTHF/BTFF (C and D) and $(BTHH/BTHH)_2$ (E and F).

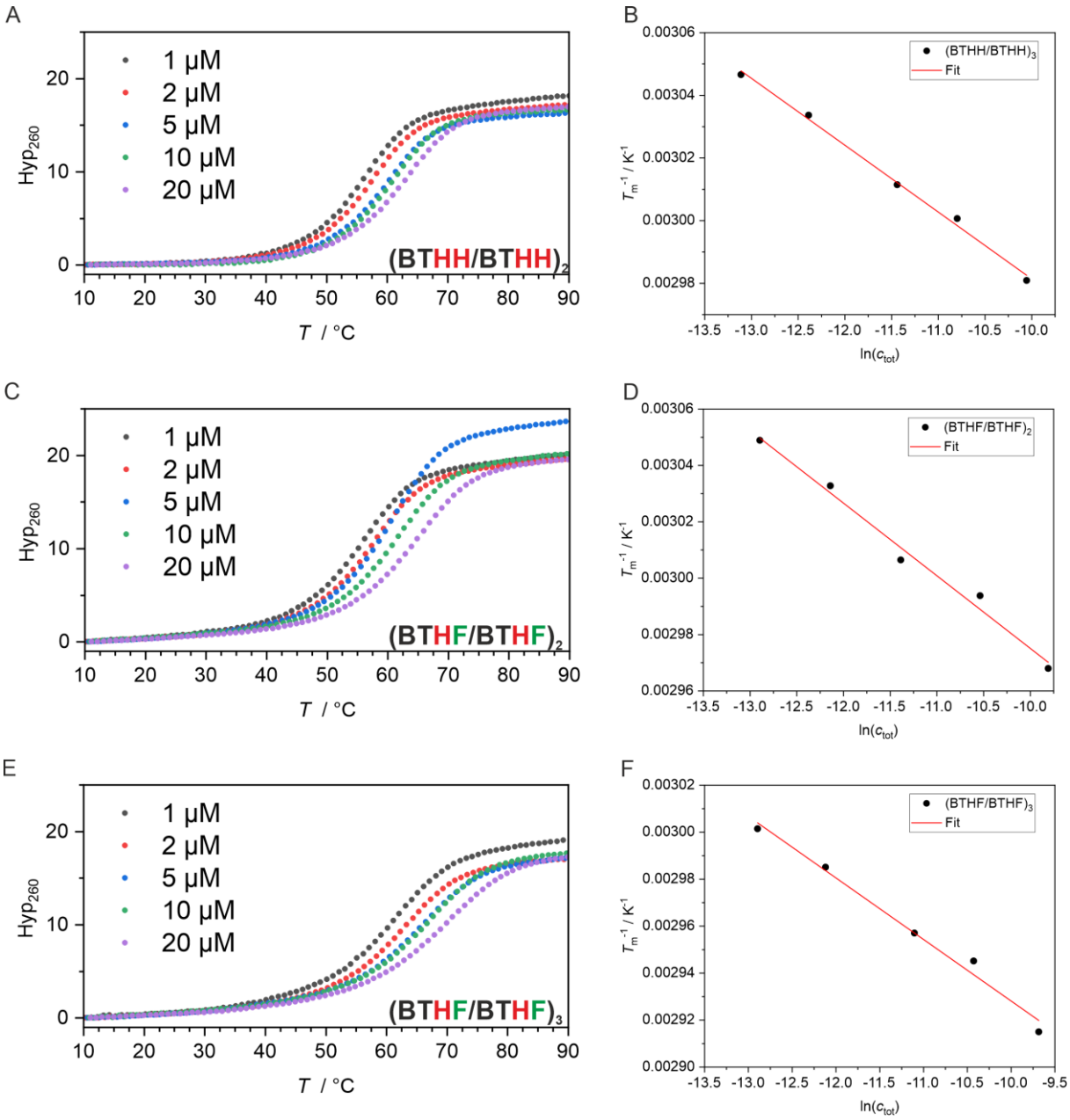


Figure S4.23 UV-thermal denaturation curves at different duplex concentrations in phosphate buffer (100 mM NaCl, 10 mM phosphate, pH 7.0) and the corresponding van't Hoff plot for $(BTHH/BTHH)_3$ (A and B), $(BTHF/BTHF)_2$ and $(BTHF/BTHF)_3$ (E and F).

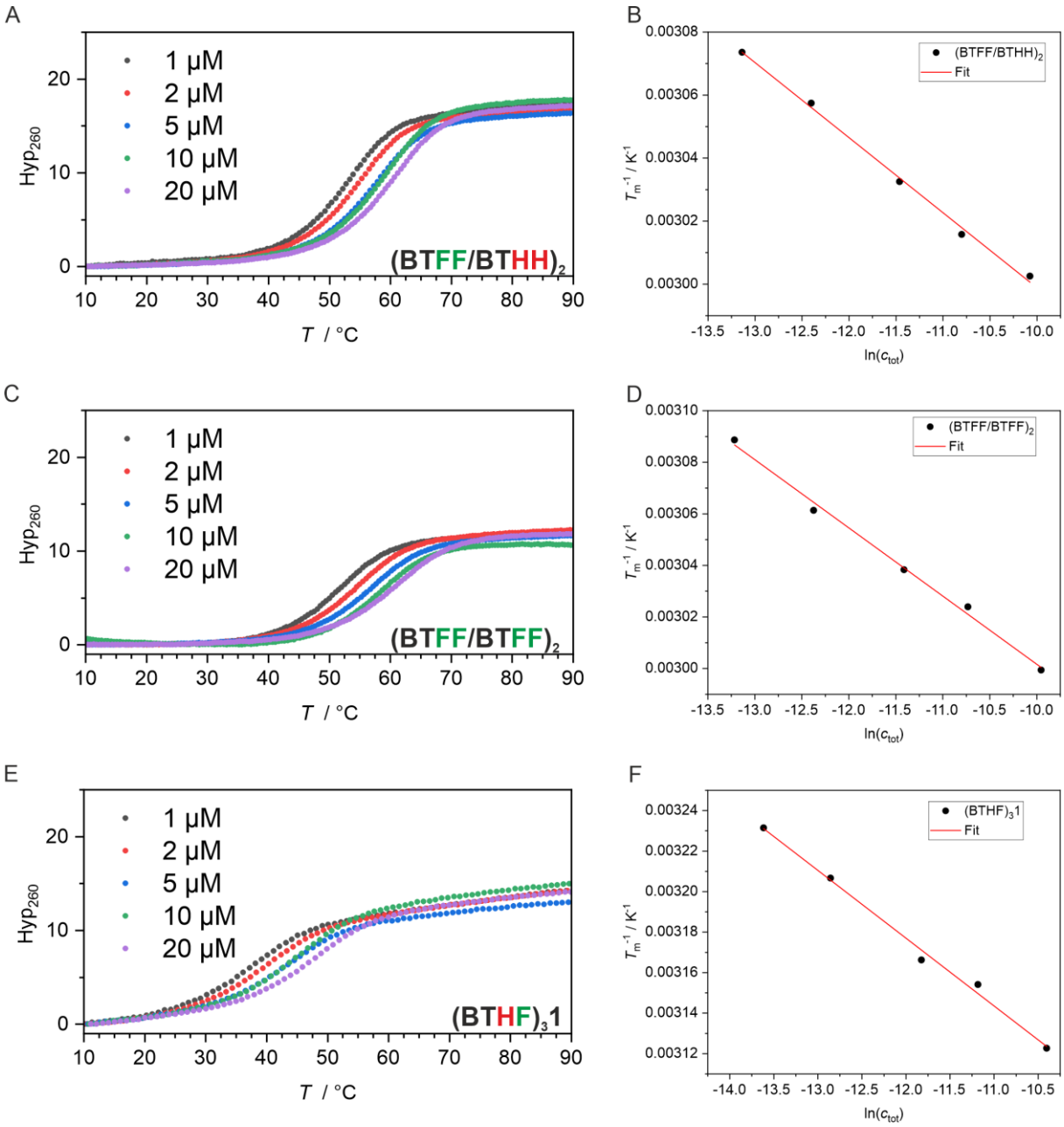


Figure S4.24 UV-thermal denaturation curves at different duplex concentrations in phosphate buffer (100 mM NaCl, 10 mM phosphate, pH 7.0) and the corresponding van't Hoff plot for $(BTFF/BTHH)_2$ (A and B), $(BTFF/BTFF)_2$ and $(BTHF)_31$ (E and F).

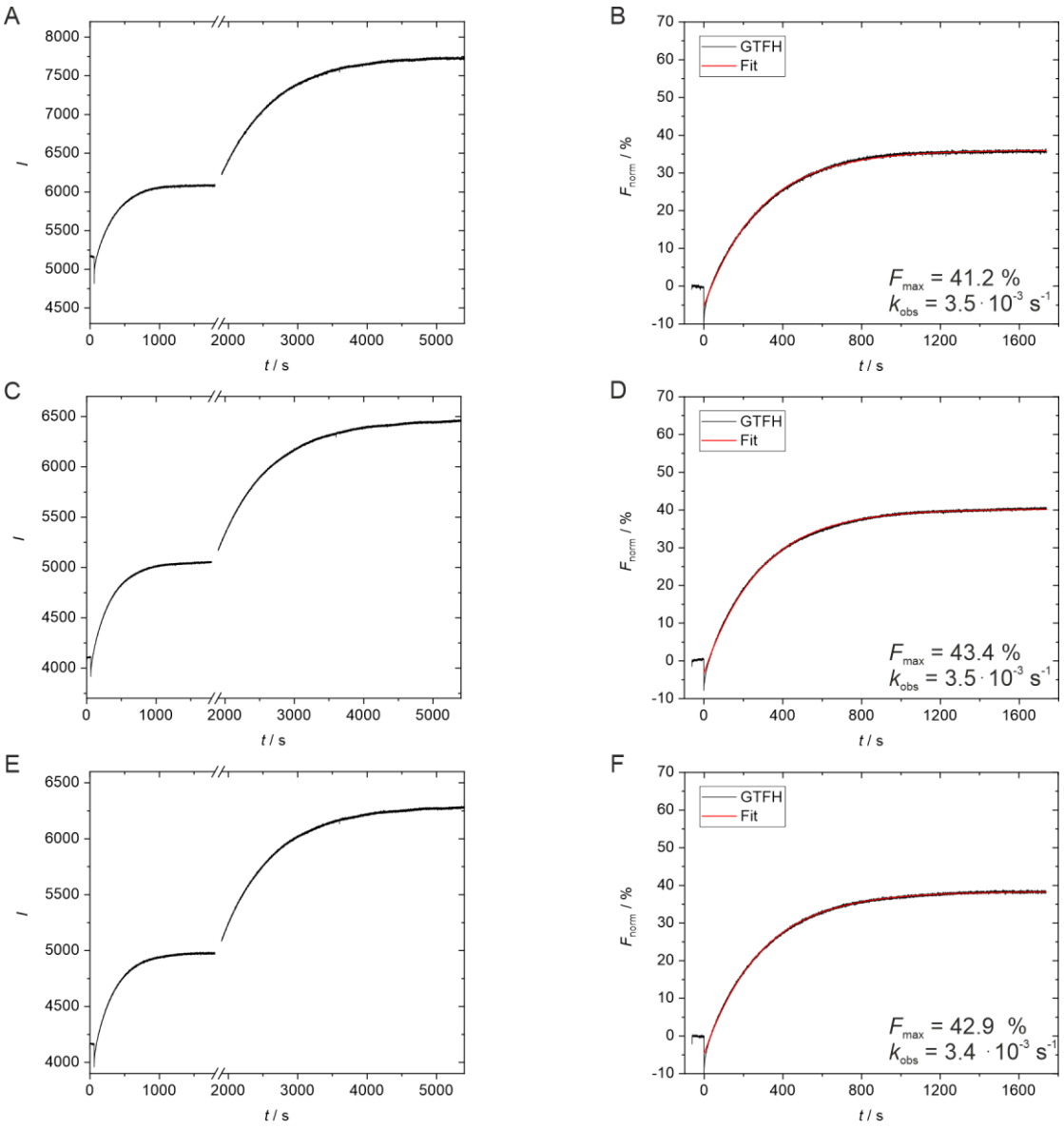


Figure S4.25 Replicates of the FRET exchange experiment with labeled GTHH/GTHH duplex and unlabeled GTFH single strand. (A), (C), and (E) show the change in Cy3 fluorescence observed for the DNA duplex upon strand displacement. The corresponding normalized curves are shown in (B), (D) and (F).

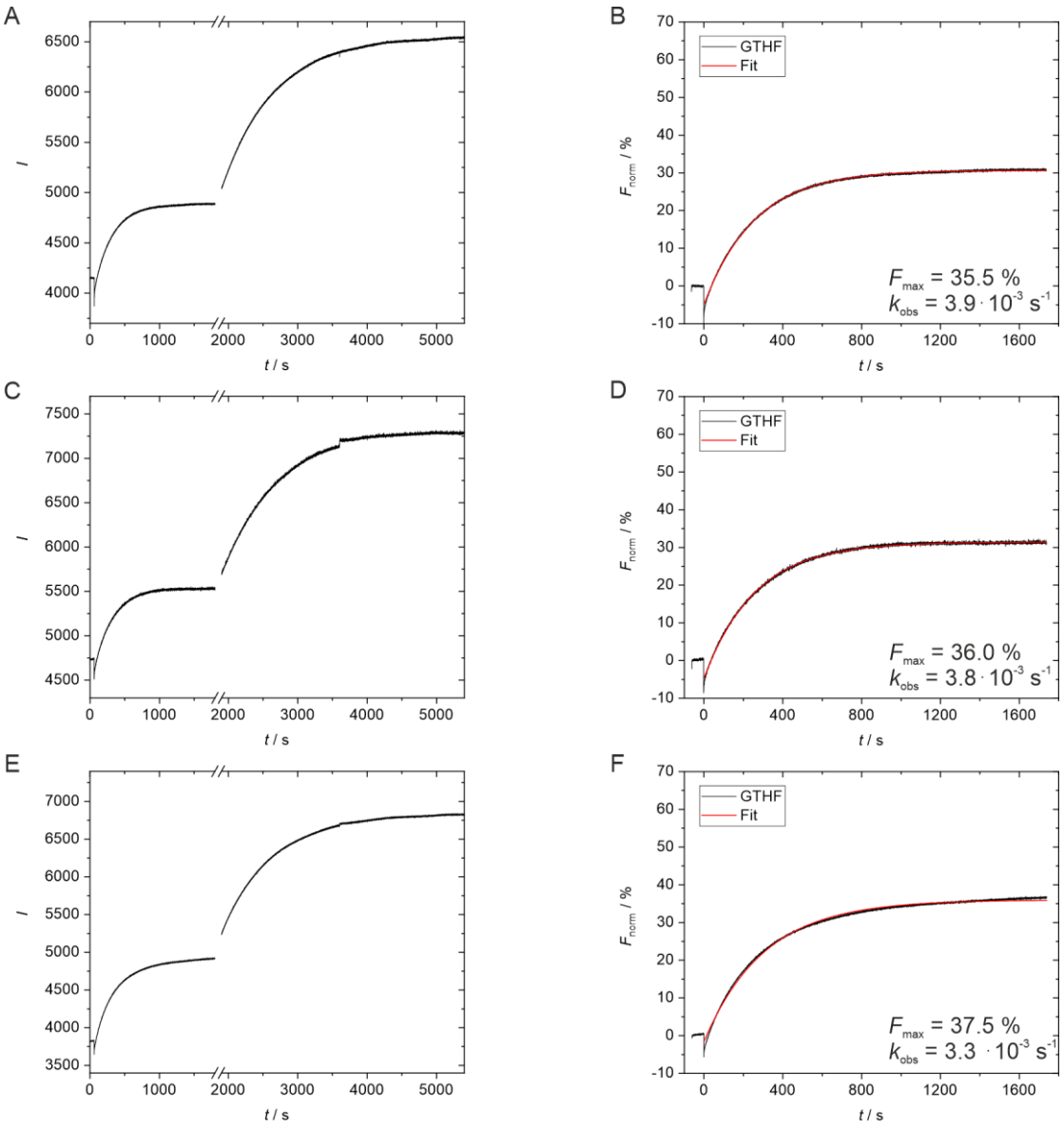


Figure S4.26 Replicates of the FRET exchange experiment with labeled GTHH/GTHH duplex and unlabeled GTHF single strand. (A), (C) and (E) show the change in Cy3 fluorescence observed for the DNA duplex upon strand displacement. The corresponding normalized curves are shown in (B), (D) and (F).

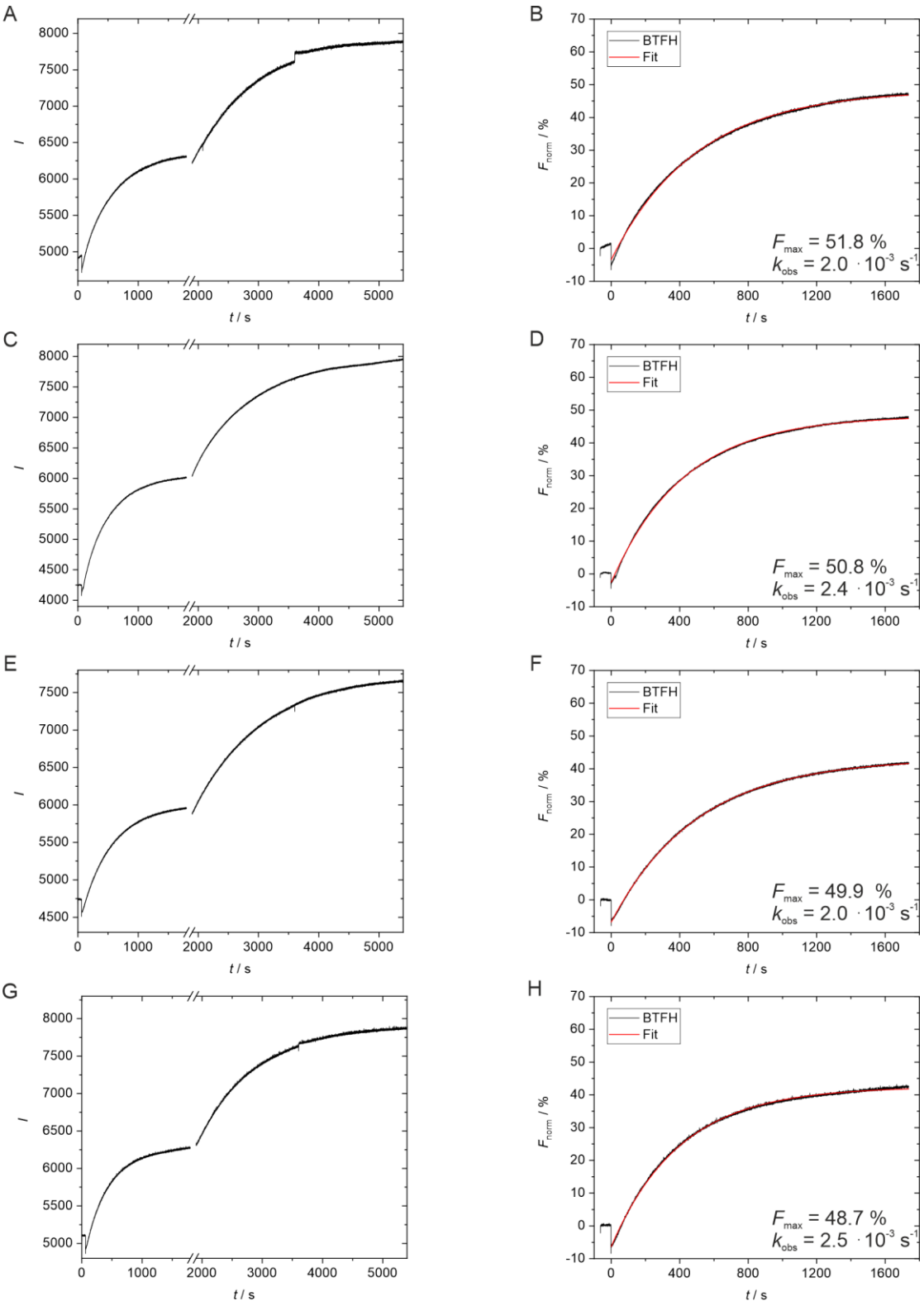


Figure S4.27 Replicates of the FRET exchange experiment with labeled BTHH/BTHH duplex and unlabeled BTFH single strand. (A), (C), (E) and (G) show the change in Cy3 fluorescence observed for the DNA duplex upon strand displacement. The corresponding normalized curves are shown in (B), (D), (F) and (H).

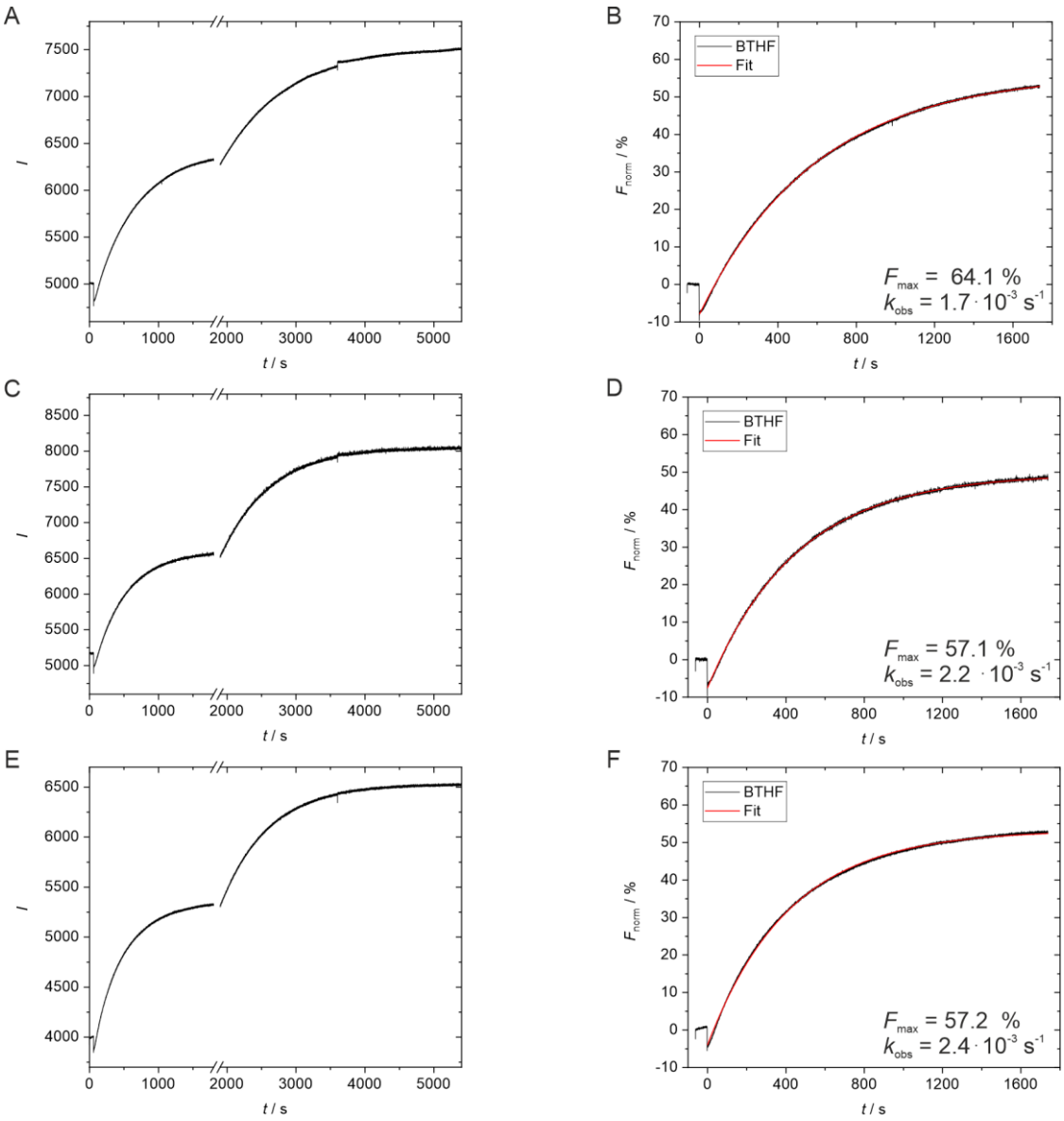


Figure S4.28 Replicates of the FRET exchange experiment with labeled BTHH/BTHH duplex and unlabeled BTHF single strand. (A), (C) and (E) show the change in Cy3 fluorescence observed for the DNA duplex upon strand displacement. The corresponding normalized curves are shown in (B), (D), and (F).

Cartesian coordinates of tolanes:

TFH

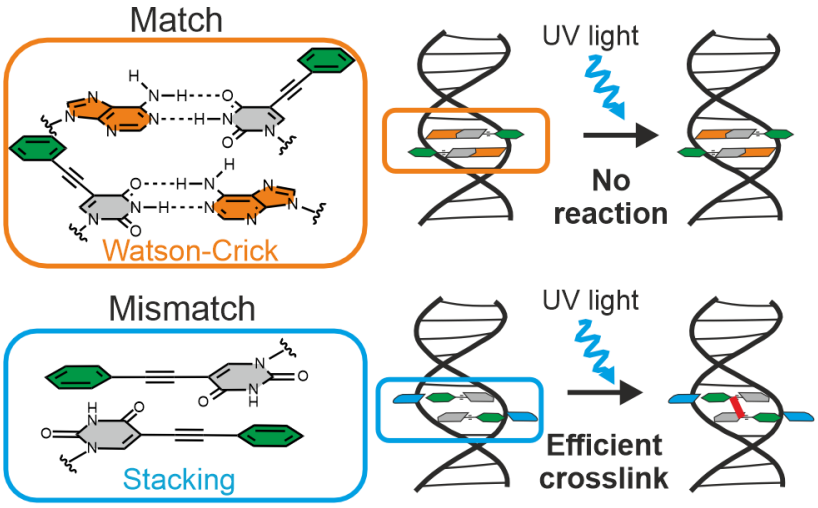
C	0.126156	2.838673	-1.032988
C	0.635411	1.585006	-0.671974
C	-0.201687	0.539188	-0.338967
C	-1.596359	0.677994	-0.332560
C	-2.096598	1.931735	-0.696087
C	-1.264337	2.980676	-1.047223
C	-2.437395	-0.398445	0.019045
C	-3.126790	-1.340883	0.323415
C	-3.920969	-2.462023	0.676315

C	-3.314444	-3.632852	1.156267
C	-4.088498	-4.731866	1.493588
C	-5.472540	-4.683324	1.358889
C	-6.082056	-3.525403	0.885730
C	-5.316863	-2.420494	0.546489
O	1.014287	3.787934	-1.382317
C	0.757666	5.156024	-1.030283
F	-1.820578	4.144140	-1.428350
F	-3.417169	2.131901	-0.726929
F	0.339676	-0.633216	0.001438
F	1.959179	1.402733	-0.646555
H	-2.237975	-3.664916	1.257164
H	-3.610540	-5.630364	1.862263
H	-6.073528	-5.544219	1.621727
H	-7.158513	-3.483537	0.780447
H	-5.785551	-1.518406	0.177139
H	1.724731	5.650424	-1.076786
H	0.066622	5.620022	-1.731745
H	0.357688	5.223499	-0.017152

THF

C	0.408272	2.866338	-0.736468
C	0.970222	1.646361	-0.344079
C	0.161842	0.564672	-0.068760
C	-1.236776	0.667145	-0.175205
C	-1.786197	1.891641	-0.570959
C	-0.978038	2.983558	-0.851212
C	-2.070361	-0.439946	0.115227
C	-2.785237	-1.380602	0.363436
C	-3.635313	-2.467988	0.652854
C	-3.132589	-3.748549	0.918278
C	-3.966796	-4.816669	1.200706
C	-5.342784	-4.628936	1.228017
C	-5.874269	-3.371988	0.968936
C	-5.027583	-2.314034	0.687203
O	1.284759	3.870126	-0.981308
C	0.776858	5.133178	-1.384766
F	-1.814867	-3.957194	0.900170
F	-3.458725	-6.025799	1.445682
F	-6.149631	-5.652075	1.502182
F	-7.196078	-3.191638	0.993990
F	-5.563538	-1.116614	0.443840
H	2.046467	1.575196	-0.262493
H	0.600521	-0.376182	0.234838
H	-2.860429	1.983981	-0.657964
H	-1.434621	3.913663	-1.154432
H	1.644524	5.773792	-1.520435
H	0.229263	5.057243	-2.328469
H	0.124183	5.561320	-0.618813

5 Programmable DNA interstrand crosslinking by alkene-alkyne [2+2] photocycloaddition



This chapter and the corresponding supporting information were published in:

H. Neitz, I. Bessi, J. Kuper, C. Kisker, C. Höbartner, *J. Am. Chem. Soc.* **2023**, *145*, 9428–9433. (https://pubs.acs.org/doi/10.1021/jacs.3c01611)

Published by American Chemical Society.

Abstract: Covalent crosslinking of DNA strands provides a useful tool for medical, biochemical, and DNA nanotechnology applications. Here we present a light-induced interstrand DNA crosslinking reaction using the modified nucleoside 5-phenylethynyl-2'-deoxyuridine (^{Phe}dU). The crosslinking ability of ^{Phe}dU was programmed by base pairing and by metal ion interaction at the Watson–Crick base pairing site. Rotation to intrahelical positions was favored by hydrophobic stacking and enabled an unexpected photochemical alkene-alkyne [2+2] cycloaddition within the DNA duplex, resulting in efficient formation of a ^{Phe}dU dimer after short irradiation times of a few seconds. A ^{Phe}dU-dimer-containing DNA was shown to efficiently bind a helicase complex, but the covalent crosslink completely prevented DNA unwinding, suggesting possible applications in biochemistry or structural biology.

5.1 Introduction

The DNA double helix is an outstanding scaffold for the guided arrangement of functional groups. Precise control of chromophores or reactive groups in DNA can lead to functional materials that offer new opportunities for chemical biology and nanotechnology.[31, 71, 454-455] The formation of a covalent bond between correctly oriented functional groups in single strands leads to the formation of an interstrand crosslink (ICL).[456] ICLs are of interest in biology because they inhibit replication and transcription, leading to cellular toxicity.[457] Examples of small-molecule-induced ICL are formed by anticancer drugs such as cisplatin[458] and nitrogen mustards.[459] Other common strategies are based on functional groups connected via disulfide[460] or imine formation,[461-462] nucleophilic reactions,[463-465] or click reactions.[466-467] Applications aiming at the stabilization of DNA nanostructures[185] often use photochemical reactions that provide an excellent level of control for ICL formation in space and time.[374, 379, 387, 468-470] Prominent examples include classical [2+2] cycloadditions, such as the formation of thymidine dimers,[372-373, 471] fast and reversible photocycloadditions between 3-cyanovinylcarbazole[382, 384-385, 472] or pyranocarbazole[383] with pyrimidine nucleotides, and photochemical cyclobutane formation between modified styryl units.[391, 473] Crosslinking by [2+2+2] cycloadditions of two alkynes with oxygen was used by Kool *et al.* for photoswitching of oligodeoxyfluoroside dyes[388] and by Nagatsugi and co-workers to implement an alkyne-alkyne photo-crosslinker using alkyne-modified pyridone groups as nucleobase replacement.[389]

5.2 Results and discussion

5.2.1 Design of the study

Here we explored 5-phenylethynyl-2'-deoxyuridine (^{Phe}dU) (see Figure 5.1A) as building block for programmable DNA interstrand crosslinking. The ^{Phe}dU nucleoside has previously been incorporated into DNA to explore its fluorescence response upon hybridization[474] and to modulate electron transport in DNA.[475] Upon photoexcitation of the extended π system, we expect different reactivity of the ^{Phe}dU moiety depending on its orientation in matched or mismatched base pairs (see Figure 5.1B).

The ^{Phe}dU nucleoside, which showed a UV absorption maximum at 306 nm (see Figure S5.1), was incorporated into DNA oligonucleotides by solid-phase synthesis using phosphoramidite chemistry (see Scheme S5.1). The ^{Phe}dU-containing duplex showed a UV absorption band at 320 nm (see Figure 5.1C and Figure S5.2). Duplex stability was measured using UV melting

curves recorded at 260 nm. The base pairing between ^{Phe}dU and adenine (T_m = 40.7 °C) resulted in a slight stabilization compared to the unmodified A/T base pair (T_m = 40.0 °C). Other nucleotides or an abasic site mimic opposite to ^{Phe}dU strongly destabilized the duplex (ΔT_m > 10 °C; see Figure 5.1D).

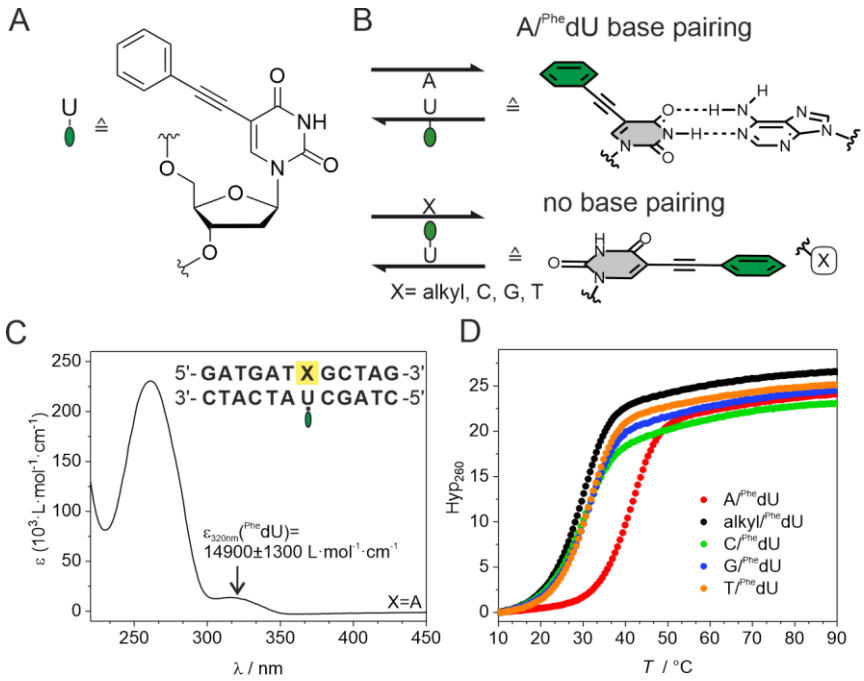


Figure 5.1 (A) Structure of ^{Phe}dU. (B) Concept of different orientations of ^{Phe}dU in Watson–Crick base pair (top) versus mismatched situation (bottom). (C) UV absorption spectrum of a DNA duplex containing the ^{Phe}dU modification (X = A). (D) UV melting curves of DNA duplexes containing different nucleobases or an alkyl linker opposite of PhedU (X = A, C, G, T, or *n*-propyl), 1 µM DNA in sodium phosphate buffer, pH 7.0.

We designed a duplex containing two ^{Phe}dU nucleosides in neighboring base pairs by placing one ^{Phe}dU in each single strand (see Figure 5.2A). Upon formation of Watson-Crick base pairs with adenine (X = A), the hydrophobic moieties of the two ^{Phe}dU residues point toward the major groove and are located too far apart to undergo any photoreaction with each other (see Figure 5.2B). In contrast, when the Watson-Crick face is not occupied by hydrogen bonding (*i.e.*, in the mismatched situation with X = C, G, T), the preferred orientation may be flipped in order to increase aromatic stacking interactions of the hydrophobic moiety.[443, 476] Irradiation with UV light could then result in a photoaddition between the modifications, forming an ICL between the two strands (see Figure 5.2C).

5.2.2 Programmability of the crosslinking reaction

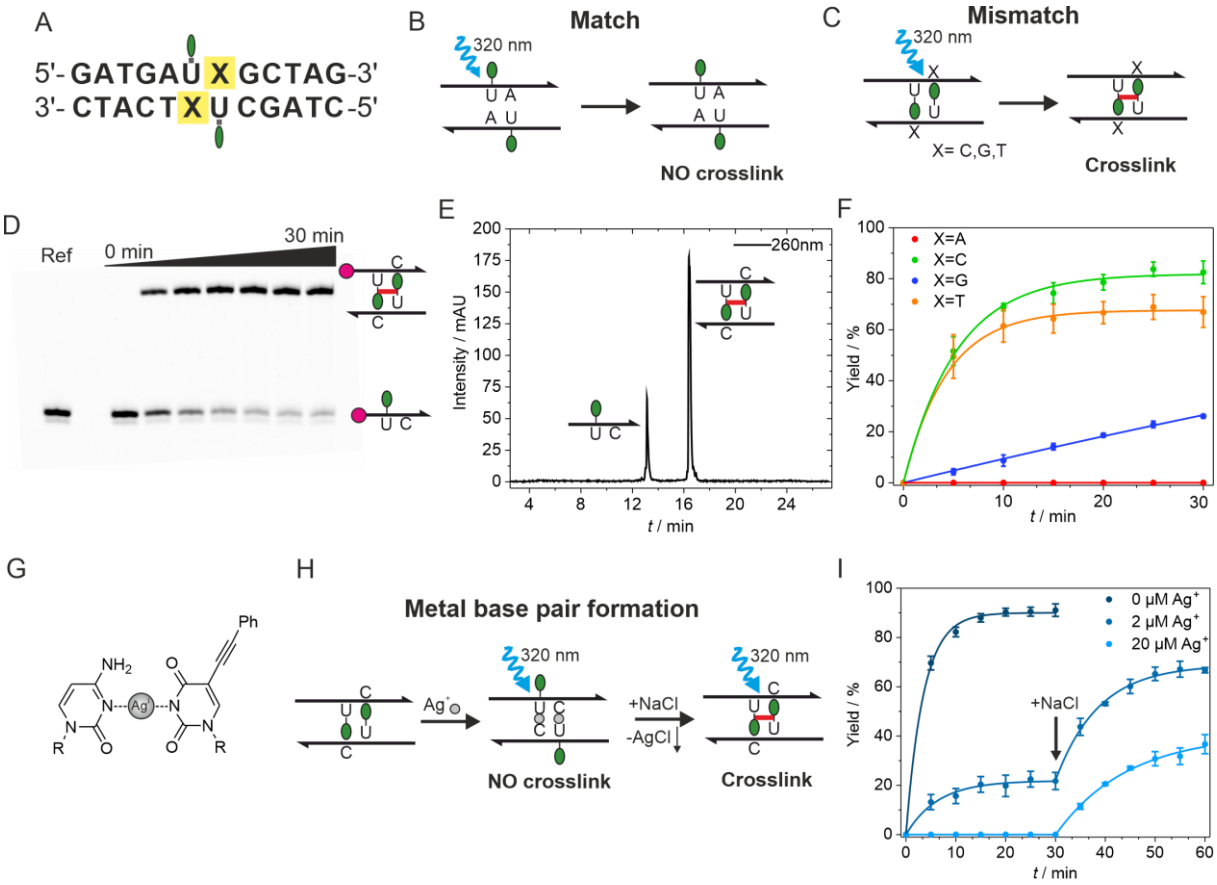


Figure 5.2 (A) DNA duplex used for ICL (X = A, C, G, T). (B, C) Schematic representations of the (B) match (X = A) and (C) mismatch (X = C, G, T) situations. (D, E) Exemplary (D) PAGE and (E) anion exchange HPLC analyses of ICL formation with the ^{Phe}dU/dC mismatch. Cy3 is depicted as magenta circle. (F) Kinetics of ICL formation (1 µM DNA in 10 mM sodium phosphate, 100 mM NaCl, pH 7.0). (G) Silver-mediated metal base pair between dC and ^{Phe}dU. (H) Schematic representation of ICL inhibition by metal base pair formation and reactivation by precipitation of Ag^+ as AgCl. (I) Kinetics of ICL formation at different [Ag^+] (1 µM DNA in 10 mM MOPS, 100 mM $NaClO_4$, pH 7.1).

Indeed, we observed ICL formation upon irradiation of the mismatched duplex samples but not when the Watson-Crick base pair was intact. The kinetics of the crosslink formation was measured by denaturing PAGE using 5'-Cy3-labeled DNA (see Figure 5.2D). In addition, the experiment was performed without a Cy3 fluorophore and analyzed by anion exchange HPLC to show that the crosslinking is independent of the cyanine dye (see Figure 5.2E). Best ICL efficiency was observed when ^{Phe}dU was placed in mismatches with pyrimidine nucleotides. The duplex containing ^{Phe}dU–deoxycytidine mismatches reached a conversion of about 83% after 30 min of UV irradiation in a stirred quartz fluorescence microcell in a fluorescence spectrometer at 15 °C. The yield for the thymidine-^{Phe}dU mismatches was reduced to about 67%, and the ^{Phe}dU-dG wobble base pairs gave only 26% ICL (see Figure 5.2F and Figure S5.3). A transilluminator as light source resulted in faster crosslinking (see Figure S5.4), but the overall

ICL yield was lower, likely due to the low T_m of the double-mismatched 12-mer duplexes (see Figure S5.5). When the duplexes were extended to 20-mers, only 30 s of irradiation on the transilluminator was required to achieve more than 80% ICL yield (see Figure S5.6).

To establish an additional level of control, we turned our attention to metal-ion-mediated stabilization of pyrimidine–pyrimidine base pairs.[338] We hypothesized that binding a transition metal between ^{Phe}dU and dC or dT would keep the phenylacetylene group in the major groove and prevent crosslinking, while absence or removal of the metals would favor the reactive stacked orientation. We screened a series of metal ions and examined the crosslinking efficiency by denaturing PAGE (see Figure S5.7). The majority of tested transition metals did not show any effect, but the presence of silver ions clearly inhibited the reaction (see Figure 5.2G-I). The addition of 2 µM Ag^+ to the dC/^{Phe}dU sample increased the T_m by 4.3 °C (see Figure S5.8) and slowed down the crosslink reaction, and 20 µM Ag^+ completely blocked the formation of the ICL. This observation suggests that ^{Phe}dU undergoes a similar interaction with Ag^+ as in the known metal base pair of thymidine and cytidine.[336] Upon addition of sodium chloride, the Ag^+ was removed by precipitation as silver chloride, and the ability to form the ICL was recovered (see Figure 5.2H and Figure S5.7).

5.2.3 Characterization of the crosslinking reaction

Since the UV-induced reaction between two ^{Phe}dU was unprecedented, we studied the nature of the crosslinked product in detail. The ICL-containing duplex was isolated in preparative scale and characterized by mass spectrometry and NMR spectroscopy. ESI-MS showed that the irradiated duplex consists of the two ^{Phe}dU-modified single strands and does not contain any addition of another compound (see Figure S5.9). Considering the previously reported [2+2+2] cycloaddition of alkyne-modified DNA,[388-389] the incorporation of O_2 could have been expected but was not observed. Temperature-dependent NMR spectroscopy of the crosslinked 12-mer duplex showed that the ICL provided a thermal stabilization of the duplex, which was also confirmed by UV melting curve analysis of the crosslinked 20-mer duplex, which showed a stabilization by 22 °C (see Figure S5.10). The NMR spectra indicated significant signal overlap (see Figure S5.11), which makes assignment of the crosslink resonances challenging and determination of the structure almost impossible. Therefore, the duplex was enzymatically digested, and the resulting nucleotide mixture was analyzed by LC/MS. Besides the natural

nucleosides dC, dG, dT, and dA, two almost coeluting new products were observed. The extracted ion chromatogram confirmed that both have the same molecular mass, originating from the addition of two ^{Phe}dU units, without any incorporation of oxygen (see Figure 5.3B and Figure S5.12). The peak corresponding to the ICL unit was isolated by RP-HPLC, and the product was fully characterized. The NMR spectra showed two sets of peaks due to the formation of two isomers. Complete NMR assignment was obtained (see Table S5.3 and Figure 5.3C and Figure S5.13-S5.15), and the structure determination revealed the formation of a cyclobutene linkage between the ^{Phe}dU moieties. The photocycloaddition occurred between the C5–C6 double bond of one ^{Phe}dU and the alkyne of the ^{Phe}dU unit in the other strand (see Figure 5.3A).

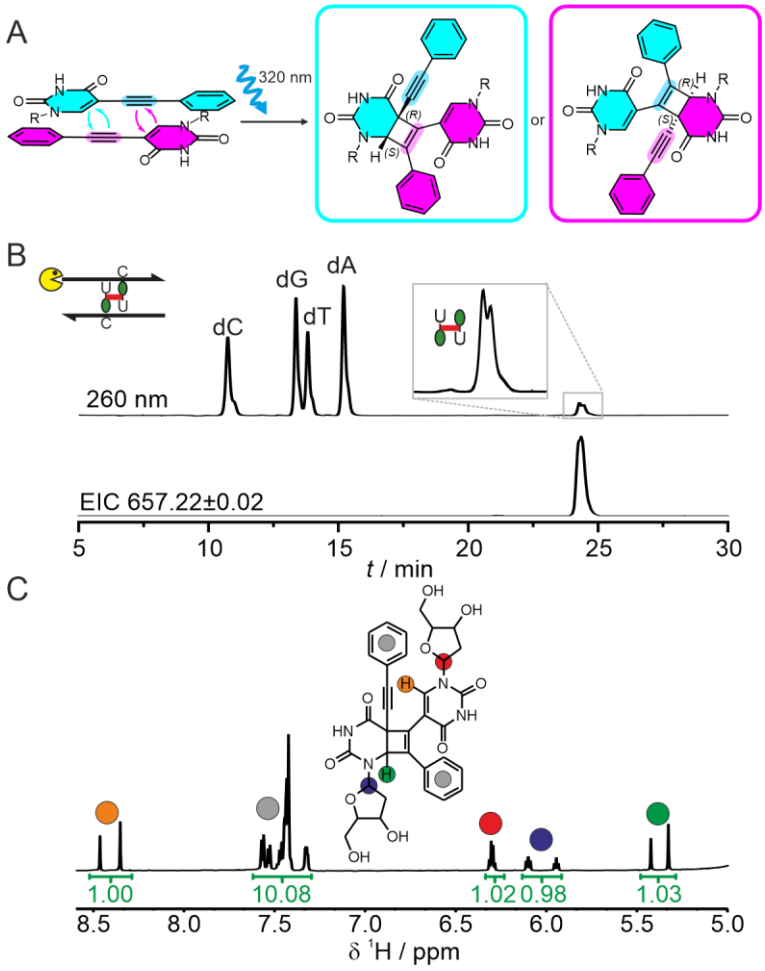


Figure 5.3 (A) ^{Phe}dU dimerization by photochemical alkene-alkyne [2+2] cycloaddition. (B) LC/MS analysis of the enzymatically digested DNA duplex containing the ICL. Extracted ion chromatogram (EIC) of ^{Phe}dU dimer. (C) Structure of the crosslinked product with assignment of the most relevant protons in the ^{1}H NMR spectrum (in D_2O, 600 MHz, 25 °C).

The stacking arrangement of the two ^{Phe}dU units in the duplex places two alkene-alkyne pairs in a symmetrical orientation, which leads to the formation of two diastereomeric ^{Phe}dU dimers

in almost equimolar ratio. The formation of substituted cyclobutenes by photoinduced [2+2] cycloaddition of alkynes and alkenes has been described for various types of activated alkenes with electron-rich alkynes,[477-479] but the reaction has not been observed or used in the context of DNA.

The NMR assignment was fully consistent with the cyclobutene ^{Phe}dU dimer formation and the presence of two intact phenyl rings. Nevertheless, other possible cycloadditions such as photochemical hexadehydro-Diels-Alder reactions[480] or radical cyclization reactions were also considered, which could potentially involve the ortho positions of the phenyl rings. To demonstrate that the phenyl groups were not directly involved in the crosslink reaction, analogous studies were explored with 5-(pyrimidin-2-yl-ethynyl)-2'-deoxyuridine (^{Pym}dU). The corresponding ^{Pym}dU phosphoramidite was synthesized and incorporated into DNA strands (see Scheme S5.2). The results of the crosslinking reaction with ^{Pym}dU-modified strands were comparable to the results with ^{Phe}dU, as observed by PAGE shift and HPLC of the crosslinked duplex and further confirmed by LC/MS and ^{1}H NMR of the ^{Pym}dU dimer after digestion (see Figure S5.16). These results showed that the phenyl group is not directly involved in the reaction and can be replaced by another aromatic/heteroaromatic residue that maintains similar hydrophobic interactions and π–π stacking geometry. Together with the results of the mixed ^{Phe}dU/^{Pym}dU duplex (see Figure S5.17), these observations support the alkene-alkyne [2+2] photocycloaddition reaction as a plausible mechanism for ^{Phe}dU and ^{Pym}dU dimerization as novel type of DNA interstrand crosslink.

5.2.4 XPD helicase inhibition

ICLs are DNA damages that prevent unwinding of the DNA duplex and inhibit replication and transcription.[481-482] We expect the ^{Phe}dU dimer to be recognized as DNA damage and at the same time to act as a roadblock for DNA helicases. Thus, ^{Phe}dU could serve as tool for biochemical and biophysical characterization of DNA damage response and DNA repair *in vitro*. To explore this possibility, we incorporated ^{Phe}dU into the stem region of an open fork DNA substrate (Y-fork) suitable for helicase assays with 5' to 3' polarity, such as the nucleotide excision repair helicase XPD[483] and its complex with the structural subunits p44 and p62 of TFIIH.[484-485] The crosslinked version of the Y-fork was formed by UV irradiation (CL Y-fork) (see Figure S5.18), and the PAGE-purified DNA was used for analysis of the interaction with

proteins. The binding affinity was examined by fluorescence anisotropy and compared to unmodified and ^{Phe}dU containing DNA Y-forks before crosslinking. All three variants showed similar K_D values in the nM range (see Figure 5.4A).

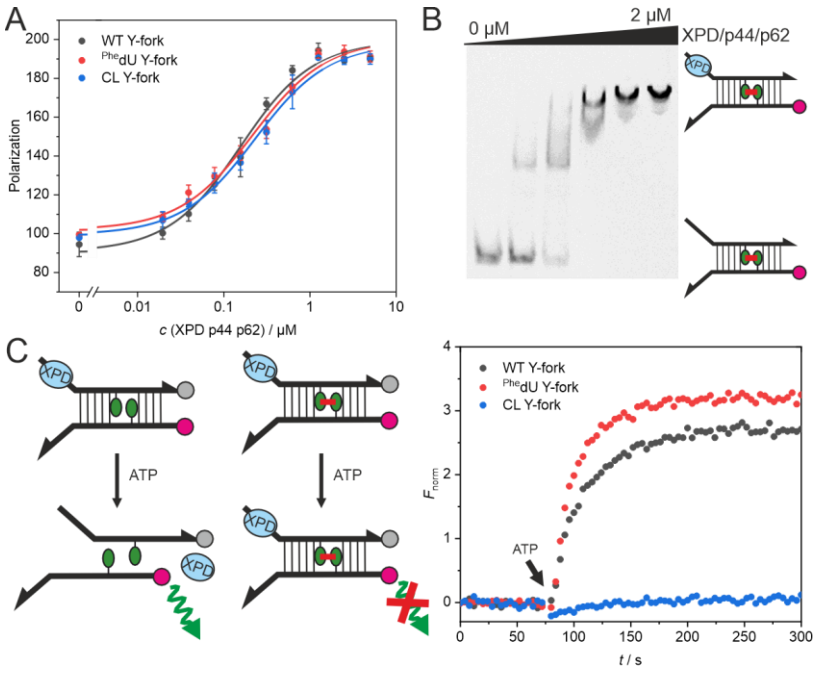


Figure 5.4 (A) Fluorescence anisotropy and (B) electrophoretic mobility shift assay showing comparable XPD/p44/p62 binding of unmodified and ^{Phe}dU-containing Y-fork DNA before and after crosslinking. Cy3 label (magenta circle). (C) Helicase assay using Cy3 and DABCYL (gray circle)-labeled DNA substrate upon addition of ATP, showing efficient inhibition of unwinding by the crosslinked ^{Phe}dU dimer.

In addition, formation of the protein–DNA complex was confirmed by an electrophoretic mobility shift assay (see Figure 5.4B and Figure S5.19). Helicase activity was measured using a fluorophor/quencher system with DABCYL added at the 3' end and Cy3 located at the 5' end of the second strand. After formation of the protein-DNA complexes, helicase activity was initiated by addition of ATP. An increase in Cy3 emission was observed for the unmodified and ^{Phe}dU-containing Y-fork, consistent with unwinding of the duplex by the helicase. In strong contrast, no change in Cy3 emission was detected in the case of the crosslinked Y-fork, indicating effective inhibition of helicase activity (see Figure 5.4C and Figure S5.20-S5.21). These results demonstrate that the ^{Phe}dU dimer ICL can be effectively used to lock helicase progression on DNA *in vitro*.

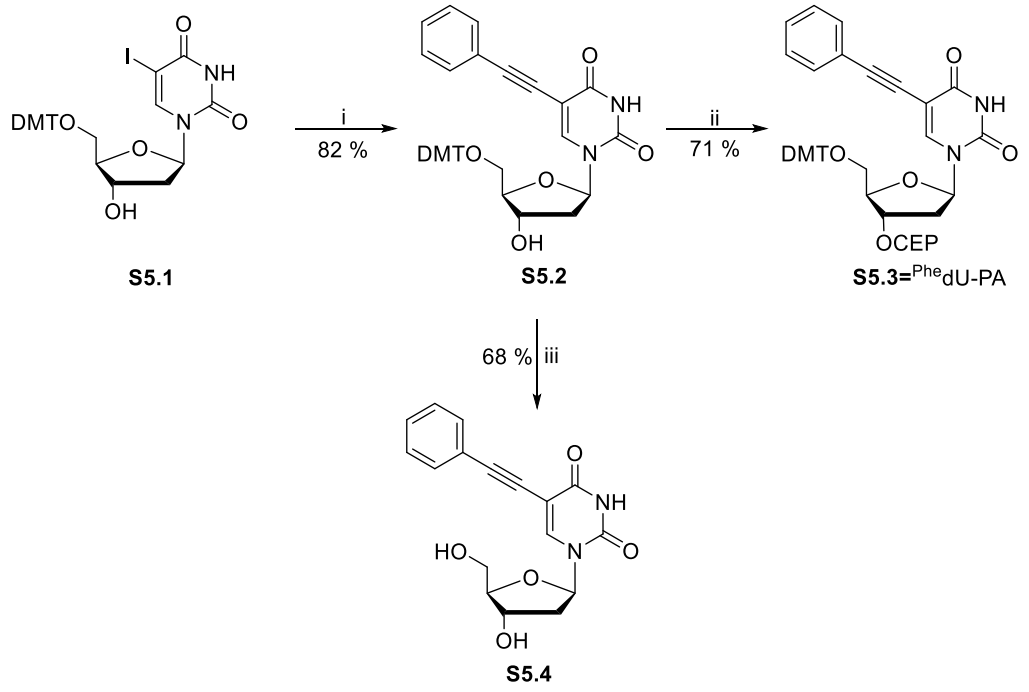
5.3 Conclusion

In summary, we have presented a new effective and fast photoinduced interstrand crosslink based on an unexpected alkene-alkyne [2+2] cycloaddition using the known ^{Phe}dU nucleoside.

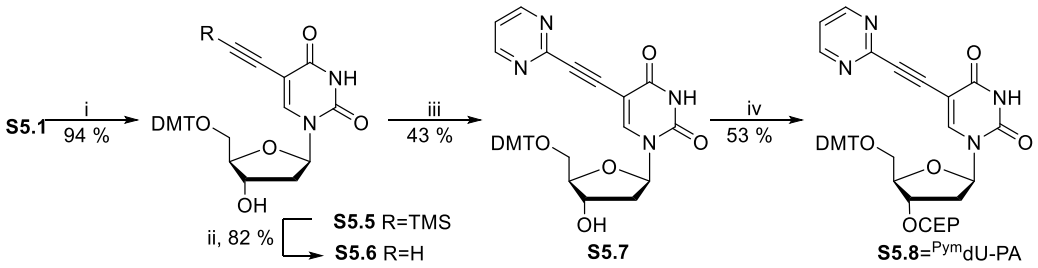
We have demonstrated that the mutual orientation of two ^{Phe}dU units in a DNA duplex can be switched between inactive and crosslink-competent conformations. The programmability was exploited by Watson-Crick hydrogen bonding or metal-mediated base pairing, in contrast to hydrophobic stacking, to control the DNA crosslink formation. The modified nucleoside is easily accessible and can be precisely incorporated into DNA. Therefore, we expect the new photochemical crosslink to be used for mechanistic and structural studies of DNA-processing enzymes or for the creation of DNA-based nanostructures.

5.4 Supporting Information

5.4.1 Additional data



Scheme S5.1 Synthesis of the ^{Phe}dU phosphoramidite. i) Phenylacetylene, CuI, $Pd(PPh_3)_4$, Et_3N/DMF, 60 °C, 23 h; ii) CEP-Cl, DIPEA, DCM, r.t., 3 h; iii) TFA, DCM, r. t.,15 min.



Scheme S5.2 Synthesis of the ^{Pym}dU phosphoramidite. i) Trimethylsilylacetylene, CuI, $Pd(PPh_3)_2Cl_2$, Et_3N/THF, 40 °C, 23 h; ii) K_2CO_3, THF/MeOH, r.t., 16 h; iii.) 2-bromopyrimidine, CuI, $Pd(PPh_3)_2Cl_2$, Et_3N/THF, 50 °C, 20 h; iv) CEP-Cl, DIPEA, DCM, r.t., 3 h.

Table S5.1 Sequences and ESI-MS results of the used DNA oligonucleotides.

Oligo	5'-Sequence-3'	Chemical formula	Calc. Mass	Measured Mass
ODN1_A	GATGATAGCTAG	$C_{119}H_{148}N_{49}O_{69}P_{11}$	3707.66977	3707.69796
ODN1_C3	GATGAT(C3)GCTAG	$C_{112}H_{143}N_{44}O_{68}P_{11}$	3532.61981	3532.62184
ODN1_C	GATGATCGCTAG	$C_{118}H_{148}N_{47}O_{70}P_{11}$	3683.65853	3683.68621
ODN1_G	GATGATGGCTAG	$C_{119}H_{148}N_{49}O_{70}P_{11}$	3723.66468	3723.66893
ODN1_T	GATGATTGCTAG	$C_{119}H_{149}N_{46}O_{71}P_{11}$	3698.65820	3698.70495
^{Phe}dU1_A	GATGA(^{Phe}dU)AGCTAG	$C_{126}H_{150}N_{49}O_{69}P_{11}$	3793.68487	3793.68611
Alk-^{Phe}dU1_A	Alk-GATGA(^{Phe}dU)AGCTAG	$C_{132}H_{159}N_{49}O_{72}P_{12}$	3953.71380	3953.72620
^{Phe}dU2_A	CTAGC(^{Phe}dU)ATCATC	$C_{123}H_{151}N_{40}O_{71}P_{11}$	3664.65486	3664.66590
^{Phe}dU1_C	GATGA(^{Phe}dU)CGCTAG	$C_{125}H_{150}N_{47}O_{70}P_{11}$	3769.67363	3769.67487
Alk-^{Phe}dU1_C	Alk-GATGA(^{Phe}dU)CGCTAG	$C_{131}H_{159}N_{47}O_{73}P_{12}$	3929.70257	3929.69139
^{Phe}dU2_C	CTAGC(^{Phe}dU)CTCATC	$C_{122}H_{151}N_{38}O_{72}P_{11}$	3640.64362	3640.64008
^{Phe}dU1_G	GATGA(^{Phe}dU)GGCTAG	$C_{126}H_{150}N_{49}O_{70}P_{11}$	3809.67978	3809.67368
Alk-^{Phe}dU1_G	Alk-GATGA(^{Phe}dU)GGCTAG	$C_{132}H_{159}N_{49}O_{73}P_{12}$	3969.70871	3969.70759
^{Phe}dU2_G	CTAGC(^{Phe}dU)GTCATC	$C_{123}H_{151}N_{40}O_{72}P_{11}$	3680.64977	3680.64948
^{Phe}dU1_T	GATGA(^{Phe}dU)TGCTAG	$C_{126}H_{151}N_{46}O_{71}P_{11}$	3785.67603	3785.66950
Alk-^{Phe}dU1_T	Alk-GATGA(^{Phe}dU)TGCTAG	$C_{132}H_{160}N_{46}O_{74}P_{12}$	3944.70223	3944.70224
^{Phe}dU2_T	CTAGC(^{Phe}dU)TTCATC	$C_{123}H_{153}N_{37}O_{73}P_{11}$	3655.64329	3655.64463
^{Pym}dU1_C	GATGA(^{Pym}dU)CGCTAG	$C_{123}H_{148}N_{49}O_{70}P_{11}$	3771.66413	3771.66284
Alk-^{Pym}dU1_C	Alk-GATGA(^{Pym}dU)CGCTAG	$C_{129}H_{157}N_{49}O_{73}P_{12}$	3931.69306	3931.70221
^{Pym}dU2_C	CTAGC(^{Pym}dU)CTCATC	$C_{120}H_{149}N_{40}O_{72}P_{11}$	3642.63412	3642.62381
^{Phe}dU1_C_20mer	CTGTGATGA(^{Phe}dU)CGCTAGACCT	$C_{202}H_{249}N_{72}O_{120}P_{19}$	6191.06099	6191.05430
Alk-^{Phe}dU1_C_20mer	Alk-CTGTGATGA(^{Phe}dU)CGCTAGACCT	$C_{208}H_{258}N_{72}O_{123}P_{20}$	6351.08992	6351.13736
^{Phe}dU2_C_20mer	AGGTCTAGC(^{Phe}dU)CTCATCACAG	$C_{201}H_{248}N_{73}O_{118}P_{19}$	6160.06641	6160.05504
^{Phe}dU1_T_20mer	CTGTGATGA(^{Phe}dU)TGCTAGACCT	$C_{203}H_{250}N_{71}O_{121}P_{19}$	6206.06065	6206.08814
^{Phe}dU2_T_20mer	AGGTCTAGC(^{Phe}dU)TTCATCACAG	$C_{202}H_{249}N_{72}O_{119}P_{19}$	6175.06607	6175.06742
Y-fork1	AGCTACCATGCCTGCACGAATTAAGCA(^{Phe}dU)CG CGTAATCATGGTCATAG-NH_2	$C_{471}H_{593}N_{179}O_{282}P_{47}$	14728.53898	14728.62746
Y-fork2	Alk-CTATGACCATGATTACGC(^{Phe}dU)CTGCTTGGAAT CCTGACGAACTGTAGA	$C_{462}H_{577}N_{169}O_{678}P_{46}$	14369.42973	14369.58916

Table S5.2. Thermodynamic data for DNA duplexes in phosphate buffer (100 mM NaCl, 10 mM sodium phosphate, pH 7.0).

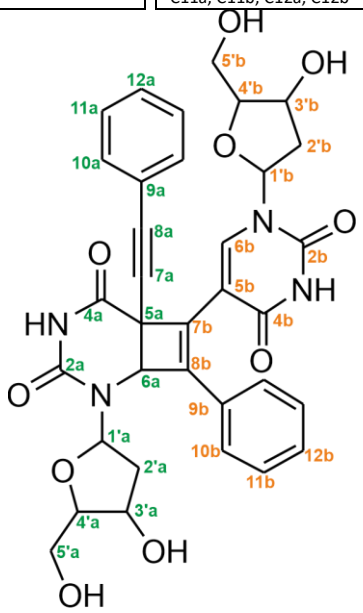
Duplex	Oligo	Sequence	c_{total}[a] [µM]	T_m [°C]	ΔH^0 [kcal mol^{-1}]	ΔS^0 [cal mol^{-1} K^{-1}]	ΔG^{298}[b] [kcal mol^{-1}]
dA/^{Phe}dU	ODN1_A ^{Phe}dU2_A	5'…TAG…3' 3'…A^{Phe}dUC…5'	2.3 4.7 11.8 22.0 46.6	40.7 42.3 44.4 45.6 47.5	-89.5±0.7	-257±2.1	-12.9±1.0
C3/^{Phe}dU	ODN1_C3 ^{Phe}dU2_A	5'…TC3G…3' 3'…A^{Phe}dUC…5'	2.2 4.3 10.9 20.8 43.4	28.7 30.4 32.7 34.3 36.2	-74.5±0.2	-219±0.5	-9.4±0.2
dC/^{Phe}dU	ODN1_C ^{Phe}dU2_A	5'…TCG…3' 3'…A^{Phe}dUC…5'	2.4 4.4 11.2 21.4 44.1	28.3 30.6 33.1 35.0 36.9	-64.2±1.3	-184±3.7	-9.3±1.7
dG/^{Phe}dU	ODN1_G ^{Phe}dU2_A	5'…TGG…3' 3'…A^{Phe}dUC…5'	2.2 4.4 11.1 22.0 44.3	30.2 31.9 34.2 35.8 37.5	-76.2±0.4	-223±1.1	-9.8±0.5
dT/^{Phe}dU	ODN1_T ^{Phe}dU2_A	5'…TTG…3' 3'…A^{Phe}dUC…5'	2.3 4.5 11.2 22.2 44.4	30.6 32.2 34.4 36.0 37.8	-78.2±0.1	-229±0.3	-10.0±0.1

[a] Total concentration of DNA strands. Values were calculated with the absorption at 260 nm of the melted duplex [b] Calculated for T = 25 °C.

Table S5.3 Complete NMR assignment of the two isomers of the ^{Phe}dU CL-unit. See structure below for nomenclature.

1st isomer	
Atom	ppm
H1‘a	5.94, t, *J*=7.0 Hz
H2‘a	2.47-2.41; 2.14-2.08
H3‘a	4.39-4.36
H4‘a	3.92-3.89
H5‘a	3.77-3.64
H6a	5.42, s
H10a	7.55-7.51
H1‘b	6.33-6.27
H2‘b	2.42-2.37, 2.26-2.20
H3‘b	4.27-4.23
H4‘b	4.02, q, *J*=4.6 Hz
H5‘b	3.81-3.77; 3.74-3.70
H6b	8.46, s
H10b	7.50-7.40
H11a, H11b, H12a, H12b	7.50-7.40
C1’a	86.9
C2’a	38
C3’a	70.3
C4’a	85.8
C5’a	61.5
C2a	152.5
C4a	171.8
C5a	50.2
C6a	58.0
C7a	81.8
C8a	87.9
C9a	120.4
C10a	131.9
C1’b	85.4
C2’b	38.0
C3’b	70.7
C4’b	86.9
C5’b	61.5
C2b	150.7
C4b	162.3
C5b	106.9
C6b	141.2
C7b	131.2
C8b	129.0
C9b	145.7
C10b	129
C11a, C11b, C12a, C12b	129

2nd isomer	
Atom	ppm
H1‘a	6.10, t, *J*=7.6 Hz
H2‘a	2.21-2.15; 2.04-1.99
H3‘a	4.31-4.28
H4‘a	3.90-3.87
H5‘a	3.77-3.64
H6a	5.33, s
H10a	7.58-7.55
H1‘b	6.33-6.27
H2‘b	2.46-2.41, 2.35-2.28
H3‘b	4.31-4.28
H4‘b	3.94-3.90
H5‘b	3.56, dd, *J*=12.3, 4.5 Hz; 3.41, dd, *J*=12.2, 6.0 Hz
H6b	8.35, s
H10b	7.35-7.30
H11a, H11b, H12a, H12b	7.50-7.40
C1’a	86.5
C2’a	38.0
C3’a	70.7
C4’a	85.8
C5’a	61.5
C2a	153.1
C4a	171.8
C5a	49.5
C6a	58.7
C7a	82.0
C8a	87.9
C9a	120.4
C10a	131.9
C1’b	85.4
C2’b	38.0
C3’b	70.7
C4’b	85.8
C5’b	61.5
C2b	150.7
C4b	162.3
C5b	106.5
C6b	140.9
C7b	131.9
C8b	129
C9b	145.7
C10b	128
C11a, C11b, C12a, C12b	129



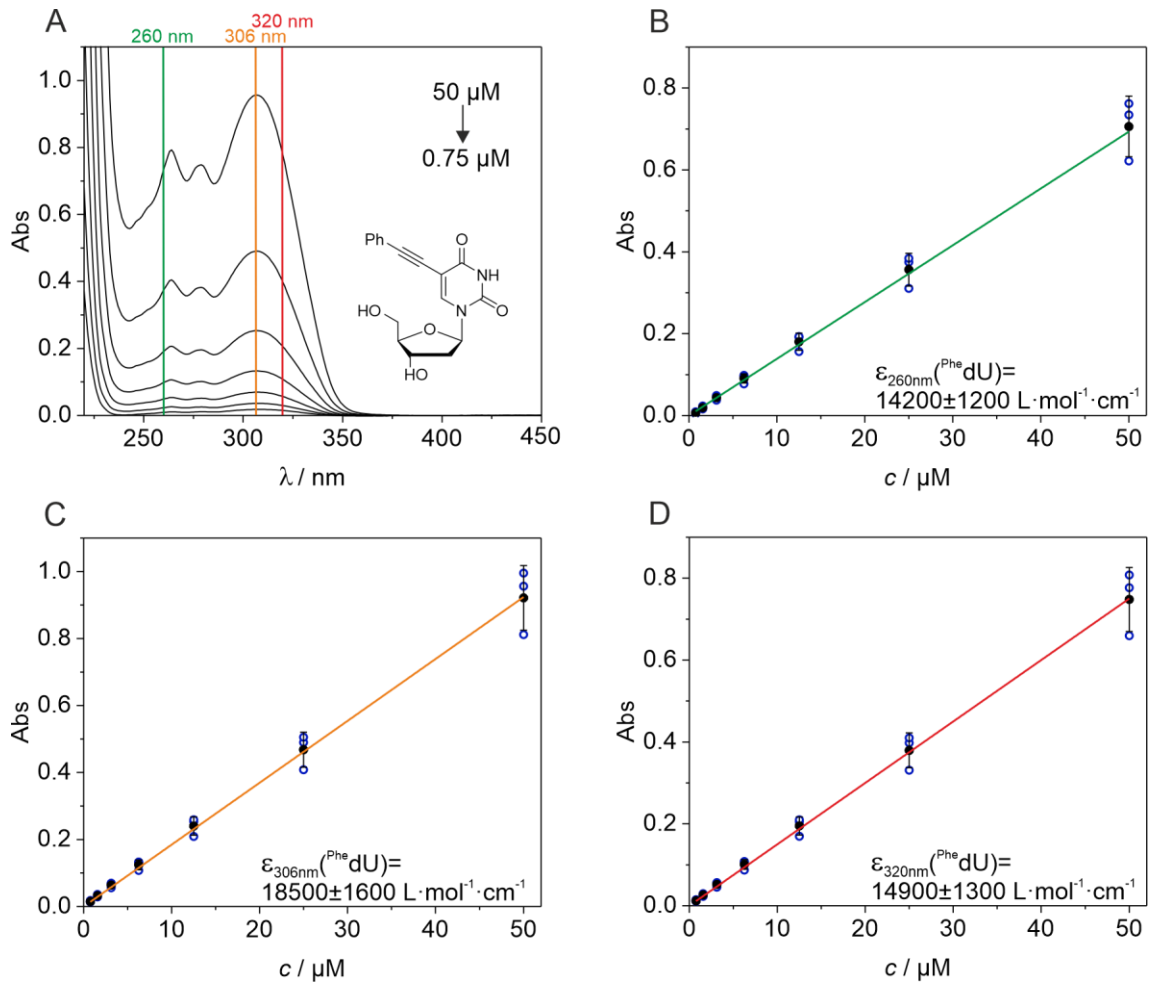


Figure S5.1 (A) UV/Vis spectrum of ^{Phe}dU in water starting from a 10 mM stock solution in DMSO. The extinction coefficient ε was determined at (B) 260 nm, (C) 306 nm and (D) 320 nm from a linear fit of the absorbance versus concentration.

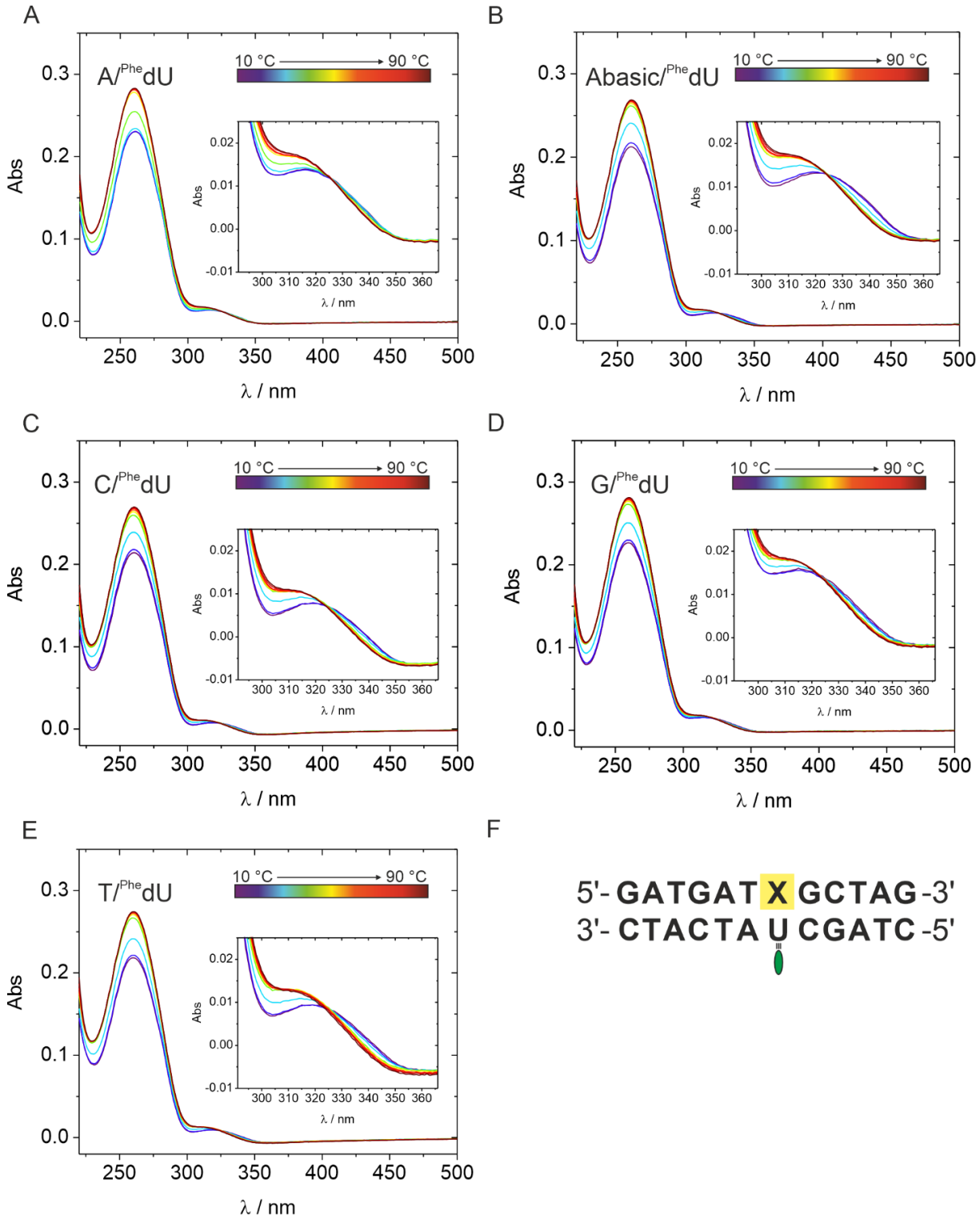


Figure S5.2 UV/Vis spectra of ^{Phe}dU opposite of (A) deoxyadenosine, (B) abasic linker, (C) deoxycytidine, (D) deoxyguanosine and (E) thymidine in a DNA duplex at different temperatures (1 µM DNA in 100 mM NaCl, 10 mM sodium phosphate, pH 7.0). The inset shows the absorption of ^{Phe}dU. (F) Sequence of the DNA duplex. Different nucleosides have been incorporated at the yellow highlighted positon X opposite of ^{Phe}dU unit.

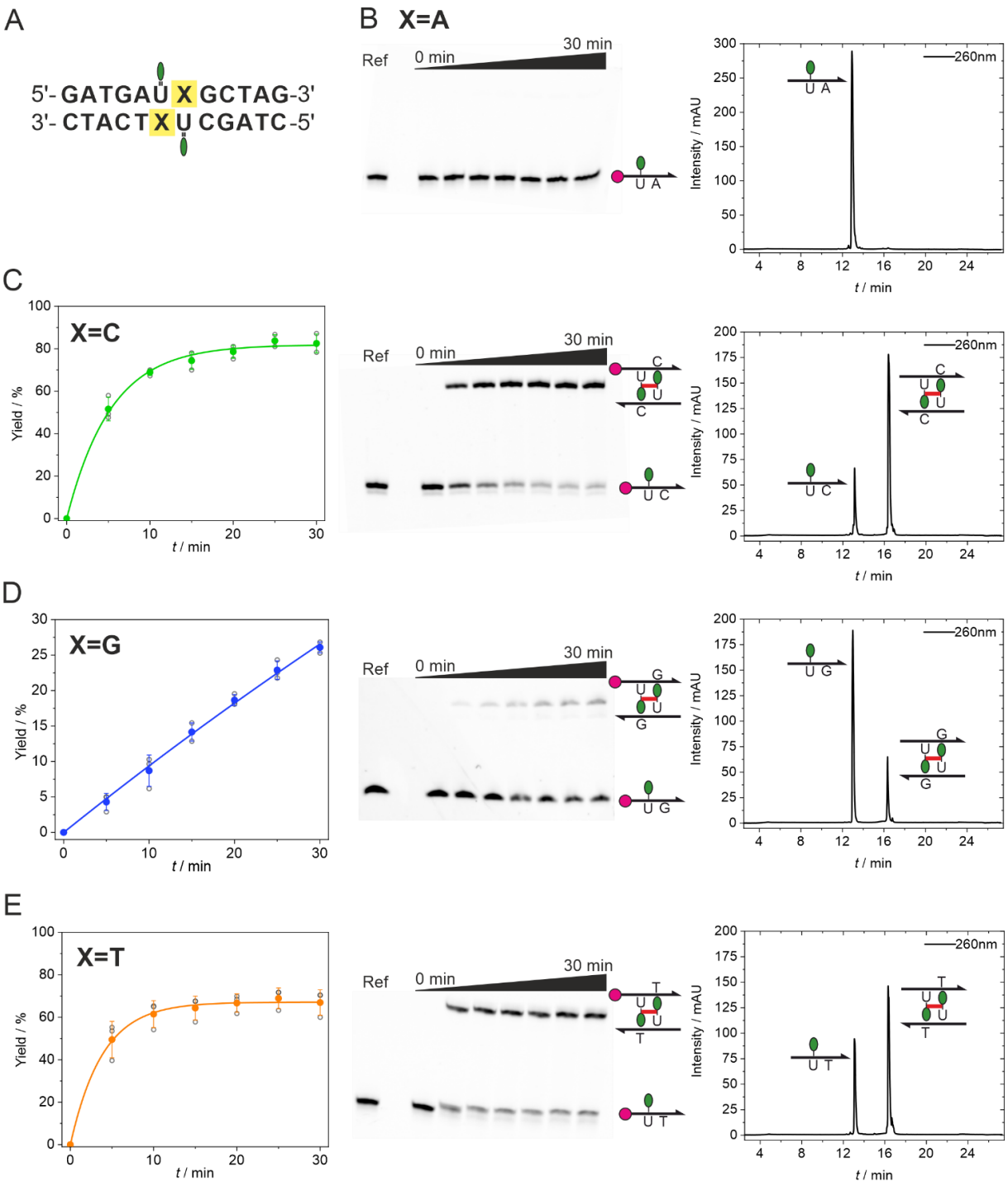


Figure S5.3 (A) Sequence of the DNA duplex used for ICL (X=A, C, G, T). Kinetic characterization of ICL formation using the fluorescence spectrometer light source with PAGE analysis and anion exchange HPLC of the DNA mixture after UV irradiation for (B) X=A, (C) X=C, (D) X=G and (E) X=T. Cy3 for quantification is shown as a magenta circle (1 µM DNA in 10 mM sodium phosphate, 100 mM NaCl, pH 7.0).

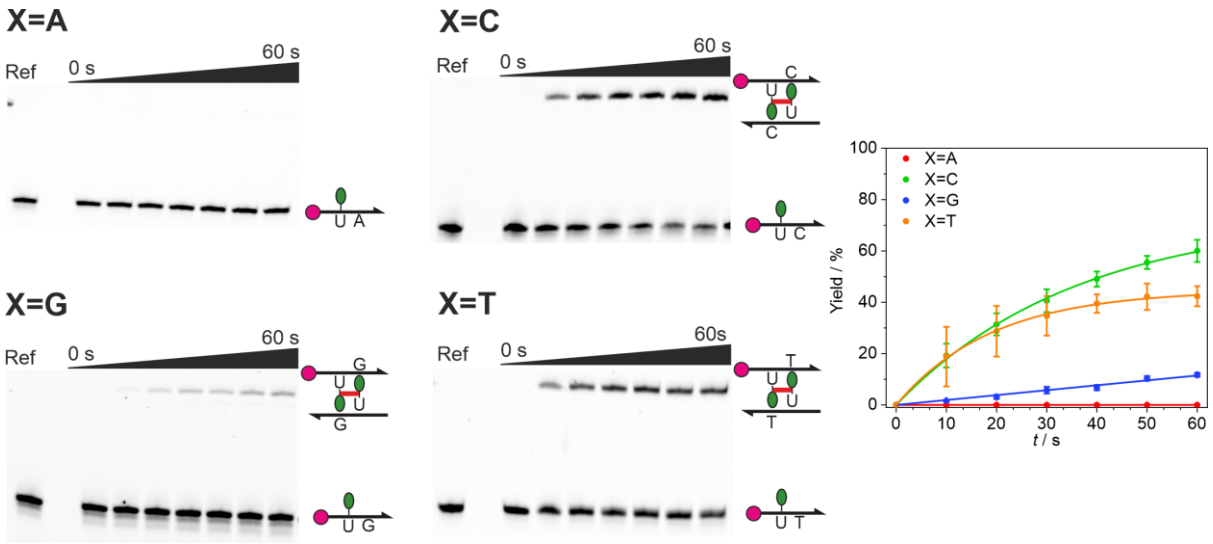


Figure S5.4 Kinetic characterization of ICL formation using a transilluminator as a light source. Cy3 for quantification is shown as a magenta circle (1 µM DNA in 10 mM sodium phosphate, 100 mM NaCl, pH 7.0).

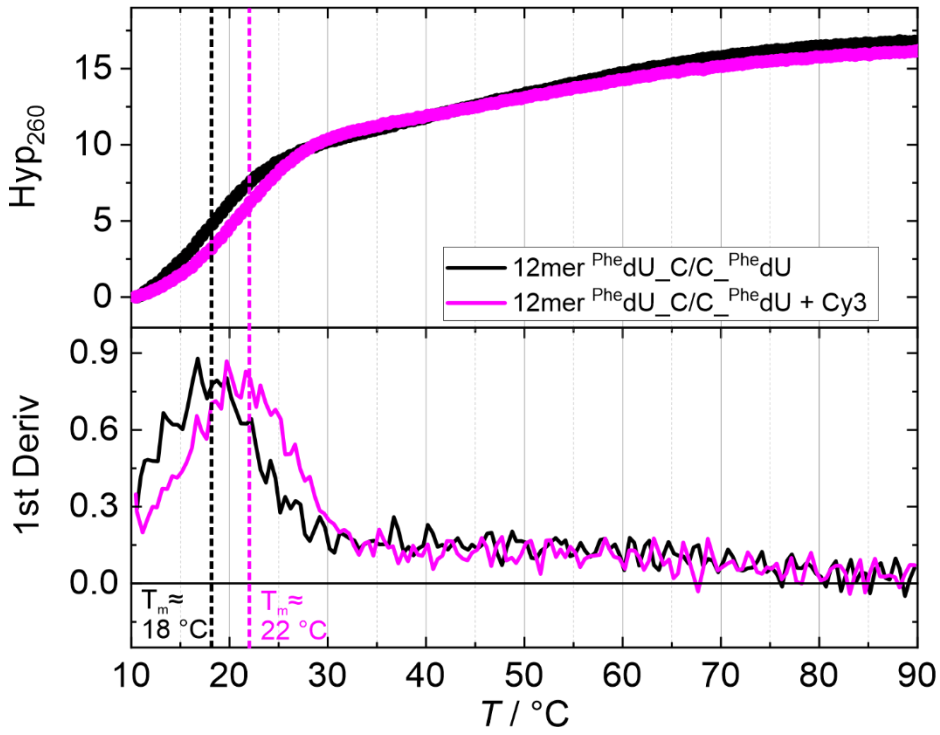


Figure S5.5 UV melting curves of the DNA duplexes where X=C without (black) and with (magenta) Cy3 label. Melting temperatures were determined by estimating the maxima of the first derivative (1 µM DNA in 10 mM sodium phosphate, 100 mM NaCl, pH 7.0).

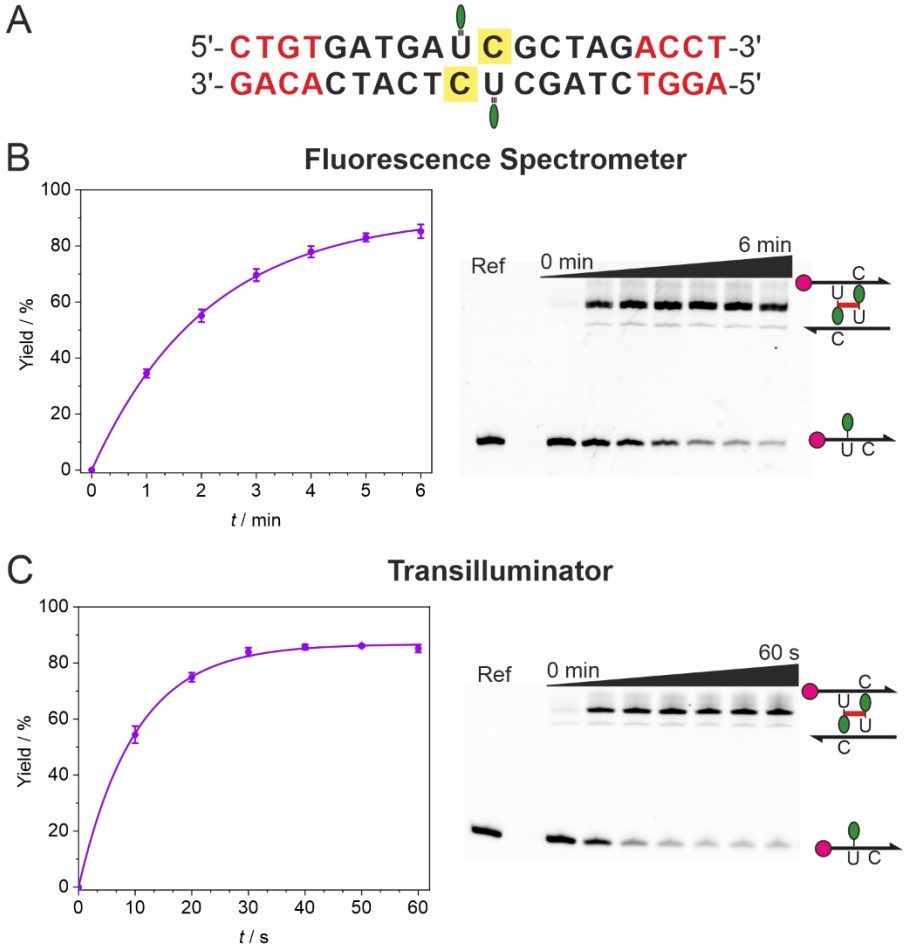


Figure S5.6 (A) Sequence of a 20mer DNA duplex containing the ^{Phe}dU crosslinking motif. Kinetic characterization of ICL formation using a (B) fluorescence spectrometer and a (C) transilluminator as light source. Cy3 for quantification is shown as magenta circle (1 µM DNA in 10 mM sodium phosphate, 100 mM NaCl, pH 7.0).

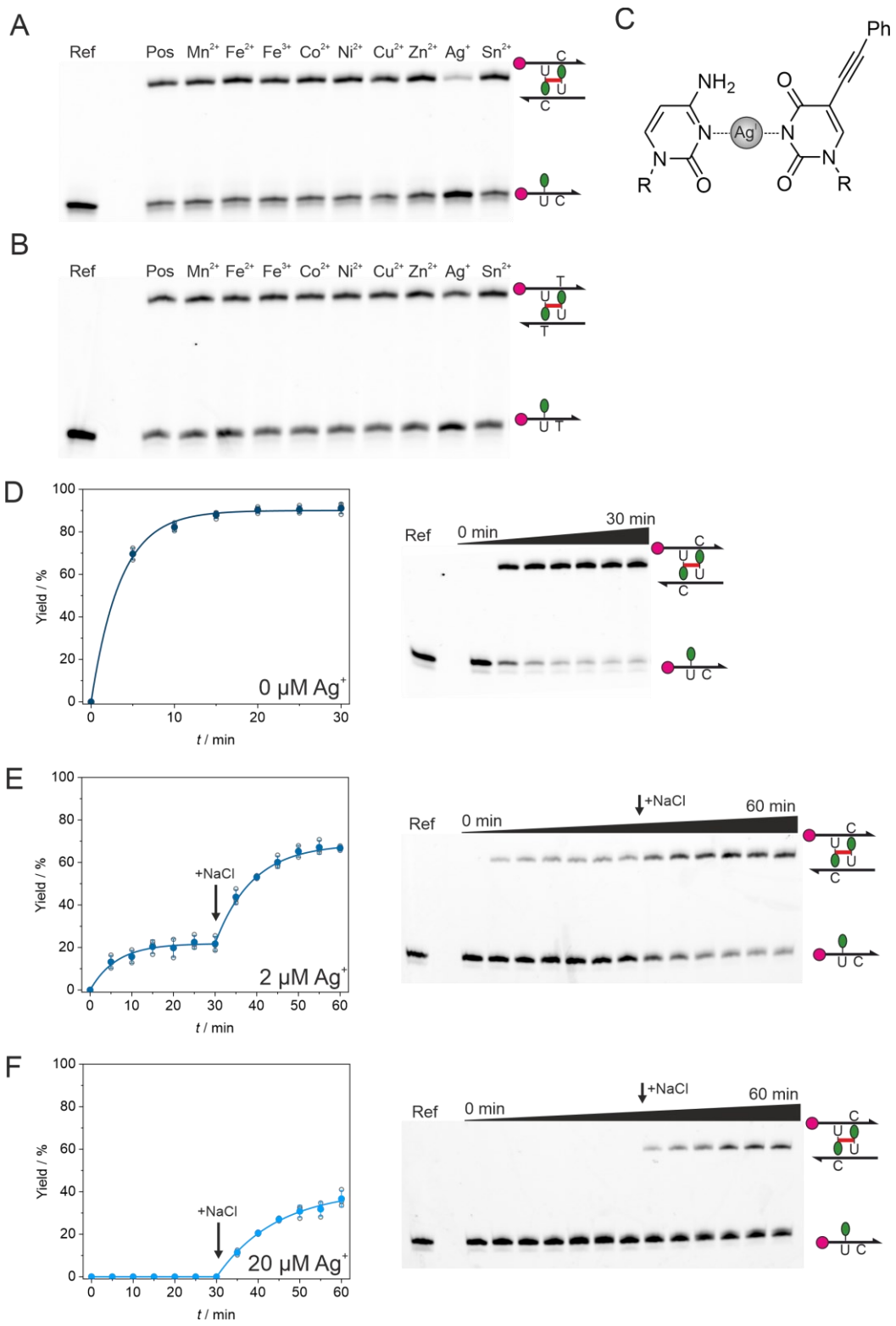


Figure S5.7 PAGE-based scouting of different metal ions (2 µM) for the inhibition of the crosslink reaction for (A) dC/^{Phe}dU mismatches and (B) dT/^{Phe}dU mismatches. A transilluminator was used as light source. (C) Silver mediated metal base pair between deoxycytidine and ^{Phe}dU. Kinetic characterization of ICL inhibition using fluorescence spectrometer light at 320 nm for (D) 0 µM, (E) 2 µM and (F) 20 µM silver ions. After 30 min NaCl was added for AgCl precipitation. Cy3 for quantification is shown as magenta circle (1 µM DNA in 10 mM MOPS, 100 mM $NaClO_4$, pH 7.1).

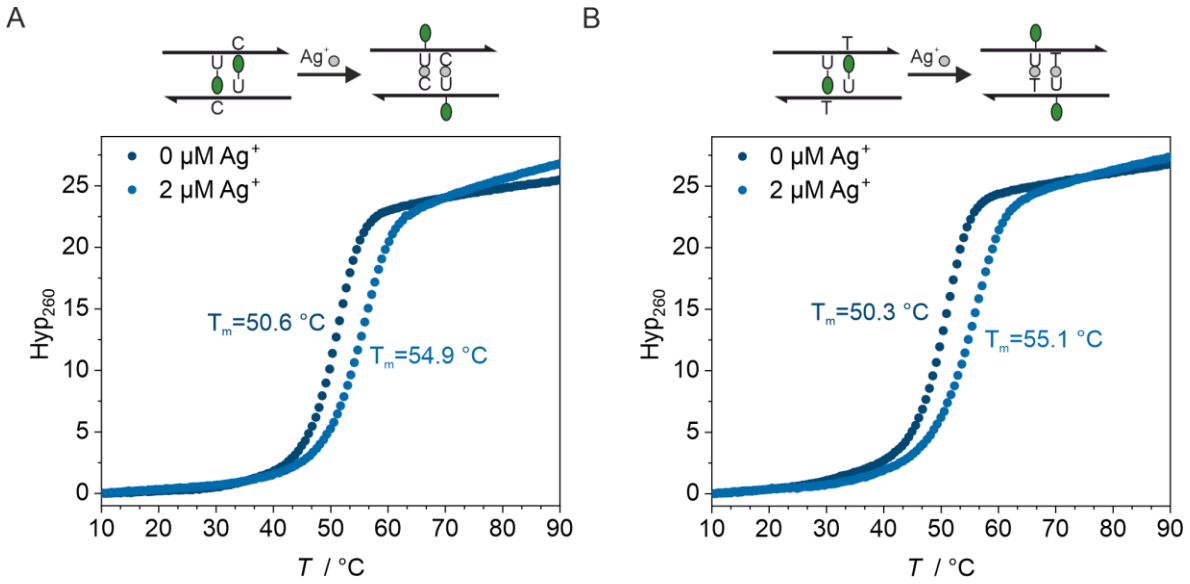


Figure S5.8 UV melting curves of the 20mer DNA duplex containing the ^{Phe}dU crosslinking motif without (dark blue) and with 2 µM $AgNO_3$ (blue) containing (A) dC/^{Phe}dU and (B) dT/^{Phe}dU mismatch motif (1 µM DNA in 10 mM MOPS, 100 mM $NaClO_4$, pH 7.1).

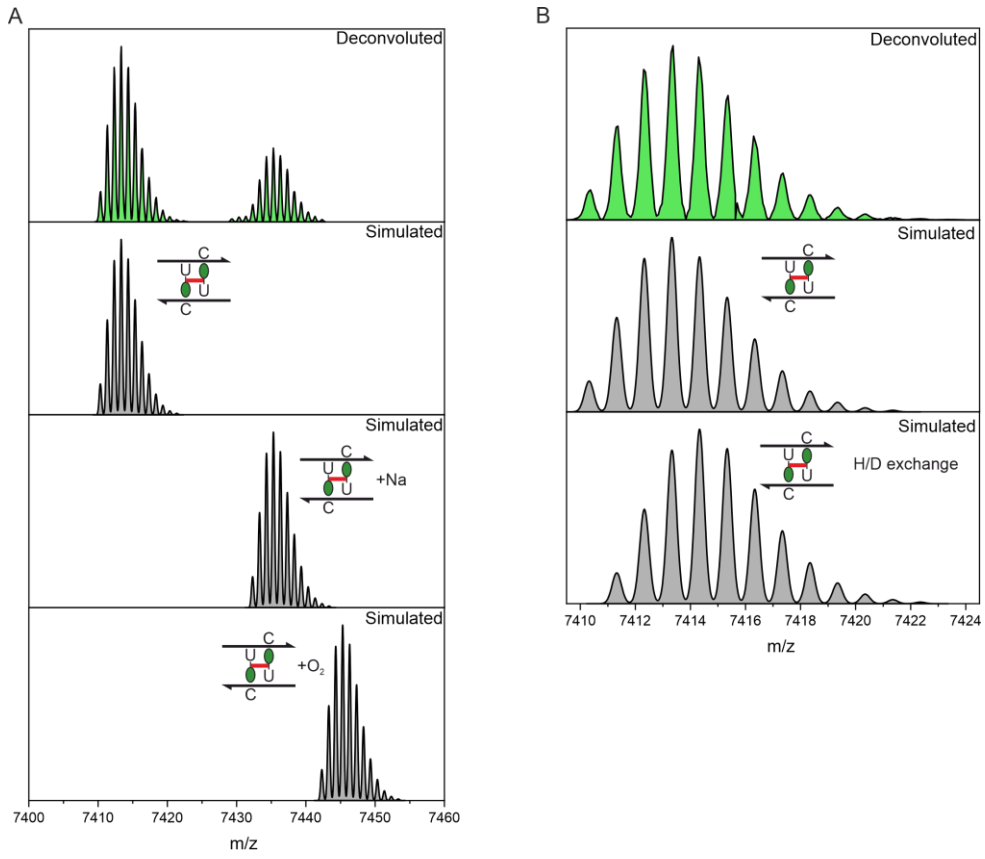


Figure S5.9 (A) HR-ESI-MS analysis of the crosslinked 12mer DNA duplex. The simulated spectra correspond to the mass of the combined single strands, with additional Na or O_2. (B) HR-ESI-MS analyses of 12mer DNA duplex crosslinked in D_2O. The simulated spectra correspond to the mass of the combined single strands without and with H/D exchange.

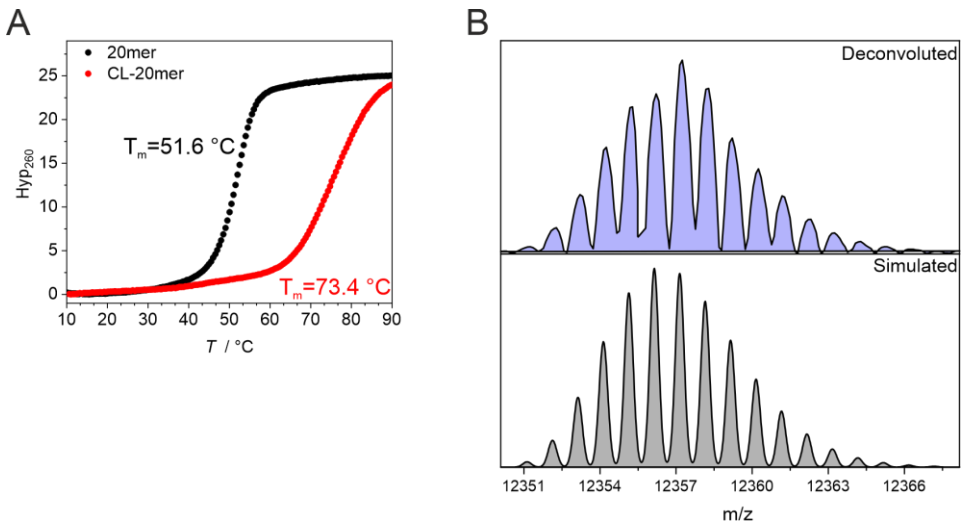


Figure S5.10 (A) UV melting curves of the 20mer DNA duplex containing the ^{Phe}dU crosslinking motif without (black) and with (red) crosslinking (1 µM DNA in 10 mM sodium phosphate buffer, 100 mM NaCl, pH 7.0). (B) HR-ESI-MS analysis of the crosslinked 20mer DNA duplex. The simulated spectrum corresponds to the mass of the combined single strands.

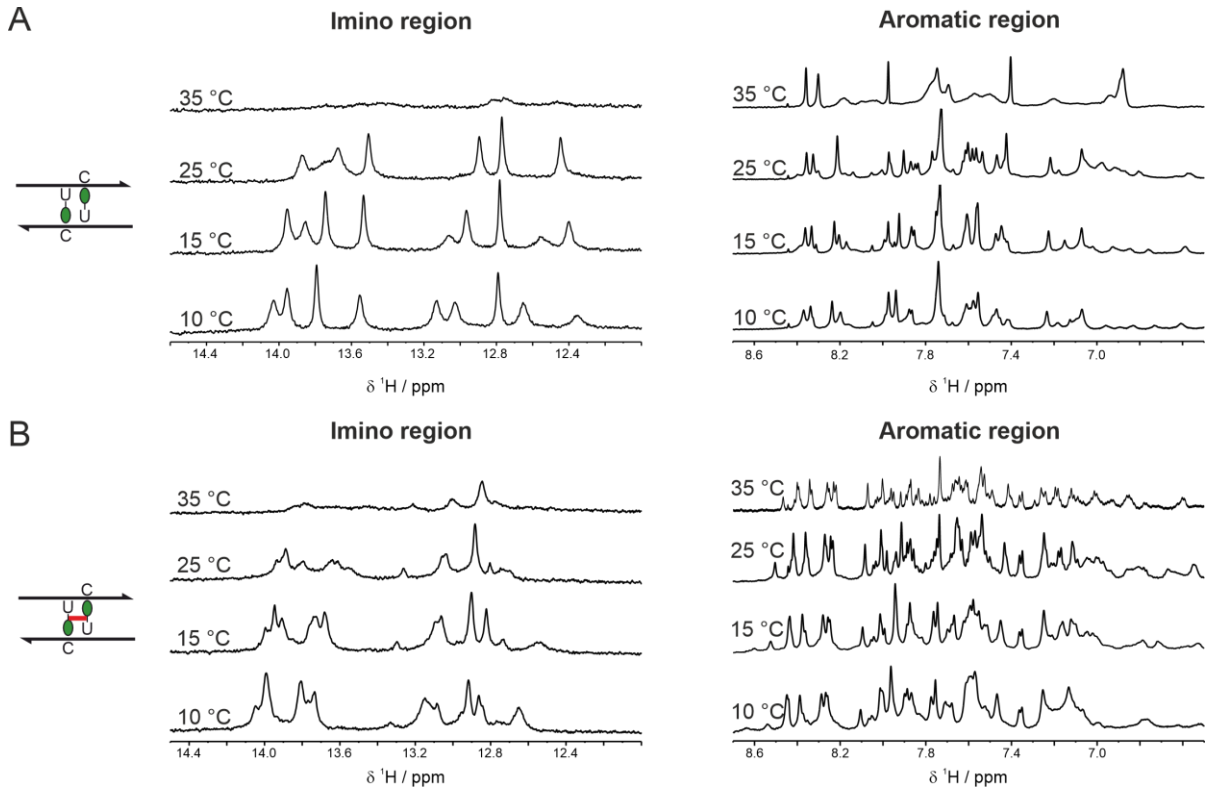


Figure S5.11 Imino and aromatic regions of the 1D ^{1}H NMR spectrum of the dC/^{Phe}dU mismatched duplex (A) without and (B) with ICL at various temperatures. Spectra recorded on 200 µM 12mer DNA duplex in in 10 mM sodium phosphate buffer, 100 mM NaCl, pH 7.0, 90%H_2O/10%D_2O, 600 MHz.

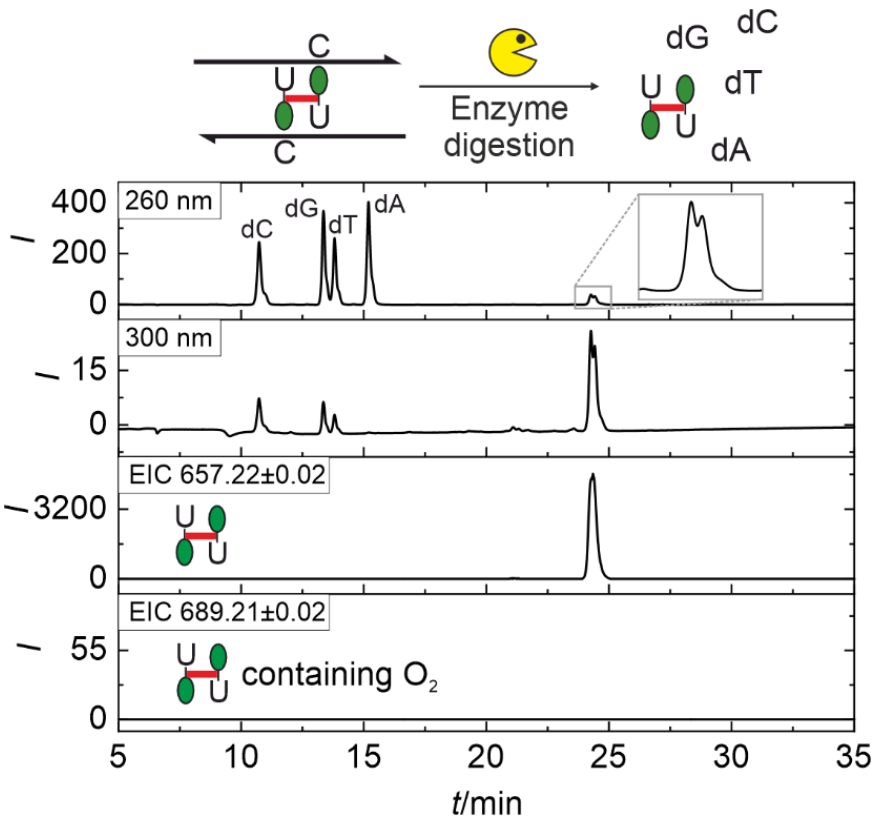


Figure S5.12 LC-MS analysis of the enzymatically digested DNA duplex containing the ICL. UV absorbance traces at 260 nm and 300 nm. Extracted ion chromatograms (EIC) showing MH^+ (m/z 657.22 ± 0.02) corresponding to two ^{Phe}dU units, and MH^+ (m/z 689.21 ± 0.02) corresponding to two ^{Phe}dU units with additional O_2. Incorporation of O_2 is not observed.

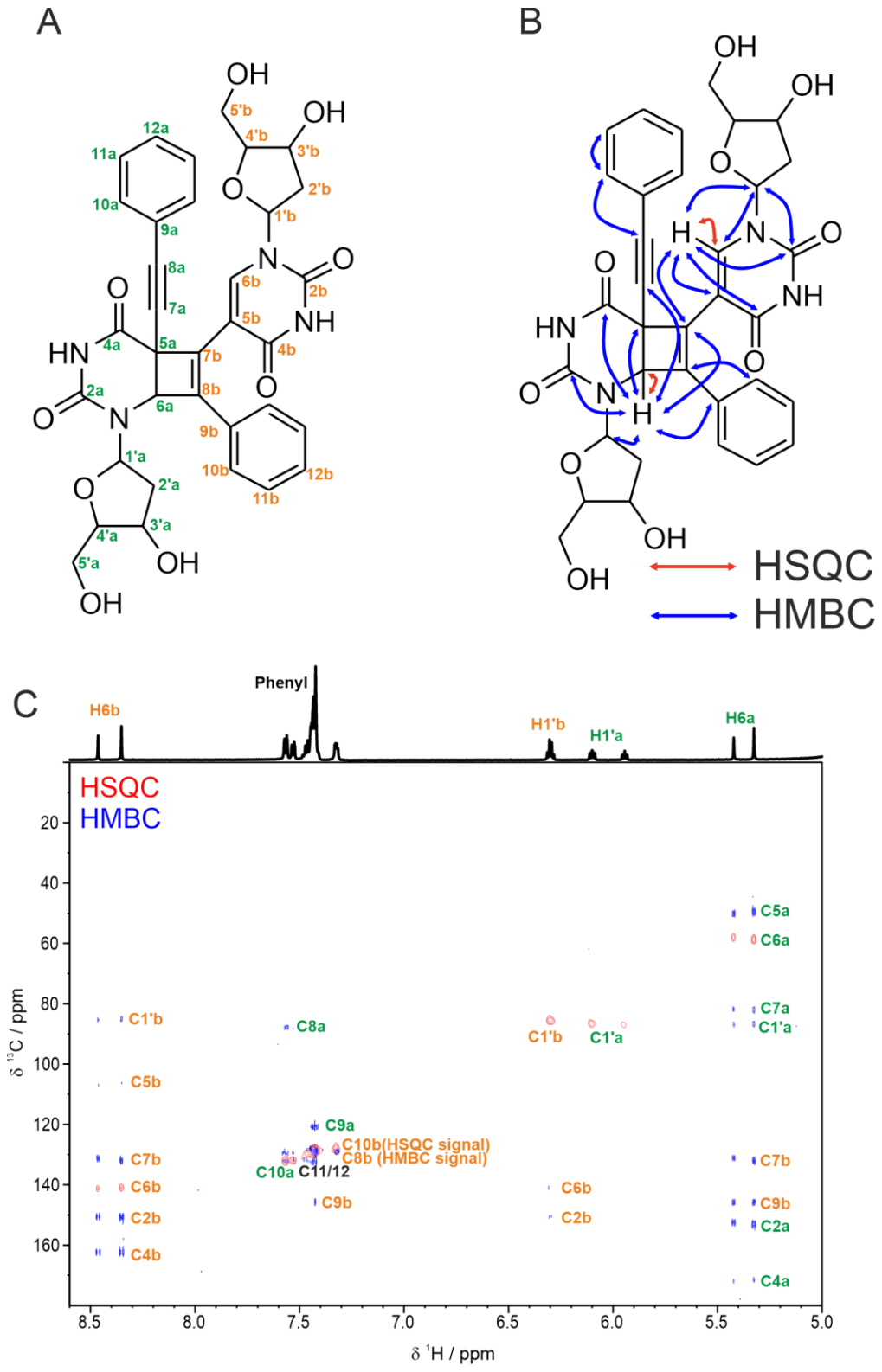


Figure S5.13 (A) Nomenclature and (B) key HSQC and HMBC correlations in the crosslinked product. Overlay of (C) HSQC (red) and HMBC (blue) spectra and the corresponding assignment. Spectra were recorded in 100% D_2O, 600 MHz, 25 °C.

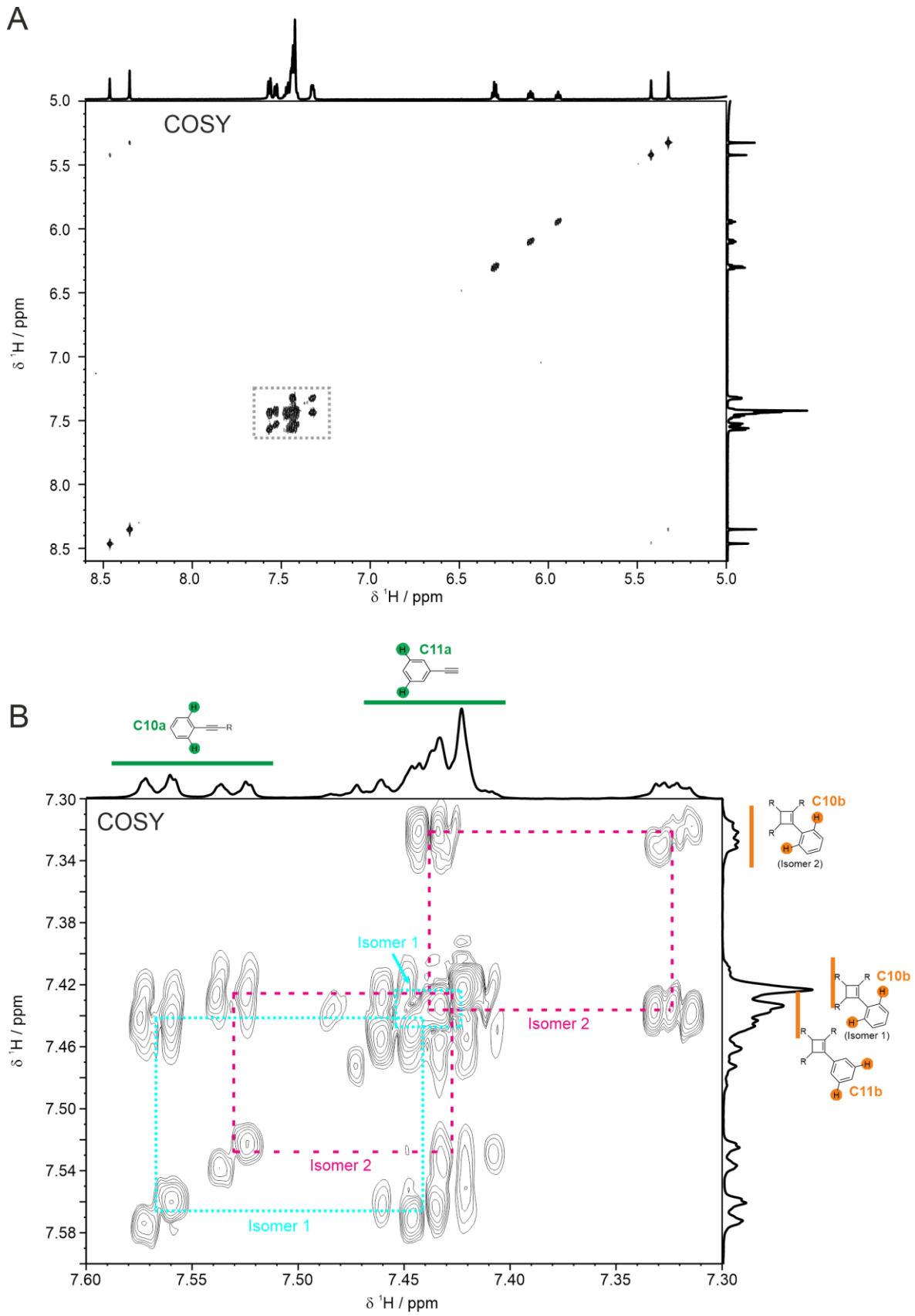


Figure S5.14 (A) COSY spectrum of the digested ICL unit. (B) Phenyl region of the COSY spectrum with corresponding assignment. The absolute configuration of the two isomers was not determined. The spectrum was recorded in 100% D_2O, 600 MHz, 25 °C.

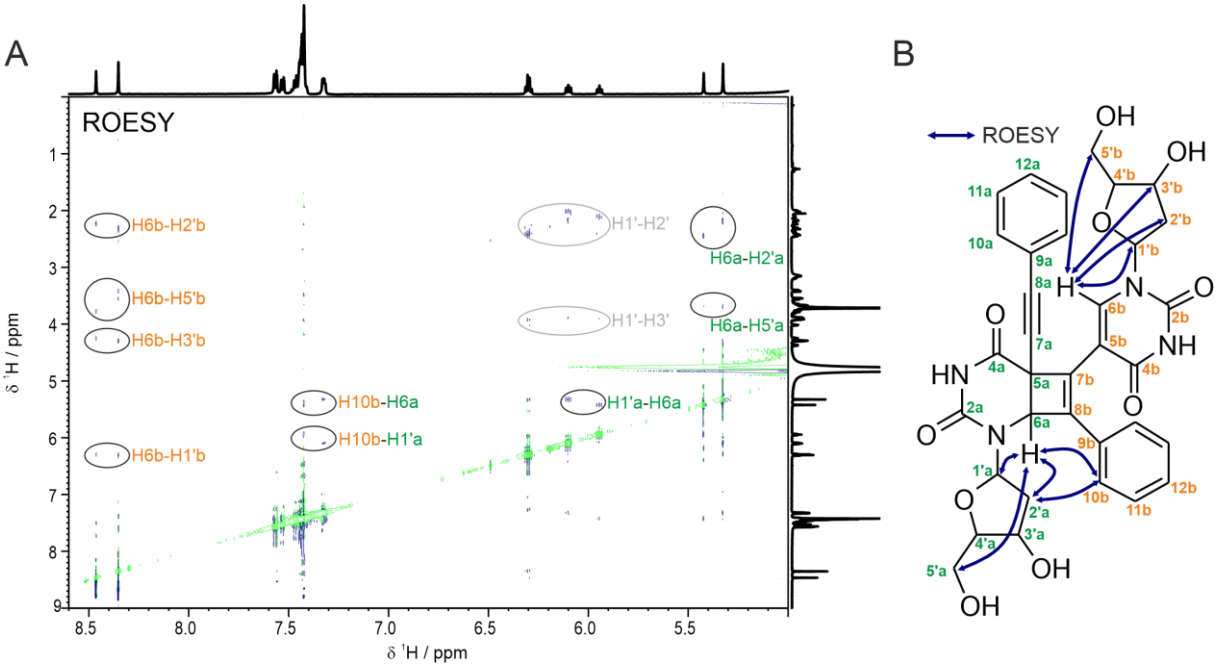


Figure S5.15 (A) ROESY spectrum and (B) structure of the formed crosslink product with nomenclature and key ROESY correlations. The spectrum was recorded in 100% D_2O, 600 MHz, 25 °C.

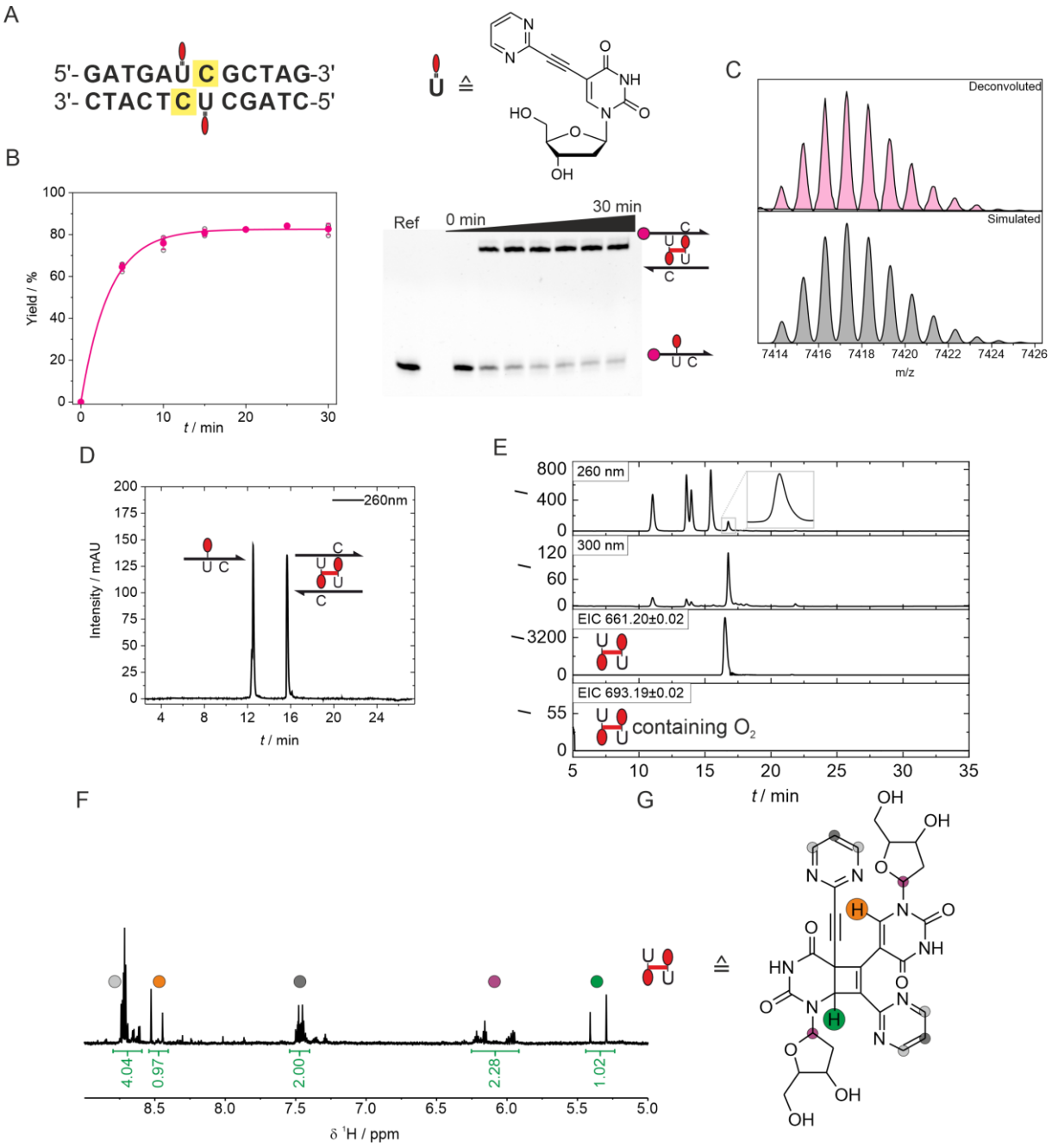


Figure S5.16 (A) Sequence of the DNA duplex and structure of the used ^{Pym}dU. (B) Kinetic characterization of ICL formation with ^{Pym}dU moiety using the fluorescence spectrometer light source with PAGE analysis. Cy3 for quantification is shown as magenta circle. (C) HR-ESI-MS analysis of crosslinked DNA duplex with ^{Pym}dU. The simulated spectrum shows the mass peaks for the combined single strands (D) HPLC of the DNA mixture after UV irradiation (1 µM DNA in 10 mM sodium phosphate buffer, 100 mM NaCl, pH 7.0). (E) LC-MS analysis of the enzymatically digested DNA duplex containing the ICL. Extracted ion chromatograms (EIC) showing MH+ (m/z 661.20 ± 0.02) corresponding to two ^{Pym}dU units and MH+ (m/z 693.19 ± 0.02) corresponding to two ^{Pym}dU units with additional O_2. (F) Aromatic-anomeric region of the 1D ^{1}H NMR spectrum and (G) putative structure of the formed crosslinked product with assignment of the most relevant protons. The spectrum was recorded in D_2O, 600 MHz, 25 °C.

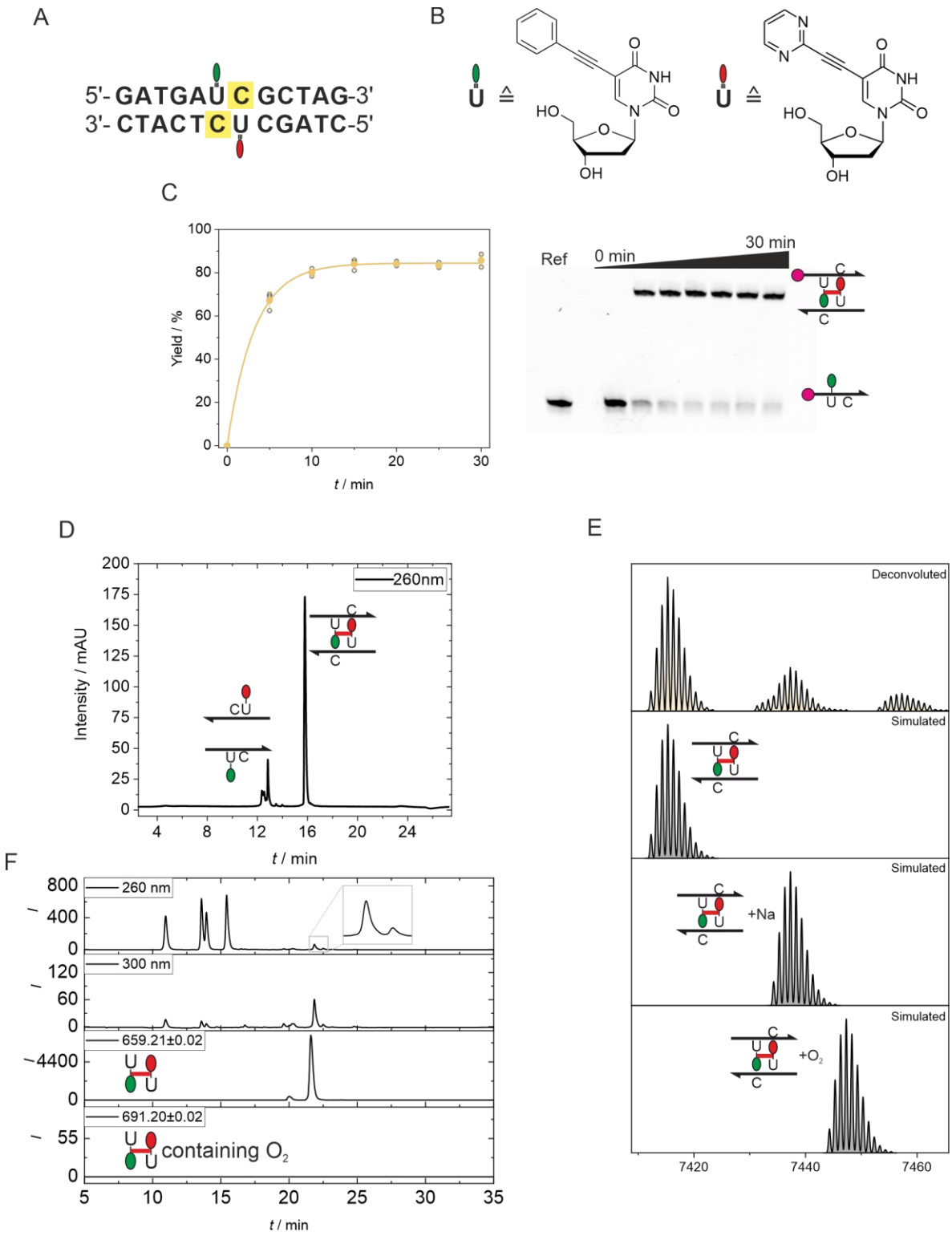


Figure S5.17 (A) Sequence of the DNA duplex and (B) structure of the used ^{Phe}dU and ^{Pym}dU. (C) Kinetic characterization of ICL formation with ^{Phe}dU and ^{Pym}dU moieties using the fluorescence spectrometer light source with PAGE analysis. Cy3 for quantification is shown as magenta circle. (D) HPLC of the DNA mixture after UV irradiation (1 µM DNA in 10 mM sodium phosphate buffer, 100 mM NaCl, pH 7.0) (E) HR-ESI-MS analysis of crosslinked DNA duplex with ^{Phe}dU and ^{Pym}dU. The simulated spectra showing the mass peaks for the combined single strands, combined single strands with sodium and combined single strands with oxygen. (F) LC-MS analysis of the enzymatically digested DNA duplex containing the ICL. Extracted ion chromatograms (EIC) showing MH+ (m/z

659.21 ± 0.02) corresponding to ^{Phe}dU and ^{Pym}dU units and MH+ (m/z 691.20 ± 0.02) corresponding to ^{Phe}dU and ^{Pym}dU units with additional O_2.

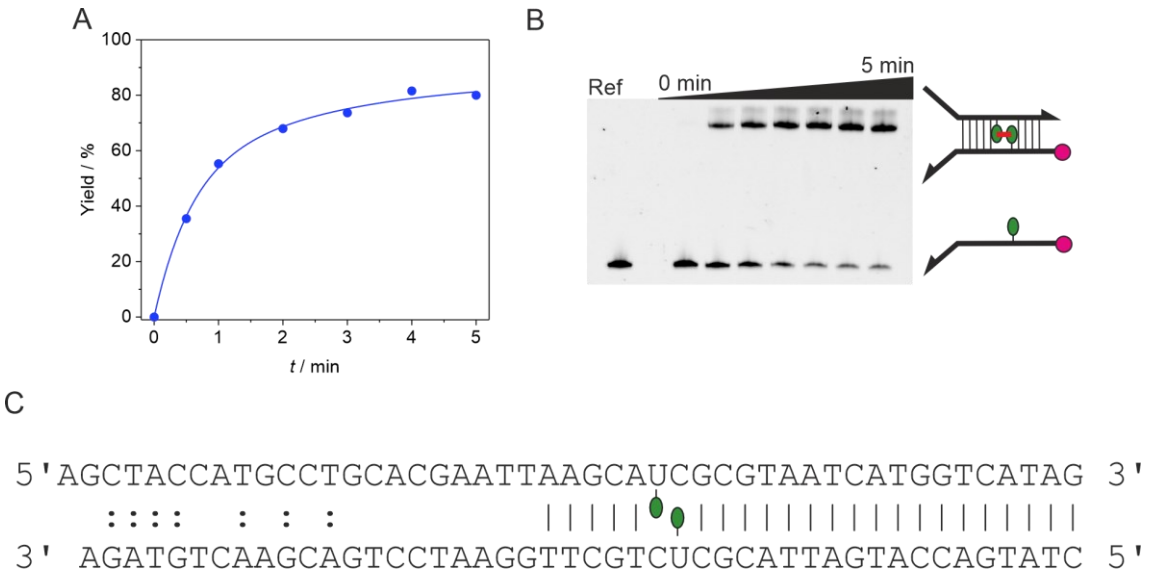


Figure S5.18 (A) Kinetic characterization of CL Y-fork formation with ^{Phe}dU moiety using the fluorescence spectrometer light source and (B) the corresponding PAGE analysis. Cy3 for quantification is shown as a magenta circle (1 µM DNA in 10 mM sodium phosphate buffer, 100 mM NaCl, pH 7.0). (C) Sequence of the Y-fork with CL motif.

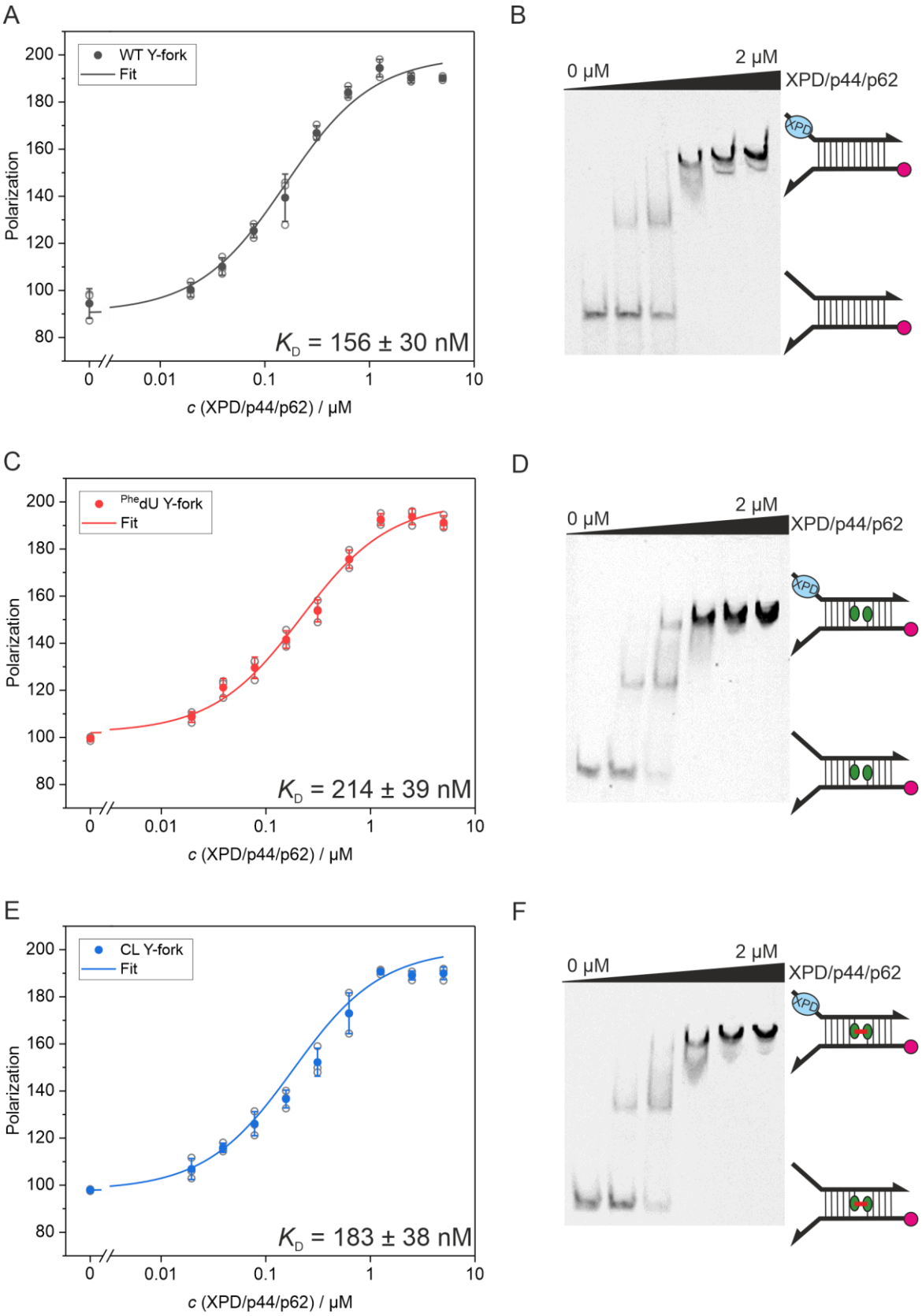


Figure S5.19 Determination of the binding affinity by fluorescence anisotropy and electrophoretic mobility shift assay (EMSA) with increasing concentration of the XPD/p44/p62 protein complex. Cy3 fluorescence was used for detection in both methods. WT Y-fork (A and B), ^{Phe}dU Y-fork (C and D) and CL Y-fork (E and F) were used as substrates.

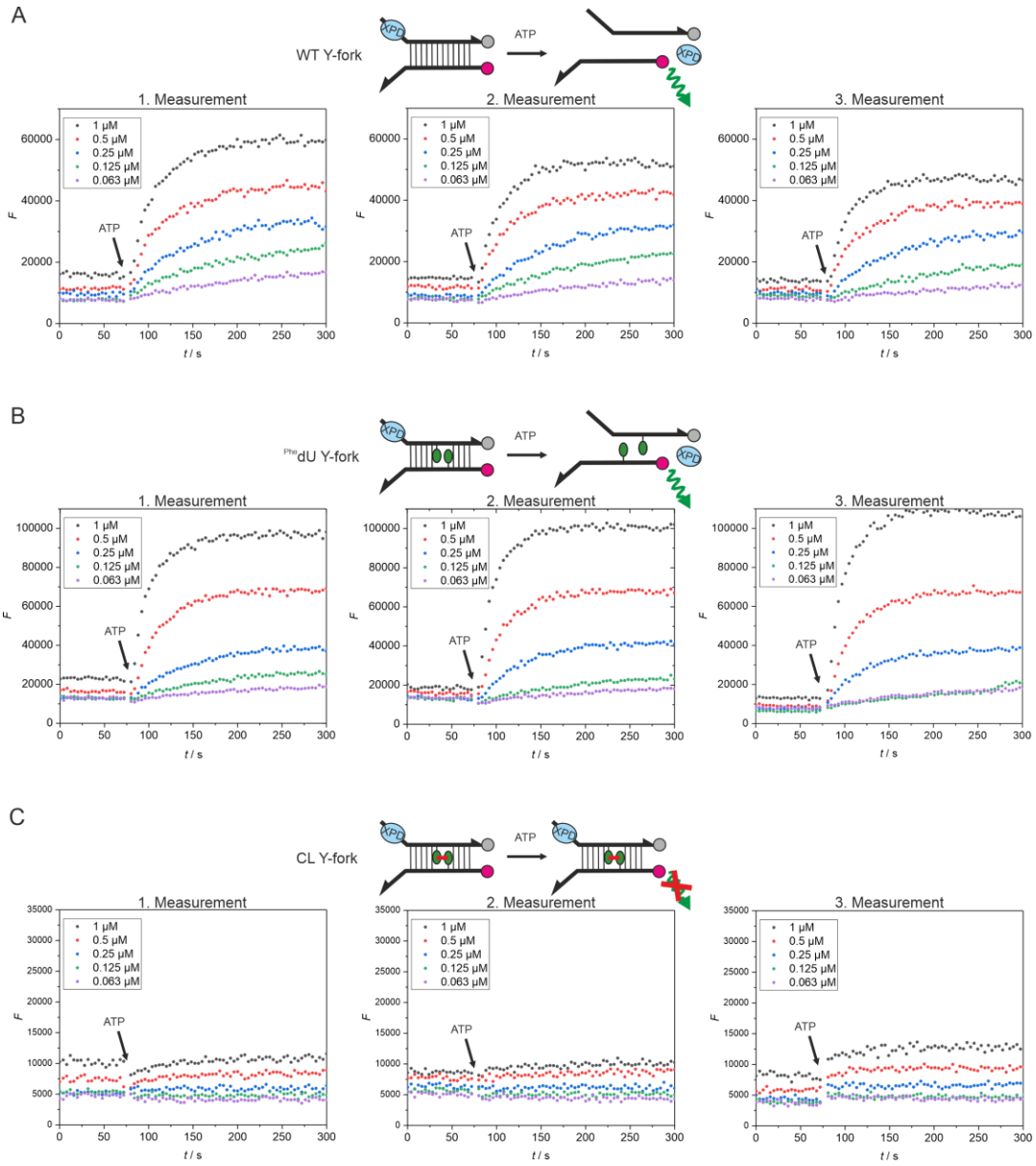


Figure S5.20 (A) Results of the helicase assay with different concentrations of the XPD/p44/p62 helicase complex for (A) WT Y-fork, (B) ^{Phe}dU Y-fork and (C) CL Y-fork as substrate. (125 nM DNA in 20 mM HEPES buffer, 50 mM KCl, 5 mM $MgCl_2$, 1 mM TCEP, pH 7.5).

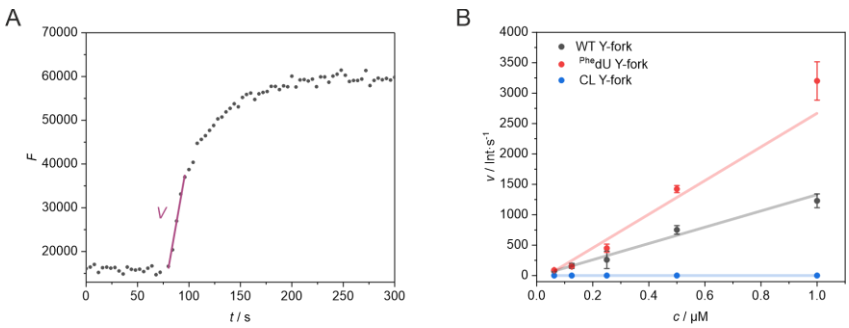


Figure S5.21 (A) Example of a helicase experiment to determine the kinetics of the helicase activity (1 µM XPD/p44/p62, 125 nM WT Y-fork in 20 mM HEPES buffer, 50 mM KCl, 5 mM $MgCl_2$, 1 mM TCEP, pH 7.5). (B) Estimated helicase activity with different substrates for XPD/p44/p62.

5.4.2 Experimental procedures

5.4.2.1 General material and methods

All standard chemicals were purchased from commercial suppliers. Trimethylsilylacetylene was obtained from Tokyo Chemical Industry and 2'-deoxy-5-iodouridine was purchased from Biosynth. 2-Bromopyrimidine, phenylacetylene, bis(triphenylphosphine)palladium(II) dichloride and tetrakis(triphenylphosphine)palladium(0) were obtained from Sigma Aldrich. 4,4'-Dimethoxytrityl chloride and 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite were obtained from Chem Genes Corporation.

2'-Deoxy-5'-*O*-(4,4'-dimethoxytrityl)-5-iodouridine (**S5.1**) and 2'-deoxy-*5*'-O-(4,4'-dimethoxytrityl)-5-(trimethylsilylethynyl)uridine (**S5.5**) were synthesized according to previously published protocols.[486]

Dry solvents DCM, DMF and THF were obtained via a solvent purification system (SPS) from Inert Corporation. Pyridine for DMT-protection was obtained from Acros Organics and dried over activated molecular sieves. Technical grade solvents were used for column chromatography and distilled prior to use. All other organic solvents were used in pro analysis or for synthesis quality without further purification.

Column chromatography was carried out on silica gel (Kieselgel 60, Merck, 0.063 – 0.200 mm). Thin layer chromatography (TLC) was performed on aluminum-backed plates coated with silica gel and a fluorescent indicator (Alugram SIL G/UV254, Macherey-Nagel, UV visualization, 254 nm).

5.4.2.2 NMR-Spectroscopy and mass spectrometry

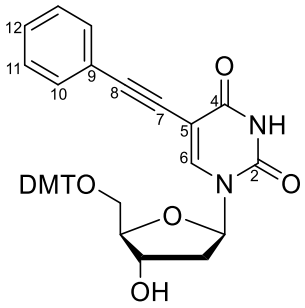
^{1}H-, ^{13}C- and ^{31}P-NMR spectra after chemical synthesis were recorded on a Bruker Avance HD III spectrometer at 400 MHz. Spectra were calibrated to the residual solvent peak of $CDCl_3$ (δ = 7.26 (^{1}H) and δ = 77.16 (^{13}C)) or CD_3OD (δ = 3.31 (^{1}H) and δ = 49.00 (^{13}C)). Chemical shifts δ are given in ppm and coupling constants *J* are given in Hz. Multiplicities are denoted as follows: s (singlet), d (doublet), t (triplet), q (quartet), p (pentet), dd (doublet of doublet), ddd (doublet of doublet of doublet), td (triplet of doublet), m (multiplet).

NMR spectra of the building blocks were evaluated using MestReNova v12.0.4.

High resolution ESI mass spectra were measured on a Bruker micrOTOF-Q III spectrometer.

5.4.2.3 Synthetic procedures

Compound **S5.2**



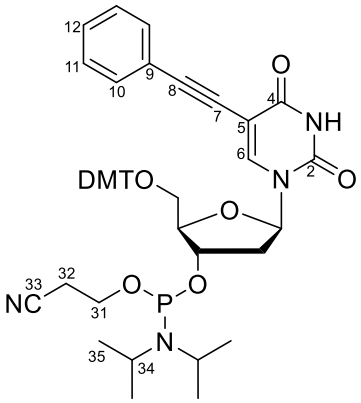
Under nitrogen atmosphere, compound **S5.1** (1.00 g, 1.52 mmol, 1.00 eq), CuI (29.0 mg, 152 µmol, 10 mol%) and $Pd(PPh_3)_4$ (176 mg, 152 µmol, 10.0 mol%) were dissolved in argon-purged Et_3N/DMF (1:1) (50 mL). Phenylacetylene (251 mg, 2.28 mmol, 1.50 eq.) was added and stirring was continued for 23 h at 60°C. The reaction mixture was washed with a saturated solution of ammonium chloride (120 mL). The aqueous phase was extracted with DCM (3 x 150 mL). The combined organic phases were washed with water (80 mL) and brine (80 mL), dried over Na_2SO_4, evaporated and the residue was purified by column chromatography (DCM + 1 % Et_3N → DCM + 2 % MeOH + 1 % Et_3N) to afford compound **S5.2** as a colorless foam (780 mg, 1.24 mmol, 82%). Characterization data were consistent with those previously reported in the literature.[474]

^{1}H NMR (400 MHz, $CDCl_3$): δ (ppm) = 8.16 (s, 1H, 6-H), 7.48 – 7.40 (m, 2H, DMT-H), 7.39 – 7.30 (m, 4H, DMT-H), 7.30 – 7.18 (m, 3H, 12-H, DMT-H), 7.20 – 7.10 (m, 3H, 11-H, DMT-H), 7.10 – 7.02 (m, 2H, 10-H), 6.83 – 6.73 (m, 4H, DMT-H), 6.36 (dd, *J* = 7.7, 5.7 Hz, 1H, 1'-H), 4.56 (dt, *J* = 5.5, 2.6 Hz, 1H, 3'-H), 4.13 (q, *J* = 3.0 Hz, 1H, 4'-H), 3.68 (s, 3H, DMT-H), 3.67 (s, 3H, DMT-H), 3.50 – 3.30 (m, 2H, 5'-H), 2.54 (ddd, *J* = 13.6, 5.8, 2.6 Hz, 1H, 2'-H), 2.32 (ddd, *J* = 13.6, 7.8, 5.9 Hz, 1H, 2'-H);

^{13}C{^{1}H} NMR (100 MHz, $CDCl_3$): δ (ppm) = 161.73 (4-C), 158.73 (DMT-C), 158.71 (DMT-C), 149.56 (2-C), 144.48 (DMT-C), 142.17 (6-C), 135.66 (DMT-C), 135.55 (DMT-C), 131.80 (10-C), 130.11 (DMT-C), 130.02 (DMT-C), 128.40 (12-C), 128.19 (11-C), 128.05 (DMT-C), 127.17 (DMT-C), 122.57 (9-C), 113.47 (DMT-C), 100.75 (5-C), 93.90 (8-C), 87.21 (DMT-C), 86.78 (4'-C), 85.95 (1'-C), 80.07 (7-C), 72.54 (3'-C), 63.64 (5'-C), 55.30 (DMT-C), 55.29 (DMT-C), 41.77 (2'-C);

HR-MS (ESI+): *m/z* calc. ($C_{38}H_{34}N_2O_7Na$, $[M+Na]^+$): 653.22582, found: 653.22666.

Compound **S5.3**



Under nitrogen atmosphere, compound **S5.2** (400 mg, 632 µmol, 1.00 eq.) was dissolved with DIPEA (648 µL, 481 mg, 3.79 mmol, 6.00 eq.) in anhydrous DCM (16 mL). After 10 min CEP-Cl (183 µL, 195 mg, 822 µmol, 1.30 eq.) was added. After 3 h stirring at room temperature additional CEP-Cl (55.0 µL, 58.4 mg, 247 µmol, 0.39 eq.) was added in two portions in a 1 h interval. The reaction mixture was stirred additionally at ambient temperature for 1 h. The solvent was removed under reduced pressure and the residue was purified by column chromatography (EtOAc + 1% Et_3N) to afford compound **S5.3** as a colorless foam (373 mg, 449 µmol, 71%). Characterization data were consistent with those previously reported in the literature.[474]

^{1}H NMR (400 MHz, $CDCl_3$): δ (ppm) = 8.22 (s, 1H, 6-H), 8.18 (s, 1H, 6-H), 7.48 – 7.40 (m, 4H, DMT-H), 7.40 – 7.30 (m, 8H, DMT-H), 7.30 – 7.17 (m, 6H, 12-H, DMT-H), 7.19 – 7.08 (m, 6H, 11-H, DMT-H), 7.06 – 6.95 (m, 4H, 10-H), 6.83 – 6.72 (m, 8H, DMT-H), 6.39 – 6.26 (m, 2H, 1'-H), 4.68 – 4.56 (m, 2H, 3'-H), 4.27 – 4.22 (m, 1H, 4'-H), 4.22 – 4.17 (m, 1H, 4'-H), 3.91 – 3.50 (m, 20H, 31-H, 34-H, DMT-H), 3.53 – 3.38 (m, 2H, 5'-H), 3.38 – 3.26 (m, 2H, 5'-H), 2.71 – 2.53 (m, 4H, 2'-H, 32-H), 2.49 – 2.40 (m, 2H, 32-H), 2.40 – 2.31 (m, 2H, 2'-H), 1.22 – 1.05 (m, 24H, 35-H);

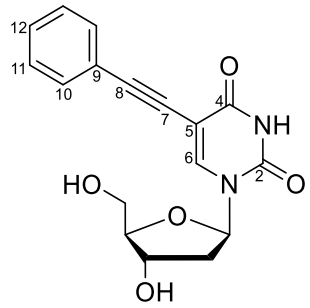
^{13}C{^{1}H} NMR (100 MHz, $CDCl_3$): δ (ppm) = 161.10 (4-C), 161.09 (4-C), 158.75 (DMT-C), 158.72 (DMT-C), 149.12 (2-C), 149.09 (2-C), 144.48 (DMT-C), 144.42 (DMT-C), 142.19 (6-C), 135.66 (DMT-C), 135.62 (DMT-C), 135.54 (DMT-C), 135.47 (DMT-C), 131.83 (10-C), 131.81 (10-C), 130.19 (DMT-C), 130.14 (DMT-C), 130.07 (DMT-C), 130.04 (DMT-C), 128.38, 128.17, 128.13, 128.07, 128.00, 127.98, 127.18 (DMT-C), 122.54 (9-C), 122.51 (9-C), 117.68 (33-C), 117.50 (33-

C), 113.46 (DMT-C), 113.44 (DMT-C), 100.79 (5-C), 100.75 (5-C), 93.96 (8-C), 93.93 (8-C), 87.22 (DMT-C), 86.42 (4'-C), 86.05 (1'-C), 86.01 (1'-C), 85.91 (1'-C), 79.85 (7-C), 74.15 (3'-C), 73.98 (3'-C), 73.82 (3'-C), 73.66 (3'-C), 63.42 (5'-C), 63.28 (5'-C), 58.57 (31-C), 58.41 (31-C), 58.22 (31-C), 55.32 (DMT-C), 55.30 (DMT-C), 55.28 (DMT-C), 43.49 (34-C), 43.42 (34-C), 43.36 (34-C), 43.29 (34-C), 40.90 (2'-C), 24.80 (35-C), 24.74 (35-C), 24.69 (35-C), 24.67 (35-C), 24.62 (35-C), 20.59 (32-C), 20.52 (32-C), 20.40 (32-C), 20.33 (32-C);

^{31}P{^{1}H} NMR (162 MHz, $CDCl_3$): δ (ppm) = 148.94, 148.50;

HR-MS (ESI+): *m/z* calc. ($C_{47}H_{51}N_4O_8PNa$, $[M+Na]^+$): 853.33367, found: 853.33344.

Compound **S5.4**



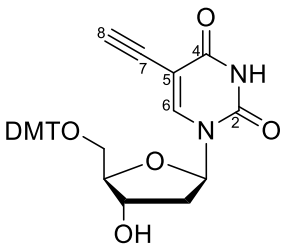
Compound **S5.2** (200 mg, 371 µmol, 1.00 eq.) was dissolved in DCM (22 mL). TFA (425 µL, 633 mg, 5.55 mmol, 15.0 eq.) was added and the mixture was stirred for 15 min at room temperature. Nitrogen was bubbled through the solution to remove TFA. MeOH (2 mL) was added and removed under reduced pressure four times. The residue was purified by column chromatography (DCM + 10 % MeOH) to afford compound **S5.4** as a colorless foam (83.4 mg, 254 µmol, 68%).

^{1}H NMR (400 MHz, CD_3OD): δ (ppm) = 8.44 (s, 1H, 6-H), 7.57 – 7.46 (m, 2H, 10-H), 7.38 – 7.32 (m, 3H, 11-H, 12-H), 6.27 (t, *J* = 6.5 Hz, 1H, 1'-H), 4.43 (dt, *J* = 6.2, 3.7 Hz, 1H, 3'-H), 3.95 (q, *J* = 3.3 Hz, 1H, 4'-H), 3.88 – 3.71 (m, 2H, 5'-H), 2.38 – 2.22 (m, 2H, 2'-H);

^{13}C{^{1}H} NMR (100 MHz, CD_3OD): δ (ppm) = 164.33 (4-C), 151.18 (2-C), 145.08 (6-C), 132.47 (10-C), 129.53 (11-C), 129.47 (12-C), 124.32 (9-C), 100.54 (5-C), 93.85 (8-C), 89.12 (4'-C), 87.04 (1'-C), 81.74 (7-C), 71.91 (3'-C), 62.50 (5'-C), 41.78 (2'-C);

HR-MS (ESI+): *m/z* calc. ($C_{17}H_{16}N_2O_5Na$, $[M+Na]^+$): 351.09382, found: 351.09514.

Compound **S5.6**



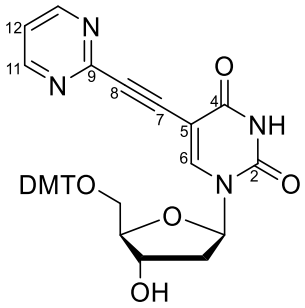
Compound **S5.5** (500 mg, 798 µmol, 1.00 eq.) was dissolved in THF/MeOH (1:1, 10 mL). K_2CO_3 (220 mg, 1.59 mmol, 2.00 eq.) was added and the mixture was stirred for 16 h at room temperature. The suspension was diluted with water (25 mL), extracted with EtOAc (3x150 mL) and washed with water and brine. The organic phase was dried over Na_2SO_4, evaporated and the residue was purified by column chromatography (DCM + 2 % MeOH + 1 % Et_3N → DCM + 5 % MeOH + 1 % Et_3N) to afford compound **S5.6** as yellow foam (364 mg, 656 µmol, 82%). Characterization data was in accordance with that previously reported in literature.[486]

^{1}H NMR (400 MHz, $CDCl_3$): δ (ppm) = 8.07 (s, 1H, 6-H), 7.45 – 7.37 (m, 2H, DMT-H), 7.37 – 7.26 (m, 6H, DMT-H), 7.24 – 7.19 (m, 1H, DMT-H), 6.89 – 6.80 (m, 4H, DMT-H), 6.28 (dd, *J* = 7.3, 5.9 Hz, 1H, 1'-H), 4.55 (dd, *J* = 6.2, 3.1 Hz, 1H, 3'-H), 4.08 (q, *J* = 3.4 Hz, 1H, 4'-H), 3.79 (s, 6H, DMT-H), 3.45 – 3.33 (m, 2H, 5'-H), 2.90 (s, 1H,8-H), 2.50 (ddd, *J* = 13.7, 5.9, 3.1 Hz, 1H, 2'-H), 2.29 (ddd, *J* = 13.5, 7.2, 6.0 Hz, 1H, 2'-H);

^{13}C{^{1}H} NMR (100 MHz, $CDCl_3$): δ (ppm) = 161.19 (4-C), 158.81 (DMT-C), 158.80 (DMT-C), 148.98 (2-C), 144.49 (DMT-C), 143.74 (6-C), 135.57 (DMT-C), 135.40 (DMT-C), 130.16 (DMT-C), 130.14 (DMT-C), 128.24 (DMT-C), 128.04 (DMT-C), 113.50 (DMT-C), 99.34 (5-C), 87.30 (DMT-C), 86.52 (4'-C), 85.81 (1'-C), 82.18 (8-C), 74.16 (7-C), 72.42 (3'-C), 63.45 (5'-C), 55.41 (DMT-C), 41.65 (2'-C);

HR-MS (ESI+): *m/z* calc. ($C_{32}H_{30}N_2O_7Na$, $[M+Na]^+$): 577.19452, found: 577.19708.

Compound **S5.7**

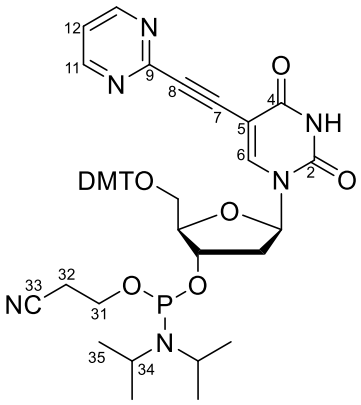


Under nitrogen atmosphere, compound **S5.6** (295 mg, 532 µmol, 1.00 eq), CuI (5.20 mg, 27.3 µmol, 5 mol%) and $Pd(PPh_3)_2Cl_2$ (38.0 mg, 54.1 µmol, 10 mol%) were dissolved in argon-purged Et_3N/THF (1:1) (3 mL). 2-Bromopyrimidine (86.0 mg, 541 µmol, 1.02 eq.) was added and stirring was continued for 20 h at 50°C. The reaction mixture was diluted with THF and EtOAc and filtrated over celite. The combined organic solvents were evaporated, and the residue was purified by column chromatography (DCM + 1 % MeOH + 1 % Et_3N → DCM + 6 % MeOH + 1 % Et_3N) to afford compound **S5.7** as a brown foam (145 mg, 229 µmol, 43%).

^{1}H NMR (400 MHz, $CDCl_3$): δ (ppm) = 8.69 (d, *J* = 5.0 Hz, 2H, 11-H), 8.12 (s, 1H, 6-H), 7.42 – 7.37 (m, 2H, DMT-H), 7.34 – 7.28 (m, 4H, DMT-H), 7.25 – 7.18 (m, 3H, 12-H, DMT-H), 7.14 – 7.07 (m, 1H, DMT-H), 6.83 – 6.75 (m, 4H, DMT-H), 6.24 (t, *J* = 6.4 Hz, 1H, 1'-H), 4.55 (dt, *J* = 7.2, 3.9 Hz, 1H, 3'-H), 4.14 – 4.06 (m, 1H, 4'-H), 3.71 – 3.69 (m, 6H, DMT-H), 3.46 (dd, *J* = 10.4, 4.8 Hz, 1H, 5'-H), 3.30 (dd, *J* = 10.4, 4.1 Hz, 1H, 5'-H), 2.65 – 2.51 (m, 1H, 2'-H), 2.38 – 2.25 (m, 1H, 2'-H);

^{13}C{^{1}H} NMR (100 MHz, $CDCl_3$): δ (ppm) = 161.07 (4-C), 158.62 (DMT-C), 158.59 (DMT-C), 157.29 (11-C), 152.68 (9-C), 149.35 (2-C), 145.28 (6-C), 144.60 (DMT-C), 135.81 (DMT-C), 135.50 (DMT-C), 130.20 (DMT-C), 130.12 (DMT-C), 128.09 (DMT-C), 128.05 (DMT-C), 127.00 (DMT-C), 119.89 (12-C), 113.36 (DMT-C), 98.44 (5-C), 91.91 (8-C), 86.90 (DMT-C), 86.18 (4'-C), 86.10 (1'-C), 80.21 (7-C), 71.97 (3'-C), 63.91 (5'-C), 55.30 (DMT-C), 41.19 (2'-C);
HR-MS (ESI+): *m/z* calc. ($C_{36}H_{33}N_4O_7$, $[M+H]^+$): 633.23438, found: 633.23847.

Compound **S5.8=^{Pym}dU-PA**



Under nitrogen atmosphere, compound **S5.7** (140 mg, 222 µmol, 1.00 eq.) was dissolved with DIPEA (227 µL, 172 mg, 1.33 mmol, 6.00 eq.) in anhydrous DCM (6 mL). After 10 min CEP-

Cl (75.0 mg, 317 µmol, 1.43 eq.) was added. The reaction mixture was stirred at ambient temperature for 3 h. The solvent was removed under reduced pressure and the residue was purified by column chromatography (EtOAc + 1% Et_3N) to afford compound **S5.8** as a brownish foam (98.4 mg, 118 µmol, 53%).

^{1}H NMR (400 MHz, $CDCl_3$): δ (ppm) = 8.71 – 8.65 (m, 4H, 11-H), 8.13 – 8.04 (m, 2H, 6-H), 7.44 – 7.39 (m, 4H, DMT-H), 7.36 – 7.30 (m, 8H, DMT-H), 7.30 – 7.21 (m, 4H, DMT-H), 7.22 – 7.18 (m, 2H, 12-H), 7.17 – 7.11 (m, 2H, DMT-H), 6.87 – 6.78 (m, 8H, DMT-H), 6.23 – 6.16 (m, 2H, 1'-H), 4.62 – 4.51 (m, 2H, 3'-H), 4.29 – 4.18 (m, 2H, 4'-H), 3.95 – 3.51 (m, 20H, 31-H, 34-H, DMT-H), 3.46 – 3.23 (m, 4H, 5'-H), 2.80 – 2.58 (m, 4H, 2'-H, 32-H), 2.48 (t, J = 6.4 Hz, 2H, 32-H), 2.33 – 2.22 (m, 2H, 2'-H), 1.22 – 1.07 (m, 24H, 35-H).

^{13}C{^{1}H} NMR (100 MHz, $CDCl_3$): δ (ppm) = 160.69 (4-C), 158.72 (DMT-C), 158.70 (DMT-C), 158.68 (DMT-C), 157.18 (11-C), 152.95 (9-C), 149.00 (2-C), 148.97 (2-C), 144.48 (DMT-C), 144.44 (DMT-C), 144.37 (6-C), 135.67 (DMT-C), 135.33 (DMT-C), 135.30 (DMT-C), 130.30 (DMT-C), 130.25 (DMT-C), 130.16 (DMT-C), 130.12 (DMT-C), 128.10 (DMT-C), 128.05 (DMT-C), 127.07 (DMT-C), 119.84 (12-C), 117.70 (33-C), 117.56 (33-C), 113.43 (DMT-C), 113.41 (DMT-C), 98.68 (5-C), 98.62 (5-C), 92.34 (8-C), 87.09 (DMT-C), 87.07 (DMT-C), 86.61 (1'-C), 86.48 (1'-C), 86.37 (4'-C), 86.03 (4'-C), 79.12 (7-C), 74.19 (3'-C), 74.02 (3'-C), 73.82 (3'-C), 73.65 (3'-C), 63.54 (5'-C), 63.38 (5'-C), 58.50 (31-C), 58.31 (31-C), 55.35 (DMT-C), 55.33 (DMT-C), 43.50 (34-C), 43.45 (34-C), 43.38 (34-C), 43.32 (34-C), 40.68 (2'-C), 24.79 (35-C), 24.72 (35-C), 24.69 (35-C), 24.65 (35-C), 20.58 (32-C), 20.50 (32-C), 20.41 (32-C), 20.34 (32-C);

^{31}P{^{1}H} NMR (162 MHz, $CDCl_3$): δ (ppm) = 149.06, 148.68;

HR-MS (ESI+): m/z calc. ($C_{45}H_{49}N_6O_8PNa$, $[M+Na]^+$): 855.32417, found: 855.32473.

5.4.2.4 Oligonucleotide synthesis

Oligonucleotides were synthesized on an Applied Biosystems ABI 392 DNA/RNA synthesizer on a 0.6 µmol scale using standard phosphoramidite chemistry. DMT-dA(bz)-CPG, DMT-dC(ac)-CPG, DMT-dG(dmf)-CPG and DMT-dT-CPG with a pore size of 1000 Å and loading density of 25 – 35 µmol/g were used as solid supports and were obtained from Sigma Aldrich. For 3'-amino functionalized oligonucleotides 3' amino (TFA) modifier C-7 lcaa CPG (500 Å) from

Chem Genes Corporation was used. DMT-dA(bz), DMT-dC(ac), DMT-dG(dmf), DMT-dT were purchased from Chem Genes Corporation. Hex-5-yn-1-*O*-(2-cyanoethyl-*N*,*N*-diisopropyl)phosphoramidite[418] and 3-(4,4'-dimethoxytrityloxy)propyl-1-(2-cyanoethyl-N,N-diisopropyl)phosphoramidite[419] were synthesized as described previously. Solutions of standard phosphoramidites (70 mM) and synthesized phosphoramidites (80 mM) were prepared in anhydrous MeCN. The following solutions were used in the synthesis:

- Activator: 0.25 M ethylthiotetrazole (ETT) in anhydrous acetonitrile,
- Oxidation: 20 mM iodine in THF/water/pyridine (66:12:22, *v*/*v*/*v*)
- Cap A: pyridine/acetic anhydride/THF (10/10/80, *v*/*v*/*v*)
- Cap B: NMI in THF (84/16, *v*/*v*).
- Deprotection: 3% trichloro acetic acid in 1,2-dichloroethane

Cleavage from the solid support and removal of the base labile protecting groups was performed by treatment with concentrated ammonium hydroxide (33% NH_3) at 25°C overnight for ^{Phe}dU and 6 h for ^{Pym}dU containing oligonucleotides in a 1.5 mL screw-capped tube. ^{Pym}dU deprotection was monitored by HPLC to ensure complete deprotection. Short deprotection times were used to avoid side product formation with ^{Pym}dU. The solid support was filtered off and the solvent was removed under reduced pressure. The residue was dissolved in water. The crude product was purified by denaturing PAGE. Gels (0.7x200x300 mm) were prepared using a 15% or 20% acrylamide solution containing 7 M urea in 1x TBE. After polymerization the gels were run in 1x TBE buffer (89 mM Tris, 89 mM boric acid, 2 mM EDTA, pH 8.3) for 2.5 h at a constant power (35 W). Product bands were visualized using a TLC plate under UV illumination. Oligonucleotides were extracted with TEN buffer (10 mM Tris-HCl, 0.1 mM EDTA, 300 mM NaCl, pH 8.0) and recovered after precipitation with ethanol. DNA strands that could not be precipitated with ethanol were desalted by size exclusion chromatography using two HiTRAP desalting columns (5 mL each) from GE Healthcare on an ÄKTAstart purification system. Oligonucleotides were eluted with water (1 mL/min). After lyophilization the oligonucleotides were dissolved in a defined volume of water.

Purified oligonucleotides were analyzed by anion exchange HPLC using a GE Healthcare ÄKTApurifier with a DNAPac PA 200 column (2x250 mm) from Thermo Scientific at a flow rate of 0.5 mL min^{-1}. Linear gradients of 0–48% B over 24 min of buffer A (25 mM Tris-HCl, 6 M urea, pH 8.0) and buffer B (25 mM Tris-HCl, 0.5 M $NaClO_4$, 6 M urea, pH 8.0) were used for analysis. The chromatograms were monitored at 260 nm. All analyses were performed at

60°C. High resolution ESI mass spectra were recorded on a Bruker micrOTOF-Q III spectrometer.

5.4.2.5 Labeling of 5'-alkyne functionalized oligonucleotides

5'-Alkyne functionalized oligonucleotides were fluorescently labeled using copper(I)-catalyzed alkyne-azide cycloaddition (CuAAC). Sulfo-Cy3-azide was obtained by Lumiprobe GmbH. Freeze-dried DNA oligonucleotide (5 nmol) was dissolved in water (5 µL) and mixed with a DMSO/^{t}BuOH mixture (3:1, 3 µL). A solution of azide (0.63 µL, 50 mM) in DMSO/^{t}BuOH (3:1, *v*/*v*) was added. A freshly prepared solution of CuBr (0.63 µL, 100 mM) in DMSO/ ^{t}BuOH (3:1) was combined with a solution of tris(benzyltriazolylmethyl)amine (1.26 µL, 100 mM) in DMSO/^{t}BuOH (3:1) and then added to the reaction mixture. After incubation for 3 h in the dark at 37 °C, the reaction mixture was purified by PAGE (15% or 20% polyacrylamide).

5.4.2.6 Labeling of 3'-amino functionalized oligonucleotides

3'-Amino functionalized oligonucleotides were labeled with DABCYL using amine-reactive NHS esters. DABCYL-*N*-succinimidyl ester was purchased from Sigma Aldrich. Freeze-dried DNA oligonucleotide (10 nmol) was dissolved in sodium carbonate buffer (80 µL, 100 mM, pH 9.0) and DABCYL-*N*-succinimidyl ester in DMF (20 µL, 15 mM) was added. After 5 h of incubation in the dark at 37 °C, the reaction mixture was prepurified by PAGE (15% polyacrylamide). After ethanol precipitation, the labeled oligonucleotides were separated by RP-HPLC from the unreacted DNA. Purification was performed on a HPLC system (AS-4050, PU-4180, CO-4060, MD-4010) from JASCO using a NUCLEOSIL reversed-phase column (C18, 250x4.6 mm, 100 Å, 5 µm) from Machery Nagel at a flow rate of 1 mL min^{-1}. Linear gradients from 2–30% B over 25 min of buffer A (10 mM NH_4OAc, pH 6.6) and buffer B (MeCN) were used for purification. Chromatograms were monitored at 260 nm and 450 nm.

5.4.2.7 UV/Vis spectroscopy / Thermal denaturing experiments

UV/Vis spectra were measured in 10 mm quartz cuvettes from Agilent using an Agilent Cary 3500 with following settings:

- Wavelength: 200-750 nm
- Averaging time: 0.02 s
- Data interval: 0.2 nm

- Scan rate: 300 nm min^{-1}
- Spectral bandwidth: 2 nm

For the determination of the extinction coefficient of the ^{Phe}dU nucleoside a stock solution in DMSO (10 mM) was prepared. Spectra measurement for extinction coefficient determination started with a 50 µM dilution in water which was serially diluted six times with water in a 1:1 ratio (25.0, 12.5, 6.25, 3.13, 1.56, 0.78 µM) to prepare a concentration series. Spectra were recorded at 20 °C.

Temperature dependent UV/Vis spectra were recorded every 10 °C in a range of 10-90 °C. 1 µM DNA samples in phosphate buffer (100 mM NaCl, 10 mM sodium phosphate, pH 7.0) were heated for 1 min at 95 °C and stored at least 30 min at room temperature. The samples in the cuvettes were overlaid with silicone oil.

UV denaturing melting experiments were recorded on Varian Cary100 equipped with a 6x6 Multicell Block Peltier Series II cell changer and a VARIAN CARY Temperature Controller. Absorption was measured at 250, 260 and 280 nm. The absorption was recorded with a spectral bandwidth of 1 nm and the averaging time was set to 2 s. The temperature cycle was programmed as follows: 1. 20 °C to 90 °C; 2. 90 °C to 10 °C; 3.10 °C to 90 °C; 4. 90 °C to 10 °C; 5. 10 °C to 90 °C

The first ramp was performed for annealing and was not considered for further melting temperature analysis. The heating rate was set to 0.5 °C min^{-1}. 500 µL sample in phosphate buffer (100 mM NaCl, 10 mM sodium phosphate, pH 7.0) or MOPS buffer (100 mM $NaClO_4$, 10 mM MOPS, pH 7.1) for metal-mediated base formation with a duplex concentration of 1 µM, 2 µM and 5 µM was measured in 10 mm quartz cuvettes from VARIAN and 300µL sample in phosphate buffer (100 mM NaCl, 10 mM sodium phosphate, pH 7.0 with a duplex concentration of 10 µM and 20 µM were measured in 1 mm quartz cuvettes from HELLMA. The samples in the cuvettes were overlaid with silicone oil.

The obtained curves were fitted in a two-state transition model with upper and lower limit to obtain the melting temperature T_m. Absorption of the melted duplexes was used to recalculate $c_{total.}$ For the estimation of the thermodynamic parameters melting curves were analyzed according to Breslauer et al:[420]

$$\frac{1}{T_m} = \frac{R(n-1)}{\Delta H^0} \ln c_{total} + \frac{\Delta S^0 - (n-1) \cdot R \cdot \ln 2n}{\Delta H^0} \tag{5.1}$$

Assuming a bimolecular association of two non-self-complementary strands leads to a molecularity $n = 2$:

$$\frac{1}{T_m} = \frac{\mathrm{R} \ln c_{total}}{\Delta H^0} + \frac{\Delta S^0 - R \cdot \ln 4}{\Delta H^0} \tag{5.2}$$

Using van't Hoff analysis enthalpy ΔH^0 could be obtained from the slope and entropy ΔS^0 from the intercept of a linear fit of ln c_{total} vs $1/T_{\text{m}}$. Afterwards the free energy ΔG^{298} was calculated using the Gibbs-Helmholtz equation with T = 298 K:

$$\Delta G^{298} = \Delta H^0 - T\Delta S^0 \tag{5.3}$$

5.4.2.8 DNA-Crosslinking procedures for $^{\text{Phe}}$dU and $^{\text{Pym}}$dU modified DNA

For analytic analysis, a 1 µM DNA duplex solution was prepared in phosphate buffer (100 mM NaCl, 10 mM sodium phosphate, pH 7.0). The duplex was annealed in the dark by heating to 95°C for 1 min and cooling down on ice for 60 min.

Irradiation with a Xe-lamp (150 W) was performed on a JASCO FP-8300 spectrofluorometer equipped with an FCT-817S cell changer and a F12 temperature controller device from Julabo. All measurements were performed in an FMM-200 5 mm quartz fluorescence microcell from JASCO with a magnetic stir bar using an FMH-802 5 mm microcell jacket from JASCO as cell adapter. The temperature was set at 10 °C or 15 °C if the duplex contains a Cy3 label. After 5 min equilibration in the spectrometer the reaction mixture (300 µL) was irradiated with the following settings:

- Ex wavelength: 320 nm
- Em wavelength: 410 nm
- Ex bandwidth: 5 nm
- Em bandwidth: 5 nm
- Response: 1 s
- PMT:voltage 350 V
- Data interval: 1 s

Aliquots (5 µL) were taken at different time points (0, 5, 10, 15, 20, 25 and 30 min for 12mer DNA; 0, 1, 2, 3, 4, 5 and 6 min for 20mer DNA or 0.5, 1, 2, 3, 4 and 5 min for the Y-fork) and quenched with loading buffer (5 µL).

For the irradiation with the transilluminator a device from HEROLAB GmbH with an irradiation maximum of 312 nm was used. The reaction mixture (60 µL) was transferred to a PCR tube and cooled again on ice. The DNA duplex was irradiated immediately after placing the reaction tube on the transilluminator. Aliquots (3 µL) were taken at different time points (0, 10, 20, 30, 40, 50 and 60 sec) and quenched with loading buffer (3 µL).

For HPLC characterization, 300 µL of a 1 µM unlabeled DNA duplex solution was lyophilized after 30 min irradiation with a Xe-lamp and analyzed by anion exchange HPLC.

The kinetics were analyzed by denaturing PAGE (0.4x200x200 mm, 10-20% acrylamide/ bisacrylamide 19:1, 7 M urea) with 1×TBE and run at 25 W power for 1 hour. The gels were imaged on a Chemidoc device, and the crosslink yield was determined from the band intensities of the Cy3 fluorescence quantified by ImageLab software. Kinetic data were obtained by fitting the resulting curve to the following equation:

$$Y = Y_{max} \cdot (1 - e^{-k_{obs} \cdot t}) \tag{5.4}$$

For a preparative crosslinking reaction, a concentration up to 40 µM was used in 1mL phosphate buffer (100 mM NaCl, 10 mM sodium phosphate, pH 7.0). After the annealing step the sample was irradiated with a Xe-lamp (150 W) in the fluorescence spectrometer. Anion exchange HPLC showed a conversion of 84 % after 2.5 h of irradiation. The reaction mixture was lyophilized and purified by PAGE (20%).

The CL Y-fork substrate was prepared using a 5 µM in 300 µL phosphate buffer (100 mM NaCl, 10 mM sodium phosphate, pH 7.0). The Y-fork was annealed in the dark by heating to 95°C for 5 min and cooling down to room temperature for 60 min. After 5 min irradiation with an Xe Lamp (150 W) in the fluorescence spectrometer the sample was lyophilized and purified by PAGE (10%).

5.4.2.9 Metal induced inhibition of ICL

All experiments with expected metal-mediated base formation were performed in a MOPS buffer (100 mM $NaClO_4$, 10 mM MOPS, pH 7.1) to prevent the formation of water insoluble salts.

A 1 µM DNA duplex solution of the oligonucleotides Cy3-^{Phe}dU1_C and ^{Phe}dU2_C or Cy3-^{Phe}dU1_T and ^{Phe}dU2_T was prepared in MOPS buffer containing 2 µM of the following salts to screen for possible metal-induced inhibition: $MnCl_4$ $4H_2O$, $FeCl_2$, $Fe(ClO_4)_3$, $CoCl_2$ $6H_2O$, $NiCl_2$ $6H_2O$, $CuSO_4$ $5H_2O$, $ZnCl_2$, $AgNO_3$ and $SnCl_2$ $2H_2O$. The duplex was annealed in the dark by heating to 95°C for 1 min and cooling down on ice for 60 min. 10 µL of the reaction mixtures were irradiated with the transilluminator for 1 min in a PCR tube and analyzed by denaturing PAGE (20%).

For silver mediated inhibition analyses, a 1 µM DNA duplex solutions of the oligonucleotides Cy3-^{Phe}dU1_C and ^{Phe}dU2_C were prepared in MOPS buffer containing 0, 2 or 20 µM $AgNO_3$ and annealed by heating to 95°C for 1 min and cooling down on ice for 60 min. 300 µL of the reaction sample was irradiated with a Xe-lamp (150W) in the fluorescence spectrometer at 15 °C. Aliquots (5 µL) were taken at different time points (0, 5, 10, 15, 20, 25 and 30 min) and quenched with loading buffer (5 µL). After 30 min of irradiation 35 µL of a NaCl solution (771 µM in MOPS buffer) was added and the sample was irradiated for another 30 min. Aliquots (5 µL) were taken at different time points (35, 40, 45, 50, 55 and 60 min), quenched with loading buffer (5 µL) and analyzed by denaturing PAGE (20%).

5.4.2.10 NMR spectroscopy of DNA duplexes

After PAGE purification, all the oligonucleotides were precipitated overnight at -20°C with five volumes of $LiClO_4$ 2% w/v in acetone. The pellet was then dissolved in ddH_2O and lyophilized to remove residual acetone. The complementary oligonucleotides were annealed in the final volume of the NMR sample (180 µL) by heating to 95°C for 5 minutes and slowly cooling down to room temperature overnight. Samples used for NMR assignments contained 0.2 mM DNA duplex and were dissolved in 90% H_2O/10% D_2O NMR buffer (10 mM sodium phosphate buffer at pH 7.0, 100 mM NaCl). 3-(Trimethylsilyl)-1-propanesulfonic acid (DSS) was added to all the samples as an internal reference. Water suppression was achieved using either jump-return-

echo (excitation maximum on the mid of the imino region) or excitation sculpting with gradients.[425-426] All the NMR experiments were performed on a Bruker Avance III 600 NMR spectrometer equipped with a DCH ^{13}C / ^{1}H cryoprobe.

The NMR spectra were acquired and processed using the Topspin 3.2 software (Bruker BioSpin, Germany).

5.4.2.11 Enzymatic Digestion and LC-MS analysis

For LC-MS characterization PAGE purified and lyophilized crosslinked DNA duplex (500 pmol) was dissolved in antarctic phosphatase buffer (10 µL, 50 mM Bis-Tris-Propane-HCl,1 mM $MgCl_2$, 0.1 mM $ZnCl_2$, pH 6, New England Biolabs) and Nuclease P1 (0.25 µL, 25 U, New England Biolabs), Phosphodiesterase I (0.5 µL, 0.05 U, Sigma Aldrich) and antarctic phosphatase (0.2 µL, 1 U, New England Biolabs) were added and the mixture was kept at 37 °C overnight. The digested product was diluted with water (40 µL), extracted with chloroform (2x25 µL) and analyzed by LC-MS using a RP-18 column (Synergi, 4-µm Fusion-RP C18 80 Å, 250x2mm) from Phenomenex at 25 °C with aqueous mobile phase A (10 mM NH_4OAc, pH 5.3) and organic mobile phase B (100% acetonitrile). The flow rate was 0.2 mL/min with a gradient of 0–80% B in 40 min. The micrOTOF-Q III with an ESI ion source was operated in positive-ion mode, with a capillary voltage of 4.5 kV, an end plate offset of 500 V, a nitrogen nebulizer pressure of 1.4 bar, a dry gas flow of 9 L/min and dry temperature of 200 °C. Data were analyzed using Data Analysis software DA 4.2 (Bruker Daltonics).

For NMR analysis of the digested crosslink unit the enzymatic digestion was upscaled. PAGE purified and lyophilized crosslinked DNA duplex (11 nmol) was dissolved in Antarctic phosphatase buffer (210 µL, 71.4 mM bis-tris-propane-HCl,1.4 mM $MgCl_2$, 0.14 mM $ZnCl_2$, pH 6, New England Biolabs) and Nuclease P1 (6 µL, 600 U, New England Biolabs), phosphodiesterase I (12 µL, 1.2 U, Sigma Aldrich) and Antarctic phosphatase (5 µL, 25 U, New England Biolabs) were added and the mixture was kept at 37 °C overnight. The digested product was diluted with water (800 µL) and extracted with chloroform (2x200 µL). If RP-HPLC indicated incomplete digestion the sample was lyophilized, and digestion was repeated. To obtain a suitable amount for 2D NMR experiments the digestion was performed 7 times. Purification was performed on a HPLC system (AS-4050, PU-4180, CO-4060, MD-4010) from JASCO using a NUCLEOSIL reversed-phase column (C18, 250x4.6 mm, 100 Å, 5 µm) from Machery Nagel at a flow rate of 1 mL min^{-1}. Linear gradients of 2–40% B over 35 min of buffer A (water) and buffer

B (MeCN) were used for purification. Chromatograms were monitored at 260 nm and 320 nm. The purified crosslink units were lyophilized and dissolved in D_2O and analyzed on a Bruker Avance III 600 NMR spectrometer equipped with a DCH ^{13}C / ^{1}H cryoprobe. NMR spectra were evaluated using MestReNova v12.0.4 or Topspin 3.2 (Bruker BioSpin).

5.4.2.12 Fluorescence anisotropy

Fluorescence anisotropy measurements were performed with 5 nM Cy3 labeled Y-fork substrate in HEPES buffer (20 mM HEPES, 50 mM KCl, 5 mM $MgCl_2$,1 mM TCEP, pH 7.5) at room temperature. The XPD/p44/p62 protein complex was prepared as described previously[484] and used at a protein concentration of 5000 nM, followed by serial dilution eight times with HEPES buffer in a 1:1 ratio (resulting in 2500, 1250, 625, 313, 156, 78.1, 39.1, 19.5 nM protein complex). In addition, fluorescence anisotropy was measured in the absence of protein. A sample volume of 50 µL was used for each well in a 96-well plate. Fluorescence was detected at an excitation wavelength of 540 nm and an emission wavelength of 590 nm using a Clariostar plate reader (BMG labtech). The gain was adjusted to a well containing buffer and DNA but no protein. K_D was obtained by fitting the resulting curve to the following equation:

$$A = A_{min} + (A_{max} - A_{min}) \cdot \frac{1}{2\,[L]} \cdot (([L] + K_D + [R]) - \sqrt{([L] + K_D + [R])^2 - 4 \cdot [L] \cdot [R]}) \quad (5.5)$$

With:

$\boldsymbol{A}$= anisotropy

A_{min} = minimal anisotropy

A_{max} = maximal anisotropy

$[\boldsymbol{L}]$ = ligand concentration

$[\boldsymbol{R}]$ = receptor concentration

5.4.2.13 Electrophoretic Mobility Shift Assay (EMSA)

Samples (10 µL) were prepared in a HEPES buffer (20 mM HEPES, 50 mM KCl, 5 mM $MgCl_2$,1 mM TCEP, pH 7.5). The concentration of the DNA substrate was 100 nM and the concentration

of the protein complex started at 2 µM, which was serially diluted four times with HEPES buffer in a 1:1 ratio (1, 0.5, 0.25, 0.13 µM) to prepare the concentration series. A reference containing only the DNA substrates was also prepared. After 30 min at room temperature loading dye (3 µL) was added and analyzed by EMSA. EMSA gels were prepared using a 6% acrylamide solution containing 25 mM TRIS, 192 mM glycine, 0.1 % APS and 0.1 % TEMED. After polymerization, the gels were run with in native PAGE running buffer (25 mM Tris, 192 mM glycine) for 40 min with 100 V.

5.4.2.14 *In vitro* helicase assay

For the helicase assay, fluorophore/quencher labeled DNA Y-fork substrates were used at a constant concentration of 125 nM. The experiments were performed in HEPES buffer (20 mM HEPES, 50 mM KCl, 5 mM $MgCl_2$,1 mM TCEP, pH 7.5). Protein concentration started with a 1000 nM dilution which was serially diluted five times with HEPES buffer in a 1:1 ratio (500, 250, 125, 62.5 nM) to prepare a concentration series. Samples were preincubated in a 96-well plate in a FLUOstar Optima plate reader (BMG labtech) at 37 °C. Helicase activity was initiated by the addition of a 25 mM ATP solution (10 µL) to give a concentration of 5 mM ATP in 50 µL reaction volume. Fluorescence was detected with excitation at 540 nm and an emission wavelength of 590 nm. Results were normalized using following equation:

$$F_{\mathrm{norm}} = \frac{F - F_0}{F_0} \tag{5.6}$$

With:

F_{norm} = normalized fluorescence

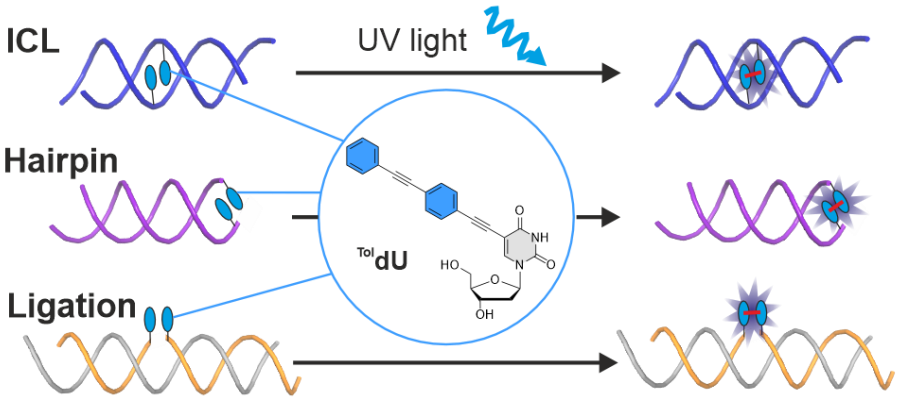
F = measured fluorescence

F_0 = starting fluorescence

5.4.3 Appendix

The characterization of the synthesized products (NMR, HPLC) and the melting curves can be found in the published supporting information.[487]

6 A tolane-modified 5-ethynyluridine as universal and fluorogenic photochemical DNA crosslinker



This chapter and the corresponding supporting information were published in:

H. Neitz, C. Höbartner, Chem. Commun. 2023, *59*, 12003–12006.
(https://pubs.rsc.org/en/content/articlelanding/2023/cc/d3cc03796g)

Published by The Royal Society of Chemistry.

Abstract: We report the fluorescent nucleoside ^{Tol}dU and its application as a photoresponsive crosslinker in three different DNA architectures with enhanced fluorescence emission of the crosslinked products. The fluorogenic ^{Tol}dU crosslinking reaction enables the assembly of DNA polymers in a hybridization chain reaction for the concentration-dependent detection of a specific DNA sequence.

6.1 Introduction

Fluorescent nucleobase analogues are of great interest in nucleic acid chemistry.[218, 488] The lack of fluorescence in natural nucleosides provides the opportunity for developing novel reporters,[489] which are sensitive to the local environment, and are useful as biochemical tools for structural characterization,[343, 490-491] detection of single nucleotide polymorphisms (SNPs)[357, 492] or enzyme activities.[361-362] The attachment of fluorescent dyes to natural nucleosides offers the advantage of a precise placement of the chromophore in the DNA/RNA helix,[454] which is also a requirement for proximity induced reactions.[185] The DNA double helix is an outstanding template for organizing reactive groups in suitable arrangements for the formation of a covalent bond between two strands, either in the form of interstrand crosslinking or by sealing a nick in the templated ligation of two single strands hybridized to a matched complementary strand.[219] Of particular interest are photo-controlled crosslinking and ligation reactions, which do not require any additional reagents and enable precise control of product formation.[468] Several examples have been reported for photocontrolled ligation[386, 471, 493] and interstrand crosslinking (ICL) reactions.[383, 385, 390, 494]

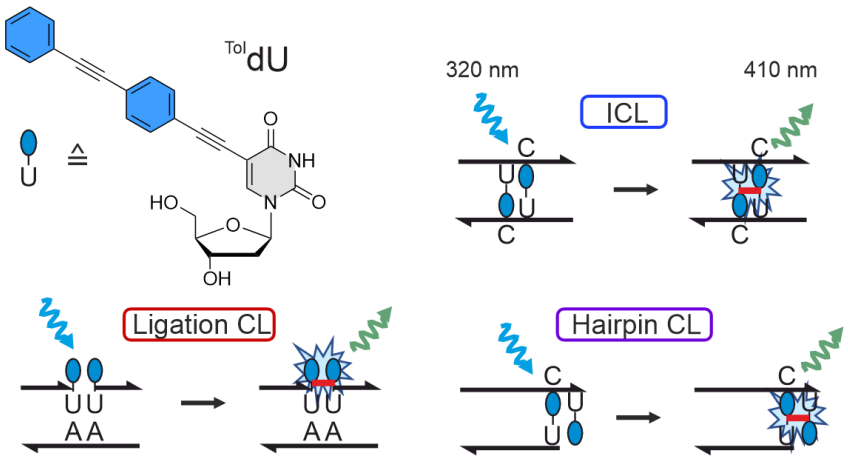


Figure 6.1 Structure of ^{Tol}dU nucleoside and three DNA architectures for ^{Tol}dU crosslinking investigated in this work: internal interstrand crosslink (ICL), terminal interstrand crosslink for hairpin formation, and linear ligation crosslink.

However, such reactions rarely lead to increased fluorescence.[365-366] A fluorescence-enhancing templated reaction was recently reported that enabled the multiple-turnover transformation of profluorophores for the detection of DNA sequences.[367] In this work, we report a fluorogenic photoinduced DNA crosslinking reaction that is compatible with concentration-dependent signal amplification in a hybridization chain reaction using 5-[4-

(phenylethynyl)phenylethynyl]-2'-deoxyuridine (^{Tol}dU) as a novel fluorescent nucleoside analogue. When incorporated into three different DNA architectures, we found that two adequately arranged ^{Tol}dU units undergo an efficient light-induced intermolecular crosslinking reaction that results in enhanced fluorescence emission (see Figure 6.1).

6.2 Results and discussion

6.2.1 ^{Tol}dU as fluorescent DNA nucleotide

Phenylene ethynylene units have been introduced as rigid hydrophobic moieties into oligonucleotides for the synthesis of DNA amphiphiles[203] and as energy-transfer spacers for the attachment of fluorophores in the major groove of DNA.[495] Here we investigated ^{Tol}dU as new fluorescent nucleoside, which formally has a tolane unit linearly attached to 5-ethynyl-2'-deoxyuridine. The ^{Tol}dU building blocks for solid-phase synthesis were prepared via Sonogashira crosscoupling of 1-ethynyl-4-(2-phenylethynyl)benzene with 5'-O-DMT protected 5-iodo-2'-deoxyuridine (see Scheme S6.1), and converted to the 3'-cyanoethyl diisopropyl phosphoramidite or attached as succinyl ester to long-chain alkyl amino controlled pore glass solid support. The free ^{Tol}dU nucleoside was obtained by removal of the DMT group and showed an absorption maximum at 328 nm and a fluorescence emission maximum at 397 nm with a quantum yield of 38% in DMSO (see Figure S6.1-S6.2).

The ^{Tol}dU moiety was incorporated into DNA strands using standard DNA phosphoramidite chemistry. The DNA sequences and HR-ESI MS data are given in Table S6.1. The presence of the ^{Tol}dU unit was also confirmed by the additional absorption peak around 330 nm in the UV spectra of single strands and ^{Tol}dU containing duplexes (see Figure S6.3). Thermal melting curves were monitored at 260 nm and revealed a melting temperature of 37.2 °C for the fully complementary duplex, *i.e.* with adenosine opposite to ^{Tol}dU (see Figure S6.4 and Table S6.2). This is a slight destabilization of 2.8 °C compared to the analogous duplex with an unmodified dA/dT base pair, but circa 10 °C more stable than the duplexes with mismatched nucleobases or a non-nucleosidic alkyl linker facing ^{Tol}dU (see Figure S6.4A). The fluorescence of ^{Tol}dU was examined in matched and mismatched duplexes. The highest fluorescence intensity was observed when ^{Tol}dU was paired with adenosine (see Figure S6.4B), suggesting a rigid orientation in the base-pair, while the mismatched cases resulted in more flexible orientations reflected

in a lower emission intensity. The same trend was confirmed by relative quantum yields and fluorescence lifetimes (see Figure S6.5-S6.6and Table S6.3-S6.4).

6.2.2 Photochemical ^{Tol}dU interstrand crosslinking (ICL)

The observed mismatch sensitivity of ^{Tol}dU suggested that hydrophobic stacking interactions between tolane units could be favored when two ^{Tol}dU nucleosides are incorporated in neighbouring mismatched base pairs.

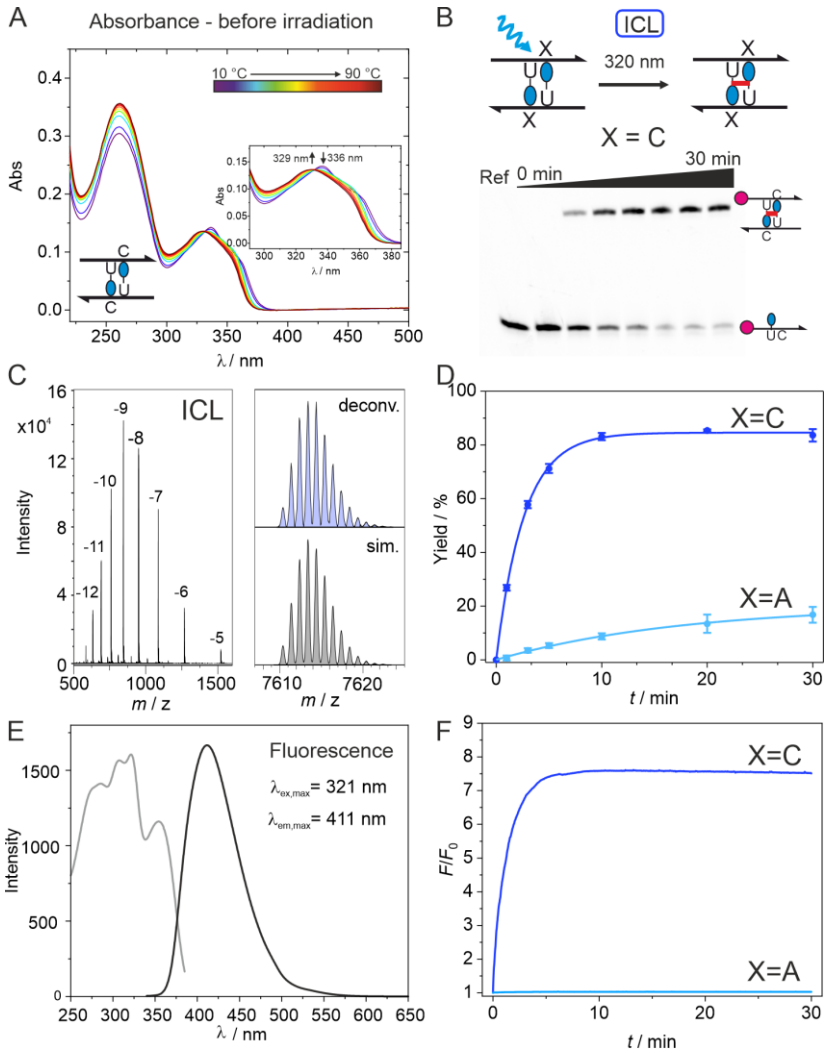


Figure 6.2 (A) Temperature-dependent UV/Vis spectra of double-modified ^{Tol}dU/dC duplex reveal hydrophobic stacking interactions at low temperatures in the folded duplex. (B) Schematic representations of the duplex for ICL formation and denaturing PAGE analysis of aliquots removed at specific time intervals. The magenta sphere represents the Cy3 label. (C) ESI-MS of isolated ICL product (left), deconvoluted and simulated mass spectra (right). Calculated monoisotopic mass: 7610.3804 Da, found 7610.3826 Da (error −0.3 ppm). (D) Kinetics of ICL formation *via* PAGE analysis (see gel image in B). (E) Fluorescence excitation and emission spectra of isolated ICL product (*i.e.*, ICL-containing DNA duplex without Cy3 label). (F) Kinetics of fluorescence increase at 410 nm upon irradiation reports the formation of fluorescent ICL when X = C, but not when X = A. Conditions for kinetics in B, D, F: 1 µM DNA in 10 mM sodium phosphate, 100 mM NaCl, pH 7.0.

Indeed, stacking interactions were observed by temperature-dependent UV/Vis spectroscopy, in analogy to earlier reports with diphenylacetylene[443] or 4,4'-substituted tolanes[406] incorporated into the DNA backbone. For interacting ^{Tol}dU units in our DNA duplexes, the absorption maximum shifted slightly to shorter wavelengths (from 336 nm to 329 nm) upon denaturation of the interaction at higher temperatures (see Figure 6.2A), while no temperature-dependent change was observed when both ^{Tol}dU units were base-paired to adenosine (see Figure S6.7A).Upon irradiation of the duplexes containing two hydrophobic ^{Tol}dU nucleosides in pyrimidine mismatches, we observed efficient crosslink formation. With one of the ^{Tol}dU DNA strands containing a Cy3 label for visualization on denaturing PAGE, more than 80% conversion to the slower-migrating crosslinked product occurred after 10 min of irradiation at 320 nm (see Figure 6.2B and Figure S6.8). The reaction was scaled up, and the crosslinked product was isolated by PAGE. ESI-MS confirmed the formation of the crosslinked product consisting of the two single strands without addition of any other compound (see Figure 6.2C and Figure S6.9). This observation is consistent with our recent findings with 5-phenylethynyl-2'-deoxyuridine (^{Phe}dU) nucleosides, for which we found that a stacked dimer in a pyrimidine mismatched orientation easily forms an interstrand crosslink via an alkene–alkyne [2+2] photocycloaddition.[487]

Interestingly, the new ^{Tol}dU ICL product showed an emission maximum at 411 nm with a quantum yield of 3.9% in the duplex (see Figure 6.2E and Figure S6.10). The time-dependent increase at the emission wavelength of 410 nm (see Figure 6.2F) showed similar kinetics to the PAGE analysis, and the final fluorescence intensity was up to seven times higher than before crosslinking. We interpret this increase in fluorescence as a result of the enhanced rigidity in the crosslinked product, thus reducing the degree of freedom for thermal non-radiative deactivation of the excited state. To prove the necessity of the mismatched dC/^{Tol}dU motif, we examined a duplex sample containing dA/^{Tol}dU base pairs instead of the mismatch. The base-paired situation keeps the tolane units oriented towards the major groove of the helix, thus preventing inter-tolane stacking interactions. Indeed, the crosslinking yield was strongly reduced (only low conversion <20% after 30 min), and no significant fluorescence increase was observed for the fully base-paired ^{Tol}dU duplex (see Figure 6.2D and Figure 6.2F, X = A).

To further analyse the ICL product, the crosslinked duplex was enzymatically digested, and the nucleoside mixture was analyzed by LC-MS after extraction of the hydrophobic nucleoside dimer into the organic phase (see Figure S6.11-S6.12). Both MALDI and ESI-MS confirmed the formation of the product containing two ^{Tol}dU units.

6.2.3 DNA hairpin formation via terminal DNA crosslinking

Encouraged by the efficient ICL formation with ^{Tol}dU in the middle of the duplex, we explored the generality of the approach, and moved the ^{Tol}dU units to the end of the duplex. Photoinduced crosslinking would then result in the formation of an artificial DNA hairpin that is capped by a covalent ^{Tol}dU dimer. We synthesized one DNA strand on ^{Tol}dU-CPG, *i.e.* placing the modification at the 3' end. In a complementary DNA oligonucleotide, we incorporated the ^{Tol}dU nucleoside at the 5' end.

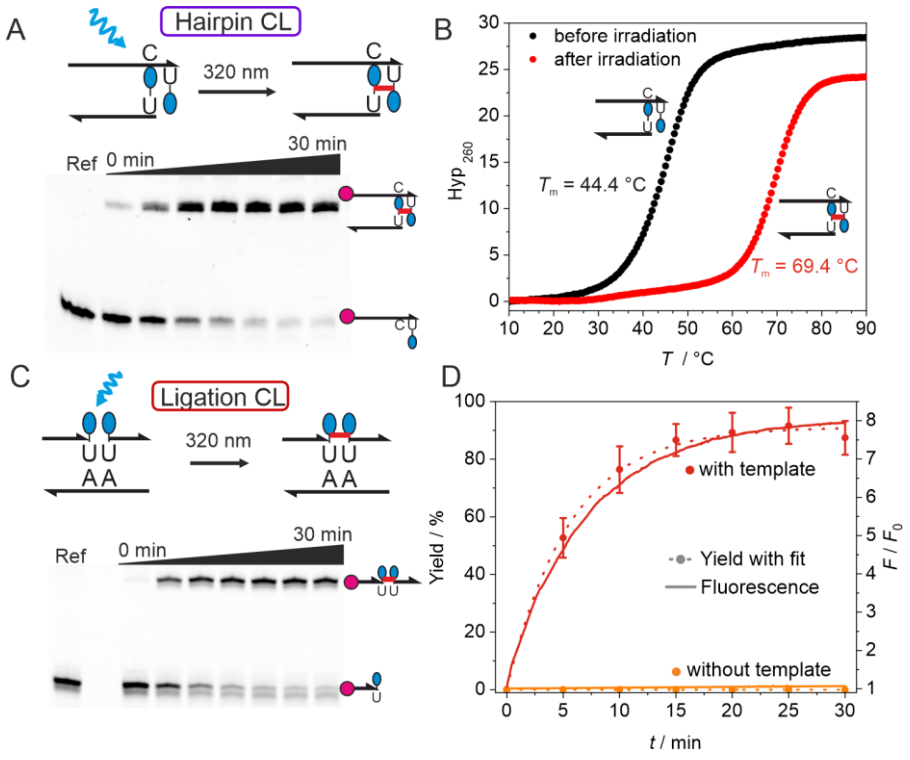


Figure 6.3 (A) PAGE analysis of terminal crosslinking reaction resulting in a DNA hairpin. (B) UV thermal melting curves before and after crosslinking monitored at 260 nm. (C) Schematic representations of the nicked duplex for photoinduced ligation of ^{Tol}dU. (D) Kinetics of ligation analyzed via PAGE (left *y*-axis) and fluorescence increase at 410 nm (right *y*-axis), with or without template. (1 µM DNA in 10 mM sodium phosphate, 100 mM NaCl, pH 7.0)

The two strands were hybridized and crosslink formation upon irradiation at 320 nm was analyzed by PAGE and fluorescence spectroscopy (see Figure 6.3 and Figure S6.8). Full conversion to a slower migrating product was observed within 10 min, showing again an eight times

higher fluorescence emission at 410 nm. The crosslinked product was isolated and characterized by ESI-MS (see Figure S6.9), revealing the expected mass of the DNA hairpin as the sum of the two ^{Tol}dU containing single strands. The thermal melting curve of the crosslinked ^{Tol}dU hairpin showed an impressive stabilization by 25 °C (see Figure 6.3B), comparable to artificial aromatic hairpin caps, such as for example a perylenebisimide covalently connected to the DNA phosphodiester backbone.[496]

6.2.4 DNA-templated photochemical DNA ligation

In a third DNA architecture for ^{Tol}dU crosslinking, we placed the ^{Tol}dU nucleosides at the 5' and 3' ends of a 12 and a 13-mer oligonucleotide, respectively. The two strands were hybridized to a fully complementary 25-nt long DNA template (see Figure 6.3C). In this setup, the tolane units were arranged directly next to each other, both facing the major groove of the nicked duplex. Upon irradiation at 320 nm, a linear ligation product was formed by sealing the nicked 25-mer duplex. This is also confirmed by ESI-MS of the isolated ligation product (see Figure S6.9), and by a single melting transition at 59.6 °C of the DNA duplex after irradiation, *i.e.* containing the ligated product (see Figure S6.14). Again, the crosslinking reaction could also be monitored by fluorescence spectroscopy, and no ligation reaction was observed in the absence of the template (see Figure 6.3D and Figure S6.15). A structural model of the ^{Tol}dU units in a nicked double helix suggests a possible orientation for formation of a parallel ^{Tol}dU dimer by alkene alkyne photocycloaddition (see Figure S6.13).

6.2.5 ^{Tol}dU as fluorogenic crosslinker for HCR assemblies

The demonstrated versatility of the ^{Tol}dU crosslinking reactions suggests many different applications in bionanotechnology, including the synthesis of DNA nanomaterials.[185] Here we explored if ^{Tol}dU and the dC/^{Tol}dU ICL motif are compatible with a hybridization chain reaction (HCR). HCR is an isothermal assembly process of two DNA hairpins into a nicked double helix by a cascade of hybridization events that is triggered by an initiator strand.[130, 136, 354-355] To investigate ^{Tol}dU-HCR, we designed two hairpin structures which contained two dC/^{Tol}dU mismatches in the stem region.[132, 268] With the addition of an initiator strand a DNA duplex polymer will be formed, in which two ^{Tol}dU units originally placed in the two different hairpins will form the ^{Tol}dU dimer crosslink motif (see Figure 6.4A). This will allow the crosslinking reaction to take place upon irradiation and an increase in fluorescence should be observed.

The setup was first designed and tested with unmodified DNA (see Figure S6.16-S6.17), and then we incorporated ^{Tol}dU into the hairpin strands (see Figure S6.18). Analysis by agarose gel electrophoresis showed the successful formation of an approximately 2000 bp long product upon addition of 10% of the initiator strand, while almost no duplex formation was observed for the hairpins alone in the absence of initiator. Upon irradiation of the HCR sample on a transilluminator light source, the covalent crosslinking reaction was initiated, and the supramolecular DNA polymer was transformed to a covalent polymer, that was further analyzed by denaturing PAGE and visualized by staining with Sybr gold. We investigated various initiator concentrations and irradiation times and analyzed HCR and crosslinking efficiency by denaturing PAGE. Covalently linked products with a size of more than 1000 bp were observed, indicating the effective crosslinking reaction of at least 40 dimeric dC/^{Tol}dU motifs have taken place after short irradiation of only two minutes (see Figure 6.4B).

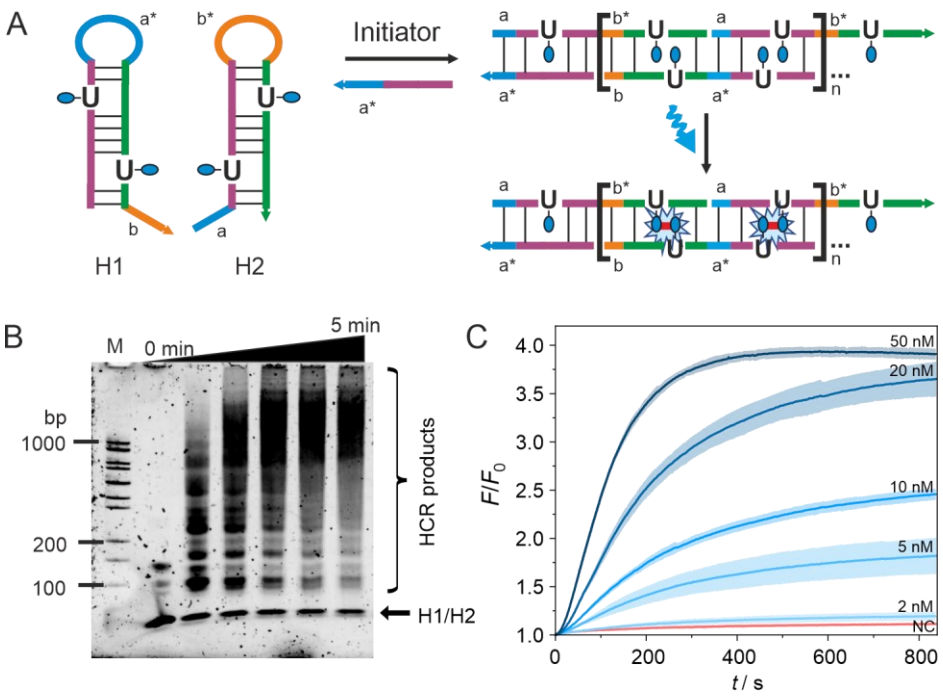


Figure 6.4 (A) Illustration of the HCR with ^{Tol}dU containing hairpins H1 and H2. (B) Denaturing PAGE (5%) of the irradiated HCR product at different timepoints (500 nM hairpins with 50 nM initiator in 75 mM sodium citrate, 750 mM NaCl, pH 7.0). (C) Fluorescence increase at 410 nm after addition of initiator DNA strand at 2 5, 10, 20 or 50 nM concentration, NC is negative control strand at 50 nM (100 nM hairpins in 75 mM sodium citrate buffer, 750 mM NaCl, pH 7.0).

6.2.6 DNA sensing via fluorogenic DNA polymerization

Next, we examined the combined ability of covalent crosslink formation and fluorescence enhancement for the sensitive detection of DNA sequences by HCR. In order to use ^{Tol}dU-HCR for DNA sequence detection, the fluorescence increase after addition of the initiator strand

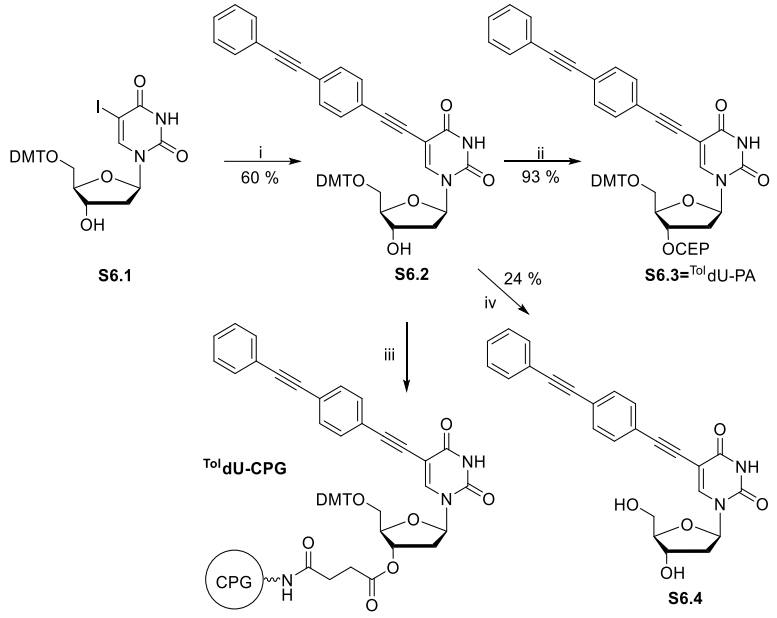
(*i.e.* the analyte) was measured under constant irradiation at 320 nm. The data in Figure 6.4C confirm the concentration-dependent response of the ^{Tol}dU-HCR DNA sensing platform. At an initiator concentration of 5 nM, the fluorescence slowly increased to *ca.* 180% within 15 min, while increasingly faster rates and higher final intensities were recorded at 10, 20 and 50 nM analyte concentration, reaching up to 400% within 5 min (see Figure 6.4C). Importantly, no fluorescence increase was observed upon addition of a random DNA sequence that was not complementary to the toehold in H2 (negative control NC, red in Figure 6.4C). These results demonstrate a new variant of fluorescence-based DNA detection is solution, that has a shorter response time and works at lower hairpin probe concentrations than a previously reported HCR setup based on pyrene excimer formation.[354]

6.3 Conclusion

In summary, we presented ^{Tol}dU as new nucleobase analogue with distinct fluorescent properties depending on Watson-Crick base pairing status and arrangement in the DNA duplex. The hydrophobic stacking interactions of two properly positioned ^{Tol}dU moieties enabled the photoinduced crosslinking reaction in three different arrangements at internal or terminal positions of the hybridized DNA duplex. In addition, enhanced fluorescence emission was observed upon crosslink formation, suggesting that the new ^{Tol}dU dimer is rigidly embedded in the resulting DNA structure. Our analytical data are consistent with a photoinduced [2+2] cycloaddition of an alkyne of one ^{Tol}dU unit and the C5-C6 double bond of the other ^{Tol}dU nucleobase. As a proof-of-principle application we showed that ^{Tol}dU photo crosslinking is compatible with HCR and the fluorescence enhancement can be used for sensitive DNA sequence detection. Moreover, we anticipate that the universal nature of the new crosslink makes it a useful tool in biochemistry and beyond, for example as an alternative to enzymatic ligation or to build covalently linked DNA nanostructures.[497-498]

6.4 Supporting information

6.4.1 Additional data



Scheme S6.1 Synthesis of the ^{Tol}dU nucleoside, phosphoramidite and solid support. i) 1-Ethynyl-4-(2-phenylethynyl)benzene, CuI, $Pd(PPh_3)_4$, Et_3N/DMF, 60 °C, over night; ii) CEP-Cl, DIPEA, DCM, r.t., 4 h; iii) 1.succinic anhydride, DMAP, DCM, r.t., 4 h, 2. LCAA-CPG, BOP, *N*-methylimidazole, MeCN r.t. over night; iv) TFA, DCM, r. t.,15 min.

Table S6.1 Sequences and ESI-MS results of the used DNA oligonucleotides.

Oligo	5'-Sequence-3'	Chemical formula	Calc. Mass	Measured Mass
ODN1_A	GATGATAGCTAG	$C_{119}H_{148}N_{49}O_{69}P_{11}$	3707.66977	3707.69796
ODN1_C3	GATGAT(C3)GCTAG	$C_{112}H_{143}N_{44}O_{68}P_{11}$	3532.61981	3532.62184
ODN1_C	GATGATCGCTAG	$C_{118}H_{148}N_{47}O_{70}P_{11}$	3683.65853	3683.68621
ODN1_G	GATGATGGCTAG	$C_{119}H_{148}N_{49}O_{70}P_{11}$	3723.66468	3723.66893
ODN1_T	GATGATTGCTAG	$C_{119}H_{149}N_{46}O_{71}P_{11}$	3698.65820	3698.70495
^{Tol}dU1_A	GATGA(^{Tol}dU)AGCTAG	$C_{134}H_{154}N_{49}O_{69}P_{11}$	3893.71617	3893.73041
Alk-^{Tol}dU1_A	Alk-GATGA(^{Tol}dU)AGCTAG	$C_{140}H_{163}N_{49}O_{72}P_{12}$	4053.74510	4053.74251
^{Tol}dU2_A	CTAGC(^{Tol}dU)ATCATC	$C_{131}H_{155}N_{40}O_{71}P_{11}$	3764.68616	3764.70338
^{Tol}dU1_C	GATGA(^{Tol}dU)CGCTAG	$C_{133}H_{154}N_{47}O_{70}P_{11}$	3869.70493	3869.70721
Alk-^{Tol}dU1_C	Alk-GATGA(^{Tol}dU)CGCTAG	$C_{139}H_{163}N_{47}O_{73}P_{12}$	4029.73387	4029.74938
^{Tol}dU2_C	CTAGC(^{Tol}dU)CTCATC	$C_{130}H_{155}N_{38}O_{72}P_{11}$	3740.67492	3740.68243
5'-^{Tol}dU-Lig/Hairpin	(^{Tol}dU)GCTATCATATG	$C_{133}H_{156}N_{41}O_{72}P_{11}$	3819.69197	3819.70602
Hairpin-3'-^{Tol}dU	Alk-CATATGATAGCC(^{Tol}dU)	$C_{148}H_{176}N_{47}O_{79}P_{13}$	4277.77939	4277.79279
Lig-3'-^{Tol}dU	Alk-AGGTCCGGACTG(^{Tol}dU)	$C_{148}H_{175}N_{50}O_{80}P_{13}$	4334.77515	4334.77052
Lig-Template	CATATGATAGCAACAGTCCGGACCT	$C_{243}H_{306}N_{96}O_{145}P_{24}$	7631.32246	7631.34777
H1-HCR	TAGCTTATCAGACCTATGTTGATCGGTCTGATAATCTAAAAGTC	$C_{432}H_{544}N_{159}O_{264}P_{43}$	13519.29063	13519.38774
H2-HCR	TCAACATCTGTCTGATAAGCTAGACTTTTATCTTATCAGACCGA	$C_{430}H_{544}N_{155}O_{264}P_{43}$	13433.26305	13433.26331
Initiator-HCR	TAGCTTATCAGACCGATGTTGA	$C_{216}H_{272}N_{81}O_{131}P_{21}$	6746.16076	6746.16580
H1-^{Tol}dU-HCR	TAGCTTATCAGACC(^{Tol}dU)ATGTTGATCGGTCTGATAA(^{Tol}dU)CTAAAAGTC	$C_{462}H_{556}N_{159}O_{264}P_{43}$	13891.38477	13892.57158
H2-^{Tol}dU-HCR	TCAACATC(^{Tol}dU)GTCTGATAAGCTAGACTTTTA(^{Tol}dU)CTTATCAGACCGA	$C_{460}H_{556}N_{155}O_{264}P_{43}$	13811.37254	13809.49174

Table S6.2 Thermodynamic data for DNA duplexes in phosphate buffer (100 mM NaCl, 10 mM sodium phosphate, pH 7.0).

Duplex	Oligo	Sequence	c_{total}[a] [µM]	T_m [°C]	ΔH^0 [kcal mol^{-1}]	ΔS^0 [cal mol^{-1} K^{-1}]	ΔG^{298}[b] [kcal mol^{-1}]
dA/^{Tol}dU	ODN1_A ^{Tol}dU2_A	5'...TAG...3' 3'...A^{Tol}dUC...5'	2.4 4.9 12.4 24.0 49.6	37.2 38.9 41.1 42.5 44.4	-83.3±0.5	-240±1.5	-11.9±0.7
C3/^{Tol}dU	ODN1_C3 ^{Tol}dU2_A	5'...TC3G...3' 3'...A^{Tol}dUC...5'	2.3 5.2 12.2 23.6 48.9	27.9 29.4 32.4 34.2 36.1	-65.9±1.5	-190±4.3	-9.2±1.9
dC/^{Tol}dU	ODN1_C ^{Tol}dU2_A	5'...TCG...3' 3'...A^{Tol}dUC...5'	2.4 5.0 13.0 24.6 50.5	23.8 26.4 28.6 31.7 33.8	-54.9±1.4	-157±4.1	-8.3±1.8
dG/^{Tol}dU	ODN1_G ^{Tol}dU2_A	5'...TGG...3' 3'...A^{Tol}dUC...5'	2.3 5.0 12.3 24.2 48.1	27.3 29.5 32.0 33.7 35.5	-68.5±0.5	-199±1.4	-9.1±0.6
dT/^{Tol}dU	ODN1_T ^{Tol}dU2_A	5'...TTG...3' 3'...A^{Tol}dUC...5'	2.3 4.7 11.9 23.2 47.4	26.7 28.6 31.1 32.8 34.8	-68.3±3.4	-199±1.0	-8.8±0.5

[a] Total concentration of DNA strands. Values were calculated with the absorption at 260 nm of the melted duplex [b] Calculated for T = 25 °C.

Table S6.3 Fluorescence data for DNA duplexes in phosphate buffer (100 mM NaCl, 10 mM sodium phosphate, pH 7.0).

Duplex	Oligo	Sequence	$\lambda_{ex,\,max}$ [a] [nm]	$\lambda_{em,\,max}$ [b] [nm]	Φ
dA/^{Tol}dU	ODN1_A/^{Tol}dU2_A	5'...TAG...3'/3'...A^{Tol}dUC...5'	333	405	0.037
C3/^{Tol}dU	ODN1_C3/^{Tol}dU2_A	5'...TC3G...3'/3'...A^{Tol}dUC...5'	328	419	0.010
dC/^{Tol}dU	ODN1_C/^{Tol}dU2_A	5'...TCG...3'/3'...A^{Tol}dUC...5'	331	413	0.017
dG/^{Tol}dU	ODN1_G/^{Tol}dU2_A	5'...TGG...3'/3'...A^{Tol}dUC...5'	328	423	0.021
dT/^{Tol}dU	ODN1_T/^{Tol}dU2_A	5'...TTG...3'/3'...A^{Tol}dUC...5'	334	415	0.022
ssDNA ^{Tol}dU	^{Tol}dU2_A	5'...C^{Tol}dUA...3'	331	420	0.021

[a] λ_{em}=405 nm , [b] λ_{ex}=320 nm.

Table S6.4 Lifetime data for DNA duplexes in phosphate buffer (1 µM DNA in 100 mM NaCl, 10 mM sodium phosphate, pH 7.0).

Duplex	Oligo	Sequence	a_1	τ_1 [ns]	a_2	τ_2 [ns]	a_3	τ_3 [ns]	$<\tau>$ [ns]
dA/^{Tol}dU	ODN1_A/^{Tol}dU2_A	5'...TAG...3' / 3'...A^{Tol}dUC...5'	0.96	0.34	0.04	1.27	-	-	0.37
C3/^{Tol}dU	ODN1_C3/^{Tol}dU2_A	5'...TC3G...3' / 3'...A^{Tol}dUC...5'	0.67	0.21	0.30	0.87	0.02	2.32	0.46
dC/^{Tol}dU	ODN1_C/^{Tol}dU2_A	5'...TCG...3' / 3'...A^{Tol}dUC...5'	0.75	0.21	0.23	0.81	0.02	2.02	0.39
dG/^{Tol}dU	ODN1_G/^{Tol}dU2_A	5'...TGG...3' / 3'...A^{Tol}dUC...5'	0.64	0.23	0.34	0.71	0.02	2.07	0.43
dT/^{Tol}dU	ODN1_T/^{Tol}dU2_A	5'...TTG...3' / 3'...A^{Tol}dUC...5'	0.73	0.18	0.25	0.76	0.02	2.00	0.36
ssDNA ^{Tol}dU	^{Tol}dU2_A	5'...C^{Tol}dUA...3'	0.65	0.17	0.33	0.78	0.03	1.95	0.42

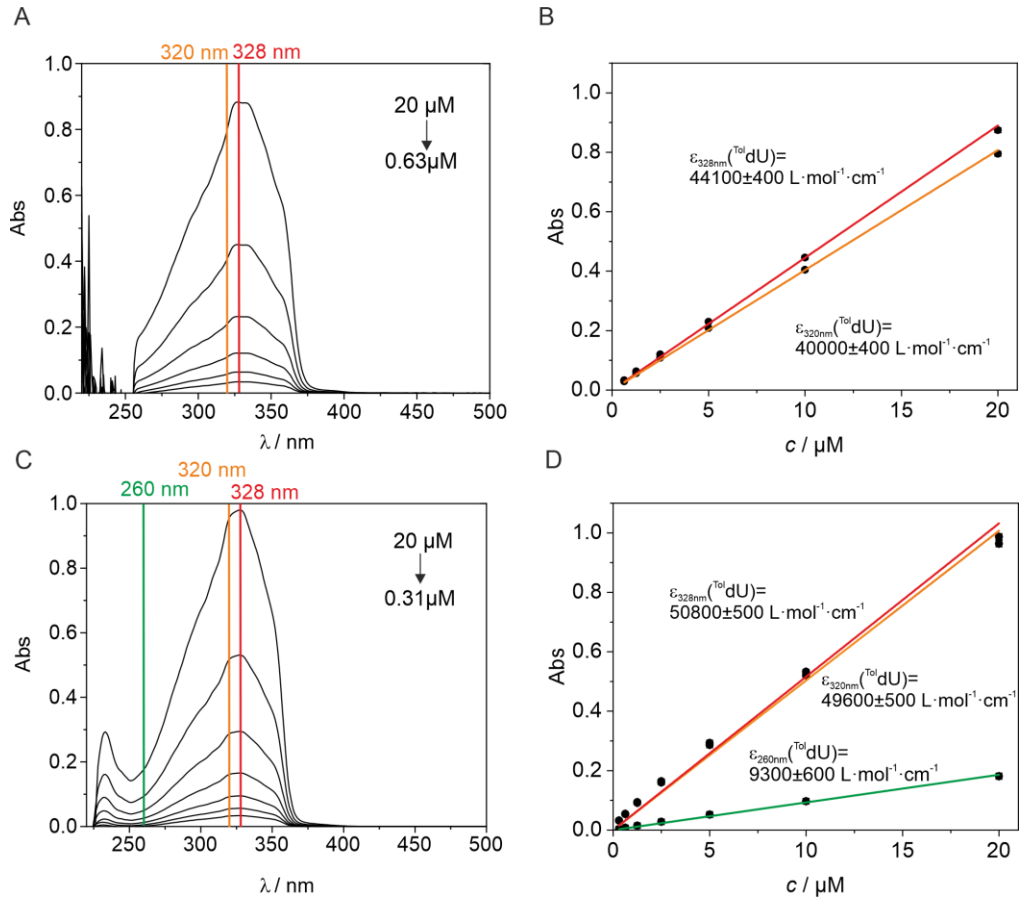


Figure S6.1 (A) UV/Vis spectrum of ^{Tol}dU in DMSO upon serial dilution starting from a 10 mM stock solution in DMSO. (B) The extinction coefficient ε was determined at 320 nm and 328 nm from a linear fit of the absorbance versus concentration. (C) UV/Vis spectrum of ^{Tol}dU in methanol starting from a 10 mM stock solution in DMSO. (D) The extinction coefficient ε was determined at 260 nm 320 nm and 328 nm from a linear fit of the absorbance versus concentration.

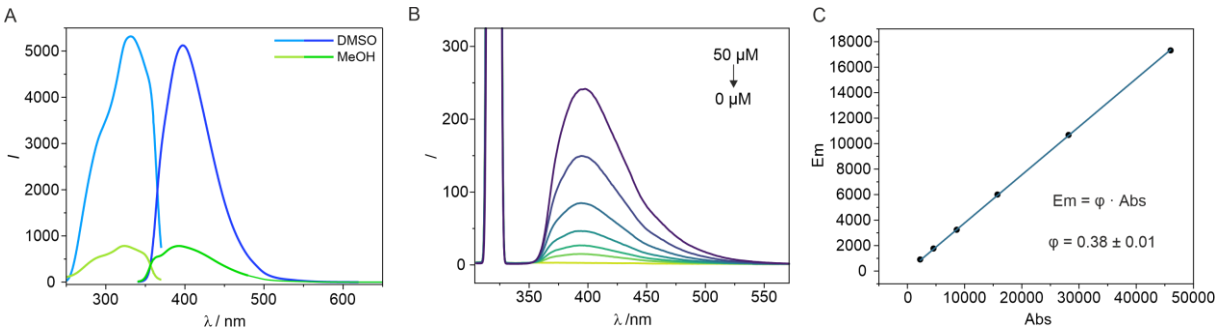


Figure S6.2 (A) Fluorescence excitation (λ_{em} = 390 nm) and emission (λ_{ex} = 320 nm) spectra of the nucleoside in DMSO and MeOH. (B) Fluorescence spectra for absolute quantum yield determination (λ_{ex}=320 nm) of the ^{Tol}dU nucleoside in DMSO at different concentrations. The scatter peak has been cropped for clarity. (C) Determination of the absolute quantum yield from a linear fit of the integrated emission vs absorbed light.

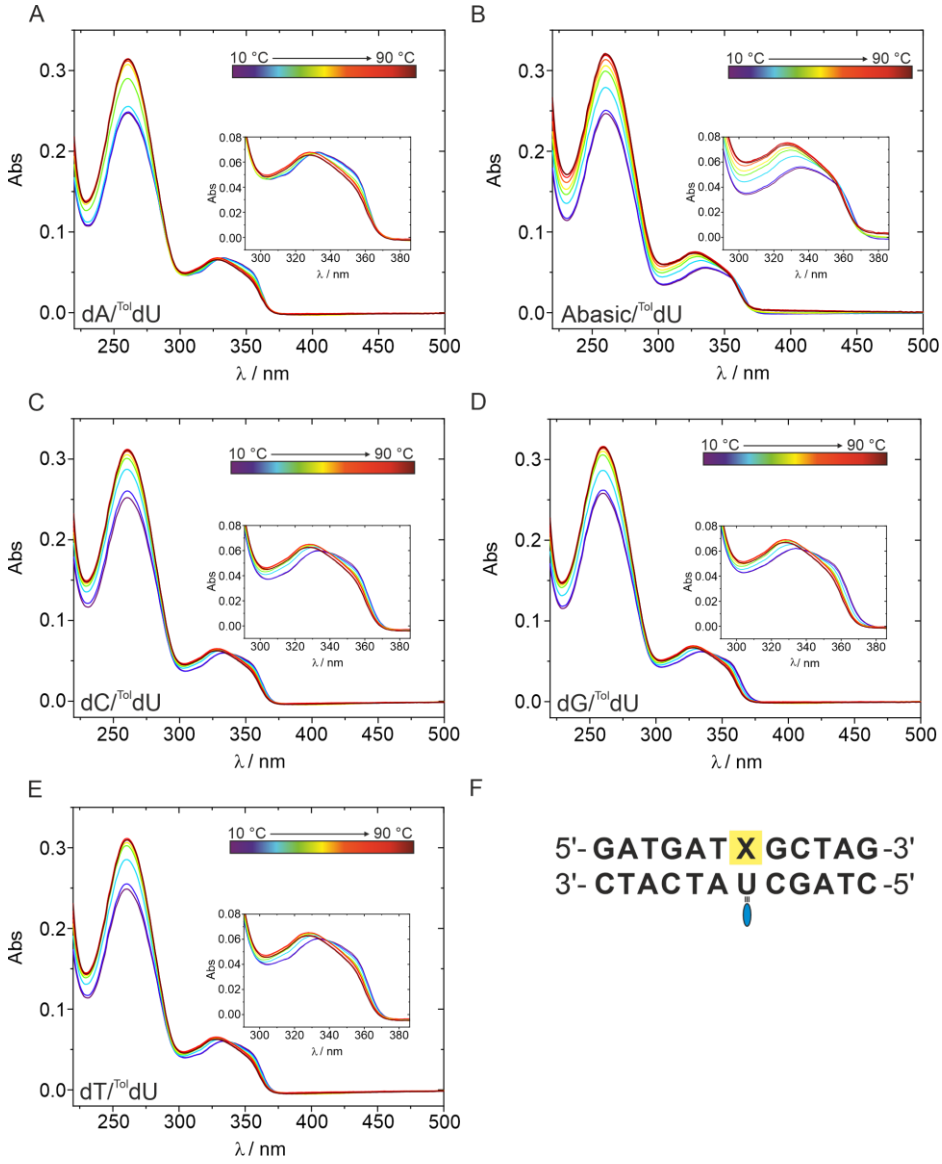


Figure S6.3 UV/Vis spectrum of ^{Tol}dU opposite of (A) deoxyadenosine, (B) abasic linker, (C) deoxycytidine, (D) deoxyguanosine and (E) thymidine in a DNA duplex at different temperatures (1 µM DNA in 100 mM NaCl, 10 mM sodium phosphate, pH 7.0). The inset shows the absorption of ^{Tol}dU. (F) Sequence of the DNA duplex. Different nucleosides have been incorporated at the yellow highlighted positon X opposite of ^{Tol}dU unit.

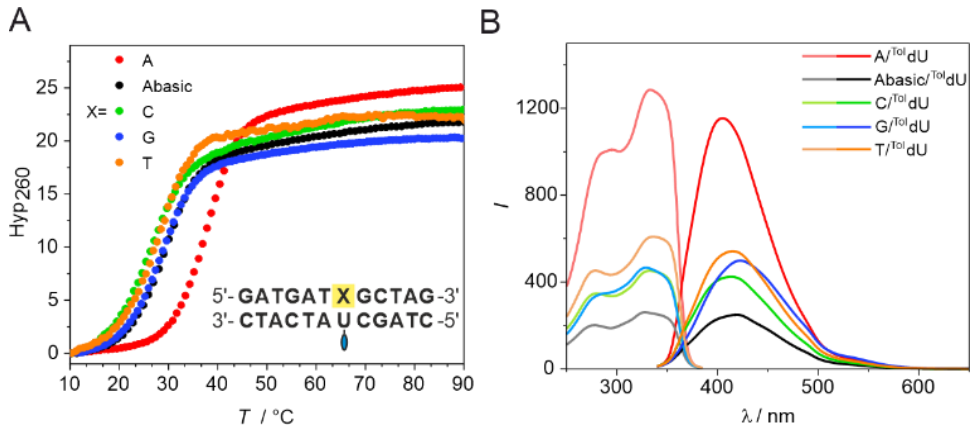


Figure S6.4 (A) UV melting curves and (B) fluorescence spectra of ^{Tol}dU DNA duplexes containing different nucleobases or an alkyl-linker opposite of ^{Tol}dU (X=A, C, G, T, or *n*-propyl), (1 µM DNA in 10 mM sodium phosphate, 100 mM NaCl, pH 7.0).

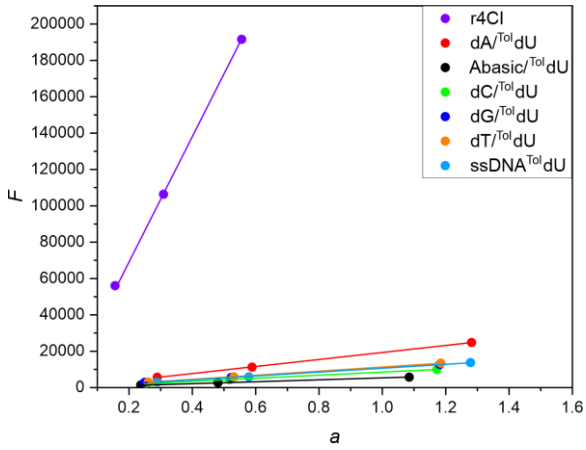


Figure S6.5 Relative quantum yield determination of ^{Tol}dU containing DNA duplexes and single strand with r4Cl as comparison (1 µM DNA in 10 mM sodium phosphate, 100 mM NaCl, pH 7.0). Results are summarized in Table S6.3.

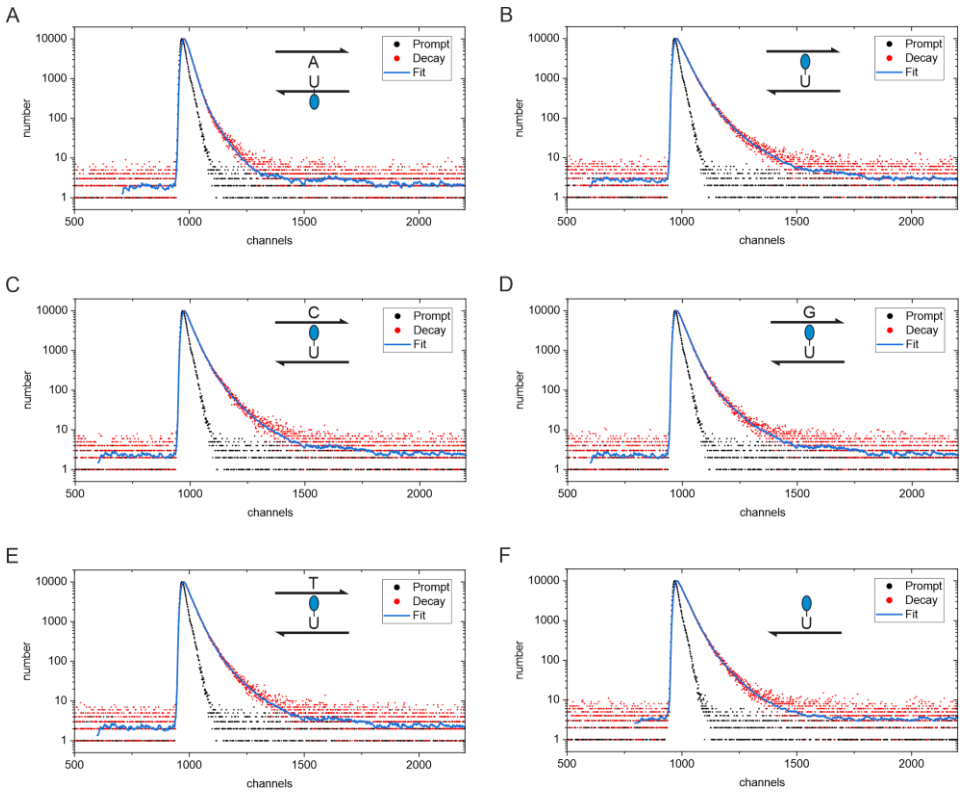


Figure S6.6 Fluorescence lifetime of ^{Tol}dU opposite of (A) deoxyadenosine, (B) abasic linker, (C) deoxycytidine, (D) deoxyguanosine, (E) thymidine in a DNA duplex and (F) in a ssDNA (1 µM DNA in 10 mM sodium phosphate, 100 mM NaCl, pH 7.0). Results are summarized in Table S6.4. Time calibration = 0.0257 ns channel^{-1}.

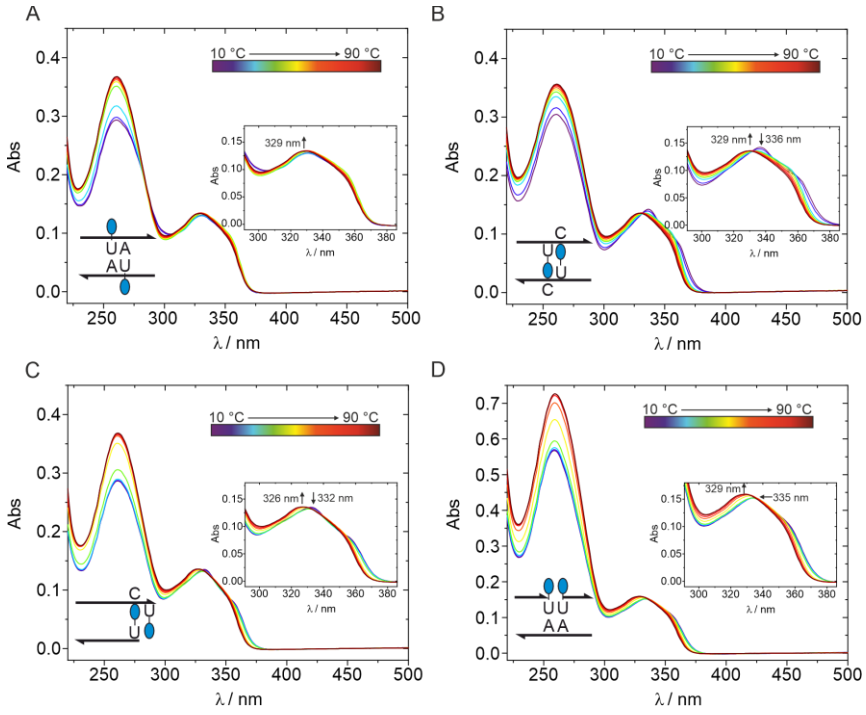


Figure S6.7 UV/Vis spectrum of two ^{Tol}dU units incorporated in DNA scaffold. ^{Tol}dU units arranged in a (A) ICL dA/^{Tol}dU, (B) ICL dC/^{Tol}dU, (C) hairpin CL and (D) ligation CL arrangement at different temperatures (1 µM DNA in 100 mM NaCl, 10 mM sodium phosphate, pH 7.0). The inset shows the absorption of ^{Tol}dU. Shifted maxima are indicated with an arrow.

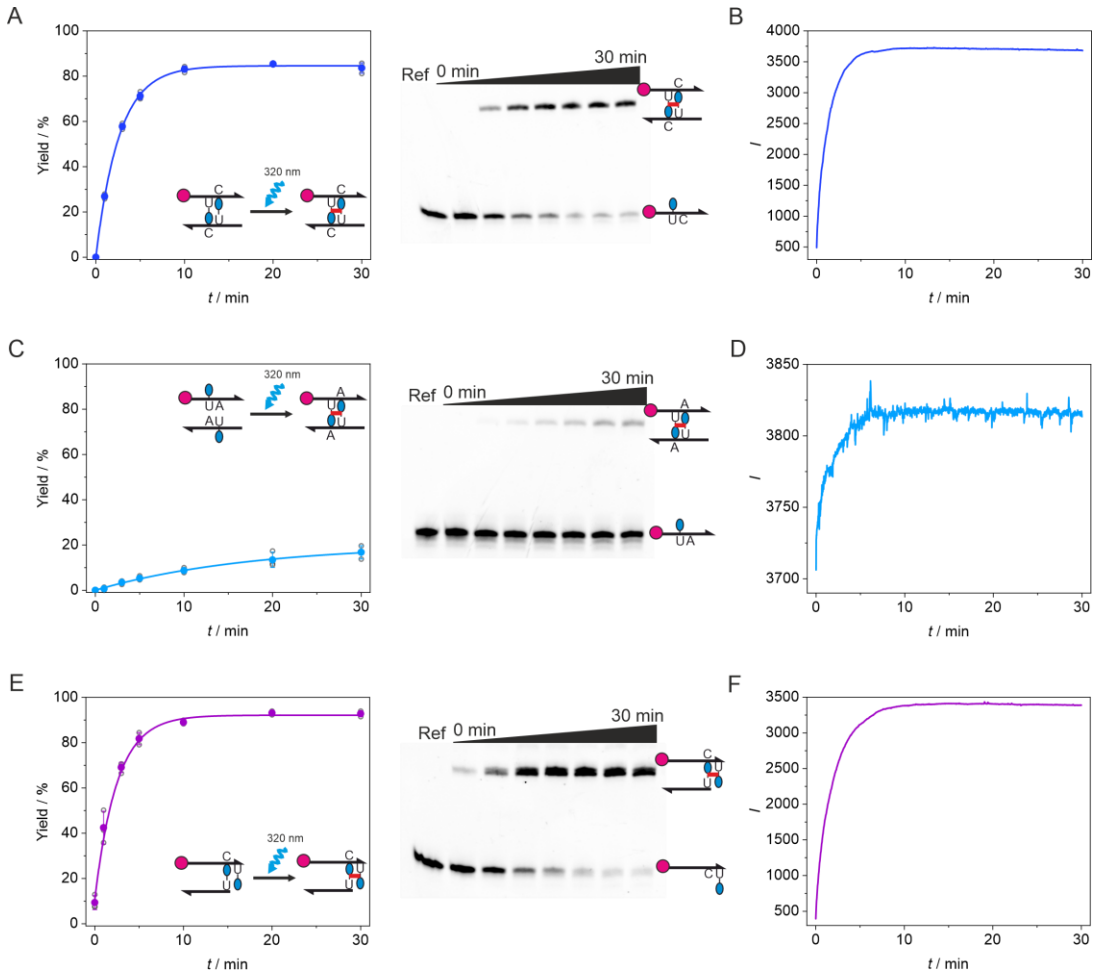


Figure S6.8 Kinetic characterization of CL formation using the fluorescence spectrometer light source with PAGE analysis and fluorescence increasement at 410 nm of the DNA mixture for (A and B) ICL dC/^{Tol}dU, (C and D) ICL dA/^{Tol}dU and (E and F) hairpin CL arrangement. Cy3 for quantification is shown as a magenta circle (1 µM DNA in 10 mM sodium phosphate, 100 mM NaCl, pH 7.0).

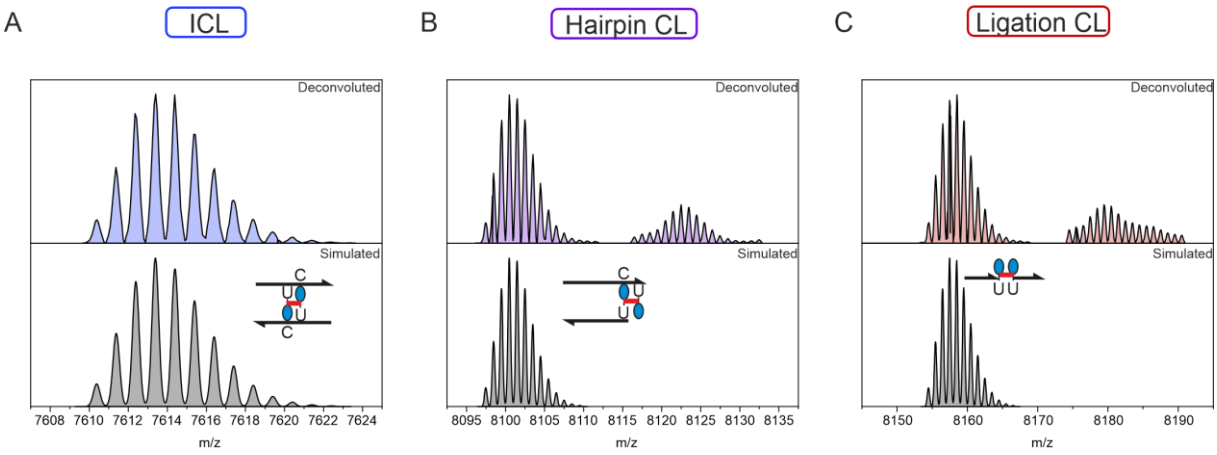


Figure S6.9 HR-ESI-MS analysis of the crosslinked DNA duplex for (A) ICL, (B) hairpin CL and (C) ligation CL. The simulated spectra correspond to the mass of the combined single strands.

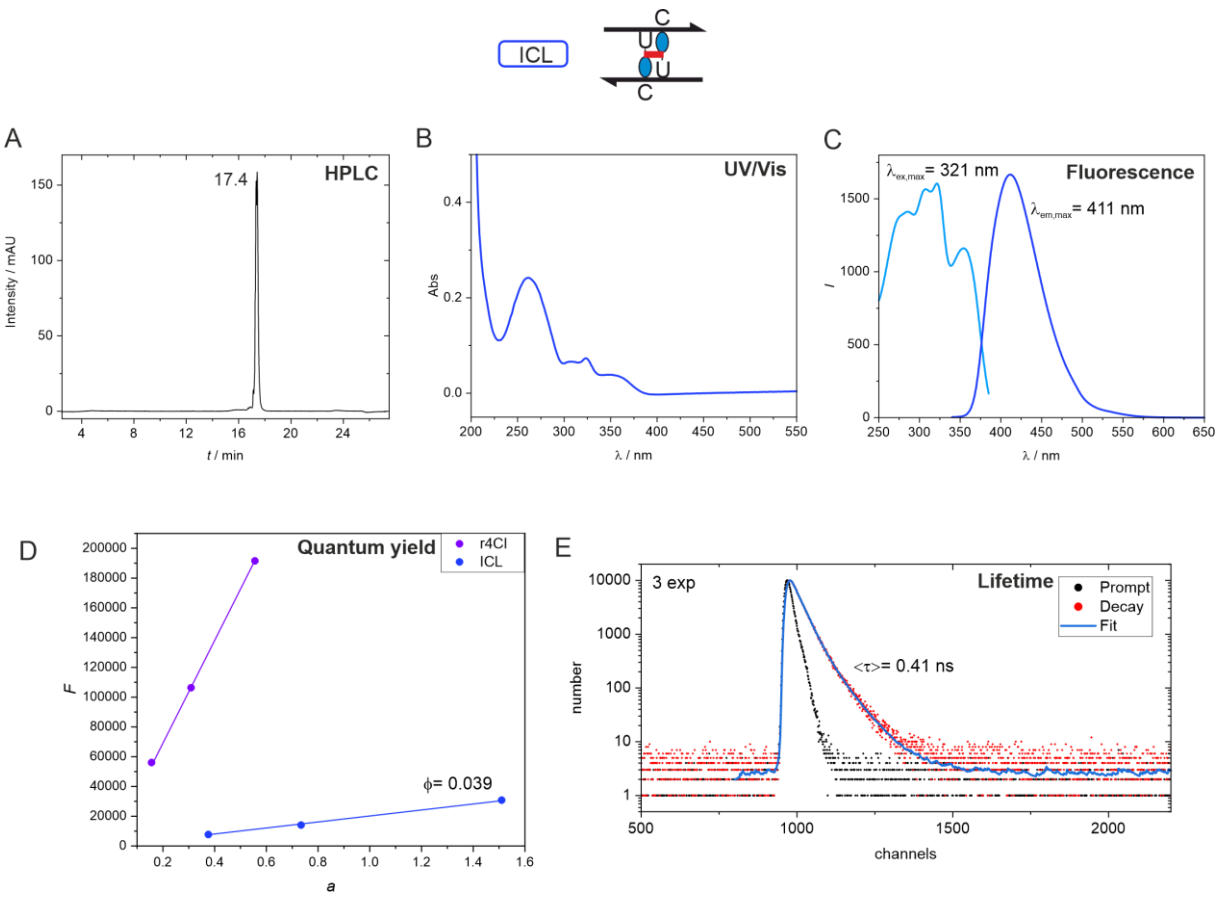


Figure S6.10 (A) Analytical anion exchange HPLC chromatogram, (B) UV/Vis spectra, (C) fluorescence excitation (λ_{em} = 405 nm) and emission (λ_{ex}=320 nm) spectra, (D) quantum yield determination and (E) life time measurement of the ICL containing duplex (10 mM sodium phosphate buffer, 100 mM NaCl, pH 7.0).

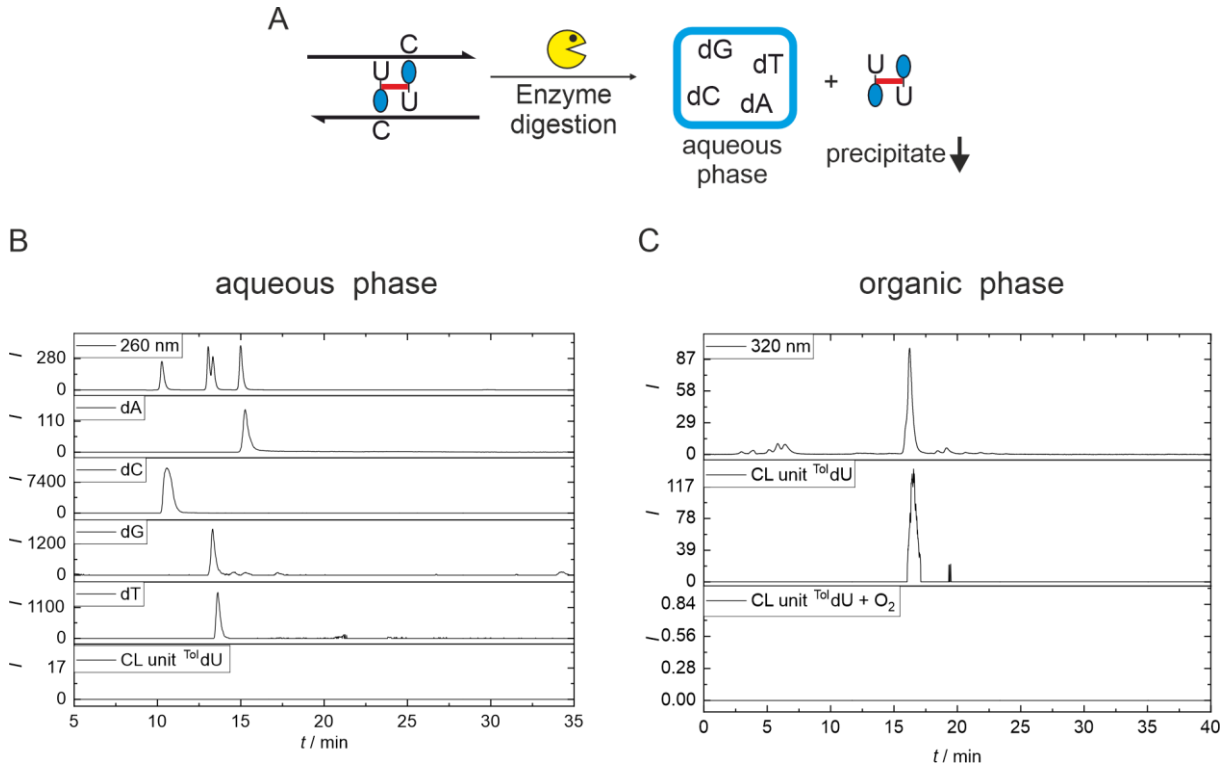


Figure S6.11 (A). Schematic drawing of the enzymatic digesting of the DNA duplex containing the ICL. The formed CL unit is water insoluble and was dissolved in MeCN/10 mM NH_4OAc (4/6, *v/v*) (B). LC-MS analysis of the aqueous phase. Extracted ion chromatograms (EIC) for CL unit ^{Tol}dU showing MH+ (m/z 857.28 ± 0.05) or MNa+ (m/z 879.26 ± 0.05) corresponding to two ^{Tol}dU units. (C) LC-MS analysis of the organic phase. EIC for CL unit ^{Tol}dU showing MH^+ (m/z 857.28 ± 0.05) or MNa^+ (m/z 879.26 ± 0.05) corresponding to two ^{Tol}dU units and EIC for CL unit ^{Tol}dU+O_2 showing MH^+ (m/z 889.271 ± 0.05) or MNa^+ (m/z 911.25 ± 0.05) corresponding to two ^{Tol}dU units with additional O_2. Incorporation of O_2 is not observed.

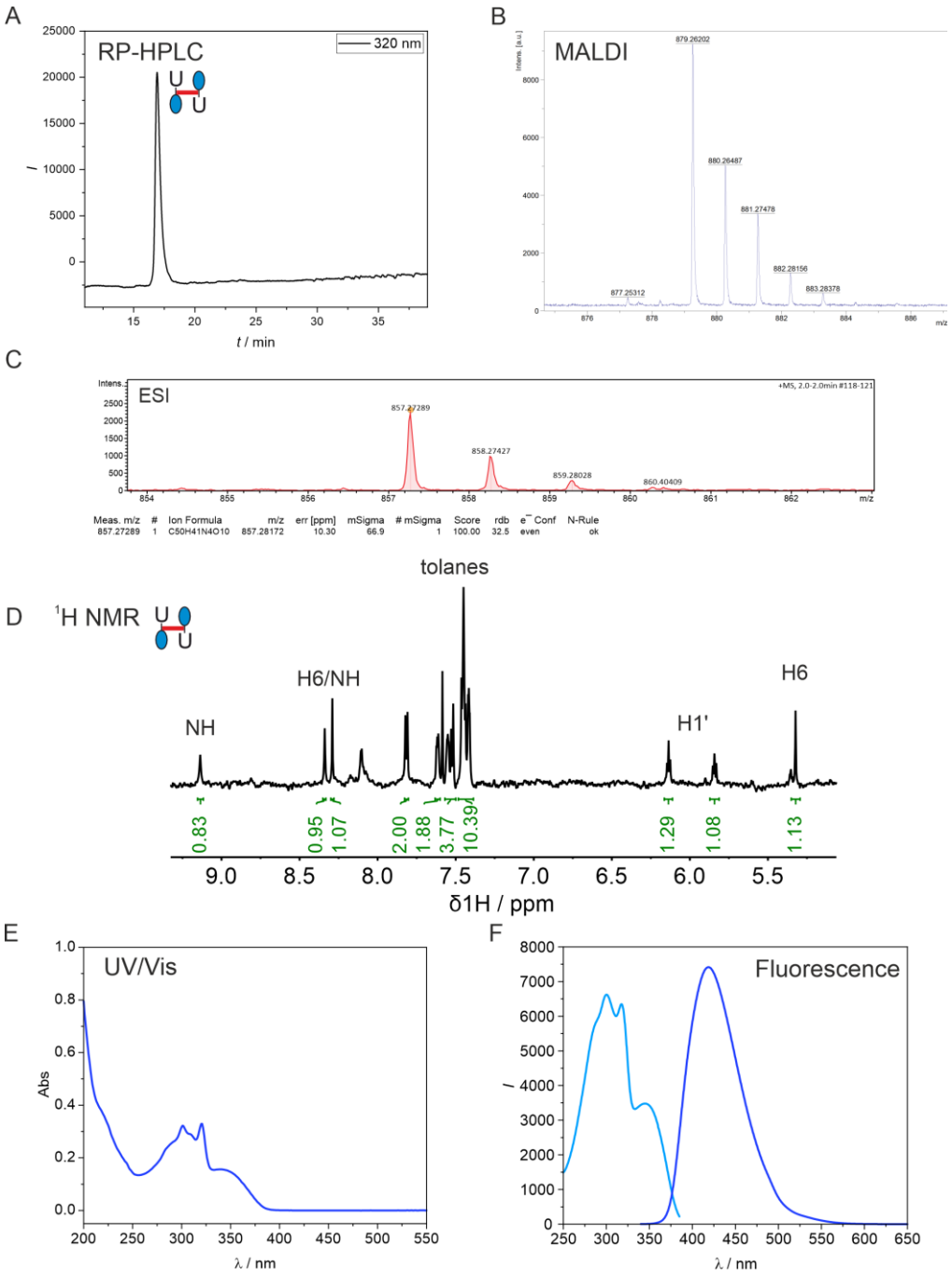
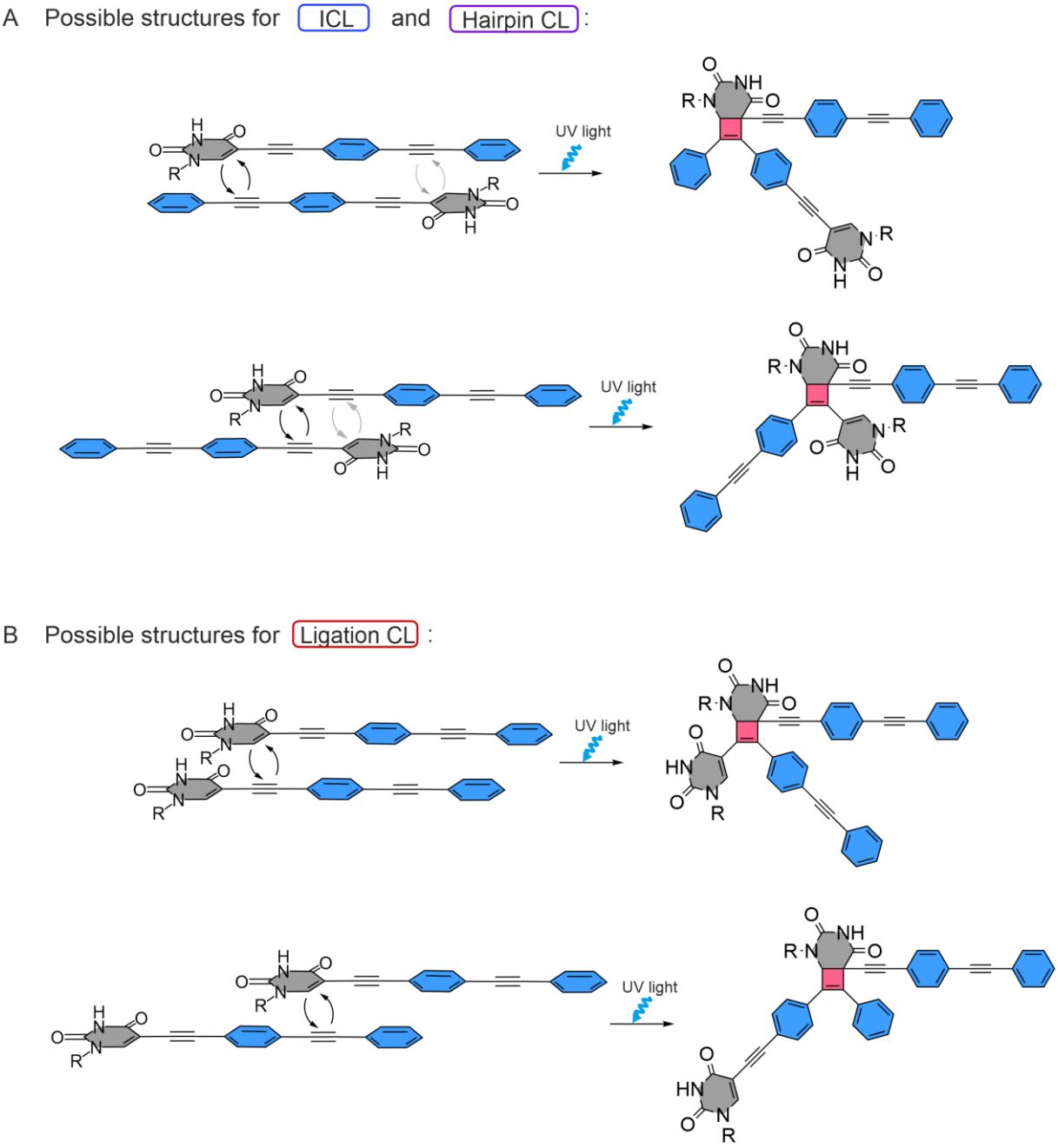


Figure S6.12 (A). Characterization of the RP-HPLC-purified crosslinked dimer unit by (A) analytical RP-HPLC using RP-C18 column (Synergi, 4-µm Fusion-RP C18 80 Å, 250x2 mm) from Phenomenex on a JASCO HPLC system, (B) MALDI MS showing Na^+-adduct (879.262 Da), (C) HR-ESI-MS (857.272 Da), (D) ^{1}H-NMR (600 MHz in MeCN-d_3) with tentative assignment, (E) UV/Vis and (F) fluorescence excitation (λ_{em} = 405 nm) and emission (λ_{ex}=320 nm) spectra in MeCN.



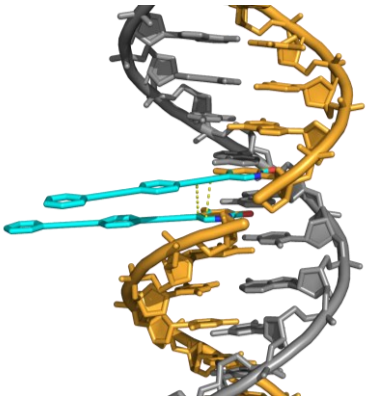


Figure S6.13 (A) Schematic drawing of possible structures for ICL and hairpin CL. The C5-C6 double bond could react with the inner or outer triple bond of the other ^{Tol}dU. In each case two isomers are possible (indicated by the light gray arrows for the second option). However, exact arrangement of the ^{Tol}dU units is not known in either case, and the spectroscopic results of the crosslinked products do not allow us to unambiguously distinguish the constitutional isomers.

(B) Schematic drawing of possible structures for nicked duplex ligation CL, involving either inner or outer triple bond. Left: A model of a nicked duplex with two ^{Tol}dU units (generated from an idealized B-form duplex) suggests a plausible parallel arrangement for the alkyne alkene cycloaddition.

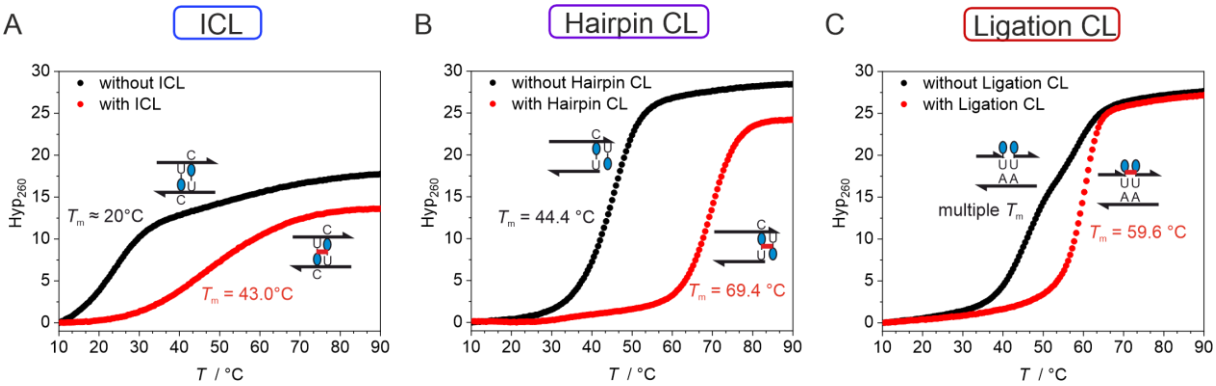


Figure S6.14 UV melting curves of the DNA duplex containing the ^{Tol}dU crosslinking motif without (black) and with (red) crosslinking for (A) ICL, (B) hairpin CL and (C) ligation CL (1 µM DNA in 10 mM sodium phosphate buffer, 100 mM NaCl, pH 7.0).

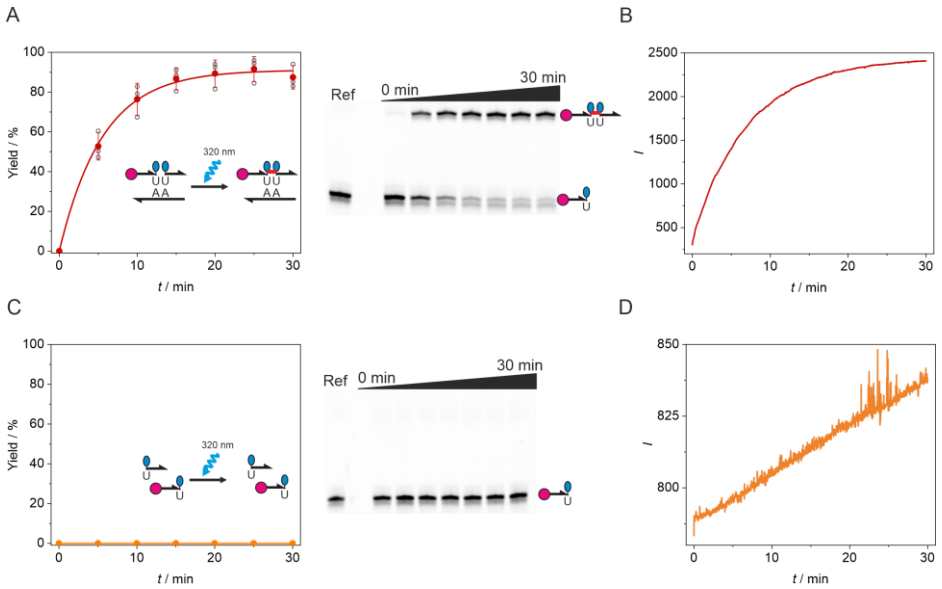


Figure S6.15 (A) Kinetic characterization of CL formation using the fluorescence spectrometer light source with PAGE analysis and fluorescence increasement at 410 nm of the DNA mixture for ligation CL (A and B) with and (C and D) without template. Cy3 for quantification is shown as a magenta circle (1 µM DNA in 10 mM sodium phosphate, 100 mM NaCl, pH 7.0).

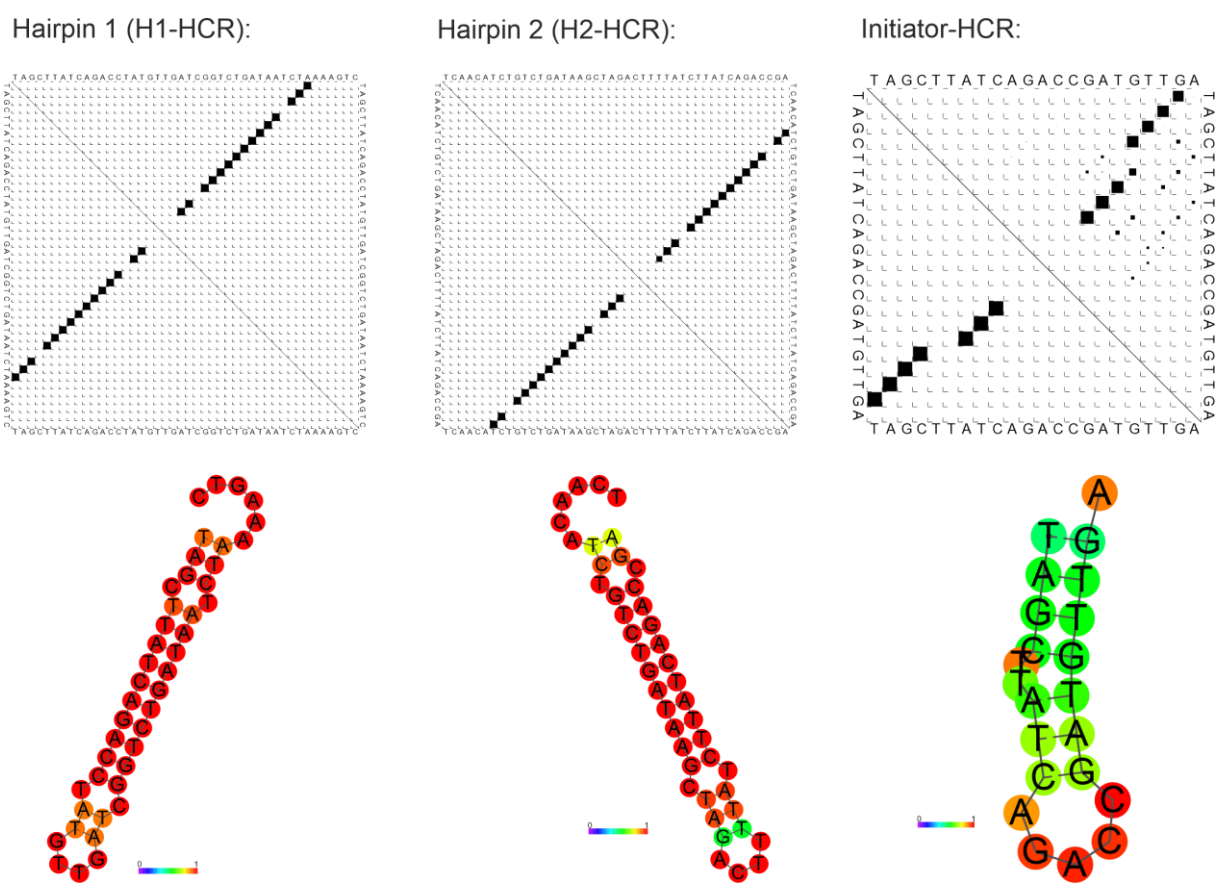
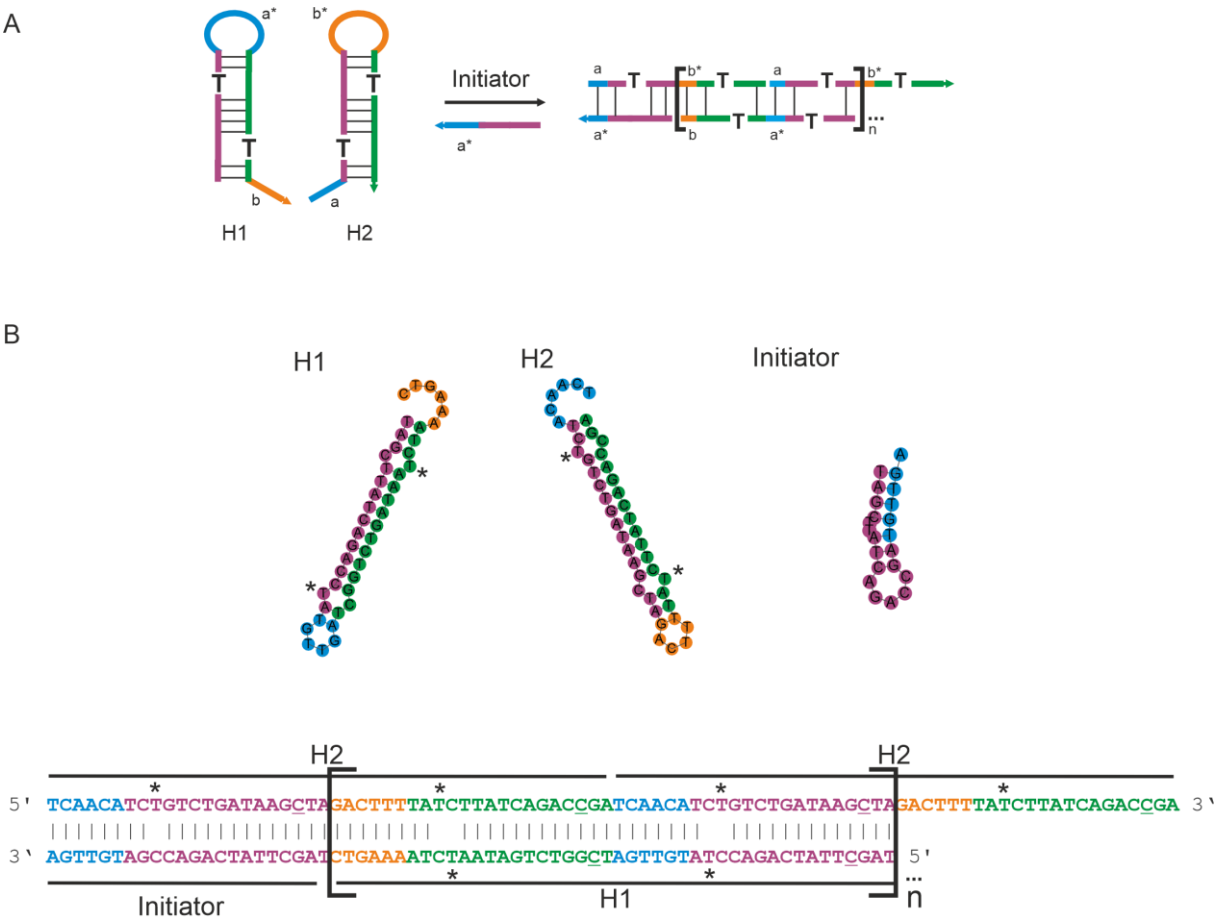


Figure S6.16 Secondary structure prediction of the DNA hairpins H1 and H2 and the initiator strand used for the HCR. Structure predictions and dot plots were generated using the ViennaRNA Web Service (http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAprobing.cgi).



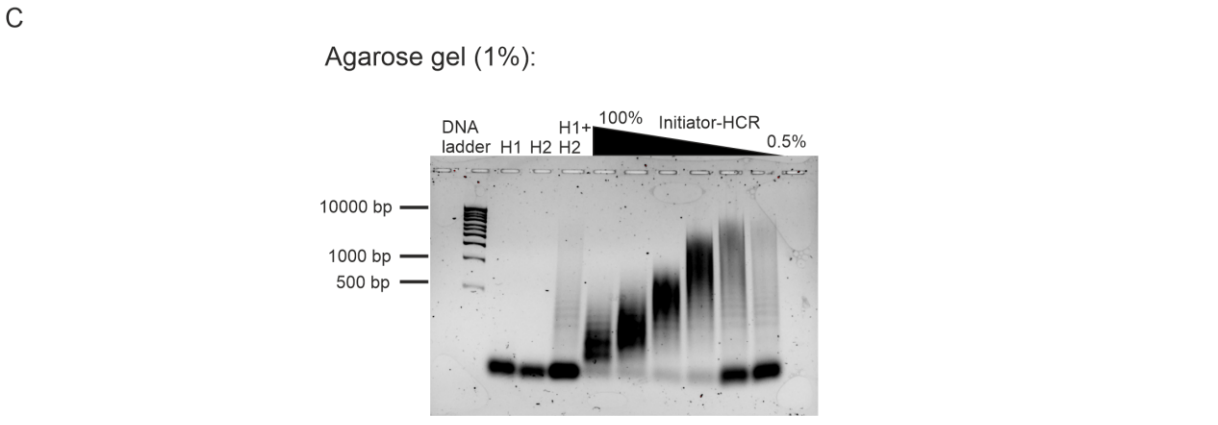


Figure S6.17 (A) Illustration of HCR with dC/dT mismatch in H1 and H2 and (B) sequence of hairpins, initiator, and HCR product. Nucleotides are labelled according to the schematic representation. Mismatched dC, which are base paired in the HCR product are underlined. Asterisks indicate mismatched Ts in the hairpins and in the HCR product. (C) The test of hairpins with two mismatches and no ^{Tol}dU modifications shows the formation of HCR products (0.5 µM DNA in 75 mM sodium citrate and 750 mM NaCl, pH 7.0).

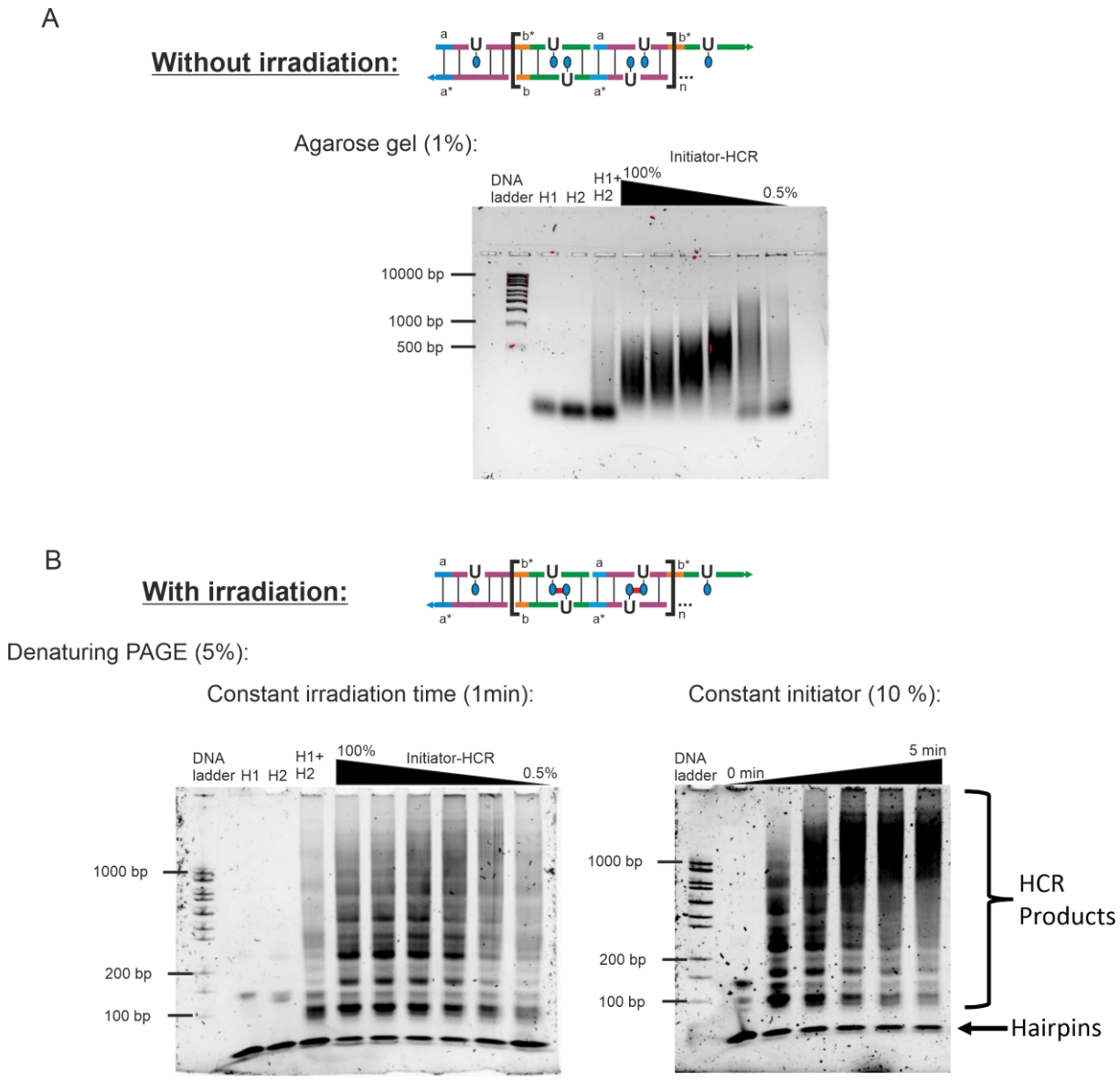


Figure S6.18 (A) HCR reaction with ^{Tol}dU containing crosslinks. (B) Analysis of HCR products after light irradiation with a transilluminator (312 nm). Denaturing PAGE showed the formation of covalently linked HCR products. (0.5 µM DNA in 75mM sodium citrate and 750 mM NaCl, pH 7.0).

6.4.2 Experimental procedures

6.4.2.1 General material and methods

All standard chemicals were purchased from commercial suppliers. Trimethylsilylacetylene was obtained from Tokyo Chemical Industry and 2'-deoxy-5-iodouridine was purchased from Biosynth. Phenylacetylene, 1-bromo-4-iodobenzene and tetrakis(triphenylphosphine)palladium(0) were obtained from Sigma Aldrich. 4,4'-Dimethoxytrityl chloride and 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite were obtained from ChemGenes Corporation.

2'-Deoxy-5'-O-(4,4'-dimethoxytrityl)-5-iodouridine (**S6.1**)[486] and 1-ethynyl-4-(2-phenylethynyl)benzene[499] were synthesized according to previously published protocols.

Dry solvents DCM, DMF and MeCN were obtained via a solvent purification system (SPS) from Inert Corporation. Pyridine for DMT-protection was obtained from Acros Organics and dried over activated molecular sieves. Technical grade solvents were used for column chromatography and distilled prior to use. All other organic solvents were used in pro analysis or for synthesis quality without further purification.

Column chromatography was carried out on silica gel (Kieselgel 60, Merck, 0.063 – 0.200 mm). Thin layer chromatography (TLC) was performed on aluminum-backed plates coated with silica gel and a fluorescent indicator (Alugram SIL G/UV254, Macherey-Nagel, UV visualization, 254 nm).

6.4.2.2 NMR-Spectroscopy and mass spectrometry

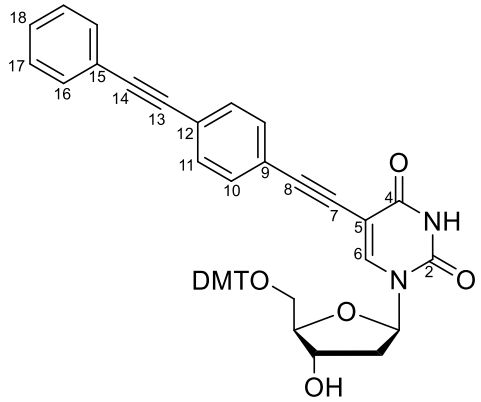
^{1}H-, ^{13}C- and ^{31}P-NMR spectra after chemical synthesis were recorded on a Bruker Avance HD III spectrometer at 400 MHz. Spectra were calibrated to the residual solvent peak of $CDCl_3$ (δ = 7.26 (^{1}H) and δ = 77.16 (^{13}C)) or DMSO-d_6 (δ = 2.50 (^{1}H) and δ = 39.52 (^{13}C)). Chemical shifts δ are given in ppm and coupling constants *J* are given in Hz. Multiplicities are denoted as follows: s (singlet), d (doublet), t (triplet), q (quartet), dd (doublet of doublet), ddd (doublet of doublet of doublet), m (multiplet).

NMR spectra of the building blocks were evaluated using MestReNova v12.0.4.

High resolution ESI mass spectra were measured on a Bruker micrOTOF-Q III spectrometer.

6.4.2.3 Synthetic procedures

Compound **S6.2**



Under nitrogen atmosphere, compound **S6.1** (649 mg, 989 µmol, 1.00 eq), CuI (18.8 mg, 96.7 µmol, 10 mol%) and $Pd(PPh_3)_4$ (114 mg, 98.7 µmol, 10.0 mol%) were dissolved in argon-purged Et_3N/DMF (1:1) (50 mL). 1-Ethynyl-4-(2-phenylethynyl)benzene (300 mg, 1.50 mmol,

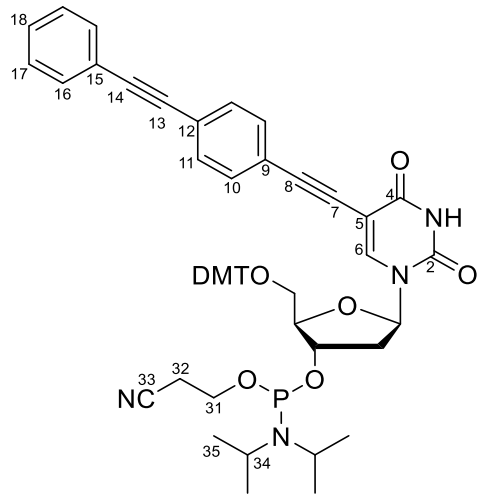
1.50 eq.) was added and stirring was continued over night at 60 °C. The reaction mixture was washed with a saturated solution of ammonium chloride (80 mL). The aqueous phase was extracted with DCM (4 x 150 mL). The combined organic phases were washed with water (80 mL) and brine (80 mL), dried over Na_2SO_4, evaporated and the residue was purified by column chromatography (DCM + 1 % Et_3N → DCM + 2 % MeOH + 1 % Et_3N) to afford compound **S6.2** as a yellow foam (437 mg, 598 µmol, 60 %).

^{1}H NMR (400 MHz, $CDCl_3$): δ (ppm) = 8.24 (s, 1H, 6-H), 7.56 – 7.47 (m, 2H, 16-H), 7.47 – 7.41 (m, 2H, DMT-H), 7.39 – 7.31 (m, 7H, 17-H, 18-H, DMT-H), 7.33 – 7.23 (m, 4H, 11-H, DMT-H), 7.18 – 7.12 (m, 1H, DMT-H), 6.98 – 6.92 (m, 2H, 10-H), 6.82 – 6.77 (m, 4H, DMT-H), 6.37 (dd, *J* = 7.6, 5.7 Hz, 1H, 1'-H), 4.64 – 4.51 (m, 1H, 3'-H), 4.13 (q, *J* = 2.9 Hz, 1H, 4'-H), 3.70 (s, 3H, DMT-H), 3.70 (s, 3H, DMT-H), 3.49 (dd, *J* = 10.8, 2.9 Hz, 1H, 5'-H), 3.32 (dd, *J* = 10.8, 3.3 Hz, 1H, 5'-H), 2.54 (ddd, *J* = 13.7, 5.9, 2.7 Hz, 1H, 2'-H), 2.35 (ddd, *J* = 13.6, 7.7, 5.9 Hz, 1H, 2'-H)

^{13}C{^{1}H} NMR (100 MHz, $CDCl_3$): δ (ppm) = 161.22 (4-C), 158.77 (DMT-C), 149.22 (2-C), 144.49 (DMT-C), 142.46 (6-C), 135.57 (DMT-C), 131.74 (16-C), 131.67 (10-C), 131.21 (11-C), 130.08 (DMT-C), 130.05 (DMT-C), 128.60 (18-C), 128.53 (17-C), 128.23 (DMT-C), 128.03 (DMT-C), 127.23 (DMT-C), 123.17 (12-C), 123.12 (15-C), 122.33 (9-C), 113.50 (DMT -C), 100.59 (5-C), 93.66 (8-C), 91.34 (14-C), 89.23 (13-C), 87.28 (DMT-C), 86.84 (4'-C), 85.98 (1'-C), 81.79 (7-C), 72.55 (3'-C), 63.56 (5'-C), 55.34 (DMT-C), 41.87 (2'-C);

HR-MS (ESI+): *m/z* calc. ($C_{46}H_{38}N_2O_7Na$, $[M+Na]^+$): 753.25712, found: 753.25754.

Compound **S6.3**



Under nitrogen atmosphere, compound **S6.2** (350 mg, 479 µmol, 1.00 eq.) was dissolved with DIPEA (491 µL, 372 mg, 2.88 mmol, 6.00 eq.) in anhydrous DCM (14 mL). After 10 min CEP-Cl (162 mg, 684 µmol, 1.30 eq.) was added. After 1.5 h stirring at room temperature additional CEP-Cl (40.0 mg, 169 µmol, 0.35 eq.) was added in two portions in a 1.5 h interval. The reaction mixture was stirred additionally at ambient temperature for 1 h. The solvent was removed under reduced pressure and the residue was purified by column chromatography (EtOAc + 1% Et_3N) to afford compound **S6.3** as a yellow foam (415 mg, 445 µmol, 93%).

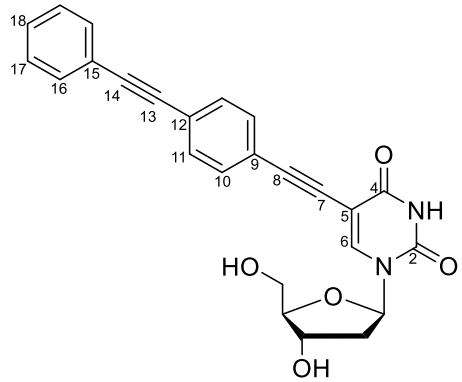
^{1}H NMR (400 MHz, $CDCl_3$): δ (ppm) = 8.41 – 8.12 (m, 2H, 6-H), 7.55 – 7.46 (m, 4H, 16-H), 7.49 – 7.41 (m, 4H, DMT-H), 7.40 – 7.29 (m, 14H, 17-H, 18-H, DMT-H), 7.32 – 7.21 (m, 6H, 11-H, DMT-H), 7.20 – 7.11 (m, 2H, DMT-H), 6.95 – 6.84 (m, 4H, 10-H), 6.84 – 6.74 (m, 8H, DMT-H), 6.39 – 6.30 (m, 2H, 1'-H), 4.67 – 4.60 (m, 2H, 3'-H), 4.29 – 4.17 (m, 2H, 4'H), 3.90 – 3.44 (m, 22H, 5'-H, 31-H, 34-H, DMT-H), 3.34 – 3.25 (m, 2H, 5'-H), 2.71 – 2.54 (m, 4H, 2'-H, 32-H), 2.50 – 2.31 (m, 4H, 2'-H, 32-H), 1.30 – 0.98 (m, 24H, 35-H);

^{13}C{^{1}H} NMR (100 MHz, $CDCl_3$): δ (ppm) = 160.98 (4-C), 160.97 (4-C), 158.77 (DMT-C), 149.06 (2-C), 149.03 (2-C), 144.51 (DMT-C), 144.45 (DMT-C), 142.49 (6-C), 135.60 (DMT-C), 135.56 (DMT-C), 135.53 (DMT-C), 131.74 (16-C), 131.67 (10-C), 131.65 (10-C), 131.16 (11-C), 131.14 (11-C), 130.17 (DMT-C), 130.11 (DMT-C), 130.11 (DMT-C), 130.07 (DMT-C), 128.59 (18-C), 128.53 (17-C), 128.21 (DMT-C), 128.12 (DMT-C), 128.06 (DMT-C), 127.21 (DMT-C), 123.13 (12-C), 123.11 (15-C), 123.09 (15-C), 122.37 (9-C), 122.34 (9-C), 117.67 (33-C), 117.50 (33-C), 113.47 (DMT-C), 100.60 (5-C), 100.57 (5-C), 93.62 (8-C), 93.59 (8-C), 91.29 (14-C), 89.26 (13-C), 87.25 (DMT-C), 86.49 (4'-C), 86.46 (4'-C), 86.14 (4'-C), 86.09 (4'-C), 86.04 (1'-C), 85.94 (1'-C), 81.75 (7-C), 81.73 (7-C), 74.03 (3'-C), 73.86 (3'-C), 73.70 (3'-C), 73.54 (3'-C), 63.36 (5'-C), 63.23 (5'-C), 58.56 (31-C), 58.40 (31-C), 58.37 (31-C), 58.21 (31-C), 55.35 (DMT-C), 55.34 (DMT-C), 43.49 (34-C), 43.42 (34-C), 43.36 (34-C), 43.29 (34-C), 41.05 (2'-C), 41.01 (2'-C), 40.96 (2'-C), 40.92 (2'-C), 24.80 (35-C), 24.76 (35-C), 24.75 (35-C), 24.69 (35-C), 24.67 (35-C), 24.62 (35-C), 20.60 (32-C), 20.52 (32-C), 20.40 (32-C), 20.33 (32-C);

^{31}P{^{1}H} NMR (162 MHz, $CDCl_3$): δ (ppm) = 148.99, 148.54;

HR-MS (ESI+): *m/z* calc. ($C_{55}H_{55}N_4O_8PNa$, $[M+Na]^+$): 953.36497, found: 953.36499.

Compound **S6.4**



Compound **S6.2** (200 mg, 274 µmol, 1.00 eq.) was dissolved in DCM (22 mL). TFA (425 µL, 633 mg, 5.55 mmol, 20.3 eq.) was added and the mixture was stirred for 15 min at room temperature. Nitrogen was bubbled through the solution to remove TFA. MeOH (2 mL) was added and removed under reduced pressure four times. The residue was purified by column chromatography (DCM + 2 % MeOH → DCM + 10 % MeOH). Additionally, the residue was washed with MeCN to afford compound **S6.4** as a yellow solid (65.4 mg, 254 µmol, 24%).

^{1}H NMR (400 MHz, DMSO-d_6): δ (ppm) = 11.73 (s, 1H, 3-H), 8.44 (s, 1H, 6-H), 7.62 – 7.53 (m, 4H, 11-H, 16-H), 7.54 – 7.47 (m, 2H, 10-H), 7.48 – 7.40 (m, 3H, 17-H, 18-H), 6.13 (t, J = 6.5 Hz, 1H, 1'-H), 5.27 (d, J = 4.3 Hz, 1H, 3'-OH), 5.20 (t, J = 4.9 Hz, 1H 5'-OH), 4.30 – 4.20 (m, 1H, 3'H), 3.82 (q, J = 3.4 Hz, 1H, 4'-H), 3.75 – 3.55 (m, 2H, 5'-H), 2.24 – 2.09 (m, 2H, 2'-H);

^{13}C{^{1}H} NMR (100 MHz, DMSO-d_6): δ (ppm) = 161.38 (4-C), 149.43 (2-C), 144.30 (6-C), 131.70 , 131.46, 131.43, 129.10 (18-C), 128.85 (17-C), 122.64 (9-C), 122.23 (12-C), 122.01 (15-C), 97.90 (5-C), 91.39 (14-C), 91.34 (13-C), 88.91 (8-C), 87.62 (4'-C), 84.93 (1'-C), 84.81 (7-C), 69.87 (3'-C), 60.78 (5'-C), 40.27 (2'-C);

HR-MS (ESI+): m/z calc. ($C_{25}H_{20}N_2O_5Na$, $[M+Na]^+$): 451.12644, found: 451.12733.

6.4.2.4 Oligonucleotide synthesis

Oligonucleotides were synthesized on an Applied Biosystems ABI 392 DNA/RNA synthesizer on a 0.6 µmol scale using standard phosphoramidite chemistry. DMT-dA(bz)-CPG, DMT-dC(ac)-CPG, DMT-dG(dmf)-CPG and DMT-dT-CPG with a pore size of 1000 Å and loading

density of 25 – 35 µmol/g were used as solid supports and were obtained from Sigma Aldrich. For 3'-^{Tol}dU functionalized oligonucleotides a ^{Tol}dU-modified solid support was prepared using native amino lcaa CPG (1000 Å) from ChemGenes Corporation. Compound **S6.2** was attached with BOP after functionalization with a succinyl ester to a lcaa CPG support. The loading density of the CPG support was determined via UV/Vis absorption of the cleaved DMT cation at 498 nm.[206] DMT-dA(bz), DMT-dC(ac), DMT-dG(dmf) and DMT-dT were purchased from ChemGenes Corporation. Hex-5-yn-1-*O*-(2-cyanoethyl-*N*,*N*-diisopropyl)phosphoramidite[418] and 3-(4,4'-dimethoxytrityloxy)propyl-1-(2-cyanoethyl-*N*,*N*-diisopropyl)phosphoramidite[419] were synthesized as described previously. Solutions of standard phosphoramidites (70 mM) and synthesized phosphoramidites (80 mM) were prepared in anhydrous MeCN. The following solutions were used in the synthesis:

Activator: 0.25 M ethylthiotetrazole (ETT) in anhydrous acetonitrile,

Oxidation: 20 mM iodine in THF/water/pyridine (66:12:22, *v*/*v*/*v*)

Cap A: pyridine/acetic anhydride/THF (10/10/80, *v*/*v*/*v*)

Cap B: NMI in THF (84/16, *v*/*v*).

Deprotection: 3% trichloro acetic acid in 1,2-dichloroethane

Cleavage from the solid support and removal of the base labile protecting groups was performed by treatment with concentrated ammonium hydroxide (33% NH_3) at 25 °C or 50 °C for 3' ^{Tol}dU-modified oligonucleotides overnight. The solid support was filtered off and the solvent was removed under reduced pressure. The residue was dissolved in water. The crude product was purified by denaturing PAGE. Gels (0.7x200x300 mm) were prepared using a 15% or 20% acrylamide solution containing 7 M urea in 1x TBE. After polymerization the gels were run in 1x TBE buffer (89 mM Tris, 89 mM boric acid, 2 mM EDTA, pH 8.3) for 2.5 h at a constant power (35 W). Product bands were visualized using a TLC plate under UV illumination. Oligonucleotides were extracted with TEN buffer (10 mM Tris-HCl, 0.1 mM EDTA, 300 mM NaCl, pH 8.0) and recovered after precipitation with ethanol. DNA strands that could not be precipitated with ethanol were desalted by size exclusion chromatography using two HiTRAP desalting columns (5 mL each) from GE Healthcare on an ÄKTAstart purification system. Oligonucleotides were eluted with water (1 mL/min). After lyophilization the oligonucleotides were dissolved in a defined volume of water.

Purified oligonucleotides were analyzed by anion exchange HPLC using a GE Healthcare ÄKTApurifier with a DNAPac PA 200 column (2x250 mm) from Thermo Scientific at a flow rate

of 0.5 mL min^{-1}. Linear gradients of 0–48 % B over 24 min of buffer A (25 mM Tris-HCl, 6 M urea, pH 8.0) and buffer B (25 mM Tris-HCl, 0.5 M $NaClO_4$, 6 M urea, pH 8.0) were used for analysis. The chromatograms were monitored at 260 nm. All analyses were performed at 60 °C. High resolution ESI mass spectra were recorded on a Bruker micrOTOF-Q III spectrometer.

6.4.2.5 Labeling of 5'-alkyne functionalized oligonucleotides

5'-Alkyne functionalized oligonucleotides were fluorescently labeled using copper(I)-catalyzed alkyne-azide cycloaddition (CuAAC). Sulfo-Cy3-azide was obtained by Lumiprobe GmbH. Freeze-dried DNA oligonucleotide (5 nmol) was dissolved in water (5 µL) and mixed with a DMSO/^{t}BuOH mixture (3:1, 3 µL). A solution of azide (0.63 µL, 50 mM) in DMSO/^{t}BuOH (3:1, *v*/*v*) was added. A freshly prepared solution of CuBr (0.63 µL, 100 mM) in DMSO/ ^{t}BuOH (3:1) was combined with a solution of tris(benzyltriazolylmethyl)amine (1.26 µL, 100 mM) in DMSO/ ^{t}BuOH (3:1) and then added to the reaction mixture. After incubation for 3 h in the dark at 37 °C, the reaction mixture was purified by PAGE (20% polyacrylamide).

6.4.2.6 UV/Vis spectroscopy / Thermal denaturing experiments

UV/Vis spectra were measured in 10 mm quartz cuvettes from Agilent using an Agilent Cary 3500 with following settings:

- Wavelength: 200-750 nm
- Averaging time: 0.02 s
- Data interval: 0.2 nm
- Scan rate: 300 nm min^{-1}
- Spectral bandwidth: 2 nm

For the determination of the extinction coefficient of the ^{Tol}dU nucleoside a stock solution in DMSO (10 mM) was prepared. Spectra measurement for extinction coefficient determination started with a 20 µM dilution in DMSO or MeOH which were serially diluted six times with DMSO or seven times with MeOH in a 1:1 ratio (20.0, 10.0, 5.00, 2.5, 1.25, 0.63, 0.31 µM) to prepare a concentration series. ^{Tol}dU crosslinking unit was measured in pure acetonitrile. Spectra were recorded at 20 °C.

Temperature dependent UV/Vis spectra were recorded every 10 °C in a range of 10-90 °C. 1 µM DNA samples in phosphate buffer (100 mM NaCl, 10 mM sodium phosphate, pH 7.0)

were heated for 1 min at 95 °C and stored at least 30 min at room temperature prior to measurement. The samples in the cuvettes were overlaid with silicone oil.

UV denaturing melting experiments were recorded on Varian Cary100 equipped with a 6x6 Multicell Block Peltier Series II cell changer and a VARIAN CARY Temperature Controller. Absorption was measured at 250, 260 and 280 nm. The absorption was recorded with a spectral bandwidth of 1 nm and the averaging time was set to 2 s. The temperature cycle was programmed as follows: 1. 20 °C to 90 °C; 2. 90 °C to 10 °C; 3.10 °C to 90 °C; 4. 90 °C to 10 °C; 5. 10 °C to 90 °C

The first ramp was performed for annealing and was not considered for further melting temperature analysis. The heating rate was set to 0.5 °C min^{-1}. 500 µL sample in phosphate buffer (100 mM NaCl, 10 mM sodium phosphate, pH 7.0) with a duplex concentration of 1 µM, 2 µM and 5 µM was measured in 10 mm quartz cuvettes from VARIAN and 300 µL sample in phosphate buffer (100 mM NaCl, 10 mM sodium phosphate, pH 7.0) with a duplex concentration of 10 µM and 20 µM was measured in 1 mm quartz cuvettes from HELLMA. The samples in the cuvettes were overlaid with silicone oil.

The obtained curves were fitted in a two-state transition model with upper and lower limit to obtain the melting temperature T_m. Absorption of the melted duplexes was used to recalculate $c_{total.}$ For the estimation of the thermodynamic parameters melting curves were analyzed according to Breslauer *et al*:[420]

$$\frac{1}{T_m} = \frac{R(n-1)}{\Delta H^0} \ln c_{total} + \frac{\Delta S^0 - (n-1) \cdot R \cdot \ln 2n}{\Delta H^0} \quad (6.1)$$

Assuming a bimolecular association of two non-self-complementary strands leads to a molecularity $n = 2$:

$$\frac{1}{T_m} = \frac{R \ln c_{total}}{\Delta H^0} + \frac{\Delta S^0 - R \cdot \ln 4}{\Delta H^0} \quad (6.2)$$

Using van't Hoff analysis enthalpy ΔH^0 could be obtained from the slope and entropy ΔS^0 from the intercept of a linear fit of ln c_{total} vs $1/T_m$. Afterwards the free energy ΔG^{298} was calculated using the Gibbs-Helmholtz equation with T = 298 K:

$$\Delta G^{298} = \Delta H^0 - T\Delta S^0 \quad (6.3)$$

6.4.2.7 Fluorescence spectra

Fluorescence spectra were recorded on a JASCO FP-8300 spectrofluorometer equipped with an FCT-817S cell changer and a F12 temperature controller device from Julabo. 1 µM samples were measured in a high precision cell from Hellma (made of quartz SUPRASIL, 10x2 mm). Excitation and emission bandwidth were set to 5 nm. Fluorescence of the samples was measured with a PMT voltage of 350 V and a response time of 1 s. Data were recorded with a scan speed of 500 nm/min with a data interval of 0.2 nm. Spectra for the ^{Tol}dU were recorded in DMSO or MeOH (dilution of a 10 mM DMSO stock solution) with an λ_{ex} = 320 nm or λ_{em} = 390 nm and annealed DNA samples in phosphate buffer (100 mM NaCl, 10 mM sodium phosphate, pH 7.0) or ^{Tol}dU crosslinking unit in pure acetonitrile (1:10 dilution from the UV/Vis sample) with an λ_{ex} = 320 nm or λ_{em} = 405 nm at 20 °C.

6.4.2.8 Absolute quantum yield

Absolute quantum yield for the ^{Tol}dU nucleosid started with a 50 µM in solution DMSO which was serially diluted in a 1:1 ratio (25.0, 12.5, 6.25, 3.13, 1.56 µM) in a Jasco 6808-H150A liquid cell (1x10x25 mm). Spectra were recorded on a JASCO FP-8300 spectrofluorometer equipped with an ILF-835 integrating sphere with the following settings:

- Ex wavelength: 320 nm
- Ex bandwidth: 5 nm
- Em bandwidth: 5 nm
- Response: 1 s
- PMT voltage: 350 V
- Data interval: 0.2 nm
- Scan speed: 100 nm/min
- Measurement range: 305-600 nm

To obtain the scattering values for the blank (s_0) and the analyte (s_x) the area of scattering peak (λ_{ex}±15 nm) was integrated. For the values of the fluorescence peak of the sample (f_x) and the blank (f_0) the rest of the spectrum (λ_{ex}+15 nm-600nm) was used for integration. The absolute quantum yield was determined according following formula:

$$\phi_{\mathrm{abs}} = \frac{f_{\mathrm{x}} - f_0}{s_0 - s_{\mathrm{x}}} \tag{6.4}$$

6.4.2.9 Relative quantum yield

Relative quantum yields were determined using a 1'-(4-cyano-1H-indol-1-yl)-β-D-ribofuranose (r4CI, Φ=0.67).[500] Samples for r4CI (6.25, 12.5 and 25 µM) were prepared in phosphate buffer (100 mM NaCl, 10 mM sodium phosphate, pH 7.0) using a 10 mM stock solution in DMSO. DNA samples (1, 2 and 5 µM) were annealed in a phosphate buffer (100 mM NaCl, 10 mM sodium phosphate, pH 7.0).

UV/Vis spectra were measured in a10 mm quartz cuvettes from Agilent using an Agilent Cary 3500 with following settings:

- Wavelength: 200-750 nm
- Averaging time: 0.032 s
- Data interval: 0.2 nm
- Scan rate: 375 nm min^{-1}
- Spectral bandwidth: 1 nm

Fluorescence spectra were recorded in a high precision cell from Hellma (made of quartz SUPRASIL, 10x2 mm) with a JASCO FP-8300 spectrofluorometer:

- Ex wavelength: 320 nm
- Ex bandwidth: 2.5 nm
- Em bandwidth: 2.5 nm
- Response: 1 s
- PMT voltage: 350 V
- Data interval: 0.2 nm
- Scan speed: 100 nm/min
- Measurement range: 340-750 nm

The measured absorbance (A) is then converted to an absorption factor (a) according following formula:

$$a = \int_{\lambda_{\mathrm{ex}} - \Delta\lambda_{\mathrm{ex}}/2}^{\lambda_{\mathrm{ex}} + \Delta\lambda_{\mathrm{ex}}/2} 1 - 10^{-A}\,\mathrm{d}\lambda \tag{6.5}$$

The emission spectra were integrated between 340-600 nm to get f_x. Comparison with the standard r4Cl led to the relative quantum yield:

$$\phi_{\mathrm{rel}} = \phi_{\mathrm{r4Cl}} \cdot \frac{f_{\mathrm{x}}}{a_{\mathrm{x}}} \cdot \frac{a_{\mathrm{r4Cl}}}{f_{\mathrm{r4Cl}}} \tag{6.6}$$

6.4.2.10 Fluorescence lifetime

Fluorescence lifetimes were determined using a time-correlated single photon counting setup at λ_{em} = 405 nm for a 1 µM DNA sample in phosphate buffer (100 mM NaCl, 10 mM sodium phosphate, pH 7.0) on a Horiba DeltaFlex spectrometer with a DeltaDiode excitation source DD-320 (λ_{ex} = 318 nm). Samples were measured after annealing in a in a high precision cell from Hellma (made of quartz SUPRASIL, 10x2 mm). The emission bandwidth has been set for a stop rate close to 2%. The photon arrival time was set to 4096 channels, resulting in a width of 0.026 ns/channel. Data collection continued until 10000 counts had accumulated in the peak channel. The instrument response function was recorded from a buffer sample by setting the emission monochromator to the actual peak wavelength of the excitation source. Fluorescence decay curves were analyzed by iterative reconvolution fitting with the instrument response function using a proprietary algorithm implemented in Horiba DAS6 (version 6.8).

6.4.2.11 DNA-Crosslinking procedures

For analytic analysis, a 1 µM DNA duplex solution was prepared in phosphate buffer (100 mM NaCl, 10 mM sodium phosphate, pH 7.0). The duplex was annealed in the dark by heating to 95 °C for 1 min and cooling down on ice for 60 min.

Irradiation with a Xe-lamp (150 W) was performed on a JASCO FP-8300 spectrofluorometer equipped with an FCT-817S cell changer and a F12 temperature controller device from Julabo. All measurements were performed in an FMM-200 5 mm quartz fluorescence microcell from JASCO with a magnetic stir bar using an FMH-802 5 mm microcell jacket from JASCO as cell adapter. The temperature was set at 10 °C or 15 °C if the duplex contains a Cy3 label. After 5 min equilibration in the spectrometer the reaction mixture (300 µL) was irradiated with the following settings:

- Ex wavelength: 320 nm
- Em wavelength: 410 nm
- Ex bandwidth: 5 nm

- Em bandwidth: 5 nm
- Response: 1 s
- PMT voltage 350 V
- Data interval: 1 s

Aliquots (5 µL) were taken at different time points (0, 1, 3, 5, 10, 20 and 30 min for ICL and hairpin CL; 0, 5, 10, 15, 20, 25 and 30 min for ligation CL) and quenched with loading buffer (5 µL).

Fluorescence increase at 410 nm and anion exchange HPLC were measured for oligonucleotides without Cy3 label.

The kinetics were analyzed by denaturing PAGE (0.4x200x200 mm, 20% acrylamide/ bisacrylamide 19:1, 7 M urea) with 1×TBE and run at 25 W power for 1 hour. The gels were imaged on a Chemidoc device, and the crosslink yield was determined from the band intensities of the Cy3 fluorescence quantified by ImageLab software. Kinetic data were obtained by fitting the resulting curve to the following equation:

$$Y = Y_{\max} \cdot (1 - e^{-k_{\text{obs}} \cdot t}) + y0 \tag{5.7}$$

For a preparative crosslinking reaction, a concentration of 20 µM was used in 300 µL phosphate buffer (100 mM NaCl, 10 mM sodium phosphate, pH 7.0). After the annealing step the sample was irradiated with a Xe-lamp (150 W) at 350 nm in the fluorescence spectrometer for 30 min in the case of ICL and hairpin CL and 60 min for ligation CL. The reaction mixture was lyophilized and purified by PAGE (20%).

Enzymatic digestion and LC-MS analysis

For LC-MS characterization PAGE purified and lyophilized crosslinked DNA duplex (500 pmol) was dissolved in antarctic phosphatase buffer (10 µL, 50 mM Bis-Tris-Propane-HCl,1 mM $MgCl_2$, 0.1 mM $ZnCl_2$, pH 6, New England Biolabs) and Nuclease P1 (0.25 µL, 25 U, New England Biolabs), Phosphodiesterase I (0.5 µL, 0.05 U, Sigma Aldrich) and antarctic phosphatase (0.2 µL, 1 U, New England Biolabs) were added and the mixture was kept at 37 °C overnight. The digested product was diluted with water (40 µL), extracted with chloroform (2x25 µL) and analyzed by LC-MS using a RP-C18 column (Synergi, 4-µm Fusion-RP C18 80 Å, 250x2 mm) from

Phenomenex at 25 °C with aqueous mobile phase A (10mM NH_4OAc, pH 5.3) and organic mobile phase B (100% acetonitrile). The flow rate was 0.2 mL/min with a gradient of 0–80% B in 40 min.The micrOTOF-Q III with an ESI ion source was operated in positive-ion mode, with a capillary voltage of 4.5 kV, an end plate offset of 500 V, a nitrogen nebulizer pressure of 1.4 bar, a dry gas flow of 9 L/min and dry temperature of 200 °C. Data were analyzed using Data Analysis software DA 4.2 (Bruker Daltonics).

The chloroform phases were combined in the reaction vessel of the digestion and the solvent was carefully removed under reduced pressure. Water (100 µL) was added, and the suspension was heated to 65 °C for 5 min to denature remaining proteins. After cooling down to room temperature MeCN (100 µL) was added, and the mixture was centrifuged for 3 min. The supernatant was transferred to another reaction vessel. The precipitate was discarded. After lyophilization of the supernatant the residue was resuspended in water (100 µL) and centrifuged for 30 min at 4 °C. The supernatant was discarded, and the pellet was dissolved in MeCN/10mM NH_4OAc buffer (2/3, *v*/*v*). The organic phase was analyzed by LC-MS using a gradient of 40–90 % B in 40 min.

For NMR analysis of the digested crosslink unit the enzymatic digestion was upscaled. PAGE purified and lyophilized crosslinked DNA duplex (11 nmol) was dissolved in Antarctic phosphatase buffer (210 µL, 71.4 mM bis-tris-propane-HCl,1.4 mM $MgCl_2$, 0.14 mM $ZnCl_2$, pH 6, New England Biolabs) and Nuclease P1 (6 µL, 600 U, New England Biolabs), phosphodiesterase I (12 µL, 1.2 U, Sigma Aldrich) and Antarctic phosphatase (5 µL, 25 U, New England Biolabs) were added and the mixture was kept at 37 °C overnight. The digested product was diluted with water (800 µL) and extracted with chloroform (2x200 µL). The chloroform phases were combined in the reaction vessel of the digestion and the solvent was carefully removed under reduced pressure. Water (200 µL) was added, and the suspension was heated to 65 °C for 5 min to denature remaining proteins. After cooling down to room temperature MeCN (200 µL) was added, and the mixture was centrifuged for 3 min. The supernatant was transferred to another reaction vessel. The precipitate was discarded. After lyophilization of the supernatant the residue was dissolved in MeCN/H_2O (500 µL, 2/3, *v*/*v*).

Purification was performed on a HPLC system (ÄKTAmicro) from GE Healthcare using a NUCLEOSIL reversed-phase column (C18, 250x4.6 mm, 100 Å, 5 µm) from Machery Nagel at a

flow rate of 1 mL min^{-1}. Linear gradients of 40–80 % B over 35 min of buffer A (water) and buffer B (MeCN) were used for purification. Chromatograms were monitored at 260 nm and 320 nm. The purified crosslink units were lyophilized and dissolved in MeCN-d_3 and analyzed on a Bruker Avance III 600 NMR spectrometer equipped with a DCH ^{13}C / ^{1}H cryoprobe. NMR spectra were evaluated using MestReNova v12.0.4 or Topspin 3.2 (Bruker BioSpin).

6.4.2.12 Hybridization chain reaction (HCR)

For hybridization chain reaction hairpin (3 µM) and initiator (1 µM) stock solutions in 5xSSC buffer (750 mM NaCl, 75 mM sodium citrate, pH 7.0) were prepared. The DNA was annealed in the dark by heating to 95 °C for 2 min and cooling down on ice for 30 min. To start the HCR both hairpins (500 nM) and the initiator (500, 250, 125, 50, 10, 2.5 nM) were combined in a reaction volume of 12 µL in 5xSCC buffer. The reaction incubated 2 h in the dark and were afterwards loaded on a 1% agarose gel containing 0.01% SybrGold for staining (90 V for 45 min or 60 V for 70 min).

Initiator dependent crosslink experiments were irradiated 1 min on a transilluminator device from HEROLAB GmbH with an irradiation maximum of 312 nm. For time dependent crosslink experiments, a constant initiator concentration of 50 nM was used, and the samples were irradiated 0.5, 1, 2, 3 and 5 min. The crosslinked product was analyzed via denaturing PAGE (5%, 200 V for 30 min).

For fluorescence measurements 280 µL of a solution containing H1- ^{Tol}dU-HCR and H2- ^{Tol}dU-HCR (final concentration 100 nM) in 5xSSC buffer was prepared. Fluorescence measurements were performed like described in “DNA-Crosslinking procedures” for 15 min at 20 °C. After 1 min irradiation 20 µL of initiator solution in 5xSCC were used to start the HCR. Fluorescence increase was measured for final initiator concentration of 50, 20, 10, 5 and 2 nM in 300 µL reaction volume. For the negative control the oligonucleotide Lig-Template with a concentration of 50 nM in 300 µL reaction volume was used. Results were normalized using following equation:

$$\boldsymbol{F_{\mathrm{norm}} = \frac{F}{F_0}} \tag{6.8}$$

With: F_{norm}= normalized fluorescence, F = measured fluorescence, F_0 = starting fluorescence

6.4.3 Appendix

The characterization of the synthesized products (NMR, HPLC, mass spectra), melting curves, UV/Vis and fluorescence spectra for quantum yield determination can be found in the published supporting information.[501]

7 Conclusion and outlook

In this thesis, the stacking of artificial nucleobases was used to extend the properties of DNA with hydrophobic aromatic moieties. For this purpose, nucleobase surrogates have been synthesized that remove or modify natural nucleobases with hydrophobic aromatic moieties and extend the function of a DNA double helix through specific recognition and reactivity. Phenylacetylene units were used as a common motif in the different approaches. With the help of phosphoramidite chemistry, several DNA strands with a hydrophobic modification have been obtained. The interactions of the incorporated building blocks in the DNA context were studied in detail with thermodynamic and NMR spectroscopic analyses.

The first goal of the thesis was to implement an artificial base pair by polar π-interaction. Known systems for unnatural base pairs exploit different approaches like alternative hydrogen bonding patterns or hydrophobic shape complementarity, but also controlled stacking was achieved using electron-rich and -poor aromatic rings. In our case specific recognition of hydrophobic moieties has been achieved through the use of arene-fluoroarene interactions. The tolane derivate THH and its fluorinated analog TFF were introduced into the DNA double helix via acyclic backbone surrogates. The positive quadrupole moment of TFF by fluorine substituents allowed specific recognition with the aromatic hydrocarbon of THH through polar π-interactions with a stabilization energy of 1.4 kcal mol^{-1}. The preferred heteropair formation was further confirmed by FRET-tracked strand exchange experiments. For the establishment of the interaction based on a stacked conformation, the choice of the acyclic backbone proved to be crucial. Thermodynamic stabilization of the DNA duplex through arene-fluoroarene interactions was only observed with the BuNA backbone but could not be achieved with the GNA backbone. Detailed NMR studies of the modified duplexes revealed a greater distortion of the DNA backbone in the case of GNA incorporated building blocks, resulting in a less stable arrangement.

Encouraged by the successful implementation of polar π-interactions and preferential stacking by the tolane/perfluorotolane heteropair, the partially fluorinated tolane analogues TFH and THF were synthesized and incorporated into DNA to obtain self-complementary units.

These building blocks complete a supramolecular language based on the fluorotolane moiety. Again, the choice of the acyclic backbone was critical for a highly stabilizing artificial base pair. The combination of the tolane THF and the BuNA backbone resulted in the BTHF moiety, which raised the melting point of a DNA duplex by more than 5 °C compared to a natural A/T base pair. NMR studies of a BTHF dimer inside a DNA duplex confirmed the stacked arrangement in a head-to-tail fashion of the tolane units. The high level of stabilization suggested that the sequential incorporation of multiple BTHF units may be used as a bioorthogonal a recognition motif in water. Indeed, when three BTHF units were incorporated into a non-self-complementary DNA strand, the formation of a dimer was observed. When BTHF was incorporated at the end of a DNA duplex, there was no visible aggregation under the conditions studied, suggesting that initial Watson-Crick base pairing of the natural nucleobases is necessary to bring the BTHF units in close proximity, thereby enhancing the local concentration to observe BTHF-induced aggregation.

The ability to tune the stacking of the aromatic rings for precise recognition by arene-fluoroarene interactions is demonstrated by this newly implemented system. However, small changes, such as an additional methylene group in the backbone, can be critical to achieve optimal stacking within the DNA, which supports careful selection of the setup of chromophore arrangements in the DNA. Further analysis by structural characterization could lead to a better understanding of the interstrand stacking arrangements within the DNA duplex, which could be helpful in the design of chromophore assemblies. Importantly, the introduction of arene-fluoroarene interactions brings programmability to hydrophobic systems, which can be used to control the assembly of DNA-organic amphiphiles using for example oligo(*p*-phenylene-ethynylene). The aggregation in these systems is mostly driven by the hydrophobic effect. By introducing arene-fluoroarene interactions, a more precise assembly could be achieved, resulting in greater structural control.

The second goal was to use aromatic stacking interactions for the controlled placement of functional groups that enable their reactivity. For this purpose, 2'-deoxyuridine was combined with a hydrophobic phenylacetylene moiety and incorporated into DNA. The artificial nucleoside ^{Phe}dU still showed Watson-Crick base pairing properties with a deoxyadenosine. When the deoxyadenosine was replaced by any of the other non-complementary nucleosides, a

mismatch situation was generated, in which the preferred conformation of the ^{Phe}dU units was changed. A conformational change by rotation around the glycosidic bond to the syn orientation resulted in a stacked conformation of the phenylacetylene units. After UV irradiation in the mismatched situation, efficient interstrand crosslinking was observed, covalently linking both DNA strands in the duplex. Base pairing of ^{Phe}dU or a metal-mediated base pair resulted in inhibition of the reaction, demonstrating that the stacked arrangement is necessary for crosslinking. Detailed analysis by LC-MS, NMR and atomic mutagenesis revealed an unexpected alkene-alkyne [2+2] cycloaddition. The crosslinked DNA substrate was used in a collaborative project that aims at structural analyses of protein complexes involved in DNA damage recognition and nucleotide excision repair. As expected, the covalent interstrand crosslink prevented strand separation by the XPD helicase activity. The stalled helicase complex enabled more detailed structural analyses by cryogenic electron microscopy in the Kisker group, which revealed novel insights into the function of the so far underexplored arch domain of XPD (unpublished results not discussed in this thesis).

The addition of a phenylacetylene unit to the hydrophobic moiety of ^{Phe}dU resulted in the fluorescent nucleoside ^{Tol}dU. Apart from detecting single nucleotide polymorphisms, the modification is capable of crosslinking two DNA strands in three different architectures. Importantly, an increase in fluorescence was observed during the crosslinking reaction, allowing easy tracking of successful crosslinking. The analytical data are consistent with the alkene-alkyne [2+2] cycloaddition observed in the case of ^{Phe}dU. The ^{Tol}dU nucleoside has been successfully used in a DNA nanotechnology approach such as the hybridization chain reaction (HCR). The HCR allowed the formation of the crosslinking motif after hybridization, resulting in covalently linked DNA duplex polymers. As a proof-of-principle application of the fluorogenic crosslinking reaction, we demonstrated successful sequence detection by fluorescence enhancement in the HCR approach.

Further applications in the growing field of DNA nanotechnology are possible, e.g. stabilization of the derived nanostructures by covalent bonds. The possibility of modification of the aromatic ring for successful crosslinking reactions opens the door for further optimization of the system. While no unspecific DNA damage was observed, the use of UV light for inducing the covalent crosslinking reaction is a potential concern, that can possibly be alleviated by shifting the absorption to visible light through a further functionalized aromatic ring.

The examples presented demonstrate the potential of DNA modified with hydrophobic aromatic moieties. However, the functionality is mostly based on a DNA duplex as a scaffold. Future directions may focus on tolane building blocks not incorporated into DNA strands. The coupling of multiple tolane units would lead to amphiphilic tolane oligomers. The supramolecular language developed in this work, which is based on arene-fluoroarene interactions of tolanes, makes it possible to control the recognition of different tolane oligomers and the structure of their assemblies (see Figure 7.1A). A long-term goal would be the controlled assembly of tolane-oligomer structures in water by means of arene-fluoroarene interactions. For this purpose, it's important to have a better understanding of the synthesis and purification of tolane oligomers. Due to their amphiphilic nature, even the incorporation of three terminal tolane units proved to be challenging.

The possibilities of the newly developed supramolecular language can be further extended by bilingual monomers. These monomers are a combination of the Watson-Crick and arene-fluoroarene base pairing and would allow translation between the two languages. Figure 7.1B schematically shows how bilingual molecules form a supramolecular assembly that mediates the interaction of a DNA strand and a tolane oligomer, with the bilingual molecule acting as a translator between two strands. The first example of such a translator molecule would be ^{Tol}dU, which contains a canonical nucleobase and the hydrophobic tolane motif. As shown in this work, an optimal arrangement is necessary to allow a stacked arrangement of tolanes. Therefore, further fine-tuning of the backbone units will be necessary in order to enable triplex formation via hydrophobic recognition in the major groove of the DNA duplex.

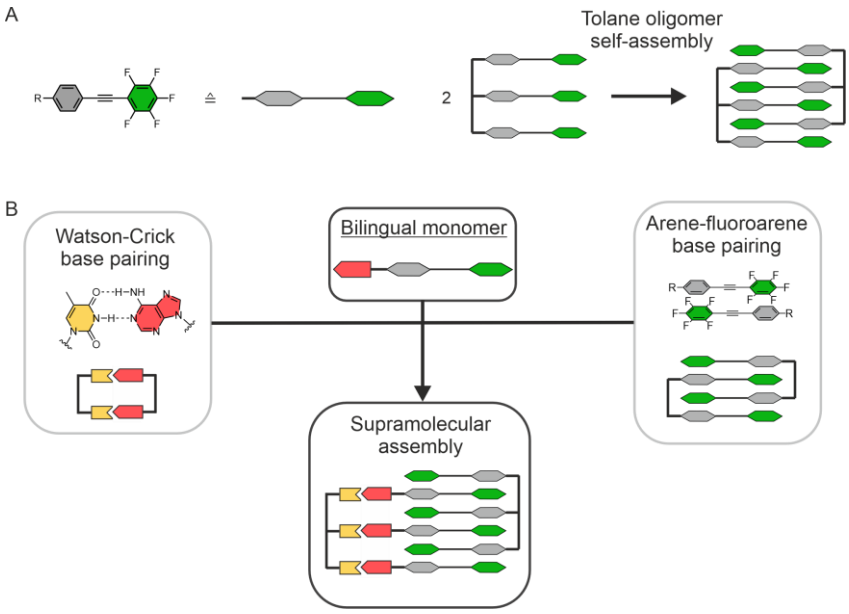


Figure 7.1 (A) Schematic illustration of the self-assembly of a tolane oligomer by arene-fluoroarene interactions. The strands consist only of tolane building blocks that form a dimer. (B) Combination of Watson-Crick and Arene-Fluoroarene base pairing, resulting in a bilingual monomer that allows the formation of supramolecular assemblies. The bilingual monomer acts as a translator between the two different languages.

Numerous future applications and further developments can be envisioned base on the results presented in this thesis. Three different areas may briefly be outlined here. One direction to explore is the non-covalent assembly of ^{Phe}dU or ^{Tol}dU nucleosides on a DNA template instead of directly incorporating them via phosphoramidite chemistry into a DNA strand. Upon irradiation, the supramolecular oligomer would be converted into covalent ^{Phe}dU or ^{Tol}dU oligomers, that lack the negatively charged phosphodiester backbone. Studying hybridization and self-assembly of such novel oligomers would be of particualr interest due to their fluorescent properties. (see Figure 7.2).

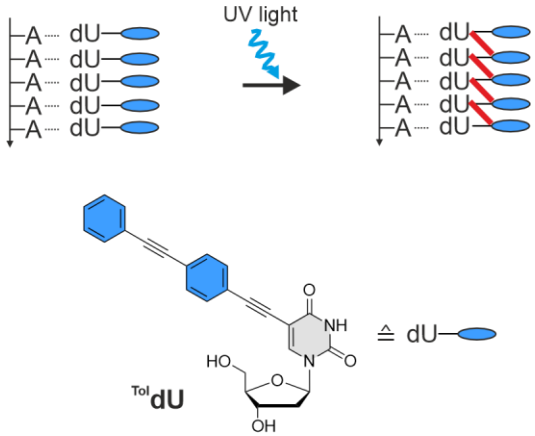


Figure 7.2 Illustration of the template-assisted reaction of ^{Tol}dU nucleosides to generate ^{Tol}dU oligomers.

As an interstrand crosslink, the alkene-alkyne [2+2] cycloaddition reaction product can be described as a DNA damage, making it attractive to study in the context of DNA repair. The ease of incorporation of a ^{Phe}dU crosslink allows incorporation into a variety of DNA substrates, leading to better characterization of enzymes involved in the DNA repair. Other *in vitro* applications in nucleic acid chemistry are also possible. One example is the *in vitro* selection of ribozymes. Many selection strategies use a loop to connect the substrate strands to the potential ribozyme sequences. The hairpin crosslink can be used to introduce a covalent bond between the substrate and the ribozyme strand after the incubation step, allowing selection of ribozymes that do not require the connecting loop to be active.

In conclusion, this work provides several examples of how DNA can be functionalized, including a novel recognition motif and controlled reactivity, and demonstrates that understanding of stacking interactions provides useful insights into nucleic acid chemistry.

8 Zusammenfassung und Ausblick

Ziel dieser Arbeit war es, durch die Nutzung von Stapelwechselwirkungen künstlicher Nucleobasen die Eigenschaften der DNA zu erweitern. Zu diesem Zweck wurden hydrophobe aromatische Einheiten verwendet, die entweder die Nucleobasen zusätzlich verlängern oder sie vollständig ersetzen, was zu neuen Funktionen in der DNA-Doppelhelix führt, wie z.B. die spezifische Erkennung oder die Einführung neuer Reaktivitäten. Dabei wurden Phenylacetylen-Einheiten als gemeinsames Motiv für die verschiedenen Ansätze verwendet. Mit Hilfe der Phosphoramidit-Chemie war es möglich, eine Vielzahl von verschiedenen DNA-Strängen mit hydrophober Modifikation zu erhalten. Die Wechselwirkungen der eingebrachten Moleküle innerhalb der DNA wurden mit thermodynamischen und NMR-spektroskopischen Analysen untersucht.

Das erste Ziel der Arbeit war die Einführung eines artifiziellen Basenpaars durch polare π-Wechselwirkungen. Bekannte Systeme für künstliche Basenpaare nutzen verschiedene Ansätze wie alternative Muster von Wasserstoffbrückenbindungen oder hydrophobe komplementäre Strukturen von Molekülen, aber auch elektronenreiche und -arme Aromaten können als Ansatz für künstliche Basenpaare genutzt werden. In dieser Arbeit wurde die spezifische Erkennung von hydrophoben Einheiten durch die Wechselwirkung von Arenen mit Fluorarenen ausgenutzt. Die 2'-Desoxyribose im Phosphodiesterrückgrat wurde durch azyklische Alkyleinheiten ersetzt um das Tolanderivat THH und dessen fluoriertes Analog TFF in die DNA-Doppelhelix einzuführen. Das durch die Fluorierung veränderte Quadrupolmoment von TFF erlaubte eine spezifische Erkennung mit dem aromatischen Kohlenwasserstoff von THH durch polare π-Wechselwirkungen mit einer Stabilisierungsenergie von 1.4 kcal mol^{-1}. Die Präferenz für die Ausbildung eines Heteropaares wurde durch FRET unterstützte Austauschexperimente weiter bestätigt. Für die Etablierung von Wechselwirkungen auf Basis von Stapelwechselwirkungen ist die Wahl des azyklischen Rückgrats entscheidend. Nur bei Verwendung von BuNA wurde eine Stabilisierung durch Wechselwirkung von Arenen mit Fluorarenen beobachtet, die bei GNA nicht auftrat. NMR-Untersuchungen zeigten eine stärkere Verformung des DNA-Gerüsts bei Verwendung von GNA, was zu einer weniger stabilisierenden Anordnung führte.

Aufbauend auf der erfolgreichen Einführung von polaren π-Wechselwirkungen durch ein Tolan-Heteropaar wurden die teilfluorierten Tolan-Analoga TFH und THF synthetisiert und DNA-Stränge eingebaut, um selbstkomplementäre Einheiten zu erhalten. Diese Bausteine vervollständigen eine supramolekulare Sprache, die auf dem Tolan-Motiv basiert. Auch hier zeigte sich, dass die Wahl des azyklischen Rückgrats entscheidend war, um eine hohe Stabilisierung durch das künstliche Basenpaar zu erreichen. Aus der Kombination vom Tolan THF und dem BuNA-Rückgrat resultierte die BTHF Einheit, dessen Dimer die Stabilität einer DNA-Duplex um 5 °C im Vergleich zu einem natürlichen A/T Basenpaar steigerte. NMR-Untersuchungen des in die DNA eingebauten BTHF-Dimers bestätigen eine gestapelte Anordnung mit antiparalleler Orientierung der Tolan-Einheiten. Der hohe Grad der Stabilisierung ließ vermuten, dass sich durch mehrmaligen sukzessiven Einbau der BTHF-Einheit eine eigenständige Erkennungseinheit im Wasser bilden könnte. Tatsächlich konnte die Bildung eines Dimers aus nicht-selbstkomplementären Strängen beobachtet werden, wenn ein BTHF-Trimer eingebaut war. Wurde dieses Trimer jedoch an das Ende eines DNA-Duplexes gesetzt, konnte keine Aggregation beobachtet werden. Dies deutet darauf hin, dass die natürliche Basenpaarung zusätzlich notwendig ist, um eine BTHF-induzierte Aggregation zu erhalten.

Mit dem hier untersuchten System konnten wir zeigen, dass die Stapelung von Aromaten durch die Wechselwirkung von Arenen mit Fluorarenen kontrolliert werden kann. Dennoch können scheinbar kleine Änderungen, wie eine zusätzliche Methylengruppe im Rückgrat, entscheidend für eine optimale Anordnung für Stapelwechselwirkungen sein. Dies unterstreicht die Notwendigkeit einer sorgfältigen Auswahl des Systems für die Anordnung von Chromophoren in der DNA. Weitere Analysen durch strukturelle Charakterisierung der gestapelten Anordnung von zwei gegenüberliegenden Tolan-Chromophoren in der DNA können weitere Erkenntnisse für das Design solcher Chromophorstrukturen liefern. Die Einführung der Wechselwirkung von Arenen mit Fluorarenen führt zur Programmierbarkeit in hydrophoben aromatischen Systemen, die zur Kontrolle von Aggregaten in DNA-organischen amphiphilen Systemen, z.B. mit Oligo(*p*-phenylethinyl), genutzt werden kann. Der hydrophobe Effekt ist die Hauptursache für die Aggregation. Durch die Einführung von Wechselwirkungen zwischen Arenen und Fluorarenen können präzisere Anordnungen erreicht werden, die zu einer größeren strukturellen Kontrolle führen.

Das zweite Ziel der Arbeit bestand darin, aromatische Stapelwechselwirkungen für die kontrollierte Platzierung reaktiver Gruppen zu nutzen, um deren Reaktivität zu ermöglichen. Dazu wurde 2'-Desoxyuridin mit einem hydrophoben Phenylacetylenrest kombiniert und in die DNA eingebaut. Das erhaltene künstliche Nukleosid ^{Phe}dU zeigte weiterhin Watson-Crick-Basenpaarungseigenschaften mit einem Desoxyadenosin. Es wurde ein Design gewählt, bei dem zwei gegenüberliegende ^{Phe}dU-Einheiten bei Fehlpaarung eine Stapelkonfiguration einnehmen. Nach UV-Bestrahlung in der Fehlpaarungssituation konnte eine effiziente kovalente Strangquervernetzung beobachtet werden. Eine Basenpaarung von ^{Phe}dU oder eine metallvermittelte Basenpaarung führte zu einer Inhibierung der Reaktion, was zeigt, dass die gestapelte Anordnung für die Vernetzung notwendig ist. Eine unerwartete Alken-Alkin [2+2] Cycloaddition wurde durch eine detaillierte Analyse mittels LC-MS, NMR und Atommutagenese entdeckt. Die Verwendung eines quervernetzten DNA-Substrats führte zu einer vollständigen Hemmung der XPD-Helikase-Aktivität und ermöglichte einen präzisen Stopp der Helikase auf dem Substrat.

Durch Anhängen einer weiteren Phenylacetylen-Einheit an den hydrophoben Teil von ^{Phe}dU entstand das fluoreszierende Nukleosid ^{Tol}dU. Neben dem Nachweis von Einzelnukleotid-Polymorphismen ist diese Modifikation in der Lage, zwei DNA-Stränge in drei verschiedenen Architekturen zu vernetzen. Bemerkenswert ist, dass während der Vernetzungsreaktion ein Anstieg der Fluoreszenz zu beobachten ist, sodass die erfolgreiche Vernetzung leicht verfolgt werden konnte. Die analytischen Daten stimmen mit der bei ^{Phe}dU beobachteten [2+2] Cycloaddition zwischen einem Alken und Alkin überein. Das Nukleosid ^{Tol}dU wurde erfolgreich in einem DNA-Nanotechnologie-Ansatz verwendet. Die Hybridisierungskettenreaktion (HCR) ermöglichte die Bildung des Vernetzungsmotivs nach der Hybridisierung. Dies führte zu kovalent verknüpften DNA-Duplex-Polymeren. Der beobachtete Anstieg der Fluoreszenz nach Zugabe des Initiatorstrangs in der HCR-Methode ist ein Beweis für die prinzipielle Anwendbarkeit des Systems zum Nachweis von DNA-Sequenzen.

Die einfache Synthese verschiedener 5-Ethinyl-2’-desoxyuridin-Derivate und die hohe Skalierbarkeit ermöglichten eine detaillierte Analyse der Vernetzungsreaktion. Außerdem ermöglichte die Menge des isolierten vernetzten Produkts die Verwendung in biochemischen Studien. Parallel zu dieser Arbeit waren Versuche zur strukturellen Charakterisierung von Enzymen, die an der menschlichen DNA-Reparatur beteiligt sind, mit Hilfe des ^{Phe}dU-vernetzten

DNA-Substrats erfolgreich und lieferten weitere Einblicke in den DNA-Reparaturmechanismus unter Verwendung von XPD. Dies zeigt, dass die entdeckte Alken-Alkin [2+2] Cycloaddition für die Analyse von DNA-verarbeitenden Enzymen verwendet werden kann. Auch auf dem wachsenden Gebiet der DNA-Nanotechnologie sind Anwendungen denkbar, zum Beispiel die Stabilisierung der abgeleiteten Nanostrukturen durch kovalente Bindungen. Die Möglichkeit, den aromatischen Ring für erfolgreiche Vernetzungsreaktionen zu modifizieren, öffnet die Tür für eine weitere Optimierung des Systems. Das verwendete UV-Licht kann unerwünschte DNA-Schäden verursachen, die durch eine Veränderung der Absorption für sichtbares Licht durch einen weiteren funktionalisierten aromatischen Ring verhindert werden könnten. Es wurden zwar keine unspezifischen DNA-Schäden beobachtet, aber die Verwendung von UV-Licht stellt ein potenzielles Problem dar, welches durch die Veränderung der Absorption für sichtbares Licht durch einen weiteren funktionalisierten aromatischen Ring verhindert werden könnte.

Die hier vorgestellten Beispiele zeigen das Potential von DNA, die mit hydrophoben aromatischen Einheiten modifiziert worden ist. Die Eigenschaften der DNA wurden erweitert, aber beruhen auf der DNA-Duplex als Gerüst. Tolan-Bausteine, die nicht in DNA-Strängen eingebaut sind könnten in Zukunft im Mittelpunkt der Forschung stehen. Die Verknüpfung mehrerer Tolan-Einheiten würde zu amphiphilen Tolan-Oligomeren führen. Die in dieser Arbeit entwickelte supramolekulare Sprache, die auf Aren-Fluoraren-Wechselwirkungen von Tolanen beruht, ermöglicht es, die Erkennung verschiedener Tolan-Oligomere und vermutlich auch die Struktur ihrer Assemblierungen zu kontrollieren (siehe Abbildung 8.1A). Ein langfristiges Ziel wäre der kontrollierte Aufbau von Tolan-Oligomer-Strukturen in Wasser mit Hilfe von Aren-Fluoraren-Wechselwirkungen. Zu diesem Zweck ist es wichtig, die Synthese und Reinigung von Tolan-Oligomeren weiter zu optimieren. Aufgrund ihres amphiphilen Charakters war selbst der Einbau von drei endständigen Tolaneinheiten eine Herausforderung.

Durch den Einsatz von bilingualen Monomeren können die Möglichkeiten der neu entwickelten supramolekularen Sprache noch erweitert werden. Diese Monomere ermöglichen eine Kombination aus Watson-Crick-Basenpaarung und Aren-Fluoraren-Basenpaarung für eine Übersetzung zwischen den beiden Sprachen. Abbildung 8.1B zeigt schematisch, wie bilinguale Moleküle eine supramolekulare Struktur bilden, die die Wechselwirkung zwischen einem DNA-Strang und einem Tolan-Oligomer vermittelt, wobei das bilinguale Molekül als

Übersetzer zwischen zwei Strängen fungiert. Ein erstes Beispiel für ein solches Translatormolekül wäre ^{Tol}dU. Es enthält eine kanonische Nukleobase und das hydrophobe Tolan-Motiv. Wie in dieser Arbeit gezeigt wurde, ist eine optimale Anordnung notwendig, um eine gestapelte Anordnung der Tolane zu ermöglichen. Daher wird eine weitere Optimierung des Rückgrates notwendig sein, um die Triplexbildung durch hydrophobe Erkennung in der Hauptfurche der DNA-Duplex zu ermöglichen.

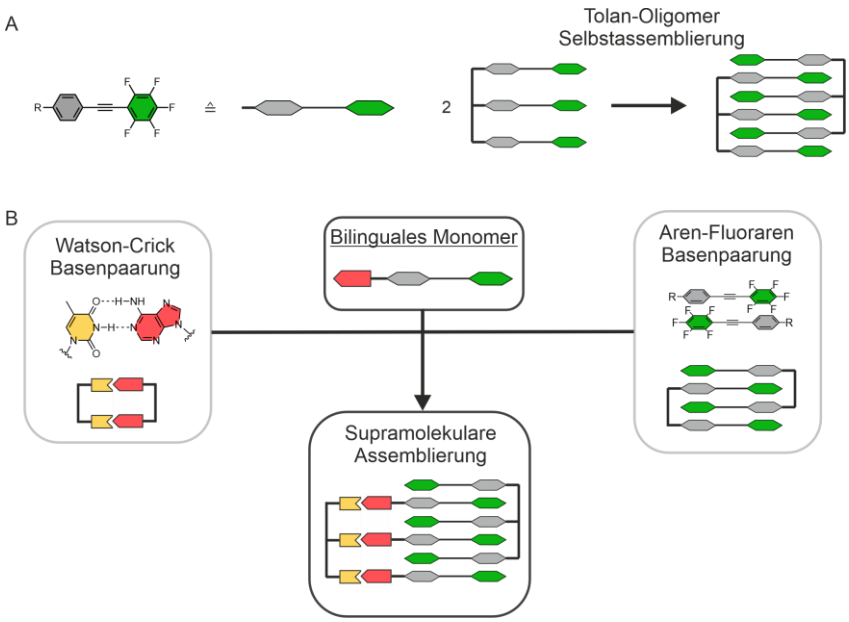


Abbildung 8.1 (A) Schematische Darstellung der Selbstassemblierung eines Tolan-Oligomers durch Aren-Fluoraren-Wechselwirkungen. Die Stränge sind nur aus Tolanbausteinen aufgebaut, die ein Dimer bilden. (B) Kombination einer Watson-Crick-Basenpaarung und einer Aren-Fluoraren-Basenpaarung, die zu einem bilingualen Monomer führt, das die Bildung supramolekularer Aggregate ermöglicht. Das bilinguale Monomer fungiert als Übersetzer zwischen den beiden unterschiedlichen Sprachen.

Auch mit der hier vorgestellten neuen Vernetzungsreaktion sind weitere Entwicklungen möglich. Eine Richtung, die es zu erforschen gilt, ist die Assemblierung von ^{Phe}dU- oder ^{Tol}dU-Nukleosiden auf einem DNA-Templat anstelle des direkten Einbaus in einem DNA-Strang durch Phosphoramidit-Chemie. Bei erfolgreicher Reaktion entstehen ^{Phe}dU- oder ^{Tol}dU-Oligomere, wobei die ^{Tol}dU-Oligomere aufgrund ihrer fluoreszierenden Eigenschaften von besonderem Interesse sind (siehe Abbildung 8.2). Da es sich um eine Strangverknüpfung handelt, kann das Reaktionsprodukt als DNA-Schaden beschrieben werden, was es für Untersuchungen im Zusammenhang mit der DNA-Reparatur interessant macht. Der einfache Einbau einer ^{Phe}dU-Vernetzung ermöglicht den Einbau in verschiedene DNA-Substrate und damit eine bessere Charakterisierung von Enzymen, die an der DNA-Reparatur beteiligt sind. Andere *in vitro* Anwendungen in der Nukleinsäurechemie sind ebenfalls möglich. Die *in vitro* Selektion von

Desoxyribozymen ist ein Beispiel dafür. Viele Selektionsstrategien verwenden eine Haarnadelstruktur, um die Stränge des Substrats mit den potentiellen Sequenzen des Ribozyms zu verbinden. Die Haarnadelstruktur-Vernetzung kann verwendet werden, um erst nach dem Inkubationsschritt eine kovalente Bindung zwischen dem Substrat und dem Ribozymstrang herzustellen, was die Selektion von Ribozymen ermöglicht, die die Haarnadelstruktur nicht benötigen, um aktiv zu sein.

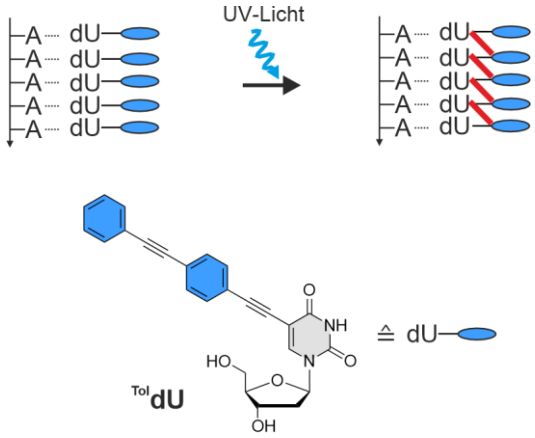


Abbildung 8.2 Darstellung der templatgestützten Reaktion von ^{Tol}dU-Nukleosiden zu ^{Tol}dU-Oligomeren.

Zusammenfassend lässt sich sagen, dass diese Arbeit mehrere Beispiele für die Funktionalisierung von DNA liefert, einschließlich eines neuartigen Erkennungsmotivs und kontrollierter Reaktivität, und zeigt, dass das Verständnis der Stapelwechselwirkung nützliche Einblicke in die Nukleinsäurechemie liefern kann.

Individual contributions

The coauthors of the publications included in this cumulative thesis are informed and agree with the reprint and the individual contributions as stated below.

Tailored Tolane-Perfluorotolane Assembly as Supramolecular Base Pair Replacement in DNA

Neitz, H.; Bessi, I.; Kachler, V.; Michel, M.; Höbartner, C. *Angew. Chem., Int. Ed.,* **2023**, *62*, e202214456.

Angabe Autoren/innen (ggf. Haupt- / Ko- / korrespondierende/r Autor/in) mit Vorname Nachname (Initialen): Hermann Neitz (Hauptautor, HN), Irene Bessi (IB), Valentin Kachler (VK), Manuela Michel (MM), Claudia Höbartner (korrespondierender Autor, CH).

Detaillierte Darstellung der Anteile an der Veröffentlichung (in %)

Autor	HN	IB	VK	MM	CH	∑ **in %**
Research design	2%				3%	5%
Synthesis and characterization	8%		6%	6%		20%
Thermodynamic analysis	15%		5%			20%
Kinetic analysis	10%					10%
NMR characterization	10%	10%				20%
Publication writing	5%	5%			5%	15%
Publication correction		1%			4%	5%
Publication coordination					5%	5%
Total	50%	16%	11%	6%	17%	100%

Programmable DNA Interstrand Crosslinking by Alkene–Alkyne [2 + 2] Photocycloaddition

Neitz, H.; Bessi, I.; Kuper, J.; Kisker, C.; Höbartner, C. *J. Am. Chem. Soc.* **2023**, *145*, 9428–9433.

Angabe Autoren/innen (ggf. Haupt- / Ko- / korrespondierende/r Autor/in) mit Vorname Nachname (Initialen): Hermann Neitz (Hauptautor, HN), Irene Bessi (IB), Jochen Kuper (JK), Caroline Kisker(CK), Claudia Höbartner (korrespondierender Autor, CH).

Detaillierte Darstellung der Anteile an der Veröffentlichung (in %)

Autor	HN	IB	JK	CK	CH	∑ **in %**
Research design	4%		1%	1%	4%	10%
Synthesis and characterization	15%					15%
Kinetic analysis	20%					20%
Product characterization	15%	5%				20%
Enzyme kinetics	12%		3%			15%
Publication writing	6%				4%	10%
Publication correction			1%	1%	3%	5%
Publication coordination					5%	5%
Total	72%	5%	5%	2%	16%	100%

A tolane-modified 5-ethynyluridine as a universal and fluorogenic photochemical DNA crosslinker

Neitz, H.; Höbartner, C. *Chem. Commun.* **2023**, *59*, 12003–12006.

Angabe Autoren/innen (ggf. Haupt- / Ko- / korrespondierende/r Autor/in) mit Vorname Nachname (Initialen): Hermann Neitz (Hauptautor, HN), Claudia Höbartner (korrespondierender Autor, CH).

Detaillierte Darstellung der Anteile an der Veröffentlichung (in %)

Autor	HN	CH	**∑ in %**
Research design	5%	5%	10%
Synthesis and characterization	15%		15%
Kinetic analysis	20%		20%
Product characterization	20%		20%
Hybridization chain reaction	15%		15%
Publication writing	5%	5%	10%
Publication correction		5%	5%
Publication coordination		5%	5%
Total	80%	20%	100%

Abbreviations

2D	two-dimensional
3D	three-dimensional
A	adenine
A	anisotropy
Abs	absorption
ac	acetyl group
aeg	aminoethylglycl
AFM	atomic force microscopy
ALKBH3	alpha-ketoglutarate-dependent dioxygenase AlkB homolog 3
ANA	arabinose nucleic acid
APS	ammonium persulfate
ATP	adenosine triphosphate
BuNA	butyl nucleic acid
bz	benzoyl group
C	cytosine
c	concentration
Calc.	calculated
CeNA	cyclohexene nucleic acid
CEP-Cl	2-cyanoethyl *N*,*N*-diisopropylchlorophosphoramidite
CLEANEX-PM	phase-modulated clean chemical exchange
COSY	correlated spectroscopy
CPD	cyclobutane pyrimidine dimer
CPG	controlled pore glass
CSP	chemical shift perturbation
CuAAC	copper(I)-catalyzed alkyne-azide cycloaddition
Cy3	cyanine 3
Cy5	cyanine 5
δ	chemical shift
DABCYL	4-((4-(dimethylamino)phenyl)azo)benzoic acid
DCM	dichloromethane
DIAD	diisopropyl azodicarboxylate
DMSO	dimethyl sulfoxide
DIPEA	*N*,*N*-diisopropylethylamine
DMF	dimethylformamide
dmf	dimethylformamide group
DMT	4,4'-dimethoxytrityl
DNA	deoxyribonucleic acid
DQF	double quantum filtered
DSS	sodium trimethylsilylpropanesulfonate
DX	double crossover

EDTA	ethylenediamine tetraacetate
Em	emission
EMSA	electrophoretic mobility shift assay
EIC	extracted-ion chromatogram
Ex	excitation
EtOAc	ethyl acetate
ETT	5-ethylthio-1*H*-tetrazole
F	fluorescence
FANA	2'-fluoro-arabinose nucleic acid
FNA	flexible nucleic acid
FRET	fluorescence resonance energy transfer
FRNA	2'-fluoro ribonucleic acid
G	guanine
G	free energy
GFP	green fluorescent protein
GNA	glycol nucleic acid
H	enthalpy
HCR	hybridization chain reaction
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HMBC	heteronuclear multiple-bond correlation
HNA	hexitol nucleic acid
HOESY	heteronuclear Overhauser effect spectroscopy
HPLC	high performance liquid chromatography
HR-ESI-MS	high resolution electrospray ionization mass spectrometry
HSQC	heteronuclear single quantum coherence
Hyp	hyperchromicity
I	intensity
ICL	interstrand crosslink
J	coupling constant
k	rate constant
K_D	dissociation constant
λ	wavelength
LC-MS	liquid-chromatography mass-spectrometry
LNA	locked nucleic acid
m1A	1-methyladenine
MeCN	acetonitrile
MeOH	methanol
mRNA	messenger ribonucleic acid
NOESY	nuclear Overhauser effect spectroscopy
NTP	nucleoside triphosphate
PA	phosphoramidite
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PDB	Protein Data Bank
PNA	peptide nucleic acid
^{Phe}dU	5-phenylethynyl-2'-deoxyuridine
ppm	parts per million
*p*TsOH	*p*-toluenesulfonic acid

MOPS	3-(*N*-morpholino)propanesulfonic acid
NMR	nuclear magnetic resonance
NTH1	endonuclease III-like protein 1
r4CI	1'-(4-cyano-1*H*-indol-1-yl)-β-D-ribofuranose
Ref	reference
RNA	ribonucleic acid
RP	reversed phase
ROESY	rotating-frame nuclear Overhauser effect spectroscopy
rRNA	ribosomal ribonucleic acid
S	entropy
SARS-CoV-2	severe acute respiratory syndrome coronavirus 2
SELEX	systematic evolution of ligands by exponential enrichment
SNP	single nucleotide polymorphism
SPS	solvent purification systems
T	thymine
T	temperature
T_m	melting temperature
t	time
^{t}BuOH	*tert*-butyl alcohol
TEAA	triethylammonium acetate
TEMED	tetramethylethylenediamine
TFIIH	transcription factor II H
TFO	triplex-forming oligonucleotide
THF	tetrahydrofuran
TLC	thin-layer chromatography
TMS	trimethylsilyl group
TNA	threose nucleic acid
^{Tol}dU	5-[4-(phenylethynyl)phenylethynyl]-2'-deoxyuridine
TRIS	tris(hydroxymethyl)aminomethane
tRNA	transfer ribonucleic acid
TX	triple crossover
Tyr	tyrosine
UBER	universal base excision reporter
UBP	unnatural base pair
UNA	unlocked nucleic acid
UV	ultraviolet
Vis	visible
XNA	xeno nucleic acid

References

1. Watson, J. D.; Crick, F. H., Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature* **1953,** *171*, 737-738.

2. Aldaye, F. A.; Palmer, A. L.; Sleiman, H. F., Assembling materials with DNA as the guide. *Science* **2008,** *321*, 1795-1799.

3. McLaughlin, C. K.; Hamblin, G. D.; Sleiman, H. F., Supramolecular DNA assembly. *Chem. Soc. Rev.* **2011,** *40*, 5647-5656.

4. Kool, E. T., Hydrogen bonding, base stacking, and steric effects in dna replication. *Annu. Rev. Biophys. Biomol. Struct.* **2001,** *30*, 1-22.

5. Jurečka, P.; Hobza, P., True stabilization energies for the optimal planar hydrogen-bonded and stacked structures of guanine...cytosine, adenine...thymine, and their 9- and 1-methyl derivatives: complete basis set calculations at the MP2 and CCSD(T) levels and comparison with experiment. *J. Am. Chem. Soc.* **2003,** *125*, 15608-15613.

6. Liu, D.; Wyttenbach, T.; Bowers, M. T., Hydration of mononucleotides. *J. Am. Chem. Soc.* **2006,** *128*, 15155-15163.

7. Guckian, K. M.; Krugh, T. R.; Kool, E. T., Solution Structure of a Nonpolar, Non-Hydrogen-Bonded Base Pair Surrogate in DNA. *J. Am. Chem. Soc.* **2000,** *122*, 6841-6847.

8. Hunter, C. A., Sequence-dependent DNA structure. The role of base stacking interactions. *J. Mol. Biol.* **1993,** *230*, 1025-1054.

9. Riley, K. E.; Hobza, P., On the importance and origin of aromatic interactions in chemistry and biodisciplines. *Acc. Chem. Res.* **2013,** *46*, 927-936.

10. Wojciechowski, F.; Leumann, C. J., Alternative DNA base-pairs: from efforts to expand the genetic code to potential material applications. *Chem. Soc. Rev.* **2011,** *40*, 5669-5679.

11. Lindman, B.; Medronho, B.; Alves, L.; Norgren, M.; Nordenskiold, L., Hydrophobic interactions control the self-assembly of DNA and cellulose. *Q. Rev. Biophys.* **2021,** *54*, e3.

12. Barone, G.; Fonseca Guerra, C.; Bickelhaupt, F. M., B-DNA Structure and Stability as Function of Nucleic Acid Composition: Dispersion-Corrected DFT Study of Dinucleoside Monophosphate Single and Double Strands. *ChemistryOpen* **2013,** *2*, 186-193.

13. Dickerson, R. E.; Drew, H. R.; Conner, B. N.; Wing, R. M.; Fratini, A. V.; Kopka, M. L., The anatomy of A-, B-, and Z-DNA. *Science* **1982,** *216*, 475-485.

14. Pohl, F. M.; Jovin, T. M., Salt-induced co-operative conformational change of a synthetic DNA: equilibrium and kinetic studies with poly (dG-dC). *J. Mol. Biol.* **1972,** *67*, 375-396.

15. Wang, A. H.; Quigley, G. J.; Kolpak, F. J.; Crawford, J. L.; van Boom, J. H.; van der Marel, G.; Rich, A., Molecular structure of a left-handed double helical DNA fragment at atomic resolution. *Nature* **1979,** *282*, 680-686.

16. Choi, J.; Majima, T., Conformational changes of non-B DNA. *Chem. Soc. Rev.* **2011,** *40*, 5893-5909.

17. Jaekel, A.; Lill, P.; Whitelam, S.; Saccà, B., Insights into the Structure and Energy of DNA Nanoassemblies. *Molecules* **2020,** *25*, 5466.

18. Hirao, I.; Nishimura, Y.; Tagawa, Y.; Watanabe, K.; Miura, K., Extraordinarily stable mini-hairpins: electrophoretical and thermal properties of the various sequence variants of d(GCGAAAGC) and their effect on DNA sequencing. *Nucleic Acids Res.* **1992,** *20*, 3891-3896.

19. Yoshizawa, S.; Kawai, G.; Watanabe, K.; Miura, K.; Hirao, I., GNA trinucleotide loop sequences producing extraordinarily stable DNA minihairpins. *Biochemistry* **1997,** *36*, 4761-4767.

20. Kannan, S.; Zacharias, M., Role of the closing base pair for d(GCA) hairpin stability: free energy analysis and folding simulations. *Nucleic Acids Res.* **2011,** *39*, 8271-8280.

21. Nayak, R. K.; Peersen, O. B.; Hall, K. B.; Van Orden, A., Millisecond time-scale folding and unfolding of DNA hairpins using rapid-mixing stopped-flow kinetics. *J. Am. Chem. Soc.* **2012,** *134*, 2453-2456.

22. Liu, Y.; West, S. C., Happy Hollidays: 40th anniversary of the Holliday junction. *Nat. Rev. Mol. Cell Biol.* **2004,** *5*, 937-944.

23. Kallenbach, N. R.; Ma, R.-I.; Seeman, N. C., An immobile nucleic acid junction constructed from oligonucleotides. *Nature* **1983,** *305*, 829-831.

24. Hohng, S.; Zhou, R.; Nahas, M. K.; Yu, J.; Schulten, K.; Lilley, D. M.; Ha, T., Fluorescence-force spectroscopy maps two-dimensional reaction landscape of the holliday junction. *Science* **2007,** *318*, 279-283.

25. Dalla Pozza, M.; Abdullrahman, A.; Cardin, C. J.; Gasser, G.; Hall, J. P., Three's a crowd - stabilisation, structure, and applications of DNA triplexes. *Chem. Sci.* **2022,** *13*, 10193-10215.

26. Plum, G. E.; Park, Y. W.; Singleton, S. F.; Dervan, P. B.; Breslauer, K. J., Thermodynamic characterization of the stability and the melting behavior of a DNA triplex: a spectroscopic and calorimetric study. *Proc. Natl. Acad. Sci. USA* **1990,** *87*, 9436-9440.

27. Varshney, D.; Spiegel, J.; Zyner, K.; Tannahill, D.; Balasubramanian, S., The regulation and functions of DNA and RNA G-quadruplexes. *Nat. Rev. Mol. Cell Biol.* **2020,** *21*, 459-474.

28. Abou Assi, H.; Garavís, M.; González, C.; Damha, M. J., i-Motif DNA: structural features and significance to cell biology. *Nucleic Acids Res.* **2018,** *46*, 8038-8056.

29. Wright, E. P.; Huppert, J. L.; Waller, Z. A. E., Identification of multiple genomic DNA sequences which form i-motif structures at neutral pH. *Nucleic Acids Res.* **2017,** *45*, 2951-2959.

30. Wang, G.; Vasquez, K. M., Dynamic alternative DNA structures in biology and disease. *Nat. Rev. Genet.* **2023,** *24*, 211-234.

31. Micura, R.; Höbartner, C., Fundamental studies of functional nucleic acids: aptamers, riboswitches, ribozymes and DNAzymes. *Chem. Soc. Rev.* **2020,** *49*, 7331-7353.

32. Ellington, A. D.; Szostak, J. W., In vitro selection of RNA molecules that bind specific ligands. *Nature* **1990,** *346*, 818-822.

33. Tuerk, C.; Gold, L., Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science* **1990,** *249*, 505-510.

34. Wang, T.; Chen, C.; Larcher, L. M.; Barrero, R. A.; Veedu, R. N., Three decades of nucleic acid aptamer technologies: Lessons learned, progress and opportunities on aptamer development. *Biotechnol. Adv.* **2019,** *37*, 28-50.

35. Yang, L. F.; Ling, M.; Kacherovsky, N.; Pun, S. H., Aptamers 101: aptamer discovery and in vitro applications in biosensors and separations. *Chem. Sci.* **2023,** *14*, 4961-4978.

36. Keefe, A. D.; Pai, S.; Ellington, A., Aptamers as therapeutics. *Nat. Rev. Drug Discov.* **2010,** *9*, 537-550.

37. Dunn, M. R.; Jimenez, R. M.; Chaput, J. C., Analysis of aptamer discovery and technology. *Nat. Rev. Chem.* **2017,** *1*, 0076.

38. Bock, L. C.; Griffin, L. C.; Latham, J. A.; Vermaas, E. H.; Toole, J. J., Selection of single-stranded DNA molecules that bind and inhibit human thrombin. *Nature* **1992,** *355*, 564-566.

39. Ellington, A. D.; Szostak, J. W., Selection in vitro of single-stranded DNA molecules that fold into specific ligand-binding structures. *Nature* **1992,** *355*, 850-852.

40. Schultze, P.; Macaya, R. F.; Feigon, J., Three-dimensional solution structure of the thrombin-binding DNA aptamer d(GGTTGGTGTGGTTGG). *J. Mol. Biol.* **1994,** *235*, 1532-1547.

41. Cai, S.; Yan, J.; Xiong, H.; Liu, Y.; Peng, D.; Liu, Z., Investigations on the interface of nucleic acid aptamers and binding targets. *Analyst* **2018,** *143*, 5317-5338.

42. Yang, L. F.; Kacherovsky, N.; Liang, J.; Salipante, S. J.; Pun, S. H., SCORe: SARS-CoV-2 Omicron Variant RBD-Binding DNA Aptamer for Multiplexed Rapid Detection and Pseudovirus Neutralization. *Anal. Chem.* **2022,** *94*, 12683-12690.

43. Paige, J. S.; Wu, K. Y.; Jaffrey, S. R., RNA mimics of green fluorescent protein. *Science* **2011,** *333*, 642-646.

44. VarnBuhler, B. S.; Moon, J.; Dey, S. K.; Wu, J.; Jaffrey, S. R., Detection of SARS-CoV-2 RNA Using a DNA Aptamer Mimic of Green Fluorescent Protein. *ACS Chem. Biol.* **2022,** *17*, 840-853.

45. Passalacqua, L. F. M.; Banco, M. T.; Moon, J. D.; Li, X.; Jaffrey, S. R.; Ferré-D'Amaré, A. R., Intricate 3D architecture of a DNA mimic of GFP. *Nature* **2023,** *618*, 1078-1084.

46. Silverman, S. K.; Baum, D. A., Use of deoxyribozymes in RNA research. *Methods Enzymol.* **2009,** *469*, 95-117.

47. Breaker, R. R.; Joyce, G. F., A DNA enzyme that cleaves RNA. *Chem. Biol.* **1994,** *1*, 223-229.

48. Carmi, N.; Balkhi, S. R.; Breaker, R. R., Cleaving DNA with DNA. *Proc. Natl. Acad. Sci. USA* **1998,** *95*, 2233-2237.

49. Gong, L.; Zhao, Z.; Lv, Y. F.; Huan, S. Y.; Fu, T.; Zhang, X. B.; Shen, G. L.; Yu, R. Q., DNAzyme-based biosensors and nanodevices. *Chem. Commun.* **2015,** *51*, 979-995.

50. Sednev, M. V.; Mykhailiuk, V.; Choudhury, P.; Halang, J.; Sloan, K. E.; Bohnsack, M. T.; Höbartner, C., N(6) -Methyladenosine-Sensitive RNA-Cleaving Deoxyribozymes. *Angew. Chem.; Int. Ed.* **2018,** *57*, 15117-15121.

51. Liaqat, A.; Stiller, C.; Michel, M.; Sednev, M. V.; Höbartner, C., N(6) -Isopentenyladenosine in RNA Determines the Cleavage Site of Endonuclease Deoxyribozymes. *Angew. Chem., Int. Ed.* **2020,** *59*, 18627-18631.

52. Liaqat, A.; Sednev, M. V.; Stiller, C.; Höbartner, C., RNA-Cleaving Deoxyribozymes Differentiate Methylated Cytidine Isomers in RNA. *Angew. Chem., Int. Ed.* **2021,** *60*, 19058-19062.

53. Purtha, W. E.; Coppins, R. L.; Smalley, M. K.; Silverman, S. K., General deoxyribozyme-catalyzed synthesis of native 3'-5' RNA linkages. *J. Am. Chem. Soc.* **2005,** *127*, 13124-13125.

54. Ponce-Salvatierra, A.; Wawrzyniak-Turek, K.; Steuerwald, U.; Höbartner, C.; Pena, V., Crystal structure of a DNA catalyst. *Nature* **2016,** *529*, 231-234.

55. Liu, H.; Yu, X.; Chen, Y.; Zhang, J.; Wu, B.; Zheng, L.; Haruehanroengra, P.; Wang, R.; Li, S.; Lin, J.; Li, J.; Sheng, J.; Huang, Z.; Ma, J.; Gan, J., Crystal structure of an RNA-cleaving DNAzyme. *Nat. Commun.* **2017,** *8*, 2006.

56. Borggrafe, J.; Victor, J.; Rosenbach, H.; Viegas, A.; Gertzen, C. G. W.; Wuebben, C.; Kovacs, H.; Gopalswamy, M.; Riesner, D.; Steger, G.; Schiemann, O.; Gohlke, H.; Span, I.; Etzkorn, M., Time-resolved structural analysis of an RNA-cleaving DNA catalyst. *Nature* **2022,** *601*, 144-149.

57. Pradeepkumar, P. I.; Höbartner, C.; Baum, D. A.; Silverman, S. K., DNA-catalyzed formation of nucleopeptide linkages. *Angew. Chem., Int. Ed.* **2008,** *47*, 1753-7.

58. Zhou, C.; Avins, J. L.; Klauser, P. C.; Brandsen, B. M.; Lee, Y.; Silverman, S. K., DNA-Catalyzed Amide Hydrolysis. *J. Am. Chem. Soc.* **2016,** *138*, 2106-2109.

59. Chinnapen, D. J.; Sen, D., A deoxyribozyme that harnesses light to repair thymine dimers in DNA. *Proc. Natl. Acad. Sci. USA* **2004,** *101*, 65-69.

60. Li, Y.; Breaker, R. R., Phosphorylating DNA with DNA. *Proc. Natl. Acad. Sci. USA* **1999,** *96*, 2746-2751.

61. Svehlova, K.; Lukšan, O.; Jakubec, M.; Curtis, E. A., Supernova: A Deoxyribozyme that Catalyzes a Chemiluminescent Reaction. *Angew. Chem., Int. Ed.* **2022,** *61*, e202109347.

62. Seeman, N. C., Nucleic acid junctions and lattices. *J. Theor. Biol.* **1982,** *99*, 237-247.

63. Fu, T. J.; Seeman, N. C., DNA double-crossover molecules. *Biochemistry* **1993,** *32*, 3211-3220.

64. Winfree, E.; Liu, F.; Wenzler, L. A.; Seeman, N. C., Design and self-assembly of two-dimensional DNA crystals. *Nature* **1998,** *394*, 539-544.

65. LaBean, T. H.; Yan, H.; Kopatsch, J.; Liu, F.; Winfree, E.; Reif, J. H.; Seeman, N. C., Construction, Analysis, Ligation, and Self-Assembly of DNA Triple Crossover Complexes. *J. Am. Chem. Soc.* **2000,** *122*, 1848-1860.

66. Mao, C.; LaBean, T. H.; Relf, J. H.; Seeman, N. C., Logical computation using algorithmic self-assembly of DNA triple-crossover molecules. *Nature* **2000,** *407*, 493-496.

67. Liu, D.; Park, S. H.; Reif, J. H.; LaBean, T. H., DNA nanotubes self-assembled from triple-crossover tiles as templates for conductive nanowires. *Proc. Natl. Acad. Sci. USA* **2004,** *101*, 717-722.

68. He, Y.; Chen, Y.; Liu, H.; Ribbe, A. E.; Mao, C., Self-assembly of hexagonal DNA two-dimensional (2D) arrays. *J. Am. Chem. Soc.* **2005,** *127*, 12202-12203.

69. He, Y.; Tian, Y.; Ribbe, A. E.; Mao, C., Highly connected two-dimensional crystals of DNA six-point-stars. *J. Am. Chem. Soc.* **2006,** *128*, 15978-15979.

70. Yan, H.; Park, S. H.; Finkelstein, G.; Reif, J. H.; LaBean, T. H., DNA-templated self-assembly of protein arrays and highly conductive nanowires. *Science* **2003,** *301*, 1882-1884.

71. Seeman, N. C.; Sleiman, H. F., DNA nanotechnology. *Nat. Rev. Mat.* **2017,** *3*, 17068.

72. Sacca, B.; Meyer, R.; Feldkamp, U.; Schroeder, H.; Niemeyer, C. M., High-throughput, real-time monitoring of the self-assembly of DNA nanostructures by FRET spectroscopy. *Angew. Chem., Int. Ed.* **2008,** *47*, 2135-2137.

73. Chen, J. H.; Seeman, N. C., Synthesis from DNA of a molecule with the connectivity of a cube. *Nature* **1991,** *350*, 631-633.

74. Zhang, Y.; Seeman, N. C., Construction of a DNA-Truncated Octahedron. *J. Am. Chem. Soc.* **1994,** *116*, 1661-1669.

75. Goodman, R. P.; Schaap, I. A.; Tardin, C. F.; Erben, C. M.; Berry, R. M.; Schmidt, C. F.; Turberfield, A. J., Rapid chiral assembly of rigid DNA building blocks for molecular nanofabrication. *Science* **2005,** *310*, 1661-1665.

76. He, Y.; Ye, T.; Su, M.; Zhang, C.; Ribbe, A. E.; Jiang, W.; Mao, C., Hierarchical self-assembly of DNA into symmetric supramolecular polyhedra. *Nature* **2008,** *452*, 198-201.

77. Bhatia, D.; Mehtab, S.; Krishnan, R.; Indi, S. S.; Basu, A.; Krishnan, Y., Icosahedral DNA nanocapsules by modular assembly. *Angew. Chem., Int. Ed.* **2009,** *48*, 4134-4137.

78. Seeman, N. C., DNA in a material world. *Nature* **2003,** *421*, 427-431.

79. Zheng, J.; Birktoft, J. J.; Chen, Y.; Wang, T.; Sha, R.; Constantinou, P. E.; Ginell, S. L.; Mao, C.; Seeman, N. C., From molecular to macroscopic via the rational design of a self-assembled 3D DNA crystal. *Nature* **2009,** *461*, 74-77.

80. Shih, W. M.; Quispe, J. D.; Joyce, G. F., A 1.7-kilobase single-stranded DNA that folds into a nanoscale octahedron. *Nature* **2004,** *427*, 618-621.

81. Lu, B.; Vecchioni, S.; Ohayon, Y. P.; Sha, R.; Woloszyn, K.; Yang, B.; Mao, C.; Seeman, N. C., 3D Hexagonal Arrangement of DNA Tensegrity Triangles. *ACS Nano* **2021,** *15*, 16788-16793.

82. Rothemund, P. W., Folding DNA to create nanoscale shapes and patterns. *Nature* **2006,** *440*, 297-302.

83. Li, Z.; Liu, M.; Wang, L.; Nangreave, J.; Yan, H.; Liu, Y., Molecular Behavior of DNA Origami in Higher-Order Self-Assembly. *J. Am. Chem. Soc.* **2010,** *132*, 13545-13552.

84. Liu, W.; Zhong, H.; Wang, R.; Seeman, N. C., Crystalline two-dimensional DNA-origami arrays. *Angew. Chem., Int. Ed.* **2011,** *50*, 264-267.

85. Woo, S.; Rothemund, P. W., Programmable molecular recognition based on the geometry of DNA nanostructures. *Nat. Chem.* **2011,** *3*, 620-627.

86. Tikhomirov, G.; Petersen, P.; Qian, L., Programmable disorder in random DNA tilings. *Nat. Nanotechnol.* **2017,** *12*, 251-259.

87. Tikhomirov, G.; Petersen, P.; Qian, L., Fractal assembly of micrometre-scale DNA origami arrays with arbitrary patterns. *Nature* **2017,** *552*, 67-71.

88. Hong, F.; Zhang, F.; Liu, Y.; Yan, H., DNA Origami: Scaffolds for Creating Higher Order Structures. *Chem. Rev.* **2017,** *117*, 12584-12640.

89. Andersen, E. S.; Dong, M.; Nielsen, M. M.; Jahn, K.; Subramani, R.; Mamdouh, W.; Golas, M. M.; Sander, B.; Stark, H.; Oliveira, C. L.; Pedersen, J. S.; Birkedal, V.; Besenbacher, F.; Gothelf, K. V.; Kjems, J., Self-assembly of a nanoscale DNA box with a controllable lid. *Nature* **2009,** *459*, 73-76.

90. Douglas, S. M.; Dietz, H.; Liedl, T.; Högberg, B.; Graf, F.; Shih, W. M., Self-assembly of DNA into nanoscale three-dimensional shapes. *Nature* **2009,** *459*, 414-418.

91. Dietz, H.; Douglas, S. M.; Shih, W. M., Folding DNA into twisted and curved nanoscale shapes. *Science* **2009,** *325*, 725-730.

92. Han, D.; Pal, S.; Nangreave, J.; Deng, Z.; Liu, Y.; Yan, H., DNA origami with complex curvatures in three-dimensional space. *Science* **2011,** *332*, 342-346.

93. Benson, E.; Mohammed, A.; Gardell, J.; Masich, S.; Czeizler, E.; Orponen, P.; Högberg, B., DNA rendering of polyhedral meshes at the nanoscale. *Nature* **2015,** *523*, 441-444.

94. Wei, B.; Dai, M.; Yin, P., Complex shapes self-assembled from single-stranded DNA tiles. *Nature* **2012,** *485*, 623-626.

95. Rothemund, P. W. K.; Andersen, E. S., The importance of being modular. *Nature* **2012,** *485*, 584-585.

96. Ke, Y.; Ong, L. L.; Shih, W. M.; Yin, P., Three-dimensional structures self-assembled from DNA bricks. *Science* **2012,** *338*, 1177-1183.

97. Gothelf, K. V., Materials science. LEGO-like DNA structures. *Science* **2012,** *338*, 1159-1160.

98. Ong, L. L.; Hanikel, N.; Yaghi, O. K.; Grun, C.; Strauss, M. T.; Bron, P.; Lai-Kee-Him, J.; Schueder, F.; Wang, B.; Wang, P.; Kishi, J. Y.; Myhrvold, C.; Zhu, A.; Jungmann, R.; Bellot, G.; Ke, Y.; Yin, P., Programmable self-assembly of three-dimensional nanostructures from 10,000 unique components. *Nature* **2017,** *552*, 72-77.

99. Han, D.; Qi, X.; Myhrvold, C.; Wang, B.; Dai, M.; Jiang, S.; Bates, M.; Liu, Y.; An, B.; Zhang, F.; Yan, H.; Yin, P., Single-stranded DNA and RNA origami. *Science* **2017,** *358*, eaao2648.

100. Praetorius, F.; Kick, B.; Behler, K. L.; Honemann, M. N.; Weuster-Botz, D.; Dietz, H., Biotechnological mass production of DNA origami. *Nature* **2017,** *552*, 84-87.

101. Liu, W.; Halverson, J.; Tian, Y.; Tkachenko, A. V.; Gang, O., Self-organized architectures from assorted DNA-framed nanoparticles. *Nat. Chem.* **2016,** *8*, 867-873.

102. Nickels, P. C.; Wunsch, B.; Holzmeister, P.; Bae, W.; Kneer, L. M.; Grohmann, D.; Tinnefeld, P.; Liedl, T., Molecular force spectroscopy with a DNA origami-based nanoscopic force clamp. *Science* **2016,** *354*, 305-307.

103. Funke, J. J.; Ketterer, P.; Lieleg, C.; Schunter, S.; Korber, P.; Dietz, H., Uncovering the forces between nucleosomes using DNA origami. *Sci. Adv.* **2016,** *2*, e160097.

104. Kilchherr, F.; Wachauf, C.; Pelz, B.; Rief, M.; Zacharias, M.; Dietz, H., Single-molecule dissection of stacking forces in DNA. *Science* **2016,** *353*, aaf5508.

105. Pumm, A. K.; Engelen, W.; Kopperger, E.; Isensee, J.; Vogt, M.; Kozina, V.; Kube, M.; Honemann, M. N.; Bertosin, E.; Langecker, M.; Golestanian, R.; Simmel, F. C.; Dietz, H., A DNA origami rotary ratchet motor. *Nature* **2022,** *607*, 492-498.

106. Zhan, P.; Jahnke, K.; Liu, N.; Gopfrich, K., Functional DNA-based cytoskeletons for synthetic cells. *Nat. Chem.* **2022,** *14*, 958-963.

107. Mao, C.; Sun, W.; Shen, Z.; Seeman, N. C., A nanomechanical device based on the B-Z transition of DNA. *Nature* **1999,** *397*, 144-146.

108. Zhang, D. Y.; Seelig, G., Dynamic DNA nanotechnology using strand-displacement reactions. *Nat. Chem.* **2011,** *3*, 103-113.

109. Simmel, F. C.; Yurke, B.; Singh, H. R., Principles and Applications of Nucleic Acid Strand Displacement Reactions. *Chem. Rev.* **2019,** *119*, 6326-6369.

110. Reynaldo, L. P.; Vologodskii, A. V.; Neri, B. P.; Lyamichev, V. I., The kinetics of oligonucleotide replacements. *J. Mol. Biol.* **2000,** *297*, 511-520.

111. Tyagi, S.; Kramer, F. R., Molecular Beacons: Probes that Fluoresce upon Hybridization. *Nat. Biotechnol.* **1996,** *14*, 303-308.

112. Zheng, J.; Yang, R.; Shi, M.; Wu, C.; Fang, X.; Li, Y.; Li, J.; Tan, W., Rationally designed molecular beacons for bioanalytical and biomedical applications. *Chem. Soc. Rev.* **2015,** *44*, 3036-3055.

113. Zhang, X. B.; Wang, Z.; Xing, H.; Xiang, Y.; Lu, Y., Catalytic and molecular beacons for amplified detection of metal ions and organic molecules with high sensitivity. *Anal. Chem.* **2010,** *82*, 5005-5011.

114. Thurley, S.; Roglin, L.; Seitz, O., Hairpin peptide beacon: dual-labeled PNA-peptide-hybrids for protein detection. *J. Am. Chem. Soc.* **2007,** *129*, 12693-12695.

115. Irmisch, P.; Ouldridge, T. E.; Seidel, R., Modeling DNA-Strand Displacement Reactions in the Presence of Base-Pair Mismatches. *J. Am. Chem. Soc.* **2020,** *142*, 11451-11463.

116. Yurke, B.; Turberfield, A. J.; Mills, A. P., Jr.; Simmel, F. C.; Neumann, J. L., A DNA-fuelled molecular machine made of DNA. *Nature* **2000,** *406*, 605-608.

117. Yan, H.; Zhang, X.; Shen, Z.; Seeman, N. C., A robust DNA mechanical device controlled by hybridization topology. *Nature* **2002,** *415*, 62-65.

118. Feng, L.; Park, S. H.; Reif, J. H.; Yan, H., A two-state DNA lattice switched by DNA nanoactuator. *Angew. Chem., Int. Ed.* **2003,** *42*, 4342-4346.

119. Sherman, W. B.; Seeman, N. C., A Precisely Controlled DNA Biped Walking Device. *Nano Letters* **2004,** *4*, 1203-1207.

120. Omabegho, T.; Sha, R.; Seeman, N. C., A bipedal DNA Brownian motor with coordinated legs. *Science* **2009,** *324*, 67-71.

121. Lund, K.; Manzo, A. J.; Dabby, N.; Michelotti, N.; Johnson-Buck, A.; Nangreave, J.; Taylor, S.; Pei, R.; Stojanovic, M. N.; Walter, N. G.; Winfree, E.; Yan, H., Molecular robots guided by prescriptive landscapes. *Nature* **2010,** *465*, 206-210.

122. Gu, H.; Chao, J.; Xiao, S. J.; Seeman, N. C., A proximity-based programmable DNA nanoscale assembly line. *Nature* **2010,** *465*, 202-205.

123. Thubagere, A. J.; Li, W.; Johnson, R. F.; Chen, Z.; Doroudi, S.; Lee, Y. L.; Izatt, G.; Wittman, S.; Srinivas, N.; Woods, D.; Winfree, E.; Qian, L., A cargo-sorting DNA robot. *Science* **2017,** *357*, eaan6558.

124. Seelig, G.; Soloveichik, D.; Zhang, D. Y.; Winfree, E., Enzyme-free nucleic acid logic circuits. *Science* **2006,** *314*, 1585-1588.

125. Qian, L.; Winfree, E., Scaling up digital circuit computation with DNA strand displacement cascades. *Science* **2011,** *332*, 1196-1201.

126. Cherry, K. M.; Qian, L., Scaling up molecular pattern recognition with DNA-based winner-take-all neural networks. *Nature* **2018,** *559*, 370-376.

127. Okumura, S.; Gines, G.; Lobato-Dauzier, N.; Baccouche, A.; Deteix, R.; Fujii, T.; Rondelez, Y.; Genot, A. J., Nonlinear decision-making with enzymatic neural networks. *Nature* **2022,** *610*, 496-501.

128. Yin, P.; Choi, H. M.; Calvert, C. R.; Pierce, N. A., Programming biomolecular self-assembly pathways. *Nature* **2008,** *451*, 318-322.

129. Glynn, A. T.; Davidson, S. R.; Qian, L., Developmental Self-Assembly of a DNA Ring with Stimulus-Responsive Size and Growth Direction. *J. Am. Chem. Soc.* **2022,** *144*, 10075-10079.

130. Dirks, R. M.; Pierce, N. A., Triggered amplification by hybridization chain reaction. *Proc. Natl. Acad. Sci. USA* **2004,** *101*, 15275-15278.

131. Ang, Y. S.; Yung, L. Y., Rational design of hybridization chain reaction monomers for robust signal amplification. *Chem. Commun.* **2016,** *52*, 4219-4222.

132. Figg, C. A.; Winegar, P. H.; Hayes, O. G.; Mirkin, C. A., Controlling the DNA Hybridization Chain Reaction. *J. Am. Chem. Soc.* **2020,** *142*, 8596-8601.

133. Idili, A.; Porchetta, A.; Amodio, A.; Vallee-Belisle, A.; Ricci, F., Controlling Hybridization Chain Reactions with pH. *Nano Lett.* **2015,** *15*, 5539-5544.

134. Xuan, F.; Hsing, I. M., Triggering hairpin-free chain-branching growth of fluorescent DNA dendrimers for nonlinear hybridization chain reaction. *J. Am. Chem. Soc.* **2014,** *136*, 9810-9813.

135. Bi, S.; Chen, M.; Jia, X.; Dong, Y.; Wang, Z., Hyperbranched Hybridization Chain Reaction for Triggered Signal Amplification and Concatenated Logic Circuits. *Angew. Chem., Int. Ed.* **2015,** *54*, 8144-8148.

136. Bi, S.; Yue, S.; Zhang, S., Hybridization chain reaction: a versatile molecular tool for biosensing, bioimaging, and biomedicine. *Chem. Soc. Rev.* **2017,** *46*, 4281-4298.

137. Wang, F.; Elbaz, J.; Orbach, R.; Magen, N.; Willner, I., Amplified analysis of DNA by the autonomous assembly of polymers consisting of DNAzyme wires. *J. Am. Chem. Soc.* **2011,** *133*, 17149-17151.

138. Beaucage, S. L.; Caruthers, M. H., Deoxynucleoside phosphoramidites—A new class of key intermediates for deoxypolynucleotide synthesis. *Tetrahedron Lett.* **1981,** *22*, 1859-1862.

139. Rumney, S. I.; Kool, E. T., DNA Recognition by Hybrid Oligoether-Oligodeoxynucleotide Macrocycles. *Angew. Chem., Int. Ed.* **1992,** *31*, 1617-1619.

140. Salunkhe, M.; Wu, T.; Letsinger, R. L., Control of folding and binding of oligonucleotides by use of a nonnucleotide linker. *J. Am. Chem. Soc.* **1992,** *114*, 8768-8772.

141. Lewis, F. D.; Wu, T.; Zhang, Y.; Letsinger, R. L.; Greenfield, S. R.; Wasielewski, M. R., Distance-dependent electron transfer in DNA hairpins. *Science* **1997,** *277*, 673-676.

142. Lewis, F. D.; Liu, X.; Miller, S. E.; Wasielewski, M. R., Electronic Interactions between π-Stacked DNA Base Pairs and Diphenylacetylene-4,4‘-dicarboxamide in Hairpin DNA. *J. Am. Chem. Soc.* **1999,** *121*, 9746-9747.

143. Olenyuk, B.; Fechtenkötter, A.; J. Stang, P., Molecular architecture of cyclic nanostructures: use of co-ordination chemistry in the building of supermolecules with predefined geometric shapes. *J. Chem. Soc., Dalton Trans.* **1998**, 1707-1728.

144. Shi, J.; Bergstrom, D. E., Assembly of Novel DNA Cycles with Rigid Tetrahedral Linkers. *Angew. Chem., Int. Ed.* **1997,** *36*, 111-113.

145. Aldaye, F. A.; Sleiman, H. F., Sequential self-assembly of a DNA hexagon as a template for the organization of gold nanoparticles. *Angew. Chem., Int. Ed.* **2006,** *45*, 2204-2209.

146. Aldaye, F. A.; Sleiman, H. F., Dynamic DNA templates for discrete gold nanoparticle assemblies: control of geometry, modularity, write/erase and structural switching. *J. Am. Chem. Soc.* **2007,** *129*, 4130-4131.

147. Aldaye, F. A.; Sleiman, H. F., Modular access to structurally switchable 3D discrete DNA assemblies. *J. Am. Chem. Soc.* **2007,** *129*, 13376-13377.

148. Scheffler, M.; Dorenbeck, A.; Jordan, S.; Wustefeld, M.; von Kiedrowski, G., Self-Assembly of Trisoligonucleotidyls: The Case for Nano-Acetylene and Nano-Cyclobutadiene. *Angew. Chem., Int. Ed.* **1999,** *38*, 3311-3315.

149. Kuroda, T.; Sakurai, Y.; Suzuki, Y.; Nakamura, A. O.; Kuwahara, M.; Ozaki, H.; Sawai, H., Assembly of DNA nanostructures with branched tris-DNA. *Chem. Asian. J.* **2006,** *1*, 575-580.

150. Zimmermann, J.; Cebulla, M. P.; Monninghoff, S.; von Kiedrowski, G., Self-assembly of a DNA dodecahedron from 20 trisoligonucleotides with C(3h) linkers. *Angew. Chem., Int. Ed.* **2008,** *47*, 3626-3630.

151. Endo, M.; Seeman, N. C.; Majima, T., DNA tube structures controlled by a four-way-branched DNA connector. *Angew. Chem., Int. Ed.* **2005,** *44*, 6074-6077.

152. Li, Y.; Pei, J.; Lu, X.; Jiao, Y.; Liu, F.; Wu, X.; Liu, J.; Ding, B., Hierarchical Assembly of Super-DNA Origami Based on a Flexible and Covalent-Bound Branched DNA Structure. *J. Am. Chem. Soc.* **2021,** *143*, 19893-19900.

153. Hong, B. J.; Eryazici, I.; Bleher, R.; Thaner, R. V.; Mirkin, C. A.; Nguyen, S. T., Directed Assembly of Nucleic Acid-Based Polymeric Nanoparticles from Molecular Tetravalent Cores. *J. Am. Chem. Soc.* **2015,** *137*, 8184-8191.

154. Leininger, S.; Olenyuk, B.; Stang, P. J., Self-assembly of discrete cyclic nanostructures mediated by transition metals. *Chem. Rev.* **2000,** *100*, 853-908.

155. Stewart, K. M.; Rojo, J.; McLaughlin, L. W., Ru(II) tris(bipyridyl) complexes with six oligonucleotide arms as precursors for the generation of supramolecular assemblies. *Angew. Chem., Int. Ed.* **2004,** *43*, 5808-5811.

156. Draksharapu, A.; Boersma, A. J.; Leising, M.; Meetsma, A.; Browne, W. R.; Roelfes, G., Binding of copper(II) polypyridyl complexes to DNA and consequences for DNA-based asymmetric catalysis. *Dalton. Trans.* **2015,** *44*, 3647-3655.

157. Goritz, M.; Kramer, R., Allosteric control of oligonucleotide hybridization by metal-induced cyclization. *J. Am. Chem. Soc.* **2005,** *127*, 18016-18017.

158. Yang, H.; Sleiman, H. F., Templated synthesis of highly stable, electroactive, and dynamic metal-DNA branched junctions. *Angew. Chem., Int. Ed.* **2008,** *47*, 2443-2446.

159. Gothelf, K. V.; Thomsen, A.; Nielsen, M.; Clo, E.; Brown, R. S., Modular DNA-programmed assembly of linear and branched conjugated nanostructures. *J. Am. Chem. Soc.* **2004,** *126*, 1044-1046.

160. Yang, H.; Rys, A. Z.; McLaughlin, C. K.; Sleiman, H. F., Templated ligand environments for the selective incorporation of different metals into DNA. *Angew. Chem., Int. Ed.* **2009,** *48*, 9919-9923.

161. Yang, H.; Altvater, F.; de Bruijn, A. D.; McLaughlin, C. K.; Lo, P. K.; Sleiman, H. F., Chiral metal-DNA four-arm junctions and metalated nanotubular structures. *Angew. Chem., Int. Ed.* **2011,** *50*, 4620-4623.

162. Engelhard, D. M.; Pievo, R.; Clever, G. H., Reversible stabilization of transition-metal-binding DNA G-quadruplexes. *Angew. Chem., Int. Ed.* **2013,** *52*, 12843-12847.

163. Abdelhamid, M. A.; Fabian, L.; MacDonald, C. J.; Cheesman, M. R.; Gates, A. J.; Waller, Z. A., Redox-dependent control of i-Motif DNA structure using copper cations. *Nucleic Acids Res.* **2018,** *46*, 5886-5893.

164. Surin, M.; Ulrich, S., From Interaction to Function in DNA-Templated Supramolecular Self-Assemblies. *ChemistryOpen* **2020,** *9*, 480-498.

165. Nakatani, K.; Hagihara, S.; Goto, Y.; Kobori, A.; Hagihara, M.; Hayashi, G.; Kyo, M.; Nomura, M.; Mishima, M.; Kojima, C., Small-molecule ligand induces nucleotide flipping in (CAG)n trinucleotide repeats. *Nat. Chem. Biol.* **2005,** *1*, 39-43.

166. Iwaura, R.; Hoeben, F. J.; Masuda, M.; Schenning, A. P.; Meijer, E. W.; Shimizu, T., Molecular-level helical stack of a nucleotide-appended oligo(p-phenylenevinylene) directed by supramolecular self-assembly with a complementary oligonucleotide as a template. *J. Am. Chem. Soc.* **2006,** *128*, 13298-13304.

167. Hofsass, R.; Sinn, S.; Biedermann, F.; Wagenknecht, H. A., Programmable and Sequence-Selective Supramolecular Assembly of Two Different Chromophores along DNA Templates. *Chem. Eur. J.* **2018,** *24*, 16257-16261.

168. Anderson, B. A.; Fauché, K.; Karunakaran, S. C.; Yerabolu, J. R.; Hud, N. V.; Krishnamurthy, R., The Unexpected Base-Pairing Behavior of Cyanuric Acid in RNA and Ribose versus Cyanuric Acid Induced Helicene Assembly of Nucleic Acids: Implications for the Pre-RNA Paradigm. *Chem. Eur. J.* **2021,** *27*, 4033-4042.

169. Lachance-Brais, C.; Rammal, M.; Asohan, J.; Katolik, A.; Luo, X.; Saliba, D.; Jonderian, A.; Damha, M. J.; Harrington, M. J.; Sleiman, H. F., Small Molecule-Templated DNA Hydrogel with Record Stiffness Integrates and Releases DNA Nanostructures and Gene Silencing Nucleic Acids. *Adv. Sci.* **2023,** *10*, e2205713.

170. Peng, T.; Dohno, C.; Nakatani, K., Mismatch-binding ligands function as a molecular glue for DNA. *Angew. Chem., Int. Ed.* **2006,** *45*, 5623-5626.

171. Dohno, C.; Atsumi, H.; Nakatani, K., Ligand inducible assembly of a DNA tetrahedron. *Chem. Commun.* **2011,** *47*, 3499-3501.

172. Iwaura, R.; Yoshida, K.; Masuda, M.; Ohnishi-Kameyama, M.; Yoshida, M.; Shimizu, T., Oligonucleotide-templated self-assembly of nucleotide bolaamphiphiles: DNA-like nanofibers edged by a double-helical arrangement of A-T base pairs. *Angew. Chem., Int. Ed.* **2003,** *42*, 1009-1012.

173. Janssen, P. G.; Vandenbergh, J.; van Dongen, J. L.; Meijer, E. W.; Schenning, A. P., ssDNA templated self-assembly of chromophores. *J. Am. Chem. Soc.* **2007,** *129*, 6078-6079.

174. Janssen, P. G. A.; Jabbari-Farouji, S.; Surin, M.; Vila, X.; Gielen, J. C.; de Greef, T. F. A.; Vos, M. R. J.; Bomans, P. H. H.; Sommerdijk, N. A. J. M.; Christianen, P. C. M.; Leclère, P.; Lazzaroni, R.; van der Schoot, P.; Meijer, E. W.; Schenning, A. P. H. J., Insights into Templated Supramolecular Polymerization: Binding of Naphthalene Derivatives to ssDNA Templates of Different Lengths. *J. Am. Chem. Soc.* **2009,** *131*, 1222-1231.

175. Sezi, S.; Wagenknecht, H. A., DNA-templated formation of fluorescent self-assembly of ethynyl pyrenes. *Chem. Commun.* **2013,** *49*, 9257-9259.

176. Ucar, H.; Wagenknecht, H. A., DNA-templated control of chirality and efficient energy transport in supramolecular DNA architectures with aggregation-induced emission. *Chem. Sci.* **2021,** *12*, 10048-10053.

177. Avakyan, N.; Greschner, A. A.; Aldaye, F.; Serpell, C. J.; Toader, V.; Petitjean, A.; Sleiman, H. F., Reprogramming the assembly of unmodified DNA with a small molecule. *Nat. Chem.* **2016,** *8*, 368-376.

178. Lachance-Brais, C.; Hennecker, C. D.; Alenaizan, A.; Luo, X.; Toader, V.; Taing, M.; Sherrill, C. D.; Mittermaier, A. K.; Sleiman, H. F., Tuning DNA Supramolecular Polymers by the Addition of Small, Functionalized Nucleobase Mimics. *J. Am. Chem. Soc.* **2021,** *143*, 19824-19833.

179. Rizzuto, F. J.; Platnich, C. M.; Luo, X.; Shen, Y.; Dore, M. D.; Lachance-Brais, C.; Guarne, A.; Cosa, G.; Sleiman, H. F., A dissipative pathway for the structural evolution of DNA fibres. *Nat. Chem.* **2021,** *13*, 843-849.

180. Shimizu, T.; Masuda, M.; Minamikawa, H., Supramolecular nanotube architectures based on amphiphilic molecules. *Chem. Rev.* **2005,** *105*, 1401-1443.

181. Vybornyi, M.; Vyborna, Y.; Häner, R., DNA-inspired oligomers: from oligophosphates to functional materials. *Chem. Soc. Rev.* **2019,** *48*, 4347-4360.

182. Menacher, F.; Stepanenko, V.; Würthner, F.; Wagenknecht, H. A., Assembly of DNA triangles mediated by perylene bisimide caps. *Chem. Eur. J.* **2011,** *17*, 6683-6688.

183. Hariharan, M.; Zheng, Y.; Long, H.; Zeidan, T. A.; Schatz, G. C.; Vura-Weis, J.; Wasielewski, M. R.; Zuo, X.; Tiede, D. M.; Lewis, F. D., Hydrophobic dimerization and thermal dissociation of perylenediimide-linked DNA hairpins. *J. Am. Chem. Soc.* **2009,** *131*, 5920-5929.

184. Eryazici, I.; Yildirim, I.; Schatz, G. C.; Nguyen, S. T., Enhancing the melting properties of small molecule-DNA hybrids through designed hydrophobic interactions: an experimental-computational study. *J. Am. Chem. Soc.* **2012,** *134*, 7450-7458.

185. Madsen, M.; Gothelf, K. V., Chemistries for DNA Nanotechnology. *Chem. Rev.* **2019,** *119*, 6384-6458.

186. Chien, M. P.; Rush, A. M.; Thompson, M. P.; Gianneschi, N. C., Programmable shape-shifting micelles. *Angew. Chem., Int. Ed.* **2010,** *49*, 5076-5080.

187. Edwardson, T. G.; Carneiro, K. M.; Serpell, C. J.; Sleiman, H. F., An efficient and modular route to sequence-defined polymers appended to DNA. *Angew. Chem., Int. Ed.* **2014,** *53*, 4567-4571.

188. Bousmail, D.; Chidchob, P.; Sleiman, H. F., Cyanine-Mediated DNA Nanofiber Growth with Controlled Dimensionality. *J. Am. Chem. Soc.* **2018,** *140*, 9518-9530.

189. Rizzuto, F. J.; Dore, M. D.; Rafique, M. G.; Luo, X.; Sleiman, H. F., DNA Sequence and Length Dictate the Assembly of Nucleic Acid Block Copolymers. *J. Am. Chem. Soc.* **2022,** *144*, 12272-12279.

190. Chidchob, P.; Edwardson, T. G.; Serpell, C. J.; Sleiman, H. F., Synergy of Two Assembly Languages in DNA Nanostructures: Self-Assembly of Sequence-Defined Polymers on DNA Cages. *J. Am. Chem. Soc.* **2016,** *138*, 4416-4425.

191. List, J.; Weber, M.; Simmel, F. C., Hydrophobic actuation of a DNA origami bilayer structure. *Angew. Chem., Int. Ed.* **2014,** *53*, 4236-4239.

192. Albert, S. K.; Golla, M.; Krishnan, N.; Perumal, D.; Varghese, R., DNA-pi Amphiphiles: A Unique Building Block for the Crafting of DNA-Decorated Unilamellar Nanostructures. *Acc. Chem. Res.* **2020,** *53*, 2668-2679.

193. Kashida, H.; Hayashi, T.; Fujii, T.; Asanuma, H., A cationic dye triplet as a unique "glue" that can connect fully matched termini of DNA duplexes. *Chem. Eur. J.* **2011,** *17*, 2614-2622.

194. Wang, W.; Wan, W.; Zhou, H. H.; Niu, S.; Li, A. D., Alternating DNA and pi-conjugated sequences. Thermophilic foldable polymers. *J. Am. Chem. Soc.* **2003,** *125*, 5248-5249.

195. Malinovskii, V. L.; Samain, F.; Häner, R., Helical arrangement of interstrand stacked pyrenes in a DNA framework. *Angew. Chem., Int. Ed.* **2007,** *46*, 4464-4467.

196. Häner, R.; Garo, F.; Wenger, D.; Malinovskii, V. L., Oligopyrenotides: abiotic, polyanionic oligomers with nucleic acid-like structural properties. *J. Am. Chem. Soc.* **2010,** *132*, 7466-7471.

197. Nussbaumer, A. L.; Studer, D.; Malinovskii, V. L.; Häner, R., Amplification of chirality by supramolecular polymerization of pyrene oligomers. *Angew. Chem., Int. Ed.* **2011,** *50*, 5490-5494.

198. Vyborna, Y.; Vybornyi, M.; Rudnev, A. V.; Häner, R., DNA-Grafted Supramolecular Polymers: Helical Ribbon Structures Formed by Self-Assembly of Pyrene-DNA Chimeric Oligomers. *Angew. Chem., Int. Ed.* **2015,** *54*, 7934-7938.

199. Winiger, C. B.; Li, S.; Kumar, G. R.; Langenegger, S. M.; Häner, R., Long-distance electronic energy transfer in light-harvesting supramolecular polymers. *Angew. Chem., Int. Ed.* **2014,** *53*, 13609-13613.

200. Kownacki, M.; Langenegger, S. M.; Liu, S. X.; Häner, R., Integrating DNA Photonic Wires into Light-Harvesting Supramolecular Polymers. *Angew. Chem., Int. Ed.* **2019,** *58*, 751-755.

201. Bosch, C. D.; Langenegger, S. M.; Häner, R., Light-Harvesting Nanotubes Formed by Supramolecular Assembly of Aromatic Oligophosphates. *Angew. Chem., Int. Ed.* **2016,** *55*, 9961-9964.

202. Vybornyi, M.; Rudnev, A. V.; Langenegger, S. M.; Wandlowski, T.; Calzaferri, G.; Häner, R., Formation of two-dimensional supramolecular polymers by amphiphilic pyrene oligomers. *Angew. Chem., Int. Ed.* **2013,** *52*, 11488-11493.

203. Albert, S. K.; Thelu, H. V. P.; Golla, M.; Krishnan, N.; Chaudhary, S.; Varghese, R., Self-assembly of DNA-oligo(p-phenylene-ethynylene) hybrid amphiphiles into surface-engineered vesicles with enhanced emission. *Angew. Chem., Int. Ed.* **2014,** *53*, 8352-8357.

204. Albert, S. K.; Golla, M.; Thelu, H. V. P.; Krishnan, N.; Varghese, R., A pH-Responsive DNAsome from the Self-Assembly of DNA-Phenyleneethynylene Hybrid Amphiphile. *Chem. Eur. J.* **2017,** *23*, 8348-8352.

205. Bösch, C. D.; Jevric, J.; Bürki, N.; Probst, M.; Langenegger, S. M.; Häner, R., Supramolecular Assembly of DNA-Phenanthrene Conjugates into Vesicles with Light-Harvesting Properties. *Bioconjug. Chem.* **2018,** *29*, 1505-1509.

206. Rothenbühler, S.; Iacovache, I.; Langenegger, S. M.; Zuber, B.; Häner, R., Supramolecular assembly of DNA-constructed vesicles. *Nanoscale* **2020,** *12*, 21118-21123.

207. Rothenbühler, S.; Iacovache, I.; Langenegger, S. M.; Zuber, B.; Häner, R., Complex DNA Architectonics horizontal line Self-Assembly of Amphiphilic Oligonucleotides into Ribbons, Vesicles, and Asterosomes. *Bioconjug. Chem.* **2023,** *34*, 70-77.

208. Albert, S. K.; Sivakumar, I.; Golla, M.; Thelu, H. V. P.; Krishnan, N.; K, L. J.; Ashish; Varghese, R., DNA-Decorated Two-Dimensional Crystalline Nanosheets. *J. Am. Chem. Soc.* **2017,** *139*, 17799-17802.

209. Krishnan, N.; Golla, M.; Thelu, H. V. P.; Albert, S. K.; Atchimnaidu, S.; Perumal, D.; Varghese, R., Self-assembly of DNA-tetraphenylethylene amphiphiles into DNA-grafted nanosheets as a support for the immobilization of gold nanoparticles: a recyclable catalyst with enhanced activity. *Nanoscale* **2018,** *10*, 17174-17181.

210. Golla, M.; Albert, S. K.; Atchimnaidu, S.; Perumal, D.; Krishnan, N.; Varghese, R., DNA-Decorated, Helically Twisted Nanoribbons: A Scaffold for the Fabrication of One-Dimensional, Chiral, Plasmonic Nanostructures. *Angew. Chem., Int. Ed.* **2019,** *58*, 3865-3869.

211. Raiber, E.-A.; Hardisty, R.; van Delft, P.; Balasubramanian, S., Mapping and elucidating the function of modified bases in DNA. *Nat. Rev. Chem.* **2017,** *1*, 0069.

212. Lopez-Tena, M.; Chen, S. K.; Winssinger, N., Supernatural: Artificial Nucleobases and Backbones to Program Hybridization-Based Assemblies and Circuits. *Bioconjug. Chem.* **2023,** *34*, 111-123.

213. Epple, S.; El-Sagheer, A. H.; Brown, T., Artificial nucleic acid backbones and their applications in therapeutics, synthetic biology and biotechnology. *Emerg. Top. Life Sci.* **2021,** *5*, 691-697.

214. Duffy, K.; Arangundy-Franklin, S.; Holliger, P., Modified nucleic acids: replication, evolution, and next-generation therapeutics. *BMC Biol.* **2020,** *18*, 112.

215. Liczner, C.; Duke, K.; Juneau, G.; Egli, M.; Wilds, C. J., Beyond ribose and phosphate: Selected nucleic acid modifications for structure-function investigations and therapeutic applications. *Beilstein J. Org. Chem.* **2021,** *17*, 908-931.

216. Asanuma, H.; Kamiya, Y.; Kashida, H.; Murayama, K., Xeno nucleic acids (XNAs) having non-ribose scaffolds with unique supramolecular properties. *Chem. Commun.* **2022,** *58*, 3993-4004.

217. Malyshev, D. A.; Romesberg, F. E., The expanded genetic alphabet. *Angew. Chem., Int. Ed.* **2015,** *54*, 11930-11944.

218. Xu, W.; Chan, K. M.; Kool, E. T., Fluorescent nucleobases as tools for studying DNA and RNA. *Nat. Chem.* **2017,** *9*, 1043-1055.

219. Gorska, K.; Winssinger, N., Reactions templated by nucleic acids: more ways to translate oligonucleotide-based instructions into emerging function. *Angew. Chem., Int. Ed.* **2013,** *52*, 6820-6843.

220. Ayadi, L.; Galvanin, A.; Pichot, F.; Marchand, V.; Motorin, Y., RNA ribose methylation (2'-O-methylation): Occurrence, biosynthesis and biological functions. *Biochim. Biophys. Acta Gene Regul. Mech.* **2019,** *1862*, 253-269.

221. Elliott, B. A.; Ho, H. T.; Ranganathan, S. V.; Vangaveti, S.; Ilkayeva, O.; Abou Assi, H.; Choi, A. K.; Agris, P. F.; Holley, C. L., Modification of messenger RNA by 2'-O-methylation regulates gene expression in vivo. *Nat. Commun.* **2019,** *10*, 3401.

222. Cummins, L. L.; Owens, S. R.; Risen, L. M.; Lesnik, E. A.; Freier, S. M.; McGee, D.; Guinosso, C. J.; Cook, P. D., Characterization of fully 2'-modified oligoribonucleotide hetero- and homoduplex hybridization and nuclease sensitivity. *Nucleic Acids Res.* **1995,** *23*, 2019-2024.

223. Kawai, G.; Yamamoto, Y.; Kamimura, T.; Masegi, T.; Sekine, M.; Hata, T.; Iimori, T.; Watanabe, T.; Miyazawa, T.; Yokoyama, S., Conformational rigidity of specific pyrimidine residues in tRNA arises from posttranscriptional modifications that enhance steric interaction between the base and the 2'-hydroxyl group. *Biochemistry* **1992,** *31*, 1040-1046.

224. Bondensgaard, K.; Petersen, M.; Singh, S. K.; Rajwanshi, V. K.; Kumar, R.; Wengel, J.; Jacobsen, J. P., Structural studies of LNA:RNA duplexes by NMR: conformations and implications for RNase H activity. *Chem. Eur. J.* **2000,** *6*, 2687-2695.

225. Obika, S.; Nanbu, D.; Hari, Y.; Morio, K.-i.; In, Y.; Ishida, T.; Imanishi, T., Synthesis of 2'-O,4'-C-methyleneuridine and -cytidine. Novel bicyclic nucleosides having a fixed C3, -endo sugar puckering. *Tetrahedron Lett.* **1997,** *38*, 8735-8738.

226. Crooke, S. T.; Liang, X. H.; Baker, B. F.; Crooke, R. M., Antisense technology: A review. *J. Biol. Chem.* **2021,** *296*, 100416.

227. Nguyen, K.; Malik, T. N.; Chaput, J. C., Chemical evolution of an autonomous DNAzyme with allele-specific gene silencing activity. *Nat. Commun.* **2023,** *14*, 2413.

228. Pallan, P. S.; Greene, E. M.; Jicman, P. A.; Pandey, R. K.; Manoharan, M.; Rozners, E.; Egli, M., Unexpected origins of the enhanced pairing affinity of 2'-fluoro-modified RNA. *Nucleic Acids Res.* **2011,** *39*, 3482-3495.

229. Denisov, A. Y.; Noronha, A. M.; Wilds, C. J.; Trempe, J. F.; Pon, R. T.; Gehring, K.; Damha, M. J., Solution structure of an arabinonucleic acid (ANA)/RNA duplex in a chimeric hairpin: comparison with 2'-fluoro-ANA/RNA and DNA/RNA hybrids. *Nucleic Acids Res.* **2001,** *29*, 4284-4293.

230. Kreutz, C.; Kahlig, H.; Konrat, R.; Micura, R., Ribose 2'-F labeling: a simple tool for the characterization of RNA secondary structure equilibria by 19F NMR spectroscopy. *J. Am. Chem. Soc.* **2005,** *127*, 11558-11559.

231. Du, Q.; Carrasco, N.; Teplova, M.; Wilds, C. J.; Egli, M.; Huang, Z., Internal derivatization of oligonucleotides with selenium for X-ray crystallography using MAD. *J. Am. Chem. Soc.* **2002,** *124*, 24-25.

232. Eschenmoser, A., Chemical etiology of nucleic acid structure. *Science* **1999,** *284*, 2118-2124.

233. Hunziker, J.; Roth, H.-J.; Böhringer, M.; Giger, A.; Diederichsen, U.; Göbel, M.; Krishnan, R.; Jaun, B.; Leumann, C.; Eschenmoser, A., Warum pentose-und nicht hexose-nucleinsäuren? Teil III. Oligo(2',3'-dideoxy-β-D-glucopyranosyl) nucleotide ('homo-DNS'): Paarungesigenschaften. *Helv. Chim. Acta* **1993,** *76*, 259-352.

234. Aerschot Van, A.; Verheggen, I.; Hendrix, C.; Herdewijn, P., 1,5-Anhydrohexitol Nucleic Acids, a New Promising Antisense Construct. *Angew. Chem., Int. Ed.* **1995,** *34*, 1338-1339.

235. Hendrix, C.; Rosemeyer, H.; De Bouvere, B.; Van Aerschot, A.; Seela, F.; Herdewijn, P., 1',5'-Anhydrohexitol Oligonucleotides: Hybridisation and Strand Displacement with Oligoribonucleotides, Interaction with RNase H and HIV Reverse Transcriptase. *Chem. Eur. J.* **1997,** *3*, 1513-1520.

236. Wang, J.; Verbeure, B.; Luyten, I.; Lescrinier, E.; Froeyen, M.; Hendrix, C.; Rosemeyer, H.; Seela, F.; Van Aerschot, A.; Herdewijn, P., Cyclohexene Nucleic Acids (CeNA): Serum Stable

Oligonucleotides that Activate RNase H and Increase Duplex Stability with Complementary RNA. *J. Am. Chem. Soc.* **2000,** *122*, 8595-8602.

237. Verbeure, B.; Lescrinier, E.; Wang, J.; Herdewijn, P., RNase H mediated cleavage of RNA by cyclohexene nucleic acid (CeNA). *Nucleic Acids Res.* **2001,** *29*, 4941-4947.

238. Schöning, K.; Scholz, P.; Guntha, S.; Wu, X.; Krishnamurthy, R.; Eschenmoser, A., Chemical etiology of nucleic acid structure: the alpha-threofuranosyl-(3'-->2') oligonucleotide system. *Science* **2000,** *290*, 1347-1351.

239. Steffens, R.; Leumann, C. J., Tricyclo-DNA: A Phosphodiester-Backbone Based DNA Analog Exhibiting Strong Complementary Base-Pairing Properties. *J. Am. Chem. Soc.* **1997,** *119*, 11548-11549.

240. Murayama, K.; Asanuma, H., Design and Hybridization Properties of Acyclic Xeno Nucleic Acid Oligomers. *ChemBioChem* **2021,** *22*, 2507-2515.

241. Schneider, K. C.; Benner, S. A., Oligonucleotides containing flexible nucleoside analogs. *J. Am. Chem. Soc.* **1990,** *112*, 453-455.

242. Merle, Y.; Bonneil, E.; Merle, L.; Sagi, J.; Szemzo, A., Acyclic oligonucleotide analogues. *Int. J. Biol. Macromol.* **1995,** *17*, 239-246.

243. Campbell, M. A.; Wengel, J., Locked vs. unlocked nucleic acids (LNA vs. UNA): contrasting structures work towards common therapeutic goals. *Chem. Soc. Rev.* **2011,** *40*, 5680-5689.

244. Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O., Sequence-selective recognition of DNA by strand displacement with a thymine-substituted polyamide. *Science* **1991,** *254*, 1497-1500.

245. 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., PNA hybridizes to complementary oligonucleotides obeying the Watson-Crick hydrogen-bonding rules. *Nature* **1993,** *365*, 566-568.

246. Rasmussen, H.; Kastrup, J. S.; Nielsen, J. N.; Nielsen, J. M.; Nielsen, P. E., Crystal structure of a peptide nucleic acid (PNA) duplex at 1.7 A resolution. *Nat. Struct. Biol.* **1997,** *4*, 98-101.

247. Diederichsen, U., Pairing Properties of Alanyl Peptide Nucleic Acids Containing an Amino Acid Backbone with Alternating Configuration. *Angew. Chem., Int. Ed.* **1996,** *35*, 445-448.

248. Saarbach, J.; Sabale, P. M.; Winssinger, N., Peptide nucleic acid (PNA) and its applications in chemical biology, diagnostics, and therapeutics. *Curr. Opin. Chem. Biol.* **2019,** *52*, 112-124.

249. Meggers, E.; Zhang, L., Synthesis and properties of the simplified nucleic acid glycol nucleic acid. *Acc. Chem. Res.* **2010,** *43*, 1092-1102.

250. Ueda, N.; Kawabata, T.; Takemoto, K., Synthesis of N-(2,3-dihydroxypropyl) derivatives of nucleic bases. *J. Heterocycl. Chem.* **1971,** *8*, 827-829.

251. Zhang, L.; Peritz, A.; Meggers, E., A simple glycol nucleic acid. *J. Am. Chem. Soc.* **2005,** *127*, 4174-4175.

252. Schlegel, M. K.; Xie, X.; Zhang, L.; Meggers, E., Insight into the high duplex stability of the simplified nucleic acid GNA. *Angew. Chem., Int. Ed.* **2009,** *48*, 960-963.

253. Schlegel, M. K.; Essen, L.-O.; Meggers, E., Atomic resolution duplex structure of the simplified nucleic acid GNA. *Chem. Commun.* **2010,** *46*, 1094-1096.

254. Schlegel, M. K.; Essen, L. O.; Meggers, E., Duplex structure of a minimal nucleic acid. *J. Am. Chem. Soc.* **2008,** *130*, 8158-8159.

255. Schlegel, M. K.; Foster, D. J.; Kel'in, A. V.; Zlatev, I.; Bisbe, A.; Jayaraman, M.; Lackey, J. G.; Rajeev, K. G.; Charisse, K.; Harp, J.; Pallan, P. S.; Maier, M. A.; Egli, M.; Manoharan, M., Chirality Dependent Potency Enhancement and Structural Impact of Glycol Nucleic Acid Modification on siRNA. *J. Am. Chem. Soc.* **2017,** *139*, 8537-8546.

256. Baumstark, D.; Wagenknecht, H. A., Perylene bisimide dimers as fluorescent "glue" for DNA and for base-mismatch detection. *Angew. Chem., Int. Ed.* **2008,** *47*, 2612-2614.

257. Zhou, H.; Johnson, A. T.; Wiest, O.; Zhang, L. L., Incorporation of porphyrin acetylides into duplexes of the simplified nucleic acid GNA. *Org. Biomol. Chem.* **2011,** *9*, 2840-2849.

258. Kumar, V.; Gore, K. R.; Pradeepkumar, P. I.; Kesavan, V., Design, synthesis, biophysical and primer extension studies of novel acyclic butyl nucleic acid (BuNA). *Org. Biomol. Chem.* **2013,** *11*, 5853-5865.

259. Kumar, V.; Kesavan, V., Acyclic butyl nucleic acid (BuNA): a novel scaffold for A-switch. *RSC Adv.* **2013,** *3*, 19330-19340.

260. Rich, A., *Horizons in Biochemistry*. Academic Press, New York: 1962.

261. Switzer, C.; Moroney, S. E.; Benner, S. A., Enzymatic incorporation of a new base pair into DNA and RNA. *J. Am. Chem. Soc.* **1989,** *111*, 8322-8323.

262. Switzer, C. Y.; Moroney, S. E.; Benner, S. A., Enzymatic recognition of the base pair between isocytidine and isoguanosine. *Biochemistry* **1993,** *32*, 10489-10496.

263. Yang, Z.; Sismour, A. M.; Sheng, P.; Puskar, N. L.; Benner, S. A., Enzymatic incorporation of a third nucleobase pair. *Nucleic Acids Res.* **2007,** *35*, 4238-4249.

264. Hutter, D.; Benner, S. A., Expanding the genetic alphabet: non-epimerizing nucleoside with the pyDDA hydrogen-bonding pattern. *J. Org. Chem.* **2003,** *68*, 9839-9842.

265. Hoshika, S.; Leal, N. A.; Kim, M. J.; Kim, M. S.; Karalkar, N. B.; Kim, H. J.; Bates, A. M.; Watkins, N. E., Jr.; SantaLucia, H. A.; Meyer, A. J.; DasGupta, S.; Piccirilli, J. A.; Ellington, A. D.; SantaLucia, J., Jr.; Georgiadis, M. M.; Benner, S. A., Hachimoji DNA and RNA: A genetic system with eight building blocks. *Science* **2019,** *363*, 884-887.

266. Hirao, I.; Ohtsuki, T.; Fujiwara, T.; Mitsui, T.; Yokogawa, T.; Okuni, T.; Nakayama, H.; Takio, K.; Yabuki, T.; Kigawa, T.; Kodama, K.; Yokogawa, T.; Nishikawa, K.; Yokoyama, S., An

unnatural base pair for incorporating amino acid analogs into proteins. *Nat. Biotechnol.* **2002,** *20*, 177-182.

267. Hirao, I.; Harada, Y.; Kimoto, M.; Mitsui, T.; Fujiwara, T.; Yokoyama, S., A two-unnatural-base-pair system toward the expansion of the genetic code. *J. Am. Chem. Soc.* **2004,** *126*, 13298-13305.

268. Morihiro, K.; Moriyama, Y.; Nemoto, Y.; Osumi, H.; Okamoto, A., anti-syn Unnatural Base Pair Enables Alphabet-Expanded DNA Self-Assembly. *J. Am. Chem. Soc.* **2021,** *143*, 14207-14217.

269. Okamura, H.; Trinh, G. H.; Dong, Z.; Masaki, Y.; Seio, K.; Nagatsugi, F., Selective and stable base pairing by alkynylated nucleosides featuring a spatially-separated recognition interface. *Nucleic Acids Res.* **2022,** *50*, 3042-3055.

270. Moran, S.; Ren, R. X.; Rumney, S.; Kool, E. T., Difluorotoluene, a Nonpolar Isostere for Thymine, Codes Specifically and Efficiently for Adenine in DNA Replication. *J. Am. Chem. Soc.* **1997,** *119*, 2056-2057.

271. Morales, J. C.; Kool, E. T., Minor Groove Interactions between Polymerase and DNA: More Essential to Replication than Watson-Crick Hydrogen Bonds? *J. Am. Chem. Soc.* **1999,** *121*, 2323-2324.

272. Kimoto, M.; Hirao, I., Genetic alphabet expansion technology by creating unnatural base pairs. *Chem. Soc. Rev.* **2020,** *49*, 7602-7626.

273. Mitsui, T.; Kitamura, A.; Kimoto, M.; To, T.; Sato, A.; Hirao, I.; Yokoyama, S., An unnatural hydrophobic base pair with shape complementarity between pyrrole-2-carbaldehyde and 9-methylimidazo[(4,5)-b]pyridine. *J. Am. Chem. Soc.* **2003,** *125*, 5298-5307.

274. Hirao, I.; Kimoto, M.; Mitsui, T.; Fujiwara, T.; Kawai, R.; Sato, A.; Harada, Y.; Yokoyama, S., An unnatural hydrophobic base pair system: site-specific incorporation of nucleotide analogs into DNA and RNA. *Nat. Methods* **2006,** *3*, 729-735.

275. Kimoto, M.; Kawai, R.; Mitsui, T.; Yokoyama, S.; Hirao, I., An unnatural base pair system for efficient PCR amplification and functionalization of DNA molecules. *Nucleic Acids Res.* **2009,** *37*, e14.

276. Yamashige, R.; Kimoto, M.; Takezawa, Y.; Sato, A.; Mitsui, T.; Yokoyama, S.; Hirao, I., Highly specific unnatural base pair systems as a third base pair for PCR amplification. *Nucleic Acids Res.* **2012,** *40*, 2793-2806.

277. Hyun Lee, K.; Kimoto, M.; Kawai, G.; Okamoto, I.; Fin, A.; Hirao, I., Dye-Conjugated Spinach RNA by Genetic Alphabet Expansion. *Chem. Eur. J.* **2022,** *28*, e202104396.

278. Betz, K.; Kimoto, M.; Diederichs, K.; Hirao, I.; Marx, A., Structural Basis for Expansion of the Genetic Alphabet with an Artificial Nucleobase Pair. *Angew. Chem., Int. Ed.* **2017,** *56*, 12000-12003.

279. McMinn, D. L.; Ogawa, A. K.; Wu, Y.; Liu, J.; Schultz, P. G.; Romesberg, F. E., Efforts toward Expansion of the Genetic Alphabet: DNA Polymerase Recognition of a Highly Stable, Self-Pairing Hydrophobic Base. *J. Am. Chem. Soc.* **1999,** *121*, 11585-11586.

280. Matsuda, S.; Fillo, J. D.; Henry, A. A.; Rai, P.; Wilkens, S. J.; Dwyer, T. J.; Geierstanger, B. H.; Wemmer, D. E.; Schultz, P. G.; Spraggon, G.; Romesberg, F. E., Efforts toward Expansion of the Genetic Alphabet: Structure and Replication of Unnatural Base Pairs. *J. Am. Chem. Soc.* **2007,** *129*, 10466-10473.

281. Marx, A.; Betz, K., The Structural Basis for Processing of Unnatural Base Pairs by DNA Polymerases. *Chem. Eur. J.* **2020,** *26*, 3446-3463.

282. Leconte, A. M.; Hwang, G. T.; Matsuda, S.; Capek, P.; Hari, Y.; Romesberg, F. E., Discovery, characterization, and optimization of an unnatural base pair for expansion of the genetic alphabet. *J. Am. Chem. Soc.* **2008,** *130*, 2336-2243.

283. Malyshev, D. A.; Seo, Y. J.; Ordoukhanian, P.; Romesberg, F. E., PCR with an expanded genetic alphabet. *J. Am. Chem. Soc.* **2009,** *131*, 14620-14621.

284. Malyshev, D. A.; Dhami, K.; Lavergne, T.; Chen, T.; Dai, N.; Foster, J. M.; Correa, I. R., Jr.; Romesberg, F. E., A semi-synthetic organism with an expanded genetic alphabet. *Nature* **2014,** *509*, 385-388.

285. Li, L.; Degardin, M.; Lavergne, T.; Malyshev, D. A.; Dhami, K.; Ordoukhanian, P.; Romesberg, F. E., Natural-like Replication of an Unnatural Base Pair for the Expansion of the Genetic Alphabet and Biotechnology Applications. *J. Am. Chem. Soc.* **2014,** *136*, 826-829.

286. Zhu, W.; Wang, H.; Li, X.; Tie, W.; Huo, B.; Zhu, A.; Li, L., Amplification, Enrichment, and Sequencing of Mutagenic Methylated DNA Adduct through Specifically Pairing with Unnatural Nucleobases. *J. Am. Chem. Soc.* **2022,** *144*, 20165-20170.

287. Zhang, Y.; Ptacin, J. L.; Fischer, E. C.; Aerni, H. R.; Caffaro, C. E.; San Jose, K.; Feldman, A. W.; Turner, C. R.; Romesberg, F. E., A semi-synthetic organism that stores and retrieves increased genetic information. *Nature* **2017,** *551*, 644-647.

288. Depmeier, H.; Hoffmann, E.; Bornewasser, L.; Kath-Schorr, S., Strategies for Covalent Labeling of Long RNAs. *ChemBioChem* **2021,** *22*, 2826-2847.

289. Meyer, E. A.; Castellano, R. K.; Diederich, F., Interactions with aromatic rings in chemical and biological recognition. *Angew. Chem., Int. Ed.* **2003,** *42*, 1210-1250.

290. Salonen, L. M.; Ellermann, M.; Diederich, F., Aromatic rings in chemical and biological recognition: energetics and structures. *Angew. Chem., Int. Ed.* **2011,** *50*, 4808-4842.

291. Chipot, C.; Jaffe, R.; Maigret, B.; Pearlman, D. A.; Kollman, P. A., Benzene Dimer: A Good Model for π−π Interactions in Proteins? A Comparison between the Benzene and the Toluene Dimers in the Gas Phase and in an Aqueous Solution. *J. Am. Chem. Soc.* **1996,** *118*, 11217-11224.

292. Tsuzuki, S.; Mikami, M.; Yamada, S., Origin of attraction, magnitude, and directionality of interactions in benzene complexes with pyridinium cations. *J. Am. Chem. Soc.* **2007,** *129*, 8656-8662.

293. Tsuzuki, S.; Honda, K.; Uchimaru, T.; Mikami, M.; Tanabe, K., Origin of Attraction and Directionality of the π/π Interaction: Model Chemistry Calculations of Benzene Dimer Interaction. *J. Am. Chem. Soc.* **2002,** *124*, 104-112.

294. Hunter, C. A.; Lawson, K. R.; Perkins, J.; Urch, C. J., Aromatic interactions. *J. Chem. Soc., Perkin Trans. 2* **2001**, 651-669.

295. Hunter, C. A.; Sanders, J. K. M., The nature of .pi.-.pi. interactions. *J. Am. Chem. Soc.* **1990,** *112*, 5525-5534.

296. Lee, G. Y.; Hu, E.; Rheingold, A. L.; Houk, K. N.; Sletten, E. M., Arene-Perfluoroarene Interactions in Solution. *J. Org. Chem.* **2021,** *86*, 8425-8436.

297. Kool, E. T., Replacing the nucleobases in DNA with designer molecules. *Acc. Chem. Res.* **2002,** *35*, 936-943.

298. Matray, T. J.; Kool, E. T., A specific partner for abasic damage in DNA. *Nature* **1999,** *399*, 704-708.

299. Matray, T. J.; Kool, E. T., Selective and Stable DNA Base Pairing without Hydrogen Bonds. *J. Am. Chem. Soc.* **1998,** *120*, 6191-6192.

300. Nielsen, C. B.; Petersen, M.; Pedersen, E. B.; Hansen, P. E.; Christensen, U. B., NMR structure determination of a modified DNA oligonucleotide containing a new intercalating nucleic acid. *Bioconjug. Chem.* **2004,** *15*, 260-269.

301. Brotschi, C.; Leumann, C. J., DNA with hydrophobic base substitutes: a stable, zipperlike recognition motif based on interstrand-stacking interactions. *Angew. Chem., Int. Ed.* **2003,** *42*, 1655-1658.

302. Roethlisberger, P.; Istrate, A.; Marcaida Lopez, M. J.; Visini, R.; Stocker, A.; Reymond, J. L.; Leumann, C. J., X-ray structure of a lectin-bound DNA duplex containing an unnatural phenanthrenyl pair. *Chem. Commun.* **2016,** *52*, 4749-4752.

303. Grigorenko, N. A.; Leumann, C. J., 2-Phenanthrenyl-DNA: synthesis, pairing, and fluorescence properties. *Chem. Eur. J.* **2009,** *15*, 639-645.

304. Hainke, S.; Seitz, O., Binaphthyl-DNA: stacking and fluorescence of a nonplanar aromatic base surrogate in DNA. *Angew. Chem., Int. Ed.* **2009,** *48*, 8250-8253.

305. Zahn, A.; Leumann, C. J., Recognition properties of donor- and acceptor-modified biphenyl-DNA. *Chem. Eur. J.* **2008,** *14*, 1087-1094.

306. Johar, Z.; Zahn, A.; Leumann, C. J.; Jaun, B., Solution structure of a DNA duplex containing a biphenyl pair. *Chem. Eur. J.* **2008,** *14*, 1080-1086.

307. Rothlisberger, P.; Levi-Acobas, F.; Leumann, C. J.; Hollenstein, M., Enzymatic synthesis of biphenyl-DNA oligonucleotides. *Bioorg. Med. Chem.* **2020,** *28*, 115487.

308. Winiger, C. B.; Langenegger, S. M.; Khorev, O.; Häner, R., Influence of perylenediimide-pyrene supramolecular interactions on the stability of DNA-based hybrids: Importance of electrostatic complementarity. *Beilstein J. Org. Chem.* **2014,** *10*, 1589-1595.

309. Doi, T.; Sakakibara, T.; Kashida, H.; Araki, Y.; Wada, T.; Asanuma, H., Hetero-Selective DNA-Like Duplex Stabilized by Donor-Acceptor Interactions. *Chem. Eur. J.* **2015,** *21*, 15974-15980.

310. Ikkanda, B. A.; Samuel, S. A.; Iverson, B. L., NDI and DAN DNA: nucleic acid-directed assembly of NDI and DAN. *J. Org. Chem.* **2014,** *79*, 2029-2037.

311. Perez de Carvasal, K.; Riccardi, C.; Russo Krauss, I.; Cavasso, D.; Vasseur, J. J.; Smietana, M.; Morvan, F.; Montesarchio, D., Charge-Transfer Interactions Stabilize G-Quadruplex-Forming Thrombin Binding Aptamers and Can Improve Their Anticoagulant Activity. *Int. J. Mol. Sci.* **2021,** *22*, 9510.

312. Perez de Carvasal, K.; Nicollet, L.; Smietana, M.; Morvan, F., Stabilization of DNA Duplexes and Hairpins by Charge-Transfer Interactions Using DAN:NDI Pairs. *Langmuir* **2023,** *39*, 7418-7425.

313. Patrick, C. R.; Prosser, G. S., A Molecular Complex of Benzene and Hexafluorobenzene. *Nature* **1960,** *187*, 1021-1021.

314. Tsuzuki, S.; Uchimaru, T.; Mikami, M., Intermolecular Interaction between Hexafluorobenzene and Benzene: Ab Initio Calculations Including CCSD(T) Level Electron Correlation Correction. *J. Phys. Chem. A* **2006,** *110*, 2027-2033.

315. Carter-Fenk, K.; Herbert, J. M., Electrostatics does not dictate the slip-stacked arrangement of aromatic π–π interactions. *Chem. Sci.* **2020,** *11*, 6758-6765.

316. Liu, B.; Gao, J.; Hao, A.; Xing, P., Arene-Perfluoroarene Force Driven Sublimation-Removable Chiral Coassemblies. *Angew. Chem., Int. Ed.* **2023,** *62*, e202305135.

317. Sharber, S. A.; Baral, R. N.; Frausto, F.; Haas, T. E.; Müller, P.; Thomas III, S. W., Substituent Effects That Control Conjugated Oligomer Conformation through Non-covalent Interactions. *J. Am. Chem. Soc.* **2017,** *139*, 5164-5174.

318. Lu, J.; ten Brummelhuis, N.; Weck, M., Intramolecular folding of triblock copolymers via quadrupole interactions between poly(styrene) and poly(pentafluorostyrene) blocks. *Chem. Commun.* **2014,** *50*, 6225-6227.

319. Huang, Z.; Chen, X.; O'Neill, S. J. K.; Wu, G.; Whitaker, D. J.; Li, J.; McCune, J. A.; Scherman, O. A., Highly compressible glass-like supramolecular polymer networks. *Nat. Mater.* **2022,** *21*, 103-109.

320. Watt, S. W.; Dai, C.; Scott, A. J.; Burke, J. M.; Thomas, R. L.; Collings, J. C.; Viney, C.; Clegg, W.; Marder, T. B., Structure and phase behavior of a 2:1 complex between arene- and fluoroarene-based conjugated rigid rods. *Angew. Chem., Int. Ed.* **2004,** *43*, 3061-3063.

321. Kissel, P.; Murray, D. J.; Wulftange, W. J.; Catalano, V. J.; King, B. T., A nanoporous two-dimensional polymer by single-crystal-to-single-crystal photopolymerization. *Nat. Chem.* **2014,** *6*, 774-778.

322. Collings, J. C.; Smith, P. S.; Yufit, D. S.; Batsanov, A. S.; Howard, J. A. K.; Marder, T. B., Arene–perfluoroarene interactions in crystal engineering. Part 10. Crystal structures of 1:1 complexes of octafluoronaphthalene with biphenyl and biphenylene. *CrystEngComm* **2004,** *6*, 25-28.

323. Yamada, S.; Mitsuda, A.; Miyano, K.; Tanaka, T.; Morita, M.; Agou, T.; Kubota, T.; Konno, T., Development of Novel Solid-State Light-Emitting Materials Based on Pentafluorinated Tolane Fluorophores. *ACS Omega* **2018,** *3*, 9105-9113.

324. Coates, G. W.; Dunn, A. R.; Henling, L. M.; Dougherty, D. A.; Grubbs, R. H., Phenyl–Perfluorophenyl Stacking Interactions: A New Strategy for Supermolecule Construction. *Angew. Chem., Int. Ed.* **1997,** *36*, 248-251.

325. Zheng, H.; Gao, J., Highly specific heterodimerization mediated by quadrupole interactions. *Angew. Chem., Int. Ed.* **2010,** *49*, 8635-8639.

326. Pace, C. J.; Zheng, H.; Mylvaganam, R.; Kim, D.; Gao, J., Stacked fluoroaromatics as supramolecular synthons for programming protein dimerization specificity. *Angew. Chem., Int. Ed.* **2012,** *51*, 103-107.

327. Zhang, C.; Welborn, M.; Zhu, T.; Yang, N. J.; Santos, M. S.; Van Voorhis, T.; Pentelute, B. L., Pi-Clamp-mediated cysteine conjugation. *Nat. Chem.* **2016,** *8*, 120-128.

328. Chen, X.; Huang, Z.; Sala, R. L.; McLean, A. M.; Wu, G.; Sokołowski, K.; King, K.; McCune, J. A.; Scherman, O. A., On-Resin Recognition of Aromatic Oligopeptides and Proteins through Host-Enhanced Heterodimerization. *J. Am. Chem. Soc.* **2022,** *144*, 8474-8479.

329. Mathis, G.; Hunziker, J., Towards a DNA-like duplex without hydrogen-bonded base pairs. *Angew. Chem., Int. Ed.* **2002,** *41*, 3203-3205.

330. Lai, J. S.; Kool, E. T., Selective pairing of polyfluorinated DNA bases. *J. Am. Chem. Soc.* **2004,** *126*, 3040-3041.

331. Zahn, A.; Brotschi, C.; Leumann, C. J., pentafluorophenyl-phenyl interactions in biphenyl-DNA. *Chem. Eur. J.* **2005,** *11*, 2125-2129.

332. Katz, S., The reversible reaction of Hg (II) and double-stranded polynucleotides. A step-function theory and its significance. *Biochim. Biophys. Acta* **1963,** *68*, 240-253.

333. Miyake, Y.; Togashi, H.; Tashiro, M.; Yamaguchi, H.; Oda, S.; Kudo, M.; Tanaka, Y.; Kondo, Y.; Sawa, R.; Fujimoto, T.; Machinami, T.; Ono, A., MercuryII-mediated formation of thymine-HgII-thymine base pairs in DNA duplexes. *J. Am. Chem. Soc.* **2006,** *128*, 2172-2173.

334. Ono, A.; Cao, S.; Togashi, H.; Tashiro, M.; Fujimoto, T.; Machinami, T.; Oda, S.; Miyake, Y.; Okamoto, I.; Tanaka, Y., Specific interactions between silver(I) ions and cytosine-cytosine pairs in DNA duplexes. *Chem. Commun.* **2008**, 4825-4827.

335. Schmidt, O. P.; Benz, A. S.; Mata, G.; Luedtke, N. W., HgII binds to C-T mismatches with high affinity. *Nucleic Acids Res.* **2018,** *46*, 6470-6479.

336. Funai, T.; Aotani, M.; Kiriu, R.; Nakamura, J.; Miyazaki, Y.; Nakagawa, O.; Wada, S. I.; Torigoe, H.; Ono, A.; Urata, H., Silver(I)-Ion-Mediated Cytosine-Containing Base Pairs: Metal

Ion Specificity for Duplex Stabilization and Susceptibility toward DNA Polymerases. *ChemBioChem* **2020,** *21*, 517-522.

337. Kondo, J.; Tada, Y.; Dairaku, T.; Hattori, Y.; Saneyoshi, H.; Ono, A.; Tanaka, Y., A metallo-DNA nanowire with uninterrupted one-dimensional silver array. *Nat. Chem.* **2017,** *9*, 956-960.

338. Clever, G. H.; Kaul, C.; Carell, T., DNA--metal base pairs. *Angew. Chem., Int. Ed.* **2007,** *46*, 6226-6236.

339. Jash, B.; Müller, J., Metal-Mediated Base Pairs: From Characterization to Application. *Chemistry* **2017,** *23*, 17166-17178.

340. Takezawa, Y.; Suzuki, A.; Nakaya, M.; Nishiyama, K.; Shionoya, M., Metal-Dependent DNA Base Pairing of 5-Carboxyuracil with Itself and All Four Canonical Nucleobases. *J. Am. Chem. Soc.* **2020,** *142*, 21640-21644.

341. Sun, Q.; Xie, X.; Song, Y.; Sun, L., A review on silver-mediated DNA base pairs: methodology and application. *Biomater. Res.* **2022,** *26*, 9.

342. Wang, L. L.; Zhang, Q. L.; Wang, Y.; Liu, Y.; Lin, J.; Xie, F.; Xu, L., Controllable DNA strand displacement by independent metal-ligand complexation. *Chem. Sci.* **2021,** *12*, 8698-8705.

343. Wachowius, F.; Höbartner, C., Chemical RNA modifications for studies of RNA structure and dynamics. *ChemBioChem* **2010,** *11*, 469-480.

344. Guo, F.; Li, Q.; Zhou, C., Synthesis and biological applications of fluoro-modified nucleic acids. *Org. Biomol. Chem.* **2017,** *15*, 9552-9565.

345. Nelissen, F. H. T.; Tessari, M.; Wijmenga, S. S.; Heus, H. A., Stable isotope labeling methods for DNA. *Prog. Nucl. Magn. Reson. Spectrosc.* **2016,** *96*, 89-108.

346. Shelke, S. A.; Sigurdsson, S. T., Site-Directed Spin Labelling of Nucleic Acids. *Eur. J. Org. Chem.* **2012,** *2012*, 2291-2301.

347. Kodr, D.; Yenice, C. P.; Simonova, A.; Saftic, D. P.; Pohl, R.; Sykorova, V.; Ortiz, M.; Havran, L.; Fojta, M.; Lesnikowski, Z. J.; O'Sullivan, C. K.; Hocek, M., Carborane- or Metallacarborane-Linked Nucleotides for Redox Labeling. Orthogonal Multipotential Coding of all Four DNA Bases for Electrochemical Analysis and Sequencing. *J. Am. Chem. Soc.* **2021,** *143*, 7124-7134.

348. Fantoni, N. Z.; El-Sagheer, A. H.; Brown, T., A Hitchhiker's Guide to Click-Chemistry with Nucleic Acids. *Chem. Rev.* **2021,** *121*, 7122-7154.

349. Tretyakova, N. Y.; Groehler, A. t.; Ji, S., DNA-Protein Cross-Links: Formation, Structural Identities, and Biological Outcomes. *Acc. Chem. Res.* **2015,** *48*, 1631-1644.

350. Dziuba, D.; Didier, P.; Ciaco, S.; Barth, A.; Seidel, C. A. M.; Mely, Y., Fundamental photophysics of isomorphic and expanded fluorescent nucleoside analogues. *Chem. Soc. Rev.* **2021,** *50*, 7062-7107.

351. Hovelmann, F.; Seitz, O., DNA Stains as Surrogate Nucleobases in Fluorogenic Hybridization Probes. *Acc. Chem. Res.* **2016,** *49*, 714-723.

352. Wang, K.; Tang, Z.; Yang, C. J.; Kim, Y.; Fang, X.; Li, W.; Wu, Y.; Medley, C. D.; Cao, Z.; Li, J.; Colon, P.; Lin, H.; Tan, W., Molecular engineering of DNA: molecular beacons. *Angew. Chem., Int. Ed.* **2009,** *48*, 856-870.

353. Häner, R.; Biner, S. M.; Langenegger, S. M.; Meng, T.; Malinovskii, V. L., A highly sensitive, excimer-controlled molecular beacon. *Angew. Chem., Int. Ed.* **2010,** *49*, 1227-1230.

354. Huang, J.; Wu, Y.; Chen, Y.; Zhu, Z.; Yang, X.; Yang, C. J.; Wang, K.; Tan, W., Pyrene-excimer probes based on the hybridization chain reaction for the detection of nucleic acids in complex biological fluids. *Angew. Chem., Int. Ed.* **2011,** *50*, 401-404.

355. Prinzen, A. L.; Saliba, D.; Hennecker, C.; Trinh, T.; Mittermaier, A.; Sleiman, H. F., Amplified Self-Immolative Release of Small Molecules by Spatial Isolation of Reactive Groups on DNA-Minimal Architectures. *Angew. Chem., Int. Ed.* **2020,** *59*, 12900-12908.

356. Okamoto, A.; Tanaka, K.; Fukuta, T.; Saito, I., Design of base-discriminating fluorescent nucleoside and its application to t/c SNP typing. *J. Am. Chem. Soc.* **2003,** *125*, 9296-9297.

357. Karimi, A.; Börner, R.; Mata, G.; Luedtke, N. W., A Highly Fluorescent Nucleobase Molecular Rotor. *J. Am. Chem. Soc.* **2020,** *142*, 14422-14426.

358. Karimi, A.; Wang, K.; Basran, K.; Copp, W.; Luedtke, N. W., A Bright and Ionizable Cytosine Mimic for i-Motif Structures. *Bioconjug. Chem.* **2023,** *34*, 972-976.

359. Khatik, S. Y.; Sudhakar, S.; Mishra, S.; Kalia, J.; Pradeepkumar, P. I.; Srivatsan, S. G., Probing juxtaposed G-quadruplex and hairpin motifs using a responsive nucleoside probe: a unique scaffold for chemotherapy. *Chem. Sci.* **2023,** *14*, 5627-5637.

360. Khatik, S. Y.; Srivatsan, S. G., Environment-Sensitive Nucleoside Probe Unravels the Complex Structural Dynamics of i-Motif DNAs. *Bioconjug. Chem.* **2022,** *33*, 1515-1526.

361. Ghosh, P.; Kropp, H. M.; Betz, K.; Ludmann, S.; Diederichs, K.; Marx, A.; Srivatsan, S. G., Microenvironment-Sensitive Fluorescent Nucleotide Probes from Benzofuran, Benzothiophene, and Selenophene as Substrates for DNA Polymerases. *J. Am. Chem. Soc.* **2022,** *144*, 10556-10569.

362. Jun, Y. W.; Wilson, D. L.; Kietrys, A. M.; Lotsof, E. R.; Conlon, S. G.; David, S. S.; Kool, E. T., An Excimer Clamp for Measuring Damaged-Base Excision by the DNA Repair Enzyme NTH1. *Angew. Chem., Int. Ed.* **2020,** *59*, 7450-7455.

363. Beharry, A. A.; Lacoste, S.; O'Connor, T. R.; Kool, E. T., Fluorescence Monitoring of the Oxidative Repair of DNA Alkylation Damage by ALKBH3, a Prostate Cancer Marker. *J. Am. Chem. Soc.* **2016,** *138*, 3647-3650.

364. Wilson, D. L.; Kool, E. T., Ultrafast Oxime Formation Enables Efficient Fluorescence Light-up Measurement of DNA Base Excision. *J. Am. Chem. Soc.* **2019,** *141*, 19379-19388.

365. Meguellati, K.; Koripelly, G.; Ladame, S., DNA-templated synthesis of trimethine cyanine dyes: a versatile fluorogenic reaction for sensing G-quadruplex formation. *Angew. Chem., Int. Ed.* **2010,** *49*, 2738-2742.

366. Prusty, D. K.; Herrmann, A., A fluorogenic reaction based on heavy-atom removal for ultrasensitive DNA detection. *J. Am. Chem. Soc.* **2010,** *132*, 12197-12199.

367. von Krüchten, D. G.; Roth, M.; Seitz, O., DNA-Templated Reactions with High Catalytic Efficiency Achieved by a Loss-of-Affinity Principle. *J. Am. Chem. Soc.* **2022,** *144*, 10700-10704.

368. Klöcker, N.; Weissenboeck, F. P.; van Dülmen, M.; Špaček, P.; Huwel, S.; Rentmeister, A., Photocaged 5' cap analogues for optical control of mRNA translation in cells. *Nat. Chem.* **2022,** *14*, 905-913.

369. Schärer, O. D., Chemistry and biology of DNA repair. *Angew. Chem., Int. Ed.* **2003,** *42*, 2946-2974.

370. Häcker, S.; Schrödter, M.; Kuhlmann, A.; Wagenknecht, H. A., Probing of DNA Photochemistry with C-Nucleosides of Xanthones and Triphenylene as Photosensitizers To Study the Formation of Cyclobutane Pyrimidine Dimers. *JACS Au* **2023,** *3*, 1843-1850.

371. Wagenknecht, H. A., Remote Photodamaging of DNA by Photoinduced Energy Transport. *ChemBioChem* **2022,** *23*, e202100265.

372. Gerling, T.; Kube, M.; Kick, B.; Dietz, H., Sequence-programmable covalent bonding of designed DNA assemblies. *Sci. Adv.* **2018,** *4*, eaau1157.

373. Brown, T. M.; Fakih, H. H.; Saliba, D.; Asohan, J.; Sleiman, H. F., Stabilization of Functional DNA Structures with Mild Photochemical Methods. *J. Am. Chem. Soc.* **2023,** *145*, 2142-2151.

374. Qiu, Z.; Lu, L.; Jian, X.; He, C., A diazirine-based nucleoside analogue for efficient DNA interstrand photocross-linking. *J. Am. Chem. Soc.* **2008,** *130*, 14398-14399.

375. Winnacker, M.; Welzmiller, V.; Strasser, R.; Carell, T., Development of a DNA photoaffinity probe for the analysis of 8-OxodG-binding proteins in a human proteome. *ChemBioChem* **2010,** *11*, 1345-1349.

376. Buchmueller, K. L.; Hill, B. T.; Platz, M. S.; Weeks, K. M., RNA-tethered phenyl azide photocrosslinking via a short-lived indiscriminant electrophile. *J. Am. Chem. Soc.* **2003,** *125*, 10850-10861.

377. Hong, I. S.; Greenberg, M. M., DNA Interstrand Cross-Link Formation Initiated by Reaction between Singlet Oxygen and a Modified Nucleotide. *J. Am. Chem. Soc.* **2005,** *127*, 10510-10511.

378. Guo, A. D.; Yan, K. N.; Hu, H.; Zhai, L.; Hu, T. F.; Su, H.; Chi, Y.; Zha, J.; Xu, Y.; Zhao, D.; Lu, X.; Xu, Y. J.; Zhang, J.; Tan, M.; Chen, X. H., Spatiotemporal and global profiling of DNA-protein interactions enables discovery of low-affinity transcription factors. *Nat. Chem.* **2023,** *15*, 803-814.

379. Op de Beeck, M.; Madder, A., Sequence specific DNA cross-linking triggered by visible light. *J. Am. Chem. Soc.* **2012,** *134*, 10737-10740.

380. Haque, M. M.; Sun, H.; Liu, S.; Wang, Y.; Peng, X., Photoswitchable formation of a DNA interstrand cross-link by a coumarin-modified nucleotide. *Angew. Chem., Int. Ed.* **2014,** *53*, 7001-7005.

381. Cimino, G. D.; Gamper, H. B.; Isaacs, S. T.; Hearst, J. E., Psoralens as photoactive probes of nucleic acid structure and function: organic chemistry, photochemistry, and biochemistry. *Annu. Rev. Biochem.* **1985,** *54*, 1151-1193.

382. Yoshimura, Y.; Fujimoto, K., Ultrafast reversible photo-cross-linking reaction: toward in situ DNA manipulation. *Org. Lett.* **2008,** *10*, 3227-3230.

383. Fujimoto, K.; Sasago, S.; Mihara, J.; Nakamura, S., DNA Photo-cross-linking Using Pyranocarbazole and Visible Light. *Org. Lett.* **2018,** *20*, 2802-2805.

384. Kishi, J. Y.; Liu, N.; West, E. R.; Sheng, K.; Jordanides, J. J.; Serrata, M.; Cepko, C. L.; Saka, S. K.; Yin, P., Light-Seq: light-directed in situ barcoding of biomolecules in fixed cells and tissues for spatially indexed sequencing. *Nat. Methods* **2022,** *19*, 1393-1402.

385. Gerling, T.; Dietz, H., Reversible Covalent Stabilization of Stacking Contacts in DNA Assemblies. *Angew. Chem., Int. Ed.* **2019,** *58*, 2680-2684.

386. Fujimoto, K.; Matsuda, S.; Takahashi, N.; Saito, I., Template-Directed Photoreversible Ligation of Deoxyoligonucleotides via 5-Vinyldeoxyuridine. *J. Am. Chem. Soc.* **2000,** *122*, 5646-5647.

387. Hentschel, S.; Alzeer, J.; Angelov, T.; Schärer, O. D.; Luedtke, N. W., Synthesis of DNA interstrand cross-links using a photocaged nucleobase. *Angew. Chem., Int. Ed.* **2012,** *51*, 3466-3469.

388. Chan, K. M.; Kolmel, D. K.; Wang, S.; Kool, E. T., Color-Change Photoswitching of an Alkynylpyrene Excimer Dye. *Angew. Chem., Int. Ed.* **2017,** *56*, 6497-6501.

389. Onizuka, K.; Ishida, K.; Mano, E.; Nagatsugi, F., Alkyne–Alkyne Photo-cross-linking on the Flipping-out Field. *Org. Lett.* **2019,** *21*, 2833-2837.

390. Kashida, H.; Doi, T.; Sakakibara, T.; Hayashi, T.; Asanuma, H., p-Stilbazole Moieties As Artificial Base Pairs for Photo-Cross-Linking of DNA Duplex. *J. Am. Chem. Soc.* **2013,** *135*, 7960-7966.

391. Abdelhady, A. M.; Onizuka, K.; Ishida, K.; Yajima, S.; Mano, E.; Nagatsugi, F., Rapid Alkene-Alkene Photo-Cross-Linking on the Base-Flipping-Out Field in Duplex DNA. *J. Org. Chem.* **2022,** *87*, 2267-2276.

392. Chaput, J. C.; Herdewijn, P., What Is XNA? *Angew. Chem., Int. Ed.* **2019,** *58*, 11570-11572.

393. Naskar, S.; Guha, R.; Müller, J., Metal-Modified Nucleic Acids: Metal-Mediated Base Pairs, Triples, and Tetrads. *Angew. Chem., Int. Ed.* **2020,** *59*, 1397-1406.

394. Pinheiro, V. B.; Holliger, P., Towards XNA nanotechnology: new materials from synthetic genetic polymers. *Trends Biotechnol.* **2014,** *32*, 321-328.

395. Samanta, B.; Höbartner, C., Combinatorial nucleoside-deletion-scanning mutagenesis of functional DNA. *Angew. Chem., Int. Ed.* **2013,** *52*, 2995-2999.

396. Zhang, H.; Han, J.; Jin, X.; Duan, P., Improving the Overall Properties of Circularly Polarized Luminescent Materials Through Arene–Perfluoroarene Interactions. *Angew. Chem., Int. Ed.* **2021,** *60*, 4575-4580.

397. Roesner, E. K.; Asheghali, D.; Kirillova, A.; Strauss, M. J.; Evans, A. M.; Becker, M. L.; Dichtel, W. R., Arene-perfluoroarene interactions confer enhanced mechanical properties to synthetic nanotubes. *Chem. Sci.* **2022,** *13*, 2475-2480.

398. Grossmann, L.; King, B. T.; Reichlmaier, S.; Hartmann, N.; Rosen, J.; Heckl, W. M.; Bjork, J.; Lackinger, M., On-surface photopolymerization of two-dimensional polymers ordered on the mesoscale. *Nat. Chem.* **2021,** *13*, 730-736.

399. Ponzini, F.; Zagha, R.; Hardcastle, K.; Siegel, J. S., Phenyl/Pentafluorophenyl Interactions and the Generation of Ordered Mixed Crystals: sym-Triphenethynylbenzene and sym-Tris(perfluorophenethynyl)benzene This work was supported by the US National Science Foundation (CHE-9904275). *Angew. Chem., Int. Ed.* **2000,** *39*, 2323-2325.

400. Friedrich, A.; Collings, I. E.; Dziubek, K. F.; Fanetti, S.; Radacki, K.; Ruiz-Fuertes, J.; Pellicer-Porres, J.; Hanfland, M.; Sieh, D.; Bini, R.; Clark, S. J.; Marder, T. B., Pressure-Induced Polymerization of Polycyclic Arene-Perfluoroarene Cocrystals: Single Crystal X-ray Diffraction Studies, Reaction Kinetics, and Design of Columnar Hydrofluorocarbons. *J. Am. Chem. Soc.* **2020,** *142*, 18907-18923.

401. Biedermann, F.; Nau, W. M.; Schneider, H. J., The hydrophobic effect revisited--studies with supramolecular complexes imply high-energy water as a noncovalent driving force. *Angew. Chem., Int. Ed.* **2014,** *53*, 11158-11171.

402. Huang, Z.; Chen, X.; Wu, G.; Metrangolo, P.; Whitaker, D.; McCune, J. A.; Scherman, O. A., Host-Enhanced Phenyl-Perfluorophenyl Polar-pi Interactions. *J. Am. Chem. Soc.* **2020,** *142*, 7356-7361.

403. Woll, M. G.; Hadley, E. B.; Mecozzi, S.; Gellman, S. H., Stabilizing and destabilizing effects of phenylalanine --> F5-phenylalanine mutations on the folding of a small protein. *J. Am. Chem. Soc.* **2006,** *128*, 15932-15933.

404. Moran, S.; Ren, R. X.; Kool, E. T., A thymidine triphosphate shape analog lacking Watson-Crick pairing ability is replicated with high sequence selectivity. *Proc. Natl. Acad. Sci. USA* **1997,** *94*, 10506-10511.

405. Lewis, F. D.; Liu, X.; Miller, S. E.; Hayes, R. T.; Wasielewski, M. R., Formation and decay of localized contact radical ion pairs in DNA hairpins. *J. Am. Chem. Soc.* **2002,** *124*, 14020-14066.

406. Letsinger, R. L.; Wu, T.; Yang, J. S.; Lewis, F. D., DNA-templated formation and luminescence of diphenylacetylene dimeric and trimeric complexes. *Photochem. Photobiol. Sci.* **2008,** *7*, 854-859.

407. Cockroft, S. L.; Hunter, C. A., Chemical double-mutant cycles: dissecting non-covalent interactions. *Chem. Soc. Rev.* **2007,** *36*, 172-188.

408. Weiser, M.; Wagenknecht, H. A., Dynamic DNA architectures: spontaneous DNA strand exchange and self-sorting driven by perylene bisimide interactions. *Chem. Commun.* **2015,** *51*, 16530-16533.

409. Bereiter, R.; Himmelstoss, M.; Renard, E.; Mairhofer, E.; Egger, M.; Breuker, K.; Kreutz, C.; Ennifar, E.; Micura, R., Impact of 3-deazapurine nucleobases on RNA properties. *Nucleic Acids Res.* **2021,** *49*, 4281-4293.

410. Hwang, T. L.; van Zijl, P. C.; Mori, S., Accurate quantitation of water-amide proton exchange rates using the phase-modulated CLEAN chemical EXchange (CLEANEX-PM) approach with a Fast-HSQC (FHSQC) detection scheme. *J. Biomol. NMR* **1998,** *11*, 221-226.

411. Heddi, B.; Foloppe, N.; Bouchemal, N.; Hantz, E.; Hartmann, B., Quantification of DNA BI/BII backbone states in solution. Implications for DNA overall structure and recognition. *J. Am. Chem. Soc.* **2006,** *128*, 9170-9177.

412. Gorenstein, D. G., Conformation and Dynamics of DNA and Protein-DNA Complexes by 31P NMR. *Chem. Rev.* **1994,** *94*, 1315-1338.

413. Phetrak, N.; Rukkijakan, T.; Sirijaraensre, J.; Prabpai, S.; Kongsaeree, P.; Klinchan, C.; Chuawong, P., Regioselectivity of Larock heteroannulation: a contribution from electronic properties of diarylacetylenes. *J. Org. Chem.* **2013,** *78*, 12703-12709.

414. Zhang, S.; Chaput, J. C., Synthesis of glycerol nucleic acid (GNA) phosphoramidite monomers and oligonucleotide polymers. *Curr. Protoc. Nucleic Acid Chem.* **2010,** *Chapter 4*, Unit4 40.

415. Flogel, O.; Okala Amombo, M. G.; Reissig, H. U.; Zahn, G.; Brudgam, I.; Hartl, H., A stereoselective and short total synthesis of the polyhydroxylated gamma-amino acid (-)-detoxinine, based on stereoselective preparation of dihydropyrrole derivatives from lithiated alkoxyallenes. *Chem. Eur. J.* **2003,** *9*, 1405-1415.

416. Lepore, S. D.; He, Y., Use of sonication for the coupling of sterically hindered substrates in the phenolic Mitsunobu reaction. *J. Org. Chem.* **2003,** *68*, 8261-8263.

417. Poppe, M.; Chen, C.; Liu, F.; Poppe, S.; Tschierske, C., Formation of a Cubic Liquid Crystalline Nanostructure with pi-Conjugated Fluorinated Rods on the Gyroid Minimal Surface. *Chem. Eur. J.* **2017,** *23*, 7196-7200.

418. El-Sagheer, A. H.; Kumar, R.; Findlow, S.; Werner, J. M.; Lane, A. N.; Brown, T., A very stable cyclic DNA miniduplex with just two base pairs. *ChemBioChem* **2008,** *9*, 50-52.

419. Wawrzyniak-Turek, K.; Höbartner, C., Enzymatic combinatorial nucleoside deletion scanning mutagenesis of functional RNA. *Chem. Commun.* **2014,** *50*, 10937-10940.

420. Marky, L. A.; Breslauer, K. J., Calculating thermodynamic data for transitions of any molecularity from equilibrium melting curves. *Biopolymers* **1987,** *26*, 1601-1620.

421. Maurer, T.; Kalbitzer, H. R., Indirect Referencing of 31P and 19F NMR Spectra. *J. Magn. Reson. B* **1996,** *113*, 177-178.

422. Wishart, D. S.; Bigam, C. G.; Yao, J.; Abildgaard, F.; Dyson, H. J.; Oldfield, E.; Markley, J. L.; Sykes, B. D., 1H, 13C and 15N chemical shift referencing in biomolecular NMR. *J. Biomol. NMR* **1995,** *6*, 135-140.

423. Goddard, T. D.; Kneller, D. G., *University of California, San Francisco*.

424. Lee, W.; Tonelli, M.; Markley, J. L., NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. *Bioinformatics* **2015,** *31*, 1325-1327.

425. Sklenář, V.; Bax, A., Spin-echo water suppression for the generation of pure-phase two-dimensional NMR spectra. *J. Magn. Res.* **1987,** *74*, 469-479.

426. Hwang, T. L.; Shaka, A. J., Water Suppression That Works. Excitation Sculpting Using Arbitrary Wave-Forms and Pulsed-Field Gradients. *J. Magn. Reson., Series A* **1995,** *112*, 275-279.

427. Derome, A. E.; Williamson, M. P., Rapid-Pulsing Artifacts in Double-Quantum-Filtered Cosy. *J. Magn. Reson.* **1990,** *88*, 177-185.

428. Shaka, A. J.; Lee, C. J.; Pines, A., Iterative schemes for bilinear operators; application to spin decoupling. *J. Magn. Reson.* **1988,** *77*, 274-293.

429. Jeener, J.; Meier, B. H.; Bachmann, P.; Ernst, R. R., Investigation of exchange processes by two-dimensional NMR spectroscopy. *J. Chem. Phys.* **1979,** *71*, 4546-4553.

430. Wagner, R.; Berger, S., Gradient-Selected NOESY-A Fourfold Reduction of the Measurement Time for the NOESY Experiment. *J. Magn. Reson. A* **1996,** *123*, 119-121.

431. Hwang, T. L.; Mori, S.; Shaka, A. J.; van Zijl, P. C. M., Application of phase-modulated CLEAN chemical EXchange spectroscopy (CLEANEX-PM) to detect water-protein proton exchange and intermolecular NOEs. *J. Am. Chem. Soc.* **1997,** *119*, 6203-6204.

432. Lee, Y.-M.; Lee, E.-H.; Seo, Y.-J.; Kang, Y.-M.; Ha, J.-H.; Kim, H.-E.; Lee, J.-H., Measurement of hydrogen exchange times of the RNA imino protons using by phase-modulated CLEAN chemical exchange spectroscopy. *Bull. Korean Chem.* **2009,** *30*, 2197-2198.

433. Strebitzer, E.; Rangadurai, A.; Plangger, R.; Kremser, J.; Juen, M. A.; Tollinger, M.; Al-Hashimi, H. M.; Kreutz, C., 5-Oxyacetic Acid Modification Destabilizes Double Helical Stem Structures and Favors Anionic Watson-Crick like cmo(5) U-G Base Pairs. *Chem. Eur. J.* **2018,** *24*, 18903-18906.

434. Szulik, M. W.; Voehler, M.; Stone, M. P., NMR analysis of base-pair opening kinetics in DNA. *Curr. Protoc. Nucl. Acid Chem.* **2014,** *59*, 7 20 1-18.

435. Neese, F., The ORCA program system. *Wires Comput Mol Sci* **2012,** *2*, 73-78.

436. Neese, F., Software update: the ORCA program system, version 4.0. *Wires Comput Mol Sci* **2018,** *8*, e1327.

437. Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H., A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **2010,** *132*, 154104.

438. Grimme, S.; Ehrlich, S.; Goerigk, L., Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.* **2011,** *32*, 1456-1465.

439. Schäfer, A.; Horn, H.; Ahlrichs, R., Fully optimized contracted Gaussian basis sets for atoms Li to Kr. *J. Chem. Phys.* **1992,** *97*, 2571-2577.

440. Weigend, F.; Ahlrichs, R., Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Phys. Chem. Chem. Phys.* **2005,** *7*, 3297-3305.

441. Weigend, F., Accurate Coulomb-fitting basis sets for H to Rn. *Phys. Chem. Chem. Phys.* **2006,** *8*, 1057-1065.

442. Retegan, M., **2019**, "https://gist.github.com/mretegan/5501553".

443. Neitz, H.; Bessi, I.; Kachler, V.; Michel, M.; Höbartner, C., Tailored Tolane-Perfluorotolane Assembly as Supramolecular Base Pair Replacement in DNA. *Angew. Chem., Int. Ed.* **2023,** *62*, e202214456.

444. Kashida, H.; Ito, H.; Fujii, T.; Hayashi, T.; Asanuma, H., Positively charged base surrogate for highly stable "base pairing" through electrostatic and stacking interactions. *J. Am. Chem. Soc.* **2009,** *131*, 9928-9930.

445. Brotschi, C.; Häberli, A.; Leumann, C. J., A stable DNA duplex containing a non-hydrogen-bonding and non-shape-complementary base couple: interstrand stacking as the stability determining factor. *Angew. Chem., Int. Ed.* **2001,** *40*, 3012-3014.

446. Kashida, H.; Sekiguchi, K.; Asanuma, H., Insulator base pairs for lighting-up perylenediimide in a DNA duplex. *Chem. Eur. J.* **2010,** *16*, 11554-11557.

447. Ishizuka, T.; Kimoto, M.; Sato, A.; Hirao, I., Site-specific functionalization of RNA molecules by an unnatural base pair transcription system via click chemistry. *Chem. Commun.* **2012,** *48*, 10835-10837.

448. Domnick, C.; Eggert, F.; Wuebben, C.; Bornewasser, L.; Hagelueken, G.; Schiemann, O.; Kath-Schorr, S., EPR Distance Measurements on Long Non-coding RNAs Empowered by Genetic Alphabet Expansion Transcription. *Angew. Chem., Int. Ed.* **2020,** *59*, 7891-7896.

449. Romesberg, F. E., Creation, Optimization, and Use of Semi-Synthetic Organisms that Store and Retrieve Increased Genetic Information. *J. Mol. Biol.* **2022,** *434*, 167331.

450. Malinovskii, V. L.; Wenger, D.; Häner, R., Nucleic acid-guided assembly of aromatic chromophores. *Chem. Soc. Rev.* **2010,** *39*, 410-422.

451. Smith, C. E.; Smith, P. S.; Thomas, R. L.; Robins, E. G.; Collings, J. C.; Dai, C.; Scott, A. J.; Borwick, S.; Batsanov, A. S.; Watt, S. W.; Clark, S. J.; Viney, C.; Howard, J. A. K.; Clegg, W.; Marder, T. B., Arene-perfluoroarene interactions in crystal engineering: structural preferences in polyfluorinated tolans. *J. Mater. Chem.* **2004,** *14*, 413-420.

452. Yamada, S.; Konno, T., Development of Donor-π-Acceptor-Type Fluorinated Tolanes as Compact Condensed Phase Luminophores and Applications in Photoluminescent Liquid-Crystalline Molecules. *Chem. Rec. 23*, e202300094.

453. Johnson, S. A.; Liu, F. Q.; Suh, M. C.; Zurcher, S.; Haufe, M.; Mao, S. S.; Tilley, T. D., Regioselective coupling of pentafluorophenyl substituted alkynes: mechanistic insight into the zirconocene coupling of alkynes and a facile route to conjugated polymers bearing electron-withdrawing pentafluorophenyl substituents. *J. Am. Chem. Soc.* **2003,** *125*, 4199-4211.

454. Teo, Y. N.; Kool, E. T., DNA-multichromophore systems. *Chem. Rev.* **2012,** *112*, 4221-4245.

455. O'Reilly, R. K.; Turberfield, A. J.; Wilks, T. R., The Evolution of DNA-Templated Synthesis as a Tool for Materials Discovery. *Acc. Chem. Res.* **2017,** *50*, 2496-2509.

456. Onizuka, K.; Yamano, Y.; Abdelhady, A. M.; Nagatsugi, F., Hybridization-specific chemical reactions to create interstrand crosslinking and threaded structures of nucleic acids. *Org. Biomol. Chem.* **2022,** *20*, 4699-4708.

457. Schärer, O. D., DNA Interstrand Crosslinks: Natural and Drug-Induced DNA Adducts that Induce Unique Cellular Responses. *ChemBioChem* **2005,** *6*, 27-32.

458. Eastman, A., Interstrand cross-links and sequence specificity in the reaction of cis-dichloro(ethylenediamine)platinum(II) with DNA. *Biochemistry* **1985,** *24*, 5027-5032.

459. Rink, S. M.; Solomon, M. S.; Taylor, M. J.; Rajur, S. B.; McLaughlin, L. W.; Hopkins, P. B., Covalent structure of a nitrogen mustard-induced DNA interstrand cross-link: an N7-to-N7 linkage of deoxyguanosine residues at the duplex sequence 5'-d(GNC). *J. Am. Chem. Soc.* **1993,** *115*, 2551-2557.

460. Erlanson, D. A.; Glover, J. N. M.; Verdine, G. L., Disulfide Cross-linking as a Mechanistic Probe for the B ↔ Z Transition in DNA. *J. Am. Chem. Soc.* **1997,** *119*, 6927-6928.

461. Tomás-Gamasa, M.; Serdjukow, S.; Su, M.; Müller, M.; Carell, T., "Post-it" type connected DNA created with a reversible covalent cross-link. *Angew. Chem., Int. Ed.* **2015,** *54*, 796-800.

462. Dutta, S.; Chowdhury, G.; Gates, K. S., Interstrand Cross-Links Generated by Abasic Sites in Duplex DNA. *J. Am. Chem. Soc.* **2007,** *129*, 1852-1853.

463. Kusano, S.; Ishiyama, S.; Lam, S. L.; Mashima, T.; Katahira, M.; Miyamoto, K.; Aida, M.; Nagatsugi, F., Crosslinking reactions of 4-amino-6-oxo-2-vinylpyrimidine with guanine derivatives and structural analysis of the adducts. *Nucleic Acids Res.* **2015,** *43*, 7717-7730.

464. Peng, X.; Hong, I. S.; Li, H.; Seidman, M. M.; Greenberg, M. M., Interstrand Cross-Link Formation in Duplex and Triplex DNA by Modified Pyrimidines. *J. Am. Chem. Soc.* **2008,** *130*, 10299-10306.

465. Morton, S. B.; Finger, L. D.; van der Sluijs, R.; Mulcrone, W. D.; Hodskinson, M.; Millington, C. L.; Vanhinsbergh, C.; Patel, K. J.; Dickman, M. J.; Knipscheer, P.; Grasby, J. A.;

Williams, D. M., Efficient Synthesis of DNA Duplexes Containing Reduced Acetaldehyde Interstrand Cross-Links. *J. Am. Chem. Soc.* **2023,** *145*, 953-959.

466. Kočalka, P.; El-Sagheer, A. H.; Brown, T., Rapid and efficient DNA strand cross-linking by click chemistry. *ChemBioChem* **2008,** *9*, 1280-1285.

467. Tera, M.; Harati Taji, Z.; Luedtke, N. W., Intercalation-enhanced "Click" Crosslinking of DNA. *Angew. Chem., Int. Ed.* **2018,** *57*, 15405-15409.

468. Lechner, V. M.; Nappi, M.; Deneny, P. J.; Folliet, S.; Chu, J. C. K.; Gaunt, M. J., Visible-Light-Mediated Modification and Manipulation of Biomacromolecules. *Chem. Rev.* **2022,** *122*, 1752-1829.

469. Kobertz, W. R.; Essigmann, J. M., Solid-Phase Synthesis of Oligonucleotides Containing a Site-Specific Psoralen Derivative. *J. Am. Chem. Soc.* **1997,** *119*, 5960-5961.

470. Rajendran, A.; Endo, M.; Katsuda, Y.; Hidaka, K.; Sugiyama, H., Photo-cross-linking-assisted thermal stability of DNA origami structures and its application for higher-temperature self-assembly. *J. Am. Chem. Soc.* **2011,** *133*, 14488-14491.

471. Antusch, L.; Gass, N.; Wagenknecht, H. A., Elucidation of the Dexter-Type Energy Transfer in DNA by Thymine-Thymine Dimer Formation Using Photosensitizers as Artificial Nucleosides. *Angew. Chem., Int. Ed.* **2017,** *56*, 1385-1389.

472. Fujimoto, K.; Yamada, A.; Yoshimura, Y.; Tsukaguchi, T.; Sakamoto, T., Details of the ultrafast DNA photo-cross-linking reaction of 3-cyanovinylcarbazole nucleoside: cis-trans isomeric effect and the application for SNP-based genotyping. *J. Am. Chem. Soc.* **2013,** *135*, 16161-16167.

473. Doi, T.; Kawai, H.; Murayama, K.; Kashida, H.; Asanuma, H., Visible-Light-Triggered Cross-Linking of DNA Duplexes by Reversible [2+2] Photocycloaddition of Styrylpyrene. *Chem. Eur. J.* **2016,** *22*, 10533-10538.

474. Hudson, R. H.; Ghorbani-Choghamarani, A., Oligodeoxynucleotides incorporating structurally simple 5-alkynyl-2'-deoxyuridines fluorometrically respond to hybridization. *Org. Biomol. Chem.* **2007,** *5*, 1845-1848.

475. Tanaka, M.; Oguma, K.; Saito, Y.; Saito, I., Drastic enhancement of excess electron-transfer efficiency through DNA by inserting consecutive 5-phenylethynyl-2’-deoxyuridines as a modulator. *Chem. Commun.* **2012,** *48*, 9394-9396.

476. Dietzsch, J.; Bialas, D.; Bandorf, J.; Würthner, F.; Höbartner, C., Tuning Exciton Coupling of Merocyanine Nucleoside Dimers by RNA, DNA and GNA Double Helix Conformations. *Angew. Chem., Int. Ed.* **2022,** *61*, e202116783.

477. Maturi, M. M.; Bach, T., Enantioselective catalysis of the intermolecular [2+2] photocycloaddition between 2-pyridones and acetylenedicarboxylates. *Angew. Chem., Int. Ed.* **2014,** *53*, 7661-7664.

478. Pappas, S. P.; Pappas, B. C.; Portnoy, N. A., Synthesis and solvolytic rearrangement of 1-methoxybicyclo[4.2.0]octa-3,7-diene-2,5-diones. *J. Org. Chem.* **1969,** *34*, 520-525.

479. Ha, S.; Lee, Y.; Kwak, Y.; Mishra, A.; Yu, E.; Ryou, B.; Park, C. M., Alkyne-Alkene [2 + 2] cycloaddition based on visible light photocatalysis. *Nat. Commun.* **2020,** *11*, 2509.

480. Ma, X.; Maier, J.; Wenzel, M.; Friedrich, A.; Steffen, A.; Marder, T. B.; Mitrić, R.; Brixner, T., Direct observation of o-benzyne formation in photochemical hexadehydro-Diels-Alder (hnu-HDDA) reactions. *Chem Sci* **2020,** *11*, 9198-9208.

481. Khan, I.; Sommers, J. A.; Brosh, R. M., Jr., Close encounters for the first time: Helicase interactions with DNA damage. *DNA Repair* **2015,** *33*, 43-59.

482. Pandey, M.; Patel, S. S., Helicase and polymerase move together close to the fork junction and copy DNA in one-nucleotide steps. *Cell Rep.* **2014,** *6*, 1129-1138.

483. Kuper, J.; Wolski, S. C.; Michels, G.; Kisker, C., Functional and structural studies of the nucleotide excision repair helicase XPD suggest a polarity for DNA translocation. *EMBO J.* **2012,** *31*, 494-502.

484. Barnett, J. T.; Kuper, J.; Koelmel, W.; Kisker, C.; Kad, N. M., The TFIIH subunits p44/p62 act as a damage sensor during nucleotide excision repair. *Nucleic Acids Res.* **2020,** *48*, 12689-12696.

485. Petruseva, I.; Naumenko, N.; Kuper, J.; Anarbaev, R.; Kappenberger, J.; Kisker, C.; Lavrik, O., The Interaction Efficiency of XPD-p44 With Bulky DNA Damages Depends on the Structure of the Damage. *Front. Cell Dev. Biol.* **2021,** *9*, 617160.

486. Romero-Perez, S.; Lopez-Martin, I.; Martos-Maldonado, M. C.; Somoza, A.; Gonzalez-Rodriguez, D., Synthesis of Phosphoramidite Monomers Equipped with Complementary Bases for Solid-Phase DNA Oligomerization. *Org. Lett.* **2020,** *22*, 41-45.

487. Neitz, H.; Bessi, I.; Kuper, J.; Kisker, C.; Höbartner, C., Programmable DNA Interstrand Crosslinking by Alkene-Alkyne [2 + 2] Photocycloaddition. *J. Am. Chem. Soc.* **2023,** *145*, 9428-9433.

488. Dziuba, D.; Didier, P.; Ciaco, S.; Barth, A.; Seidel, C. A. M.; Mely, Y., Fundamental photophysics of isomorphic and expanded fluorescent nucleoside analogues. *Chem Soc Rev* **2021,** *50*, 7062-7107.

489. Michel, B. Y.; Dziuba, D.; Benhida, R.; Demchenko, A. P.; Burger, A., Probing of Nucleic Acid Structures, Dynamics, and Interactions With Environment-Sensitive Fluorescent Labels. *Front. Chem.* **2020,** *8*, 112.

490. Wranne, M. S.; Füchtbauer, A. F.; Dumat, B.; Bood, M.; El-Sagheer, A. H.; Brown, T.; Gradén, H.; Grøtli, M.; Wilhelmsson, L. M., Toward Complete Sequence Flexibility of Nucleic Acid Base Analogue FRET. *J. Am. Chem. Soc.* **2017,** *139*, 9271-9280.

491. Eubanks, C. S.; Forte, J. E.; Kapral, G. J.; Hargrove, A. E., Small Molecule-Based Pattern Recognition To Classify RNA Structure. *J. Am. Chem. Soc.* **2017,** *139*, 409-416.

492. Okamoto, A.; Kanatani, K.; Saito, I., Pyrene-labeled base-discriminating fluorescent DNA probes for homogeneous SNP typing. *J. Am. Chem. Soc.* **2004,** *126*, 4820-4827.

493. Ihara, T.; Fujii, T.; Mukae, M.; Kitamura, Y.; Jyo, A., Photochemical ligation of DNA conjugates through anthracene cyclodimer formation and its fidelity to the template sequences. *J. Am. Chem. Soc.* **2004,** *126*, 8880-8881.

494. Manicardi, A.; Cadoni, E.; Madder, A., Visible-light triggered templated ligation on surface using furan-modified PNAs. *Chem. Sci.* **2020,** *11*, 11729-11739.

495. Thoresen, L. H.; Jiao, G. S.; Haaland, W. C.; Metzker, M. L.; Burgess, K., Rigid, conjugated, fluoresceinated thymidine triphosphates: syntheses and polymerase mediated incorporation into DNA analogues. *Chem. Eur. J.* **2003,** *9*, 4603-4610.

496. Carmieli, R.; Zeidan, T. A.; Kelley, R. F.; Mi, Q.; Lewis, F. D.; Wasielewski, M. R., Excited state, charge transfer, and spin dynamics in DNA hairpin conjugates with perylenediimide hairpin linkers. *J. Phys. Chem. A* **2009,** *113*, 4691-4700.

497. Zhang, K.; Yam, V. W., Platinum(II) non-covalent crosslinkers for supramolecular DNA hydrogels. *Chem. Sci.* **2020,** *11*, 3241-3249.

498. Wang, D.; Liu, P.; Luo, D., Putting DNA to Work as Generic Polymeric Materials. *Angew. Chem., Int. Ed.* **2022,** *61*, e202110666.

499. Zhang, X.; Xie, X.; Liu, Y., Nickel-Catalyzed Highly Regioselective Hydrocyanation of Terminal Alkynes with $Zn(CN)_2$ Using Water as the Hydrogen Source. *J. Am. Chem. Soc.* **2018,** *140*, 7385-7389.

500. Steinmetzger, C.; Bäuerlein, C.; Höbartner, C., Supramolecular Fluorescence Resonance Energy Transfer in Nucleobase-Modified Fluorogenic RNA Aptamers. *Angew. Chem., Int. Ed.* **2020,** *59*, 6760-6764.

501. Neitz, H.; Höbartner, C., A tolane-modified 5-ethynyluridine as a universal and fluorogenic photochemical DNA crosslinker. *Chem. Commun.* **2023,** *59*, 12003-12006.