Univerzita Karlova v Praze

Přírodovědecká fakulta

Studijní program: Organická chemie

Studijní obor: Organická chemie

Alessandro Panattoni

Synthesis and studies of modified DNA: (i) development of DNA targeting molecular scissors and (ii) competitive enzymatic incorporation of base-modified nucleotides

Syntéza a studie modifikované DNA: (i) vývoj molekulárních nůžek cílících na DNA a (ii) kompetitivní enzymová inkorporace nukleotidů modifikovaných na bázi

Disertační práce

Školitel: Prof. Ing. Michal Hocek, CSc., DSc.

Praha, 2020

Disertační práce byla vypracována na Ústavu Organické Chemie a Biochemie, Akademie věd České Republiky v Praze v období říjen 2015 – květen 2020.

Prohlášení

Prohlašuji, že jsem závěrečnou práci zpracoval samostatně a že jsem uvedl všechny použité informační zdroje a literaturu. Tato práce, ani její podstatná část, nebyla předložena k získání jiného nebo stejného akademického titulu.

V Praze, dne

Podpis: Alessandro Panattoni

Aknowledgments

In the first place, I would like to express my deepest gratitude to my supervisor Prof. Michal Hocek, for infinite opportunities to learn under his guidance, as well as for his continuous encouragement and inspiring advice. I also want to thank Dr. Hana Cahova, for her relevant help and careful co-supervision during the first few months of my Ph.D. study. I would like to thank Prof. Tom Brown and Dr. Afaf El-Sagheer, who welcomed me in their lab at the University of Oxford for a one-month secondment. Similarly, I must thank Prof. Andrew Kellett for welcoming me in his group at the Dublin City University during my one-month intership. Furthermore, I thank Dr. Radek Pohl for his help in measurement and interpretation of NMR spectra.

Many thanks are going to former and current members of the group. Among them I have to mention Anna Simonova, Jan Matyasovsky, Nemanja Milisavljevic and Mike Downey, whose precious friendship outside the lab was fundamental in more than many occasions. Now, my best wishes are going to Simon Pospisil.

Finally, my great appreciation goes to my whole family and my beloved Karolina for constant support and encouragement.

This work was part of a multidisciplinary project performed in Prof. Michal Hocek group at the Institute of Organic Chamestry and Biochemistry AS CR (IOCB, Prague) in collaboration with Prof. Tom Brown (University of Oxford) and Prof. Andrew Kellett (DCU, Dublin). All the chemical synthesis, DNA synthesis (both enzymatic and chemical), and other biochemical experiments were performed by me. Most NMR spectra were measured and interpreted by Dr. Radek Pohl. Experiments performed or designed by collaborators, are clearly stated in the thesis.

This work was supported by the H2020 Marie Skłodowska-Curie Actions. Grant Number: H2020-MSCA-ITN-2014-642023; the Czech Science Foundation: 18-03305S; Academy of Sciences of the Czech Republic: RVO: 61388963 and Praemium Academiae to Michal Hocek; the Biotechnology and Biological Sciences Research Council. Grant Numbers: BB/J001694/2, BB/ R008655/1; and the European Regional Development Fund; OP RDE (No. CZ.02.1.01/0.0/0.0/16_019/0000729 to Michal Hocek.

Abstract

In the first part of this work, a series of site-specific artificial metallonucleases (AMNs) was developed conjugating clamped-phenanthroline (Clip-Phen) copper complexes to triplex-forming oligonucleotides (TFOs). Several synthetic routes were explored for the synthesis of the TFO-AMNs hybrids, all sharing a copper-catalyzed alkyne-azide cycloaddition (CuAAC) reaction as the key step. As a consequence, building blocks for enzymatic or chemical synthesis of oligonucleotides (ONs) containing clickable groups, or already conjugated to the Clip-Phen ligand via CuAAC, were prepared.

Two new alkynyl-linked nucleoside-5'-*O*-triphosphates (dNTPs) were designed and developed in order to obtain an efficient polymerase incorporation of clickable alkynyl-tethers into ONs and, at the same time, enhance the efficiency of CuAAC reactions on modified DNA. The relative 3'-*O*-phosphormaidites were also prepared in order to insert the same alkynyl-linkers into ONs via solid-phase synthesis.

The AMN was linked at the 5'- or 3'-ends or in the middle of the TFO stretch, using diverse likers. The hybridization of all the synthesized TFOs with a target DNA duplex was studied. Finally, an extensive study of cleavage efficiency and specificity of the TFO-AMN conjugates towards the target DNA was performed, exploring the influence of the hybrid nature and reaction parameters.

In the second part of this work, a systematic study of competitive polymerase incorporation of 5-substituted pyrimidine and 7-substituted 7-deazapurine dNTPs, in the presence of their natural counterparts (natural dNTPs) was performed. The base-modified dN^RTPs were incorporated into ONs by competitive primer extension (PEX) or polymerase chain reaction (PCR), in different ratios with natural dNTPs. A study of enzymatic kinetics of single-nucleotide incorporation was performed to explore the substrate affinity to the Bst DNA polymerase and incorporation rate of all base-modified dN^RTPs.

All studied dNTPs were successfully incorporated into DNA in the presence of natural dNTPs. 7-deazapurine dNTPs bearing π-electron-containing substituents, and 5-phenyl pyrimidine dNTPs, were found to be even better substrates of the Bst polymerase than their natural counterparts.

Souhrn

V první části této práce byla vyvinuta série specifických umělých metalonukleáz (AMN), které konjugují sevřené fenanthrolinové měděné komplexy (Clip-Phen) s oligonukleotidy vytvářejícími triplex (TFO). Pro syntézu hybridů TFO-AMN bylo zkoumáno několik syntetických cest, přičemž všechny sdílely mědí katalyzovanou alkyn-azidovou cykloadici (CuAAC) jako klíčový krok. V důsledku toho byly připraveny stavební bloky pro enzymatickou nebo chemickou syntézu oligonukleotidů (ON) obsahujících klikací skupiny nebo již konjugované s ligandem Clip-Phen přes CuAAC. Byly navrženy a vyvinuty dva nové alkynyl modifikované nukleosid-5'-*O*-trifosfáty (dNTP), aby bylo dosaženo účinné polymerázové inkorporace klikacích alkynyl řetězců do ON a zároveň zvýšení účinnosti CuAAC reakcí na modifikované DNA. Byly také připraveny odpovídající 3'-*O*-fosforamidity, aby se pomocí chemické syntézy na pevné fázi vložily stejné alkynylové linkery do ON.

AMN byla připojena na 5’- nebo 3’- koncích nebo uprostřed TFO úseku pomocí různých linkerů. Byla studována hybridizace všech syntetizovaných TFO s duplexem DNA obsahujícím cílovou sekvenci. Nakonec byla provedena rozsáhlá studie účinnosti štěpení a specificity konjugátů TFO-AMN vůči cílové DNA, přičemž se zkoumal vliv vlastností hybridů a reakčních parametrů.

Ve druhé části této práce bylo provedeno systematické studium kompetitivní polymerázové inkorporace 5-substituovaných pyrimidinových a 7-substituovaných 7-deazapurinových dNTP v přítomnosti jejich přirozených protějšků (přírodních dNTP). dN^RTP s modifikovanými nukleobázemi byly inkorporovány do ON kompetitivní extenzi primeru (PEX) nebo polymerázovou řetězovou reakcí (PCR), v různých poměrech s přírodními dNTP. Byla studována enzymatická kinetika inkorporace jednoho nukleotidu, aby se prozkoumala afinita substrátu k Bst DNA polymeráze a rychlost inkorporace všech dN^RTP s modifikovanými nukleobázemi.

Všechny studované dNTP byly úspěšně začleněny do DNA v přítomnosti přírodních dNTP. Bylo zjištěno, že 7-deazapurinové dNTP nesoucí substituenty bohaté na π-elektrony a 5-fenylpyrimidinové dNTP jsou ještě lepšími substráty pro Bst polymerázu, než jejich přirozené protějšky.

List of publications of the author, related to the Thesis

1. H. Cahová, A. Panattoni, P. Kielkowski, J. Fanfrlík and M. Hocek; 5-Substituted Pyrimidine and 7-Substituted 7-Deazapurine dNTPs as Substrates for DNA Polymerases in Competitive Primer Extension in Presence of Natural dNTPs; *ACS Chem. Biol.*, 2016, 11, 3165–3171 (DOI: 10.1021/acschembio.6b00714).

2. A. Panattoni, R. Pohl and M. Hocek; Flexible Alkyne-Linked Thymidine Phosphoramidites and Triphosphates for Chemical or Polymerase Synthesis and Fast Postsynthetic DNA Functionalization through Copper-Catalyzed Alkyne–Azide 1,3-Dipolar Cycloaddition; *Org. Lett.*, 2018, 20, 3962–3965 (DOI: 10.1021/acs.orglett.8b01533).

3. A. Panattoni, A. H. El-Sagheer, T. Brown, A. Kellett and M. Hocek; Oxidative DNA Cleavage with Clip-Phenanthroline Triplex Forming Oligonucleotide Hybrids; *ChemBioChem*, 2020, 21, 991–1000. (DOI: 10.1002/cbic.201900670).

Table of Contents

1. Introduction

1.1. DNA function and structure

In all known organisms, cells need to be able to store and hand down all the functional and structural information they possess, the genetic information, containing all the instructions for the growth, functioning and reproduction of the organism. These instructions are enclosed into the so called genome (or genetic material), consisting of deoxyribonucleic acid (DNA). The genetic information is used for the biosynthesis of RNA and, consequently, the biosynthesis of proteins, proving the crucial role and importance of DNA. Chemically, the DNA molecule is a polymer in which monomers (2'-deoxyribonucleotides, dNs) contain a nucleic base (a purine base or a pyrimidine base) connected to the 1' carbon of a ribose sugar. The ribose backbone is formed through phosphodiester bonds which connects the 5' oxygen of a nucleotide (nt) to the 3' oxygen of the following nucleotide (Figure 1A). The elucidation of its structure by Watson and Crick,[1] and independently by Rosalind Franklin[2] and Maurice Wilkins[3], showed that DNA is constructed from four different nucleotides: 2'-deoxyadenosine (**dA**), 2'-deoxycytidine (**dC**), 2'-deoxyguanosine (**dG**) and 2'-deoxythymidine (**dT**), the so-called canonical bases which order along the polymer define and constitute the genetic information. Two DNA strands coil one around each-other to form a double helix (Figure 1B). The two strands are coupled in opposite (antiparallel) direction and hold together by π–stacking between the nucleobases, which are facing the inner part of the helix. Hydrogen bonding between two bases belonging to the two different strands (Watson-Crick base pairing) occur in a canonical way, with a **dT** always displaced opposite to a **dA** and a **dC** always facing a **dG**. The two strands are complementary and store the same information. Two H-bonds form between A and T, and three form between G and C (Figure 1C).

Different forms of DNA exist in nature, in dependence on the environment. The three most common A-DNA, B-DNA and Z-DNA forms are all double stranded helices, and the Z-DNA is the only structure consisting of a left-handed helix. Among all, the B-form of DNA is the most commonly found, and exists under normal physiological conditions. The base pairs in the B-DNA are displaced on the helix axis in a perpendicular way. The major and minor grooves of the helix

in B-DNA are very similar in depth, although the former is considerably wider (22 Å wide, against the 12 Å of the minor groove).

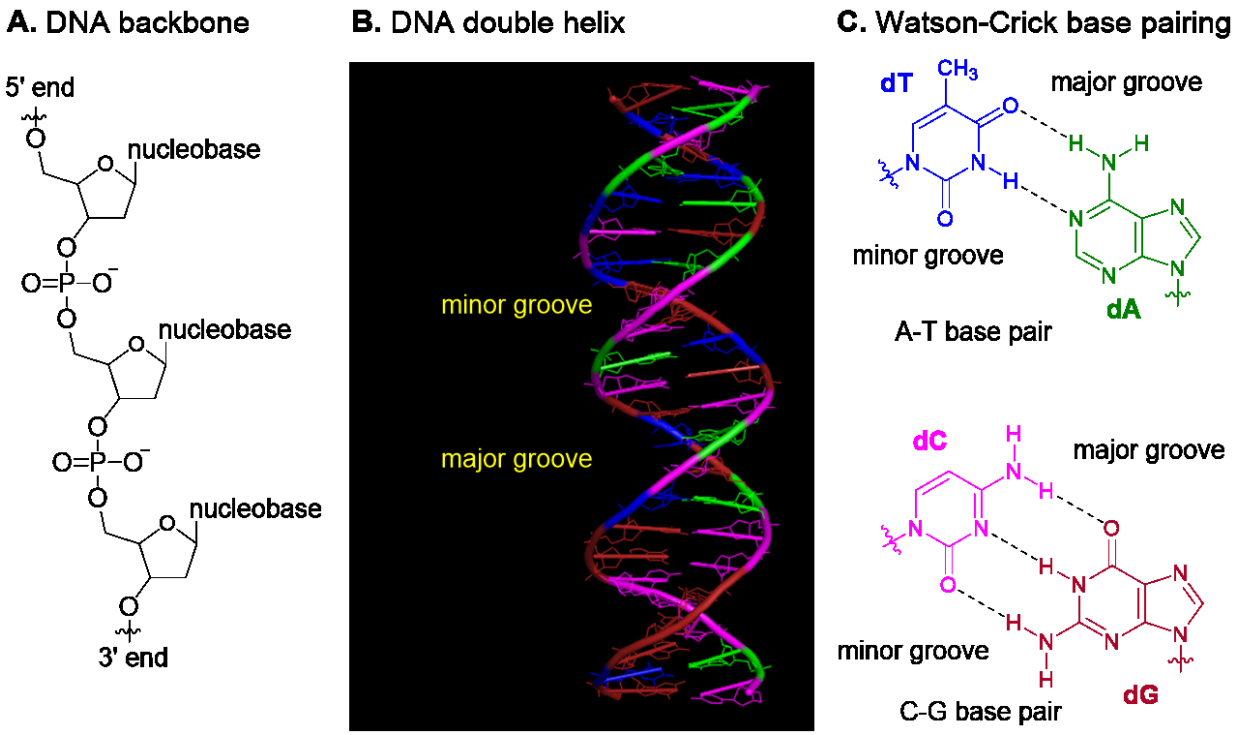


Figure 1. A. structure of the DNA backbone; **B.** Cartoon of a B-DNA double helix; **C.** Hydrogen bonding in Watson-Crick base pairing.

1.2. Enzymatic approaches to the synthesis of DNA

In nature, enzymes called DNA polymerases are in charge of the DNA synthesis. Nucleotide 5'-*O*-triphosphates (dNTPs) are the substrates of these enzymes, used as building blocks for the synthesis of new strands. The DNA polymerase adds new dNTPs to the 3'-hydroxyl group of a pre-existing polynucleotide chain (called primer). The primer is annealed to a longer DNA strand (template, longer at the 3' side of the primer), which is read by the polymerase while adding dNTPs to the primer, in order to match the addition following the Watson-Crick base pairing. As a consequence, the obtained product has a sequence complementary to the template.

Several artificial approaches to the production of ssDNA or dsDNA employing DNA polymerases were developed, whereas, the two main ones are the Primer Extension (PEX) and the Polymerase Chain Reaction (PCR).

In PEX, a short oligonucleotide (ON), generally of 20-50 nt, is used as primer annealed to a longer template, to be extended in a isothermal reaction by the DNA polymerase (Figure 2A). In a typical assay, the reaction mixtures must contain the required ONs (primer and template), the polymerase in a suitable buffer, and dNTPs as building blocks for the elongation of the primer.

On the other hand, the PCR process developed by Kary Mullis in the 1980s,[4] is slightly more complicated and consists of few isothermal steps. A rather long dsDNA is denatureated during the initiation step (at high temperatures, generally around 95 °C) to provide the two single stranded templates (Figure 2B). Two different primers are required, with sequences designed to bind the two templates at their 3' ends. Their annealing step is generally conducted at 45-55 °C, in dependence of the primers sequence and lengths. The DNA polymerase elongates (typically at 70-75 °C) the two primers to obtain two independent final dsDNA products identical to the initial starting duplex. The two products are denaturated providing two couples of identical templates, which are entering a second annealing-elongation-denaturation cycle. Every cycle exponentially amplifies the product outcome.

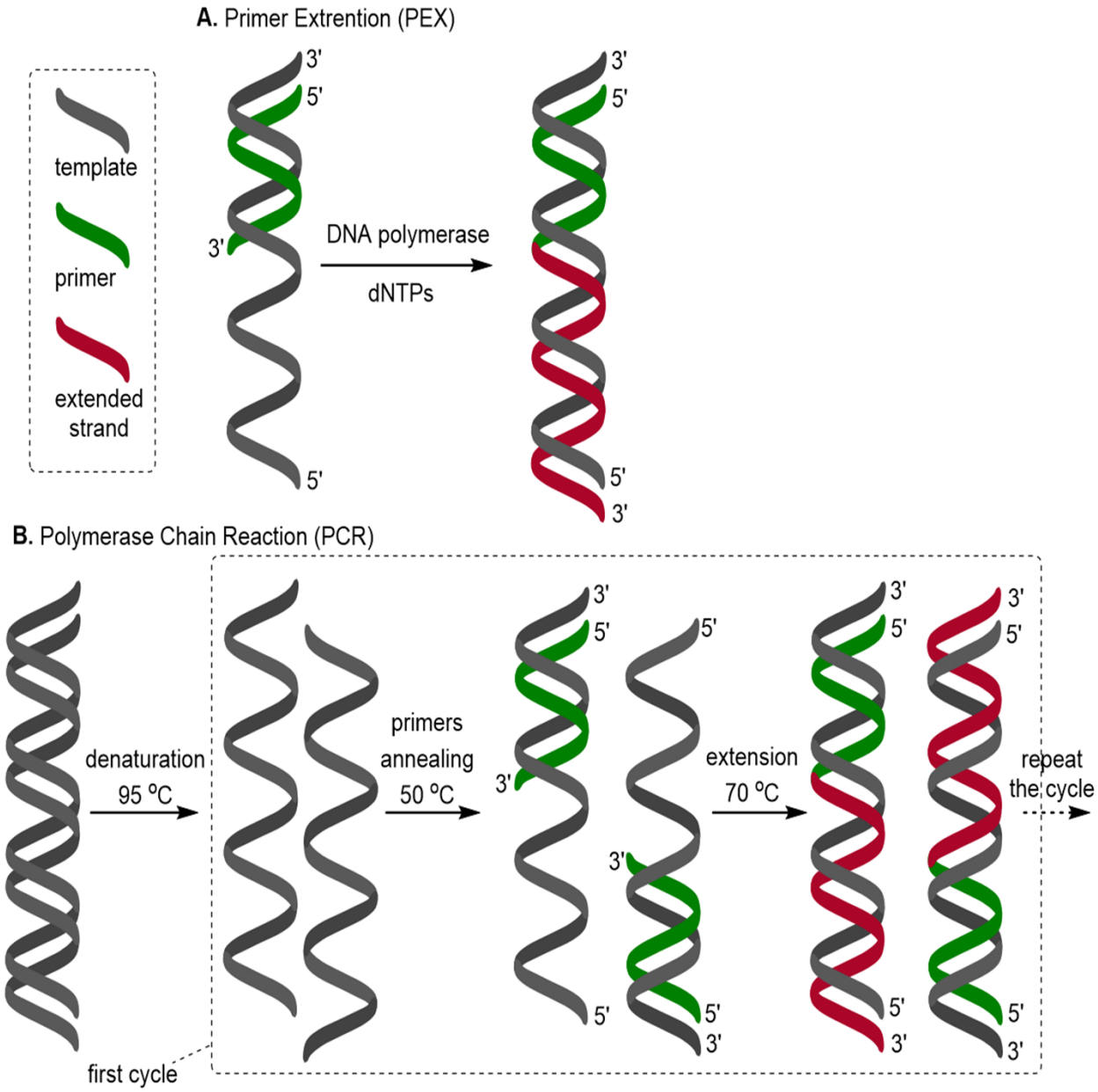


Figure 2. Enzymatic approaches to the synthesis of DNA: **A.** PEX, and **B.** PCR.

1.2.1. Enzymatic synthesis of ssDNA using the Nicking Enzyme Amplification Reaction

As previously seen, PEX and PCR methodologies for enzymatic DNA synthesis yield dsDNA. In addition, none of these techniques are possible if the synthesis of short ONs is required, since they require a stable hybridization of primer and template. However, the enzymatic production of short single-stranded ONs can be very important, for example for the production of primers

to be employed in PEX or PCR. At this purpose, the Nicking Enzyme Amplification Reaction (NEAR) is a very useful method which exploits the combination of two enzymes, particularly a DNA polymerase and a nicking enzyme, for the production and amplification of short ssDNA. Moreover, the whole process is isothermal, making NEAR a powerful tool for *in vivo* DNA replication and amplification (Figure 3).[5]

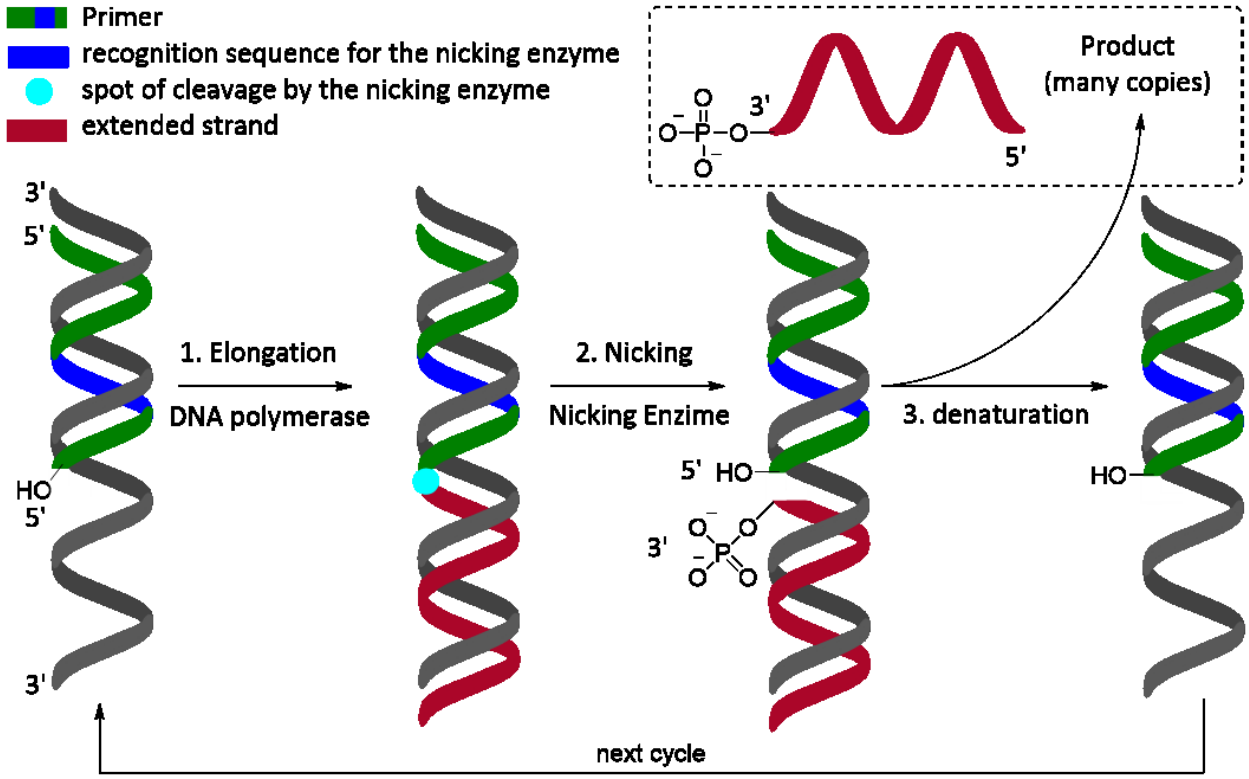


Figure 3. Catalytic cycle of the Nicking Enzyme Amplification Reaction.

The catalytic cycle starts with the hybridization of primer and template, similarly to PEX. Thus, the DNA polymerase elongates the primer till the full length product, which sequence and length are dictated by the template. The primer contains a recognition sequence for a nicking enzyme, which is able to snip a single strand of the targeted dsDNA, in proximity (but not within) the recognition sequence. The cleavage by the nicking enzyme must occur at a previously designed spot of the elongated strand, in order to restore the original primer. The second DNA segment resulting from the cleavage maintain the phosphate group in 3', and will be the NEAR product. A careful design of the product length (therefore, of the template) is very important for a success

in NEAR. The length of the product must be shorter than the one of the primer, in order to have a lower hybridization energy of product and template, with respect to the one of primer and template. In this way it is possible to find a temperature at which the primer can stay annealed to the template, but not the product, being the hybridization energy too low. Therefore the NEAR product falls off the template, leaving the DNA polymerase free to re-elongate the primer, starting the second catalytic cycle.

In this way, many copies of the product can be formed from a single pair of primer and template, proving NEAR to be an excellent method for a linear amplification of short ssDNA.

1.3. Type-II restriction endonucleases

Restriction endonucleases (REs) are enzymes present in the Restriction-Modification (RM) System of bacteria and achea, having the role of protection of their hosts from virus (bacteriophage) infection, by cleavage of foreign DNA.[6–10] REs recognise and cleave specific sequences of dsDNA, which generally are palindromic sequences of 4 to 8 bp. The cleavage occurs at both strands at the phosphodiester backbone in a symmetrical way, which can yield 5'-overhanging, 3'-overhanging or blunt ends DNA segments, most often having a 5'-phosphate terminus. Specific binding at the recognition sequence occurs thanks to several contacts between the protein and the substrate DNA. With the exception of very few cases, the formation of the specific complex between protein and DNA requires Mg^{2+} or Ca^{2+} cations, and numerous direct or water-mediated hydrogen bonds, to stabilize the specific interactions.[10]

So far thousands different REs have been discovered and studied and divided into four main classes, depending on the number and structure of subunits. Type II REs are the most relevant, due to their widespread use. Their main advantage over other types of REs is that they cleave DNA within or in close proximity of the recognition sequence, and they do not require ATP hydrolysis for the cleavage process.

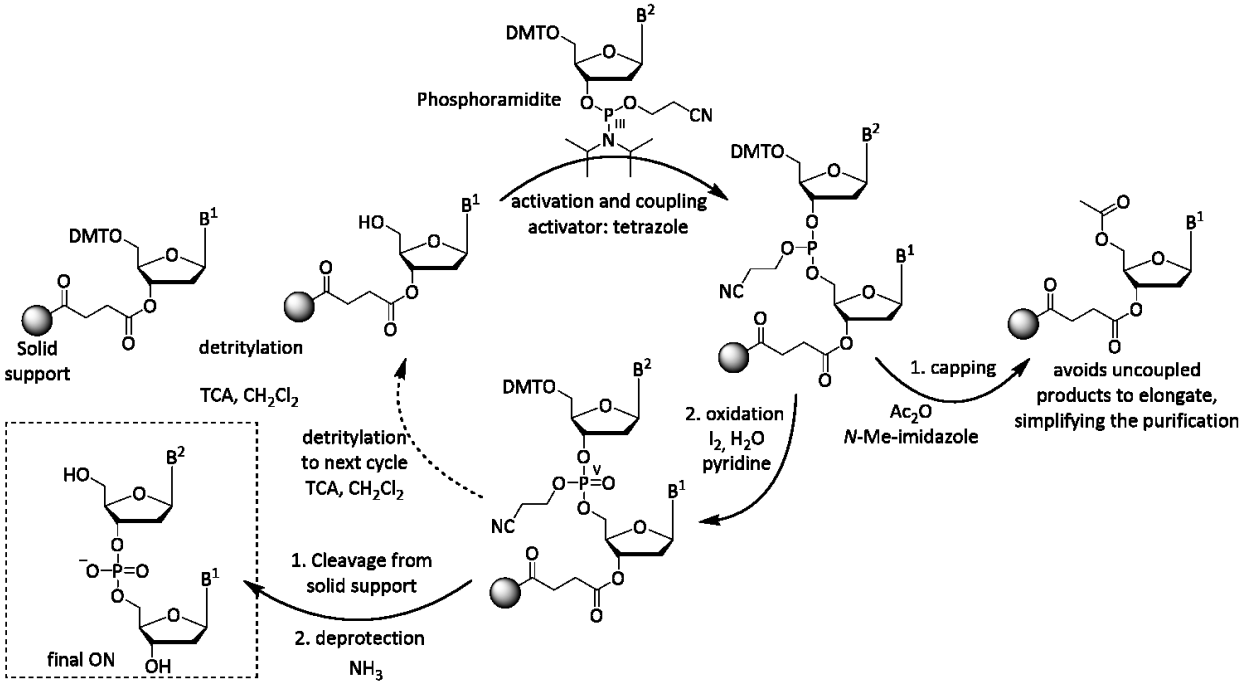
As REs are employed by bacteria to defend themselves by cleaving foreign DNA at specific sites, it is important for the hosts to protect their own genome to be hydrolysed by their own enzymes. The RM system of any organism couples a RE to a methyltransferase enzyme with DNA methylase activity, with the same sequence specificity. In this way, the host genome can be protected by methylation (generally introducing adenine-N^6-methyl groups), since methylated DNA is a very poor substrate for REs.[11]

REs are enzymes universally and widely used as reagents for the manipulation of DNA in laboratories, such as recombination or cloning. Being most of REs sensible to modifications present within their recognition sequence, these enzymes can be also exploited for DNA analysis and studies of incorporation of modifications. Several systematic studies were, hence, performed in order to understand the RE tolerance to DNA modifications of different nature and size.[12–15]

1.4. Chemical Synthesis of DNA

The chemical synthesis of ONs is conducted on a solid support (solid-phase synthesis), and performed by automated machines called DNA synthesizers.[16]

The building blocks for the synthesis are nucleotide-3'-cyanoethylphosphoramidites (CEP), presenting a reactive phosphorous (III) atom attached at the 3'-hydroxyl of the ribose moiety, a dimethoxytrityl (DMT) protection at the 5' hydroxyl group, and all other eventual amino or hydroxyl groups adequately protected. The synthesis proceeds in the 3' to 5' direction (opposite to enzymatic synthesis).



Scheme 1. General scheme of the solid-phase synthesis of ONs conducted by an automated DNA synthesizer.

Typically, the first nucleotide is attached to the solid support and the DNA synthesizer carries on all the steps of the synthesis automatically (all steps are shown in Scheme 1) and delivers all the nucleotide-phosphoramidites in the right order to obtain the final ON sequence decided by the user. ONs are then cleaved from the solid support and fully de-protected in a single step, and further purify, generally by standard desalting spin-columns or HPLC.

The solid-phase synthesis of ONs can be considered, somehow, complementary to the enzymatic approaches. The main advantage of the solid-phase synthesis is that the scale of DNA production can be much higher (e.g. µmol scale, against the nmol scale of the enzymatic methods). Moreover, some bulky groups may not be accepted by the polymerases, for example because they cannot be accommodated into their active site. In this sense, there are no limitations in the chemical synthesis.

On the other hand, it is very difficult to obtain DNA sequences via solid-phase synthesis which are longer than 100 nt, and almost impossible to prepare ONs with more than ca. 200 nt in acceptable yields. Moreover, most reactive chemical groups have to be adequaly protected, considering that the consequent deprotection step is performed on valuable and fragile synthetic DNA.

1.5. Synthesis and role of base-modified ONs

Base-modified DNA finds several applications in biotechnology, diagnostics and therapy, as well as in material science.[17–19] Most commonly, base-modifications are used to label the DNA, for imaging, or to bioconjugate or cross-link ONs to other biomolecules. Modifications are generally attached at the position 5 of pyrimidines, or at the positions 7 of 7-deazapurines. When attached to these positions, the modifications point out in the major groove of DNA where there is enough space to be accommodated. Minor groove modifications are rarer, due to the fact that the steric hindrance can prevent the double helix formation. However, examples of small minor groove modifications have been reported.[20,21]

The most common approach to insert base-modifications in DNA via enzymatic methods consists of the incorporation of base-modified dNTPs (dN^{R}TPs), through PEX or PCR. The employed DNA polymerases have to tolerate the modification present on dN^{R}TPs. Class B polymerases are generally used, being more capable to accept non-natural substrates, however, they can struggle in incorporating dNTPs bearing very bulky groups. The dN^{R}TPs building blocks are generally obtained by chemically modifying the nucleotide, to be further derivatized through 5'-*O*-triphosphorilation. A second very useful approach consists of subjecting dN^{R}TPs bearing reactive groups to a second derivatization performed directly on the triphosphate, which reaction conditions have to be tolerated by the 5'-*O*-triphosphate group. The major example of this approach is the use of a metal-catalyzed cross-coupling reaction on 5-iodo- **dC^{I}TP** and **dT^{I}TP** or 7-iodo-7-deaza- **dA^{I}TP** and **dG^{I}TP**. A direct modification of base-halogenated dN^{I}TPs though water-phase Suzuki, Heck or Sonogashira cross-coupling reactions can allow the introduction of fuctional groups which would not survive the triphosphorilation step, and facilitate the scope of many studies, using a single starting material.[22]

This approach is not possible for the production of the more sensitive base-modified phosphoramidites, for their employment in chemical synthesis of ONs. In this case, modifications have to be introduced on the nucleosides before the final 3'-*O*-phosphitylation step. The high reactivity of P^{III} in phosphoramidites further limits the incorporation of reactive groups through solid-phases synthesis, unless a protecting group is used.

1.6. CuAAC as a powerful tool for DNA post-synthetic modification

As previously seen, common approaches to the production of base-modified DNA can have important drawbacks. The incorporation of bulky motifs or substrates bearing reactive or labile groups, both via enzymatic and chemical synthesis, is often not practicable due to several inconveniences. For an efficient modification, it is often preferable to first produce nucleic acids bearing small biorthogonal chemical groups, and then perform a second chemical reaction directly on the functionalized nucleic acid; a process known as post-synthetic functionalization.[23] Therefore, it is of crucial importance to have available methods which involve an easy introduction of reactive groups into DNA, tolerated by the DNA construction process. Thus, a subsequent efficient and chemo-selective modification at these sites.

Among all the developed approaches for a post-synthetic functionalization of nucleic acids, the copper-catalysed alkyne-azide cycloaddition (CuAAC) has received an outstanding interest. It is a fast high-yielding reaction, often referred to as "click reaction".[24–26] Terminal alkynes and azides have the great advantage of being compatible and inert within living systems. Moreover, the triazole ring resulting from their reaction is chemically stable in physiological media and considered to be not toxic. Therefore, this reaction is a powerful biorthogonal tool. So far, many alkyne-containing nucleoside phosphoramidite building blocks[27–33] as well as alkyne-[12,30,34,35] or azide-linked[36,37] dNTPs were developed for the synthesis of clickable DNA.

1.6.1. Alkyne base-modifications for post-synthetic labeling of DNA via CuAAC

Generally, reactive terminal alkynes are directly attached at the nucleobase (at the C-5 position of pyrimidines or the C-7 position of 7-deazapurines), and ethynyl building blocks for both enzymatic and solid-phase synthesis of oligonucleotides are easily available. Uridine derivatives are the most accessible and, among them, 5-ethynyl-2'-deoxyuridine (**dU^{e}**, Figure 4) is the most widely employed. This nucleoside can penetrate cellular membranes, be phosphorylated intracellularly by natural kinases, and incorporated into the genome. It is, thus, an excellent tool for metabolic labelling through CuAAC.[38–40] Although the 5'-*O*-triphosphate of **dU^{e}** (**dU^{e}TP**) was shown to be an excellent substrate for DNA polymerases,[41] the click reaction on EdU-modified DNA can be rather difficult and inefficient, due to steric hindrance caused by the closed proximity

of the reactive alkyne and the nucleobase. Moreover, during the solid-phase synthesis of EdU-modified DNA, a partial hydration of the alkyne yielding the 5-acetyl uridine derivative occurs, forcing the use of a silyl-protected alkynyl-phosphoramidite. A second chemical step (deprotection) is necessary on precious modified DNA.[42]

The 5-(octa-1,7-diynyl)-modified 2'-deoxyuridine nucleoside (**dU°**, Figure 4) was developed in order to increase the CuAAC efficiency. The relative 5'-*O*-triphosphate (**dU°TP**)[34] and 3'-*O*-phosphoramidite[29,31,43] are now widely used for the production of clickable DNA. Although the CuAAC on octadiynyl-modified DNA is very efficient, and the method was used for high-density modification of DNA through click reaction,[34] it is important to notice that **dU°TP** is a not an optimal substrate for DNA polymerases. Its incorporation into DNA can be problematic due to the high hydrophobicity of the linked alkynyl group. The alkynyl bond between the tether and the nucleobase facilitates its synthesis, which can be easily performed using a Sonogashira cross-coupling reaction between 5-iodo-2'-deoxyribouridine and octa-1,7-diyne (Figure 4). However, the rigidity of this connection somehow restricts the flexibility of the whole linker, which contains only 4 aliphatic bonds with free rotation around their axes.

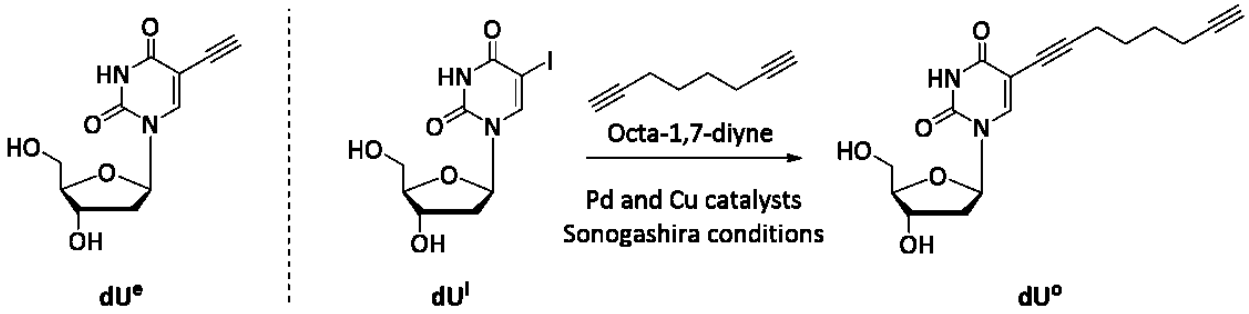


Figure 4. Structures of **dU^e**, **dU^I** and **dU°**, and scheme of synthesis of **dU°**.

1.7. Genome editing

Gene editing serves to engineer the genome of living cells, through replacement, insertion, modification or deletion of specific DNA sequences, as well as at single nucleotide level. Therefore, site-selective dsDNA cleavage became one of the most ambitious goals for biochemists, being a key step in gene editing, when followed by non-homologous end joining recombination (NHJR) or homologous recombination (HR) repair.[44–46]

HR is a DNA repair pathway, used by cells in order to repair potentially lethal dsDNA breaks. A homologous chromosome serves as a genomic template and the cell can find there the information to be copied back into the damaged DNA. Alternatively, a user-provided template dispensed for repair by HR, can lead to the introduction of desired sequence changes.

NHEJ is a repair mechanism that does not use a genomic template but directly ligates the two ends of the sniped DNA. Therefore, it is inaccurate and often leads to localized small changes in the sequence or insertions, provoking gene inactivation.

Targetable nucleases, or other systems able to accomplish site-selective DNA cleavage are often referred to as "molecular scissors". They are of great demand for direct and precise DNA mutations and sequence changes. These DNA cutters, ideally, must accomplish two different features: high cleaving ability and site-selectivity. They will be elucidated in details in the following chapters.

1.7.1. DNA-targeting specific cutters with DNA-protein interaction based selectivity

In the past three decades, most efforts in targeting gene editing, hence to introduce breaks in dsDNA with high specificity, have been focusing on engineering DNA-protein interactions. As a first example, zinc fingers nucleases (ZFNs) were developed as highly specific restriction enzymes. ZFNs are obtained combining a *Fok*I nuclease domain with a DNA recognition domain, which consists of three to six small proteins, called zinc fingers (ZFs). ZFs are able to recognize and bind DNA triplets.[47,48] *Fok*I is a non-specific Type II restriction endonuclease. At each finger, a Zinc atom is coordinated by a Cys_2His_2 domain, and the resulting aggregate is directly responsible of the protein-DNA specific interaction.[49] Most of the possible DNA triplets can now be targeted by

designed ZFs. Their combinations allow the use of these nucleases for the selective cleavage of a broad range of target sequences.[50–52]

Transcription activator-like effector (TALE) nucleases (TALENs) are very similar systems to ZFNs. The same non-specific *Fok*I nuclease domain is fused to a customizable DNA-binding domain.[53] The central recognition domain of TALE uses tandem repeats of 33 to 35 amino acid residues long for single base-pair recognition in target DNA. The so-called repeat variable di-residues consist of two amino acids situated at the 12 or 13 positions of each repeat and is responsible of base-pair specificity in DNA binding by TALE.[54,55]

More recently, it has been discovered that several bacteria and most archaea are able to detect and silence nucleic acids coming from foreign organisms, such as viruses or plasmids.[56] These microbes use CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats)/Cas (CRISPR-associated) systems, to excise short DNA segments from invading organisms and integrate them into their own genome, in between copies of a repeated sequence. At this point, the microbes are able to recognize the integrated sequence within the invader nucleic acids, and silence them, even in the case of a further invasion. The whole system relies on a single-guide-RNA (sgRNA) bound to a Cas protein. This protein contains two active nuclease domains, each cutting a strand of the target DNA. The sgRNA can be produced by transcription *in vitro* or *in vivo*, and its sequence can be engineered to bind virtually any target dsDNA.[57–61] So far, the revolutionary CRISPR/Cas systems are the most reliable gene-targeting nucleases, presenting several advantages with respect to ZFNs and TALENs. A single endonuclease protein is required, while more complicated dimer systems of endonuclease-binding proteins complexes are required in the other cases. Moreover, the guide sgRNA is easy and relatively cheap to synthesize. Its gene-recognition is based only on the well understood Watson-Crick base pairing. Clinical use of CRISPR/Cas systems as tools for manipulation of human genome in gene editing and gene therapy is now the "hot-topic" in Biology and Biochemistry.

1.7.2. Oxidative artificial metallonucleases (AMNs)

Similarly to engineered endonucleases based on DNA-protein interactions, fully synthetic AMNs have also attracted great attention as tools to introduce specific dsDNA breaks.

The study of chemical nucleases started in 1979, when Sigman *et al*. reported that a bis 1,10-Phenanthroline (Phen) Cupric complex $[Cu(Phen)_2]^{2+}$ (structure in Figure 5) was able to induce dsDNA cleavage in the presence of molecular oxygen and an external reductant, such as a thiol or ascorbate.[62,63]

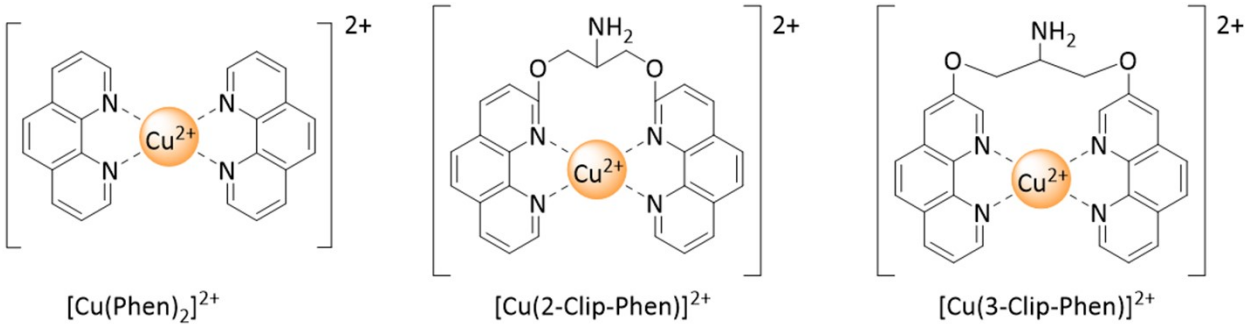


Figure 5. Structures of the main Cu-Phenanthroline based AMNs

Copper-phenanthroline complexes, and many analogues of them, immediately gained high importance and have been deeply studied as one of the most active class of chemical nucleases. The reductant is added exogenously, and reduces Cu^{II} to form a Cu^{I} complex $[Cu(Phen)_2]^{+}$, producing superoxide at the same time as a co-product. The active intermediates are formed by reaction of hydrogen peroxide with Cu^{I}, and consist of species in which hydroxyl or perhydroxyl radicals are bound to copper, such as Cu^{III}–OH or Cu^{I}– OOH species. The great advantage of these reactive species relies in their non-diffusible nature, inducing DNA damage only in close proximity to the spot where they are bound. Hydrogen peroxide can be added separately or, generally, it is formed *in situ* by the dismutation of the superoxide anion, which is generated by oxidation of the Cu^{I} complex in the presence of molecular oxygen.[64]

$[Cu(Phen)_2]^{+}$ binds DNA in the minor groove by intercalation. The interaction with DNA is reversible. The oxidative damage is directed to the sugar backbone of DNA, mainly to the C1' position[65] and, in lower amount, to the C4' and C5' positions[66] of the ribose units. DNA cleavage is started by abstraction of hydrogen atoms at these positions of the ribose moiety, and is the result of a series of further elimination reactions. The 3' and 5' termini monophosphate esters

and 5-methylenefuranone are produced as the main cleavage products in case of an initial C1' attack, while 3'-phosphoglycolate is a marker of an initial C4' oxidation.[64,67,68]

The best performances of copper-phenanthroline complexes towards dsDNA cleavage were always achieved when both phenanthroline ligands were present in the complex $[Cu(Phen)_2]^+$. Although $[Cu(Phen)]^+$, presenting a single Phen ligand in its structure, is still an active metallonuclease, its cleavage ability and specificity are significantly decreased with respect to the bis-Phen complex. Unfortunately, the association constant of the second phenanthroline to $[Cu(Phen)]^+$ is quite low, forcing the use of several equivalents of free ligand in order to have double chelation by two Phen and formation of the active species.

A solution to favour the right stoichiometry of two Phen per one Cu atom, is to link together the two needed ligands. To this aim, some clamped phenanthroline ligands (Clip-Phen) have been designed and studied, as an alternative to $[Cu(Phen)_2]^+$. Among all, 2-Clip-Phen (from now referred to as just Clip-Phen)[69] and 3-Clip-Phen[70] resulted to be the most reactive. In these complexes a serinol bridge covalently links the positions 2 (for Clip-Phen) or 3 (for 3-Clip-Phen) of the two Phen ligands through ether bonds, avoiding the dissociation of the second Phen moiety (Figure 5). As a consequence of higher stability and electron-donation by ether oxygen atoms, the oxidative nuclease activity of their Cu^I complexes is higher than the one of $[Cu(Phen)_2]^+$ (two times higher for Clip-Phen, and higher by a factor of 60 for 3-Clip-Phen).[71] However, it is important to mention at this stage, that a significantly increased cleaving activity is not always desired, as it can lead to a loss in site-specificity.

As a second important advantage, the free amino group present on the Clip-Phen serinol bridge allows for easy conjugation to systems which can improve the nuclease selectivity. Similarly to many other chemical nucleases, copper complexes of Clip-Phen are able to achieve high DNA oxidative damage. However, their sequence-selectivity was shown to be intrinsically low.

1.7.3. DNA-targeting specific cutters with chemical nucleases

In principle, whichever system that is able to specifically bind a sequence of DNA can be attached to a synthetic cleaving agent, to be converted into highly specific chemical nucleases.

Similarly to what was shown for protein nucleases in ZNFs, TALENs and CRISPR/Cas systems, artificial nucleases have also been coupled to DNA-targeting systems in order to improve their

site-specificity. Examples are metallonucleases linked to groove binders or intercalators, or short oligonucleotides, for DNA recognition through Watson-Crick base-pairing. Metal chelates such as Fe^{II}-EDTA[72–74] , Fe^{II}-porphyrines,[75,76] or Cu^{II}-Phen[77–79] have been attached to short oligonucleotides, to selectively target longer ssDNA containing the complementary sequences, achieving DNA oxidative damage in close proximity of the cutter agent. Cu^{II}-Clip-Phen complexes, on the other hand, were coupled to minor groove specific binders or intercalators, in order to target specific regions of dsDNA.[70,80–82]

A further and very interesting approach for the synthesis of site-selective AMNs consists of driving the cleaving agent with the use of a triplex forming oligonucleotide (TFO), able to target certain polypurine segments of DNA is a sequence-specific manner.

1.8. DNA triple helices

TFOs are short oligonucleotides able to bind a oligopurine segment of a DNA duplex with formation of a DNA triple helix (H-DNA).[83] In this way, a TFO acts as a highly sequence-specific DNA ligand, intercalating in the major groove of dsDNA and interacting with the oligopurine strand with formation of specific hydrogen bonds, known as Hoogsteen or reverse-Hoogsteen base-pairing. The sequence selectivity of the triplex formation lies in the specificity of the Hoogsteen base pairing formation, similarly to Watson-Crick base pairing in dsDNA. The main Hoogsteen or reverse-Hoogsteen hydrogen bonding structures are reported in Figure 6A and 6B.[84,85]

Three different classes of triplexes are known, in dependence of the type of hydrogen bonding or the orientation of the third strand (the TFO) with respect to the purine motif of the targeted DNA duplex. In a first case, an oligopyrimidine TFO binds the dsDNA purine strand in a parallel way (same 5' to 3' orientation). A thymidine nucleobase binds against an adenosine, forming a TA-T motif, and a cytidine against a Guanosine, forming a CG-C motif. In this latter motif, to enable the Hoogsteen bonding, the cytidine of the third strand must be protonated at the N^3 position, requiring then an acidic pH. The motif is thus indicated as a CG-C^+. Antiparallel triplexes form when an oligopurine TFO binds through reverse-Hoogsteen hydrogen bonds the purine strand of the dsDNA with opposite orientation and formation of TA-A and CG-G triplets. Adenosine nucleotides can be replaced by thymidine in this class of triplexes, leading to the formation of a stable TA-T antiparallel triplet.[86] Antiparallel triplex formation does not require acidic pH, since no cytidine is present within the TFO sequence. A poly-GT oligonucleotide can also act as a TFO by forming either an antiparallel triplex through reverse-Hoogsteen CG-G and TA-T triplets, or a parallel triplex by forming Hoogsteen CG-G and TA-T triplets. These triplexes are less stable and, therefore, less studied.[87]

In all cases, the binding is extremely sequence-specific and, generally, no mismatches are tolerated. Effects of mismatches in the Hoogsteen pairing were deeply investigated, by both molecular modelling and thermal dissociation experiments.[86,88,89]

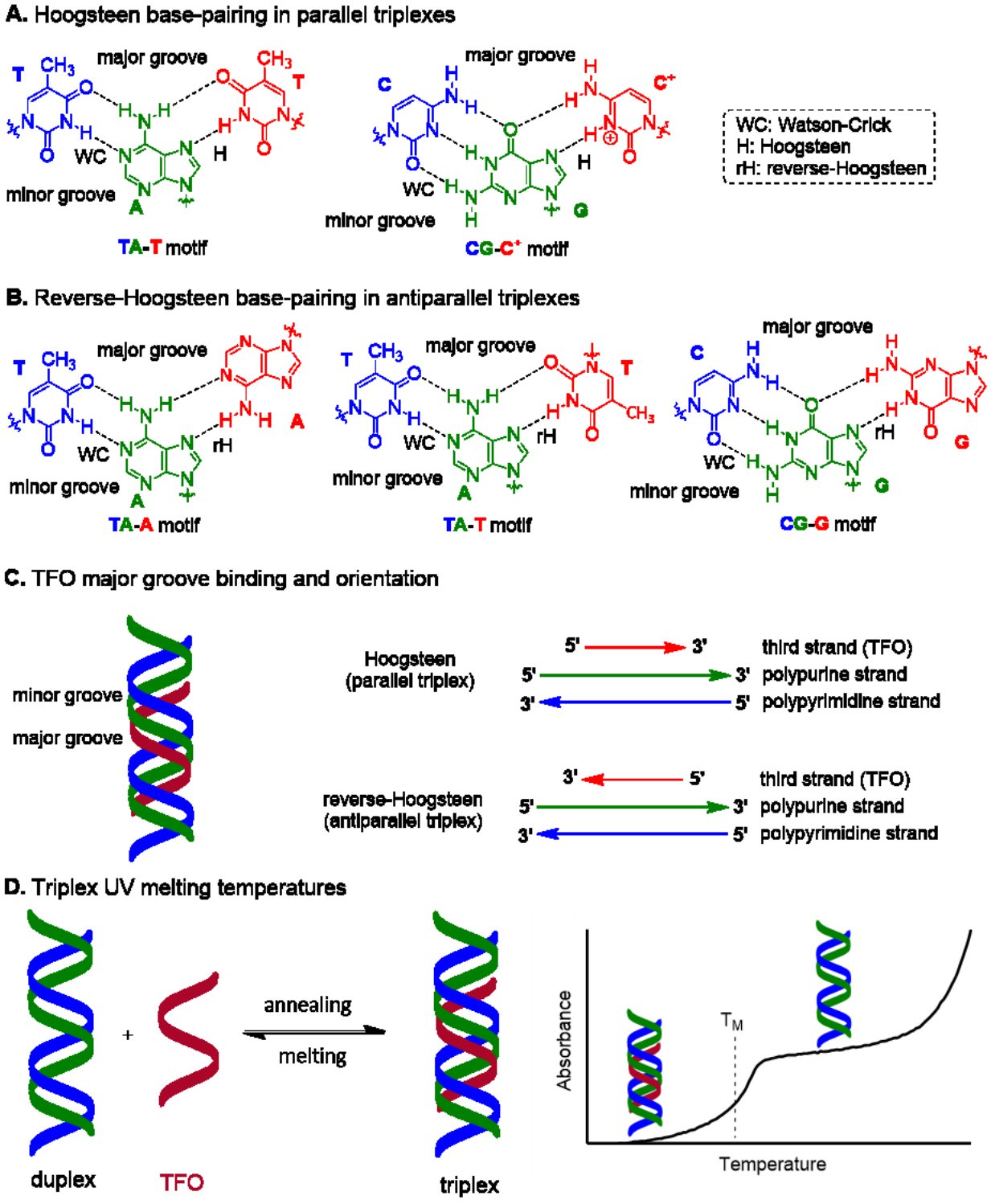


Figure 6. A. Hoogsteen (H) hydrogen bonding for the formation of TA-T and CG-C+ motifs in parallel triplexes (WC indicates the Watson-Crick base pairing inn the target duplex); **B.** Reverse-Hoogsteen (rH) hydrogen bonding for the formation of TA-A, TA-T and CG-G motifs in antiparallel triplexes (WC indicates the Watson-Crick base pairing inn the target duplex); **C.** The TFO binds in the major groove of the target

dsDNA, either with the same 5’ to 3’ orientation of the purine strand of the duplex, or with opposite orientation.

Both the length of the binding TFO, and the number of mismatches highly influence the triplex stability. The thermal stability of DNA triplexes can be measured by UV melting curves, similarly to the thermal stability of DNA duplexes. Absorbance at increasing temperatures is plotted against the temperature, in order to obtain a melting profile. The second derivative of the curve reach the value zero in the melting temperature of the triplex (T_M). At this temperature an equimolar concentration of triplex and duplex is present in solution, following the equilibrium between triplex and duplex + TFO (Figure 6D).

1.8.1. Antigene strategy

The high specificity in DNA recognition by a TFO during a triplex formation suggested the use of these tools for regulation of gene expression: a tactic now known as antigene strategy.[90–92] This technique presents some advantages with respect to the antisense oligonucleotides or ribozymes, which are targeting directly messenger RNAs (mRNAs). The inhibition of protein synthesis is very difficult to achieve without a blockage of the corresponding gene (e.g. with antisense agents), since many thousands RNA molecules are present and constantly replenished. Vice-versa, acting directly on the gene, only two alleles have to be “switched-off”.[92] Also targeting directly the catalytic action of the final protein, using conventional drugs, is difficult and less efficient.

Short oligopurine segments are quite frequent in the genome (e.g. statistically a oligopurine segment of 10 nucleotides can be found once in 1000 bases, while longer oligopurine segments, such as 25 nt, can be expected to be found once in the genome, resulting useful for selectivity in gene switch-off. The triplex formation mainly changes (or prevents) the capacity of the specific proteins needed during the transcription process to recognize and bind the targeted gene (Figure 7A), leading to a modulation of the transcription itself. The DNA triple helix stability is comparable to, or even higher than, the one of the initiation complex (forming between DNA, the activator protein binding the promoter region, and transcription factors). Mechanisms of inhibition of transcription may involve a triple helix formation in the promoter region, interfering with the activator protein. Triplex formation can also prevent the transcription factors or the RNA

polymerase to bind DNA. A last case involves the alteration of DNA flexibility or geometry, preventing the right assembling of the initiation complex (Figure 7).[92,93]

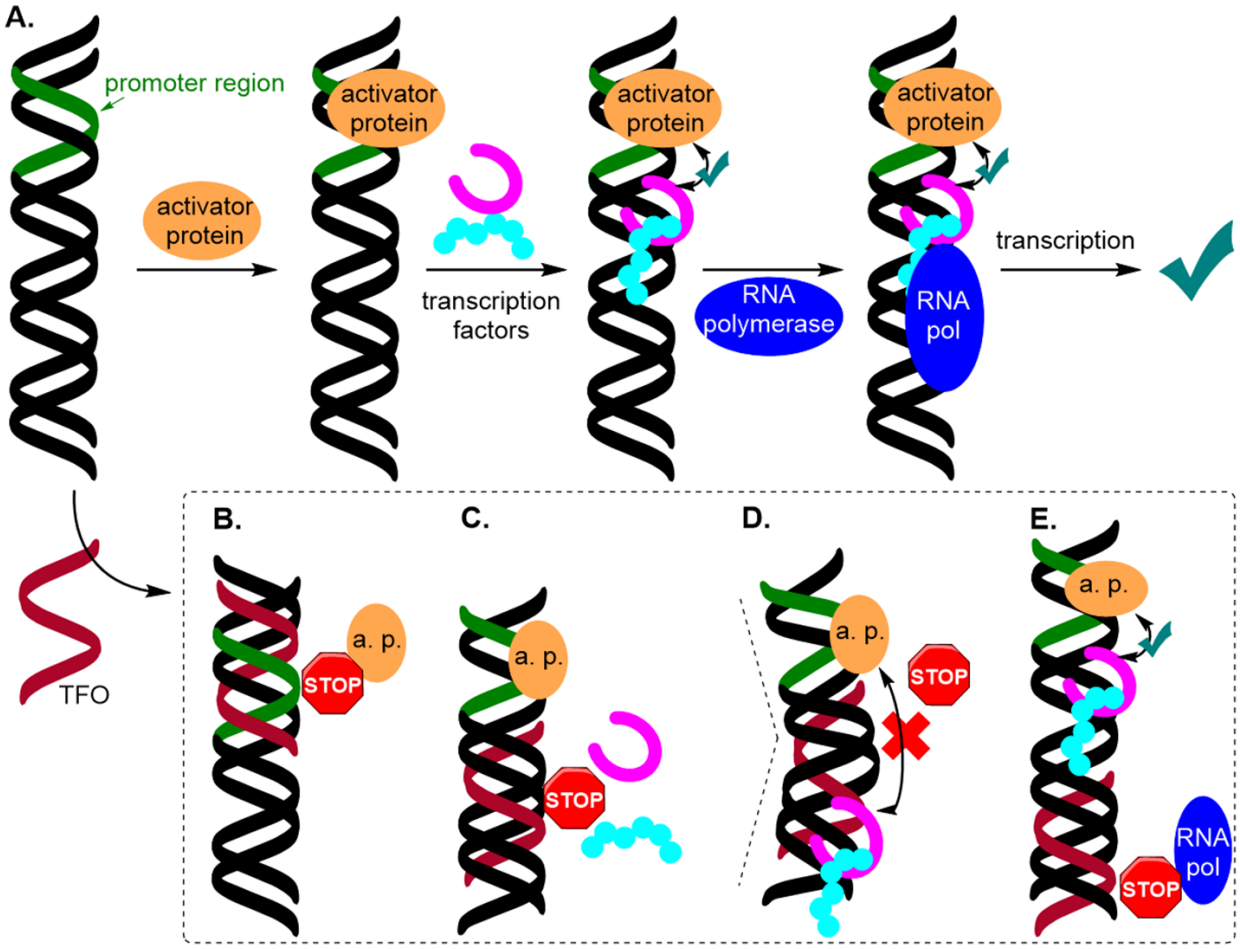


Figure 7. Physical blockage of transcription by triplex formation; **A.** Transcription process; **B.** Inhibition of transcription by triplex formation in the promoter region, preventing the binding of the activator protein; **C.** Triplex formation preventing the binding of transcriptional factors; **D.** The triplex formation alter DNA geometry, preventing a good interaction between the activator protein and transcriptional factors. **E.** The triplex formation directly prevent the binding of the RNA polymerase.

The crucial importance of the regulation of gene expression led to the development of several chemical modifications at the TFO, aiming to improve the gene knockout. Most of reported modifications enhance the triplex stability, mainly, with base-modifications which are aimed to higher the strength of the Hoogsteen hydrogen bonding. The TFO resistance to hydrolysis or cleavage by restriction enzymes can also be enhanced through modifications at the

phosphodiester backbone or at the sugar moiety.[94] Also a 2'-*O*-methylribose modification makes the TFO endonuclease-resistant, and thermodynamically stabilizes the triplex as well.[95]

Reactive TFO derivatives were widely studied as important and prominent tools for the manipulations of genes. Their ability of site-directed mutagenesis or suppression of the expression have been largely studied as an alternative or variation of the antigene strategy. To this aim, TFOs have been coupled with different classes of DNA damaging agents, such as psoralene[96] or other photoactivatable agents[97], bleomycin[98], or even enzyme nucleases.[99,100] Among all developed hybrids, TFOs conjugated to artificial metallonucleases (TFO-AMNs) were shown to be valuable tools for site-specific DNA oxidative cleavage, thus valuable synthetic DNA-targeting molecular scissors.

1.9. TFO-directed site-specific DNA oxidative cleavage

DNA triple helices form with high specificity and affinity, confirming TFOs as good targeting domains in the construction of DNA-targeting nucleases. A first attempt to exploit the sequence-selectivity of triplex formation in DNA oxidative cleavage was reported by Moser and Dervan, who connected the Fe^{II}-EDTA DNA cutter to the 5' end of a homopyrimidine probe able to form a triplex with the target dsDNA.[101] Their TFO-Fe^{II}-EDTA molecular scissor was found to produce sequence-specific cleavage patterns, with efficiency depending on probe concentration, pH, cations concentration and number of mismatches in the TFO for the triplex formation.

Subsequently, also the $[Cu(Phen)_2]^{2+}$ complex was conjugated to a TFO able to target a single site in the Simian virus 40 genome.[77,102] The Phen ligand was tethered through polymethylene thiophosphate spacers to the 5' end of the TFO, and different lengths of the linker were investigated, as well as different experimental conditions. In the best cases, DNA oxidative cleavage occurred at the specific target site, with much higher efficiency with respect to the previously described Fe^{II}-EDTA-conjugates (70% of TFO-Cu^{II}-Phen against 25% of TFO-Fe^{II}-EDTA). With a similar TFO-AMN hybrid, in which a Phen ligand was connected to the TFO through a polyamide linker directly attached to a 5'-thiophosphate group, Shimizu et al. proved the cleaving ability towards a target duplex to be directly correlated to the thermal stability of the triplex.[103]

2. Aims of the Thesis

1. Development of site-specific artificial metallonuclease (AMN) cutters of DNA, based on conjugates of triplex-forming oligonucleotides (TFOs) and Cu complexes of phenanthroline ligands

1.1. Design and synthesis of building blocks for enzymatic and chemical synthesis of oligonucleotides containing clickable groups, phenanthroline ligands or Cu-phenanthroline complexes

1.2. Design and optimization of new linkers for clickable nucleoside triphosphate (dNTP) building blocks, with improved combination of substrate activity with polymerases, and reactivity in CuAAC reactions

1.3. Design and synthesis of AMN-TFO conjugates bearing reactive complexes at the 3'- or 5'- end, or in the middle of the strand

1.4. Study of triplex formation and efficiency and specificity of cleavage of target DNA duplexes with AMN-TFO conjugates

2. Systematic study of competitive incorporation of series of 5-substituted pyrimidine and 7-substituted 7-deazapurine nucleotides in primer extension, in the presence of natural dNTPs

2.1. Development of assays and systematic study of competitive incorporations of modified nucleotides in different ratios with natural nucleotides

2.2. Detailed enzyme kinetic study of incorporation of modified nucleotides

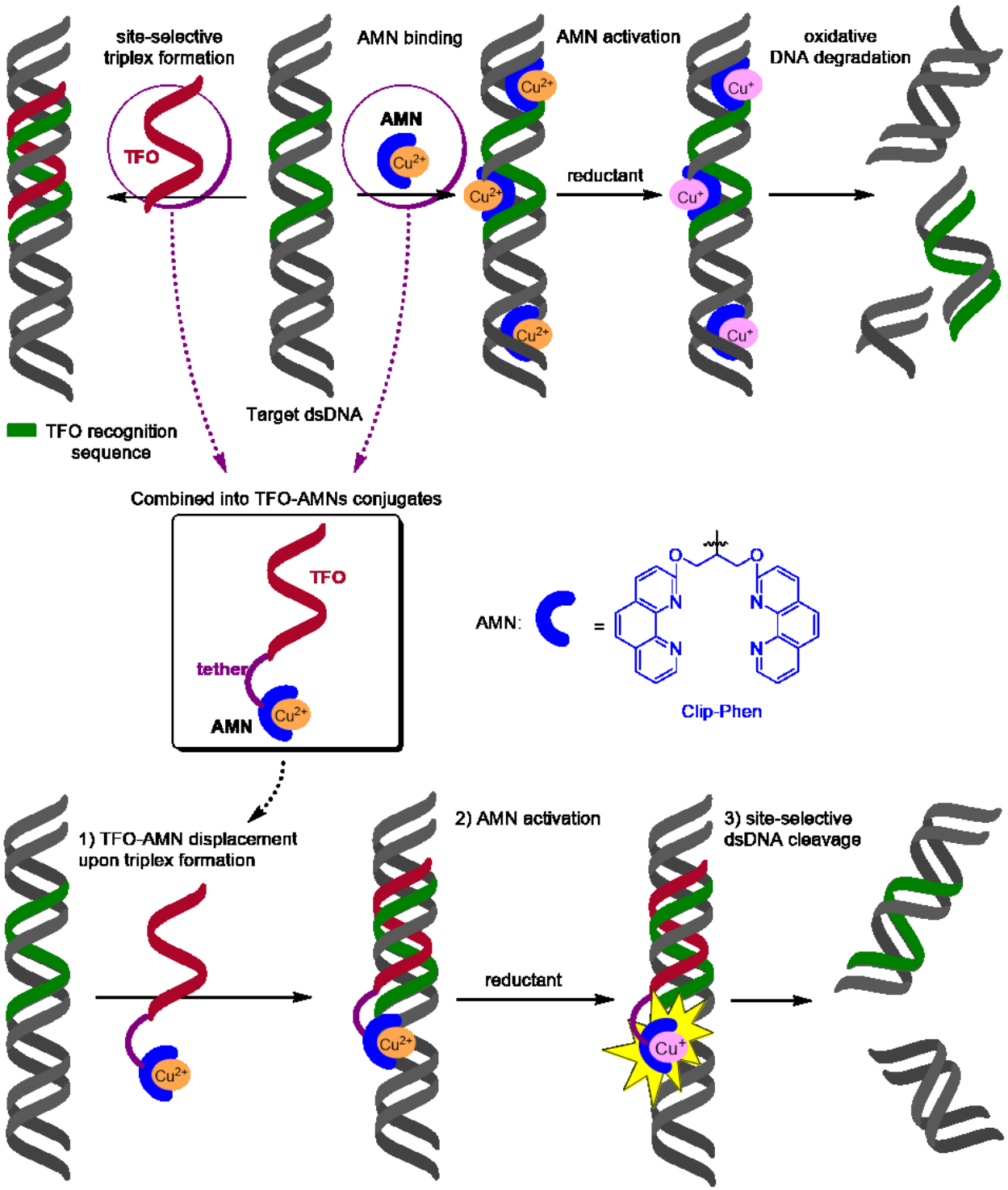
2.3. Mechanistic explanation of increased substrate activities of aryl-modified dNTPs

3. Results and discussion

3.1. DNA-targeting molecular scissors

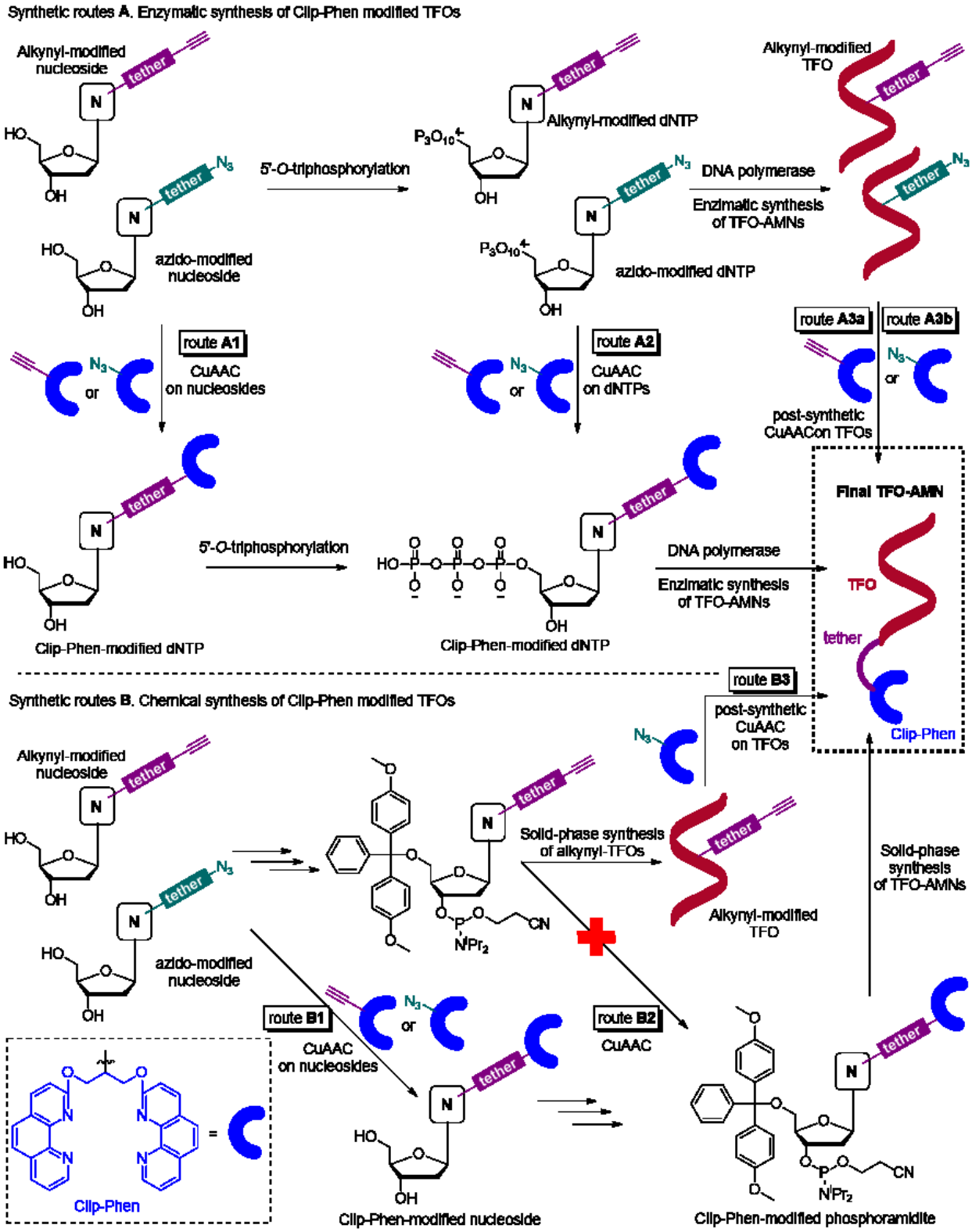
3.1.1. Strategy and design of TFO-Clip-Phen specific cutters

In order to achieve site-selective dsDNA cleavage, it was necessary to design specific cutters able to recognize certain sequences in the targeted DNA. TFOs can offer promise as tools for gene targeting, allowing for the recognition of a specific site in the target DNA, exploiting the high sequence-selectivity of triplex formation. Therefore, the selected approach for the production of DNA specific cutters, combined the metallonuclease technology with DNA triple helices (Scheme 2).[104] TFO-AMN conjugates were designed in order to allow a good displacement of the nuclease in close proximity to a selected sequence of the target dsDNA to be cleaved, by guidance of the TFO, upon triplex formation.



Scheme 2. General strategy for dsDNA cleavage by TFO-AMNs conjugates. The site-selectivity of the cleavage is achieved upon triplex formation.

3.1.1.1. Synthetic routes for the production of the TFO-AMNs conjugates



Scheme 3. All possible strategies for the synthesis of the TFO-AMN conjugates.

Copper(II)-1,10-phenanthroline complexes are suitable AMN to be employed, since they oxidative cleave dsDNA only upon previous activation, by Cu^{II} reduction to Cu^{I}. In this way it is possible to allow triple helices to form in a first moment, leaving undisturbed the targeted (but also the non-targeted) DNA. After triplex formation, the AMN can be activated by a reducing agent, when already displaced in the spot where the dsDNA cleavage is desired (Scheme 2).

The higher stability, but also cleaving activity, of Cu^{I}-Clip-Phen with respect to Cu^{I}-Phen drove the interest towards the possibility to tether this complex to a TFO, for the synthesis of DNA-targeting molecular scissors. Moreover, the free amino group on the serinol bridge of the Clip-Phen ligand allows an easy linkage to the oligonucleotides.

The 3-Clip-Phen Cu complex was thought to be too active, with increased possibility of off-target DNA damage.

As discussed in the introduction, the copper-catalysed alkyne-azide cycloaddition is a very reliable tool for efficient modification of ONs, being quick, high yielding and chemoselective. Therefore, it was considered as a first choice to covalently link the AMN to TFOs.

In a first place, Clip-Phen analogues containing a terminal alkyne or an azido group within their structures are required, to be used in CuAAC reactions. At this point several approaches can be considered in order to attach these Clip-Phen derivatives to the TFOs (Scheme 3).

A first attempted synthetic route consisted of the production of building blocks already bearing the Clip-Phen ligand for a direct synthesis of Clip-Phen-ONs, either enzymatically or chemically on solid support.

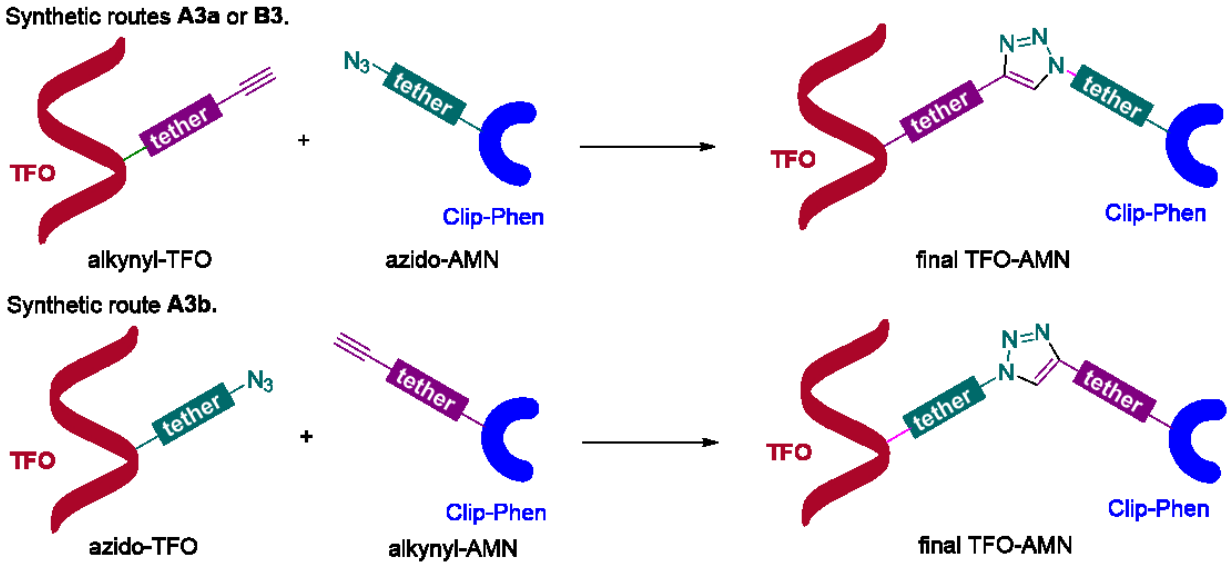
For the enzymatic synthesis, Clip-Phen modified dNTPs are necessary (synthetic routes **A1** and **A2** in Scheme 3). Two possibilities were taken into consideration for the preparation of such dNTPs, since the Clip-Phen moiety can be attached to the nucleobase either after or before the 5'-*O*-triphosphorilation step. In particular, the CuAAC step can also be performed on easily available alkynyl- or azido- modified nucleoside, to be 5'-*O*-triphosphorylated in a second step (route **A1**). On the other hand, alkynyl- or azido-containing nucleosides can be 5'-*O*-triphosphorylated in a first step, and then the alkynyl- or azido- triphosphates can undergo a further CuAAC with Clip-Phen derivatives (route **A2**).

In a similar way, the Clip-Phen modified nucleosides can be converted in a few steps into the relative 3'-*O*-phosphormaidites, which can be used as monomers in the chemical synthesis of TFOs on solid support (synthetic route **B1**, Scheme 3). The opposite approach (synthetic route **B2**), would be based on a CuAAC step on alkynyl- or azido-modified 3'-*O*-phosphormaidites to yield Clip-Phen-modified phosphoramidites (in analogy of the enzymatic route **A2**). However, it is not possible to follow this route, since amidites are unstable in CuAAC conditions.

In addition, a slightly different approach involving the use of a post-functionalization step of ONs using the CuAAC reaction is also possible. The design of this further approach was dictated by the high probability of encountering difficulties in the synthetic routes **A1** and **A2**, or **B1**. As a matter of fact, the synthesis of Clip-Phen modified triphosphates could be challenging. On one side, the CuAAC step conducted in the presence of Cu^{II} could hydrolyse or strongly bind the triphosphate group. On the other side, the 5'-*O*-triphosphorylation of a nucleoside containing basic substituents such as 1,10-phenanthroline could be impossible. Moreover, even in case of success in the synthesis of the Clip-Phen modified dNTPs, it is not certain that the DNA polymerase would accept them and incorporate them into ONs. As discussed in the introduction, DNA polymerases can struggle incorporating dNTPs bearing very bulky or reactive groups. In a similar way, the reactive 1,10-phenanthroline moiety could disturb, if not completely prevent, the proceeding of the synthesis on solid support.

Therefore, an approach based on a post-synthetic functionalization of alkynyl- or azido ONs would be very useful to circumvent all these drawbacks. In this case, the main step consists of a CuAAC either between a pre-synthesized alkynyl-modified TFO and an azido derivative of the Clip-Phen ligand (synthetic route **A3a** or **B3**), or, in an opposite way, involving a CuAAC reaction between an azido-modified TFO and a Clip-Phen derivative bearing a terminal alkynyl moiety (Scheme 4).

The synthesis of alkynyl-modified ONs can be performed both enzymatically and chemically, employing alkynyl-dNTPs in the former case, and alkynyl-3'-*O*-phosphoramidites in the latter case. It is important to note that azido-modified ONs can be prepared only enzymatically, as azides are usually incompatible with the phosphoramidite chemistry employed during the solid-phase synthesis, probably due to their Staudinger reaction with the P^{III} atom of the amidite.[105]

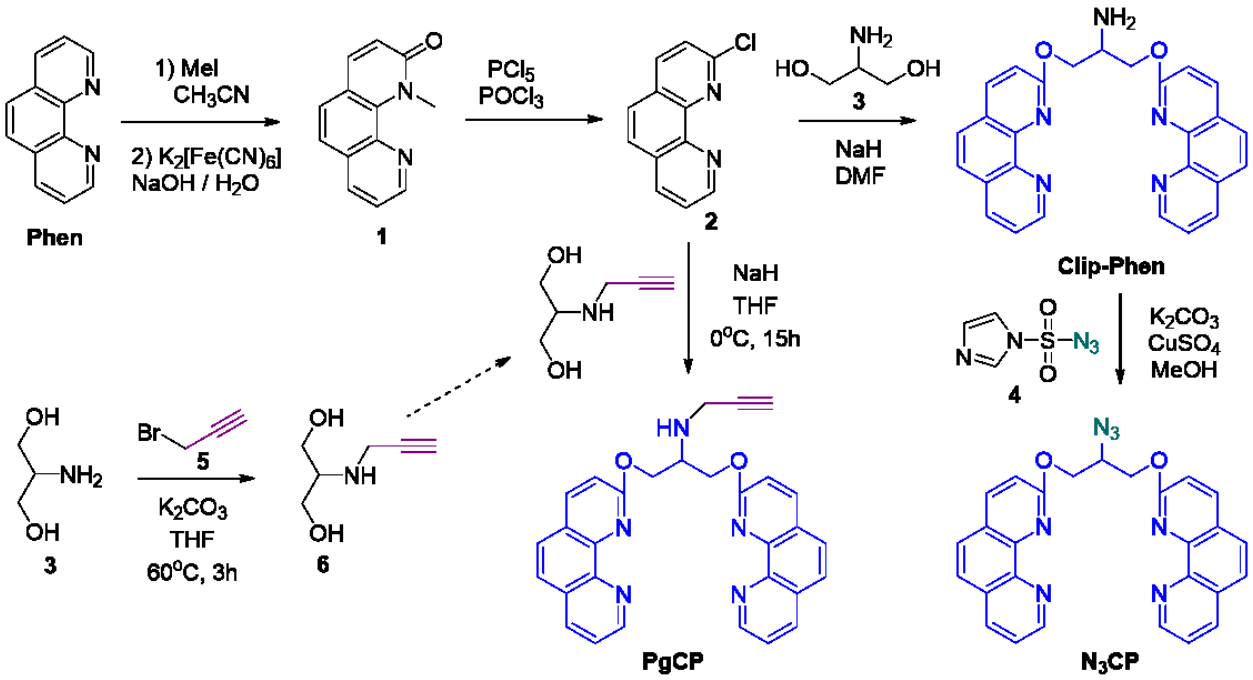


Scheme 4. Two possible approaches for the synthesis of Clip-Phen modified TFOs using a post-synthetic CuAAC reaction step.

3.1.2. Clamped-Phenanthroline derivatives for click ligation to TFOs

In order to conjugate the Clip-Phen moiety to dNTPs, phosphoramidites, or TFOs bearing alkynyl- or azido functional group via CuAAC, it was necessary to chemically modify the ligand, introducing a terminal alkyne or an azido group, useful for click chemistry.

The azido derivative **N_3CP** can be obtained by a facile and good-yielding functional group transformation on **Clip-Phen** (which synthesis was described by Pitie et al. and reported in Scheme 5)[69], using 1-*H*-imidazole sulfonyl azide (**4**) as a diazo-transfer reagent to convert the amino group on the serinol moiety to an azido group (Scheme 5).[106]



Scheme 5. Scheme of synthesis of the alkynyl- and azido- derivatives of Clip-Phen **PgCP** and **N_3CP**.

The preparation of a Clip-Phen derivative containing a terminal alkyne within its structure, useful for conjugation with azido-modified dNTPs, amidites or ONs, resulted slightly more difficult. Attempts to directly propargylate the amino group on the serinol bridge of **Clip-Phen** using propargyl bromide failed. After the reaction, a complex mixture was obtained due to the formation of double and triple substitution products on the serinol amino group, as well as

reactions of the propargyl bromide with the reactive nitrogen atoms at positions 1 or 10 of the 1,10-phenanthrolin motifs. Therefore, it was necessary to change strategy, and the Clip-Phen structure was assembled using a serinol moiety already bearing the propargyl group. For this, serinol (**3**) was reacted with a sub-stoichiometric amount of propargyl bromide (**5**) in a presence of a base, to ensure a mono-substitution. The obtained *N*-propargyl serinol (**6**) was further reacted with 2-chloro-1,10-phenanthroline (**2**) and sodium hydride to obtain the target derivative **PgCP** (Scheme 5).

3.1.2.1. Binding and cleaving abilities of N_3CP and PgCP Cu^{II} complexes

Although the Cu^{II} complex of **Clip-Phen** is known to bind very strongly DNA in its minor groove, the presence of different functional groups on its structure might influence or prevent this ability. Moreover, a difference in binding ability could also affect the activity of the complex in DNA oxidative degradation, resulting then useless for the construction of TFO-AMNs. A high improve in the cleaving ability is also not wanted, since it could affect the selectivity in the cleavage of the molecular scissors. For these reasons, it was necessary to evaluate both the binding and cleaving ability of the Cu^{I} complexes of the newly synthesized ligands **PgCP** and **N_3CP**, before proceeding to the synthesis of the TFO-AMN hybrids. These experiments were designed by Dr. Zara Molphy, and were performed by me at the Dublin Citu University (DCU), under the supervision of Prof. Andrew Kellett, during my one-month secondment in Dublin.

The binding ability was evaluated in experiments with type-1 topoisomerases: nuclear enzymes that regulate DNA topology.[107–109] Topoisomerase-I cuts one of the two strands of supercoiled DNA and induces relaxation, prior to reanneal the sniped strand. In the absence of DNA binders or intercalators, the Topoisomerase-I is able to read the DNA through all its length and fully relax a supercoiled plasmid DNA. However, if a DNA binder or a DNA intercalator is present, such as bis-1,10-phenanthroline Cu^{II} complexes, the read through of the topoisomerase is blocked and no DNA relaxation is observed. As the concentration of the complexes increases, the amount of the relaxed DNA decreases gradually, due to the inhibition of the topoisomerase by the complex binding. Under a certain complex concentration, the topoisomerase is completely inhibited, achieving no relaxation of supercoiled DNA. The binding ability of both **PgCP** and **N_3CP** Cu^{II} complexes was evaluated by incubating supercoiled DNA with increasing amounts of the two

complexes, thus leaving enough time for the intercalation or binding to occur efficiently. The topoisomerase-I was added to the solution in a second moment. A comparison to the binding ability of the standardly used bis-1,10-phenanthroline Cu^{II} complex was made.

Agarose gel analysis showed that both **PgCP** and **N_3CP** Cu^{II} complexes intercalated supercoiled DNA approximately as well as $[Cu(Phen)_2]^{2+}$ (Figure 8).

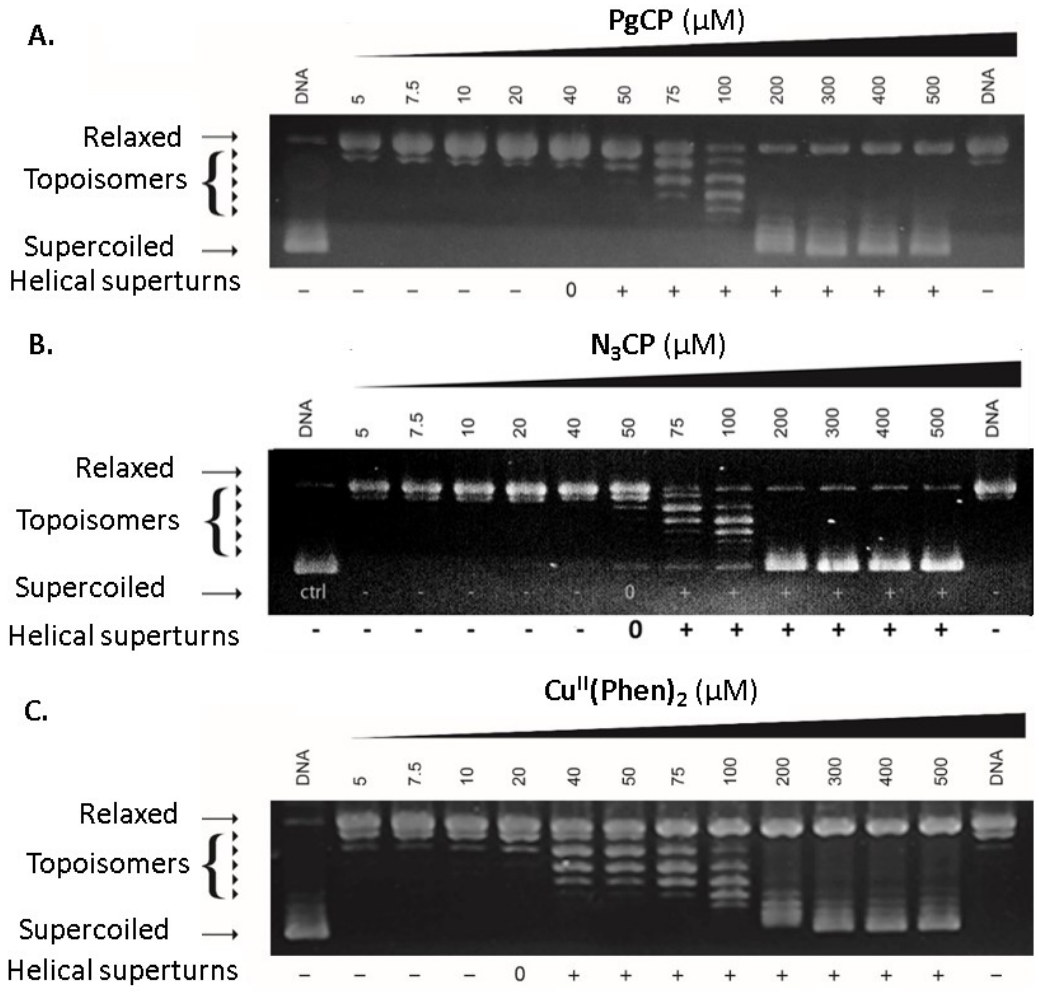


Figure 8. Agarose gel analysis of the binding ability of Cu^{II} complexes of **PgCP** and **N_3CP**.

Next, the ability of **PgCP-Cu^{I}** and **N_3CP-Cu^{I}** to oxidatively damage dsDNA was studied incubating supercoiled DNA at increasing concentration of the Cu^{II} complexes, followed by addition of Na-$_L$-ascorbate used as a reductant to activate the nucleases.

Agarose gel analysis showed that the two complexes are able to cleave dsDNA even at quite low concentrations. Moreover, controls showed that at the same concentrations of complex, no detectable DNA cleavage occurs in the absence of Na-$_L$-ascorbate, indicating that only the Cu^I complex is active (Figure 9A).

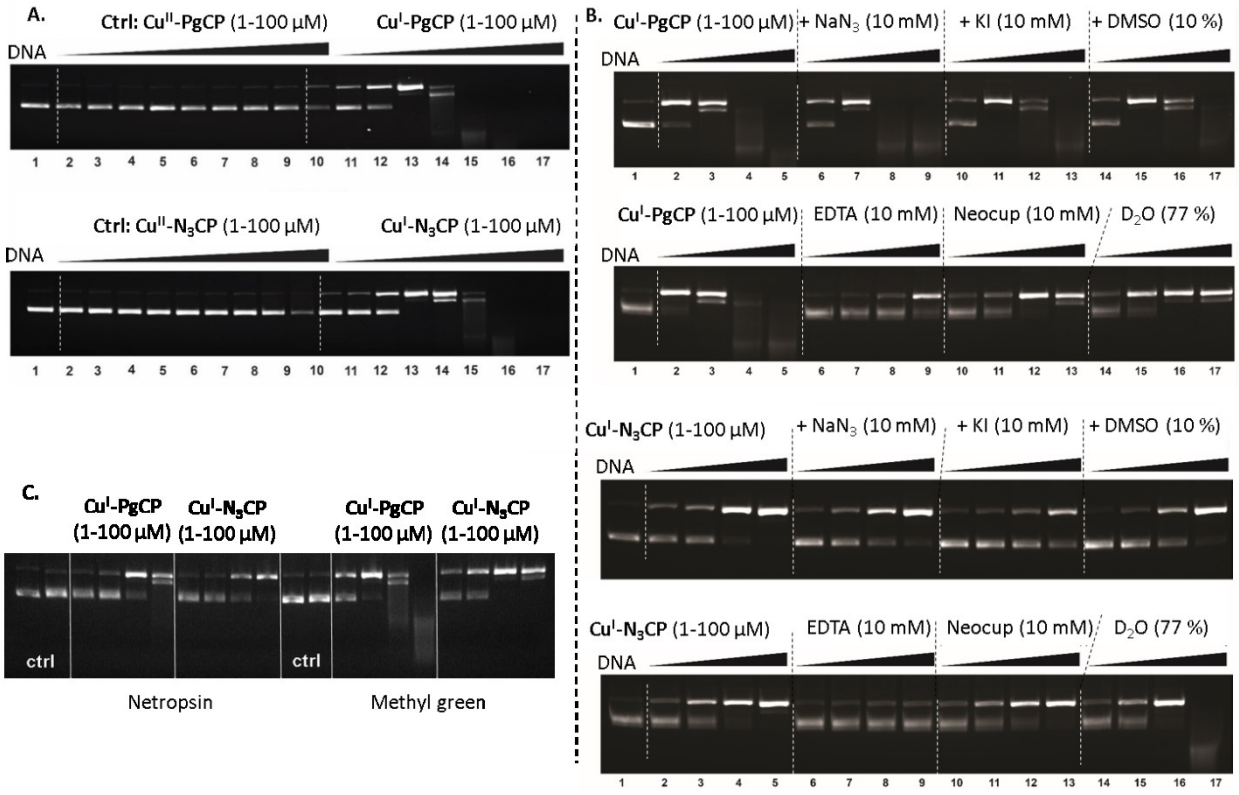


Figure 9. **A.** Agarose gel analysis of the cleaving ability of **PgCP** and **N_3CP** Cu^{II} complexes in the absence (Lanes 2-9) or presence (Lanes 10-17) of reductant; **B.** Agarose gel analysis of cleaving ability of **PgCP** and **N_3CP** Cu^I complexes in the presence of selective radical quenchers or stabilizers; **C.** Agarose gel analysis of cleaving ability of **PgCP** and **N_3CP** Cu^I complexes in the presence of specific groove binders.

To better understand the mechanism of cleavage of the two Cu^I complexes, and to acknowledge which reactive oxygen species (ROS) are involved in the dsDNA oxidative damage by the two complexes, the nuclease activity experiments were repeated in the presence of selective radical quenchers or stabilizers (Figure 9B).[107,110,111] Supercoiled DNA was, hence, incubated with the Cu^{II} complexes in the presence of NaN_3 (scavenger for singlet oxygen), KI (scavenger for hydrogen peroxide), DMSO (scavenger for •OH radicals), EDTA (scavenger for Cu^{II}), Neocuproine (scavenger

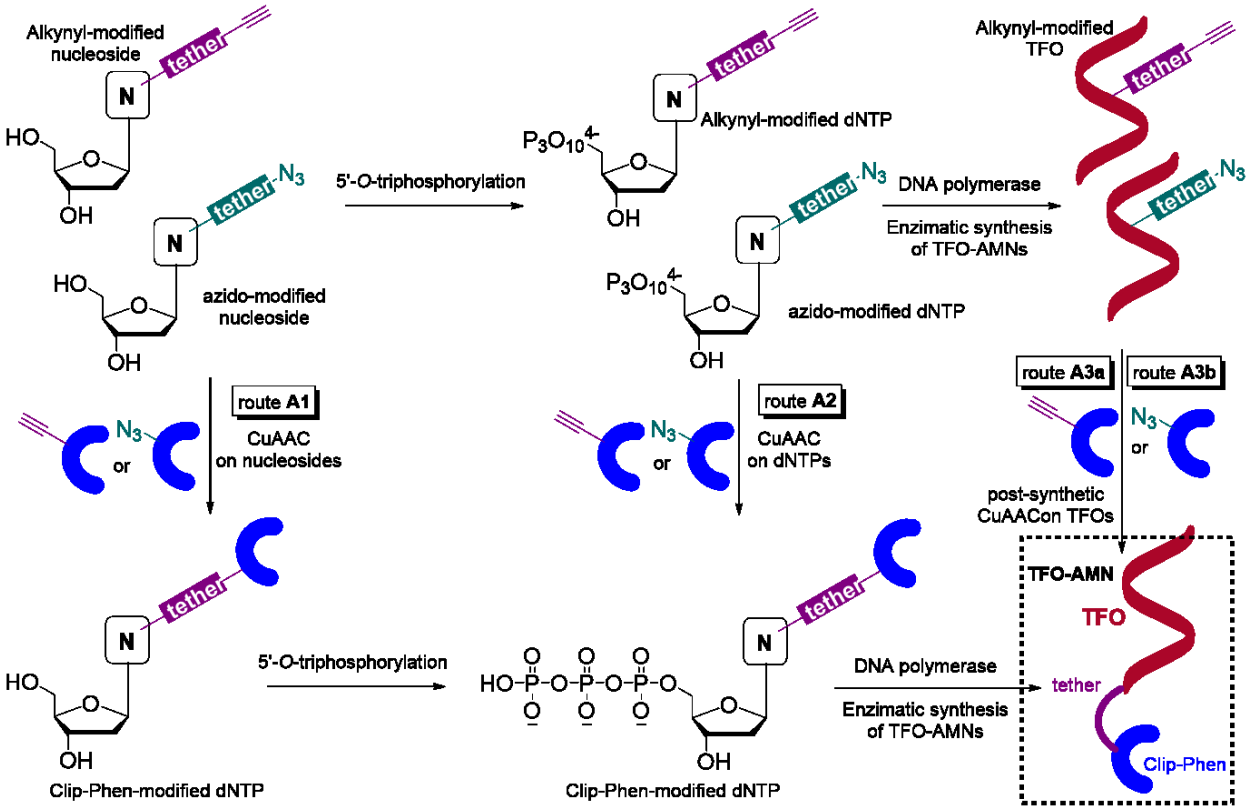
for Cu^{I}) and D_2O (singlet oxygen has longer life in deuterated water), to be then activated by addition of Na-$_L$-ascorbate. For both Cu^{I} complexes, the Neocuproine seemed to have the predominant effect, inhibiting the dsDNA damage at the highest extent. While KI and DMSO slightly slowed down the DNA cleavage by **PgCP-Cu^{I}** and **N$_3$CP-Cu^{I}**, other reagents such as NaN_3 or D_2O had no effect, implying H_2O_2 and •OH radicals to be the most effective ROS formed by the complexes while binding to the target DNA.

Finally, the metallonucleases were added to supercoiled DNA which was previously incubated in the presence of Netropsin (minor groove binder) or Methyl Green (major groove binder). These two specific groove binders block the access to one of the two grooves selectively, impeding the Clip-Phen complex to intercalate and, as a consequence, reducing its cleavage ability. This experiment in very useful to determine the preferential binding either to the minor groove or in the major groove of DNA.

Both Cu^{II} complexes of **PgCP** and **N$_3$CP** showed higher cleaving activity when Methyl Green was present, with respect to the activity in the presence of Netropsin, indicating their preference of binding into the DNA minor groove, in analogy to $[Cu(Phen)_2]^{2+}$ (Figure 9C).

This series of experiments proved **PgCP-Cu^{I}** and **N$_3$CP-Cu^{I}** to be valid AMNs to be employed in the synthesis of TFO-AMNs hybrids, being able to efficiently bind dsDNA and being very active towards the DNA oxidative damage. Moreover, the low cleaving ability of the Cu^{II} complexes of **PgCP** and **N$_3$CP** was also proved. The great importance of this result lies in the fact that TFO-AMNs bearing a Clip-Phen-Cu^{II} complex could be prepared and hybridized with the target dsDNA upon triplex formation without any DNA damage occurring until a reductant is added to the mixture, to activate the AMN.

3.1.3. Enzymatic approaches for the synthesis of the TFO-AMNs conjugates



Scheme 6. Enzymatic approaches to the synthesis of the TFO-AMNs conjugates.

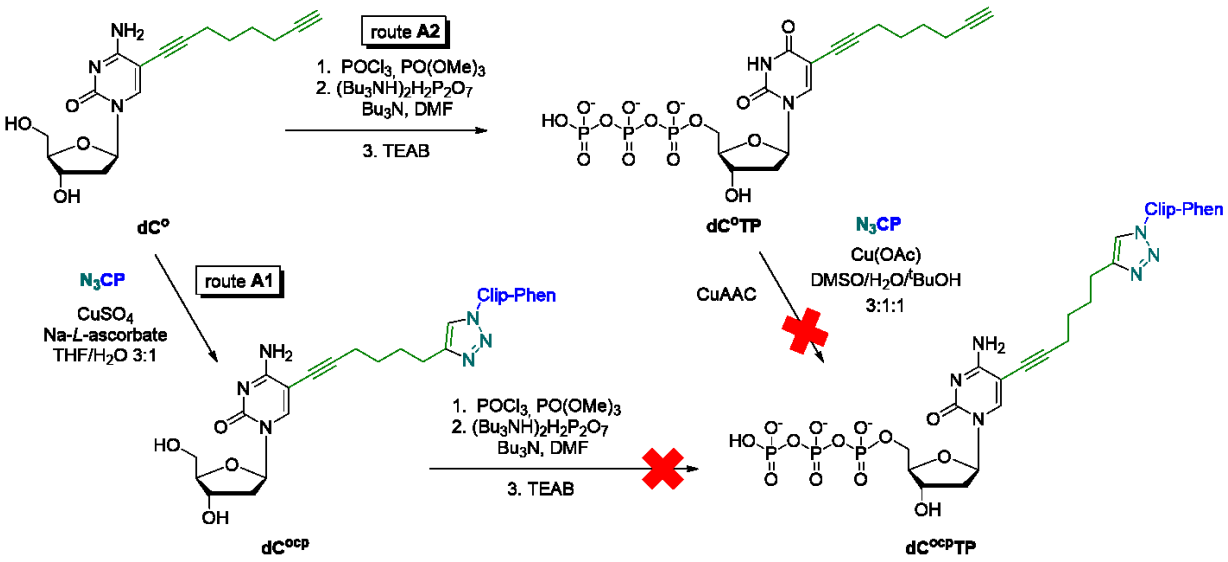
At first, an enzymatic synthesis of the TFO-AMNs was considered, since this approach was never explored before and could constitute the first small step for an *in loco* production of the selective nucleases *in vivo*. As discussed above, two main strategies are possible for the enzymatic production of TFO-AMNs. A first one introduces the Clip-Phen moiety into TFOs by incorporation of dNTPs already bearing the nuclease as a base-modification (synthetic routes **A1** and **A2**). The second one consists of the synthesis of TFOs presenting reactive alkynyl or azido functional groups, to be reacted with Clip-Phen derivatives in a post-synthetic step using the CuAAC reaction (route **A3**, Scheme 6).

3.1.3.1. Synthesis of base-modified dNTPs bearing the Clip-Phen moiety

The preparation of dN^{R}TPs already containing the Clip-Phen moiety, and their successful employment in PEX experiments, could facilitate the enzymatic synthesis of TFO-AMNs since no

post-synthetic step would be required on precious modified DNA. Many attempts to their synthesis were conducted.

To this purpose, alkynyl nucleosides, such as 5-ethynyl-2'-*O*-deaxycytidine (**dC^{e}**)[12] and 5-(octa-1,7-diynyl)-2'-*O*-deoxycytidine (**dC^{o}**)[112], as well as 5-azidomethyl-2'-*O*-deoxyuridine (**dU^{am}**)[113] suitable for an inverse route, were prepared following reported procedures, and employed as starting materials to develop Clip-Phen-modified dNTPs. 5'-*O*-triphosphorylation of these three nucleosides, to obtain the relative base-modified dNTPs, was performed treating the starting dNs with $POCl_3$ in $PO(OMe)_3$, and subsequent addition of tributylammonium pyrophosphate followed by quenching with triethylammounium bicarbonate (TEAB), following a well-established procedure.[114]



Scheme 7. Attempts of synthesis of **dU^{ocp}TP**.

Unfortunately, none of the desired products was recovered from reactions of the three mentioned alkynyl- or azido- dNTPs with **N$_3$CP** or **PgCP**. Employed CuAAC conditions were previously reported in the literature for reactions on dNTPs or ONs,[115–117] however, the desired Clip-Phen modified 5'-*O*-triphosphates **dC^{ecp}TP**, **dC^{ocp}TP** and **dU^{amcp}TP** were never isolated from the complicated reaction mixtures (an example is shown in Scheme 7, route **A2**). For these attempts, water/DMSO/^{t}BuOH mixtures were used as solvent. The presence of the DMSO could

constitute a further complication for the purification step, however, its presence was necessary to solubilize the Clip-Phen reaction partner. Different reaction conditions were tested, employing either a Cu^{II} catalysts ($CuSO_4$) in the presence of Na-$_L$-ascorbate, or a Cu^{I} catalyst in the presence of TBTA as a Cu^{I} stabilizer.[118] The formation of the desired product in all these reaction is not excluded, however the Cu^{II} cations can hydrolysis of the phosphate esters, or be strongly complexed by the oxygen atoms making the purification steps extremely challenging.

The opposite approach was also attempted (Scheme 7, route **A1**), and the click reaction was performed on the nucleosides, to be 5'-*O*-triphosphorylated in a second step. **dC^{o}** and **dU^{am}** underwent the CuAAC with their reaction partners (**N_3CP** in the former case, and **PgCP** in the latter case) using $CuSO_4$ and Na-$_L$-ascorbate in a THF / H_2O 3 : 1 mixture.[112,117] Unfortunately, the 5'-*O*-triphosphorylation of the two obtained nucleosides **dC^{ocp}** and **dU^{amcp}** resulted impossible, probably due to the high reactivity of the aromatic nitrogen atoms in the 1,10-phenanthroline moiety (Scheme 7).

3.1.3.2. Enzymatic synthesis of alkynyl- and azido-TFOs through NEAR

The failure in the preparation of Clip-Phen-modified dNTPs led to the design of the synthetic route **A3** (Scheme 6), involving a post-synthetic CuAAC step on pre-synthesized ONs containing alkynyl- or azido- reactive groups.

Ethynyl-, octadiynyl- and azidomethyl- TFOs could be produced by PEX, employing **dC^{e}TP**, **dC^{o}TP**, **dU^{e}TP**, **dU^{o}TP** or **dU^{am}TP**, obtained as described above. As discussed in the introduction, PEX experiments yield dsDNA. However, TFOs need be single stranded, since the template would strongly compete with the target dsDNA duplex during triplex formation, due to the much higher strength of Watson-Crick base pairing with respect to the Hoogsteen or reverse-Hoogsteen ones. To separate the template from the modified elongated strand (the desired TFO), some techniques are available, such as a magneto-separation, if a biotinylated templated is used in PEX,[119] or treatment with λ-exonuclease to fully digest the template, when the templating strand in PEX is 5'-*O*-phosphorylated.[120,121] However, the scale of PEX reaction is generally very low (generally 0.1 - 1 nmol of product) and, moreover, a substantial amount of precious modified DNA is lost during the process to obtain ssDNA and subsequent purification. This makes difficult and laborious to obtain enough starting alkynyl- or azido-modified TFO to be employed in the

CuAAC step with **N_3CP** or **PgCP**, and further purification, especially if an optimization of the reaction conditions is needed.

The Nicking Enzyme Amplification Reaction (NEAR) is a valid method for the enzymatic synthesis of TFOs, since, as discussed in the introduction, directly yields ssDNA.[5] Moreover, being an amplification reaction, a much higher amount of product can be obtained with respect to what possible by PEX. The main limitation of this method lies in the maximum length of ONs which can be produced. When modified nucleotides of different nature are employed in this technique, reaction yields are acceptable to good for product lengths of 10 – 16 nt, but drop substantially when longer ONs are desired.[120]

A second drawback, shared by all enzymatic methods employable in the synthesis of the target modified TFOs, consists of the problematic incorporation of a single modification into the final ONs. This problem was relevant especially during the construction of TFOs able to form stable parallel triplexes.

As discussed in the introduction, parallel TFOs are short polypyrimidine ONs, which means their sequences are composed only by the two nucleotides **dC** and **dT**. Relevant polypurine sequences present in living organisms, which can be targeted upon parallel triplex formation, always contain multiple dGs and dAs. As a consequence, a parallel TFO able to target this sequence would contain multiple dTs and dCs to form Hoogsteen hydrogen bonds. During NEAR, the polymerase would follow the template to insert dCs complementary to dGs, and dTs complementary to dAs, in the final ON. If for example, the modification within the TFO is inserted using **dC^eTP** instead of **dCTP** in NEAR experiments, the polymerase would place a modified-**dC^e** opposite to each **dG** nucleotide of the template, resulting in a highly modified final TFO (Figure 10A).

To ensure a single modification (and the right positioning of the modification) in the final TFO, the template has to have only one complementary nucleotide, complementary to that position. Thus, to incorporate the modification using **dC^eTP**, the template can contain a single **dG** nucleotide, and the rest of the sequence would be composed only by **dA** nucleotides. In this case, the final TFO would consist of a poly-dT containing a single **dC^e** nucleotide, with limited possibilities of targeting via triplex formation DNA segments present in living organisms (Figure 10B).

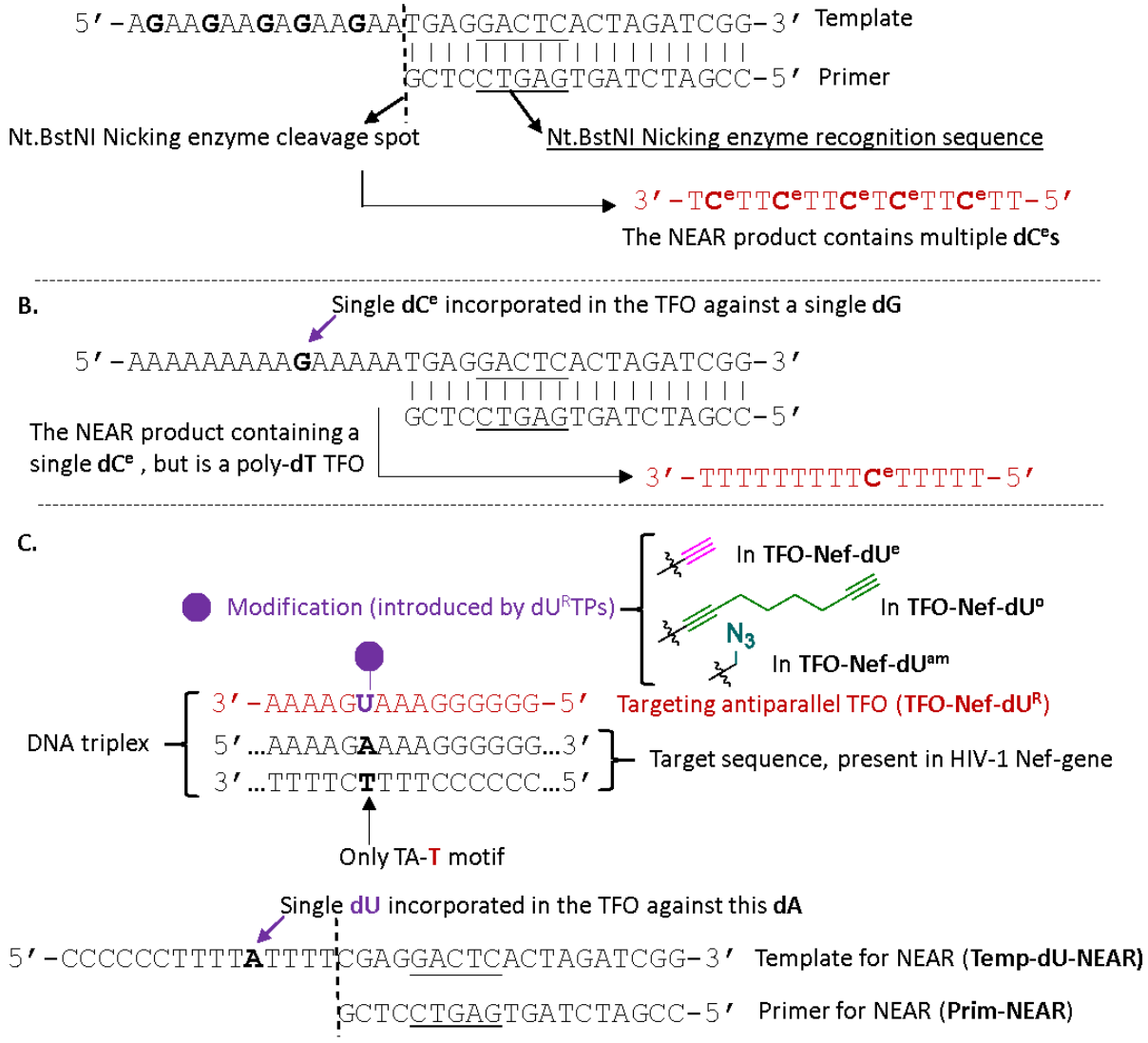


Figure 10. **A.** Multiple incorporation of modified nucleotides during the synthesis of parallel TFOs by NEAR; **B.** Incorporation of a single modification into TFOs by NEAR, yielding a poly-dN TFO with a single different nucleotide. **C.** Strategy for the incorporation of dU^{R}TPs into antiparallel TFOs with no mismatches, and sequences of the target duplex, TFOs produced by NEAR, primer for NEAR and template for NEAR.

Therefore, in order to incorporate a single modification within the sequence of a TFO which could target a relevant polypurine or polypyrimidine segment of DNA, it is necessary to introduce a third base which will lead to a mismatch in triplex formation. Just as an example, a modification could be introduced through one dA^{R} or one dG^{R} nucleotide in a polypyrimidine TFO. However, the purine dN would lead to a mismatch in the triplex, and even a single mismatch could

completely prevent the triplex formation, especially when very short TFOs are employed, such as the ones which can be obtained by NEAR.

On the other hand, when antiparallel triplexes are considered, three different reverse-Hoogsteen motifs are possible and stable (TA-A, TA-T and CG-G, see Figure 6 in the Introduction), allowing the use of **dU^{e}TP**, **dU^{o}TP** and **dU^{am}TP** for the synthesis of polypurine TFOs lacking mismatches, by having a single **dA** nucleotide in the template (Figure 10C). The polypurine sequence chosen as a target for the TFOs (Figure 10C) is a 15 nt segment present in the HIV-1 genome, in the gene encoding for the negative regulatory factor (Nef).[122] Therefore, TFOs containing the **dU^{e}** (**TFO-Nef-dU^{e}**), **dU^{o}** (**TFO-Nef-dU^{o}**) and **dU^{am}** (**TFO-Nef-dU^{am}**) nucleotides were prepared by NEAR using the primer **Prim-NEAR** and the template **Temp-dU-NEAR**, with a combination of KOD XL DNA polymerase and Nt.BstNI nicking endonuclease, to be then purified by HPLC (all sequences are shown in Figure 10C).

3.1.3.3. CuAAC of alkynyl- and azido-modified TFOs obtained by NEAR

CuAAC reactions between the prepared TFOs containing alkynyl- or azido-dU nucleotides and **N_3CP** or **PgCP** provided the target Clip-Phen-based molecular scissors. Reactions were conducted by employing 5 equiv. of the Clip-Phen reaction partner with respect to the TFO, in a water/DMSO/^{t}BuOH 3 : 1 : 1 mixture, strictly under an argon atmosphere. Cu(OAc) was used as the Cu^{I} catalyst. The main difference of this procedure, with respect to the generally used procedures for CuAAC reactions, lies in the non-aerobic atmosphere and the absence of Na-$_{L}$-ascorbate (or other reductants). In CuAAC reaction mixtures, the reductant is necessary to reduce Cu^{II} while forming during the reaction, back to the active Cu^{I} catalyst. The absence of reductant and the inert atmpsphere were necessary to avoid DNA oxidative damage, with consequent degradation of the TFOs in the reaction mixtures. In fact, the Clip-Phen derivative can complex Cu and cleave DNA in the presence of molecular oxygen.

After CuAAC, EDTA was added to the reaction mixtures to complex copper out of the Clip-Phen moiety, and the Clip-Phen-modified TFOs **TFO-Nef-dU^{ecp}**, **TFO-Nef-dU^{ocp}** and **TFO-Nef-dU^{amcp}** were obtained after spin-column purification. It was necessary to complex the copper atom out of the TFO-AMNs conjugates to avoid their self-degradation during their storage. 1 equiv. of Cu^{II}

would be then added to the TFO just before its use in DNA cleavage assays. All the structures are reported in Figure 11.

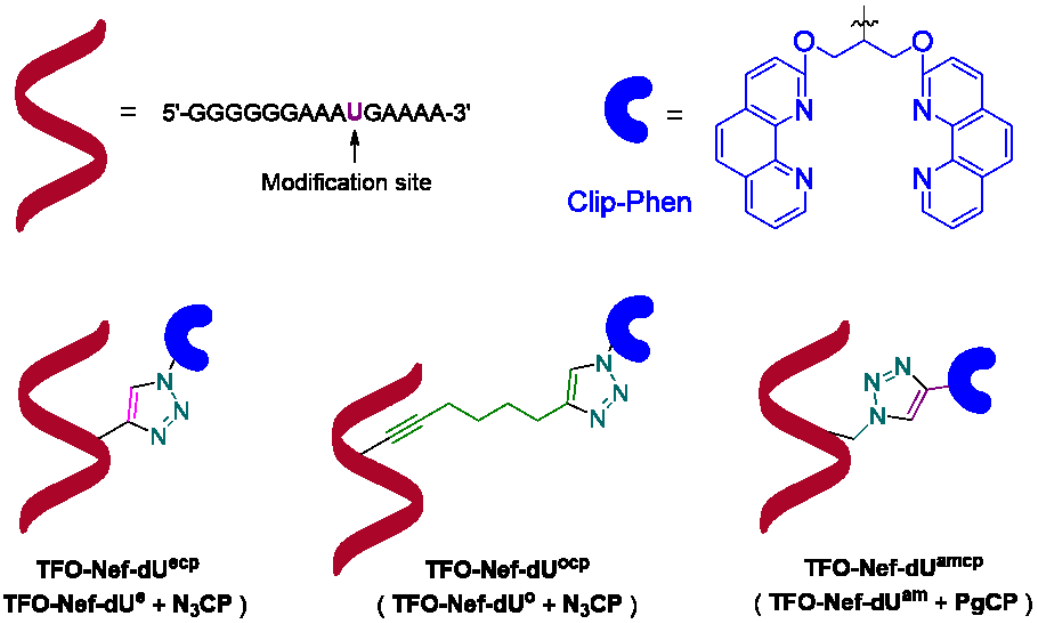
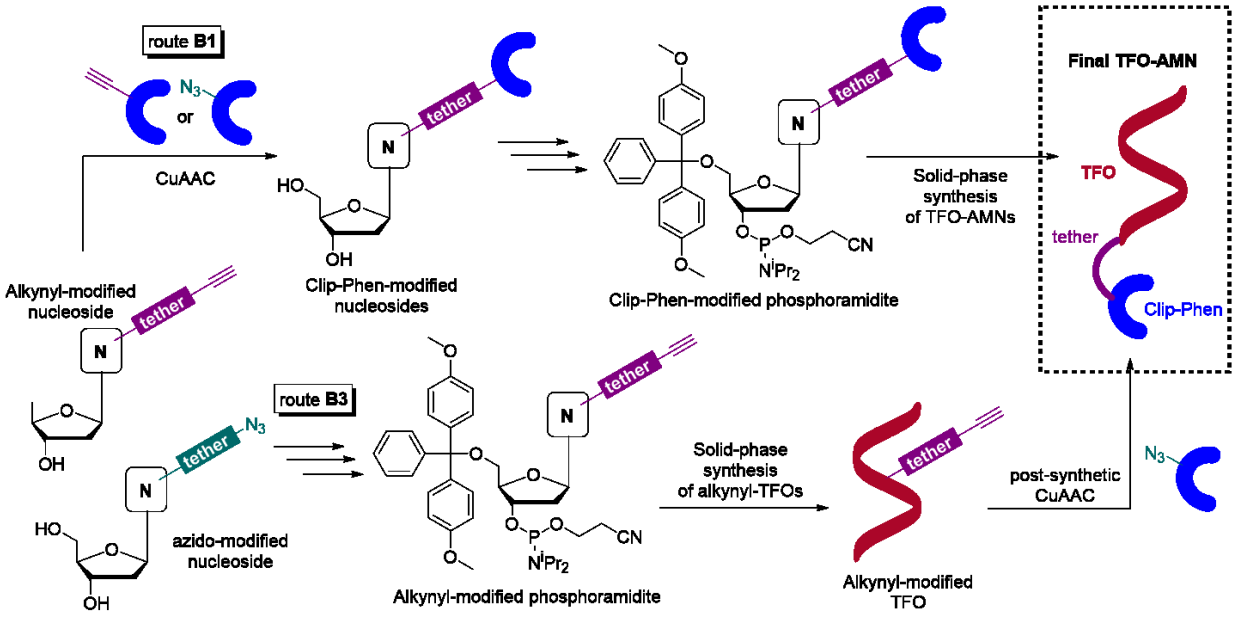


Figure 11. Structures of the TFO-AMNs conjugates obtained by CuAAC reaction between TFO-Nef-dU^{R} and **N_3CP** or **PgCP**.

3.1.4. Chemical approaches for the synthesis of the TFO-AMNs conjugates

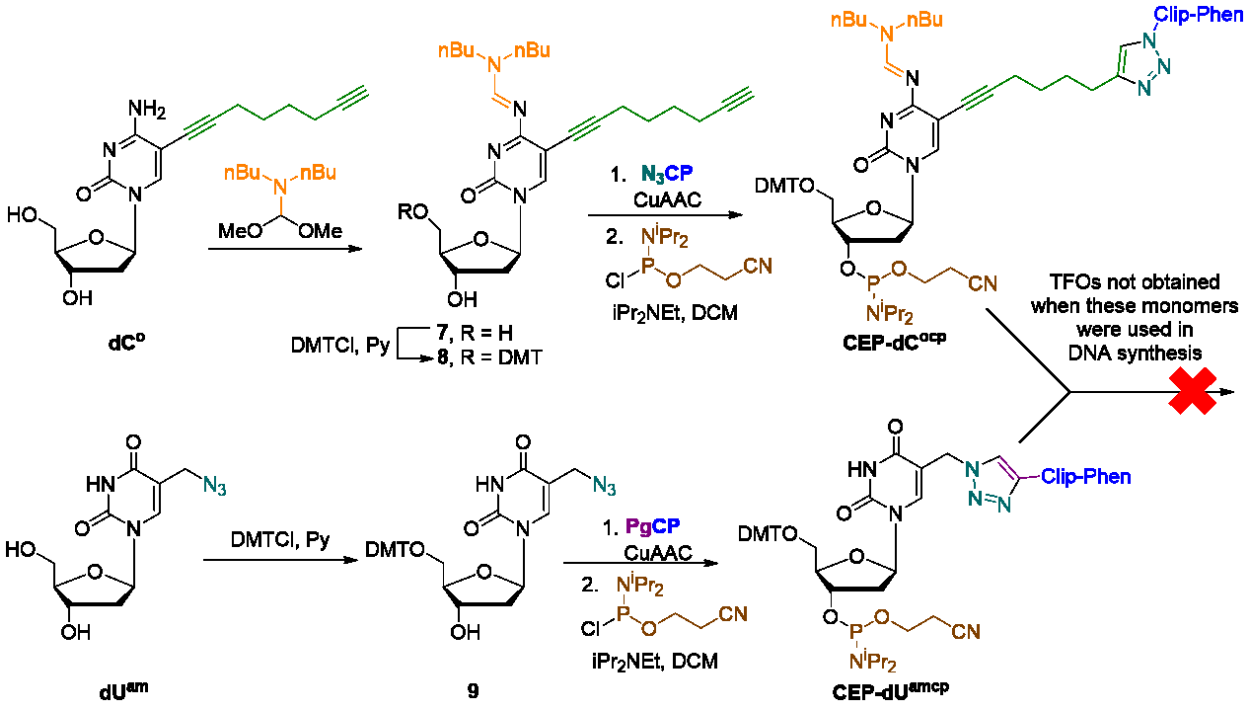


Scheme 8. Routes for the preparation of TFO-AMN conjugates involving the solid-phase synthesis of ONs.

An alternative route to the preparation of the target TFO-AMNs involves the use of an automated synthesizer to perform their synthesis on solid support (synthetic routes **B**, Scheme 8). Two big advantages are present in this route, with respect to the enzymatic approaches. In the first place, no length limitations are present in the solid-phase synthesis of ONs, leaving open the possibility of synthesis of longer TFOs, which would allow more stable triplexes. As a second advantage, it is possible and easy to insert a single modification within the sequence of the synthesized ON, without any restriction on the rest of the inserted nucleotides.

Firstly, to follow the synthetic route **B1**, it was necessary to convert the Clip-Phen-conjugated nucleosides **dC^{ocp}** and **dU^{amcp}** (chosen for this study and which synthesis is described above) into their relative 3'-*O*-cyanoethylphosphormaidites (CEP), essential monomers of the solid-phase synthesis of ONs. A slight modification of the synthetic strategy of **dC^{ocp}** was necessary, since the free amino group on the cytosine nucleobase had to be protected, before the CuAAC reaction with **N$_3$CP**. To this purpose, the octadiynyl nucleoside **dC^{o}** was reacted with *N*,*N*-di-^{n}butylformamide dimethyl acetal (Scheme 9). The obtained *N*-protected nucleoside **7**, and the

azidomethyl-nucleoside **dU^{am}**, were 5'-*O*-protected by treatment with dimethoxytritylchloride in pyridine, to yield the 5'-*O*-DMT-nucleosides **8** and **9** respectively which underwent CuAAC reactions with **N$_3$CP** or **PgCP** in the same condition described above.



Scheme 9. Scheme of synthesis of 3'-*O*-phosphormaidites **CEP-dC^{ocp}** and **CEP-dU^{amcp}**.

The Clip-Phen conjugated products underwent the 5'-*O*-phosphitylation with 2-cyanoethyl *N*,*N*-diisopropylchlorophosphoramidite in the presence of diisopropylethylamine to yield the final CEP building blocks to be loaded on the DNA synthesizer (Scheme 9). The synthesis were straightforward and all intermediates and final products were obtained in good-to-high yields, with the exception of the CuAAC step (51 % yield for the **dC^{o}** derivative, but only 23 % for **dU^{am}**). The lower yields of this step could be attributed to the highly difficult separation between the conjugated nucleosides bearing a Clip-Phen ligand, and the remaining Clip-Phen derivative CuAAC partners **N$_3$CP** and **PgCP**, in the presence of Cu^{II} ions.

The final 3'-*O*-phosphitylation steps, and attempts to couple the obtained building blocks in soli-phase synthesis, were performed by me under the supervision of Prof. Tom Brown and Dr. Afaf H. El-Sagheer at the University of Oxford (UK), during my one-month secondment in Oxford.

In addition to TFOs of the same sequences of those synthesized by NEAR (**TFO-Nef-dU^{R}**), the use of **CEP-dC^{ocp}** in solid-phase synthesis allowed to design the synthesis of polypyrimidine TFOs targeting the same HIV-1 Nef gene polypurine segment, upon more stable parallel triplexes formation, with no mismatches contained.

Unfortunately, none of the synthesized Clip-Phen-CEP building blocks was successfully employed in the solid-phase synthesis of TFO-AMNs, and Clip-Phen-TFOs were never recovered. The failure in the preparation of these ONs could be due to the relatively high reactivity of the nitrogen atoms of the Clip-Phen moiety, which could affect the coupling step of cycles following their incorporation, by reacting with the P^{III} atom of CEP building blocks.

The complementary approach, based on the solid-phase synthesis of alkynyl-modified TFOs by solid-phase synthesis followed by a post-functionalization through CuAAC (synthetic route **B3**, Scheme 8) will be described separately, in chapter 3.1.7. As mentioned above, the opposite approach is not possible because the azide group is not compatible with reagents used in the solid-phase synthesis of ONs, impeding the production of azido-modified TFOs.[105]

3.1.5. Triplex formation studies and preliminary dsDNA cleavage attempts

In order to obtain the desired antiparallel DNA triplexes, a target duplex containing the HIV-1 Nef gene polypurine segment was annealed to the Clip-Phen-modified TFOs (2 equiv.), in the suitable buffer, by heating the mixtures at 95 °C and cooling them very slowly to 4 °C (0.2 °C / sec) in a thermal block. UV melting analysis of antiparallel triplexes were not possible, due to both their very low stability, and the disruption in these analysis by the easy self-association of TFOs containing high percentage of **dA** and **dG** nucleotides.[123]

In order to confirm the triplex formation it was necessary to perform invasive non-denaturing PAGE analysis (Figure 12A). The main issue with these analysis consists of the fact that the triplex could be not stable enough to survive the analysis itself, with the risk of obtaining a false negative also in experiments in which the triplex was formed correctly. As discussed in the introduction, antiparallel triplexes are quite unstable and do not form in the absence of bivalent cations. The addition of a relevant amount of $MgCl_2$ (to 400 mM final concentration) to the ONs solutions, but also to the electrophoresis running buffer, was necessary in order to obtain triplexes which were stable enough to be analysed by PAGE. Moreover, the pH of the mixtures (and electrophoresis running buffer) also played an important role, and no triplexes were detected at pH lower than 7.5 (Figure 12B). All these issues proved antiparallel triplexes not suitable as tools to target dsDNA.

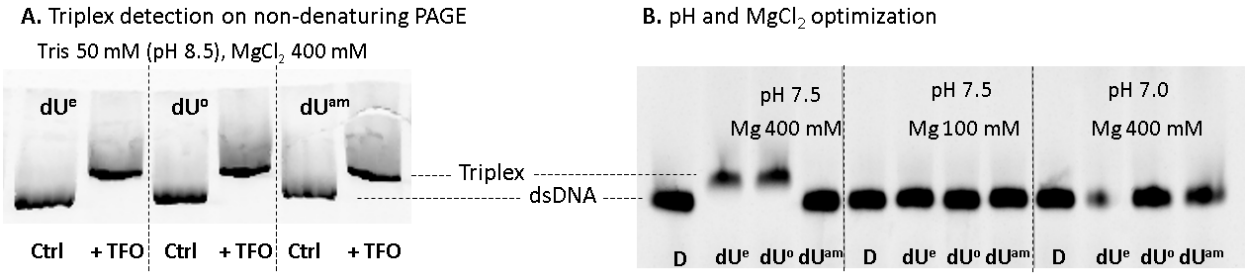


Figure 12. Non-denaturing PAGE analysis of triplexes formed between a target duplex containing a polypurine segment present in the Nef gene of HIV-1 genome, and TFOs **TFO-Nef-dU^{ecp}**, **TFO-Nef-dU^{ocp}** and **TFO-Nef-dU^{amcp}**.

However, preliminary attempts of DNA cleavage were performed by annealing the triplexes in the presence of $CuCl_2$ and Na-$_L$-ascorbate (at different concentrations in different experiments), and incubating the mixtures at different temperatures (4 – 37 °C) for different periods of time (3 – 48 h). Experiments carried on with higher amounts of Na-$_L$-ascorbate, as well as those carried out for longer time or at higher temperatures, showed a good activity of the nucleases. However, the extent of DNA cleavage provided by TFO-AMNs was always similar or comparable to the one observed in control experiments were the free **PgCP** or **N_3CP** ligands were used instead of the Clip-Phen-modified TFOs. These results suggested a complete absence of selectivity in the DNA oxidative cleavage by the developed TFO-AMNs, which could be due to two main reasons.

In a first place, the linker between the Clip-Phen moiety and the ONs is in most cases very short, especially in the case of TFOs obtained by incorporation of the ethynyl- and azidomethyl-dUTPs. On the other hand, the octadiynyl linker of **dU^o** is relatively longer, however it is tethered to the nucleobase by a triple bond, which considerably reduces the flexibility of the whole linker itself, leaving only 4 bonds with fully free rotation. The importance of the length and flexibility of the linker could be understood by analysing the mechanism of action of copper-phenathroline based nucleases, which need to intercalate into the minor groove of DNA where the cleavage is induced. As matter of fact, a long and flexible linker would allow the nuclease to reach the DNA backbone of the targeted duplex once the triplex is formed. If the nuclease is connected to the TFO by a too short or rigid linker, no cleavage could occur which is not the non-specific one due to the TFO-AMN free in solution and not annealed to the target duplex.

The second factor influencing the selectivity in the cleavage is the stability of the triplex itself. The concentration of free TFO-AMN present in the solution and not annealed to the target duplex is increasing with the decreasing of the triplex stability, and contribute to the non-specific cleavage. This problem can be easily circumvent by designing and synthesizing longer TFOs, to stabilize triplexes with additional Hoogsteen hydrogen bonds. However, due to the high difficulty of the synthesis of ONs longer than 15-18 nt by NEAR, a chemical synthesis of the target TFOs must be considered.

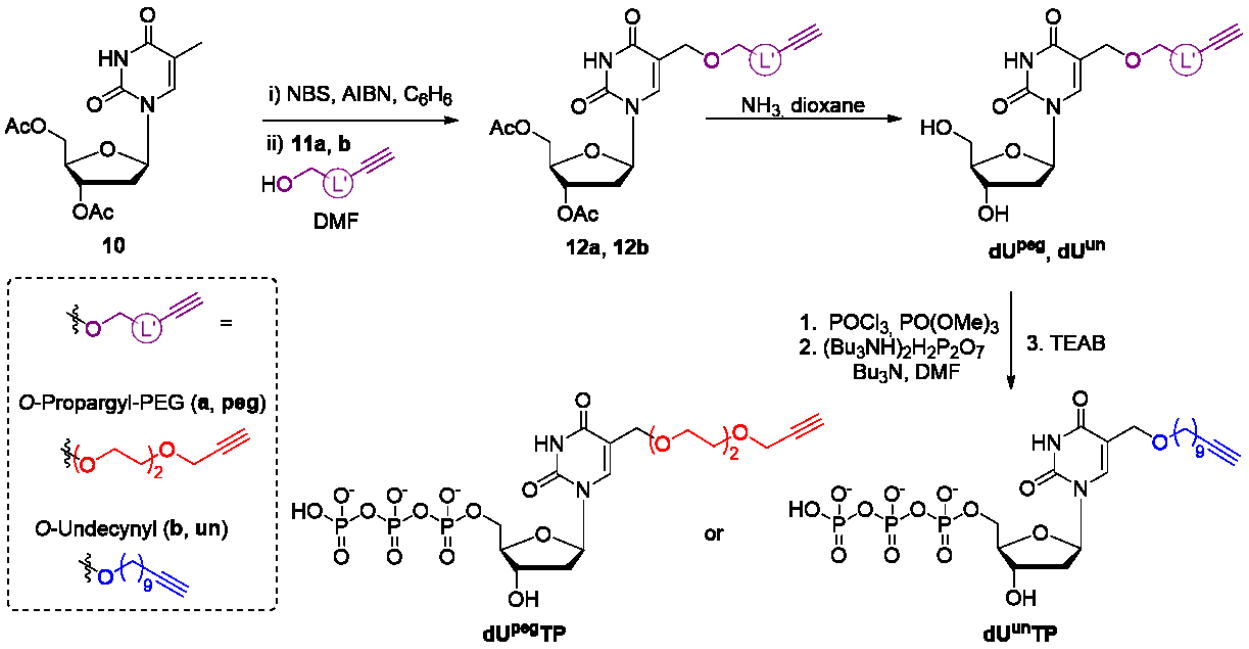
3.1.6. Synthesis of long and flexible alkyne-linked thymidine nucleotides

As described in the introduction, several building blocks for the synthesis of ONs containing terminal alkynes useful for CuAAC are readily available. Most commonly, such building blocks are uridine derivatives, because of their easy accessibility and synthesis. However, most of them present some relevant disadvantages for the enzymatic introduction of clickable alkynes into DNA.

In particular, the most common clickable nucleotide is 5-ethynyl-dU (**dU^{e}**). The nucleoside **dU^{e}** can penetrate cellular membranes and is intracellularly phosphorylated to **dU^{e}TP** which incorporates into genomic DNA. Thus, **dU^{e}** can be detected by fluorescent staining through CuAAC reactions, for metabolic labeling of DNA synthesis.[38–40] The reactive ethynyl-linked nucleotide 5'-*O*-triphosphate **dU^{e}TP** is very well incorporated into ONs by DNA polymerases, and is substrate activity is even comparable to the one of natural **dTTP**.[41] However, CuAAC reactions on **dU^{e}**-modified DNA lack in efficiency, because of the steric hindrance caused by the fact that the ethynyl group is directly linked to the nucleobase and stay well deep inside the DNA groove. On the other hand, the longer octadiyne-linked dNTP (**dU^{o}TP**) was designed to circumvent the poor reactivity of **dU^{e}TP**, and is now widely used in the enzymatic synthesis of clickable DNA.[34] Due to the much longer linker which connects the nucleobase to the reactive acetylene moiety, the efficiency of CuAAC reactions of **dU^{o}**-modified DNA is much higher than the one of **dU^{e}**-DNA. However, the partially rigid and fully hydrophobic nature of the linked alkyne in **dU^{o}TP**, makes this nucleotide a rather bad substrate of DNA polymerases. Hence, its enzymatic incorporation into ONs can be quite difficult, especially when longer DNA with multiple modifications such PCR products is desired.[124]

To circumvent the drawbacks present in the use of **dU^{e}** and **dU^{o}**, it is desirable to have in hand different alkyne-linked clickable nucleotides, able to be easily incorporated into DNA and undergo efficiently CuAAC reactions. For this purpose, two novel thymidine building blocks bearing longer and more flexible alkynyl tethers were designed, containing a hydrophilic oligoethylene glycol (peg) alkyne or a hydrophobic undecyne (un).[124]

Natural thymidine was considered to be the most suitable nucleoside to be employed in this synthesis as starting material, because of its contained cost and easy derivatization. The synthesis started from 3',5'-*O*-diacetyl protected thymidine (**10**) which underwent a radical α–bromination with NBS in benzene. The obtained α-bromothymidine derivative was not isolated, and treated *in situ* with the alkyny-linked alcohol bearing the desired alkynyl linker (Scheme 10).



Scheme 10. General scheme of synthesis of the nucleotide 5'-*O*-triphosphates **dU^{peg}TP** and **dU^{un}TP**.

O-propargyl-diethylene glycol (**11a**) was prepared following a literature procedure,[125] while the undec-10-ynol (**11b**) is commercially available. The two nucleosides **dU^{peg}** or **dU^{un}** were obtained using **11a** or **11b** as nucleophiles in the substitution reaction on 3',5'-*O*-diacetyl protected α–brominated thymidine, followed by a subsequent step of deprotection of the acetyl groups on the deoxyribose moiety using ammonia in dioxane.

Nucleosides **dU^{peg}** or **dU^{un}** were converted in their relative 5'-*O*-triphosphates after reaction with $POCl_3$ in trimethylphosphate, followed by addition of tributylammonium pyrophosphate, and, at last, treatment with triethylammonium bicarbonate (TEAB), as discussed before.[114] The obtained 5'-*O*-triphosphates **dU^{peg}TP** and **dU^{un}TP** can be used as interesting building blocks to introduce these long and flexible terminal alkynes into ONs using enzymatic synthesis.

3.1.6.1. Incorporation of the alkyne-linked thymidine nucleotides into ONs

Both **dU^{peg}TP** and **dU^{un}TP** resulted to be good substrates for KOD XL DNA polymerase (as well as for Bst large fragment, PWO and Vent(exo-) polymerases, see Figure 30 in *Appendix 1*) while testing their enzymatic incorporation into ONs. These dN^{R}TPs were incorporated into short oligonucleotides by PEX, both using a template designed to incorporate a single modification (Figure 13a-b), and a longer template allowing four modifications to be incorporated (Figure 13c-d). The formation of the correct oligonucleotides (**ON_1U^{R}** in case of single modification or **ON_4U^{R}** when four modifications were incorporated) was also proved by MALDI spectrometry.

Next, attempts to incorporate enzymatically **dU^{peg}TP** and **dU^{un}TP** into longer DNA through PCR were performed. The standardly used clickable nucleotides **dU^{e}TP** and **dU^{o}TP** were tested in parallel in PCR experiments to have a direct comparison. KOD XL DNA polymerase was chosen to incorporate the four alkynyl- dN^{R}TPs, using two different templates, a shorter 98 bp and a longer 235 bp one (Figure 13e-f). As mentioned above, the ethynyl derivative **dU^{e}TP** is known to be an excellent substrate for DNA polymerases, and, indeed, its incorporation was smooth and provided the higher amount of product (based on the spot intensity on the gel). Similarly, the use of the hydrophilic peg-linked **dU^{peg}TP** led to the formation of strong modified PCR amplicons, using both the shorter and longer templates. It is important to note that a careful tuning of experimental conditions for the incorporation of this nucleotide was needed, and the addition of extra $MgSO_4$, to a final concentration of 10 mM, consistently helped.

On the contrary, the incorporation of the hydrophobic octadiynyl- and undecynyl- dN^{R}TPs resulted more challenging. The octadiynyl-linked nucleotide **dU^{o}TP** and the undecynyl-linked **dU^{un}TP** gave only very weak PCR products when the shorter 98 bp template was used. In a similar way, when the longer 235 bp template was used, **dU^{o}TP** gave a very weak PCR amplicon while the use of **dU^{un}TP** did not lead to the formation of any product (Figure 13f). The difficult incorporation of hydrophobic nucleotides may be attributed to a struggling re-annealing step of the heavily modified template to de primers, during the amplification reaction cycles. On the

other hand, the long peg linker could complex the Mg^{2+} di-cation wrapping on itself and decreasing the actual volume of the modifications.

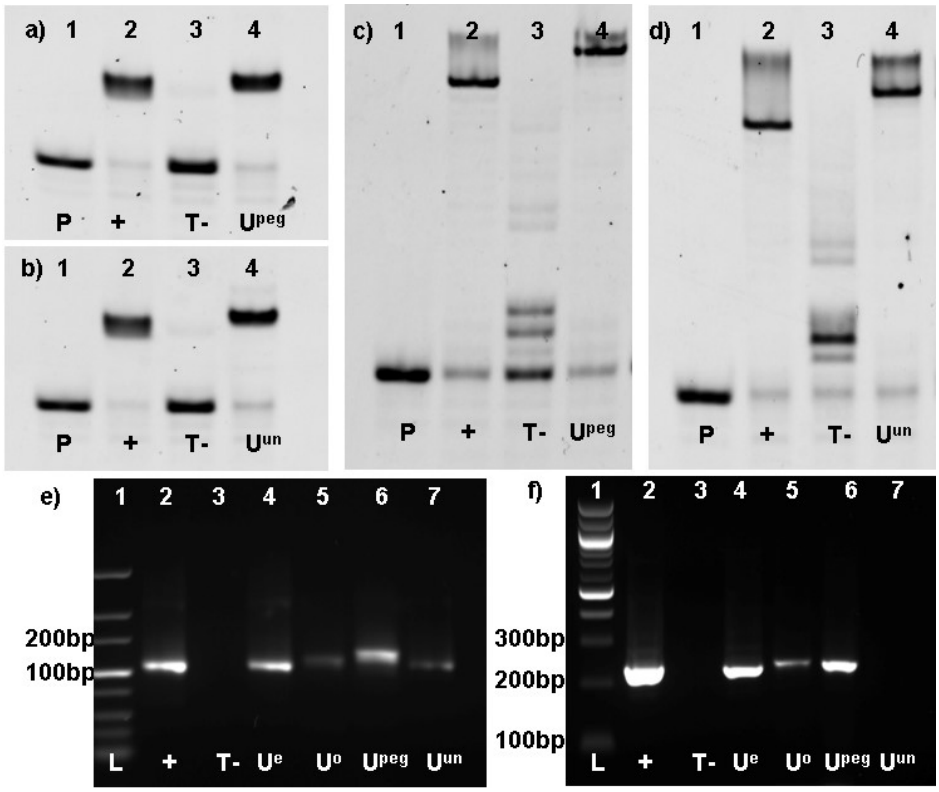


Figure 13. Denaturing PAGE analysis of PEX products containing a single modification after incorporation of (a) **dU^{peg}TP** or (b) **dU^{un}TP**, or containing 4 modifications after incorporation of (a) **dU^{peg}TP** or (b) **dU^{un}TP**. Agarose gel analysys of 98bp (e) and 235 bp (f) PCR products obtained using **dU^{e}TP**, **dU^{o}TP**, **dU^{peg}TP** or **dU^{un}TP**.

3.1.6.2. CuAAC on ONs containing long and flexible alkyne-linked thymidine nucleotides

To study the use of the newly produced clickable dU^{R}TPs, and to compare the reactivity in CuAAC of ONs containing these flexible alkyne-linked thymidine nucleotides with the reactivity of the generally used **dU^{e}**- and **dU^{o}**-modified ONs, short ONs were synthesized through PEX (**ON_dU^{R}**), as well as longer DNA using PCR (**DNA_dU^{R}**). These ONs contained all four alkynyl modified dU^{R}s, and were used as reagents in CuAAC reactions with azides of different nature. It is important to notice that **ON_dU^{R}s** contained each one a single **dU^{R}** modification, while PRC products

DNA_dU^{R}s had all thymidines replaced by alkynyl-linked nucleotides, except for those within the primers.

For a fast comparison of reactivity of all the alkynyl groups in CuAAC, a simple qualitative experiment involving a reaction between the **ON_dU^{R}** (or **DNA_dU^{R}**), and the fluorescent azide TAMRA-N_3, was designed. In this way it was possible, after spin-column purification of the products useful to remove the non-reacted excess of fluorescent azide, to qualitatively follow the reaction through steady-state fluorescence measurements.

Due to the high efficiency of the process, PEX experiments can be considered to be quantitative, whereas the primer is the limiting reagent. Thus, in all reactions performed using **ON_dU^{R}**, it was possible to employed the same exact amount of starting material, for better comparison. Reactions were carried out at 37 °C for 15 min or 1 h, using 5 equiv. of the azide in the presence of CuBr, in a water/DMSO/^{t}BuOH mixture. The peg-linked **ON_dU^{peg}** showed immediately great potentiality of its terminal alkyne as a reactive group for post-synthetic modification of DNA through CuAAC, as it gave the strongest fluorescence of the clicked conjugates (**ON_dU$^{peg\text{-}TAMRA}$**) after 15 min reaction. ONs carrying hydrophobic alkynyl modifications (**ON_dU^{o}** and **ON_dU^{un}**) also reacted very efficiently, even though not as much as the peg-modified ON, while the shorter ethynyl linker showed significantly slower formation of the fluorescent conjugate (Figure 14A). When the reaction was carried on for longer time (1 h), the same trend was observed: the three ONs bearing longer alkynes reached the completion, while **ON_dU^{e}** was found once again slower in the reaction (Figure 14B).

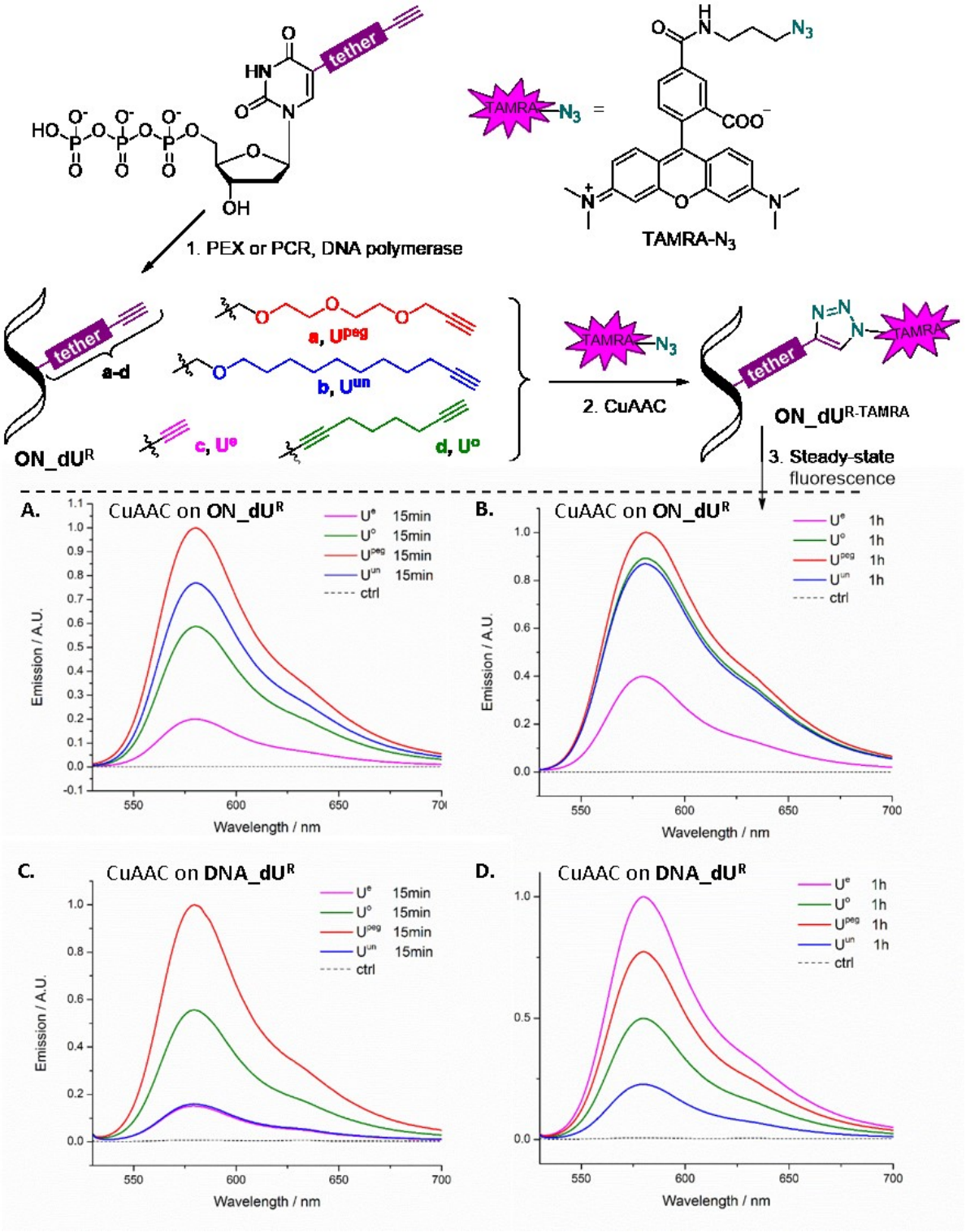


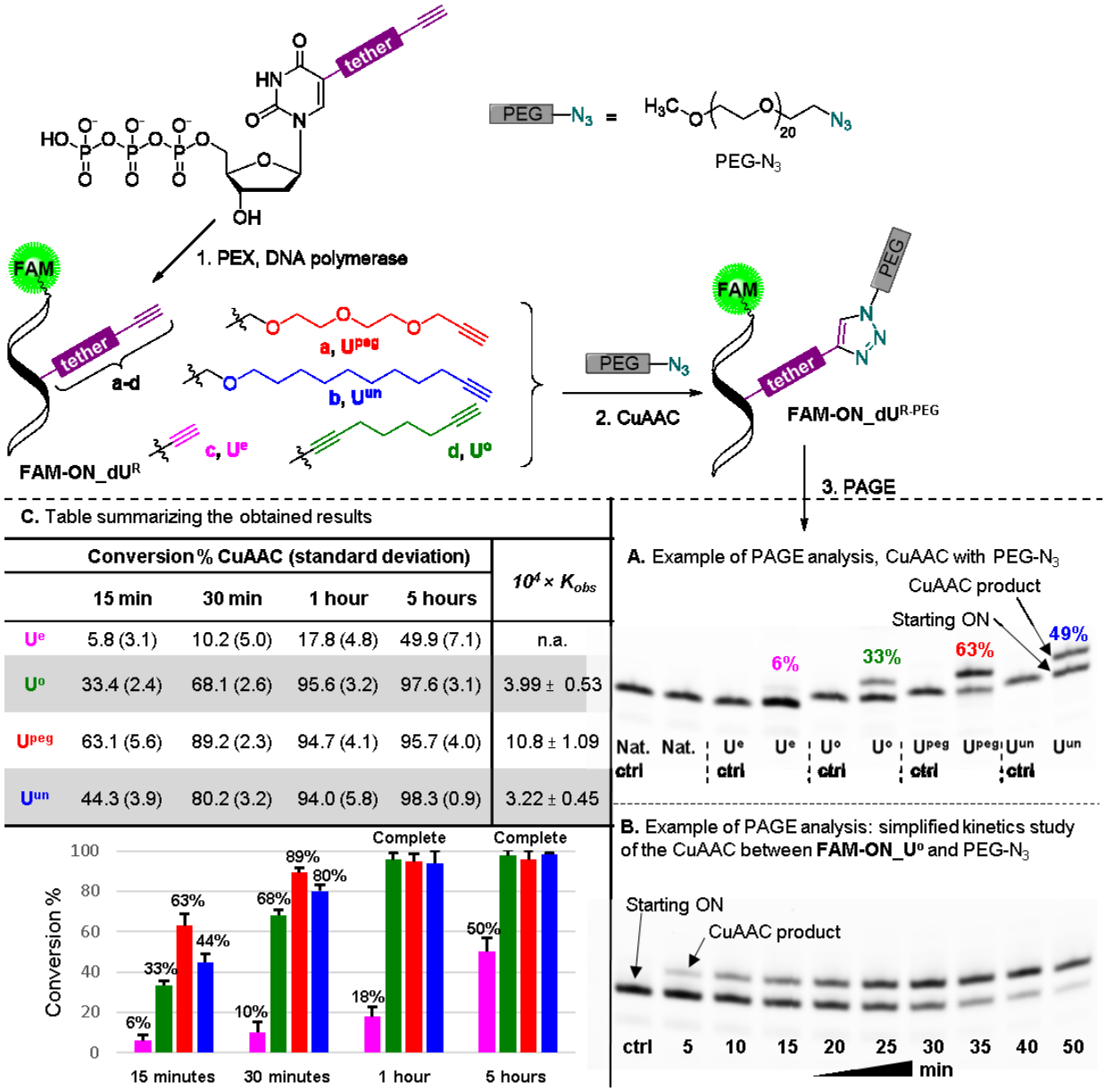
Figure 14. General scheme of preparation of alkynyl-ONs bearing **dU^{peg}**, **dU^{un}**, **dU^{e}** and **dU^{o}** groups by PEX or PCR, and their reaction with TAMRA-N_3; **A.-D.**: steady-state fluorescent measurements of products of CuAAC between TAMRA-N_3 and **A.** ONs obtained by PEX, 15 min CuAAC reaction time; **B.** ONs obtained by PEX, 1 h CuAAC reaction time; **C.** ONs obtained by PCR, 15 min CuAAC reaction time; **D.** ONs obtained by PCR, 1 h CuAAC reaction time.

The click reaction was investigated also on the 98 bp PCR products **DNA_dU^{R}**. As discussed in the previous paragraph, PCR products were obtained in different amounts depending on the employed modified dU^{R}TP. The obtained PCR amplicons were reacted with TAMRA-N_3 in the same conditions employed for the reactions on the PEX products, and, after further spin-column purification to remove the excess of fluorescent azide, the total fluorescence of the final products was measured. In this way, the overall outcomes evaluate both the efficiencies of PCR and CuAAC, coupled together as a unique process. As mentioned before, **dU^{e}TP** is an excellent substrate for DNA polymerases, and its use in PCR resulted into the higher amount of PCR product (**DNA_dU^{e}**) obtained, compared to products of incorporation of the other alkynyl dU^{R}TPs. However, the click on ethynyl-modified DNA is not very efficient and, when reaction mixtures were incubated for short time, e.g. 15 min, the clicked adduct **DNA_dU$^{e\text{-}TAMRA}$** gave a very weak fluorescence (Figure 14C). Similarly, also the undecynyl-linked **DNA_dU$^{un\text{-}TAMRA}$** showed a low fluorescence intensity, however, in this case due to the poor efficiency of the PCR with this substrate. At these reaction times, in general, the influence of the rate of the CuAAC is predominant and, once again, the peg-linked clicked product **DNA_dU$^{peg\text{-}TAMRA}$** gave the strongest fluorescence, confirming to be the best choice for an efficient post-synthetic modification using CuAAC. On the other hand, when reactions were carried on for longer time, e.g. 1 h, the amount of starting material is crucial for the final outcome, and the **DNA_dU$^{e\text{-}TAMRA}$** resulted to have the higher overall fluorescence (Figure 14D).

3.1.6.3. Quantitative study and simplified reaction kinetics

Although the previously described approach based on fluorescent spectroscopy is very fast and simple, it is not suitable for an accurate quantification of the CuAAC conversions, since the silica-membrane-based purification step, necessary to remove the unreacted TAMRA-N_3, could slightly alter the final outcome. Therefore, a fully quantitative approach was designed, and alkyne-linked ONs were reacted with a non-fluorescent azide and the reaction mixtures were analyzed by denaturing PAGE gels. For this purpose, it was necessary to prepare the starting alkynyl-ONs by PEX using a 6-carboxyfluoresceine (6-FAM) labelled primer. PEX products (**FAM-ON_dU^{R}**) were incubated at 37 °C with 5 equiv. of a polyethylene glycol azide (PEG-N_3), in the presence of CuBr, in a water/DMSO/^{t}BuOH mixture. The high molecular weight of the chosen azide leads to a

significant decreasing of mobility on the gel, ensuring a good separation between the PEG-clicked products and the starting alkynyl-ONs bands. Densitometric analysis of bands on PAGE allowed the quantification of the CuAAC conversions (Figure 15A).



	Conversion % CuAAC (standard deviation)				$10^4 \times K_{obs}$
	15 min	30 min	1 hour	5 hours	
U^e	5.8 (3.1)	10.2 (5.0)	17.8 (4.8)	49.9 (7.1)	n.a.
U^o	33.4 (2.4)	68.1 (2.6)	95.6 (3.2)	97.6 (3.1)	3.99 ± 0.53
U^{peg}	63.1 (5.6)	89.2 (2.3)	94.7 (4.1)	95.7 (4.0)	10.8 ± 1.09
U^{un}	44.3 (3.9)	80.2 (3.2)	94.0 (5.8)	98.3 (0.9)	3.22 ± 0.45

Figure 15. General scheme of the preparation of alkynyl-ONs bearing **dU^{peg}**, **dU^{un}**, **dU^{e}** and **dU^{o}** nucleotides by PEX, and their reaction with PEG-N_3; **A.**: Example of denaturing PAGE analysis for the quantification of conversions of the CuAAC between **FAM-ON_dU^{R}** and PEG-N_3; **B.**: Example of PAGE for the study of simplified reaction kinetics of the CuAAC between **FAM-ON_dU^{R}** and PEG-N_3; **C.**: Table summarizing the obtained results.

Reactions were performed for different time periods, and results are shown in the table in Figure 15C. When reacted for short time, e.g. 15 min, the peg-linked **FAM-ON_dU^{peg}** showed again the higher reactivity among the studied alkynyl-ONs, and it was converted into the clicked product with an almost two-fold conversion compared to the standard octadiynyl-modified **FAM-ON_dU^{o}** (63% conversion for **dU^{peg}** against 33% for **dU^{o}**). The undecynyl-modified DNA reacted somewhat faster than the standard octadynyl-modified DNA, however not as efficiently as the peg-modified ON (**FAM-ON_dU^{o}**, 44% conversion). In the employed conditions, a comparison between alkyne-linked ONs containing **dU^{peg}**, **dU^{un}** or **dU^{o}** nucleotides resulted difficult, since all reaction were close to completion. On the other hand, the ethynyl group showed an evident lower reactivity, with only 6% conversion after 15 min and approximatively 50% conversion after 5 hours.

For a better understanding of the differences in the rate of the CuAAC when different alkynyl-linked ONs were employed, simplified reaction kinetics were studied. For this purpose, reactions between **FAM-ON_dU^{R}** and PEG-N_3 were monitored at regular time intervals (every 2 or 5 min) quenching aliquots of the reaction mixture with a solution of EDTA (useful to strongly chelate Cu^{II}, subtracting the catalyst from the reaction mixture) and a large excess of propiolic acid (which quickly reacts with the PEG-N_3 still present in the mixture). Aliquots were analyzed by denaturing PAGE (Figure 15B), and densitometric analysis provided reaction conversions at different time intervals. Plotting these conversions against the time, and further fitting with the exponential curve $y = A_1e^{-k} + y_0$, gave the observed reaction rate constants K_{obs}, shown in the table in Figure 15C. Obtained constants confirmed an approximately two-fold higher reaction rate for peg-linked ONs (K_{obs}: 1.1×10^3) compared to undecynyl- (K_{obs}: 4.0×10^4) and octadiynyl- (K_{obs}: 3.2×10^4) modified ONs.

3.1.6.4. Discussion of the studied alkynyl linkers

To summarize this brief study, two new "clickable" building blocks (**dU^{peg}** and **dU^{un}**) for the production of alkynyl-modified DNA, suitable for a post-synthetic modification through CuAAC, were prepared from inexpensive thymidine following an easy and straightforward designed synthetic route. ONs containing such building blocks reacted efficiently with different types of azides in CuAAC conditions, and their reactions were compared to the ones of standardly used ethynyl- and octandiynyl-linked ONs. The length of the linker between the nucleobase and the

acetylene moiety clearly influenced the results, probably allowing a lower hindrance of the terminal alkynyl group, and the very short ethynyl group resulted a poorly reactive building block. The hydrophilic **dU^{peg}** modification, resulted to be a choice of great advantage when a CuAAC needs to be performed on DNA. Firstly, **dU^{peg}TP** was proved to be an excellent substrate of the employed polymerases, and it was possible to efficiently incorporate the peg-linked alkyne even in long PCR products, while **dU^{o}TP** and **dU^{un}TP** bearing hydrophobic modifications were much poorer substrates. Moreover, ONs bearing this alkyne group were converted into the relative clicked products with a twice faster rate compared to all other studied alkynes, in all cases. Solvation of the peg linker by water constituting the reaction media, mediated by the presence of cations, could advantage this linker, over the hydrophobic undecynyl- and octadiynyl- linkers, in pointing well outside the DNA groove and resulting more available for the azide in solution.[124] These findings allowed a better planning of the synthesis of TFO-AMN hybrids, making possible the conjugation of the AMN to the ONs through different tethers which were proved to eventually undergo CuAAC post-modification in a very efficient way. Mostly, the length and flexibility of the newly developed linkers increased the chances of the AMN moiety of reaching the backbone of the targeted duplex, where the oxidative DNA damage occurs. Therefore, substantial increase in the cleavage selectivity of TFO-AMN where the two main units are connected by these long and flexible linkers is expected.

3.1.7. Solid-phase synthesis of alkynyl-TFOs

As previously seen, an approach based on a post-synthetic functionalization via CuAAC of alkynyl-TFOs obtained by solid-phase synthesis, would avoid all the limitations of the enzymatic methods, such as the low scale of reactions and the difficult insertion of a single modification. At the same time, this approach circumvents the difficult incorporation of Clip-Phen-linked CEPs in solid-phase synthesis (synthetic route **B3**, Scheme 3).

Alkynyl-modified TFOs can be readily prepared by standard solid-phase synthesis using alkynyl-modified phosphoramidites.[16] Following this approach, both the nature of the linker between the ONs and the artificial nuclease, and the position of the modification within the TFO sequences, can be easily explored, and the effect of these two parameters on cleavage ability and selectivity of the molecular scissors can be screened.[104]

Improvements in the selectivity of DNA cleavage by the TFO-AMNs conjugates, with respect to the 15 mer polypurine TFOs previously built by NEAR, could be expected by targeting the DNA with longer TFO sequences, which synthesis is possible using the automated synthesizer. The choice of synthesizing polypyrimidine sequences able to form parallel triplexes, instead of a polypurine TFO, could constitute a further increase in the triplex stability, and consequently an increase of the cleavage selectivity.

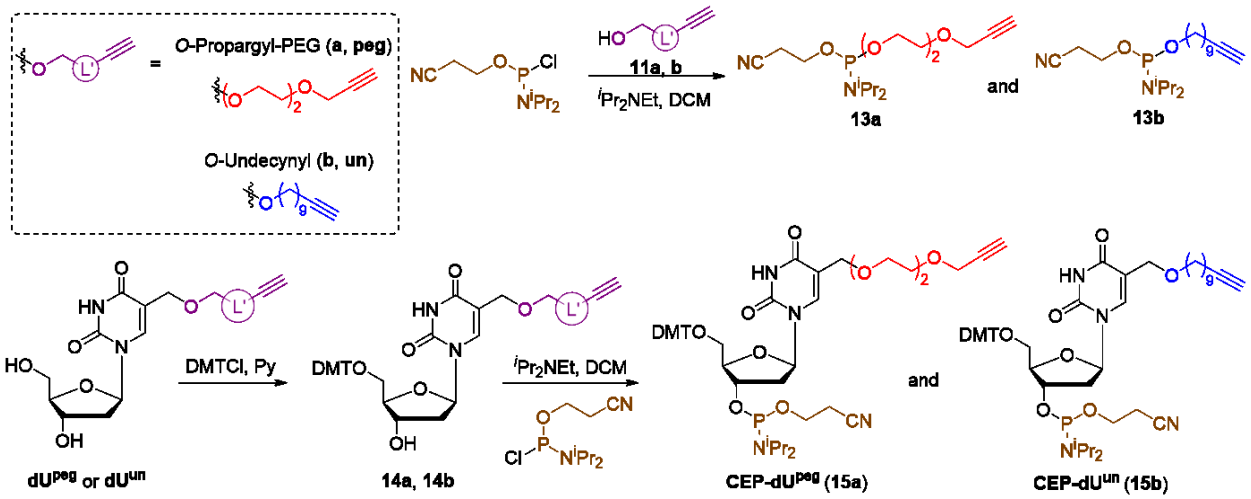
3.1.7.1. Synthesis of long and flexible alkyne-linked thymidine phosphoramidites

The great efficiency in CuAAC reactions of ONs containing the alkyne-linked thymidine nucleotides **dU^{peg}** and **dU^{un}**, in addition to the increased length and flexibility of the two alkynes with respect to the traditionally employed alkynyl-modifications, encouraged the use of these linkers to tether TFOs to the Clip-Phen AMN. However, to incorporate these alkynes into TFOs via solid-phase synthesis, it was necessary to prepare CEPs bearing these moieties.

The alkynyl alcohols **11a** and **11b** can be reacted with *O*-(cyanoethyl)-diisopropyl-chlorophosphite to obtain CEPs **13a** and **13b**, useful to introduce the alkynyl moieties at a 5'-phosphorilated position of ONs through solid-phase synthesis (Scheme 11).[104]

On the other hand, the two previously prepared nucleosides **dU^{peg}** and **dU^{un}** were converted into their 5'-*O*-dimethoxytrityl protected derivatives **14a** and **14b**, using dimethoxytrityl chloride in

pyridine. In the subsequent step, **14a** and **14b** were efficiently converted into the final 3'-*O*-phosphoramidites **15a** and **15b**, using the same 3'-*O*-phosphitylation procedure employed for the synthesis of **13a** and **13b**. In this way, building blocks useful to insert internal alkynes into TFOs in solid-phase synthesis were obtained (Scheme 11).[104]



Scheme 11. Scheme of synthesis of 3'-*O*-phosphoramidites **13a** and **13b** for the synthesis of 5'-modified TFOs, and 3'-*O*-phosphoramidites **CEP-dU^{peg}** and **CEP-dU^{un}** for the synthesis of base-modified TFOs.

All the synthesized CEPs **13a**,**b** and **15a**,**b**, as well as the commercially available 1,7-octadiynyl-linked 5'-*O*-deoxiuridine phosphoramidite (**15c**, Figure 16A) and an alkynyl-linked controlled pore glass, were employed for the incorporation of alkynyl group into TFOs through solid-phase synthesis.

3.1.8. Design and synthesis of the Clip-Phen based DNA-targeting TFO-AMNs

The gene coding for the exterior envelop glycoprotein (Env-gene)[122] in the HIV-1 genome contains a polypurine segment longer than the one present in the Nef gene, which is suitable for parallel triplex formation, upon annealing of a third pyrimidine-rich strand, suitable as a target DNA duplex for the developed DNA-targeting molecular scissors. Specifically, it is a 32-mer segment containing only two thymidine nucleotides within its sequence, which lead to inversion sites in triplex formation. Therefore, two possible TFO sequences were designed to be efficiently used to target the Env-gene of HIV-1 through triple helix formation: a 24-mer TFO, containing a single mismatch against a thymidine of the target duplex, and a longer 32-mer TFO sequence, however leading to the formation of two mismatches, (sequences in Figure 16D). In order to minimize the destabilization caused by mismatches upon triplex formation, two dG nucleotides were displaced against the thymidines of the target dsDNA, to obtain the highest possible triplex melting temperatures.[88,89]

Three different series of alkynyl-TFOs were synthesized, in which the acetylene modification was linked at the 5' terminal position through a phosphate diester, at an internal position linked to a thymidine, or at the 3' position, through an ether bond. To incorporate 5'-modifications in TFOs **A1**-**3**, amidites **13a** and **13b**, which carry flexible alkynyl tethers directly attached to the reactive phosphorous atom, were employed in solid-phase synthesis. Internal modifications in TFOs **A4-6** were introduced with the use of the prepared amidites **15a-b**, as well the commercially available octadiynyl-linked amidite **15d**, in the solid-phase synthesis, while, in order to obtain the 3'-modified TFO **A7**, a commercially available alkynyl-CPG was loaded in the column employed in the solid-phase synthesis (Figure 16C). In this way, it was possible to synthesize 7 different classes of TFOs, allowing the evaluation of the influence of TFO length, as well as modification nature and position along the sequence, in the dsDNA cleavage ability and selectivity.

Subsequently, CuAAC reactions between **N_3CP** and ONs **A1**-**7** provided Clip-Phen-modified TFOs **C1-C7**. Similarly to what described before for CuAAC reactions on NEAR products, these reactions were performed in strictly anaerobic atmosphere, and in the absence of sodium ascorbate, to avoid oxidative damage at the ONs by the Cu^{I} complex of **N_3CP** forming in the reaction mixture.

Reactions were carried out in a water/DMSO/^{t}BuOH 3 : 1 : 1 mixture. The absence of ascorbate forced the direct use a Cu^{I} source, and CuOAc was preferred to CuBr because of its higher solubility in the reaction media, moreover, TBTA was used as a ligand to stabilize Cu^{I} in solution.[118] The addition of a wide excess of EDTA served to quench the reactions and to chelate all the copper out of the clip-phenanthroline moiety. Using membrane-based spin columns for the purification of the resulting oligonucleotides **C1-C7**, it was possible to wash the EDTA-copper complex away, and then elute the products, without need of any further purification. The resulting Clip-Phen-TFOs were characterized by MALDI-TOF mass spectrometry, and no traces of the starting oligo **A1-A7** were found in the spectra, suggesting a quantitative conversion of CuAAC on all the employed substrates (see MALDI spectra in *Appendix 3*, Figures 48-54). Final yields were calculated based on the concentration of the ON solutions (measured through absorbance at 260 nm), and resulted only approximatively 30 to 45 % in all cases, most probably due to elevate loss of product during the spin-column purification step.

Finally, Cu^{2+} cation complexation, accomplished by adding one equiv. of $CuCl_2$ to Clip-Phen-TFOs **C1-C7** provided the TFO-AMNs **T1-T7**. In order to avoid a long exposure of the synthesized TFOs to a potential damage caused by copper ions, the addition of $CuCl_2$ always preceded the triplex formation with the target duplex by only short periods of time, typically 1 hour, to be used for DNA cleavage without further purification.

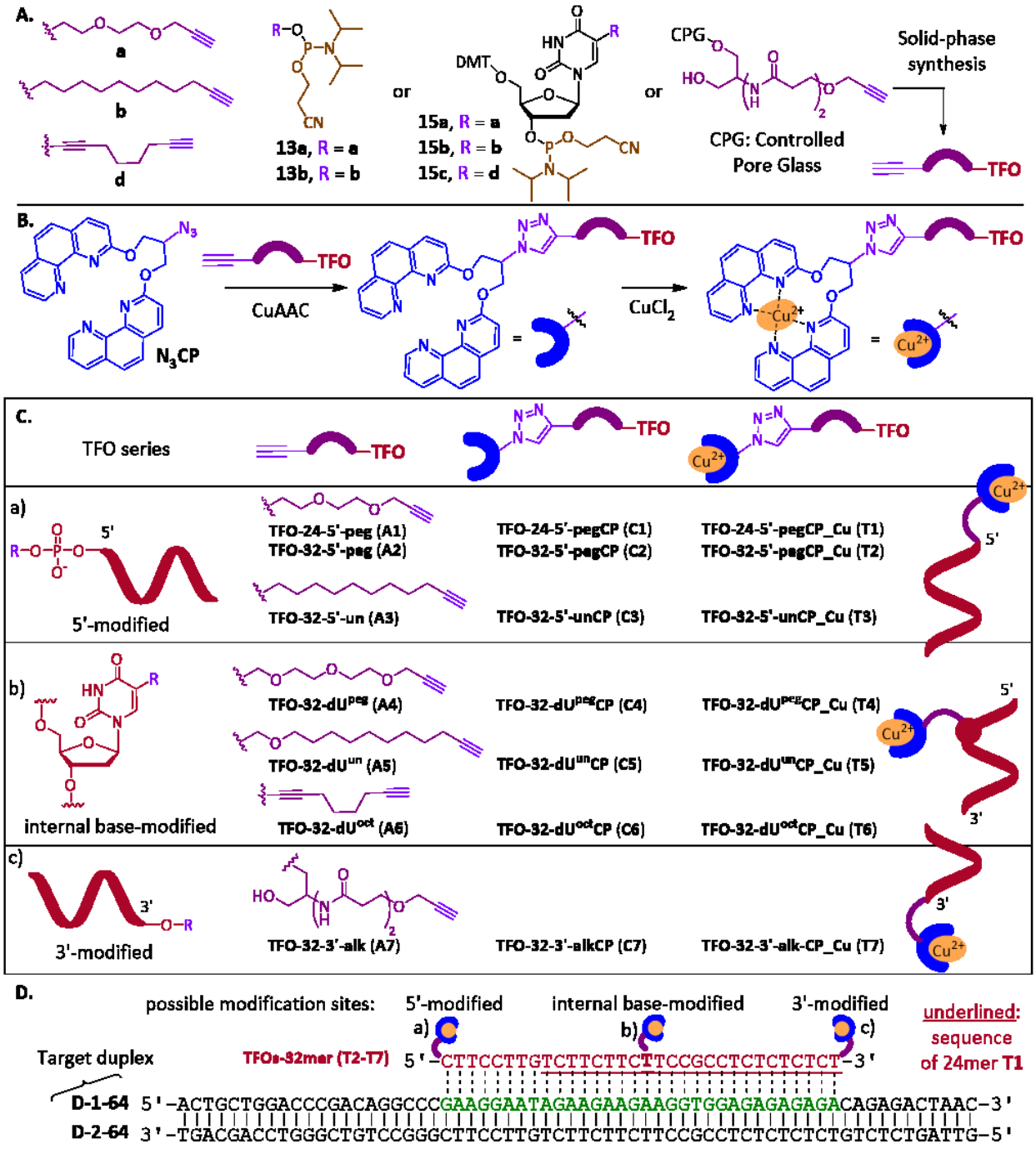
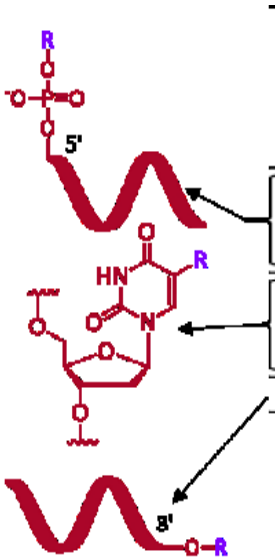


Figure 16. **A.** General scheme of the solid-phase synthesis of alkynyl-linked TFOs **A1-A7**; **B.** General scheme of the synthesis of Cu-Clip-Phen modified TFOs **T1-T7** through CuAAC; **C.** Structures of all the prepared TFOs; **D.** Sequences of the target duplex and all the prepared TFOs.

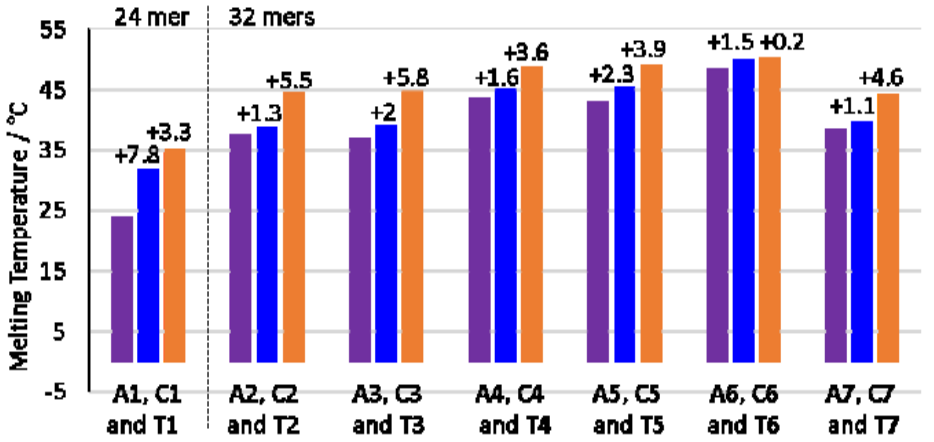
3.1.9. Triplex annealing and UV melting temperature

Triplex thermal melting curves were measured in order to ascertain a strong targeting of the target duplex by the synthesized TFO-AMNs. As discussed previously, parallel triplexes form through Hoogsteen base pairing, requiring an acidic pH for cytidine N^3 protonation. Cations, and especially divalent cations such as Mg^{2+}, play a prominent role in triplex stabilization,[126] therefore triple helices were obtained by annealing synthesized TFOs with the target duplex containing the recognition sequence for the TFOs in a 10 mM phosphate buffer (pH 6.2), containing 250 mM NaCl and 1 mM $MgCl_2$. UV melting temperatures (T_M) served to prove that the annealing between all TFOs and the target duplex successfully occurred, and are reported in Table 1.

Table 1. Triplex UV melting temperatures.



Oligo	Alkyne-TFO (A)	ClipPhen-TFO (C)	Cu-ClipPhen-TFO (T)
A1, C1 and T1	24.1	31.9	35.2
A2, C2 and T2	37.7	39.0	44.5
A3, C3 and T3	37.1	39.1	44.9
A4, C4 and T4	43.6	45.2	48.8
A5, C5 and T5	43.0	45.3	49.2
A6, C6 and T6	48.5	50.0	50.2
A7, C7 and T7	38.7	39.8	44.4



All synthesized TFOs formed triplexes stable enough to be employed in the study, with T_M between 24.1 and 50.2 °C. As expected, T_M measurements for the shorter 24-mer TFOs **A1**, **C1** and **T1** resulted in the lowest denaturing temperatures: these 24-mer TFOs, as previously seen, present a single **dG** nucleotide within their sequence, leading to a mismatch upon triplex formation. On the other hand, longer 32-mer TFOs possess a further destabilizing mismatch, however, the influence on triplex melting of the 7 additional Hoogsteen base pairs (occurring when triplexes are formed with the 32-mers TFOs) is predominant.

The weak intercalation of one 1,10-phenanthroline moiety of the Clip-Phen-modified TFOs **C1-7** slightly stabilized the triple helices, and higher T_M were recorded for this series of TFOs compared to their relative alkynyl-precursors **A1-7**. A further gain in T_M, although only modest (0.2 to 5.8 °C), was caused by prior complexation of Cu^{2+}.

The position of the modification within the TFO sequence also influenced, although not significantly. TFOs bearing internal modifications attached to thymidine nucleotides formed the most stable triplexes, with the highest T_M (50.2 °C) recorded when the octadiynyl-linked TFO-AMNs **T6** was studied. TFO-AMNs with 5'- and 3'- terminal modifications formed slightly weaker triplexes compared to internally modified TFOs, with very similar denaturation temperatures to each other (44.4 to 44.9 °C). All 32-mer TFO-AMNs, hence, formed triplexes with T_M higher than 37 °C, potentially allowing in-vivo application.

3.1.10. DNA cleavage assay and preliminary results

DNA cleavage experiments were performed annealing the target duplex to AMN-modified TFOs **T1-T7**, previous Cu^{2+} complexation. Subsequently, Na-$_L$-ascorbate was added to the resulting DNA triplexes. The ascorbate addition activates the AMN through Cu^{2+} to Cu^{+} reduction (Figure 17). Similar systems were reported to efficiently cleave DNA only when the nuclease comes into contact with the target DNA already laying in a reducing environment.[102] However, the synthesized hybrids were found to cleave DNA more efficiently when sodium ascorbate was added as the last reagent, after the triplex was annealed. For this reason, triplexes were formed in a first step using AMN-TFOs **T1-7**, and the reductant was added in a second moment. Cleavage experiments were performed at the same pH and salt concentration used for the triplex T_M studies (pH 6.2, 250 mM NaCl, and 1 mM $MgCl_2$) to ensure the correct formation of the triple helices. To quench the reactions at the desired times, it was necessary to add a stop solution containing a high excess of EDTA, in order to complex Cu out of the Clip-Phen deactivating the AMNs. As a target duplex (**D**) for this study, a 64bp DNA fragment containing the polypurine segment present in the ENV-gene of HIV-1 was employed. **D** was formed by annealing the two ONs **D-1-64-FAM** and **D-2-64** (see sequences in Figure 16D). **D-1-64-FAM** was labelled with fluorescent 6-FAM, in order to allow an easy following of the cleavage experiments by denaturing PAGE gels, simply measuring by densitometric analysis the disappearing (or weakening) of the duplex band on gel.

The aim of this preliminary study, was to discover reaction conditions in which the synthesized TFO-AMNs are able to cleave dsDNA to a certain extent, however taking into consideration the risk of unwanted off-target cleavage (non-specific cleavage, not driven by triplex formation, therefore occurring at different sites of **D**). Using **T4** as the model substrate, different reaction time periods and temperatures were examined. Similarly, Na-$_L$-ascorbate and TFO concentrations necessary were tested in order to reach relevant DNA cleavage, maintaining the risk of off-target DNA damage as low as possible (e.g. using the lowest possible amount of reductant, or carrying the reaction for short times). Summarizing the results of this preliminary study, oxidative cleavage was slowly occurring when triplexes were incubated in the presence of ascorbate at

temperatures below 25 °C, or for a short period of time (below 3 h). When less than 2 equiv. of TFO-AMN were employed, no DNA damage was observed, and when 5 equiv. of **T4** and only 1000 equiv. of reductant with respect to the target duplex were used, a significant cleavage of 21-26% was achieved. Increasing **T4** or ascorbate concentration, the DNA damage was also enhanced, however together with the risk of unwanted non-specific damage.

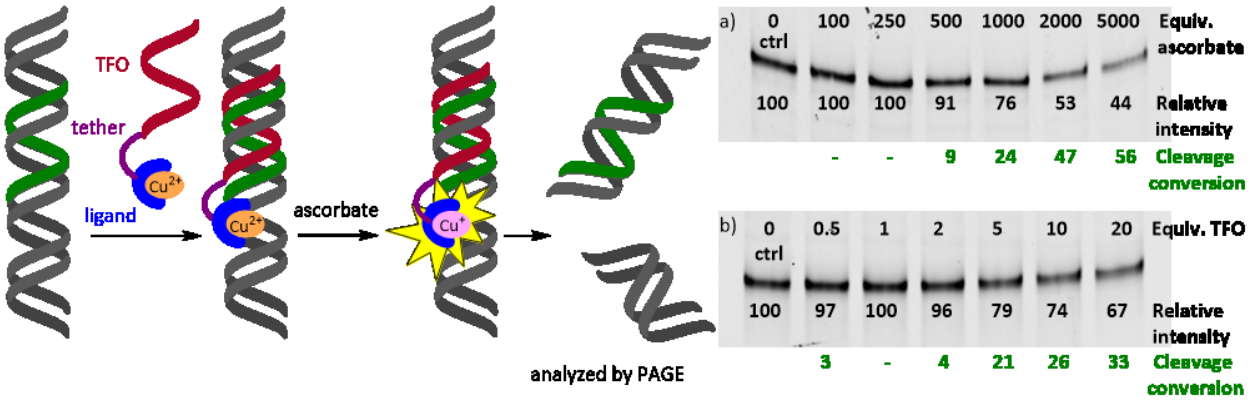


Figure 17. General strategy for dsDNA cleavage assays, using the synthesized AMN-TFO **T1-7**, and examples of PAGE analysis of preliminary screening experiments performed with **T4** to study the influence of the amount of Na-$_L$-ascorbate (a) or TFO (b) on the dsDNA cleavage. Relative intensities of the bands on gels were measured by densitometric analysis and referenced to the control (where the triplex was incubated in the absence of Na-$_L$-ascorbate (a) or TFO (b)).

3.1.10.1. Discrimination of target DNA cleavage vs off-target DNA damage

To measure the selectivity of the synthesized TFO-AMNs **T1-7** towards the target DNA, a slightly different experiment was designed, where an equimolar mixture of two DNA duplexes was treated with the TFO-AMNs in cleavage conditions (Figure 18). The strategy consisted of the fact that only one of the two employed duplexes present in the reaction mixtures contained the recognition sequence for the TFOs, allowing triplex formation (the target duplex **D**). The absence of the target polypurine sequence within the second DNA duplex (the off-target duplex **off**) would make impossible the triplex formation, and only off-target DNA damage can occur to this duplex. Both duplexes were labelled with the fluorescent 6-FAM at their 5' terminus, in order to allow PAGE analysis of the experiment outcomes. Moreover, **off** was designed to be 12 nt shorter than **D**, to allow a good separation on PAGE gels, hence, the possibility to clearly distinguish the

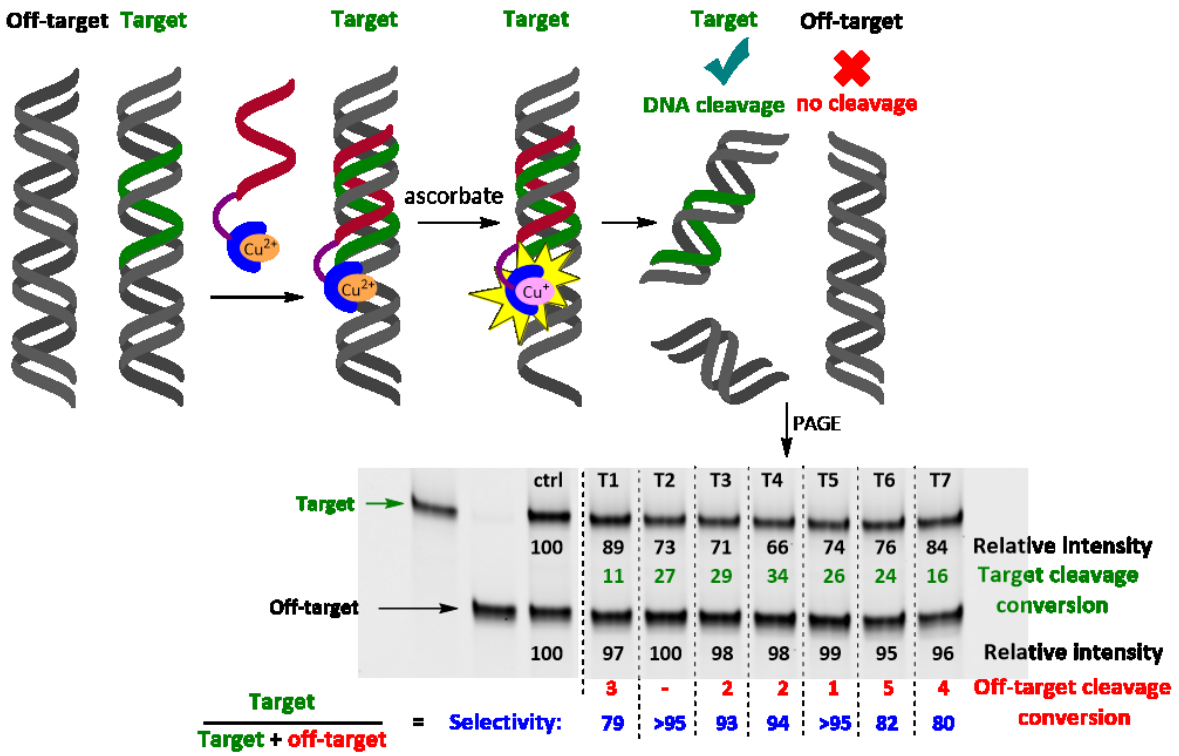
relative bands of the two DNA duplexes. **Off** was obtained annealing ONs **off-1-52-FAM** and **off-2-52** (for sequences see the experimental section, Table 14, paragraph 5.6.1, page 152). After performing the cleavage experiments, densitometric measurements of the PAGE gels were used to calculate the cleavage occurred at both **D** and **off**, independently. DNA cleavage, visible on gels as weakening or disappearing of duplex bands, was measured in comparison to a reference experiment, in which neither TFO, nor ascorbate, were employed. The cleavage selectivity could be calculated by dividing the percentage of cleavage occurred at **D**, by the total cleavage (cleavage occurred at **D** + cleavage occurred at **off**).

Starting from conditions selected during the first fast study, the cleavage of all synthesized TFOs **T1-T7** towards **D** was screened, evaluating the influence of several parameters, such as TFO nature and concentration, ascorbate concentration, as well as reaction time and temperature. The main purpose of this study was, indeed, focused on reaching the highest cleavage selectivity, ideally leaving **off** fully un-cleaved. The table in Figure 18 is summarizing the most significant results, while a comprehensive table of all the results is reported in the experimental part (Table 15, paragraph 5.6.2., pages 153-159).

In general, the cleavage occurred in very low extent at both **D** and **off** in experiment carried out with low ascorbate concentration, such as 200 or 500 equivalents. Similarly, when **D** and **off** were incubated with less than 5 equiv. of TFO, the recorded DNA cleavage was poor. On the other hand, when more than 5 equiv. of TFO and/or more than 1000 equiv. of ascorbate were employed, the system was too active and the off-target duplex **off** was also extensively damaged, with consequent loss in selectivity. A similar situation was encountered when mixtures were incubated for time periods longer than 12 h, such as 24 or 48 h, with high cleavage occurring at both **D** and **off**. Summarizing, the best conditions were found when the equimolar mixture of **D** and **off** was treated with 5 equiv. of TFO and 1000 equiv. of ascorbate, at 25 °C for 12 h. Under these conditions the target duplex **D** was efficiently cleaved by all TFOs **T1-7**, showing a cleavage extent of 12-30 % referred to the control, while no off-target cleavage was found (first column in Figure 18). If the temperature was raised to 37 °C, the cleavage was enhanced, however, unfortunately, a loss in selectivity was also observed (second column in Figure 18). Lowering the reaction time, even to a very short incubation time such as 3 h, did not help, and lower cleavage

percentage and selectivity were observed (third column in Figure 18). The shorter 24-mer TFO **T1** was found to be the least selective, probably due to the lower stability of the formed triplex. No significant difference was found between the cleaving ability and selectivity of TFOs **T2-T5**, which were able to cleave **D** to an extent of 25-30 %, leaving **off** practically untouched. The octadiynyl-modified **T6** showed a cleaving ability only slightly inferior, giving a 21% cleavage in the optimized conditions, and a lower selectivity too. Surprisingly, the 3'-modified **T7** was poorly efficient in the cleavage, showing the least activity among the 32-mer TFOs (16 % cleavage in the optimized conditions).

To prove that the presence of the nuclease is essential for the dsDNA cleavage, alkynyl-modified TFOs **A1-A7** were treated with $CuCl_2$, and employed in cleavage assays, in conditions in which **T1-T7** gave high cleavage. No cleavage was detected, neither at **D**, nor at **off**, when the mixture of duplexes was treated with 5 equiv. of **A1-A7** and subsequently incubated with 1000 equiv. of ascorbate at 37 °C for 24 h (gels are shown in *Appendix 1*). In a similar way, Clip-Phen-modified TFOs **C1-C7** were shown to be unable to cleave dsDNA in the absence of copper. After annealing the ligand-modified TFOs with the duplex, and further treatment with ascorbate, the mixture was incubated at 37 °C for 24 h, and no cleavage was observed, proving copper to be essential for the nuclease activity. With these control experiments, all the reagents, *i.e.* the Clip-Phen-modified TFO complexed to copper, and ascorbate, were proved to be required to achieve dsDNA cleavage. The equimolar mixture of **D** and **off** was also treated with the free copper complex **Cu-N_3CP** at different concentrations, in the presence of the reductant. PAGE analysis showed cleavage occurring at both **D** and **off** in an identical way, proving the triplex formation to be necessary to achieve selectivity, and, thus, the cleavage to be triplex-mediated.



	Target cleavage conversion %[a] [Selectivity][b]			
TFO:	**12 h, 25 °C**	**12 h, 37 °C**	**3 h, 37 °C**	**24 h, 25 or 37 °C**
T1	12 [83]	20 [85]	8 [71]	
T2	27 [>95]	32 [>95]	23 [>95]	
T3	28 [>95]	26 [>95]	26 [>95]	
T4	30 [>95]	31 [>95]	29 [>95]	High off-target cleavage
T5	25 [>95]	24 [>95]	23 [>95]	
T6	21 [91]	35 [82]	19 [>95]	
T7	16 [90]	10 [>95]	9 [77]	

Figure 18. General strategy designed for the measurements of cleavage selectivity towards DNA containing the TFO-recognition sequence, and example of denaturing PAGE. Relative band intensities are calculated by densitometric analysis and referred to the control. In the following table are reported the most relevant results. All reported results were recorded carrying out reactions with 2 pmol of both duplexes, 5 equiv. of TFO and 1000 equiv. of Na-L-ascorbate. (a) percentage of dsDNA cleavage occurred at **D**. (b) selectivity calculated dividing the cleavage occurred at **D** by the total cleavage (occurred at **D** and **off**).

3.1.10.2. Footprinting of cleavage products

In order to footprint the positions of dsDNA cleavage, it was necessary to visualize all potential DNA fragments on PAGE. To achieve that, a radioactively labelled 98 bp dsDNA containing the TFO target sequence was prepared by PCR, employing 1 % radiolabelled α-^{32}P-dATP (Figure 19). The polymerase incorporates radioactive α-^{32}P-dA statistically at every position against thymine bases and, therefore, all the cleavage products are detectable on denaturing PAGE gel. A radioactively labelled DNA ladder was also produced through PCR, using the same template but with primers designed to obtain products of 90, 85, 80, 75 and 70 bp lengths. TFO **T2**, presenting the nuclease attached to the 5' position, was found to cleave the target duplex 2 to 5 bp far from the TFO recognition sequence, giving a clearly visible cleavage product of approx. 76-79 bp (Figure 19).

The full-length DNA band broadening in cleavage conditions is most probably due to oxidative ssDNA damage and alkali-labile sites, induced by exposure to the AMN.

The first control consisted of the target duplex in reaction buffer (**ctrl1** on gel), the second consisted of the target duplex treated with ascorbate and $CuCl_2$ (**ctrl2**), while the third control (**ctrl3**) contained target duplex treated with **T2** in the absence of Na-$_L$-ascorbate. All controls together proved that the shorter 76-79 bp product was generated by cleavage by **T2** in the presence of ascorbate, with all the cleaving partners necessary for the formation of this product.

In the tested experimental setup, as well as in those employed previously to analyze the cleavage of FAM-labelled duplexes, it was not possible to establish whether a double-strand cleavage or single-strand nicking has occurred. Fast experiments employing a target duplex containing each of the strands labelled with a different fluorophore, showed only disappearance of both bands on gel (relative to the two strands), but no shorter fragments were detected. Nevertheless, in analogy to previous literature,[64,77,127] the nature of the Cu-Phen AMN employed suggested formation of both ss- and dsDNA breaks (formed through two successive ssDNA breaks, should be the major cleavage pathway).

The best results were obtained when only 1 equiv. of **T2** was annealed to the target duplex, and the system treated with 1000 equiv. of ascorbate. When the cleavage was pushed to higher extent, either increasing the equivalents of TFO (gels in Figure 38 and 41, *Appendix 1*) or with

higher amounts of ascorbate (gels in Figure 40, *Appendix 1*), only indiscriminate cleavage was observed, and it was impossible to detect any distinct fragment. Experiments with **T4** and **T7**, carrying the nuclease at an internal position or at the 3' position respectively, did not provide distinct cleavage products (gels in Figure 39, *Appendix 1*)

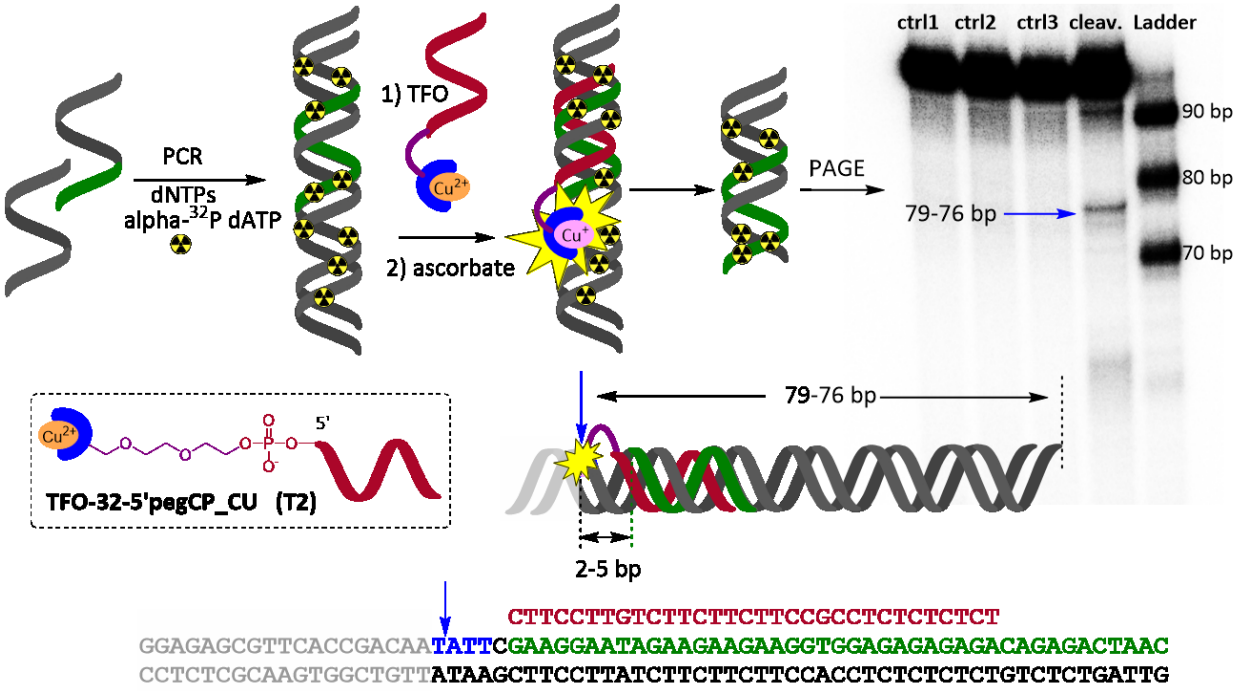


Figure 19. System employed to footprint the position of the cleavage, and cleavage pattern showed by **T2**, found by denaturing PAGE.

3.1.10.3. Discussion of TFO-AMNs activity and selectivity

A stabile triplex formation resulted to be of crucial importance for a good cleavage selectivity by the synthesized TFO-AMNs, and the length of TFO sequences was the main influencing factor in triplex stabilization. **T1** is 8 nucleotides shorter with respect to the other TFOs **T2-7**, and the T_M of the triplex formed by hybridization of **T1** and **D** (in the reaction buffer) was only 35.2 °C, suggesting that less than 50 % of the TFO is actually annealed to **D** at 25°C (employed reaction temperature in the optimized conditions). Accordingly, experiments performed with **T1** at 25 °C showed poor selectivity.

On the other hand, not much difference in cleavage activity and selectivity was induced by the position of the AMN modification within the TFO sequence. **T2** and **T3**, bearing the modification attached to the 5'- terminus, were only slightly more effective than **T4** and **T5** (which carry the AMN attached through the same linkers at an internal thymidine). This is probably due to the higher degree of freedom of the linker, which is able to reach the phosphodiester DNA backbone of **D** in an easier way (where the oxidative cleavage takes place). Surprisingly, the 3'-modified TFO-AMN **T7** showed a much lower activity and selectivity, however, the employed linker is very different from the ones of **T1-6**, thus no direct comparison is possible. Interestingly, **T6**, bearing an octadiynyl linker between the AMN and an internal thymidine of the TFO, displayed the lowest selectivity among internally-modified TFOs **T4-6**, although the triplex formed annealing **T6** to **D** was the most stable (T_M: 50.2 °C). The lower flexibility of its linker, compared to the longer and fully saturated peg- and undecynyl- tethers of **T4** and **T5**, affected the cleavage selectivity, probably due to the higher difficulty of reaching the phosphodiester backbone of **D**.

The fine tuning of the amount of Na-$_L$-ascorbate employed in experiments turned out to be of key importance. With too low amounts of this reductant, Cu^{I} would be too quickly re-oxidized to Cu^{II}, with consequent deactivation of the AMN. In contrast, high concentrations of Na-$_L$-ascorbate led to high off-target DNA damage, due to high production of diffusible free radicals.

The DNA cleavage ability of the developed TFO-AMNs was in general lower than the one of TFO-Cu^{II}-Phen hybrids[77,102] synthesized by Helen, and only comparable to the one of TFO-Fe^{II}-EDTA hybrids[101] designed by Dervan. However, a big advantage of the newly described Clip-Phen-TFO conjugates on the previously reported hybrids, lies in the straightforwardness and flexibility of their convergent synthetic approach. With a single and efficient CuAAC it was possible to attach the AMN moiety to a series of different TFOs, and explore the influence of several linkers differing in length and nature. Moreover, the nature of the employed linkers (triazole-linked TFO-AMNs, instead of 5'-thiophosphate or polymethylene thiophosphate spacers), as well as the one of the AMN itself (Clip-Phen chelate Cu-complex, instead of Phen Cu-complex), make the designed TFO-Clip-Phen conjugates more stable with respect to all hybrids which were precedently studied. As a last advantage, the selected synthetic approach allows to connect the AMN to whichever position within the TFO sequence, and it is not limited only to a 5'-end modification.

The evaluation of TFO-AMNs efficiency in site-selective cleavage of a target DNA duplex showed that an optimal displacement of the AMN, directed by a TFO upon triplex formation, was achieved. Optimum conditions were found, in which TFO-AMNs were able to successfully discriminate and substantially cleave a DNA duplex containing their recognition sequence (**D**), leaving practically intact other non-targeted DNA duplexes (**off**).

Footprinting experiments were performed in very mild conditions, for example using only 1 equiv. of TFO-AMN with respect to the target duplex, and DNA fragments resulting from the cleavage were visualized only when the 5'-terminal peg-linked **T2** was used. This experiment clearly showed that local site-selectivity in the cleavage is achieved (although only at the very beginning of the reaction), however, the flexibility of the peg linker prevented specificity at single-nucleotide level. A similar situation might have occurred when the **D** + **off** mixture is used, showing thus good selectivity at the beginning of the reaction, while after certain extent of cleavage the discrimination between **D** and **off** drops. The drop in selectivity is probably due to a combination of two factors: while the concentration of **D** decreases, a relative higher level of **off** stays in solution and, at the same time, higher relative concentration of TFO-AMN as **D** and **off** are consumed.

The low resolution in DNA cleavage exhibited by the prepared TFO-AMNs was in strong contrast with the clear footprinting observed in DNA cleavage by the TFO-Fe^{II}-EDTA hybrids designed by Dervan,[101] and the Phen-TFO hybrid by Shimizu et al.[103] On the other end, it is important to note that the DNA cleavage of the target duplex with previously prepared TFO-Fe^{II}-EDTA or TFO-Cu^{II}-Phen was never tested in the presence of a large amount of non-targeted DNA.

The poor resolution in DNA cleavage by the developed TFO-AMN conjugates, makes impossible their application if a further recombination is wanted, for example for gene editing. On the other hand, the type of non-resolved cleavage achieved can be very useful to damage specific DNA *loci* in such way that it cannot be repaired. For this purpose, the type of cleavage obtained by local radical formation induced by Cu^{I} can be superior to CRISPR-Cas systems, which cleavage is relatively easy to be repaired.

3.2. Competitive incorporations of base-modified dNTPs

A critical part of the thesis is based on the production of base-modified DNA, proving all methods available to introduce base-modification into ONs, of primary importance. In this chapter the attention will be focused on the enzymatic synthesis of base-modified ONs, which, as previously seen, is generally conducted incorporating chemically modified dN^RTPs using DNA polymerases. For many applications in biotechnology, an *in cellulo* or *in vivo* synthesis of modified DNA is desirable. However, most of the enzymatic techniques used in laboratories for DNA synthesis work in the absence of the natural dNTP corresponding to the base-modified nucleotide carrying the modification (Figure 20). The presence of a large amount of natural dNTPs in cells makes it impossible to achieve the same situation when *in vivo* experiments are performed. However, no systematic study of polymerase incorporation of base-modified dN^RTPs, in the presence of their relative natural dNTPs, was known (Figure 20E).

This study was inspired by a previous research by Dr. Pavel Kielkowski (currently at LMU, Munich), who performed a systematic study of competitive enzymatic incorporation of aryl-modified dA^RTPs and dC^RTPs in the presence of natural **dATP** and **dCTP**. Pavel Kielkowski also designed all assays for competitive PEX experiments and kinetic studies employed in here.[128] All experiments (except the direct comparison of **$dG^{7d}TP$** and **$dG^{ph}TP$** designed by me) were designed by Dr. Hana Cahova (IOCB, Prague), who co-supervised me during the whole project. All the experiment were performed by me. Dr. Jindrich Fanfrlik (IOCB, Prague) was also a coworker, and performed all the computational studies.[41]

3.2.1. Enzymatic incorporation of base-modified dNTPs

Most commonly, when a chemical modification is required onto a nucleobase in an ON, a base-modified nucleotide-5'-*O*-triphosphate (dN^{R}TP) is employed in PEX, PCR, or any other method for the enzymatic production of DNA, in the absence of its relative natural dNTP.[22]

As discussed in the introduction, the simplest way to introduce enzymatically a modification into a short dsDNA consists of a PEX experiment, in which one (or more) modified dN^{R}TP are used instead of the relative natural dNTP. If, for example, a base-modified adenosine is desired in a PEX product, a base-modified dA^{R}TP would be used in the PEX reaction instead of (and in the absence of) natural **dATP** (Figure 20A). This will ensure the presence of the modification in the final DNA, which will contain the modified dA^{R} nucleotide opposite to every complementary **dT** nucleotides originally present in the template. In such case, it is of critical importance to carefully design the template, since the number and position of modification in the final product will depend only on its sequence. Similarly, when a longer base-modified dsDNA needs to be prepared, a base-modified dN^{R}TP can be employed in PCR experiments. If its natural counterparts is not present in the reaction mixture, the products will contain the base-modified nucleotide at all positions against its complementary dN (with exception made for regions annealing to primers) (Figure 20B). The same outcome is achieved when NEAR is performed using one (or more) modified dNTPs (Figure 20C), with the difference that short single-stranded modified ONs are yielded by this amplification reaction.

On the other hand, all enzymatic methods for polymerase incorporation of dNTPs can be performed using a mixture of natural and base-modified dNTPs. In this way, the final outcome will be a product (e.g. a short ON in the case of PEX and NEAR, or longer DNA for PCR) in which a statistical dispersion of the modified and natural dNs are found, positioned in opposite to the complementary base preexisting on the templates (as an example, a PCR performed with a mixture of natural and modified **dATP** is shown in Figure 20D). In the products of these reactions, the ratio between natural and modified dNTP incorporated is depending on several factors, such as reaction parameters, relative concentration of the two dNTPs in competition, and many other.

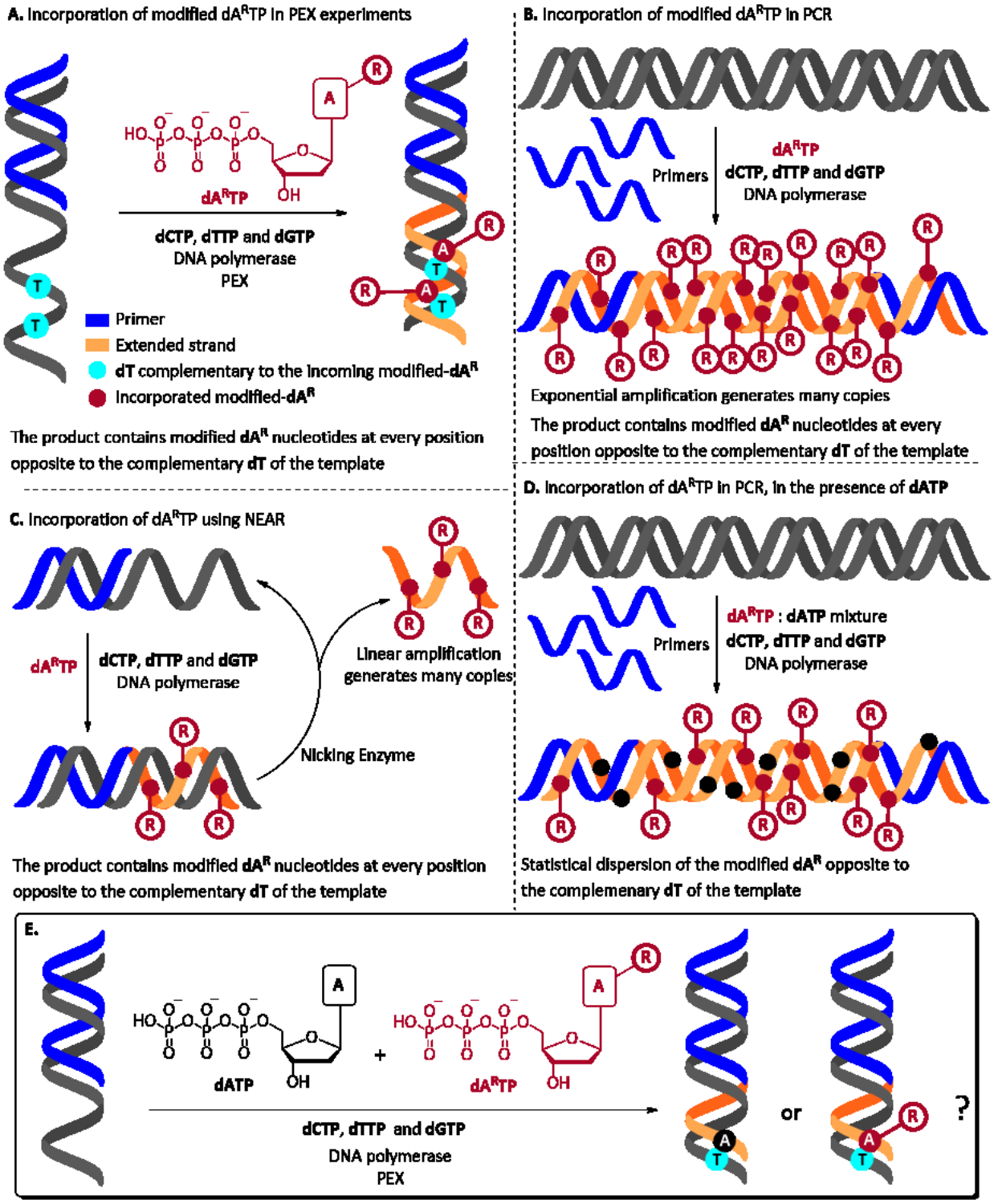


Figure 20. **A.** PEX experiment for the incorporation of a modified dNTP (dA^{R}TP in the specific example) into short dsDNA; **B.** PCR for the incorporation of a modified dA^{R}TP into long dsDNA; **C.** Incorporation of a modified dA^{R}TP into short a short ssDNA using NEAR; **D.** PCR for the incorporation of a modified dA^{R}TP, in the presence of natural **dATP**; **E.** Competitive incorporation of dN^{R}TPs into ONs.

Among all the factors, the ability of the polymerase to accept the modified dNTP as a substrate, in competition with its natural counterpart, plays a prominent role. The study of such systems, in which mixtures of natural and modified dNTPs are employed in enzymatic methods for DNA production, could be of critical importance, especially in the prospective of *in vivo* applications. High concentrations of natural dNTPs are present in live cells, and biotechnological methods to insert base-modification via polymerase incorporation of modified dNTPs must rely on the ability of the enzyme to accept and use the modified substrate also in a wide abundance of its natural counterpart. It is of crucial importance, thus, to have a good understanding of the behavior of DNA polymerases when the simultaneous presence of natural and base-modified dNTPs is needed or not avoidable (Figure 20E).

Therefore, a systematic study of polymerase incorporation of base-modified dNTPs, in the presence of all natural dNTPs, was performed, employing substrates carrying base-modifications of different size and nature.[41] Mainly, PEX and PCR experiments were conducted employing a comprehensive series of base-modified dNTPs, in the presence of all natural dNTPs. PEX assays, thus, were designed in order to allow the analysis and quantification of the relative amounts of natural and modified dNTP incorporated into the products. Similar assays were then adapted for the study of even more complex techniques, such as PCR.

Moreover, an insight into DNA polymerases mechanism and interactions with base-modified dNTPs was attempted by studying enzymatic kinetics of incorporation and performing molecular modelling.[41]

3.2.1.1. The competitive PEX assay

Several studies have been conducted to understand the behavior of REs when base-modification are present in the targeted DNA within their recognition sequence. Results showed how many REs do not tolerate bulky modifications within the recognition sequences, while the enzymatic snip of DNA can proceed when smaller modification are present. Other REs, such as EcoRI, AfeI, BglII, RsaI SacI, sphI and KpnI, are even more sensitive and tolerate only very small or no chemical modification on the base.[12–15,129]

A simple assay for the study of the competitive incorporation of modified dN^{R}TPs in the presence of natural dNTPs was developed, based on this knowledge, by Pavel Kielkowski (Figure 33).[128]

When a mixture of dNTP and dN^{R}TP is employed in the PEX assay, a mixture of modified and natural DNA is obtained, in proportions depending on the ability of the dN^{R}TP to be a substrate for the polymerase, with respect to its natural counterpart. For competitive PEX experiments it was necessary to design a template allowing the incorporation of a single modification within the recognition sequence of a RE, chosen to be unable to cleave DNA carrying the studied modification. The mixture of natural and modified DNA resulting from the competitive PEX was subsequently treated with the RE and the outcome was analyzed by PAGE. Intensities of bands on gel relative to cleaved and uncleaved products were proportional to the ratio between natural and modified DNA, therefore proportional to the ratio between dNTP and dN^{R}TP incorporated by the polymerase.

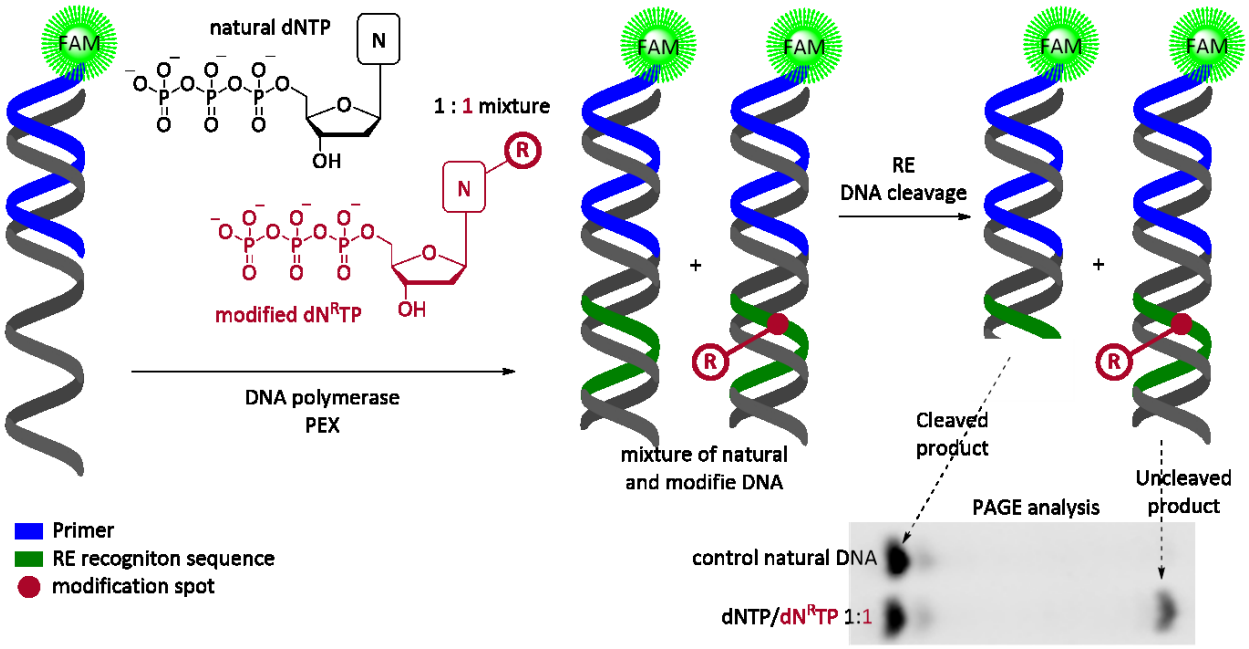


Figure 21. Designed assay for the study of competitive incorporations of base-modified dN^{R}TPs in the presence of natural dNTPs.

3.2.1.2. Criteria of choice of studied chemical modifications and synthesis of dN^{R}TPs

In the previous work performed by Pavel Kielkowski, the competitive incorporation of several 7-substituted-7-deazaadenosine-5'-*O*-triphosphates and 5-substituted-cytidine-5'-*O*-triphosphates in the presence of dATP and dCTP, were studied. With the exception of 7-

deazaadenosine (**dA^{7d}TP**) and 7-ethynyladenosine and 7-ethynylcytidine (**dA^{e}TP** and **dC^{e}TP**), all the dN^{R}TPs employed in the study bore aryl substituents. Aryl modifications were chosen of different size and electronic density, starting from a simple unsubstituted phenyl group (for **dA^{ph}TP** and **dC^{ph}TP**), and 3-aminophenyl (**dA^{nh2}TP** and **dC^{nh2}TP**) or 3-nitrophenyl substituents (**dA^{no2}TP** and **dC^{no2}TP**), to nucleotides bearing bulkier fluorescent labels such as **dA^{e}TP** and **dC^{e}TP**, or very bulky groups for **dA^{str}TP** and **dC^{str}TP**, as well as reactive groups in the case of **dC^{ft}TP** (structures in Figure 22).[128]

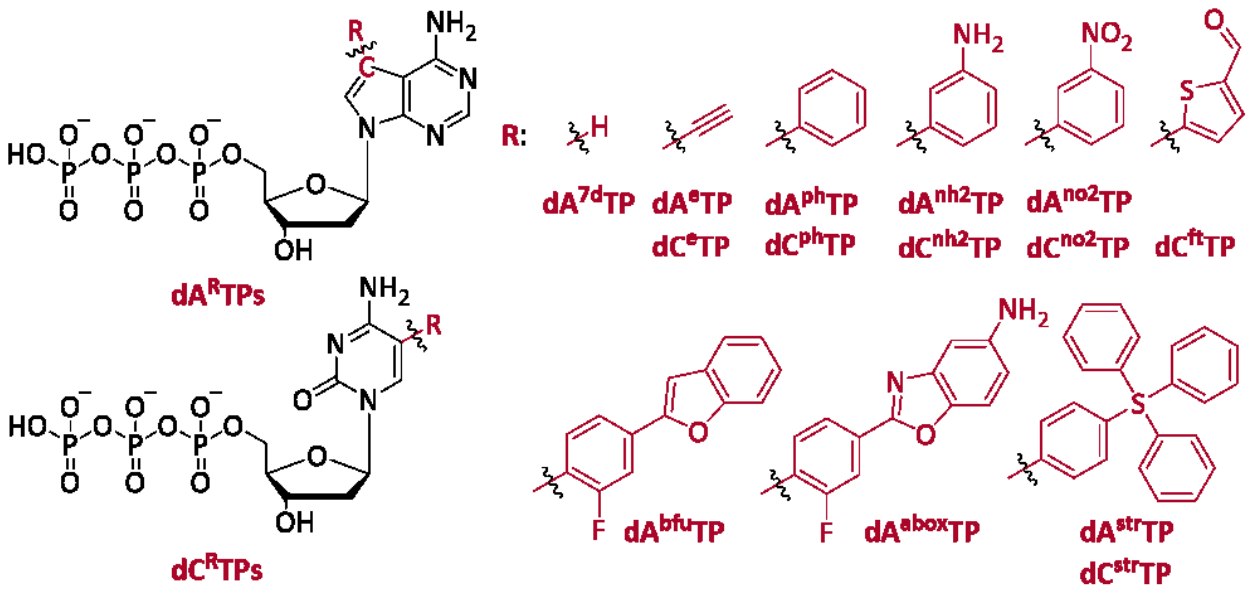


Figure 22. Structures of modified dA^{R}TPs and dC^{R}TPs employed by Pavel Kielkowski in his former work on competitive enzymatic incorporations of dNTPs into DNA.[128]

All these dN^{R}TPs were tested in competitive PEX experiments, and for most of the study the Bst DNA polymerase was used. A deeper study of **dA^{ph}TP** was performed, since this dNTP was proved to be an excellent substrate for the Bst polymerase. When an equimolar mixture of **dA^{ph}TP** and **dATP** were employied in PEX, 67 % of the product contained the phenylated nucleotide (Table 2). Its competitive incorporation was also tested using different DNA polymerases, including prokaryotic thermoresistent polymerases (such as Pwo, Taq, KOD XL Klenow and Vent(exo-) DNA polymerases), and the human polymerase α. A good acceptance of this nucleotide by all tested polymerases was shown.

Similarly, other 7-aryl-7-deaza-dATPs analogues were optimum substrates of the Bst DNA polymerase, and were often incorporated into DNA significantly better than natural **dATP**. The corresponding 5-aryl-dCTP analogues, however were only comparable or worse substrates for polymerases compared to natural **dCTP** (Table 2). Enzymatic kinetics of incorporation confirmed 7-aryl-dA^{R}TPs to be very good substrates of Bst DNA polymerase, thanks to an increased affinity of these 7-aryl nucleotides to the active site of the enzyme in comparison to **dATP**, and the measured Michaelis-Menten constants (K_M) for 7-aryl-dA^{R}TPs were often lower than the one of natural **dATP**.

Table 2. Competitive competitions of modified dA^{R}TPs and dC^{R}TPs, and enzymatic kinetics, performed by Pavel Kielkowski.[128]

dNTP	Comp.[(a)]	K_M[(b)]	dNTP	Comp.[(a)]	K_M[(b)]	dNTP	Comp.[(a)]	K_M[(b)]
dATP	-	28	dA^{no2}TP	70	6.5	dCTP	-	4.4
dA^{7d}TP	32	60	dA^{bfu}TP	67	5.7	dC^{e}TP	54	3.1
dA^{e}TP	67	13	dA^{abox}TP	73	8	dC^{ph}TP	58	2.4
dA^{ph}TP	67	6.8	dA^{str}TP	17	>1000	dC^{nh2}TP	32	21
dA^{nh2}TP	61	7						

(a): Results obtained in competitive PEX experiments employing a 1: 1 mixture of dA^{R}TP and **dATP**; (b): Michaelis-Menten constant. All data shown in the Table are taken from the reference.[128]

However, this study was not fully explanatory, since most modifications were either very small (dN^{7d}TP) or very bulky (most aryl dN^{R}TPs). Most dNTPs were selected on the basis of their role and applications (e.g. fluorescent or redox labels), and not focusing on exploring the DNA polymerase mechanism. Moreover, no base-modified dG^{R}TPs nor dU^{R}TPs were included in the study.

Expanding this study to a more systematic, comprehensive and concise selection of substrates would allow a better elucidation of the mechanism of DNA polymerases. Therefore, the choice of modifications present on dN^{R}TPs was of crucial importance, and directly connected to the desired outcome of the study.

Once the DNA replication starts and the polymerase-primer-template complex is formed, a significantly large space is present in the major groove direction of the forming DNA, where modifications can be accommodated. The aim of this study was, mainly, to understand how different types of modifications can be guested within this pocket. Moreover, functional groups present in the side chains of amino acids could stabilize (or destabilize) the complex formed between the polymerase and the incoming dN^{R}TP, by interactions with the chemical modification.

At first, dN^{R}TPs bearing modifications of increasing size were studied, to evaluate whether the bulkiness of the modification would have some influence on the rate of incorporation of the nucleotide. At the same time, modifications with increasing number of π–electrons were chosen for testing π-π, π-cation or π-anion interactions between the modification on dN^{R}TPs and the residues of amino acids present in the active pocket of the enzyme.

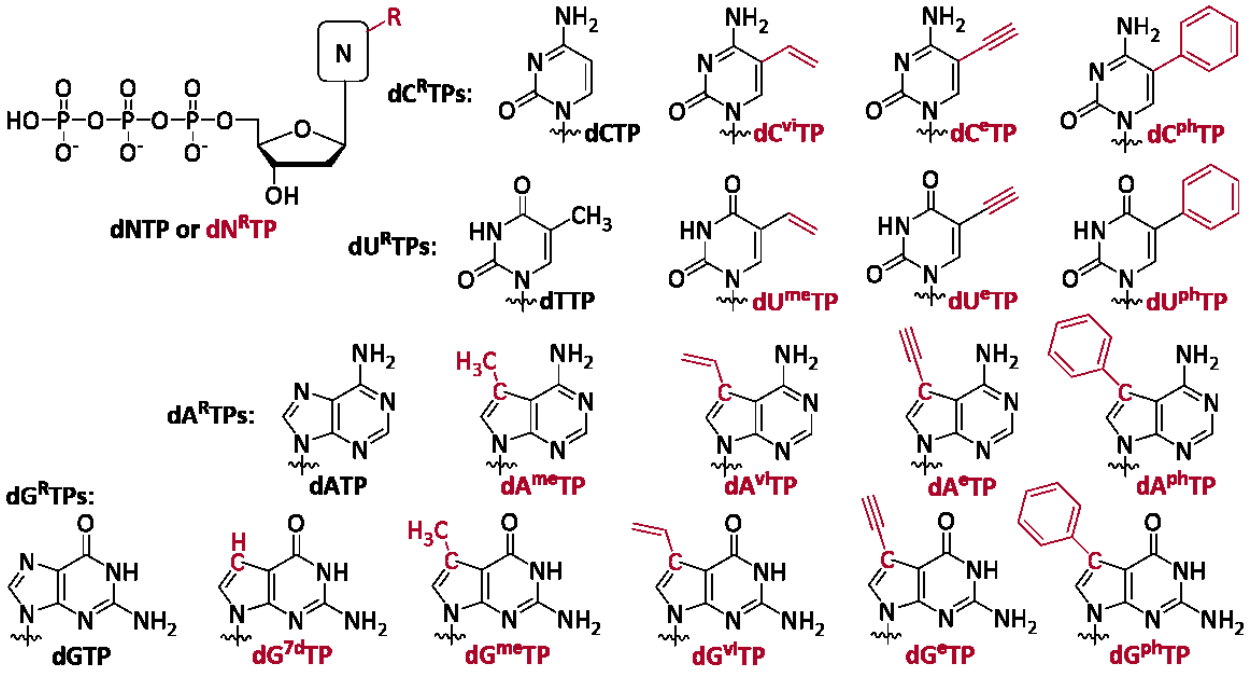


Figure 23. Structures of all the dN^{R}TPs employed in competitive PEX experiments.

All the employed dN^{R}TPs are shown in Figure 23, listed in order with increasing size and number of π–electrons of the modification. A series of pymirimines carrying vinyl (**dC^{vi}TP** and **dU^{vi}TP**)[15],

ethynyl (**dC^{e}TP** and **dU^{e}TP**)[12] or phenyl (**dC^{ph}TP** and **dU^{ph}TP**)[12] modifications were employed in the study. Similar series of modifications were studied on purines (**dA^{vi}TP**,[15] **dG^{vi}TP**,[13] **dA^{e}TP**,[14] **dG^{e}TP**,[13] **dA^{ph}TP**[14] and **dG^{ph}TP**[13]), with the addition of the 7-methyl modified substrates (**dA^{me}TP** and **dG^{me}TP**), carrying hence a smaller non-polar modification with no π-electrons. The unsubstituted **dG^{7d}TP** was also employed, to allow a comparison with **dA^{7d}TP** previously studied by Pavel Kielkowski. The only modification of this substrate consisted of having the N-7 atom of **dGTP** substituted with a carbon atom, in this way removing a possible interaction between N-7 and an amino acid side-chain within the polymerase active site.

7-Methyl-7-deazapurine-5'-*O*-triphosphates, as well as **dG^{7d}TP**, are commercially available. All other triphosphates were synthesized freshly before their employment following reported procedures.[12–15] Vinyl- and phenyl- dN^{R}TPs were synthesized starting from 5-iodopyrimidine-5'-*O*-triphosphates or 7-iodo-deazapurine-5'-*O*-triphosphates, and performing a water-phase Suzuki coupling employing potassium vinyltrifluoroborate or phenylboronic acid as the coupling partner. On the other hand, ethynyl nucleosides dN^{e} were obtained by Sonogashira coupling reactions of 5-iodo pyrimidines or 7-iodo-7-deaza purines with trimethylsilylacetylene and subsequent deprotection of the alkyne by reaction with potassium carbonate in methanol. 5'-*O*-triphosphorilation with $POCl_3$ in trimethylphosphate, followed by addition of tributylammonium pyrophosphate, and quenching with TEAB, provided the final dN^{e}TPs.

3.2.3. Competitive PEX using Bst large fragment DNA polymerase

The investigation of the effect of the size and electron density of the modification on substrates activity was performed employing mixtures of various ratios of modified dN^{R}TPs and natural dNTPs in competitive PEX (See Figure 21). Sequences of employed primers and templates for the study of each modifications were designed in order to allow a single modification within the recognition sequence of the employed RE, and are reported in Figure 24A. DNA cleavage by EcoRI is known to be blocked by most of the major groove DNA modification,[12–15,129] and for this reason it was chosen as the RE for the study of modifications on dU, dA and dG. Some modifications on dC were found not to block completely the cleavage by EcoRI, thus, they were studied employing AfeI as the RE, which recognition sequence still allows the incorporation of a single modified dC^{R}TP within it.

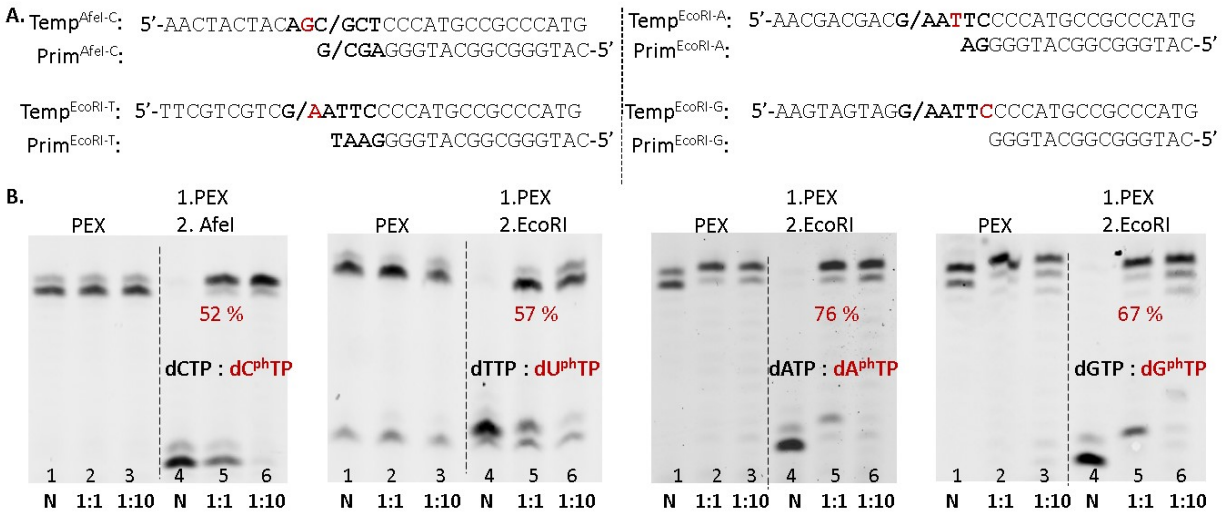


Figure 24. **A.** Sequences of primers and templates used in the competitive PEX experiments; **B.** Example of PAGE analysis of competitive PEX experiments of dN^{ph}TPs.

All modified dN^{R}TPs were studied in a 1: 1 mixture with their natural counterparts, or in higher excess such as 1: 10 ratio with respect to their relative natural dNTPs. The Bst large fragment DNA polymerase was chosen for their competitive incorporation. All results are reported in Table 3.

Generally, the size of the modification had only a slight effect on the activity of the substrates, while π–electronic systems highly influenced the rate of incorporation, and, despite their size, all dNTPs bearing a phenyl group were found to be better substrates than natural dNTPs.

7-Substituted-7-deazapurine dN^{R}TPs were found to be better substrates than 5-modified dC^{R}TPs and dU^{R}TPs bearing the same modifications. Interestingly, most of the 7-modified-7-deaza-dA^{R}TPs and dG^{R}TPs were very well accepted by the Bst DNA polymerase and were found to be better substrates than natural **dATP** and **dGTP**, respectively.

The Bst DNA polymerase showed the highest affinity for 7-phenyl-7-deaza purine triphosphates **dA^{ph}TP** and **dG^{ph}TP**, which 1: 1 mixtures with **dATP** or **dGTP** respectively resulted into the formation of 76 % or 67 % phenylated-DNA when employed in competitive PEX (Table 3 and Figure 24B). Vinyl-dATP and ethynyl-dATP and dGTP were also optimum substrates, and PEX experiments using 1: 1 mixtures of these dN^{R}TPs with natural **dATP** or **dGTP** yielded higher amounts of modified-DNA than natural-DNA. **dG^{vi}TP** and all purine dN^{R}TPs bearing modification with no π-electrons were worse substrates than their natural counterparts, although still very good substrates of the Bst large fragment DNA polymerase.

Table 3. Full table of results obtained in competitive PEX experiments of dN^{R}TPs in the presence of natural dNTPs in 1: 1 ratio (**Comp. 1:1**) or 1: 10 ratio (**Comp. 1:10**).

dC^{R}TP	Comp. 1: 1	Comp. 1: 10	dU^{R}TP	Comp. 1: 1	Comp. 1: 10	dA^{R}TP	Comp. 1: 1	Comp. 1: 10	dG^{R}TP	Comp. 1: 1	Comp. 1: 10
									dG^{7d}TP	33	77
						dA^{me}TP	35	78	dG^{me}TP	42	82
dC^{vi}TP	49	88	dU^{vi}TP	41	74	dA^{vi}TP	70	86	dG^{vi}TP	39	67
dC^{e}TP	49	86	dU^{e}TP	28	63	dA^{e}TP	72	88	dG^{e}TP	58	81
dC^{ph}TP	52	86	dU^{ph}TP	57	79	dA^{ph}TP	76	88	dG^{ph}TP	67	90

On the other hand, modified pyrimidine dNTPs were generally worse than the purine derivatives bearing the same modifications. Only phenyl-modified pyrimidines were found to be better substrates then the natural dNTPs and when the DNA obtained by PEX using 1 : 1 **dCTP** / **dC^{ph}TP** or **dTTP** / **dU^{ph}TP** mixtures was cleaved by the REs, uncleaved DNA was observed on PAGE gels in respectively 52 % and 57 %. Uridine derivatives resulted to be the worst substrate (41 %

incorporation for **dU^{vi}TP**, and only 28 % for **dU^{e}TP**). On the other hand, vinyl- and Ethynyl-dCTPs were comparable to **dCTP**.

The same trend was observed when the modified dN^{R}TPs were employed in high excess with respect to the relative dNTPs (10: 1 ratio), and competition values increased in the series for rich π–electronic systems, confirming the high activity of phenyl-modified dN^{ph}TPs as substrates for the employed polymerase. In fact, when using a 10: 1 excess of dN^{R}TP with respect to the relative natural dNTP, a large portion of modified DNA was always formed (63-90 %).

An attempt to further prove the higher affinity of the Bst large fragment polymerase towards phenylated dN^{ph}TPs, in comparison to other modified substrates, was performed by a direct comparison of **dG^{ph}TP** and **dG^{7d}TP** in competitive PEX experiments (Figure 25A). This direct competition between these two triphosphates was designed in order to obtain a strong confirmation that the higher extent of incorporation of phenyl-substituted dN^{ph}TPs is due to the effect of the phenyl substituent itself. Both **dG^{7d}TP** and **dG^{ph}TP** lack the nitrogen atom at the position 7 of the purine ring, and only **dG^{ph}TP** bears a phenyl ring at that position, being this ring the only difference between these two substrates. The differences in competitive PEX experiments between **dG^{ph}TP** and **dGTP** could be attributed also to the presence of the nitrogen atom in position 7 in the natural substrate, while the use of **dG^{7d}TP** instead of **dGTP** eliminates this doubt.

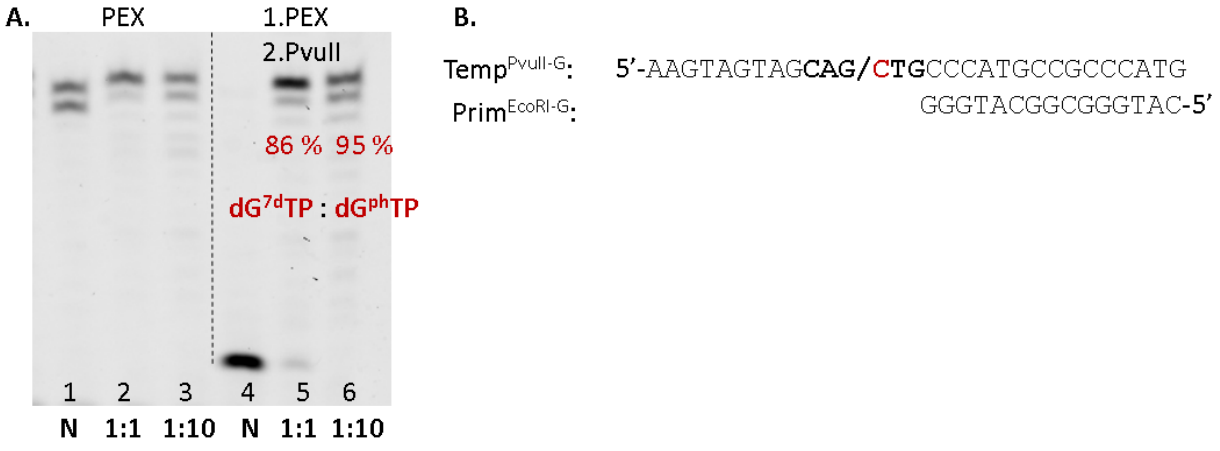


Figure 25. **A.** direct competitive PEX experiment of **dG^{7d}TP** versus **dG^{ph}TP** using PvuII; **B.** Sequences used for the direct competitive PEX experiment.

This direct competitive incorporation yielded 86 % of phenylated DNA, in full agreement with what expected based on results previously obtained (67 % **dG^{ph}TP** vs **dGTP** and 33 % **dG^{7d}TP** vs **dGTP**, Table 3). For this experiment, it was necessary to use a different RE, in particular PvuII, which is known to cleave **dG^{7d}** modified DNA, but not DNA containing **dG^{ph}** within its recognition sequence.[13] The employed primer and template are reported in Figure 25B.

3.2.3.1. Competitive PCR using Bst large fragment DNA polymerase

Furthermore, a competitive PCR experiment was performed in order to test phenylated **dG^{ph}TP** and **dU^{ph}TP** as substrates in more complex processes. It is important to note that the Bst large fragment polymerase is not a thermostable polymerase and would not survive the PCR cycles. It was, thus, necessary to replace it and employ KOD XL DNA polymerase for this study. A 339 bp template was designed in order to contain the EcoRI recognition sequence at approximately half of its sequence (to allow a good visualization on gel). **dU^{ph}TP** was a very good substrate for KOD XL in PCR experiments and, after the cleavage of the DNA obtained using a 1: 1 mixture of **dTTP** and **dU^{ph}TP** by EcoRI, the formation of 74 % of modified DNA was observed through agarose gel analysis (Figure 26A). However, the EcoRI recognition sequence (GAATTC) contains two **dA** nucleotides, against which thymidines (or, in this case, **dU^{ph}TP**) are incorporated. Thus, the 74 % ratio between uncleaved and cleaved DNA observed on gel does not represent the percentage of incorporation of a single **dU^{ph}**, since the presence of the modification at a single site (independently one to the other), or at both sites, would prevent the cleavage by EcoRI. Thus, a quantification of the competition ratio would be incorrect. However, the experiment clearly showed a significant incorporation of **dU^{ph}TP** in the presence of **dTTP** in various ratio.

On the other hand, the incorporation of **dG^{ph}TP** in PCR was unexpectedly challenging, and the right PCR amplicon was not obtained when high ratio of the modified dNTP was used (Figure 26B). However, this can be attributed to the difficult reading by the polymerase through the **dG^{ph}**-modified template in the PCR cycles, more than to a troubling incorporation (which was proved to be excellent in PEX experiments). The quantification of competition of this substrate was, however, impossible.

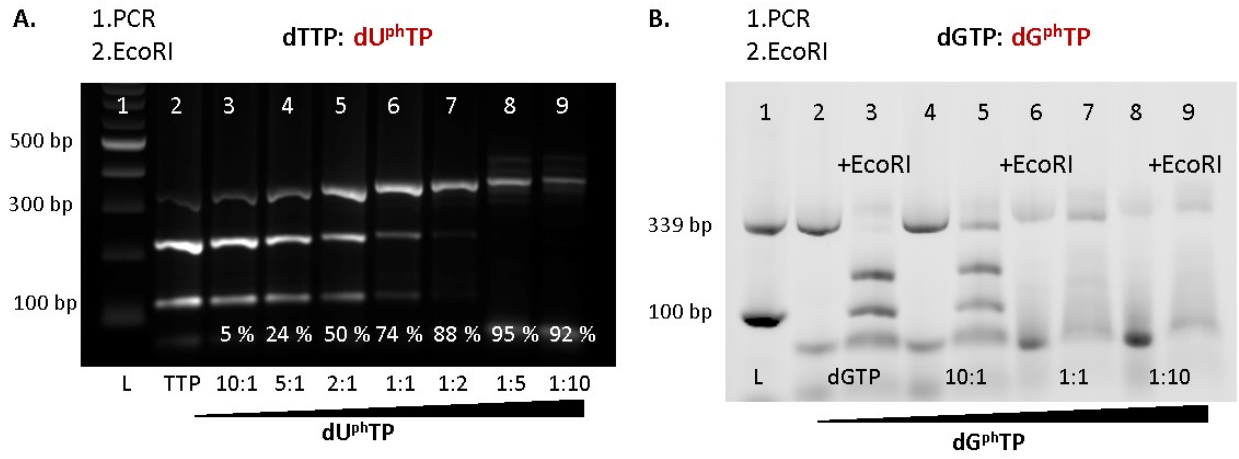


Figure 26. Agarose gel analysis of competitive PCR incorporations using different ratios of **A. dTTP** and **dU^{ph}TP**; **B. dGTP** and **dG^{ph}TP**.

3.2.3.2. Competitive PEX with different DNA polymerases

Competitive PEX experiments of ethynyl- and phenyl-modified uracil and deazaguanine (**dU^{e}TP**, **dU^{ph}TP**, **dG^{e}TP** and **dG^{ph}TP**) were also performed using different DNA polymerases, such as KOD XL, Pwo and Vent (exo-) DNA polymerases. The high tendency of DNA polymerases to incorporate the selected modified substrates was confirmed in most cases. Specifically, when equimolar mixtures of **dGTP** and **dG^{ph}TP** were used in competitive PEX experiments (Table 4, and selected examples of PAGE analysis in Figure 27), all the studied polymerases provided high percentage of modified DNA as product (41 to 66 %), with Bst large fragment and KOD XL preferring the modified dNTP to the natural one (respectively 66 % and 58 % incorporation of **dG^{ph}TP**).

Similarly, the alkynyl derivative **dG^{e}TP** was well tolerated by all polymerases, and preferred to **dGTP** when Bst large fragment or KOD XL were employed (58 % and 51 % formation of the modified product). The use of Vent (exo-) DNA polymerase led to the formation of a slightly inferior amount of modified product (24 %).

Table 4. Summary of the results obtained in competitive PEX experiments of **dU^{e}TP**, **dU^{ph}TP**, **dG^{e}TP** and **dG^{ph}TP** using different DNA polymerases.

	Bst large fragment		KOD XL		Pwo		Vent(exo-)	
	1/1	1/10	1/1	1/10	1/1	1/10	1/1	1/10
dU^{e}TP	28	63	37	88	26	53	20	60
dU^{ph}TP	57	80	17	57	6	57	3	34
dG^{e}TP	58	81	51	88	37	61	24	70
dG^{ph}TP	66	90	58	89	43	80	41	84

As shown before, the Bst large fragment polymerase tolerated uridine modification worse than modifications at any other nucleotide. **dU^{e}TP** resulted to be a worse substrate than natural **dTTP** also for the other studied polymerases (20 – 37 % formation of the modified product). **dU^{ph}TP** was found to be a very difficult substrate for Pwo and Vent (exo-) DNA polymerases with only 6 and 3 % incorporation respectively.

When a larger excess of the modified dNTP was used (1: 10 ratio dNTP/dN^{R}TP), a higher amount of modified DNA was produced in all cases. Modified uridines were confirmed to be worse substrates than modified guanosines, especially when Pwo or Vent (exo-) were employed as polymerases.

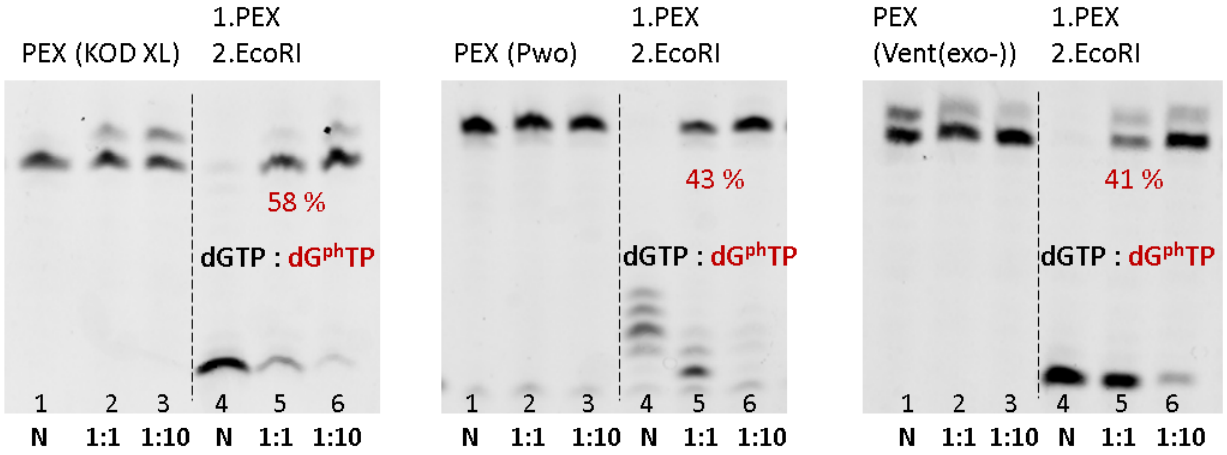


Figure 27. Examples of PAGE analysis of competitive PEX experiments using **dG^{ph}TP** with different DNA polymerases.

3.2.3.3. Sequence-dependent competitive PEX experiments

Four different series of primers and templates were designed in such way that the modified dN^{R}TP was incorporated following a different nucleobase in each case (the employed primers and templates are listed under Table 5). In this way, it was possible to determine whether the competitive incorporation of modified dNTPs was sequence dependent, and study the influence of the primer and template sequences on the incorporation. Competitive PEX experiments of phenylated **dU^{ph}TP** and **dG^{ph}TP** were performed with Bst DNA polymerase.

Almost no differences were observed for **dU^{ph}TP** incorporations. All the sequences provided 54 to 57 % of modified DNA when a 1: 1 ratio with **dTTP** was employed, and 79 to 93 % when a 10: 1 ratio was used. The incorporation of **dG^{ph}TP** with Bst large fragment polymerase resulted slightly more sequence-dependent. Using a 1: 1 starting mixture of **dGTP** and **dG^{ph}TP**, the phenylated product was obtained in 67 % yield when the modification was incorporated after a **dG** nucleotide, while the formation of only 43 % modified DNA was observed when a **dA** nucleotide preceded the incorporation of **dG^{ph}TP**.

Table 5. Results obtained in the sequence-dependent competitive PEX experiments.

Sequence	5'-...GG^{ph}...-3'		5'-...AG^{ph}...-3'		5'-...CG^{ph}...-3'		5'-...TG^{ph}...-3'	
RE	EcoRI		BglII		AfeI		RsaI	
dNTP/dN^{R}TP	1:1	1:10	1:1	1:10	1:1	1:10	1:1	1:10
Competition:	67	90	43	87	58	74	56	85
Sequence	5'-...TU^{ph}...-3'		5'-...AU^{ph}...-3'		5'-...CU^{ph}...-3'		5'-...GU^{ph}...-3'	
RE	EcoRI		SphI		SacI		KpnI	
dNTP/dN^{R}TP	1:1	1:10	1:1	1:10	1:1	1:10	1:1	1:10
Competition:	57	79	55	93	54	89	56	91

Temp$^{EcoRI\text{-}G}$: 5'-AAGTAGTAG**G/AATTC**CCCATGCCGCCCATG
Prim$^{EcoRI\text{-}G}$: GGGTACGGCGGGTAC-5'

Temp$^{BglII\text{-}G}$: 5'-AAGTAGTAG**A/GATCT**CCCATGCCGCCCATG
Prim$^{BglII\text{-}G}$: GGGTACGGCGGGTAC-5'

Temp$^{AfeI\text{-}G}$: 5'-AAGTAGTAG**AGC/GCT**CCCATGCCGCCCATG
Prim$^{AfeI\text{-}G}$: CGAGGGTACGGCGGGTAC-5'

Temp$^{RsaI\text{-}G}$: 5'-AAGTAGTAGT**GT/AC**ACCCATGCCGCCCATG
Prim$^{RsaI\text{-}G}$: GGGTACGGCGGGTAC-5'

Temp$^{EcoRI\text{-}T}$: 5'-TTCGTCGTCG**/AATTC**CCCATGCCGCCCATG
Prim$^{EcoRI\text{-}T}$: TAAGGGGTACGGCGGGTAC-5'

Temp$^{SphI\text{-}T}$: 5'-TTCGTCGTC**GCATG/C**CCCATGCCGCCCATG
Prim$^{SphI\text{-}T}$: GGGTACGGCGGGTAC-5'

Temp$^{SacI\text{-}T}$: 5'-TTCGTCGTC**GAGCT/C**CCCATGCCGCCCATG
Prim$^{SacI\text{-}T}$: GGGTACGGCGGGTAC-5'

Temp$^{KpnI\text{-}T}$: 5'-TTCGTCGTC**GGTAC/C**CCCATGCCGCCCATG
Prim$^{KpnI\text{-}T}$: GGGTACGGCGGGTAC-5'

Modifications within the sequences are highlited in red, while the nucleotides preceding the modifications in the elongation direction are highlited in blue.

3.2.4. Enzymatic Kinetics of single nucleotide incorporation

The study of kinetics of single nucleoside incorporations of modified dN^{R}TPs can be a very interesting tool to better understand the behavior that DNA polymerases adopt when facing an unusual non-natural substrate. A first important parameter to be studied is the Michaelis-Menten constant (K_M), which delineates the affinity of the substrate to the active site of the enzyme. Modified dN^{R}TPs which are incorporated more efficiently than their natural counterparts in competition assays, generally had lower values of K_M, indicating a higher affinity to the enzyme with respect to the relative natural dNTP.

On the other hand, the velocity (*V*) is describing how fast the enzyme can incorporate the substrates into DNA at a certain concentration of the substrate itself. The maximum velocity (V_{max}) of the enzyme can be mathematically extrapolated at very high concentration of the substrate. This is useful to calculate the rate of incorporation of the modified nucleotides, which can be described by the enzyme turnover number (k_{cat}), directly proportional to the V_{max} of the enzyme. High values of k_{cat} correspond to high rates of incorporation of the substrates by the polymerase.

Dividing k_{cat}/K_M it is possible to obtain a value that directly measure the catalytic efficiency of the enzyme, considering both the affinity of the substrate to the enzyme and the velocity of incorporation. This value (in here referred to as the discrimination ratio) can be calculated by normalizing the efficiency of incorporation of the modified dN^{R}TP by the efficiency of incorporation of the natural dNTP.

Enzyme steady-state kinetics were studied employing a template which was a single nucleotide longer than the primer (Figure 28A). In this way, it was not necessary to add any other dNTP in the reaction mixture, which could interfere with the study by interacting with the polymerase and occupying temporarily the active pocket of the enzyme, thus altering the substrate-enzyme affinity and restraining the incorporation of the modified dNTP. Steady-state kinetics experiments were, hence, conducted incubating primer, template and a constant concentration of Bst DNA polymerase (1.375 fM, 0.01 U) at 65 °C with an increasing amount of dN^{R}TP. Experiments were quickly quenched after 3 minutes using a stop solution containing formamide

and products were separated using 20 % denaturating PAGE. Selected examples of PAGE analysis are shown in Figure 28B, where a progressive formation of the PEX product can be noticed.

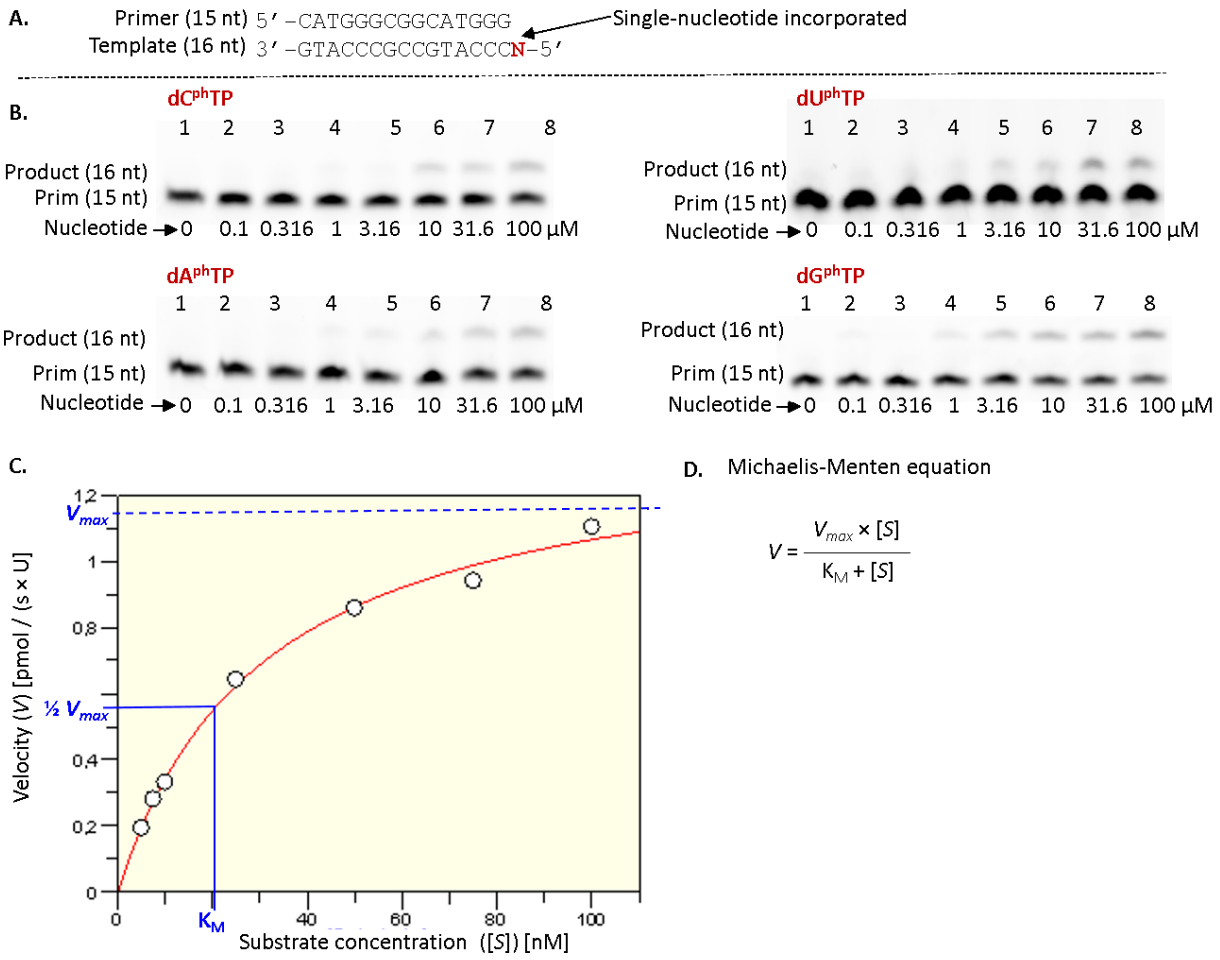


Figure 28. **A.** Sequences of primer and templates used for the study of enzymatic kinetics; **B.** PAGE analysis of kinetics of single-nucleotide incorporation of **dN^{Ph}TPs**; **C.** Relation between the velocity (*V*) and the substrate concentration ([*S*]); **D**. The Michaelis-Menten equation.

Densitometric analysis of the PAGE gels can provide the relative intensities of all the bands. Referring to the band relative to substrate concentration 0 where, consequently, no product is formed (and the whole integration is due to the starting DNA), it is possible to calculate the amount of product obtained in 3 minutes at all different starting concentrations of substrate (also knowing the initial amount of employed DNA). Dividing the amount of product (in pmol) by Enzyme concentration (in U) multiplied by the reaction time (180 sec), the velocity of the enzyme

is obtained, for that precise starting concentration of substrate (V_0). Plotting V_0 against the substrate concentration, a curve is obtained, from which the V_{max} can be extrapolated (Figure 28C). At this point, the K_M values, which mathematically indicate the concentration of substrate necessary to have ½ of V_{max}, can be determined by fitting the obtained V_{max} to the Michaelis-Menten equation (Figure 28D, this was done using the software using KaleidaGraph).

All values obtained by this study are shown in Table 6. In general, all base-modified pyrimidine nucleotides (dC^{R}TPs and dU^{R}TPs) had values of K_M higher than the relative natural dNTP, indicating a lower affinity to the enzyme by the modified substrates. The only exception was the phenyl-substituted **dC^{ph}TP**, with a K_M value of 6, with respect of 6.7 of natural **dCTP**, confirming one more time the high influence of a phenyl on the incoming nucleotide.

Purine nucleotides bearing small modifications, such as **dA^{vi}TP**, **dG^{7d}TP**, **dG^{me}TP**, **dG^{vi}TP**, also showed lower affinity to the polymerase with respect to **dATP** and **dGTP**, respectively. On the other hand, electron rich ethynyl- and phenyl-modified adenosines and guanosines had lower K_M values than **dATP** and **dGTP** showing a great affinity of these substrates to the Bst DNA polymerase. Interestingly, also the methylated **dA^{me}TP** had a low K_M value (Table 6).

Table 6. Steady-state kinetics of incorporation of natural and base-modified dNTPs.

dN^{R}TP	K_M	k_{cat}	k_{cat}/K_M	Ratio[a]	dN^{R}TP	K_M	k_{cat}	k_{cat}/K_M	Ratio[a]
dCTP	6.7	6.1	0.91	1	dTTP	3.2	6.6	2.06	1
dC^{vi}TP	14.1	15.4	1.1	1.2	dU^{vi}TP	13.1	17.7	1.35	0.7
dC^{e}TP	15.1	10.4	0.69	0.8	dU^{e}TP	6.3	4.6	0.73	0.4
dC^{ph}TP	6	7.5	1.25	1.4	dU^{ph}TP	7.9	9.3	1.18	0.7
dATP	8.7	8.0	0.99	1	dGTP	5.3	13.1	2.47	1
dA^{me}TP	6.1	5.5	0.9	0.9	dG^{7d}TP	12.5	12.1	0.97	0.4
dA^{vi}TP	10.9	7.1	0.65	0.7	dG^{me}TP	11.1	12.7	1.14	0.5
dA^{e}TP	7.3	6.3	0.88	0.9	dG^{vi}TP	9.5	13.8	1.45	0.6
dA^{ph}TP	5.6	13.5	2.41	2.4	dG^{e}TP	7.5	17.5	2.33	0.9
					dG^{ph}TP	4.1	13.7	3.34	1.4

(a) The ratio is calculated as $(k_{cat}/K_M)_{modified} / (k_{cat}/K_M)_{natural}$.

It is interesting to note that most of the modified triphosphates have a discrimination ratio lower than the one of their relative natural dNTP, with the exception made for **dC^{vi}TP** and the

phenylated nucleotides **dC^{ph}TP**, **dA^{ph}TP** and **dG^{ph}TP**. Most often, a low discrimination ratio was due to very high values of K_M, indicating that the affinity of the dN^{R}TPs to the polymerase is the parameter influencing in the larger extent the rate of incorporation.

As expected, the K_M values for the phenyl-modified **dC^{ph}TP**, **dA^{ph}TP** and **dG^{ph}TP** were lower than the ones of **dCTP**, **dATP** and **dGTP** respectively. This is in perfect agreement with the high percentage of modified nucleotide incorporated in competition assays described above (52 % for **dC^{ph}TP**, 76 % for **dA^{ph}TP**, and 67 % for **dG^{ph}TP**; see paragraph 3.2.3., Table 3), which is, therefore, probably due the higher affinity of 5-phenyl cytidine and 7-phenyl-substituted 7-deazapurine for the active site of the complex of polymerase with primer and template. The lower value obtained in competitive PEX experiments for **dC^{ph}TP**, in comparison with phenyl purines, could be attributed to a lower k_{cat}, while the even worse result of **dU^{ph}TP** can be attributed to a K_M value higher than the one of natural **dTTP**.

Either way, it is important to state that the enzymatic incorporation of nucleotides is a very complex process, where several other mechanisms are taking place, and its efficiency does not depend only on the affinity of substrates to the enzyme and the rate of incorporation. For example, the presence or absence of proofreading can highly influence the experiment outcomes. The incorporation of dNTPs following the modified dN^R, for the further extension of the primer to reach the whole length of the template, can also be influenced by the presence of the modification, and slightly alter the results.

3.2.5. Molecular modelling to study the affinity dN^{ph}TPs - Bst DNA polymerase

To explain the high affinity of 7-phenyl-substituted 7-deazapurines to the Bst DNA polymerase, especially in comparison to the affinity of 5-phenyl pyrimidine nucleotides, and the fact that they were better substrates than their natural counterparts, a molecular modelling study was performed. This study was fully performed by Dr. Jindrich Fanfrlik (IOCB, Prague).

A crystal structure of Bst DNA polymerase is available in the literature.[130] Primer, template and a dNTP were docked within this crystal structure, to obtain the complex to be modeled using as a computational tool a semiempirical quantum mechanical scoring function (Figure 29A).[131,132] The scoring function is a mathematical tool to predict with good approximation the affinity between two molecules, and, in this particular case, the obtained scores for each dNTP are an indication of the affinity of the substrate itself to the tertiary complex polymerase-primer-template: the larger the affinity, the more negative the score. 5-Phenyl-substituted uridine was confirmed to have lower affinity to the polymerase with respect to natural **dTTP**, but also to 2'-deoxyuridine-5'-*O*-triphosphate (**dUTP**), showing a higher score than both **dTTP** and **dUTP** (Table 7). On the other hand, **dC^{ph}TP** had a score only slightly more negative than **dCTP**, showing a similar affinity of the two triphosphates to the enzyme, and confirming a 52 % incorporation in competitive PEX (52 %, Table 7).

Table 7. Scores of natural dNTPs and phenylated dN^{ph}TPs.

dNTP	Score	Relative score	dNTP	Score	Relative score
dCTP	-68.4		dATP	-86.9	
dC^{ph}TP	-70.2	-1.8	dA^{ph}TP	-97.4	-10.5
dUTP	-72.4		dGTP	-108.5	
dTTP	-70.5		dG^{ph}TP	-129.0	-20.5
dU^{ph}TP	-61.8	10.5 (to **dUTP**), 8.7 (to **dTTP**)			

7-phenyl-substitute purines showed a very high affinity to the polymerase with scores strongly more negative than **dATP** and **dGTP**, especially for **dG^{ph}TP** which had the lower score among the studied dNTPs. The higher percentage of incorporation of **dA^{ph}TP** with respect to **dG^{ph}TP** in

competitive PEX experiments was confirmed to depend on a higher rate of incorporation of the phenylated adenosine (higher k_{cat}) more than on a higher affinity of the substrate.

A virtual scan was performed docking **dG^{ph}TP** within the polymerase active pocket, in order to study the effect of selected amino acids on the stabilization of the substrate, to which its high affinity to the enzyme is due. Looking at the crystal structure of the Bst DNA polymerase, it is possible to recognize three main amino acids within the active site, which possess charged or aromatic side chains, allowing strong interactions with the incoming dNTPs. In particular, the arginine in position 629 (Arg629) has a cationic side chain, while the phenylalanine in position 710 (Phe710) and the tyrosine in position 714 (Tyr714) have aromatics rings allowing possible interactions with their π-electrons. In the virtual scan, these three amino acids were substituted one by one by a glycine (Gly), which possesses no side chain. In this way it was possible to evaluate the contribution of each of the three selected amino acids to the stabilization of **dG^{ph}TP**, and compare it to natural **dGTP** (Figure 29B). When either the Phe710 or the Tyr714 were substituted with a glycine in this virtual glycine scan, no substantial change in binding energy was recorded between **dGTP** and **dG^{ph}TP**. On the other hand, when the Arg629 was replaced with a glycine, the stabilization of **dG^{ph}TP** dropped significantly in comparison with **dGTP**, suggesting a strong binding between this arginine and the modification on the nucleotide, involving thus a cation-π interaction between the guanidinium cation of Arg629 and the phenyl ring of **dG^{phT}P**.

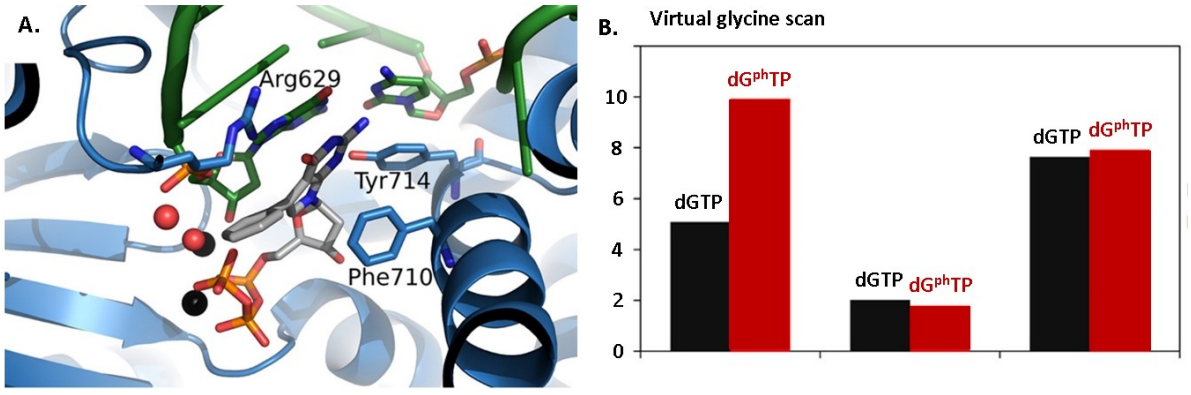


Figure 29. **A.** Molecular modelling of **dG^{ph}TP** docked inside the active site of Bst DNA polymerase with primer and template; **B.** Energy contributions of the selected Arg629, Phe710 and Tyr714 amino acids to the satbilization through binding of **dGTP** (in black) and **dG^{ph}TP** (in red) within the active pocket of Bst DNA polymerase, obtained by the virtual glycine scan.

3.2.6. Discussion of the competitive incorporation of base-modified dNTPs

The performed systematic study of competitive PEX experiments employing base-modified dN^{R}TPs bearing modifications of increasing bulkiness in the presence of their natural counterparts showed, in the first place, that all the employed modifications are well tolerated by the studied polymerases (Bst large fragment, KOD XL, Pwo and Vent(exo-) DNA polymerases). In general, modified purines were accepted better than the pyrimidines bearing the same modification. 7-deazapurine dNTPs bearing π-electron-containing substituents (ethynyl and phenyl, as well as 7-vinyl-7-deazaadenine) are generally better substrates of Bst polymerase than natural **dATP** or **dGTP**, respectively. The corresponding 5-substituted cytosine dC^{R}TPs resulted to be comparable to **dCTP**, whereas the 5-substituted uracil dU^{R}TPs are generally worse substrates than **dTTP**. Most important, all the studied base-modified dNTPs were incorporated into DNA when all natural dNTPs were present in reaction mixtures.

Enzyme kinetics helped to assign the better incorporation trend of dN^{R}TPs containing π-electron rich modifications, and in particular 7-phenyl-7-deazapurine dNTPs, to a high affinity of these substrates to the active site of the polymerase, with respect of the one of their natural counterparts. K_M values for these substrates were lower than the ones the relative natural dNTP. The performed molecular modelling also confirmed the higher affinity of the 7-phenyl-deazapurine nucleotides to the active site and revealed that the most significant contributing factor is the cation–π interaction of the modified dN^{ph}TP with Arg629.

These results help the elucidation of the mechanism of incorporation of base-modified nucleotides into DNA by DNA polymerases. The very good activity of 7-substituted-7-deazapurine dA^{R}TPs and dG^{R}TPs, as well as the only slightly worse activity of 5-substituted dC^{R}TPs and dU^{R}TPs as substrates for DNA polymerases, is very encouraging for further development of methods of polymerase synthesis of modified DNA. The formation, in all cases, of modified DNA when dN^{R}TPs were employed in the presence of their natural counterparts allows for potential *in cellulo* or even *in vivo* applications.

4. Conclusions

A series of diverse Cu-chelated Clip-Phen-linked TFO-AMN hybrids was designed and synthesized, in order to achieve desirable site-selective DNA cleavage.[104] Their synthesis was straightforward. The choice of a CuAAC reaction as the key step allowed to easy couple the two main units (TFO and AMN). In this way, it was possible and relatively easy to plan a wide range of probes, differing in the nature of the tethers between TFO and AMN, and in the positions of the AMN within the TFO stretch. This confirmed that the selected approach is of great convenience, with respect of all methods previously reported to prepare TFO-AMN conjugates.

The design of two clickable AMNs (**PgCP** and **N_3CP**) can also constitute an advantage allowing for a future fast development of conjugates targeting different polypurine segments, simply by changing the TFO sequence.

The two new clickable thymidine derivatives **dU^{peg}** and **dU^{un}** were designed as an aid to enhance the specificity in DNA cleavage by the AMN hybrids. However, their use in DNA synthesis can be of great advantage in all situations in which a post-synthetic CuAAC step is desired.[124] The two long and flexible alkynyl tethers provided a good enhancement of the efficiency of the CuAAC, especially in the case of the peg-linked alkyne which reacted approximatively twice as fast as all the other studied alkynes (including the standardly used octadiyne). This is of great importance, since it allows better conversions of reactions on valuable modified DNA, and shorter exposition of the starting DNA to oxidative damage caused by Cu^{I} ions necessary for the CuAAC.

A further relevant advantage of the use of **dU^{peg}** consists of the efficient incorporation of its relative nucleotide (**$dU^{peg}TP$**) into DNA, even in PCR reaction where long and heavily-modified products are produced, thanks to the hydrophilic nature of its linker. For comparison, PCR employing the octadiynyl-linked **$dU^{o}TP$** yielded a substantially lower amount of DNA, and **$dU^{un}TP$** provided PCR products only when a very short template was used.

All synthesized conjugates were shown to form stable triplexes performing a systematic study of the hybridization of TFO-AMNs with a dsDNA containing the TFO recognition sequence. The optimized conditions for triplex annealing were selected in order to be as close as possible to a physiologic environment, with the prospective of an *in cellulo* application. However, the low pH

(6.2) and high salt concentration (1 mM Mg^{2+}, 250 Na^{+}) were necessary to form triple helices stable enough for the study. Further improvement in this sense could consists in a further chemical modification of TFOs, to stabilize triplexes even in physiologic conditions.[133–136]

The efficiency and selectivity of cleavage of a target DNA versus an off-target DNA by all synthesized TFO-AMNs was systematically studied. In the optimized conditions, TFOs bearing the AMN tethered by the flexible linkers (either to the 5'-end, or to an internal thymine base) showed a significant cleavage of the target duplex (up to 34 %) and no off-target cleavage. Higher cleavage was also possible by tuning the reaction conditions. However, the off-target DNA cleavage was also enhanced.

In most cases, no clear footprinting cleavage bands were observed. The flexibility of the employed linkers, as well as the diffusible free radical generation by the Cu^{I}-AMNs during the cleavage, did not allow the DNA degradation to occur with a single-nucleotide resolution.

The prepared AMN–TFO conjugates, hence, do not have the cleavage specificity of the enzymatic methods, such as TALENs, ZFNs and, above all, CRISPR/Cas. This type of radical DNA damage lacking a single-nucleotide resolution prevents their use in gene editing, since a subsequent recombination process (HR or NHJR) would be impossible.

On the other hand, the synthesized Clip-Phen linked TFO-AMNs can be of great interest if no subsequent recombination is wanted. An efficient and specific targeting and degradation of genes can potentially constitute a very important tool for selective gene knock-out, for example to study and recognize the main roles of that precise gene.

In this way, systems able to selectively degrade a specific gene could constitute an alternative to other gene silencing techniques which are often difficult and inefficient, such as the antigene strategy. Therefore, potential applications of the developed Clip-Phen molecular scissors can be found in this area. The choice of a target sequence present in a viral genome (HIV-1) immediately suggests the prospect of targeting polypurine segments of viral genes, to inhibit the synthesis of vital protein in the virus life cycle (e.g. for replication, or structural proteins). Similarly, genes of the hosts, which the virus needs and uses to replicate within the cell, could be knocked-out. Moreover, a selective degradation of cancer-related genes could provide the inhibition of the synthesis of a selected protein needed by cancer cells, such as antiapoptotic proteins.

Therefore, this work constitute an important step towards a methodology of preparation and evaluation of very important tools for gene therapy, such as specific DNA cutters. In this sense, further efforts must be addressed in improving both the efficiency and specificity of the TFO-AMN conjugates.

In the second part of this work, a systematic study of competitive PEX experiments using a series of 7-substituted 7-deazapurine and 5-substituted pyrimidine dN^{R}TPs in the presence of all natural dNTPs was performed.[41] All dN^{R}TPs bore substituents of increasing bulkiness and number of π-electron pairs. All the studied modified dN^{R}TPs were good to excellent substrates for several prokaryotic DNA polymerases, such as Bst Large Fragment, KOD XL, Pwo and Vent(exo-) DNA polymerases. 7-Deazapurine dNTPs bearing π-electron-containing substituents (ethynyl and phenyl, as well as 7-vinyl-7-deazaadenine) were found to be better substrates of Bst polymerase than natural **dATP** or **dGTP**, respectively. The corresponding 5-substituted-dC^{R}TP were comparable substrates for DNA polymerases to **dCTP**, whereas the 5-substituted-dU^{R}TPs were generally worse substrates than **dTTP**. Nevertheless, all base-modified dN^{R}TPs were successfully incorporated into DNA even in the presence of natural dNTPs.

These results were confirmed by studying the kinetics of single-nucleotide incorporation. All phenylated dN^{ph}TPs had higher affinity to the Bst DNA polymerase than their natural counterparts. Molecular modelling helped attributing this high affinity to a strong π-cation interaction between the phenyl ring of the incoming dN^{ph}TP and an arginine residue present inside the active pocket of the Bst DNA polymerase. These findings were in good agreement with a previous study performed by Pavel Kielkowski, who showed how many aryl-modified-dA^{R}TPs were better substrates of several polymerases, with respect to **dATP**.[128]

Combining all these results, a wide knowledge of the polymerase incorporation of base-modified dNTPs is achieved. The successful incorporation of all modified dN^{R}TPs into ONs in the presence of their natural counterparts is of great importance. In this way, the expectation of their *in vivo* incorporation into genomic DNA is very high, if a satisfactory delivery of dN^{R}TPs to the cell nucleus is accomplished. Therefore, the outcome of this study allows for a better planning and development of methods for enzymatic synthesis of base-modified DNA *in cellulo* and *in vivo*.

5. Experimental

5.1. Chemical synthesis

Materials and methods

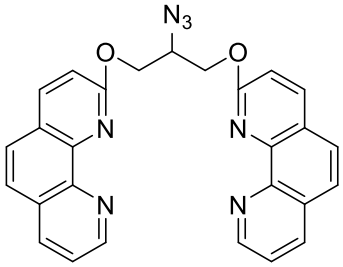
Reagents and solvents were purchased from Fluorochem, Sigma–Aldrich and AlfaAesar. Benzene, Et_3N and CH_2Cl_2 were dried by distillation over CaH_2. *N*,*N*-diisopropylethylamine (DIPEA) and *N*,*N*-dimethylformamide (DMF) were dried by distillation at reduced pressure over CaH_2. Tetrahydrofuran (THF) was dried by distillation over Na. CH_3CN was degassed under a positive stream of argon just before its use. Unless otherwise stated, all the reactions were performed under a positive atmosphere of argon by standard syringe, cannula and septa techniques. *N*-bromosuccinimide was purified by crystallization from water. 3',5'-di-*O*-acetylthymidine (**10**),[1] 2-[2-(prop-2-yn-1-yloxy)ethoxy]ethan-1-ol (**11a**),[125] 2-clip-phenanthroline (Clip-Phen),[69] 2-chloro-1,10-phenanthroline (**2**),[137] , 1-*H*-imidazole-1-sulfonyl azide (**4**),[106] **dC^{o}** and **dC^{o}TP**,[112] **dC^{vi}TP**,[15] **dC^{e}** and **dC^{e}TP**,[12] **dC^{ph}TP**,[12] **dU^{am}** and **dU^{am}TP**,[113] **dU^{o}** and **dU^{o}TP**,[112] **dU^{vi}TP**,[15] **dU^{e}** and **dU^{e}TP**,[12] **dU^{ph}TP**,[12] **dG^{me}TP**,[13] **dG^{vi}TP**,[13] **dG^{e}TP**,[13] **dG^{ph}TP**,[13] **dA^{me}TP**,[138] **dA^{vi}TP**,[15] **dA^{e}TP**,[14] **dA^{ph}TP**,[14] were prepared as described previously, or from the relative nucleoside synthesized as described in the references, followed by classic 5'-*O*-triphosphorilation.[114] **dG^{7d}TP** was purchased by Jena Bioscience.

Reactions were monitored by thin-layer chromatography (TLC) using Merck silica gel 60 F_{254} plates and visualized by UV (254 nm). Column chromatography was performed using silica gel (40–63 μm). Reverse-phase high-performance flash chromatography (HPFC) purifications were done on a Biotage SP1 apparatus with C-18 columns. Purifications of nucleoside triphosphates were performed using HPLC (Waters modular HPLC system) on a column packed with 10 μm C18 reversed phase (Phenomenex, Luna C18 (2) 100 Å). NMR spectra were measured on a Bruker AVANCE 500 (^{1}H at 500.0 MHz, ^{13}C at 125.7 MHz and ^{31}P at 202.4 MHz) in CD_3OD, CD_3CN, $CDCl_3$, DMSO-d_6 or D_2O solutions at 25 °C. Chemical shifts (in ppm, δ scale) were referenced to the residual solvent signal in ^{1}H spectra or to the solvent signal in ^{13}C spectra. Coupling constants (*J*) are given in Hz. The complete assignment of ^{1}H and ^{13}C signals was performed by an analysis of

the correlated homonuclear H,H-COSY, and heteronuclear H,C-HSQC and H,C-HMBC spectra. High resolution mass spectra were measured on a LTQ Orbitrap XL spectrometer (Thermo Fisher Scientific). IR spectra were measured on a Bruker Alpha FT-IR spectrometer using attenuated total reflection (ATR).

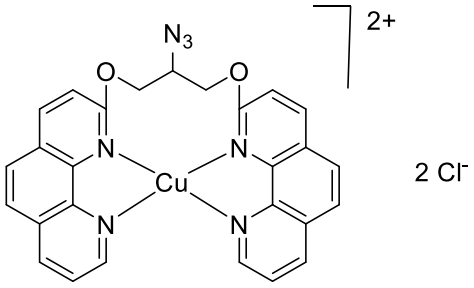
Synthesized compounds

2,2'-((2-azidopropane-1,3-diyl)bis(oxy))bis(1,10-phenanthroline), (N_3CP)



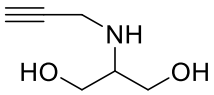
2-clip-phenanthroline (Clip-Phen, 200mg, 0.447 mmol),[69] 1*H*-imidazole-1-sulfonyl azide (**4**, 141 mg, 0.670 mmol), K_2CO_3 (46.6 mg, 0.335 mmol) and $CuSO_4$ pentahydrate (1.3 mg, 0.005 mmol) were stirred in MeOH (5 mL) at room temperature for 12 hours under aerobic atmosphere. After evaporation of the solvent under reduced pressure, the crude product was purified by flash chromatography (Eluent EtOAc : MeOH 9:1) to afford a yellowish solid. The obtained solid was suspended in hot MeOH and stirred for 20 min, after filtration compound **N_3CP** was recovered as a pale yellow solid (118 mg, 56%). m.p.: 96.1 – 98.9 °C with decomposition. ^{1}H NMR (500.0 MHz, DMSO-d_6): 4.68 (tt, 1H, ^3J = 6.7, 4.2, CHN); 4.93 (dd, 1H, ^2J = 11.4, ^3J = 6.7, CH_a**H_b**O); 5.01 (dd, 1H, ^2J = 11.4, ^3J = 6.7, CH_a**H_b**O); 7.33 (d, 2H, $J_{3,4}$ = 8.7, H-3); 7.73 (dd, 2H, $J_{8,7}$ = 8.1, $J_{8,9}$ = 4.3, H-8); 7.87 (d, 2H, $J_{6,5}$ = 8.7, H-6); 7.96 (d, 2H, $J_{5,6}$ = 8.7, H-5); 8.44 (d, 2H, $J_{4,3}$ = 8.7, H-4); 8.46 (dd, 2H, $J_{7,8}$ = 8.1, $J_{7,9}$ = 1.8, H-7); 9.10 (dd, 2H, $J_{9,8}$ = 4.3, $J_{9,7}$ = 1.8, H-9). ^{13}C NMR (125.7 MHz, DMSO-d_6): 59.32 (CHN); 65.59 (CH_2O); 113.44 (CH-3); 123.21 (CH-8); 124.40 (CH-6); 125.07 (C-4a); 126.42 (CH-5); 129.02 (C-6a); 136.32 (CH-7); 140.22 (CH-4); 143.75 (C-10b); 144.62 (C-10a); 149.90 (CH-9); 161.60 (C-2). HRMS calc. for $C_{27}H_{20}O_2N_7$: 474.16730, found: 474.16762.

2,2'-((2-azidopropane-1,3-diyl)bis(oxy))bis(1,10-phenanthroline) copper(II) complex (Cu-N_3CP)



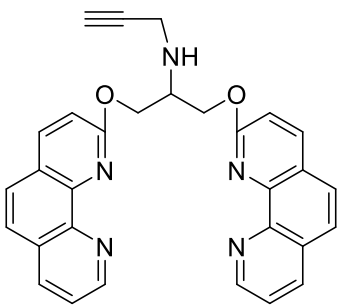
N_3CP (650 mg, 1.372 mmol) and a stoichiometric amount of $CuCl_2$ (187 mg, 1.237 mmol) were stirred in DMF (20 mL) at room temperature until a clear solution was obtained. The desired copper (II) complex **Cu-N_3CP** was thus precipitated by adding 200 mL of Et_2O and obtained after filtration as a green solid. HRMS calc. for $C_{27}H_{19}O_2N_7ClCu$: 571.05803, found: 571.05793.

***N*-propargyl serinol (6)**



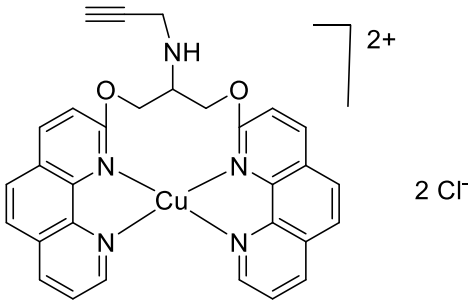
Serinol (**3**, 1g, 10.976 mmol) and K_2CO_3 (3.03g, 21.952 mmol) were suspended in anhydrous THF (40 mL) under argon atmosphere. Propargyl bromide (**5**, 80% solution in toluene, 1.22 mL, 10.976 mmol) was slowly added by syringe and the reaction mixture was stirred for 6 hours at 60°C. After evaporation of the solvent under reduced pressure, the crude product was purified by flash chromatography (Eluent EtOAc: MeOH 8:2) to afford the desired compound **6** as a yellowish oil (724 mg, 72 %). ^{1}H NMR (400 MHz, DMSO-*d6*): δ 4.44 (bs, 2H); 3.43 – 3.25 (m, 6H); 3.03 (t, *J* = 2.5 Hz, 1H); 2.68 (q, *J* = 5.7 Hz, 1H). HRMS calc. for $C_6H_{12}O_2N$: 130.08626, found: 130.08626.

***N*-(1,3-bis((1,10-phenanthroline-2-yl)oxy)propan-2-yl)prop-2-yn-2-amine (PgCP)**



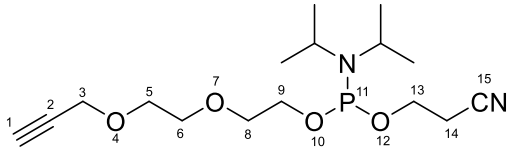
N-propargylserinol (**6**, 301 mg, 2.329 mmol) was dissolved in anhydrous THF (30 mL) under argon atmosphere and the solution was cooled at 4°C. Sodium hydride (112 mg, 4.659 mmol) was added and the mixture was stirred for 1 hour at 4°C, finally 2-chloro-1,10-phenanthroline[137] (**2**, 1g, 4.659 mmol) was added and the reaction mixture was allowed to slowly reach room temperature and vigorously stirred overnight. The reaction was carefully quenched with MeOH (10 mL) and after evaporation of the solvent under reduced pressure, the crude product was purified by flash chromatography (MeOH 10% to 30% in EtOAc) to afford the desired compound **PgCP** as a pale yellow solid (1.13g, 73%). ^{1}H NMR (400 MHz, DMSO-d_6) δ 9.13 – 9.07 (m, 2H), 8.46 (d, *J* = 8.2 Hz, 2H), 8.37 (d, *J* = 8.7 Hz, 2H), 7.93 (d, *J* = 8.7 Hz, 2H), 7.84 (d, *J* = 8.7 Hz, 2H), 7.73 (dd, *J* = 8.1, 4.3 Hz, 2H), 7.27 (d, *J* = 8.7 Hz, 2H), 4.90 – 4.72 (m, 4H), 3.73 (s, 1H), 3.70 – 3.63 (m, 2H), 3.03 (t, *J* = 2.2 Hz, 1H). HRMS calc. for $C_{30}H_{24}O_2N_5$ = 486.19245, found: 486.19260.

_N_-(1,3-bis((1,10-phenanthroline-2-yl)oxy)propan-2-yl)prop-2-yn-2-amine copper(II) complex (Cu-PgCP)



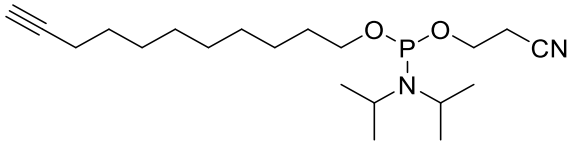
PgCP (680 mg, 1.401 mmol) and a stoichiometric amount of $CuCl_2$ (188.3 mg, 1.401 mmol) were suspended in DMF (20 mL) and stirred at room temperature until a clear solution was obtained. The desired copper(II) complex (**Cu-PgCP**, *Figure 5*) was precipitated by adding 200 mL of Et_2O and obtained after filtration as a green solid. HRMS calc. for $C_{30}H_{23}O_2N_5ClCu$: 583.08308, found: 583.08321.

***O*-2-[2-(prop-2-yn-1-yloxy)ethoxy]ethyl-*O*'-(2-cyanoethyl)-*N*,*N*-diisopropylphosphoramidite (13a)**



2-[2-(prop-2-yn-1-yloxy)ethoxy]ethan-1-ol (**11a**, 238 mg, 1.651 mmol) was co-evaporated three times with anhydrous pyridine and dissolved in 5 mL of anhydrous CH_2Cl_2. Freshly distilled diisopropylethylamine (2.5 equiv., 4.127 mmol, 719 µL) was added, followed by 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite (1.2 equiv, 1.981 mmol, 442 µL), and the reaction mixture was stirred at room temperature for 1.5 hour. The crude product was diluted with anhydrous CH_2Cl_2 (50 mL) and quickly washed with saturated aqueous KCl (20 mL), dried over Na_2SO_4. Flash chromatography (cyclohexane : EtOAc 1:3) provided the desired compound as a clear oil (332 mg, 59%). ^{1}H NMR (600.1 MHz, CD_3CN): 1.14 – 1.19 (m, 12H. (C**H$_3$**)$_2$CH); 2.60 – 2.68 (m, 2H, H-14); 2.69 (t, 1H, $J_{1,3}$ = 2.5, H-1); 3.56 – 3.64 (m, 8H, H-5,6,8, $(CH_3)_2$C**H**); 3.65 – 3.71 (m, 1H, H-9b); 3.72 – 3.84 (m, 3H, H-9a,13); 4.15 (d, 2H, $J_{3,1}$ = 2.5, H-3). ^{13}C NMR (150.9 MHz, CD_3CN): 20.90 (d, $J_{C,P}$ = 6.8, CH_2-14); 24.86, 24.88 (d, $J_{C,P}$ = 7.3, (**C**H_3)$_2$CH)); 43.68 (d, $J_{C,P}$ = 12.4, $(CH_3)_2$**C**H)); 58.61 (CH_2-3); 59.31 (d, $J_{C,P}$ = 18.9, CH_2-13); 63.51 (d, $J_{C,P}$ = 17.3, CH_2-9); 69.80 (CH_2-5); 70.86 (CH_2-6); 71.77 (d, $J_{C,P}$ = 7.5, CH_2-8); 75.58 (CH-1); 80.80 (C-2); 119.39 (C-15). ^{31}P{^{1}H} NMR (162.3 MHz, CD_3CN): 150.70. HRMS (ESI$^+$): calculated for $C_{16}H_{30}O_4N_2P$ = 345.19377; found: 345.19384. Calculated for $C_{16}H_{29}O_4N_2NaP$ = 367.17572; found: 367.17572.

***O*-10-Undecynyl-*O*'-(2-cyanoethyl)-*N*,*N*-diisopropylphosphoramidite (13b)**



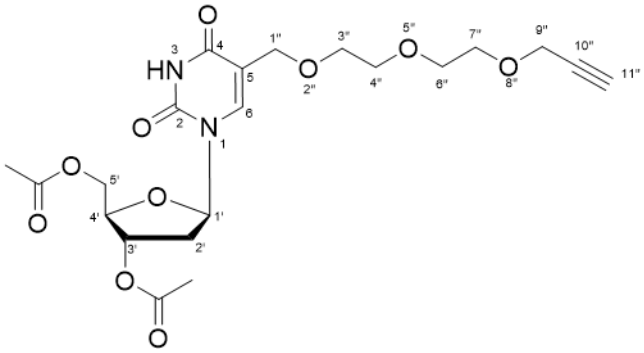
10-Undecyn-1-ol (**11b**, 320 mg, 1,902 mmol) was co-evaporated three times with anhydrous pyridine and dissolved in 5 mL of anhydrous CH_2Cl_2. Freshly distilled diisopropylethylamine (2.5 equiv., 4.754 mmol, 828 µL) was added followed by 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite (1.2 equiv, 2.282 mmol, 510 µL) and the reaction mixture was stirred at room temperature for 1.5 hours. The crude product was diluted with anhydrous CH_2Cl_2

(50 mL) and quickly washed with saturated aqueous KCl (20 mL), dried over Na_2SO_4. Flash chromatography (cyclohexane : EtOAc 85:15) provided the desired compound as a clear oil (526 mg, 75%). ^{1}H NMR (600.1 MHz, CD_3CN): 1.16, 1.17 (2 × d, 2 × 6H, (C**H$_3$**)$_2$CH); 1.27 – 1.41 (m, 10H, H-5,6,7,8,9); 1.44 – 1.53 (m, 2H, H-4); 1.54 – 1.61 (m, 2H, H-10); 2.11 (t, 1H, $J_{1,3}$ = 2.7, H-1); 2.16 (td, 2H, $J_{3,4}$ = 7.1, $J_{3,1}$ = 2.7, H-3); 2.57 – 2.68 (m, 2H, H-16); 3.54 – 3.68 (m, 4H, H-11, $(CH_3)_2$C**H**); 3.69 – 3.83 (m, 2H, H-15). ^{13}C NMR (150.9 MHz, CD_3CN): 18.73 (CH_2-3); 20.98 (d, $J_{C,P}$ = 6.9, CH_2-14); 24.90 (d, $J_{C,P}$ = 7.3, (**C**H_3)$_2$CH)); 26.59, 29.23, 29.37, 29.68, 29.89, 30.12 (CH_2-4,5,6,7,8,9); 31.89 (d, $J_{C,P}$ = 7.2, CH_2-10); 43.65 (d, $J_{C,P}$ = 12.3, $(CH_3)_2$**C**H)); 59.13 (d, $J_{C,P}$ = 19.0, CH_2-15); 64.26 (d, $J_{C,P}$ = 17.2, CH_2-11); 69.48 (CH-1); 85.33 (C-2); 119.31 (C-15). ^{31}P{^{1}H} NMR (162.3 MHz, CD_3CN): 149.60. HRMS (ESI$^+$): calculated for $C_{20}H_{37}O_2N_2P$ = 369.26654; found: 369.26666. Calculated for $C_{20}H_{38}O_2N_2NaP$ = 391.24849; found: 391.24862.

General procedure A: synthesis of 3',5'-di-*O*-acetyl-protected nucleosides 14c and 14d

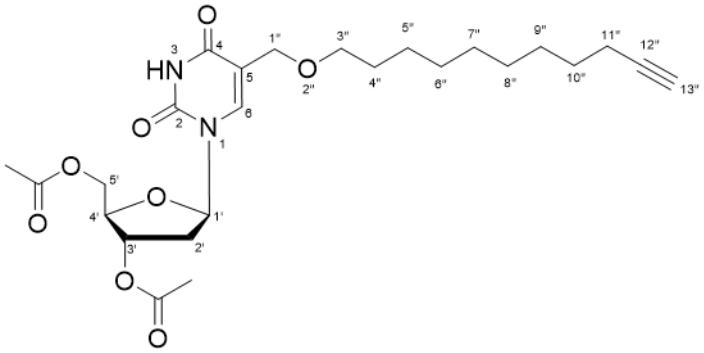
3',5'-Di-*O*-acetylthymidine was dissolved in 10mL of anhydrous benzene. Freshly recrystallized NBS (1,2 equiv.) and AIBN (0.12 equiv.) were added under argon, the reaction mixture was stirred at 60°C and the reaction proceeding was monitored by TLC (Eluent: Cyclohexane : EtOAc 1 : 1). When a complete conversion was reached (approx. 3 h), the solvent was evaporated under reduced pressure at a rotary evaporator and the residue was re-dissolved in 10 mL of anhydrous DMF, then, a solution of 2-[2-(prop-2-yn-1-yloxy)ethoxy]ethan-1-ol (**11a**, 1.2 equiv.) in anhydrous DMF (10 mL) or 10-undecyn-1-ol (**11b**,1.2 equiv.) was added via syringe and the reaction mixture was stirred at r.t. overnight. DMF was removed under reduced pressure and the residue was re-dissolved in EtOAc (200mL) to be washed with H_2O (2 × 100 mL) and brine (1 × 100 mL), and dried over Na_2SO_4. The reaction products were purified by reverse-phase HPFC.

α-[2-[2-(Prop-2-yn-1-yloxy)ethoxy]ethoxy]-3',5'-di-*O*-acetyl-thymidine (12a)



The title product was obtained by the reaction of 3',5'-di-*O*-acetylthymidine (500 mg, 1.53 mmol), NBS (1,2 equiv., 1,836 mmol, 327 mg), AIBN (0.12 equiv., 0.184 mmol, 30 mg) and a solution of 2-[2-(prop-2-yn-1-yloxy)ethoxy]ethan-1-ol (**11a**, 1.2 equiv., 1,836 mmol, 265 mg) in anhydrous DMF, following the general procedure A. Reverse phase HPFC, using a linear gradient of methanol (5→40%) in water over 30 min, afforded the desired compound as a white foam (373 mg, 52%). ^{1}H NMR (500.0 MHz, $CDCl_3$): 2.11, 2.13 (2 × s, 2 × 3H, CH_3COO); 2.22 (ddd, 1H, J_{gem} = 14.2, $J_{2'b,1'}$ = 8.6, $J_{2'b,3'}$ = 6.6, H-2'b); 2.43 (t, 1H, $J_{11'',9''}$ = 2.4, H-11''); 2.47 (ddd, 1H, J_{gem} = 14.2, $J_{2'a,1'}$ = 5.7, $J_{2'a,3'}$ = 2.0, H-2'a); 3.63 – 3.72 (m, 8H, H-3'',4'',6'',7''); 4.19 (d, 2H, $J_{9'',11''}$ = 2.4, H-9''); 4.25 (td, 1H, $J_{4',5'}$ = 3.7, $J_{4',3'}$ = 2.0, H-4'); 4.28 (dd, 1H, J_{gem} = 12.9, $J_{1''b,6}$ = 1.3, H-1''b); 4.30 – 4.37 (m, 3H, H-1''a, H-5'); 5.22 (dt, 1H, $J_{3',2'}$ = 6.6, 2.0, $J_{3',4'}$ = 2.0, H-3'); 6.31 (dd, 1H, $J_{1',2'}$ = 8.6, 5.7, H-1'); 7.56 (t, 1H, $J_{6,1''}$ = 1.3, H-6); 8.95 (s, 1H, NH-3). ^{13}C NMR (125.7 MHz, $CDCl_3$): 20.75, 20.90 (**C**H_3COO); 37.53 (CH_2-2'); 58.35 (CH_2-9''); 63.88 (CH_2-5'); 65.23 (CH_2-1''); 69.04, 70.335, 70.341, 70.36 (CH_2-3'',4'',6'',7''); 74.38 (CH-3'); 74.56 (CH-11''); 79.60 (C-10''); 82.32 (CH-4'); 85.31 (CH-1'); 112.47 (C-5); 136.65 (CH-6); 150.00 (C-2); 162.11 (C-4); 170.35 (CH_3**C**OO). HRMS (ESI$^+$): calculated for $C_{21}H_{28}O_{10}N_2Na$ = 491.16362; found: 491.16364.

α-(Undec-10-yn-1-yloxy)thymidine-3',5'-di-*O*-acetylthymidine (12b)



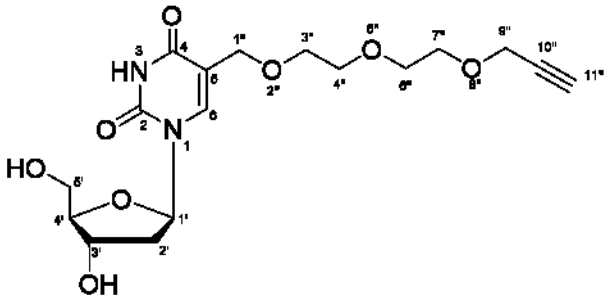
The title product was obtained by the reaction of 3',5'-di-*O*-acetylthymidine (500 mg, 1.53 mmol), NBS (1,2 equiv., 1,836 mmol, 327 mg), AIBN (0.12 equiv., 0.184 mmol, 30 mg) and 10-undecyn-1-ol (**11b**, 1.2 equiv., 1,836 mmol, 352 µL) following the general procedure A. Reverse phase HPFC, using a linear gradient of methanol (20→70%) in water over 30 min, afforded the desired compound as a white solid (429 mg, 57%). ^{1}H NMR (500.0 MHz, $CDCl_3$): 1.22 – 1.32 (m, 8H, H-5″,6″,7″,8″); 1.32 – 1.41 (m, 2H, H-9″); 1.46 – 1.53 (m, 2H, H-10″); 1.53 – 1.59 (m, 2H, H-4″); 1.92 (t, 1H, $J_{13'',11''}$ = 2.7, H-13″); 2.10, 2.11 (2 × s, 2 × 3H, CH_3COO); 2.16 (td, 2H, $J_{11'',10''}$ = 7.1, $J_{11'',13''}$ = 2.7, H-11″); 2.19 (ddd, 1H, J_{gem} = 14.1, $J_{2'b,1'}$ = 8.6, $J_{2'b,3'}$ = 6.2, H-2'b); 2.47 (ddd, 1H, J_{gem} = 14.1, $J_{2'a,1'}$ = 5.7, $J_{2'a,3'}$ = 1.8, H-2'a); 3.45 – 3.53 (m, 2H, H-3″); 4.19 (dd, 1H, J_{gem} = 13.0, $J_{1''b,6}$ = 1.3, H-1″b); 4.25 (ddd, 1H, $J_{4',5'}$ = 3.8, 3.2, $J_{4',3'}$ = 2.3, H-4'); 4.27 (dd, 1H, J_{gem} = 13.0, $J_{1''a,6}$ = 1.3, H-1″a); 4.30 (dd, 1H, J_{gem} = 12.1, $J_{5'b,4'}$ = 3.2, H-5'b); 4.34 (dd, 1H, J_{gem} = 12.1, $J_{5'a,4'}$ = 3.8, H-5'a); 5.22 (ddd, 1H, $J_{3',2'}$ = 6.2, 1.8, $J_{3',4'}$ = 2.3, H-3'); 6.32 (dd, 1H, $J_{1',2'}$ = 8.6, 5.7, H-1'); 7.53 (t, 1H, $J_{6,1''}$ = 1.3, H-6); 9.46 (s, 1H, NH-3). ^{13}C NMR (125.7 MHz, $CDCl_3$): 18.31 (CH_2-11″); 20.69, 20.87 (**C**H_3COO); 25.95, 28.38, 28.63, 28.96, 29.30, 29.34, 29.41 (CH_2-4″,5″,6″,7″,8″,9″,10″); 37.59 (CH_2-2'); 63.90 (CH_2-5'); 64.65 (CH_2-1″); 68.05 (CH-13″); 71.36 (CH_2-3″); 74.36 (CH-3'); 82.26 (CH-4'); 84.71 (C-12″); 85.17 (CH-1'); 112.84 (C-5); 136.09 (CH-6); 150.18 (C-2); 162.33 (C-4); 170.20, 170.34 (CH_3**C**OO). HRMS (ESI$^+$): calculated for $C_{25}H_{37}O_8N_2$ = 493.25444; found: 493.25466.

General procedure B: deprotection of 3',5'-di-*O*-acetyl-protected nucleosides dU^{peg} and dU^{un}

Precursors **12a** or **12b** were dissolved in 40 mL of a 1:1 mixture of 1,4-dioxane and 27% aqueous NH_3, and stirred vigorously at r.t. for 48 hours. The solvent was evacuated under reduced pressure and the crude product was re-dissolved in CH_2Cl_2 (200 mL), washed with water (2 × 100

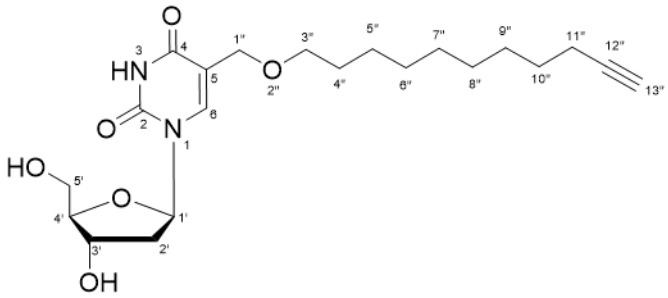
mL) and brine (1 × 100 mL), dried over Na_2SO_4 and purified by flash chromatography (DCM:MeOH 9:1).

α-[2-[2-(Prop-2-yn-1-yloxy)ethoxy]ethoxy] thymidine (dU^{peg})



The title compound **dU^{peg}** was obtained from the precursor **12a** (500mg, 1.067 mmol) following the general procedure B. The product was purified by flash chromatography (DCM:MeOH 9:1) and obtained as a white solid (365 mg, 89%). ^{1}H NMR (500.0 MHz, CD_3OD): 2.24 (ddd, 1H, J_{gem} = 13.5, $J_{2'b,1'}$ = 7.2, $J_{2'b,3'}$ = 6.1, H-2'b); 2.29 (ddd, 1H, J_{gem} = 13.5, $J_{2'a,1'}$ = 6.3, $J_{2'a,3'}$ = 3.7, H-2'a); 2.85 (t, 1H, $J_{11'',9''}$ = 2.4, H-11'', partially exchanged with deuterium); 3.64 – 3.70 (m, 8H, H-3'',4'',6'',7''); 3.74 (dd, 1H, J_{gem} = 12.0, $J_{5'b,4'}$ = 3.7, H-5'b); 3.80 (dd, 1H, J_{gem} = 12.0, $J_{5'a,4'}$ = 3.3, H-5'a); 4.93 (td, 1H, $J_{4',5'}$ = 3.7, 3.3, $J_{4',3'}$ = 3.7, H-4'); 4.19 (s, 2H, H-9''); 4.25, 4.28 (2 × dd, 2 × 1H, J_{gem} = 12.9, $J_{1'',6}$ = 1.0, H-1''); 5.41 (dt, 1H, $J_{3',2'}$ = 6.1, 3.7, $J_{3',4'}$ = 3.7, H-3'); 6.28 (dd, 1H, $J_{1',2'}$ = 7.2, 6.3, H-1'); 8.05 (t, 1H, $J_{6,1''}$ = 1.0, H-6). ^{13}C NMR (125.7 MHz, CD_3OD): 41.35 (CH_2-2'); 58.98 (CH_2-9''); 62.81 (CH_2-5'); 66.56 (CH_2-1''); 70.00, 70.77, 71.26, 71.43 (CH_2-3'',4'',6'',7''); 72.15 (CH-3'); 76.01 (CH-11''); 80.56 (C-10''); 86.61 (CH-1'); 88.85 (CH-4'); 112.22 (C-5); 140.83 (CH-6); 152.67 (C-2); 165.84 (C-4). HRMS (ESI+): calculated for $C_{17}H_{24}O_8N_2Na$ = 407.14249; found: 407.14255.

α-(Undec-10-yn-1-yloxy) thymidine (dU^{un})

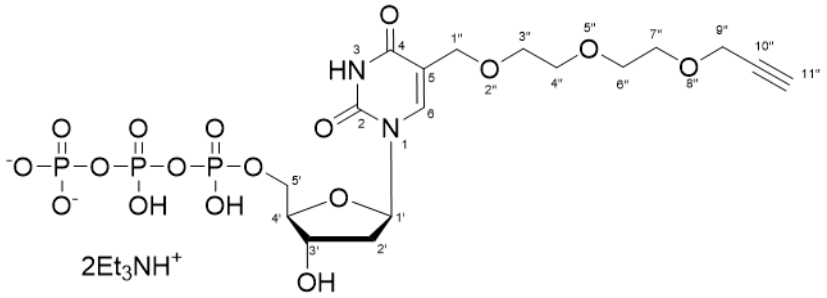


The title compound **dU^{un}** was obtained from the precursor **12b** (300 mg, 0.61 mmol) following the general procedure B. The product was purified by flash chromatography (DCM:MeOH 9:1) and obtained as a white solid (226 mg, 92%). ^{1}H NMR (500.0 MHz, CD_3OD): 1.22 – 1.45 (m, 10H, H-5″,6″,7″,8″,9″); 1.46 – 1.53 (m, 2H, H-10″); 1.54 – 1.62 (m, 2H, H-4″); 2.10 – 2.18 (m, 3H, H-11″,13″); 2.22 (ddd, 1H, J_{gem} = 13.5, $J_{2'b,1'}$ = 7.2, $J_{2'b,3'}$ = 6.2, H-2'b); 2.30 (ddd, 1H, J_{gem} = 13.5, $J_{2'a,1'}$ = 6.2, $J_{2'a,3'}$ = 3.5, H-2'a); 3.49 (t, 2H, $J_{3'',4''}$ = 6.6, H-3″); 3.73 (dd, 1H, J_{gem} = 12.0, $J_{5'b,4'}$ = 3.7, H-5'b); 3.79 (dd, 1H, J_{gem} = 12.0, $J_{5'a,4'}$ = 3.4, H-5'a); 3.93 (ddd, 1H, $J_{4',5'}$ = 3.7, 3.4, $J_{4',3'}$ = 3.5, H-4'); 4.19 (dd, 1H, J_{gem} = 13.0, $J_{1''b,6}$ = 0.8, H-1″b); 4.22 (dd, 1H, J_{gem} = 13.0, $J_{1''a,6}$ = 0.8, H-1″a); 5.22 (dt, 1H, $J_{3',2'}$ = 6.2, 3.5, $J_{3',4'}$ = 3.5, H-3'); 6.29 (dd, 1H, $J_{1',2'}$ = 7.2, 6.2, H-1'); 8.01 (t, 1H, $J_{6,1''}$ = 0.8, H-6). ^{13}C NMR (125.7 MHz, CD_3OD): 19.00 (CH_2-11″); 27.17, 29.69, 29.75, 30.15, 30.49, 30.58, 30.65 (CH_2-4″,5″,6″,7″,8″,9″,10″); 41.38 (CH_2-2'); 62.83 (CH_2-5'); 66.21 (CH_2-1″); 69.39 (CH-13″); 71.68 (CH_2-3″); 72.18 (CH-3'); 85.08 (C-12″); 86.58 (CH-1'); 88.87 (CH-4'); 112.43 (C-5); 140.76 (CH-6); 152.80 (C-2); 166.04 (C-4). HRMS (ESI$^+$): calculated for $C_{21}H_{32}O_6N_2Na$ = 431.21526; found: 431.21542.

General procedure C: 5'-*O*-triphosphorylation

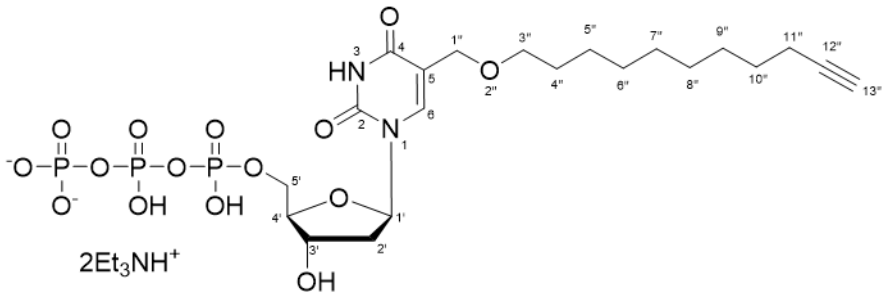
Nucleosides **dU^{peg}** or **dU^{un}** were dried under reduced pressure overnight, dissolved in dry $PO(OMe)_3$ (750 µL) and cooled at 0°C. Freshly distilled $POCl_3$ (1.2 equiv.) was added dropwise and the reaction mixture was stirred at 0°C overnight. Then, an ice-cold solution of (*n*-$Bu_3NH)_2H_2P_2O_7$ (5 equiv.) and *n*-Bu_3N (4 mmol) in anhydrous DMF (3 mL) was added dropwise to the reaction mixture at -10°C. The reaction mixture was stirred for 1.5 hour at -10°C and then quenched by dropwise addition of cold 2M TEAB (1 mL). The mixture was concentrated on a rotary evaporator and the residue was co-evaporated several times with milli-Q water. The crude product was dissolved in water (ca 3 mL), filtered and purified by semi-preparative HPLC using a linear gradient of methanol (5→100%) in 0.1 M TEAB buffer. The appropriate fractions were combined and evaporated on a rotary evaporator. The viscous oil was co-evaporated three times with milli-Q water and the product was finally freeze-dried.

α-[2-[2-(Prop-2-yn-1-yloxy)ethoxy]ethoxy]thymidine-5'-*O*-triphosphate (dU^{peg}TP)



dU^{peg}TP was obtained from the starting nucleoside **dU^{peg}** (55 mg, 0.143 mmol) following the general procedure C. Reverse phase HPLC, using a linear gradient of methanol (5→100%) in water over 3 hours, afforded the desired compound as a white solid (29 mg, 22%). ^{1}H NMR (500.0 MHz, D_2O): 1.27 (t, 18H, J_{vic} = 7.3, C**H$_3$**CH$_2$N); 2.35 – 2.44 (m, 2H, H-2'); 2.86 (t, 1H, $J_{11'',9''}$ = 2.4, H-11''); 3.20 (q, 12H, J_{vic} = 7.3, CH$_3$C**H$_2$**N); 3.67 – 3.76 (m, 8H, H-3'',4'',6'',7''); 4.17 – 4.25 (m, 5H, H-4',5',9''); 4.34, 4.37 (2 × d, 2 × 1H, J_{gem} = 11.7, H-1''); 4.66 (m, 1H, H-3'); 6.33 (t, 1H, $J_{1',2'}$ = 6.9, H-1'); 8.03 (s, 1H, H-6). ^{13}C NMR (125.7 MHz, D_2O): 10.96 (**C**H$_3$CH$_2$N); 41.56 (CH$_2$-2'); 49.37 (CH$_3$**C**H$_2$N); 60.61 (CH$_2$-9''); 67.99 (CH$_2$-1''); 68.22 (d, $J_{C,P}$ = 5.6, CH$_2$-5'); 71.36, 71.41, 72.13, 72.29 (CH$_2$-3'',4'',6'',7''); 73.58 (CH-3'); 78.64 (CH-11''); 81.95 (C-10''); 88.15 (CH-1'); 88.34 (d, $J_{C,P}$ = 9.1, CH-4'); 113.49 (C-5); 144.37 (CH-6); 154.34 (C-2); 167.97 (C-4). ^{31}P NMR (202.4 MHz, D_2O): -22.88 (dd, J = 20.2, 19.2, P$_\beta$); -11.13 (d, J = 20.2, P$_\alpha$); -10.43 (d, J = 19.2, P$_\gamma$). HRMS (MALDI-): calculated for $C_{17}H_{26}N_2O_{17}P_3$ = 623.0439; found: 623.0420.

α-(Undec-10-yn-1-yloxy)thymidine-5'-*O*-triphosphate (dU^{un}TP)



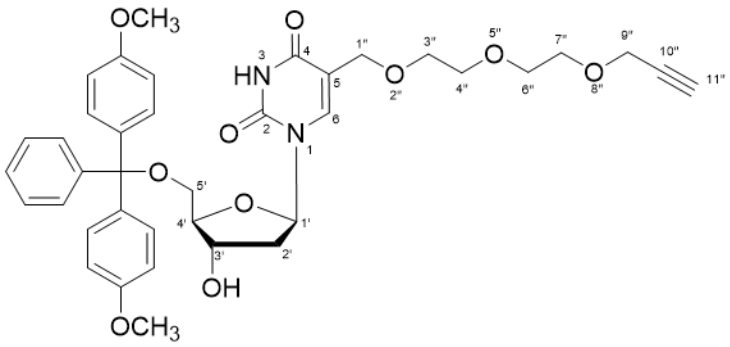
dU^{un}TP was obtained from the starting nucleoside **dU^{un}** (100 mg, 0.244 mmol) following the general procedure C. Reverse phase HPLC, using a linear gradient of methanol (5→100%) in water over 3 hours, afforded the desired compound as a white solid (124 mg, 48%). ^{1}H NMR (500.0

MHz, D_2O, ref(dioxane) = 3.75 ppm): 1.24 – 1.31 (m, 26H, H-5″,6″,7″,8″, C**H₃**CH_2N); 1.33 – 1.43 (m, 2H, H-9″); 1.46 – 1.52 (m, 2H, H-10″); 1.52 – 1.59 (m, 2H, H-4″); 2.19 (td, 2H, $J_{11'',10''}$ = 7.1, $J_{11'',13''}$ = 2.7, H-11″); 2.32 (t, 1H, $J_{13'',11''}$ = 2.7, H-13″); 2.35 – 2.44 (m, 2H, H-2′); 3.20 (q, 12H, J_{vic} = 7.3, CH_3C**H₂**N); 3.54 (t, 2H, $J_{3'',4''}$ = 6.7, H-3″); 4.17 – 4.25 (m, 3H, H-4′,5′); 4.28, 4.33 (2 × d, 2 × 1H, J_{gem} = 11.8, H-1″); 4.66 (m, 1H, H-3′); 6.33 (t, 1H, $J_{1',2'}$ = 6.9, H-1′); 8.00 (s, 1H, H-6). ^{13}C NMR (125.7 MHz, D_2O, ref(dioxane) = 69.3 ppm): 10.96 (**C**H_3CH_2N); 20.23 (CH_2-11″); 26.56, 27.83, 30.43, 30.63, 30.85, 30.98, 31.27 (CH_2-4″,5″,6″,7″,8″,9″,10″); 41.57 (CH_2-2′); 49.36 (CH_3**C**H_2N); 67.52 (CH_2-1″); 68.25 (d, $J_{C,P}$ = 5.5, CH_2-5′); 71.66 (CH-13″); 72.70 (CH_2-3″); 73.62 (CH-3′); 86.10 (CH-1′); 88.34 (d, $J_{C,P}$ = 9.2, CH-4′); 89.41 (C-12″); 113.84 (C-5); 140.10 (CH-6); 154.32 (C-2); 167.85 (C-4). ^{31}P NMR (202.4 MHz, D_2O): -22.93 (bm, P_β); -11.17 (d, J = 20.2, P_α); -10.45 (bd, J = 15.9, P_γ). HRMS (MALDI-): calculated for $C_{21}H_{34}N_2O_{15}P_3$ = 647.1172; found: 647.1181.

General procedure D: 5′-*O*-dimethoxytrityl protection of nucleosides

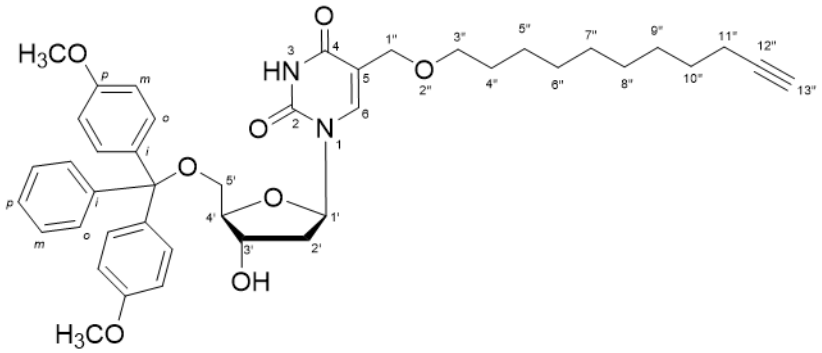
The suitable nucleosides were dried by repeated co-evaporations with anhydrous pyridine and dissolved in 20 mL of anhydrous pyridine. 4,4′-Dimethoxytrityl chloride (1.2 equiv.) was added in 4 portions over one hour and the reaction mixture was stirred at r. t. overnight. Pyridine was removed under reduced pressure at a rotary evaporator and the crude reaction mixture was re-dissolved in CH_2Cl_2 (200 mL), washed with 10% aqueous $NaHCO_3$ (100 mL), H_2O (100 mL) and brine (100 mL), and dried over Na_2SO_4. Purification by flash chromatography (CH_2Cl_2: MeOH, with 0.5% Et_3N) afforded the pure desired compounds.

5'-*O*-[Bis(4-methoxyphenyl)phenylmethyl]-α-[2-[2-(prop-2-yn-1-yloxy)ethoxy]ethoxy]-thymidine (14a)



The title product **14a** was obtained from the nucleoside **dU^{peg}** (300mg, 0.78 mmol) and 4,4'-dimethoxytrityl chloride (1.2 equiv., 0.94 mmol, 319 mg) following the general procedure D. The product was purified by flash chromatography (CH_2Cl_2: MeOH 95:5, Et_3N 0.5%) and obtained as a yellowish foam (418 mg, 78%). ^{1}H NMR (500.0 MHz, $CDCl_3$): 2.26 (ddd, 1H, J_{gem} = 13.5, $J_{2'b,1'}$ = 7.3, $J_{2'b,3'}$ = 6.5, H-2'b); 2.41 (t, 1H, $J_{11'',9''}$ = 2.4, H-11''); 2.42 (ddd, 1H, J_{gem} = 13.5, $J_{2'a,1'}$ = 6.1, $J_{2'a,3'}$ = 3.4, H-2'a); 3.34 (dd, 1H, J_{gem} = 10.4, $J_{5'b,4'}$ = 3.7, H-5'b); 3.38 – 3.45, 3.52 – 3.55, 3.60 – 3.63 (3 × m, 9H, H-5'a,3'',4'',6'',7''); 3.77 (s, 6H, CH_3O-DMTr); 3.82, 3.97 (2 × dd, 2 × 1H, J_{gem} = 12.0, $J_{1'',6}$ = 0.7, H-1''); 4.02 (td, 1H, $J_{4',5'}$ = 3.7, $J_{4',3'}$ = 3.4, H-4'); 4.14 (d, 2H, $J_{9'',11''}$ = 2.4, H-9''); 4.50 (dt, 1H, $J_{3',2'}$ = 6.5, 3.4, $J_{3',4'}$ = 3.4, H-3'); 6.34 (dd, 1H, $J_{1',2'}$ = 7.3, 6.1, H-1'); 6.80 – 6.84 (m, 4H, H-*m*-C_6H_4-DMTr); 7.21 (m, 1H, H-*p*-C_6H_5-DMTr); 7.25 – 7.31 (m, 6H, H-*o*-C_6H_4-DMTr, H-*m*-C_6H_5-DMTr); 7.38 – 7.42 (m, 2H, H-*o*-C_6H_5-DMTr); 7.71 (t, 1H, $J_{6,1''}$ = 0.7, H-6). ^{13}C NMR (125.7 MHz, $CDCl_3$): 40.67 (CH_2-2'); 55.17, 55.18 (CH_3O-DMTr); 58.24 (CH_2-9''); 63.53 (CH_2-5'); 64.98 (CH_2-1''); 68.94, 69.79, 70.12, 70.14 (CH_2-3'',4'',6'',7''); 72.01 (CH-3'); 74.56 (CH-11''); 79.59 (C-10''); 84.90 (CH-1'); 85.82 (CH-4'); 86.64 (C-DMTr); 111.90 (C-5); 113.18 (CH-*m*-C_6H_4-DMTr); 126.96 (CH-*p*-C_6H_5-DMTr); 127.91 (CH-*m*-C_6H_5-DMTr); 128.02 (CH-*o*-C_6H_5-DMTr); 130.00, 130.03 (CH-*o*-C_6H_4-DMTr); 135.44, 135.45 (C-*i*-C_6H_4-DMTr); 138.53 (CH-6); 144.45 (C-*i*-C_6H_5-DMTr); 150.28 (C-2); 158.52, 158.54 (C-*p*-C_6H_4-DMTr); 162.62 (C-4). HRMS (ESI+): calculated for $C_{38}H_{42}O_{10}N_2Na$ = 709.27317; found: 709.27350.

5'-*O*-[Bis(4-methoxyphenyl)phenylmethyl]-α-(undec-10-yn-1-yloxy)thymidine (14b)



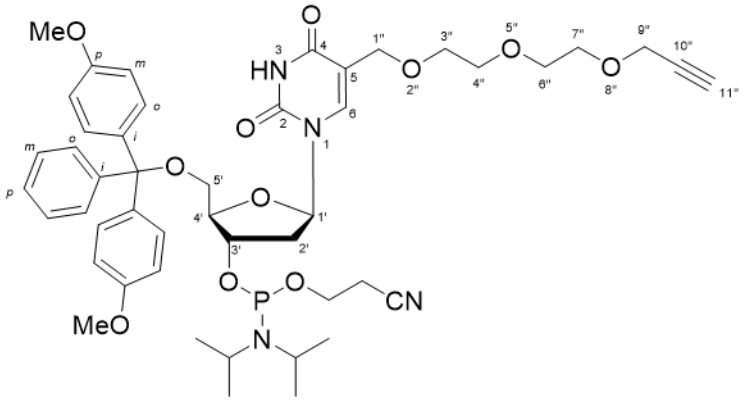
The title product **14b** was obtained from the nucleoside **dU^{un}** (420 mg, 1.023 mmol) and 4,4'-dimethoxytrityl chloride (1.2 equiv., 1.227 mmol, 416 mg) following the general procedure D. The product was purified by flash chromatography (CH_2Cl_2: MeOH 97:3, Et_3N 0.5%) and obtained as a yellowish foam (604 mg, 83%). ^{1}H NMR (500.0 MHz, CD_3CN): 1.14 – 1.29 (m, 8H, H-5'',6'',7'',8'');

1.31 – 1.38 (m, 4H, H-4″,9″); 1.43 – 1.51 (m, 2H, H-10″); 2.11 – 2.18 (m, 3H, H-11″,13″); 2.20 – 2.29 (m, 2H, H-2′); 3.11 – 3.19 (m, 2H, H-3″); 3.24 (dd, 1H, J_{gem} = 10.6, $J_{5'b,4'}$ = 3.2, H-5′b); 3.29 (dd, 1H, J_{gem} = 10.6, $J_{5'a,4'}$ = 4.2, H-5′a); 3.37 (d, 1H, $J_{OH,3'}$ = 4.2, OH-3′); 3.61 (dd, 1H, J_{gem} = 11.7, $J_{1''b,6}$ = 0.8, H-1″b); 3.76 (s, 6H, CH_3O-DMTr); 3.88 (dd, 1H, J_{gem} = 11.7, $J_{1''a,6}$ = 0.8, H-1″a); 3.92 (td, 1H, $J_{4',3'}$ = 4.2, $J_{4',5'}$ = 4.2, 3.2, H-4′); 4.43 (dq, 1H, $J_{3',2'}$ = 5.8, 4.2, $J_{3',4'}$ = $J_{3',OH}$ = 4.2, H-3′); 6.20 (t, 1H, $J_{1',2'}$ = 6.7, H-1′); 6.84 – 6.89 (m, 4H, H-*m*-C_6H_4-DMTr); 7.24 (m, 1H, H-*p*-C_6H_5-DMTr); 7.28 – 7.34 (m, 6H, H-*m*-C_6H_5-DMTr, H-*o*-C_6H_4-DMTr); 7.42 – 7.46 (m, 2H, H-*o*-C_6H_5-DMTr); 7.60 (t, 1H, $J_{6,1''}$ = 0.8, H-6), 9.04 (s, 1H, NH). ^{13}C NMR (125.7 MHz, CD_3CN): 18.74 (CH_2-11″); 26.77, 29.28, 29.43, 29.75, 30.08, 30.14, 30.29 (CH_2-4″,5″,6″,7″,8″,9″,10″); 41.01 (CH_2-2′); 55.92 (CH_3O-DMTr); 64.48 (CH_2-5′); 65.45 (CH_2-1″); 69.57 (CH_2-3″); 71.11 (CH-13″); 71.94 (CH-3′); 86.50 (CH-1′); 85.58 (C-12″); 86.82 (CH-4′); 87.36 (C-DMTr); 112.43 (C-5); 114.11 (CH-*m*-C_6H_4-DMTr); 127.91 (CH-*p*-C_6H_5-DMTr); 128.93 (CH-*m*-C_6H_5-DMTr); 129.04 (CH-*o*-C_6H_4-DMTr); 131.03, 131.06 (CH-*o*-C_6H_4-DMTr); 136.66, 136.89 (C-*i*-C_6H_4-DMTr); 139.49 (CH-6); 145.93 (C-*i*-C_6H_5-DMTr); 151.20 (C-2); 159.71 (C-*p*-C_6H_4-DMTr); 163.50 (C-4). HRMS (ESI+): calculated for $C_{42}H_{50}O_8N_2Na$ = 733.34594; found: 733.34611.

General procedure E: synthesis of 3′-*O*-phosphoramidites

The suitable 5′-*O*-DMT-substituted nucleosides were dried by repeated co-evaporations with anhydrous pyridine and subsequent co-evaporations with anhydrous CH_2Cl_2, and finally dissolved in 3 mL of anhydrous CH_2Cl_2. Freshly distilled *N*,*N*-diisopropylethylamine (DIPEA, 2.5 equiv.) was added, followed by the addition of 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite (1.2 equiv.), and the reaction mixture was stirred at r. t. until a complete conversion was observed by TLC analysis (approx. 1.5 h). The reaction mixture was diluted with anhydrous CH_2Cl_2 (50 mL), quickly washed under an argon atmosphere with saturated aqueous KCl (20 mL) and dried over Na_2SO_4. Purification by flash chromatography (cyclohexane : EtOAc, with 5% Et_3N) under an atmosphere of argon provided the pure desired products.

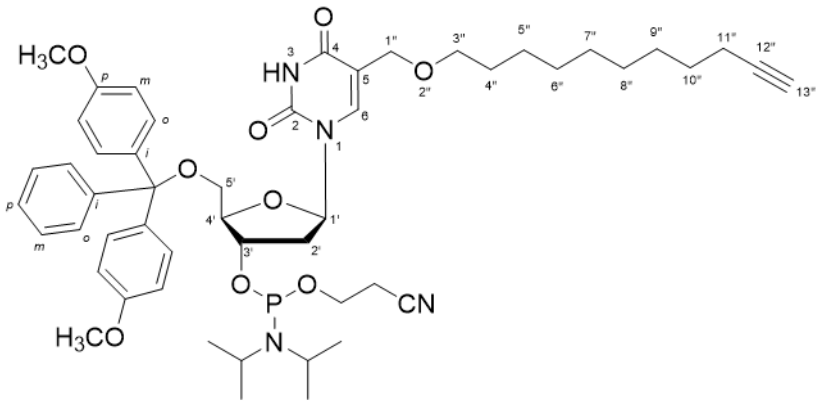
5'-*O*-[Bis(4-methoxyphenyl)phenylmethyl]-α-[2-[2-(prop-2-yn-1-yloxy)ethoxy]ethoxy]-thymidine 3'-[2-cyanoethyl bis(1-methylethyl)phosphoramidite] (CEP-dU^{peg}, 15a)



The title product **15a** was obtained from the precursor **14a** (165 mg, 0.240 mmol), *N*,*N*-diisopropylethylamine (2.5 equiv., 0.601 mmol, 104 µL) and 2-cyanoethyl-N,N-diisopropylchlorophosphoramidite (1.2 equiv, 0.288 mmol, 64 µL) following the general procedure E. The product was purified by flash chromatography (Cyclohexane:EtOAc 1:5, Et_3N 0.5%) and obtained as a yellowish foam (108 mg, 51%). ^{1}H NMR (500.0 MHz, CD_3CN): 1.15, 1.17 (2 × d, 2 × 6H, J_{vic} = 6.8, (C**H$_3$**)$_2$CH); 2.35 (ddd, 1H, J_{gem} = 13.8, $J_{2'b,1'}$ = 7.0, $J_{2'b,3'}$ = 6.4, H-2'b); 2.39 (ddd, 1H, J_{gem} = 13.8, $J_{2'a,1'}$ = 6.4, $J_{2'a,3'}$ = 4.1, H-2'a); 2.50 – 2.53 (m, 2H, OCH$_2$C**H$_2$**CN); 2.68 (t, 1H, $J_{11'',9''}$ = 2.4, H-11''); 3.28 – 3.40 (m, 6H, H-5',3'',4''); 3.44 – 3.49 (m, 2H, H-6''); 3.52 – 3.56 (m, 2H, H-7''); 3.56 – 3.59 (m, 4H, (CH$_3$)$_2$C**H**, OC**H$_2$**CH$_2$CN); 3.70 (dd, 1H, J_{gem} = 11.7, $J_{1''b,6}$ = 0.7, H-1''b); 3.76 (s, 6H, CH$_3$O-DMTr); 3.95 (dd, 1H, J_{gem} = 11.7, $J_{1''a,6}$ = 0.7, H-1''a); 4.10 (bm, 1H, H-4'); 4.11 (d, 2H, $J_{9'',11''}$ = 2.4, H-9''); 4.61 (ddt, 1H, $J_{H,P}$ = 10.4, $J_{3',2'}$ = 6.4, 4.1, $J_{3',4'}$ = 4.1, H-3'); 6.20 (d, 1H, $J_{1',2'}$ = 7.0, 6.4, H-1'); 6.85 – 6.90 (m, 4H, H-*m*-C$_6$H$_4$-DMTr); 7.24 (m, 1H, H-*p*-C$_6$H$_5$-DMTr); 7.29 – 7.36 (m, 6H, H-*m*-C$_6$H$_4$-DMTr, H-*m*-C$_6$H$_5$-DMTr); 7.43 – 7.47 (m, 2H, H-*o*-C$_6$H$_5$-DMTr); 7.65 (s, 1H, H-6); 9.27 (s, 1H, NH-3). ^{13}C NMR (125.7 MHz, CD_3CN): 20.94 (d, $J_{C,P}$ = 7.2, OCH$_2$**C**H$_2$CN); 24.86, 24.88 (dd, $J_{C,P}$ = 7.3, (**C**H$_3$)$_2$CH); 40.04 (d, $J_{C,P}$ = 4.3, CH$_2$-2'); 43.98 (d, $J_{C,P}$ = 12.3, (CH$_3$)$_2$**C**H); 55.91, 55.94 (CH$_3$O-DMTr); 58.65 (CH$_2$-9''); 59.45 (d, $J_{C,P}$ = 19.1, O**C**H$_2$CH$_2$CN); 63.97 (CH$_2$-5'); 65.82 (CH$_2$-1''); 69.85 (CH$_2$-7''); 70.31 (CH$_2$-3''); 70.84 (CH$_2$-4''); 70.91 (CH$_2$-6''); 73.76 (d, $J_{C,P}$ = 16.8, CH-3'); 75.65 (CH-11''); 80.91 (C-10''); 85.75 (CH-1'); 86.06 (CH-4'); 87.41 (C-DMTr); 112.21 (C-5); 114.12 (CH-*m*-C$_6$H$_4$-DMTr); 119.36 (CN); 127.96 (CH-*p*-C$_6$H$_5$-DMTr); 128.93 (CH-*m*-C$_6$H$_5$-DMTr); 129.03 (CH-*o*-C$_6$H$_5$-DMTr); 131.07 (CH-*o*-C$_6$H$_4$-DMTr); 136.60, 136.67 (C-*i*-C$_6$H$_4$-DMTr); 139.86 (CH-6); 145.84

(C-*i*-C_6H_5-DMTr); 151.19 (C-2); 159.72 (C-*p*-C_6H_4-DMTr); 163.63 (C-4). $^{31}P\{^1H\}$ NMR (202.4 MHz, CD_3CN): 149.30. HRMS (ESI+): calculated for $C_{47}H_{59}O_{11}N_4NaP$ = 909.38102; found: 909.38133.

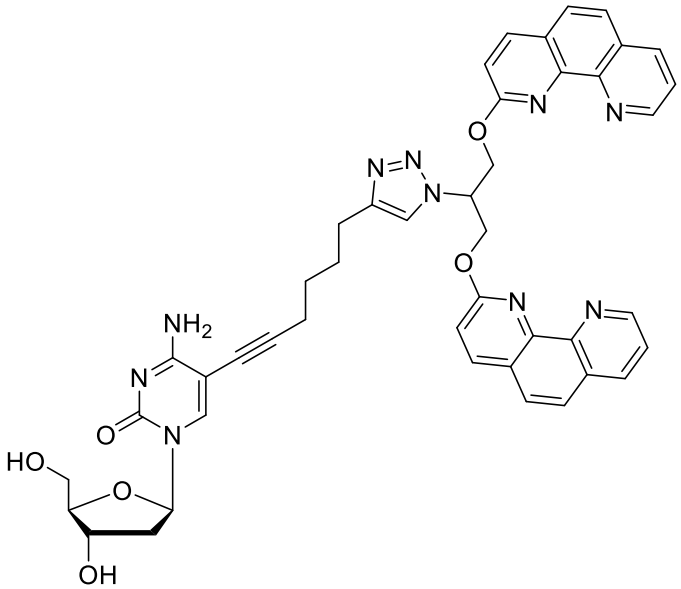
5'-*O*-[Bis(4-methoxyphenyl)phenylmethyl]-α-(undec-10-yn-1-yloxy)-thymidine 3'-[2-cyanoethyl bis(1-methylethyl)phosphoramidite] (CEP-dU^{un}, 15b)



The title product **15b** was obtained from the precursor **14b** (210 mg, 0.295 mmol), *N*,*N*-diisopropylethylamine (2.5 equiv., 0.739 mmol, 129 µL) and 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite (1.2 equiv, 0.354 mmol, 79 µL) following the general procedure E. The product was purified by flash chromatography (Cyclohexane:EtOAc 4:1, Et_3N 0.5%) and obtained as a yellowish foam (177 mg, 66%). ^1H NMR (500.0 MHz, CD_3CN): 1.13 – 1.29 (m, 20H, H-5″,6″,7″,8″, (C**H$_3$**)$_2$CH); 1.29 – 1.38 (m, 4H, H-4″,9″); 1.43 – 1.51 (m, 2H, H-10″); 2.11 – 2.17 (m, 3H, H-11″,13″); 2.33 (dt, 1H, J_{gem} = 13.7, $J_{2'b,1'}$ = $J_{2'b,3'}$ = 6.7, H-2'b); 2.39 (ddd, 1H, J_{gem} = 13.7, $J_{2'a,1'}$ = 6.3, $J_{2'a,3'}$ = 3.9, H-2'a); 2.52 (t, 2H, J_{vic} = 6.0, OCH$_2$C**H$_2$**CN); 3.10 – 3.19 (m, 2H, H-3″); 3.31 (dd, 1H, J_{gem} = 10.7, $J_{5'b,4'}$ = 4.0, H-5'b); 3.35 (dd, 1H, J_{gem} = 10.7, $J_{5'a,4'}$ = 3.1, H-5'a); 3.54 – 3.70 (m, 5H, H-1″b, OC**H$_2$**CH$_2$CN, (CH$_3$)$_2$C**H**); 3.76 (s, 6H, CH$_3$O-DMTr); 3.91 (dd, 1H, J_{gem} = 11.8, $J_{1''a,6}$ = 0.8, H-1″a); 4.10 (m, 1H, H-4'); 4.61 (ddt, 1H, $J_{H,P}$ = 10.5, $J_{3',2'}$ = 6.7, 3.9, $J_{3',4'}$ = 3.9, H-3'); 6.21 (dd, 1H, $J_{1',2'}$ = 6.7, 6.3, H-1'); 6.84 – 6.89 (m, 4H, H-*m*-C_6H_4-DMTr); 7.24 (m, 1H, H-*p*-C_6H_5-DMTr); 7.29 – 7.35 (m, 6H, H-*m*-C_6H_5-DMTr, H-*o*-C_6H_4-DMTr); 7.43 – 7.47 (m, 2H, H-*o*-C_6H_5-DMTr); 7.63 (t, 1H, $J_{6,1''}$ = 0.8, H-6), 9.21 (s, 1H, NH). ^{13}C NMR (125.7 MHz, CD_3CN): 18.75 (CH_2-11″); 20.94 (d, $J_{C,P}$ = 7.2, OCH$_2$**C**H$_2$CN); 24.87, 24.89 (2 × d, $J_{C,P}$ = 7.3, (**C**H$_3$)$_2$CH); 26.75, 29.28, 29.43, 29.75, 30.07, 30.13, 30.27 (CH_2-4″,5″,6″,7″,8″,9″,10″); 40.09 (d, $J_{C,P}$ = 4.3, CH_2-2'); 43.99 (d, $J_{C,P}$ = 12.3, (CH$_3$)$_2$**C**H); 55.92 (CH$_3$O-DMTr); 59.47 (d, $J_{C,P}$ = 19.1, O**C**H$_2$CH$_2$CN); 63.92 (CH_2-5'); 65.44 (CH_2-1″);

69.58 (CH-13''); 71.11 (CH_2-3''); 73.70 (d, $J_{C,P}$ = 16.8, CH-3'); 85.58 (CH-1', C-12''); 86.01 (d, $J_{C,P}$ = 4.2, CH-4'); 87.43 (C-DMTr); 112.56 (C-5); 114.11 (CH-*m*-C_6H_4-DMTr); 119.35 (CN); 127.94 (CH-*p*-C_6H_5-DMTr); 128.93 (CH-*m*-C_6H_5-DMTr); 129.06 (CH-*o*-C_6H_4-DMTr); 131.09 (CH-*o*-C_6H_4-DMTr); 136.57, 136.70 (C-*i*-C_6H_4-DMTr); 139.39 (CH-6); 145.84 (C-*i*-C_6H_5-DMTr); 151.20 (C-2); 159.72 (C-*p*-C_6H_4-DMTr); 163.58 (C-4). $^{31}P\{^1H\}$ NMR (202.4 MHz, CD_3CN): 149.37. HRMS (ESI+): calculated for $C_{51}H_{67}O_9N_4NaP$ = 933.45379; found: 933.45487.

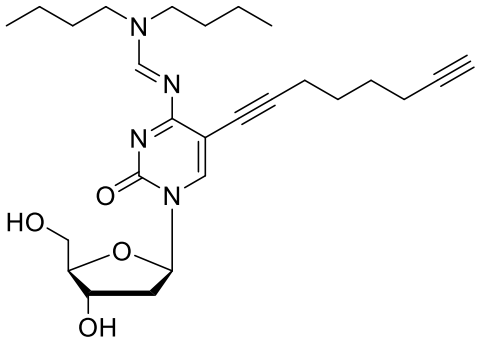
Synthesis of dC^{ocp} through CuAAC reaction between dC^{o} and N_3CP



dC^{o} (50 mg, 0.151 mmol) and **N_3CP** (86 mg, 0.181 mmol) were suspended in a THF:H_2O:^{t}BuOH 3:1:1 mixture (5 mL). 0.3 mL of a freshly prepared solution of sodium ascorbate (1 M) and $CuSO_4$ pentahydrate (7.5%) in H_2O were added by syringe and the reaction mixture was stirred at room temperature overnight. After evaporation of the solvents, the crude reaction product was purified by reverse-fase HPLC (MeOH 7% to 100% over 40 minutes in H_2O) to afford the desired product **dC^{ocp}** as a pale yellow solid (77mg, 63%). ^{1}H NMR (400 MHz, Methanol-d_4) δ 9.01 (dd, *J* = 4.4, 1.7 Hz, 2H), 8.40 (dd, *J* = 8.2, 1.7 Hz, 2H), 8.23 (d, *J* = 1.5 Hz, 2H), 8.13 (d, *J* = 8.7 Hz, 2H), 7.77 – 7.66 (m, 6H), 7.13 (d, *J* = 8.7 Hz, 2H), 6.16 (t, *J* = 6.3 Hz, 1H), 5.72 (tt, *J* = 6.6, 4.9 Hz, 1H), 5.58 – 5.40 (m, 4H), 4.33 (dt, *J* = 6.3, 4.0 Hz, 1H), 3.88 (q, *J* = 3.4 Hz, 1H), 3.83 – 3.66 (m, 2H), 2.58 (t, *J* = 7.4 Hz, 2H), 2.33 (m, 1H), 2.30 (dt, *J* = 12.1, 7.1 Hz, 2H), 2.07 (dt, *J* = 13.2, 6.4 Hz, 1H), 1.61 (p, *J* =

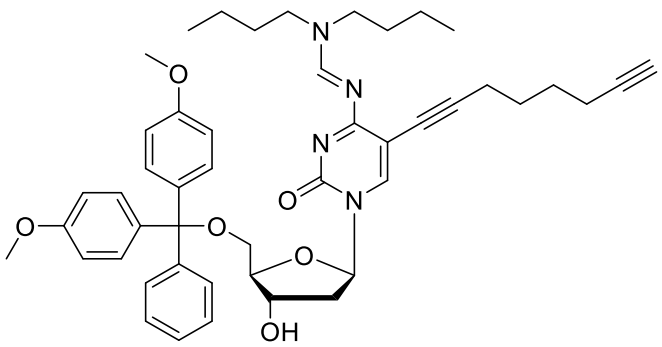
7.5 Hz, 2H), 1.41 (dq, *J* = 9.5, 6.9 Hz, 2H). HRMS calc. for $C_{44}H_{41}O_6N_{10}$: 805.32051, found: 805.32080.

Synthesis of the precursor 7 by *N*^{4}-protection of dC^o



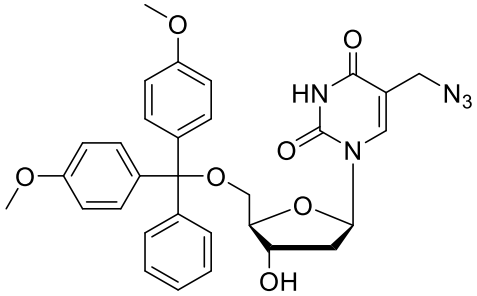
Precursor **dC^o** (1.49 g, 4.50 mmol) was dissolved in 20 mL of DMF, and dinormalbutylformamide dimethyl acetal (4.2 mL, 17.99 mmol, 4.5 equiv.) was added under an aerobic atmosphere. The reaction mixtures was stirred at r.t. overnight. The solvent was evacuated under reduced pressure and the crude product was re-dissolved in CH_2Cl_2 (200 mL), washed with water (2 × 100 mL) and brine (1 × 100 mL), dried over Na_2SO_4 and purified by flash chromatography (DCM : MeOH 9 : 5). The title compound was obtained as a pale yellow solid (1.21 g, 79 %). ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.77 (s, 1H), 7.99 (s, 1H), 6.12 (t, *J* = 6.3 Hz, 1H), 4.58 (t, *J* = 4.9 Hz, 1H), 4.06 (dq, *J* = 4.8, 2.8 Hz, 1H), 3.95 – 3.81 (m, 2H), 3.61 (td, *J* = 7.4, 3.8 Hz, 2H), 3.51 – 3.45 (m, 1H), 3.36 (t, *J* = 7.5 Hz, 2H), 2.51 – 2.37 (m, 4H), 2.24 (dp, *J* = 4.3, 1.6 Hz, 2H), 1.97 (m, *J* = 3.8, 1.8 Hz, 1H), 1.75 – 1.57 (m, 7H), 1.43 – 1.26 (m, 4H), 0.96 (m, *J* = 7.6, 7.1, 3.6 Hz, 6H). HRMS (ESI+): calculated for $C_{41}H_{48}O_6N_4Na$ = 715.34661; found: 715.34678.

Synthesis of the precursor 8 by 5'-*O*-dimethoxytrityl protection of compound 7



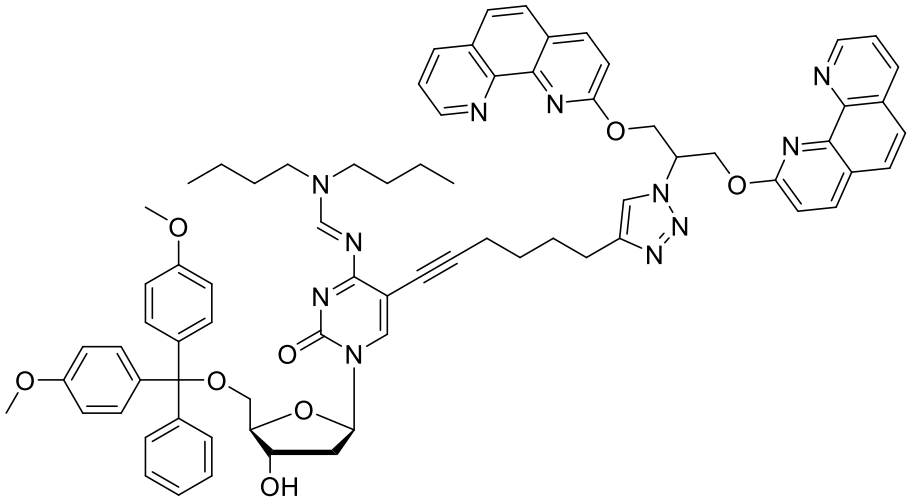
The title product **8** was obtained from the N^4-protected nucleoside **7** (690 mg, 1.47 mmol) and 4,4'-dimethoxytrityl chloride (1.2 equiv., 1.76 mmol, 596 mg) following the general procedure D. The product was purified by flash chromatography (CH_2Cl_2: MeOH 98 : 2, Et_3N 0.5%) and obtained as a white foam (640 mg, 74%).^{1}H NMR (400 MHz, $CDCl_3$) δ 8.85 (s, 1H), 8.17 (s, 1H), 7.51 – 7.44 (m, 2H), 7.41 – 7.35 (m, 4H), 7.33 (t, *J* = 0.8 Hz, 1H), 7.27 – 7.19 (m, 1H), 6.89 – 6.82 (m, 5H), 6.38 (t, *J* = 6.5 Hz, 1H), 4.49 (d, *J* = 5.1 Hz, 1H), 4.17 – 4.10 (m, 1H), 3.81 (s, 6H), 3.62 (dd, *J* = 8.4, 6.6 Hz, 2H), 3.46 – 3.30 (m, 4H), 2.73 – 2.64 (m, 1H), 2.21 (m, *J* = 18.9, 6.7 Hz, 2H), 2.06 (m, *J* = 6.8, 2.6 Hz, 2H), 1.92 (t, *J* = 2.7 Hz, 1H), 1.72 – 1.57 (m, 4H), 1.56 – 1.42 (m, 5H), 1.42 – 1.29 (m, 5H), 1.02 – 0.93 (m, 6H). HRMS (ESI+): calculated for $C_{47}H_{57}O_6N_4$ = 773.42726; found: 773.42748.

Synthesis of the precursor 9 by 5'-*O*-dimethoxytrityl protection of compound dU^{am}



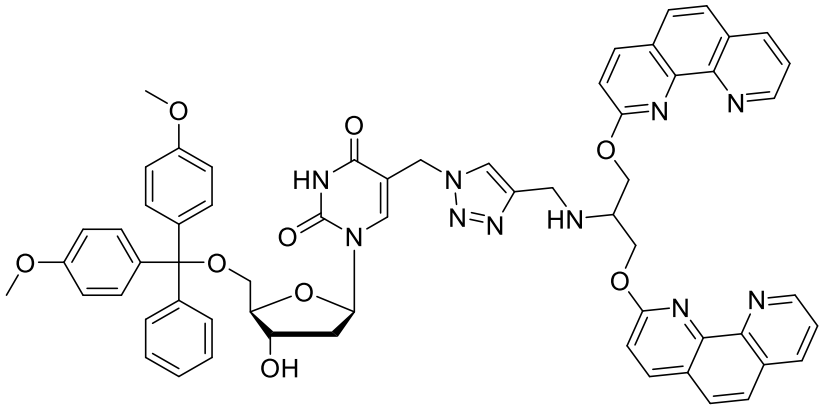
The title product **9** was obtained from the nucleoside **dU^{am}** (420mg, 1.48 mmol) and 4,4'-dimethoxytrityl chloride) (1.2 equiv., 1.78 mmol, 603 mg) following the general procedure D. The product was purified by flash chromatography (CH_2Cl_2: MeOH 95 : 5, Et_3N 0.5%) and obtained as a white solid (942 mg, 83%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.65 (s, 1.7 Hz, 1H), 7.84 (s, 1H), 7.44 – 7.27 (m, 9H), 6.92 – 6.83 (m, 4H), 6.41 (dd, *J* = 7.4, 6.0 Hz, 1H), 4.63 (dt, *J* = 6.2, 3.2 Hz, 1H), 4.09 (q, *J* = 3.1 Hz, 1H), 3.82 (s, 6H), 3.59 (dd, *J* = 13.7, 0.7 Hz, 1H), 3.52 (dd, *J* = 10.6, 3.3 Hz, 1H), 3.45 – 3.36 (m, 2H), 2.48 (ddd, *J* = 13.6, 6.0, 3.3 Hz, 1H), 2.35 (ddd, *J* = 13.6, 7.4, 6.2 Hz, 1H). HRMS (ESI+): calculated for $C_{31}H_{31}O_7N_5Na$= 608.21157; found: 608.21165.

CuAAC reaction between 8 and N_3CP



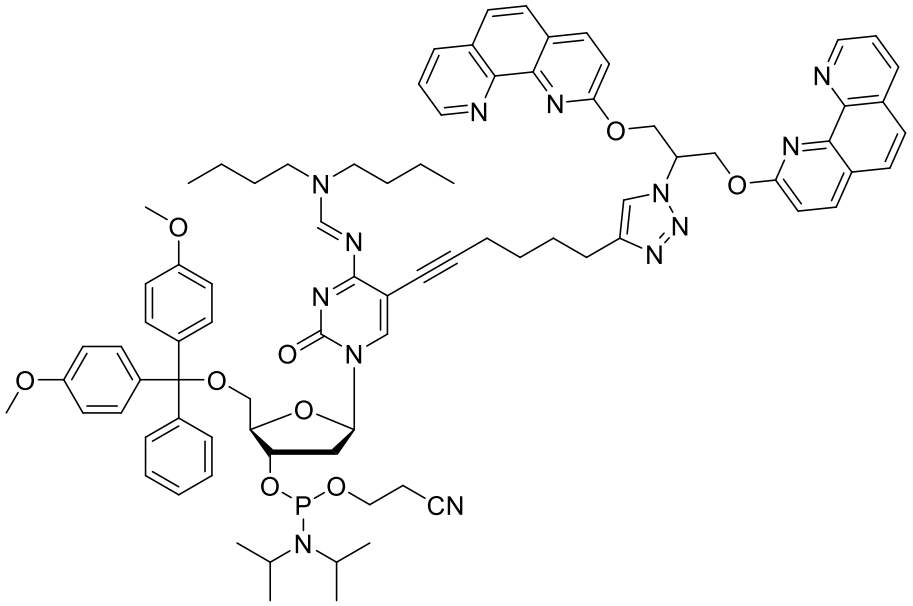
To a solution of the precursor **8** (594 mg, 0.77 mmol), **N_3CP** (400 mg, 0.84 mmol, 1.1 equiv.), $CuSO_4$ pentahydrate (384 mg, 1.54 mmol, 2 equiv.) in a mixture THF/^{t}BuOH/H_2O 3 : 1 : 1, was added Na-*L*-ascorbate (304 mg, 1.54 mmol, 2 equiv.). The reaction mixture was stirred under an argon atmosphere at r.t. overnight. The reaction was quenched by addition of a water solution containing EDTA (5 %) and $NaHCO_3$ (5 %), and the mixture was further stirred for 30 min. The organic layer was separated and the water phase further washed with CH_2Cl_2 (2 × 100 mL). The organic fractions were combined and the solvent evaporated at reduced pressure. Reverse phase high-performance flash chromatography, using a linear gradient of methanol (5 → 100 %) in water over 100 min, afforded the desired compound as a white solid (165 mg, 17 %). ^{1}H NMR (400 MHz, $CDCl_3$) δ 9.14 – 9.06 (m, 2H), 8.76 (s, 1H), 8.21 – 8.12 (m, 3H), 8.03 (d, *J* = 8.6 Hz, 2H), 7.88 (s, 1H), 7.71 – 7.58 (m, 4H), 7.54 (dd, *J* = 8.1, 4.3 Hz, 2H), 7.49 – 7.38 (m, 2H), 7.39 – 7.29 (m, 4H), 7.27 – 7.20 (m, 2H), 7.11 – 7.07 (m, 2H), 6.78 (dd, *J* = 8.9, 1.7 Hz, 4H), 6.38 (t, *J* = 6.4 Hz, 1H), 5.76 – 5.66 (m, 1H), 5.48 – 5.32 (m, 4H), 4.52 (s, 1H), 4.42 – 4.29 (m, 1H), 4.16 (s, 1H), 3.76 (s, 1H), 3.72 (s, 3H), 3.71 (s, 3H), 3.55 (m, 2H), 3.39 – 3.20 (m, 4H), 2.77 – 2.64 (m, 1H), 2.54 (t, *J* = 7.7 Hz, 2H), 2.21 – 2.05 (m, 3H), 1.66 – 1.44 (m, 6H), 1.39 – 1.17 (m, 6H), 0.99 – 0.80 (m, 6H). HRMS (ESI+): calculated for $C_{74}H_{75}O_8N_{11}Na$ = 1268.56923; found: 1268.56973.

CuAAC reaction between 9 and PgCP



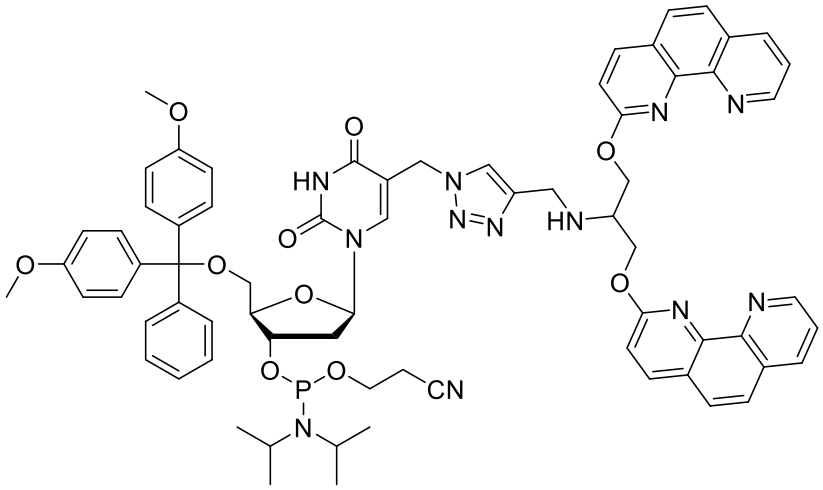
To a solution of the precursor **9** (290 mg, 0.49 mmol), **PgCP** (288 mg, 0.59 mmol, 1.2 equiv.), $CuSO_4$ pentahydrate (246 mg, 0.99 mmol, 2 equiv.) in a mixture THF/^{t}BuOH/H_2O 3 : 1 : 1, was added Na-*L*-ascorbate (196 mg, 0.99 mmol, 2 equiv.). The reaction mixture was stirred under an argon atmosphere at r.t. overnight. The reaction was quenched by addition of a water solution containing EDTA (5 %) and $NaHCO_3$ (5 %,), and the mixture was further stirred for 30 min. The organic layer was separated and the water phase further washed with CH_2Cl_2 (2 × 100 mL). The organic fractions were combined and the solvent was evaporated at reduced pressure. Reverse phase high-performance flash chromatography, using a linear gradient of methanol (5→100%) in water over 1 hour, afforded the desired compound as a white solid (201 mg, 38%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 9.03 (dq, *J* = 4.7, 2.8, 2.4 Hz, 2H), 8.15 (dd, *J* = 8.2, 1.7 Hz, 2H), 7.97 (d, *J* = 8.7 Hz, 2H), 7.87 (s, 1H), 7.82 (s, 1H), 7.67 – 7.53 (m, 3H), 7.49 (ddd, *J* = 8.1, 4.4, 1.5 Hz, 2H), 7.35 (d, *J* = 7.2 Hz, 2H), 7.23 (td, *J* = 10.8, 9.7, 5.4 Hz, 6H), 7.18 – 7.05 (m, 3H), 6.78 (dd, *J* = 9.1, 3.1 Hz, 4H), 6.23 (t, *J* = 6.5 Hz, 1H), 5.10 – 4.93 (m, 3H), 4.64 – 4.53 (m, 2H), 4.38 (d, *J* = 14.5 Hz, 1H), 4.23 (s, 2H), 4.09 (d, *J* = 3.2 Hz, 1H), 3.71 (s, 6H), 3.69 – 3.61 (m, 1H), 3.35 – 3.24 (m, 2H), 2.52 – 2.37 (m, 1H), 1.29 (d, *J* = 11.4 Hz, 4H), 0.95 – 0.81 (m, 1H). HRMS (ESI+): calculated for $C_{61}H_{55}O_9N_{10}$ = 1071.41480; found: 1071.41488. HRMS (ESI+): calculated for $C_{61}H_{55}O_9N_{10}$ = 1071.41480; found: 1071.41488.

Synthesis of CEP-dC^{ocp}



The title product **CEP-dC^{ocp}** was obtained from its precursor Clip-Phen linked 5'-*O*-DMT precursor (2.5 equiv., 130 mg, 0.089 mmol), *N*,*N*-diisopropylethylamine (2.5 equiv., 0.226 mmol, 39 µL) and 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite (1.2 equiv, 0.107 mmol, 24 µL) following the general procedure E. The product was precipitated with n-pentane from the reaction mixture and recovered by filtration as a yellow solid (85 mg, 66 %) without any further purification. NMR (400 MHz, CD_3CN) δ 8.95 – 8.76 (m, 2H), 8.73 (m, 1H), 8.45 – 8.22 (m, 3H), 8.16 – 7.93 (m, 3H), 7.74 – 7.49 (m, 6H), 7.31 – 6.44 (m, 14H), 6.42 (t, *J* = 5.5 Hz, 1H), 5.55 – 5.36 (m, 1H), 5.48 – 5.32 (m, 5H), 4.52 (s, 1H), 4.62 – 4.54 (m, 1H), 4.21 (bs, 1H), 3.70 – 3.41 (m, 14H), 3.39 – 3.20 (m, 4H), 2.77 – 2.64 (m, 1H), 2.58 – 2.46 (m, 4H), 2.21 – 2.05 (m, 3H), 1.66 – 1.44 (m, 6H), 1.43 – 1.11 (m, 18H), 1.01 – 0.90 (m, 6H) $^{31}P\{^1H\}$ NMR (202.4 MHz, CD_3CN): 150.11 and 149.86. HRMS (ESI+): calculated for $C_{74}H_{74}N_{11}NaO_{10}P$ = 1330,53740; found: 1071.53399.

Synthesis of CEP-dU^{amcp}



The title product **CEP-dU^{amcp}** was obtained from its precursor Clip-Phen linked 5'-*O*-DMT precursor (2.5 equiv., 0.601 mmol, 104 µL) and 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite (1.2 equiv, 0.288 mmol, 64 µL) following the general procedure E. The product was precipitated with n-pentane from the reaction mixture and recovered by filtration as a yellow solid (153 mg, 48%) without any further purification. $^{31}P\{^{1}H\}$ NMR (202.4 MHz, CD_3CN): 151.30 and 151.07. HRMS (ESI+): calculated for $C_{70}H_{71}N_{12}NaO_{10}P$ = 1293,51003; found: 1293,51064.

5.2. Solid-phase synthesis of oligonucleotides

Material and methods

Solvents and reagents for the solid-phase synthesis of ONs were purchased from Link Technologies, Sigma-Aldrich and Thermo Fisher-Scientifics. Natural phosphoramidites were purchased from Sigma-Aldrich. Modified ONs bearing internal modifications were synthesized through standard phosphoramidite chemistry with an automated DNA synthesizer (ABI 3400, Applied Biosystems), employing the modified synthesized amidites **13a**, **13b**, **15a** and **15b**, or the commercially available amidite **15c** (Baseclick GmbH). The 3'-modified TFO **A7** was obtained employing a commercial alkyne-modified controlled-pore glass (Baseclick GmbH). Deprotections were carried out incubating the reaction products in 30 % aqueous NH_3 at 55 °C for 6 hours. Purification of prepared ONs was performed using semi-preparative HPLC (Waters modular HPLC system) on a C18 reversed phase column (Phenomenex, Luna C18 (2) 100 Å), using a linear gradient of acetonitrile (3.5 → 40 %) in 0.1 M TEAA buffer. The products were analyzed by MALDI-TOF MS. Concentration of DNA solutions was calculated using A260 values measured on a Cary100 Bio UV-Vis spectrophotometer (Varian). Sequences of synthesized ONs are shown in the following table. All sequences are shown in Table 8.

Table 8. Sequences of modified ONs synthesized on solid support

Oligo	Sequence 5'→ 3'	Length	Amidite
ON-U^{peg}	TGTGTGTGTGG**[dU^{peg}]**GGTTGTTGTTGTTTTGGTTG	32mer	**15a**
ON-U^{un}	TGTGTGTGTGG**[dU^{un}]**GGTTGTTGTTGTTTTGGTTG	32mer	**15b**
TFO-24-5'peg (A1)	**[5'peg]** TCTTCTTCTTCCGCCTCTCTCTCT	24mer	**13a**
TFO-32-5'peg (A2)	**[5'peg]** CTTCCTTGTCTTCTTCTTCCGCCTCTCTCTCT	32mer	**13a**
TFO-32-5'un (A3)	**[5'un]** CTTCCTTGTCTTCTTCTTCCGCCTCTCTCTCT	32mer	**13a**
TFO-32-dU^{peg} (A4)	CTTCCTTGTCTTCTTC **[dU^{peg}]** TCCGCCTCTCTCTCT	32mer	**15a**
TFO-32-dU^{un} (A5)	CTTCCTTGTCTTCTTC **[dU^{un}]** TCCGCCTCTCTCTCT	32mer	**15b**
TFO-32-dU^{oct} (A6)	CTTCCTTGTCTTCTTC **[dU^{oct}]** TCCGCCTCTCTCTCT	32mer	**15c**
TFO-32-3'alk (A7)	CTTCCTTGTCTTCTTCTTCCGCCTCTCTCTCT **[3'alk]**	32-mer	Alk-CPG

5.3. Enzymatic synthesis of modified DNA

Materials and methods

Synthetic oligonucleotides were purchased from Generi Biotech. The sequences of all employed oligonucleotides are listed in Table 12, paragraph 5.3.1. KOD XL DNA polymerase was purchased from Novagen, Bst large fragment and Vent(exo-) from New England Biolabs, and Pwo from Peqlab. All restriction endonucleases were purchased from New England Biolabs. Unmodified nucleoside triphosphates (**dATP**, **dTTP**, **dCTP**, and **dGTP**) were purchased from Fermentas. **dC^{o}TP**,[112] **dC^{vi}TP**,[15] **dC^{e}TP**,[12] **dC^{ph}TP**,[12] **dU^{am}TP**,[113] **dU^{o}TP**,[112] **dU^{vi}TP**,[15] **dU^{e}TP**,[12] **dU^{ph}TP**,[12] **dG^{me}TP**,[13] **dG^{vi}TP**,[13] **dG^{e}TP**,[13] **dG^{ph}TP**,[13] **dA^{me}TP**,[138] **dA^{vi}TP**,[15] **dA^{e}TP**,[14] **dA^{ph}TP**,[14] were prepared as described previously, or from the relative nucleoside synthesized as described in the references, followed by classic 5'-*O*-triphosphorilation.[114] **dC^{me}TP** and **dG^{7d}TP** were purchased by Jena Bioscience. Other chemicals were of analytical grade. Concentration of DNA solutions was calculated using A260 values measured on a Nanodrop 1000 Spectrophotometer (Thermo Scientific) or on a Cary100 Bio UV-Vis spectrophotometer (Varian). Stop solution contained 80% (v/v) formamide, 20 mM EDTA, 0.025% (w/v) bromophenol blue, 0.025% (w/v) xylene cyanol in water. All gels were analysed with phosphorimager (Typhoon 9410, Amersham Biosciences). All solutions for biochemistry experiments were prepared in Milli-Q water. Mass spectra of oligonucleotides were measured by MALDI-TOF MS, on UltrafleXtreme MALDI-TOF/TOF mass spectrometer (Bruker Daltonics, Germany), with 1 kHz smartbeam II laser.

NEAR for the production of alkynyl- or azido-modified antiparallel TFOs

In a typical experiment, the needed template (**Temp-NEAR**, see Table 9, 0.125 µM), primer **Prim-Nea**r (0.125 µM), modified dU^{R}TP (156 µM), natural dNTPs (125 µM), 1 × ThemoPol buffer (10 mM KCl, 10 mM $(NH_4)_2SO_4$, 20mM Tris-HCl - pH 8.8-, 0.1 % Triton-X-100, and 2 mM $MgSO_4$), 0.5× NEBuffer 3 (50mM NaCl, 25 mM Tris-HCl - pH 7.9-, 5 mM $MgCl_2$, and 0.5 mM DDT), 0.10 U/µL Vent(exo-) DNA Polymerase, and 0.60 U/µL Nt.BstNBI, were incubated in a total volume of 750 µL for 3 h at 55 °C. The reactions were quenched by cooling at 4°C, concentrated under vacuum at 55 °C to 100 µL, and the products were purified by reverse phase HPLC (C18 column).

Table 9. Sequences of ONs employed in NEAR and NEAR products

Oligo	Sequence 5'→ 3'	Lenght
Prim-NEAR	CCGATCTAGT*GA GTCCTCG	19-mer
Temp-dU-NEAR	CCCCCCTTTTATTTTCGAGGACTCA*CTAGATCGG	34-mer
TFO-Nef-dU^{e}	AAAA[**U^{e}**]AAAAGGGGGG	15-mer
TFO-Nef-dU^{o}	AAAA[**U^{o}**]AAAAGGGGGG	15-mer
TFO-Nef-dU^{am}	AAAA[**U^{am}**]AAAAGGGGGG	15-mer

*: Cleavage site of the nicking enzyme; underlined: recognition sequence for the nicking enzyme.

PEX on analytical scale

Enzymatic incorporation of dU^{peg}TP by PEX on analytic scale

The reaction mixture (20 µl) contained KOD XL DNA polymerase (0.15 U), natural dNTPs (**dATP**, **dGTP** and **dCTP** 0.05 mM each), **dU^{peg}TP** (0.1 mM), 5'-6-FAM labelled primer **Prim^{PEX}–FAM** (0.15 µM), template **Temp1^{PEX}** or **Temp4^{PEX}** (0.225 µM) in KOD XL polymerase reaction buffer supplied by the manufacturer. The reaction mixtures were incubated for 20 min at 60 °C in a thermal cycler. Stop solution was added and the reaction mixture was denatured at 95 °C for 5 min and analyzed using 12.5 % denaturing PAGE. Visualization was performed on Typhoon biological imager FLA 9500 (PAGE gels are shown in Figure 30, *Appendix 1*).

Enzymatic incorporation of dU^{un}TP by PEX on analytic scale

The reaction mixture (20 µl) contained KOD XL DNA polymerase (0.25 U), natural dNTPs (**dATP**, **dGTP** and **dCTP** 0.025 mM each), **dU^{un}TP** (0.2 mM), 5'-6-FAM labelled primer **Prim^{PEX}–FAM** (0.15 µM), template **Temp1^{PEX}** or **Temp4^{PEX}** (0.225 µM) in KOD XL polymerase reaction buffer supplied by the manufacturer. The reaction mixtures were incubated for 35 min at 60 °C in a thermal cycler. Stop solution was added and the reaction mixture was denatured at 95 °C for 5 min and analyzed using 12.5 % denaturing PAGE. Visualization was performed on Typhoon biological imager FLA 9500 (PAGE gels are shown in Figure 30, *Appendix 1*).

PEX on semi-preparative scale

General procedure F: semi-preparative production of modified oligonucleotides ON_1U^{R} and ON_4U^{R} through PEX, and subsequent magnetic separation

The reaction mixture (200 µl) contained KOD XL DNA polymerase (2.5 U), natural dNTPs (**dATP**, **dGTP** and **dCTP**, 2mM each, 20 µl); **TTP**, **dU^{e}TP**, **dU^{o}TP**, **dU^{peg}TP** or **dU^{un}TP** (8 mM, 20 µl), **Prim^{PEX}** (100 µM, 10 µl), 5'-biotinylated template **Temp1^{PEX}-bio** (for the synthesis of **ON_1U^{R}**) or **Temp4^{PEX}-bio** (for the synthesis of **ON_4U^{R}**) (100 µM, 10 µl) in KOD XL polymerase reaction buffer supplied by the manufacturer. The reaction mixtures were incubated for 40 min at 60 °C in a thermal cycler and reactions were stopped by cooling at 4 °C. The reaction mixtures were added to streptavidine magnetic beads (Roche, 50 µl) previously washed with binding buffer TEN 100 (3 × 200 µl, 10 mM Tris, 1 mM EDTA, 100 mM NaCl, pH 7.5) and suspended in binding buffer TEN 100 (50 µl). The mixture was incubated at 15 °C and 1200rpm for 30 min, then the liquid phase was separated and the magnetic beads were washed with washing buffer TEN 1000 (3 × 200 µl, 10 mM Tris, 1 mM EDTA, 1 M NaCl, pH 7.5) and with milli-Q water (4 × 200 µl). 50 µl of milli-Q water were added and the mixture was incubated for 1.5 min at 65 °C to denaturate the dsDNA. The solution was quickly transferred into a clean Eppendorf and analyzed by MALDI-TOF MS.

General procedure G: semi-preparative production of modified dsDNA DNA_1U^{R} through PEX

The reaction mixture (200 µl) contained KOD XL DNA polymerase (2.5 U), natural dNTPs (**dATP**, **dGTP** and **dCTP**, 2mM each, 20 µl); **dTTP**, **dU^{e}TP**, **dU^{o}TP**, **dU^{peg}TP** or **dU^{un}TP** (8 mM, 20 µl), **Prim^{PEX}-FAM** (100 µM, 10 µl), template **Temp1^{PEX}** (100 µM, 10 µl) in KOD XL polymerase reaction buffer supplied by the manufacturer. The reaction mixtures were incubated for 40 min at 60 °C in a thermal cycler and reactions were stopped by cooling at 4 °C. Products were purified with QIAQuick Nucleotide Removal kits (QIAGEN) and eluted in milli-Q water to be employed in the next experiments.

Competitive PEX experiments: Direct Competition Assay

Primer Extension Experiment

The reaction mixture (40 µl) contained DNA polymerase (Bst large fragment (0.3 U), KOD XL (0.25 U), Pwo (0.3 U) or Vent(exo-) (0.3 U)), natural dNTPs (0.05 mM), for ratio 1 : 1 dNTPs (0.05 mM) and dN^RTP (0.05 mM), for ratio 1 : 10 dNTP (0.01 mM) and dN^RTP (0.1 mM), the suitable 5'-6-FAM-labelled primer **$Prim^{RE-N}$** (0.15 µM), the suitable template **$Temp^{RE-N}$** (0.225 µM) in the suitable polymerase reaction buffer, supplied by the manufacturer. The reaction mixtures were incubated for 30 min at 60 °C (or 65 °C, when Bst large fragment polymerase was used) in a thermal cycler and divided into two portions of volume 20 µl. In the first portion, the PEX reaction was quenched by addition of stop solution (20 µL) which was then denatured at 95 °C for 5 min and analyzed using 12.5 % denaturing PAGE. Visualization was performed on Typhoon biological imager FLA 9500. The second portion was used in following cleavage reaction. All experiments were performed at least in triplicate. PAGE gels are shown in Figures 42 and 43 *Appendix 1*.

Restriction Endonucleases cleavage

The 20 µl portion of product of PEX experiments were mixed with 1 x reaction buffer supplied by manufacturer relevant to RE and corresponding RE (20 U). Reaction mixtures were incubated at 37 °C for 60 min (or 30 min, in the case of competitive PEX experiments of **$dU^{hm}TP$** and **$dU^{f}TP$**) and then the stop solution was added (to reach 40 µl total volume). Products of cleavage by RE were denatured at 95 °C for 5 min and analyzed using 12.5 % denaturing PAGE. Visualization was performed on Typhoon biological imager FLA 9500. PAGE gels are shown in Figures 42 and 43, *Appendix 1*.

Steady state kinetics assay

Assays typically contained the suitable 5'-6-FAM labelled primer **$Prim^{PEX}$-FAM** (1 µM) and the suitable template **$TempTerm^{N}$** (1 µM), ThermoPol buffer (1 x) and Bst large fragment DNA polymerase (0.01 U), in a total volume of 10 µL. Reactions were initiated via addition of varying concentrations of natural or modified dNTPs, mixtures incubated at 65 °C for 3 min and reactions quenched with 10 µL of stop solution. Products were denatured at 95 °C for 5 min and separated

using 20 % denaturing PAGE. Kinetic parameters were determined by fitting data to the Michaelis-Menten equation using KaleidaGraph 4.0. All rates were normalized to the same final enzyme activity specified by suppliers (1 U). The reported discrimination values were determined by comparing the efficiency (k_{cat}/K_M) of incorporation for the modified to the efficiency (k_{cat}/K_M normalized to 1) of incorporation for the corresponding natural nucleotide. All the experiments were performed in duplicate (Examples of PAGE analysis are shown in Figure 44, *Appendix 1*).

Enzymatic synthesis of longer DNA through polymerase chain reaction (PCR)

Incorporation of dU^{e}TP, dU^{o}TP, dU^{peg}TP and dU^{un}TP into 98bp ONs by KOD XL polymerase

The PCR mixture (10 µL) contained KOD XL DNA polymerase (1 U for the incorporation of natural TTP and 2-4 U for **dN^{R}TPs**), natural dNTPs (**dATP**, **dGTP** and **dCTP** 0.1 - 0.2 µM), **dU^{R}TP** (0.4 µM), primers **PrimFOR-98** and **PrimREV-98** (2 µM), template **TempPCR-98** (20 nM) in KOD XL polymerase reaction buffer supplied by the manufacturer, and $MgSO_4$ (10 mM). After a preheating at 94 °C for 3 min, 30 PCR cycles were run under the following conditions: denaturation for 1 min at 95 °C, annealing (53 – 57 °C, 1 min), extension (72 °C, 1 – 2.5 min), followed by final extension step 5 min at 75 °C. PCR products were purified by QIAQUICK PCR purification kits and analyzed on a 2 % agarose gel in 0.5 x TBE buffer, followed by staining with GelRed (Biotium, 1x) (agarose gel shown in Figure 31, *Appendix 1)*.

Incorporation of dU^{e}TP, dU^{o}TP and dU^{peg}TP into 235bp ONs by KOD XL polymerase

The PCR mixture (10 µL) contained KOD XL DNA polymerase (1-4 U), natural dNTPs (**dATP**, **dGTP** and **dCTP** 0.1-0.2 µM), **dU^{R}TP** (0.1-0.4 µM), primers **PrimFOR-235** and **PrimREV-235** (2 µM), template **TempPCR-235** (10 nM) in KOD XL polymerase reaction buffer supplied by the manufacturer, and $MgSO_4$ (10 mM). After a preheating at 95 °C for 3 min, 40 PCR cycles were run under the following conditions: denaturation for 1 min at 95 °C, annealing (55 – 57 °C, 1 min), extension (72 °C, 1 – 2 min), followed by final extension step 5 min at 72 °C. PCR products were purified by QIAQUICK PCR purification kits and analyzed on a 1.3 % agarose gel in 0.5 x TBE buffer, followed by staining with GelRed (Biotium, 1x) (agarose gel shown in Figure 31, *Appendix 1*).

Table 10. Conditions for the incorporation of **dN^{R}TPs** into 98bp DNA through PCR.

dU^{R}TP	Natural dNTPs (μM)	KOD XL (U)	Annealing °C / time	Elongation °C / time
TTP (+)	100 μM	1 U	53°C / 1 min	72°C / 1 min
dU^{e}TP	200 μM	2 U	54.5°C / 1 min	72°C / 1.5 min
dU^{o}TP	200 μM	4 U	54.5°C / 1 min	72°C / 1.5 min
dU^{peg}TP	200 μM	4 U	54.5°C / 1 min	72°C / 1.5 min
dU^{un}TP	200 μM	4 U	57°C / 1 min	72°C / 2.5 min

Table 11. Conditions for the incorporation of dN^{R}TPs into 235bp DNA through PCR.

dN^{R}TP	Natural dNTPs (μM)	**dU^{R}TP** (μM)	KOD XL (U)	Annealing °C / time	Elongation °C / time
TTP (+)	100 μM	100 μM	1 U	55°C / 1 min	72°C / 1 min
dU^{e}TP	200 μM	200 μM	2 U	55°C / 1 min	72°C / 1.5 min
dU^{o}TP	200 μM	400 μM	4 U	57°C / 1 min	72°C / 2 min
dU^{peg}TP	200 μM	400 μM	4 U	57°C / 1 min	72°C / 2 min

Incorporation of dU^{ph}TP and dG^{ph}TP by KOD XL polymerase in PCR experiment, followed by cleavage by EcoRI

The PCR mixture (20 μL) contained KOD XL DNA polymerase (2 U), natural dNTPs (0.2 mM), for ratio 10/1 dNTPs (0.182 mM) and **dU^{ph}TP** or **dG^{ph}TP** (0.018 mM), for ratio 5/1 dNTP (0.166 mM) and dN^{ph}TP (0.034 mM), for ratio 2/1 dNTP (0.134 mM) and dN^{ph}TP (0.066 mM), for ratio 1/1 dNTP (0.1 mM) and dN^{ph}TP (0.1 mM), for ratio 1/2 dNTP (0.066 mM) and dN^{ph}TP (0.134 mM), for ratio 1/5 dNTP (0.034 mM) and dN^{ph}TP (0.166 mM) and for ratio 1/10 dNTP (0.018 mM) and dN^{ph}TP (0.182 mM), primers **PrimFOR-239** and **PrimREV-239** (5'-6-FAM-labeled in experiment with **dG^{ph}TP**, 2 μM), template **TempPCR-239** (2.3 ng) in KOD XL polymerase reaction buffer supplied by the manufacturer. 40 PCR cycles were run under the following conditions:

denaturation for 1 min at 94 °C, annealing for 1 min at 69 °C, extension for 1 min at 72 °C, followed by final extension step 5 min at 72 °C. PCR products were cleaved by EcoRI and analyzed on a 1.3 % agarose gel in 0.5 x TBE buffer, followed by staining with GelRed (Biotium, 1x) in the experiment with **dU^{ph}TP**. The visualization of agarose gel analysis of the experiment with **dG^{ph}TP** was performed on Typhoon biological imager FLA 9500.

PCR synthesis of radioactively labelled target duplex

The PCR mixture (50 µL) contained KOD XL DNA polymerase (1 U) natural dNTPs (32 µM), α-^{32}P dATP (0.25 mCi, 0.33 µM), primers **PrimFOR-98** and **PrimREV-98** (2 µM), template **TempPCR-98** (20 nM) in KOD XL polymerase reaction buffer supplied by the manufacturer). After a preheating at 94°C for 2 min, 24 PCR cycles were run under the following conditions: denaturation for 30 sec at 94 °C, annealing for 30 sec at 49 °C, extension for 30 sec at 72 °C , followed by final extension step 30 sec at 72 °C. PCR products were purified by QIAQUICK PCR purification kits. The PCR optimization was performed in the absence of α-^{32}P dATP. DNA concentrations were measured on non-radioactively labelled PCR products, assuming them to be identical to the radioactively labelled PCR products obtained with the same protocol. For the synthesis of the radioactively labelled DNA ladder, primers **Prim90**, **Prim85**, **Prim80**, **Prim75** and **Prim 70** were employed instead of **PrimFOR-98**, using the same template and conditions.

5.3.1. Table of ONs employed in the enzymatic synthesis

Table 12. Sequences of ONs employed in enzymatic synthesis

Oligo	Sequence 5'→ 3'	Length
Prim^{PEX}	CATGGGCGGCATGGG	15mer
Temp1^{PEX}	CCCACCCATGCCGCCCATG	19mer
Temp4^{PEX}	CTAGCATGAGCTCAGTCCCATGCCGCCCATG	31mer
Prim^{AfeI-C}	CATGGGCGGCATGGGAGCG	19mer
Temp^{AfeI-C}	AACTACTAC**AGC/GCT**CCCATGCCGCCCATG	30mer
Prim$^{EcoRI-T}$	CATGGGCGGCATGGGGAAT	18mer
Temp$^{EcoRI-T}$	TTCGTCGTC**G/AATTC**CCCATGCCGCCCATG	30mer
Prim$^{EcoRI-A}$	CATGGGCGGCATGGGGA	17mer
Temp$^{EcoRI-A}$	AACGACGAC**G/AATTC**CCCATGCCGCCCATG	30mer
Prim$^{EcoRI-G}$	Same as **Prim^{PEX}**	
Temp$^{EcoRI-G}$	AAGTAGTAG**G/AATTC**CCCATGCCGCCCATG	30mer
Prim$^{BglII-G}$	Same as **Prim^{PEX}**	
Temp$^{BglII-G}$	AAGTAGTAG**A/GATCT**CCCATGCCGCCCATG	30mer
Prim^{AfeI-G}	CATGGGCGGCATGGGAGC	18mer
Temp^{AfeI-G}	AAGTAGTAG**AGC/GCT**CCCATGCCGCCCATG	30mer
Prim^{RsaI-G}	Same as **Prim^{PEX}**	
Temp^{RsaI-G}	AAGTAGTAGT**GT/AC**ACCCATGCCGCCCATG	30mer
Prim^{SphI-T}	Same as **Prim^{PEX}**	
Temp^{SphI-T}	TTCGTCGTC**GCATG/C**CCCATGCCGCCCATG	30mer
Prim^{SacI-T}	Same as **Prim^{PEX}**	
Temp^{SacI-T}	TTCGTCGTC**GAGCT/C**CCCATGCCGCCCATG	30mer
Prim^{KpnI-T}	Same as **Prim^{PEX}**	
Temp^{KpnI-T}	TTCGTCGTC**GGTAC/C**CCCATGCCGCCCATG	30mer
Temp$^{PvuII-G}$	AAGTAGTAG**CAG/CTG**CCCATGCCGCCCATG	30mer
TempTerm^{C}	GCCCATGCCGCCCATG	16mer
TempTerm^{T}	ACCCATGCCGCCCATG	16mer

Oligo	Sequence 5'→ 3'	Length
TempTerm^{A}	TCCCATGCCGCCCATG	16mer
TempTerm^{G}	CCCCATGCCGCCCATG	16mer
PrimFOR-98	GACATCATGAGAGACATCGC	20mer
PrimREV-98	CAAGGACAAAATACCTGTATTCCTT	25mer
Prim90	TTCACCGACAATATTCGAAGGAAT	24mer
Prim85	CGACAATATTCGAAGGAATAGAAG	24mer
Prim80	ATATTCGAAGGAATAGAAGAAGAAGG	26mer
Prim75	CGAAGGAATAGAAGAAGAAGGTG	23mer
Prim70	GAATAGAAGAAGAAGGTGGAGAG	23mer
TempPCR-98	GACATCATGAGAGACATCGCCTCTGGGCTAATAGGACTACTTCTA ATCTGTACAGAGCAGATCCCTGGACAGGAAGGAATACAGGTATT TTGTCCTTG	98mer
PrimFOR-235	CGTCTTCAAGAATTCTATTTGACA	24mer
PrimREV-235	GGAGAGCGTTCACCGACA	18mer
TempPCR-235	CGTCTTCAAGAATTCTATTTGACAAAAATGGGCTCGTGTTGTACA ATAAATGTGTCTAAGCTTGGGTCCCACCTGACCCCATGCCGAACT CAGAAGTGAAACGCCGTAGCGCCGATGGTAGTGTGGGGTCTCCC CATGCGAGAGTAGGGAACTGCCAGGCATCAAATAAAACGAAAGG CTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGT GAACGCTCTCC	235mer
PrimFOR-239	TAGGGGTTCCGCGCACATTTCCCCG	25mer
PrimREV-239	GGAGAGCGTTCACCGACAAACAACAG	26mer
TempPCR-239	TAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTCT AAGAAACCATTATTATCATGACATTAACCTATAAAAATAGGCGTA TCACGAGGCCCTTTCGTCTTCAA**GAA/TTC**TATTTGACAAAAATGG GCTCGTGTTGTACAATAAATGTGTCTAAGCTTGGGTCCCACCTGA CCCCATGCCGAACTCAGAAGTGAAACGCCGTAGCGCCGATGGTA GTGTGGGGTCTCCCCATGCGAGAGTAGGGAACTGCCAGGCATCA AATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTAT CTGTTGTTTGTCGGTGAACGCTCTCC	239mer

Bold: recognition sequence for REs; /: spot of cleavage of REs

5.4. Post-synthetic modification of ONs through CuAAC

CuAAC between TAMRA-N_3 and synthesized ONs ON-U^{peg} and ON-U^{un}

To 50 µL of starting oligonucleotide (100 µM , 5nmol) were added 25 µL of a DMSO/ ^{t}BuOH 3:1 mixture and 5 µL of a DMSO solution of 5-carboxytetramethylrhodamine-azide (TAMRA-N_3, 5 mM , 5 equiv.) and 25 µL of a freshly prepared DMSO solution of CuOAc (500 µM) and TBTA (2.5 mM). The reaction mixture was stirred for 5 h at room temperature, then freeze-dried, and the crude solid was re-suspended in 100 µL of milli-Q water to be purified on QIAQUICK Nucleotide Removal kit (QIAGEN). The products were analysed by MALDI-TOF MS (spectra in Figures 46 and 47, *Appendix 3*).

CuAAC between synthesized alkynyl-TFOs (A1-A7) and N_3CP

To 50 µL of starting oligonucleotide (100 µM , 5nmol) were added 25 µL of a DMSO/ ^{t}BuOH 3:1 mixture and 5 µL of a DMSO solution of **N_3CP** (**X**, 5 mM , 5 equiv.) and 25 µL of a freshly prepared DMSO solution of CuOAc (500 µM) and TBTA (2.5 mM). The reaction mixture was stirred overnight at room temperature, and an aqueous solution of EDTA (100 µL, 100 mM) was added to chelate the copper. The mixtures was freeze-dried, and the crude solid was re-suspended in 100 µL of milli-Q water to be purified on QIAQUICK Nucleotide Removal kit (QIAGEN). The products were analyzed by MALDI-TOF MS (spectra in Figures 48-54, *Appendix 3*).

Post-synthetic CuAAC on enzymatically obtained ONs

CuAAC of PEX products and TAMRA-N_3 for steady-state fluorescence measurements

To **ON-$1U^{R}$** obtained from semi-preparative primer extension following the general procedure F, (1 nmol in 30 µL of water) were added a DMSO solution 5-carboxytetramethylrhodamine-azide (TAMRA-N_3, 2.5 mM, 2 µL), and a freshly prepared solution (10 µL) of CuBr (500 µM) and TBTA (2.5 mM) in DMSO). The reaction mixtures were incubated for 15 min or 1 hour at 37 °C, then 1 mL of PNI buffer was added and the mixtures were purified on QIAQUICK Nucleotide Removal kit (QIAGEN) and eluted in 120 µL of milli-Q water. 20 µL of the solution were employed for MALDI-

TOF MS analysis. Emission spectra were recorded in a 100 µL quartz cuvette at 25 °C with the remaining 100 µL of the solution. Excitation wavelength was 520 nm (Figure 55 in *Appendix 4*).

CuAAC of PEX products and PEG-N_3 for quantification by PAGE

DNA_1U^R was obtained from semi-preparative primer extension following the general procedure G. To a solution of 0.2 nmol of modified **DNA_1U^R** in water (30 µL), were added 4 µL of a DMSO/^{t}BuOH 3:1 mixture, and 5 µL of a freshly prepared solution of CuBr (10 mM) and TBTA (50 mM) in DMSO. A DMSO solution of poly(ethylene glycol) methyl ether azide (PEG-N_3, 1 mM, 1 µL, 5 equiv.) was added and the reaction mixtures were quickly incubated for 15 min, 30 min, 1 hour or 5 hours at 37 °C, then 1 mL of PNI buffer was added and the mixtures were purified on QIAQUICK Nucleotide Removal kit (QIAGEN). The products were analysed using 20 % PAGE gel, visualization was performed on Typhoon biological imager FLA9500. The degree of conversion of the reaction was calculated on the basis of the spot intensity (PAGE gels in Figure 32, *Appendix 1*).

CuAAC of PEX products and PEG-N_3 for the determination of the observed reaction rate constants

DNA_1U^R was obtained following the general procedure G. To a solution of 0.2 nmol of modified **DNA_1U^R** in water (30 µL), were added 4 µL of a DMSO/^{t}BuOH 3:1 mixture, and 5 µL of a freshly prepared solution of CuBr (10 mM) and TBTA (50 mM) in DMSO. A DMSO solution of PEG-N_3 (1 mM, 1 µL, 5 equiv.) was added and the reaction mixtures were quickly incubated at 37 °C, 800 rpm. Aliquots of 4 µL were taken at appropriate time intervals and added to 6 µL of a prepared solution of EDTA (100 mM) and propiolic acid (100 µM) (pH 7.6 adjusted with $NaHCO_3$ 0.1 M) to quench the reaction. The products were analysed using 20% PAGE gel, visualization was performed on Typhoon biological imager FLA9500. The degree of conversion of the reaction at every time interval was calculated on the basis of the spot intensity (PAGE gels in Figure 33, *Appendix 1*; graphs in figure 45, *Appendix 2*). The measured conversions were plotted against time and the curves were fitted employing OriginPro 2016 with the exponential equation Exp1dec ($y = A_1e^{-k} + y_0$) to obtain the observed reaction rate constants k_{obs}.

CuAAC of PCR products and TAMRA-N_3 for steady-state fluorescence measurements

98 bp modified dsDNA was obtained from PCR experiments following the general procedure G and purified by QIAQUICK PCR purification kits (QIAGEN). To 30 µL of DNA solution were added 10 µL of a 3:1 DMSO/^{t}BuOH mixture, a DMSO solution of TAMRA-N_3 (5 mM, 5 µL), and a freshly prepared solution (5 µL) of CuBr (5 mM) and TBTA (25 mM) in DMSO. The reaction mixtures were quickly incubated at 37 °C for 15 minutes or 1 hour, then 1 mL of PNI buffer was added and the mixture was purified on QIAQUICK Nucleotide Removal kit (QIAGEN) and eluted in 100 µL of milli-Q water. Emission spectra were recorded in a 100 µL quartz cuvette at 25°C. The excitation wavelength was 520 nm (Figure 56 in *Appendix 4*).

5.5. UV triplex melting curves

Melting curves of DNA triplexes were recorded on a Cary 100 Bio UV-Vis spectrophotometer (Varian), in 1 cm path-length quartz cells, monitoring absorbance at 260 nm while slowly increasing temperature.[135] All the solutions of duplexes or triplexes were prepared in 10 mM phosphate buffer, containing 250 mM NaCl and 1 mM $MgCl_2$ (pH 6.2). In a typical experiment, an equimolar solution of target duplex and TFO (1 μM each) was pre-annealed by incubation at 10 °C for 1 hour. For measurements of Cu-Clip-Phen modified TFOs **T1-T7**, solutions contained also 1 μM $CuCl_2$. The melting run was performed with a temperature increase of 1 °C/min, in the range of 15 °C - 85 °C. After each run, samples were cooled back to 15 °C with a rate of 0.2 °C/min. The denaturation-annealing process was repeated 3 times. UV melting curves are shown in Figures 57 and 58, *Appendix 5*.

5.5.1. Table of UV triplex melting temperatures.

Table 13. Triplex melting temperatures

Oligo	Alkyne	Clip-phen	Cu(II) complex
TFO-24-5'peg	24.1	31.9	35.2
TFO-32-5'peg	37.7	39.0	44.5
TFO-32-5'un	37.1	39.1	44.9
TFO-32-dU^{peg}	43.6	45.2	48.8
TFO-32-dU^{un}	43.0	45.3	49.2
TFO-32-dU^{oct}	48.5	50.0	50.2
TFO-32-3'alk	38.7	39.8	44.4

5.6. DNA cleavage assays.

Cleavage of target dsDNA in the presence of non-targeted DNA

All mother solutions of duplexes, TFOs or triplexes were prepared in 10 mM phosphate buffer, containing NaCl 250 mM and $MgCl_2$ 1 mM, at pH 6.2. In a typical experiment, 2 µL of a 1:1 mixture of target duplex (**D**, 1 µM, 2 pmol) and off-target duplex (**Off**, 1 µM, 2 pmol) were incubated for 1 h at 10 °C, in the presence of a premixed solution of Clip-Phenanthroline-modified TFO (**T1-T7**, 1 µL, concentration in dependence of the equivalents needed, see Table 15, paragraph 5.6.2.) and $CuCl_2$ (concentration in dependence of the equivalents needed, always 1 equiv. with respect to the TFO). A solution of Na-$_L$-ascorbate (1 µL, concentration in dependence of the equivalents needed, see Table 15) in buffer, was added and the mixture was incubated at the desired temperature for the desired time (see Table 15). Control experiments were performed testing: **A**) a lack of copper (control **Cu**) in the presence of Na-$_L$-ascorbate and Clip-phenanthroline-modified TFO (**T2** was chosen for the control, as being one of the most active); **B**) a lack of Na-$_L$-ascorbate (control **Asc**) in the presence of Clip-phenanthroline-modified TFO and $CuCl_2$; **C**) a lack of TFO (control **TFO** or **ctrl**) in the presence of Na-$_L$-ascorbate and $CuCl_2$.

The reactions were quenched with 10 µL of a stop solution containing 80% (v/v) formamide, 100 mM EDTA, 0.025% (w/v) bromophenol blue, 0.025% (w/v) xylene cyanol in water. The reactions were analysed by polyacrylamide electrophoresis and the cleavage selectivity was calculated based on the target duplex spot intensity, respect to the off-target duplex one. Selected examples of denaturing PAGE gels are reported in Figures 34-37, *Appendix 1*.

Control experiments: alkynyl-TFO (A1-A7) and Clip-phenanthroline TFOs (C1-C7).

All mother solutions of duplexes, TFOs or triplexes were prepared in 10 mM phosphate buffer, containing NaCl 250 mM and $MgCl_2$ 1 mM, at pH 6.2. 2 µL of a 1:1 mixture of target duplex (**D**, 1 µM) and off-target duplex (**Off**, 1 µM) were incubated for 1 h at 10 °C in the presence of **A**) a premixed solution of alkynyl-modified TFOs (**A1-A7**, 1 µL, 10 µM, 5 equiv.) and $CuCl_2$ (1 equiv. with respect to the TFO); **B**) in the presence of Clip-Phenanthroline-modified TFOs (**C1-C7**, 1 µL,

10 μM, 5 equiv.) in the absence of $CuCl_2$. Na-$_L$-ascorbate (1 μL, 2 mM, 1000 equiv.) was added and the mixture was incubated at 37 °C for 12 h. An example of PAGE gel in Figure 37.

Cleavage assays on radioactively-labelled target duplex obtained by PCR.

All mother solutions of duplexes, TFOs or triplexes were prepared in 10 mM phosphate buffer, containing NaCl 250 mM and $MgCl_2$ 1 mM, at pH 6.2. In a typical experiment, 2.5 μL of target duplex (approx. 120 ng/μL, 5 pmol) were incubated for 1 h at 10 °C, in the presence of a premixed solution of Clip-Phenanthroline-modified TFO (**T1-T7**, 1 μL, concentration in dependence of the equivalents needed, see gels in Figures 38-41, *Appendix 1*), and $CuCl_2$ (concentration in dependence of the equivalents needed, always 1 equiv. with respect to the TFO). A solution of (+)-sodium $_L$-ascorbate (1 μL, concentration in dependence of the equivalents needed, see gels in Figures 38-41) in buffer, was added and the mixture was incubated at 25 °C for 12 h. The reactions were quenched adding 10 μL of a stop solution containing 80% (v/v) formamide, 100 mM EDTA, 0.025% (w/v) bromophenol blue, 0.025% (w/v) xylene cyanol in water. For non-denaturing gel analysis the stop solution contained 50% glycerol, 100 mM $MgCl_2$ and 100 mM EDTA. The reactions were analysed by non-denaturing polyacrylamide electrophoresis; PAGE gels are shown in Figures 38-41.

5.6.1. Table of ONs employed as target or off-target duplexes, or for their synthesis

Table 14. Target and off-target duplexes

Oligo	Sequence 5'→ 3'	Length
D1-64	ACTGCTGGACCCGACAGGCCCGAAGGAATAGAAGAAGAAGGTGG AGAGAGAGACAGAGACTAAC	32-mer
D1-64-FAM	**[6FAM]** ACTGCTGGACCCGACAGGCCCGAAGGAATAGAA GAAGAAGGTGGAGAGAGAGACAGAGACTAAC	32-mer
D2-64	GTTAGTCTCTGTCTCTCTCTCCACCTTCTTCTTCTATTCCTTCGGGCCT GTCGGGTCCAGCAGT	32-mer
OFF-D1-52	ACCCGACAGGCCCAGGAAGGTAAGGAGGAAAAATGAGAAAGAGG GCAGAGAC	32-mer
OFF-D1-52-FAM	**[6FAM]** ACCCGACAGGCCCAGGAAGGTAAGGAGGAAAAA TGAGAAAGAGGGCAGAGAC	32-mer
OFF-D2-52	GTCTCTGCCCTCTTTCTCATTTTTCCTCCTTACCTTCCTGGGCCTGTC GGGT	32-mer
TempPCR	GGAGAGCGTTCACCGACAATATTCGAAGGAATAGAAGAAGAAGGT GGAGAGAGAGATGTGCGTGCCCGCTCGCCTGTCAAATAGAATTCT TGAAGACG	98-mer
PrimFOR	GGAGAGCGTTCACCGACA	18-mer
PrimREV	CGTCTTCAAGAATTCTATTTGACA	24-mer
Prim90	TTCACCGACAATATTCGAA GGA AT	24-mer
Prim85	CGACAATATTCGAAGGAATAGAAG	24-mer
Prim80	ATATTCGAAGGAATAGAAGAAGAAGG	26-mer
Prim75	CGAAGGAATAGAAGAAGAAGGTG	23-mer
Prim70	GAATAGAAGAAGAAGGTGGAGAG	23-mer

Bold: modification spot and type. Underlined: recognition sequence for the TFOs.

5.6.2. Table of dsDNA cleavage selectivity

Table 15. DNA cleavage selectivity

Entry	TFO (eq.)	Asc. (eq.)	T (°C)	Time (h)	TFO	Cleavage Target (%)[a]	Cleavage Off-Target (%)[b]	Selectivity (%)[c]
					T1			
					T2			
					T3			
1	0.2	200	25	12	**T4**	< 5%	< 5%	-
					T5			
					T6			
					T7			
					T1			
					T2			
					T3			
2	0.2	200	37	12	**T4**	< 5%	< 5%	-
					T5			
					T6			
					T7			
					T1	25	10	71
					T2	22	5	81
					T3	23	9	72
3	5	2000	25	12	**T4**	21	7	75
					T5	26	11	70
					T6	10	4	71
					T7	18	3	85

Entry	TFO (eq.)	Asc. (eq.)	T (°C)	Time (h)	TFO	Cleavage Target (%)[a]	Cleavage Off-Target (%)[b]	Selectivity (%)[c]
					T1	11		79
					T2	27		>95
					T3	29		93
4-A	5	1000	25	12	**T4**	34	< 5%	94
					T5	26		>95
					T6	24		82
					T7	16		80
					T1	13		86
					T2	27		>95
					T3	26		>95
4-B	5	1000	25	12	**T4**	25	< 5%	>95
					T5	23		>95
					T6	18		>95
					T7	16		>95
					T1	5		71
					T2	25		>95
					T3	28		>95
5-A	5	1000	37	3	**T4**	28	< 5%	>95
					T5	27		>95
					T6	19		>95
					T7	10		77

Entry	TFO (eq.)	Asc. (eq.)	T (°C)	Time (h)	TFO	Cleavage Target (%)[a]	Cleavage Off-Target (%)[b]	Selectivity (%)[c]
					T1	10		83
					T2	21		>95
					T3	24		>95
5-B	5	1000	37	3	**T4**	29	< 5%	>95
					T5	19		>95
					T6	18		>95
					T7	7		>95
					T1	18		78
					T2	33		94
					T3	27		>95
6-A	5	1000	37	12	**T4**	31	< 5%	>95
					T5	25		>95
					T6	40		80
					T7	11		>95
					T1	21		91
					T2	31		>95
					T3	25		>95
6-B	5	1000	37	12	**T4**	30	< 5%	>95
					T5	23		>95
					T6	29		80
					T7	8		>95

Entry	TFO (eq.)	Asc. (eq.)	T (°C)	Time (h)	TFO	Cleavage Target (%)[a]	Cleavage Off-Target (%)[b]	Selectivity (%)[c]
					T1	6	-	-
					T2	4	-	-
					T3	7	-	-
7-A	0.2	1000	37	24	**T4**	21	6	78
					T5	17	4	81
					T6	4	-	-
					T7	2	-	-
					T1	-	-	-
					T2	9	9	-
					T3	7	7	-
7-B	0.2	1000	37	48	**T4**	6	6	-
					T5	3	4	-
					T6	5	5	-
					T7	3	3	-
					T1	6	2	-
					T2	5	-	-
					T3	6	4	-
7-C	0.2	1000	37	48	**T4**	3	2	-
					T5	5	3	-
					T6	11	3	95
					T7	13	7	92

Entry	TFO (eq.)	Asc. (eq.)	T (°C)	Time (h)	TFO	Cleavage Target (%)[a]	Cleavage Off-Target (%)[b]	Selectivity (%)[c]
					T1	29	23	75
					T2	25	17	81
					T3	37	24	72
7-D	0.2	1000	37	48	**T4**	25	22	77
					T5	30	27	72
					T6	30	21	76
					T7	18	12	87
					T1	33	12	73
					T2	31	10	76
					T3	22	9	71
8-A	0.2	5000	25	12	**T4**	20	7	74
					T5	41	13	76
					T6	26	11	70
					T7	23	14	62
					T1	32	20	62
					T2	30	16	65
					T3	23	6	79
8-B	0.2	5000	25	12	**T4**	20	8	71
					T5	41	11	79
					T6	25	7	78
					T7	38	12	76

Entry	TFO (eq.)	Asc. (eq.)	T (°C)	Time (h)	TFO	Cleavage Target (%)[a]	Cleavage Off-Target (%)[b]	Selectivity (%)[c]
					T1	60	53	53
					T2	68	47	59
					T3	59	42	58
9-A	0.2	5000	37	12	**T4**	68	52	57
					T5	64	44	59
					T6	66	50	57
					T7	54	45	55
					T1	51	48	51
					T2	54	40	57
					T3	52	39	57
9-B	0.2	5000	37	12	**T4**	64	55	53
					T5	65	51	56
					T6	58	48	54
					T7	55	43	56
					T1			
					T2			
					T3			
10-A	5	1000	25	24	**T4**	High cleavage	High cleavage	n.a.
					T5			
					T6			
					T7			

Entry	TFO (eq.)	Asc. (eq.)	T (°C)	Time (h)	TFO	Cleavage Target (%)[a]	Cleavage Off-Target (%)[b]	Selectivity (%)[c]
					T1			
					T2			
					T3			
10-B	5	1000	37	24	**T4**	High cleavage	High cleavage	n.a.
					T5			
					T6			
					T7			
					A1			
					A2			
					A3			
11-A	5	1000	37	24	**A4**	Not cleaved	Not cleaved	n.a.
					A5			
					A6			
					A7			
					C1			
					C2			
					C3			
11-B	5	1000	37	24	**C4**	Not cleaved	Not leaved	n.a.
					C5			
					C6			
					C7			

a) dsDNA cleavage occurred at the target duplex, calculated on the base of the spot intensity on gel, with respect to the intensity of the reference spot. b) dsDNA cleavage occurred at the off-target duplex, with respect to the intensity of the reference spot. c) calculated dividing the cleavage at the target duplex by the total cleavage (target + off-target duplexes):

$$Selectivity = 100 \times \frac{Cleavage(target)}{Cleavage(target) + Cleavage\ (off-target)}.$$

Appendix 1. Copies of selected PAGE and agarose gels

Enzymatic incorporation of dU^{peg}TP and dU^{un}TP.

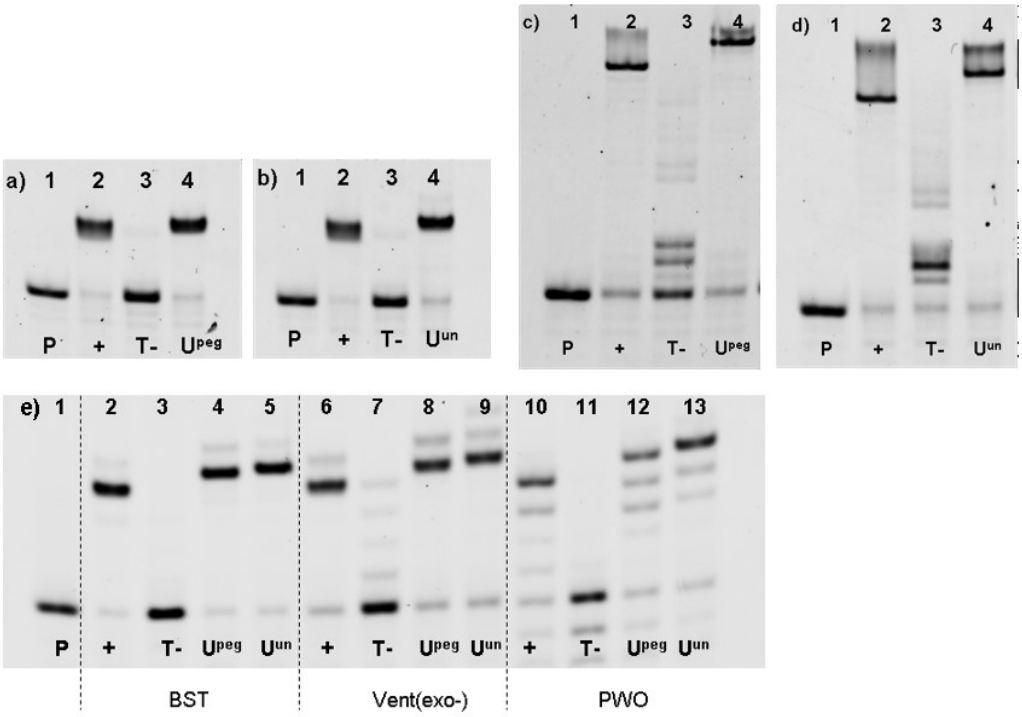


Figure 30. Denaturing PAGE analysis of PEX experiment using KOD XL DNA Polymerase, Prim^{PEX}-FAM and Temp1^{PEX} (in **a**, **b**) or Temp4^{PEX} (in **c**, **d**) the presence of **a**, **c**) **dU^{peg}TP**, and **b**, **d**) **dU^{un}TP**. Lane 1, **P**: Primer^{PEX}-FAM; Lane 2, **+**: product of PEX with natural **dTTP**; Lane 3, **T-**: product of PEX in the absence of **dTTP** and **dU^{R}TP**; Lane 4, **U^{R}**: product of PEX in the absence of **dTTP** and presence of **dU^{peg}TP**, or **dU^{un}TP**; e) PAGE analysis of PEX experiments usinf Bst, Vent(exo-) and Pwo DNA polymerases and the two 5'-*O*-triphospahates **dU^{peg}TP** and **dU^{un}TP**.

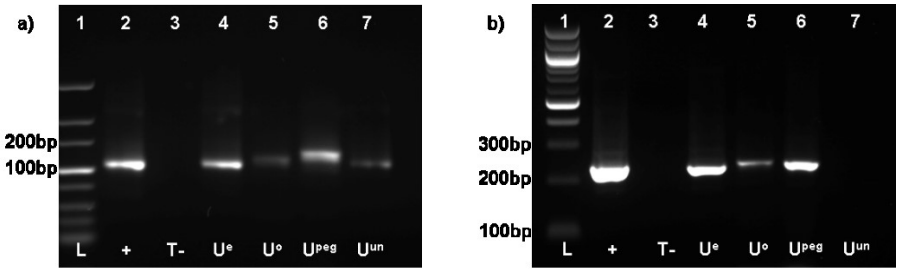


***Figure 31*.** Agarose gel analysis of PCR using KOD XL DNA Polymerase and **TempPCR-98**, **PrimFOR-98** and **PrimREV-98** (in **a**) or **TempPCR-235**, **PrimFOR-235** and **PrimREV-235** (in **b**) and PrimREV-98. Lane 1, **L**: DNA ladder; Lane 2, **+**: product of PCR with natural **dTTP**; Lane 3, **T-**: product of PCR in the absence of **dTTP** and **dU^{R}TP**; Lanes 4 - 7, **U^{R}**: product of PCR in the absence of **dTTP** and presence of **dU^{R}TP**.

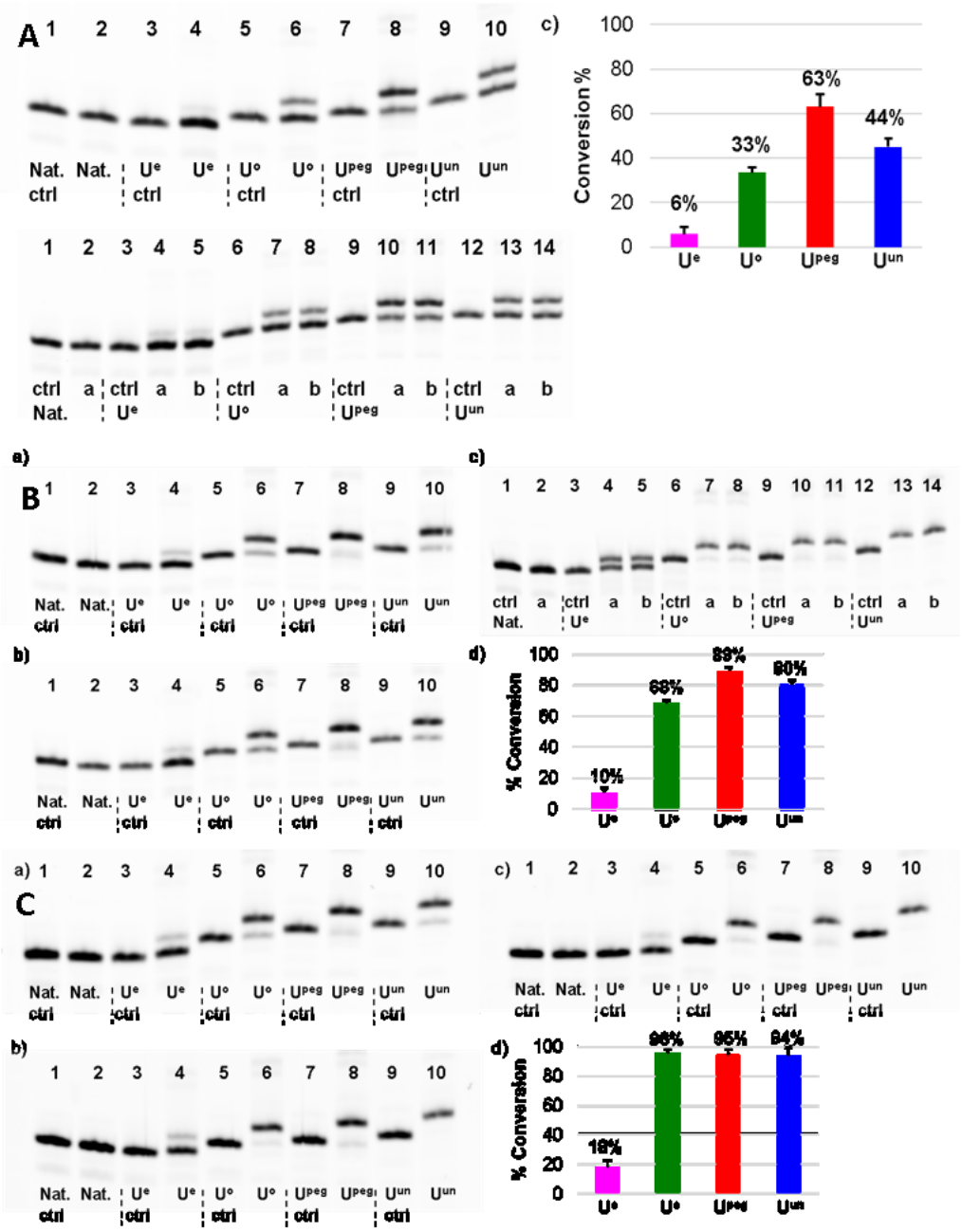


Figure 32. PAGE analysis of CuAAC performed between **DNA_1U^{R}** and PEG-N_3 for 15 min (**A**), 30 min (**B**) or 1 h (**C**) at 37°C. Lanes **Nat. ctrl**: natural DNA control before CuAAC. Lane **Nat.**: natural DNA treated in the CuAAC conditions; Lanes **U^{e} ctrl**: **DNA_dU^{e}**; Lanes **U^{e}**: CuAAC on **DNA_dU^{e}**; Lanes **U^{o} ctrl DNA_dU^{o}**; Lanes **U^{o}**: CuAAC on **DNA_dU^{o}**; Lanes **U^{peg} ctrl**: **DNA_dU^{peg}**; Lanes **U^{peg}**: CuAAC on **DNA_dU^{peg}**; Lanes **U^{un} ctrl**: **DNA_dU^{un}**; Lanes **U^{un}**: CuAAC on **DNA_dU^{un}**. Graphs are showing the percent of conversion of the CuAAC obtained on the basis of the spots intensity.

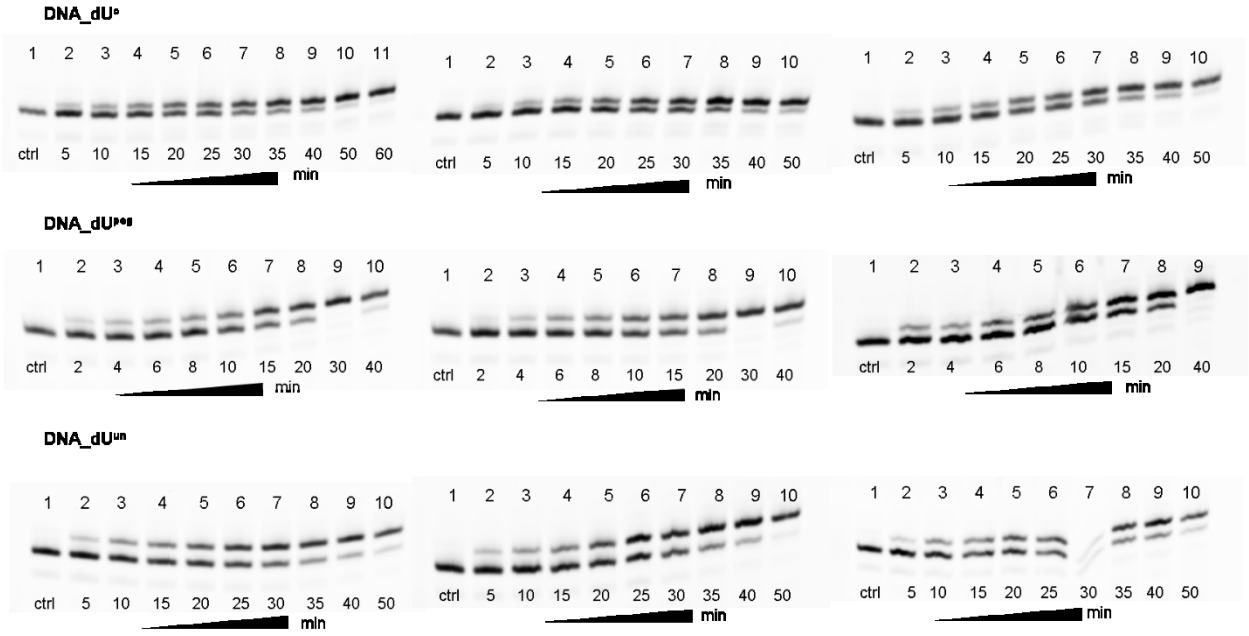


Figure 33. PAGE analysis of simplified reaction kinetics of CuAAC. CuAAC was performed between **DNA_dU°**, **DNA_dU^{peg}** or **DNA_dU^{un}**, and **PEG-N$_3$**.

Selected examples of PAGE analysis of DNA cleavage experiments

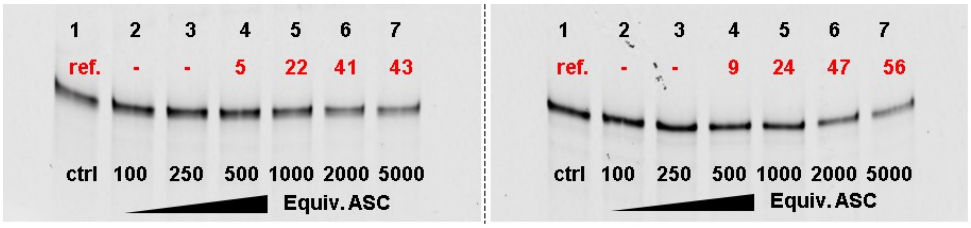


Figure 34. Denaturing PAGE analysis of cleavage experiment employing 5 equivalents of TFO **T4** (**TFO-32-dU^{pegCP}**), with increasing amount of Na-$_L$-ascorbate (from 100 equivalents to 5000 equivalents). Reaction mixtures were incubated for 12 h at 37 °C. The dsDNA cleavage percentage is calculated (based on the spot intensity) with respect to the control (**ctrl**) in which no ascorbate was added, and shown in red on the gel.

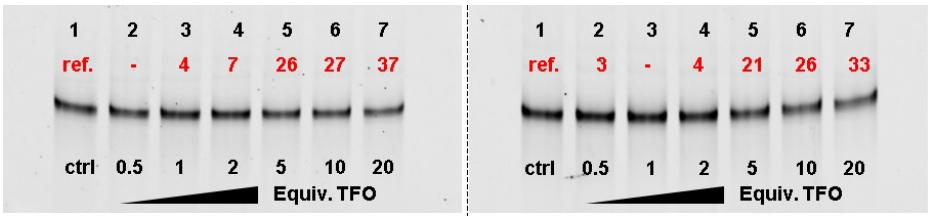


Figure 35. Denaturing PAGE analysis of cleavage experiment employing increasing equivalents of equivalents of TFO **T4** (**TFO-32-dU^{pegCP}**, from 0.5 to 20 equivalents), and 1000 equivalents of Na-$_L$-ascorbate. Reaction mixtures were incubated for 12 h at 37 °C. The dsDNA cleavage percentage is calculated (based on the spot intensity) with respect to the control (**ctrl**) in which no TFO was added, and shown in red on the gel.

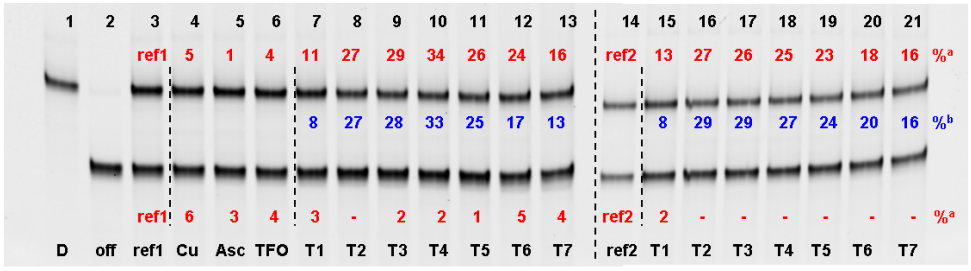


Figure 36. Example of denaturing PAGE analysis of cleavage experiment. In this experiment were employed 5 equiv. of TFOs **T1-T7**, and the mixture was incubated for 12 h at 25 °C with 1000 equiv. of Na-$_L$-ascorbate. Lane 1 (**D**): Target duplex. Lane 2 (**off**): off-target duplex. Lanes 3 and 14 (**ref**): 1:1 mixture of target duplex and off-target duplex. Lane 4 (**Cu**): control of the cleaving ability of Clip-phenanthroline-modified TFO **C2** in the absence of Cu and presence of ascorbate. Lane 5 (**Asc**): control of cleavage by the Cu-Clip-phenanthroline-modified TFO **T2** in the absence of Na-$_L$-ascorbate. Lane 6 (**TFO**): control of cleavage in the presence of $CuCl_2$ (1 µL, 10 µM) and Na-$_L$-ascorbate (1000 equiv.). Lanes 7-13 and 15-21 (**T1-T7**): cleavage experiments employing Cu-Clip-phenanthroline-modified TFOs **T1-T7**. The dsDNA cleavage (a) is shown in red; selectivity (b) is shown in blue.

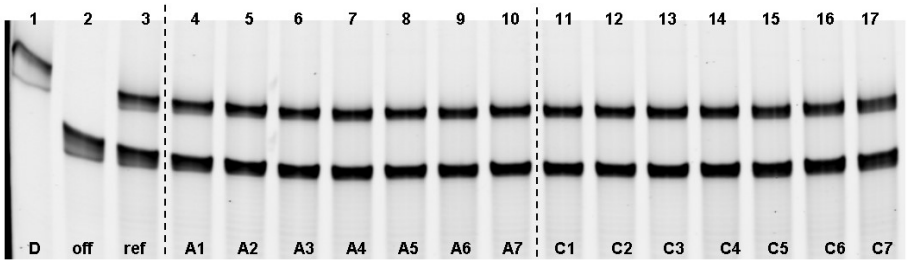


Figure 37. Denaturing PAGE analysis of dsDNA cleavage control experiment. Lane 1 (**D**): Target duplex. Lane 2 (**off**): off-target duplex. Lanes 3 and 14 (**ref**): 1:1 mixture of target duplex and off-target duplex. Lanes 4-10 (**A1-A7**): a 1:1 mixture of target duplex and off-target duplex was incubated at 37 °C for 12 h with 5 equiv. of alkyne-modified TFOs **A1-A7**, 5 equiv. of $CuCl_2$ and 1000 equiv. of Na-$_L$-ascorbate. Lanes 11-17 (**C1-C7**): a 1:1 mixture of target duplex and off-target duplex was incubated at 37 °C for 12 h with 5 equiv. of Clip-phenanthroline-modified TFOs **C1-C7** and 1000 equiv. of Na-$_L$-ascorbate, in the absence of $CuCl_2$.

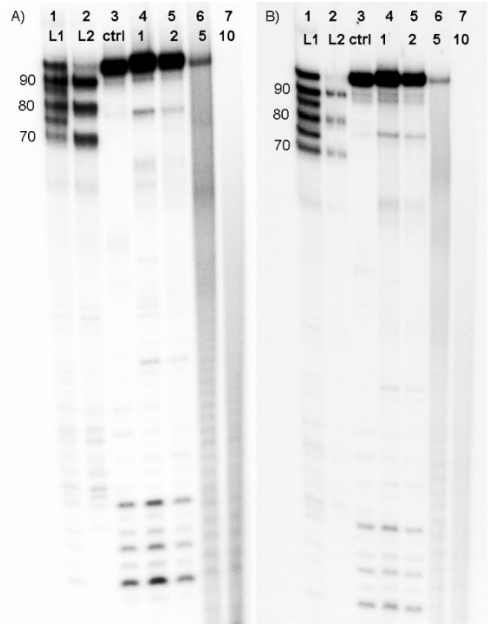


Figure 38. Non-denaturing PAGE analysis of cleavage experiments of the radioactively labelled target duplex obtained by PCR. Lane 1 (**L1**): DNA ladder containing 98, 90, 85, 80, 75 and 70 bp PCR products. Lane 2 (**L2**): DNA ladder containing 90, 80 and 70 bp PCR products. Lane 3 (**ctrl**): Target duplex. Lanes 4 to 7 (**1**, **2**, **5**, **10**): cleavage experiments employing 1, 2, 5 or 10 equiv. of **TFO-32-5'pegCP** (**T2**) and 1000 equiv. of Na-$_L$-ascorbate, and incubated at 25 °C for 12 h.

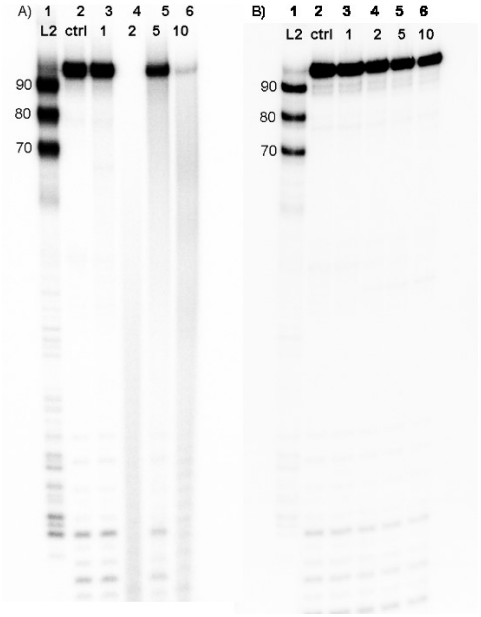


Figure 39. Non-denaturing PAGE analysis of cleavage experiments of the radioactively labelled target duplex obtained by PCR. Lane 1 (**L2**): DNA ladder containing 90, 80 and 70 bp PCR products. Lane 2 (**ctrl**): Target duplex. Lanes 3 to 6 (**1**, **2**, **5**, **10**): cleavage experiments employing 1, 2, 5 or 10 equiv. of **TFO-32-dU^{pegCP}** (**T4**, in gel A) or **TFO-32-3'alkCP** (**T7**, in gel B) and 1000 equiv. of Na-$_L$-ascorbate, and incubated at 25 °C for 12 h.

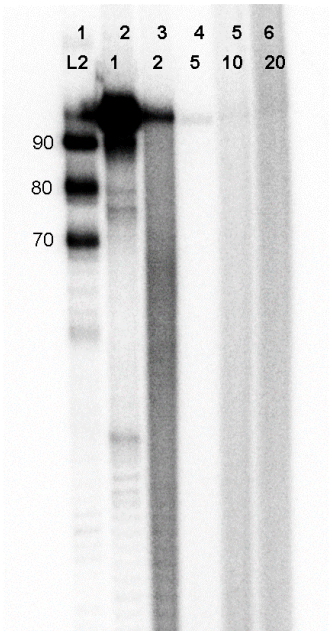


Figure 40. Non-denaturing PAGE analysis of cleavage experiments of the radioactively labelled target duplex obtained by PCR. Lane 1 (**L2**): DNA ladder containing 90, 80 and 70 bp PCR products. Lanes 2 to 6 (**1**, **2**, **5**, **10**, **20**): cleavage experiments employing 1, 2, 5 or 10 equiv. of **TFO-32-5'pegCP** (**T2**) and 1000 (**1**), 2000 (**2**), 5000 (**5**), 10000 (**10**) or 20000 (**20**) equiv. of Na-$_L$-ascorbate, and incubated at 25 °C for 12 h.

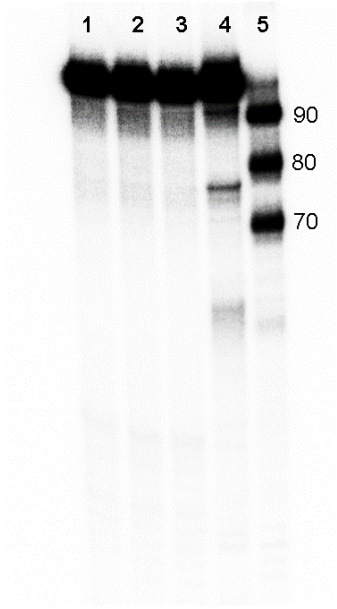


Figure 41. Non-denaturing PAGE analysis of cleavage experiments of the radioactively labelled target duplex obtained by PCR. Lane 1: Target duplex. Lane 2: Target duplex incubated in the presence of Cu-ClipPhen-TFO and absence of Na-$_L$-ascorbate. Lane 3: Target duplex incubated in the presence of Na-$_L$-ascorbate and in the absence of TFO. Lane 4: cleavage experiment employing 1 equiv. of **TFO-32-5'pegCP** (**T2**) and 1000 equiv. of Na-$_L$-ascorbate, and incubated at 25 °C for 12 h. Lane 5: DNA ladder containing 90, 80 and 70 bp PCR products.

Competitive PEX experiments.

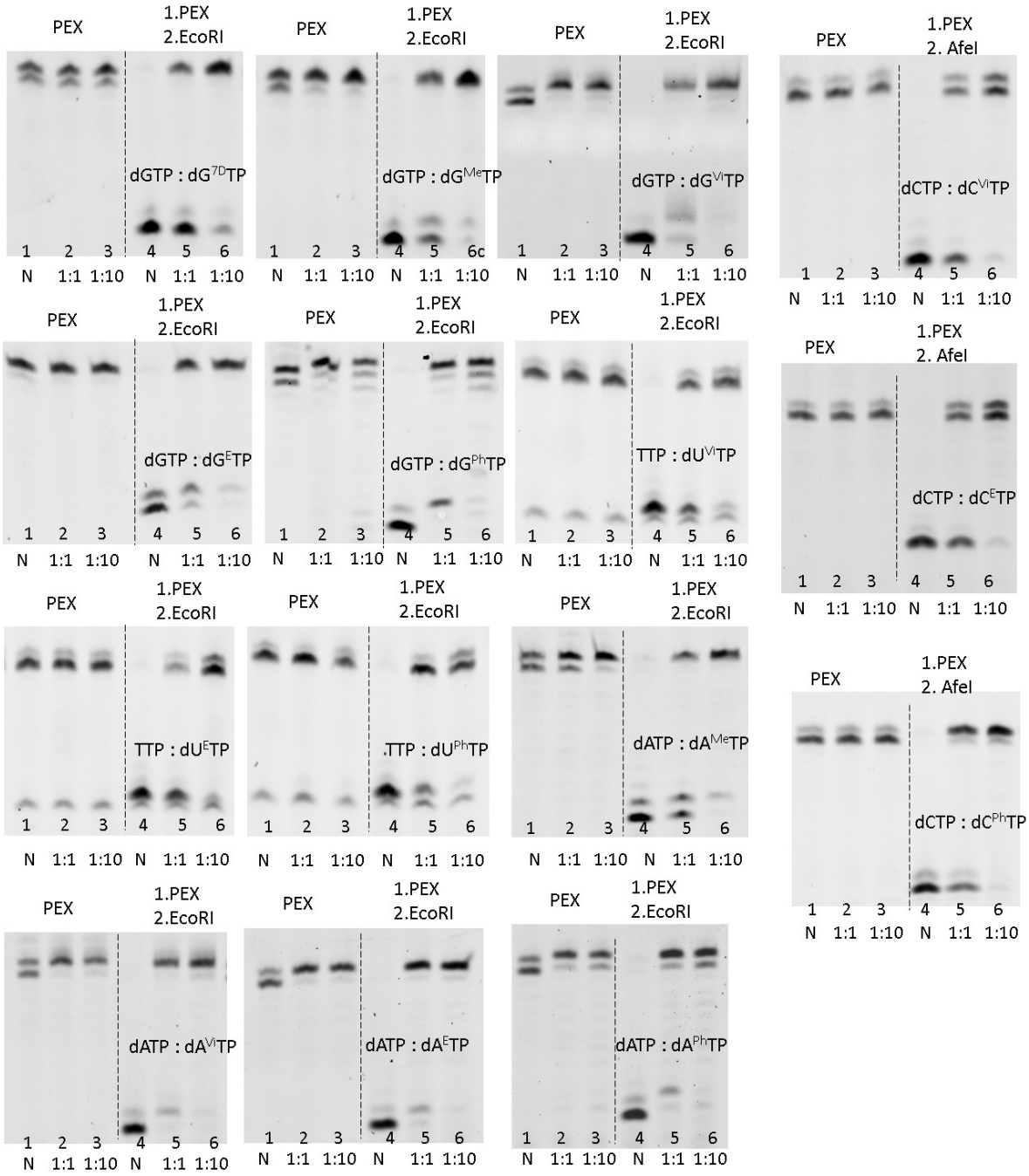


Figure 42. PAGE analyses of PEX experiments with Bst DNA polymerase and cleavage products with dN^{R}TPs. Lanes 1, **N**: product of PEX with natural dNTPs; lanes 2 and 3, **1/1** and **1/10**: product of PEX with 3 natural dNTPs and corresponding ratios of dNTP/dN^{R}TP; lanes 4, **N**: product of cleavage of unmodified DNA by RE (EcoRI or AfeI); lanes 5 and 6, **1/1** and **1/10**: products of cleavage of modified DNAs by RE.

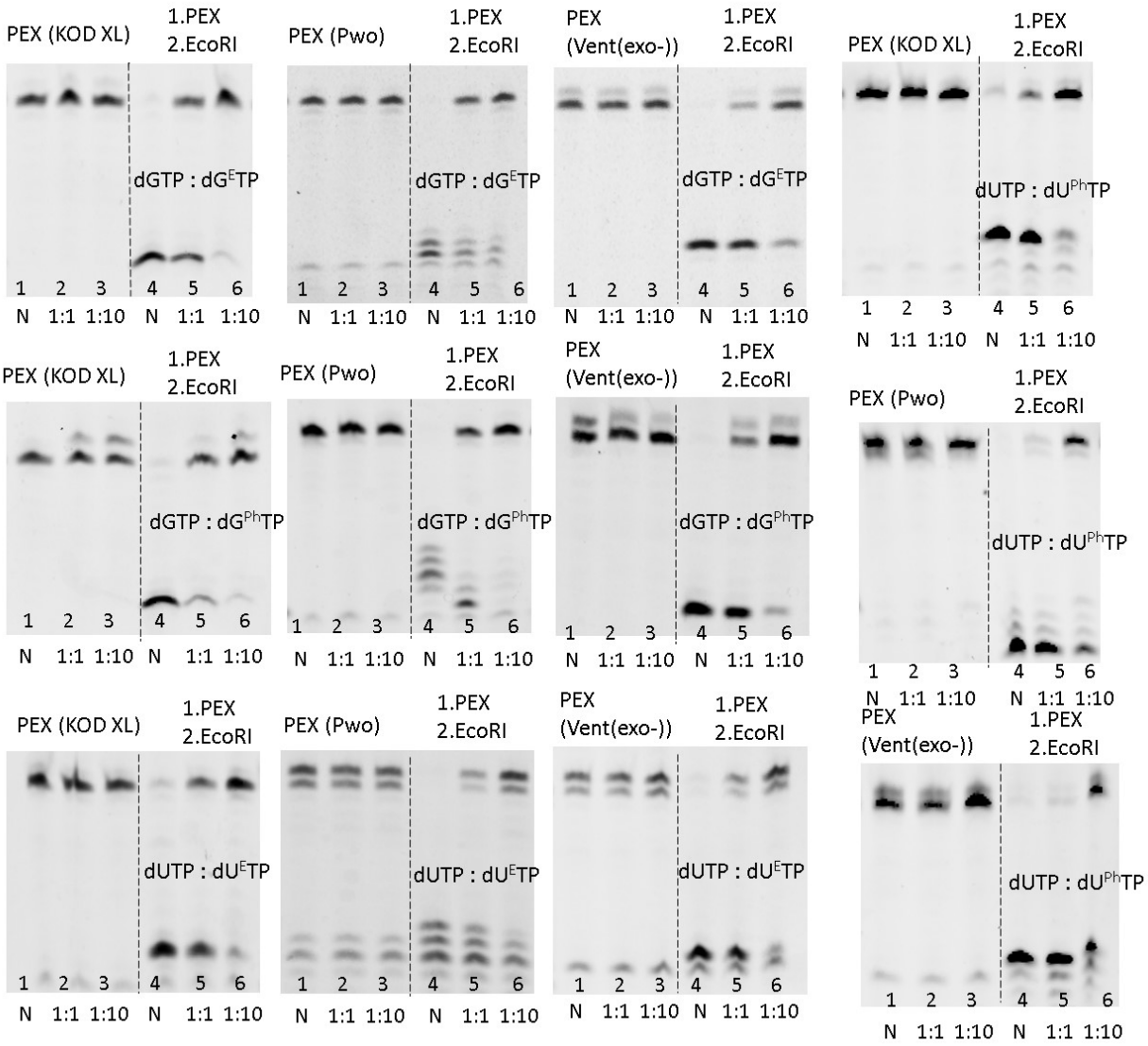


Figure 43. PAGE analyses of PEX experiments with 3 different DNA polymerases (KOD XL, Pwo and Vent(exo-)) and cleavage products with **dG^{ph}TP**, **dG^{e}TP**, **dU^{ph}TP** and **dU^{e}TP** Lanes 1, **N**: product of PEX with natural dNTPs; lanes 2 and 3, **1/1** and **1/10**: product of PEX with 3 natural dNTPs and corresponding ration of dNTP/dN^{R}TP; lanes 4, **N**: product of cleavage of unmodified DNA by EcoRI; lanes 5 and 6, **1/1** and **1/10**: products of cleavage of modified DNAs by RE.

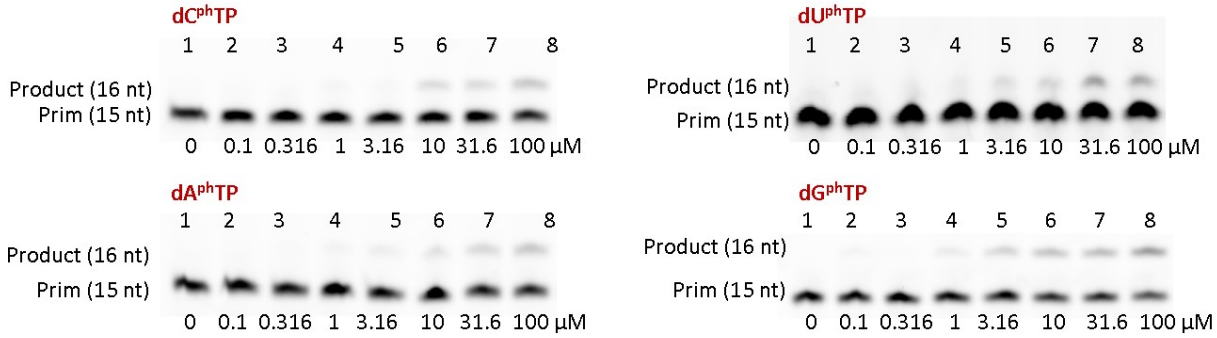


Figure 44. Examples of PAGE analysis of steady-state kinetics of incorporation of modified dN^{R}TPs.

Appendix 2. Simplified kinetics of CuAAC on ONs

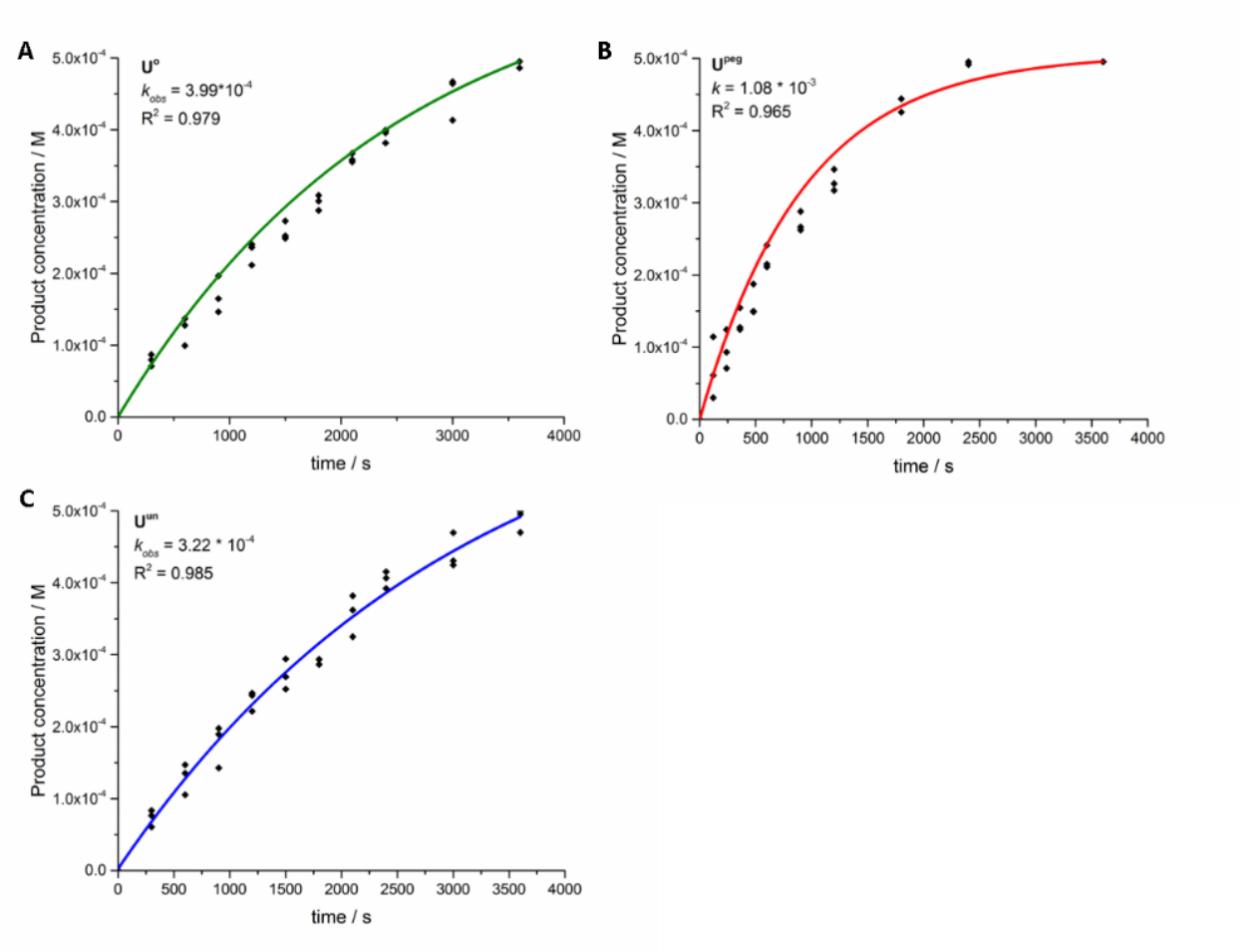


Figure 45. Graphs of CuAAC product concentration plotted against time. CuAAC was performed between **DNA_dU°** (graph **A**), **DNA_dU^{peg}** (graph **B**) or **DNA_dU^{un}** (graph **C**), and **PEG-N$_3$** and analyzed through PAGE. Data were fitted in the curve $y = A_1e^{-k} + y_0$ to obtain the observed reaction rate constant (k_{obs}).

Appendix 3. Copies of selected MALDI spectra

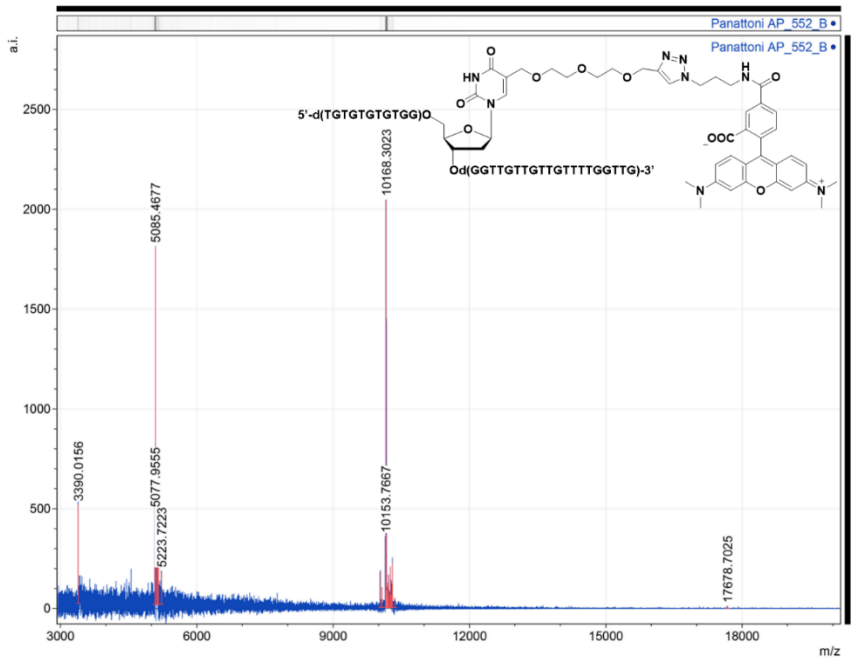


Figure 46. MALDI-TOF MS spectrum of **ON_U$^{peg\text{-}TAMRA}$** after CuAAC reaction of **ON_U^{peg}** with 5-TAMRA-N_3. Calculated for $[M+H]^+$: 10167.9 Da; found: 10168.3 Da.

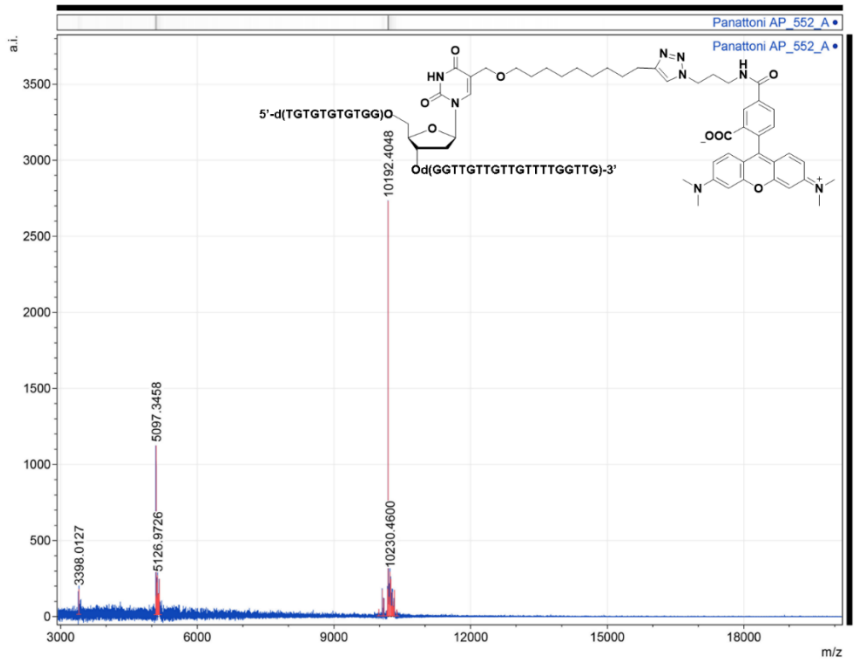


Figure 47. MALDI-TOF MS spectrum of **ON_U$^{un\text{-}TAMRA}$** after CuAAC reaction of **ON_U^{un}** with 5-TAMRA-N_3. Calculated for $[M+H]^+$: 10192.1 Da; found: 10192.4 Da.

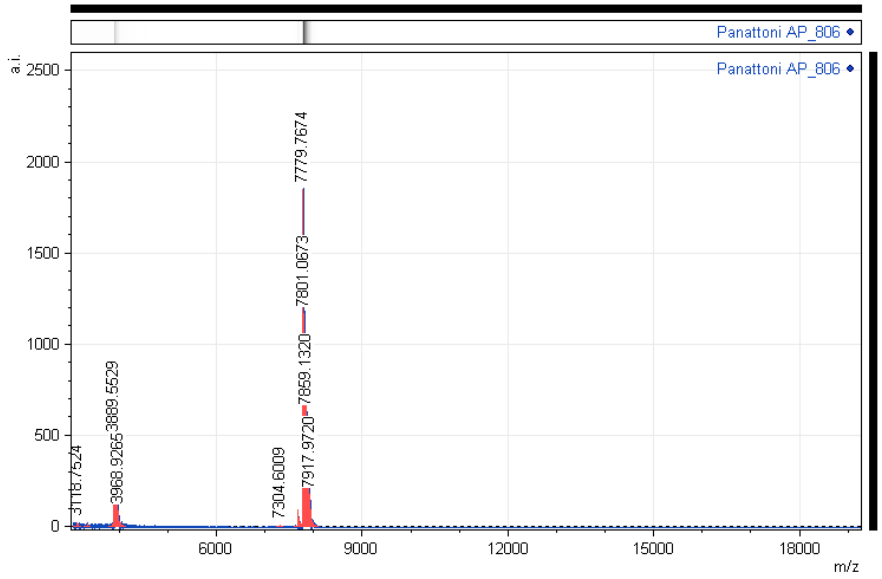


Figure 48. **TFO-24-5'pegCP.** Oligonucleotide resulted from CuAAC between **TFO-24-5'peg** and **N_3CP**. Calculated for $[M+H]^+$: 7778.9 KDa; Found: 7779.8 KDa.

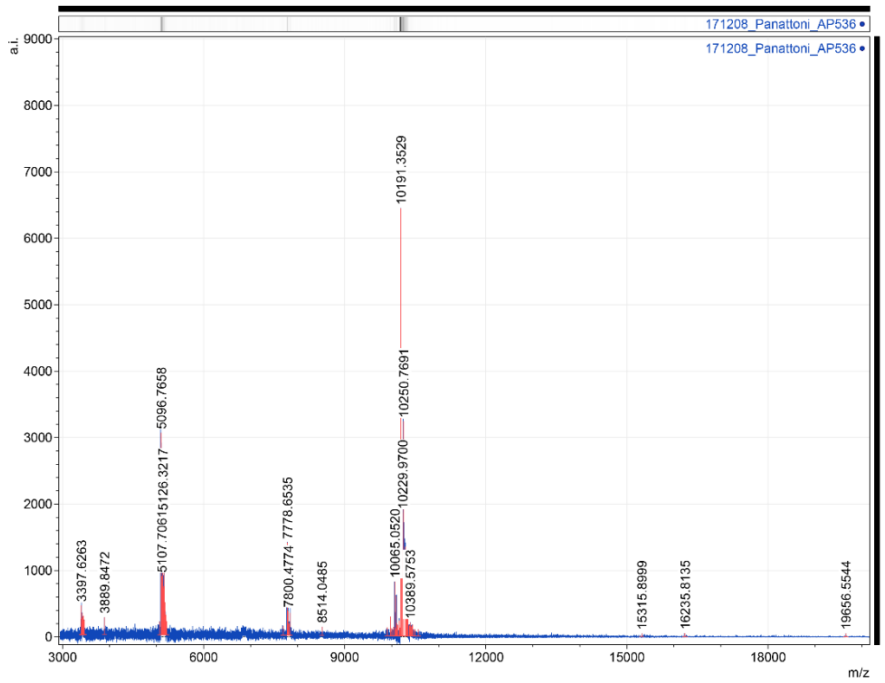


Figure 49. **TFO-32-5'pegCP.** Oligonucleotide resulted from CuAAC between **TFO-32-5'peg** and **N_3CP**. Calculated for $[M+H]^+$: 10191.8 KDa; Found: 10191.4 KDa.

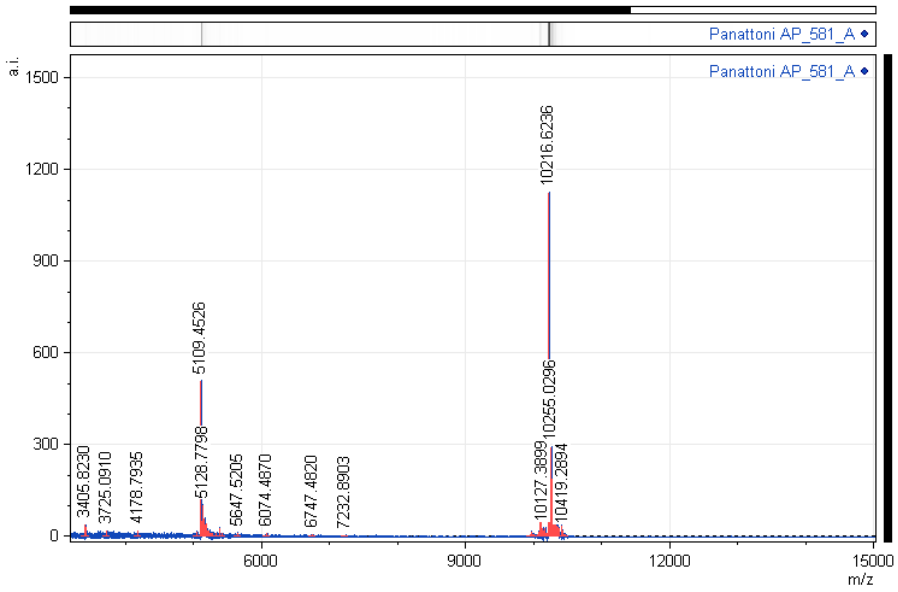


Figure 50. **TFO-32-5'unCP.** Oligonucleotide resulted from CuAAC between **TFO-32-5'un** and **N_3CP**. Calculated for $[M+H]^+$: 10216.9 KDa; Found: 10216.6 KDa

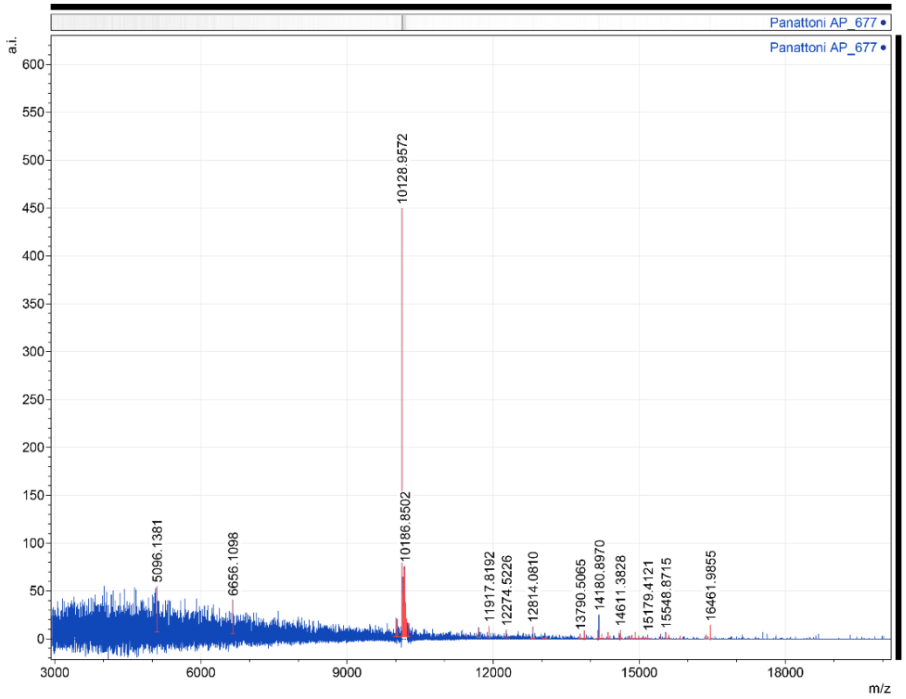


Figure 51. **TFO-32-dU^{pegCP}.** Oligonucleotide resulted from CuAAC between **TFO-32-dU^{peg}** and **N_3CP**. Calculated for $[M+H]^+$: 10128.9 KDa; Found: 10129.0 KDa.

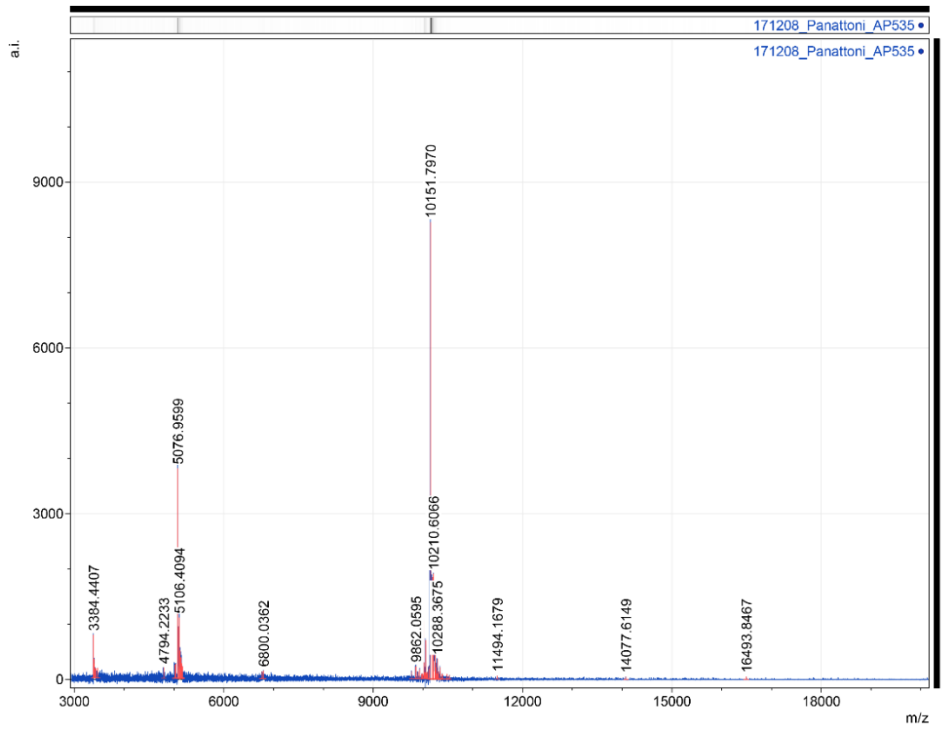


Figure 52. **TFO-32-dU^{unCP}.** Oligonucleotide resulted from CuAAC between **TFO-32-dU^{un}** and **N_3CP**. Calculated for $[M+H]^+$: 10150.8 KDa; Found: 10151.8 KDa.

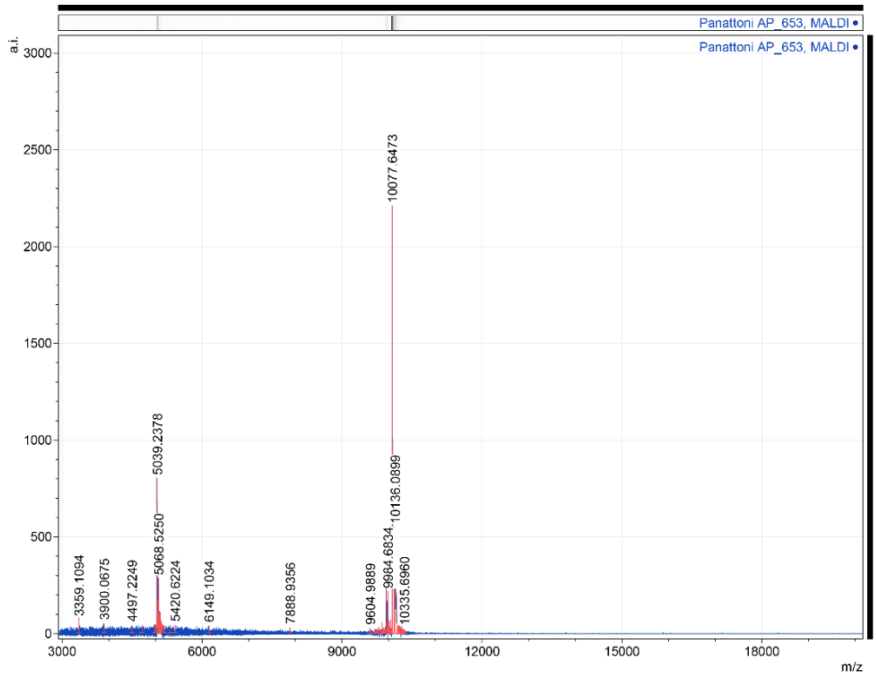


Figure 53. **TFO-32-dU^{octCP}.** Oligonucleotide resulted from CuAAC between **TFO-32-dU^{oct}** and **N_3CP**. Calculated for $[M+H]^+$: 10076.8 KDa; Found: 10077.6 KDa.

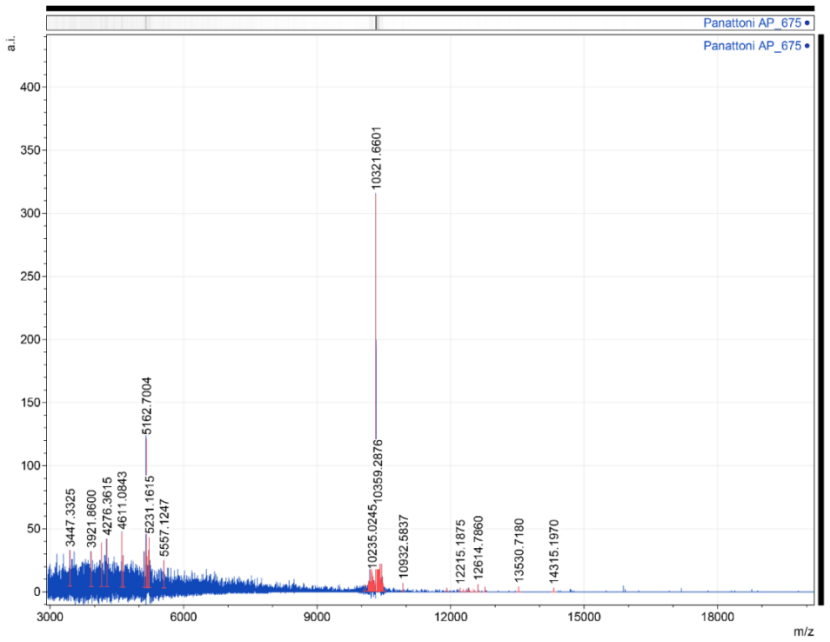


Figure 54. **TFO-32-3'alkCP.** Oligonucleotide resulted from CuAAC between **TFO-32-3'alk** and **N_3CP**. Calculated for $[M+H]^+$: 10320.1 KDa; Found: 10321.7 KDa.

Appendix 4. Steady-State fluorescence measurements.

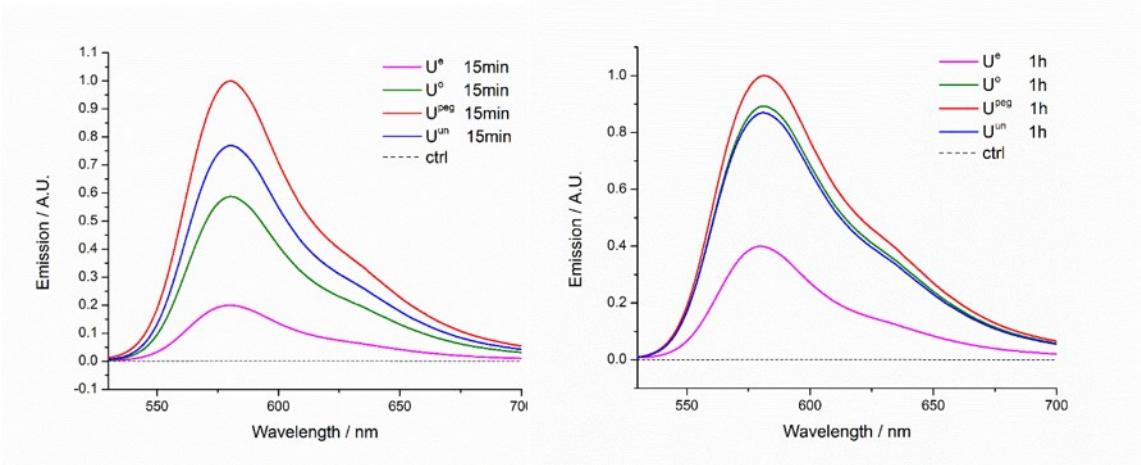


Figure 55. CuAAC carried out for 15 minutes or 1 hour between TAMRA-N_3 and **ON_1U^{e}** (**U^{e}**, magenta line), **ON_1U^{o}** (**U^{o}**, green line), **ON_1U^{peg}** (**U^{peg}**, red line) and **ON_1U^{un}** (**U^{un}**, blue line). The control line (**ctrl**, black dashed line) is obtained employing natural **TTP** in the PEX step.

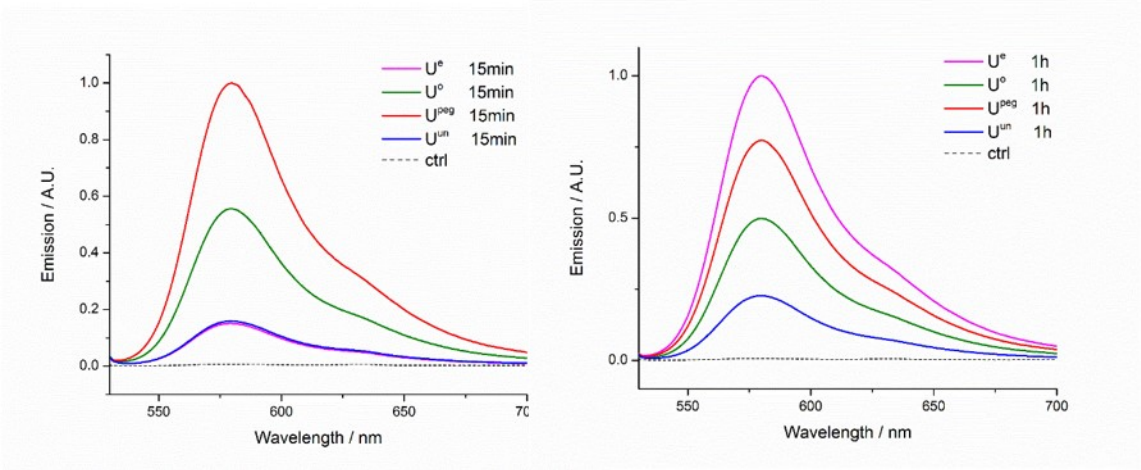


Figure 56. CuAAC carried out for 15 minutes or 1 hour between TAMRA-N_3 and 98bp modified DNA obtained by PCR with **dU^{e}TP** (**U^{e}**, magenta line), **dU^{o}TP** (**U^{o}**, green line), **dU^{peg}TP** (**U^{peg}**, red line) and **dU^{un}TP** (**U^{un}**, blue line). The control line (**ctrl**, black dashed line) is obtained employing natural **TTP** in the PEX step.

Appendix 5. UV triplex melting curves

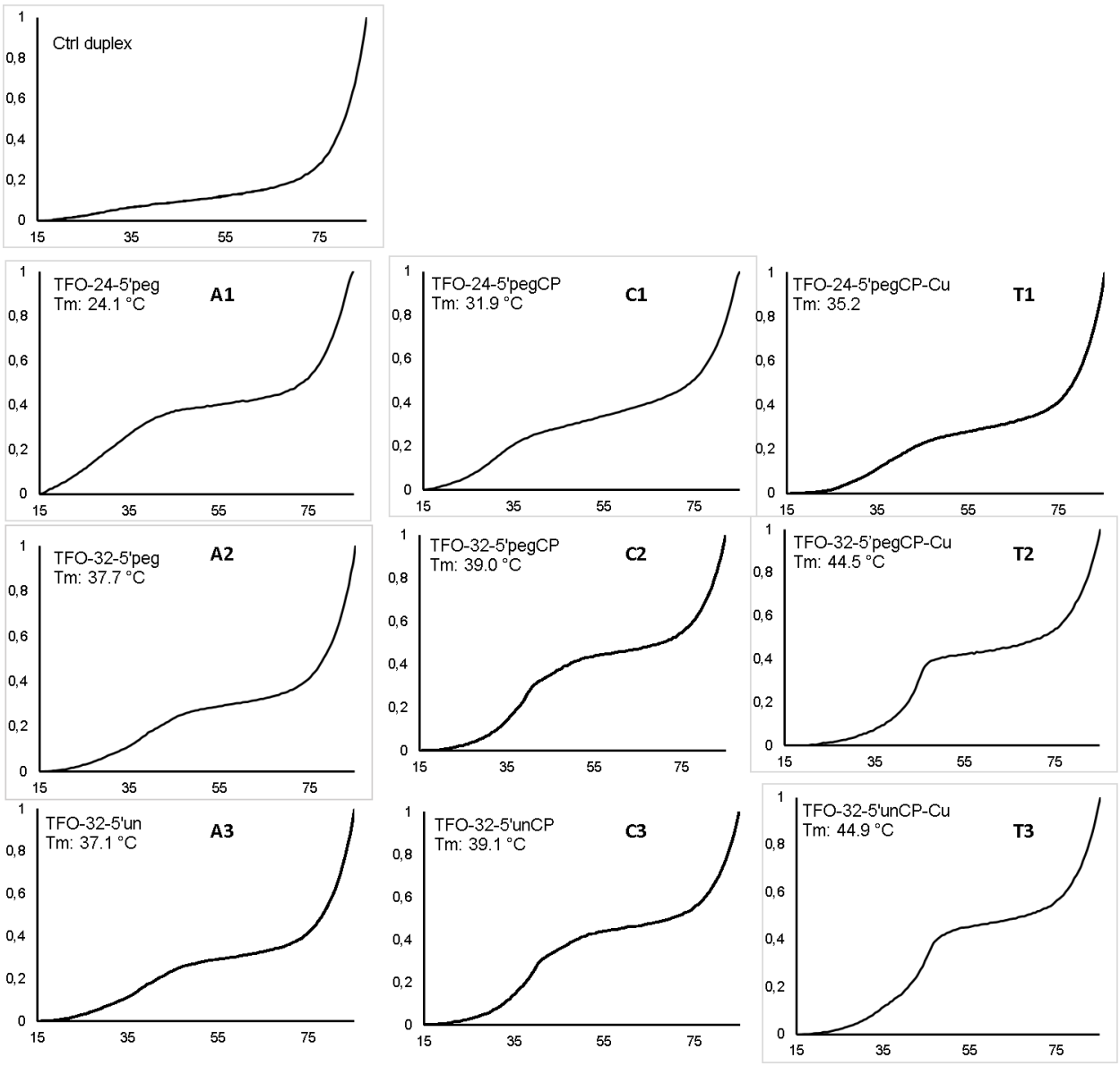


Figure 57. UV triplex melting curves of TFOs **A1-3**, **C1-3** and **T1-3**

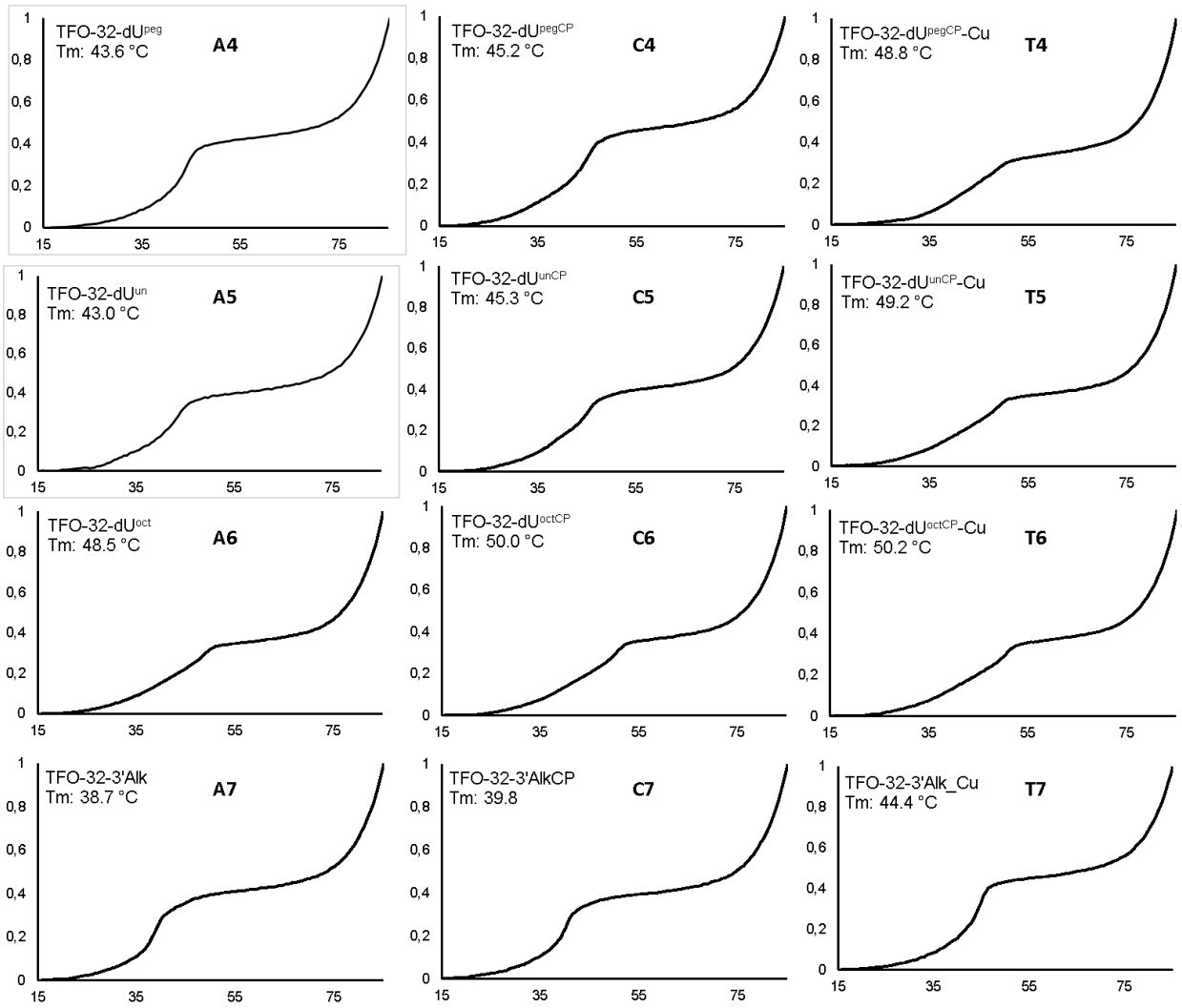


Figure 58. UV triplex melting curves of TFOs **A4-7**, **C4-7** and **T4-7**

List of abbreviations

ACN: acetonitrile

AMN: artificial metallonuclease

CEP: 3'-*O*-cyanoethyl phosphoramidite

CPG: controlled-pore glass

CuAAC: copper catalyzed alkyne-azide 1,3-dipolar cycloaddition

DCM: dichloromethane

DIPEA: *N*,*N*-diisopropylethylamine

DMF: dimethyl formamide

dNTP: 2'-deoxyribonucleotide triphosphate

dsDNA: double-stranded DNA

HR: homologous recombination

k_{cat}: Enzyme turnover number

K_M: Michaelis-Menten constant

NEAR: nicking enzyme amplification reaction

NHJR: non-homologous end joining

nt: nucleotide

ON: oligonucleotide

PCR: polymerase chain reaction

PEX: primer extension

Py: pyridine

RE: restriction endonuclease

RM: restriction-modification

ssDNA: single-stranded DNA

TEAB: tetraethyl ammonium bicarbonate

TFO: triplex-forming oligonucleotide

THF: tetrahydrofuran

T_M: melting temperature

6. References

1 J. D. WATSON and F. H. C. CRICK, *Nature*, 1953, **171**, 737–738.

2 R. E. FRANKLIN and R. G. GOSLING, *Nature*, 1953, **171**, 740–741.

3 M. H. F. WILKINS, A. R. STOKES and H. R. WILSON, *Nature*, 1953, **171**, 738–740.

4 R. Saiki, D. Gelfand, S. Stoffel, S. Scharf, R. Higuchi, G. Horn, K. Mullis and H. Erlich, *Science*, 1988, **239**, 487–491.

5 J. Van Ness, L. K. Van Ness and D. J. Galas, *Proc. Natl. Acad. Sci.*, 2003, **100**, 4504–4509.

6 E. A. Galburt and B. L. Stoddard, *Biochemistry*, 2002, **41**, 13851–13860.

7 C. M. Dupureur, *Biochemistry*, 2005, **44**, 5065–5074.

8 J. J. Perona, *Methods*, 2002, **28**, 353–364.

9 A. Pingoud, M. Fuxreiter, V. Pingoud and W. Wende, *Cell. Mol. Life Sci.*, 2005, **62**, 685–707.

10 A. Pingoud, *Nucleic Acids Res.*, 2001, **29**, 3705–3727.

11 L. Jen-Jacobson, L. E. Engler, D. R. Lesser, M. R. Kurpiewski, C. Yee and B. McVerry, *EMBO J.*, 1996, **15**, 2870–2882.

12 H. Macíčková-Cahová, R. Pohl and M. Hocek, *ChemBioChem*, 2011, **12**, 431–438.

13 M. Mačková, S. Boháčová, P. Perlíková, L. Poštová Slavětínská and M. Hocek, *ChemBioChem*, 2015, **16**, 2225–2236.

14 H. Macíčková-Cahová and M. Hocek, *Nucleic Acids Res.*, 2009, **37**, 7612–7622.

15 M. Mačková, R. Pohl and M. Hocek, *ChemBioChem*, 2014, **15**, 2306–2312.

16 S. L. Beaucage and M. H. Caruthers, *Tetrahedron Lett.*, 1981, **22**, 1859–1862.

17 M. Famulok, J. S. Hartig and G. Mayer, *Chem. Rev.*, 2007, **107**, 3715–3743.

18 G. Mayer, *Angew. Chemie Int. Ed.*, 2009, **48**, 2672–2689.

19 T. Tørring, N. V. Voigt, J. Nangreave, H. Yan and K. V. Gothelf, *Chem. Soc. Rev.*, 2011, **40**, 5636–5646.

20 J. Matyašovský, R. Pohl and M. Hocek, *Chem. - A Eur. J.*, 2018, **24**, 14938–14941.

21 J. Matyašovský, P. Perlíková, V. Malnuit, R. Pohl and M. Hocek, *Angew. Chemie Int. Ed.*, 2016, **55**, 15856–15859.

22 M. Hocek, *J. Org. Chem.*, 2014, **79**, 9914–9921.

23 J. A. Prescher and C. R. Bertozzi, *Nat. Chem. Biol.*, 2005, **1**, 13–21.

24 H. C. Kolb, M. G. Finn and K. B. Sharpless, *Angew. Chemie Int. Ed.*, 2001, **40**, 2004–2021.

25 M. Meldal and C. W. Tornøe, *Chem. Rev.*, 2008, **108**, 2952–3015.

26 E. Haldón, M. C. Nicasio and P. J. Pérez, *Org. Biomol. Chem.*, 2015, **13**, 9528–9550.

27 D. Graham, J. A. Parkinson and T. Brown, *J. Chem. Soc. Perkin Trans. 1*, 1998, 1131–1138.

28 P. Ding, D. Wunnicke, H.-J. Steinhoff and F. Seela, *Chem. - A Eur. J.*, 2010, **16**, 14385–14396.

29 P. M. E. Gramlich, C. T. Wirges, A. Manetto and T. Carell, *Angew. Chemie Int. Ed.*, 2008, **47**, 8350–8358.

30 P. M. E. Gramlich, C. T. Wirges, J. Gierlich and T. Carell, *Org. Lett.*, 2008, **10**, 249–251.

31 F. Seela and X. Ming, *Helv. Chim. Acta*, 2008, **91**, 1181–1200.

32 F. Seela, H. Xiong, P. Leonard and S. Budow, *Org. Biomol. Chem.*, 2009, **7**, 1374.

33 S. S. Pujari, S. A. Ingale and F. Seela, *Bioconjug. Chem.*, 2014, **25**, 1855–1870.

34 J. Gierlich, K. Gutsmiedl, P. M. E. Gramlich, A. Schmidt, G. A. Burley and T. Carell, *Chem. - A Eur. J.*, 2007, **13**, 9486–9494.

35 X. Ren, A. H. El-Sagheer and T. Brown, *Analyst*, 2015, **140**, 2671–2678.

36 J. Balintová, J. Špaček, R. Pohl, M. Brázdová, L. Havran, M. Fojta and M. Hocek, *Chem. Sci.*, 2015, **6**, 575–587.

37 X. Ren, A. H. El-Sagheer and T. Brown, *Nucleic Acids Res.*, 2016, **44**, e79.

38 A. Salic and T. J. Mitchison, *Proc. Natl. Acad. Sci.*, 2008, **105**, 2415–2420.

39 A. B. Neef and N. W. Luedtke, *Proc. Natl. Acad. Sci.*, 2011, **108**, 20404–20409.

40 A. B. Neef, L. Pernot, V. N. Schreier, L. Scapozza and N. W. Luedtke, *Angew. Chemie Int. Ed.*, 2015, **54**, 7911–7914.

41 H. Cahová, A. Panattoni, P. Kielkowski, J. Fanfrlík and M. Hocek, *ACS Chem. Biol.*, 2016, **11**, 3165–3171.

42 S. A. Ingale, H. Mei, P. Leonard and F. Seela, *J. Org. Chem.*, 2013, **78**, 11271–11282.

43 F. Seela, V. R. Sirivolu and P. Chittepu, *Bioconjug. Chem.*, 2008, **19**, 211–224.

44 J. H. Hu, K. M. Davis and D. R. Liu, *Cell Chem. Biol.*, 2016, **23**, 57–73.

45 D. Carroll, *Mol. Ther.*, 2016, **24**, 412–413.

46 D. Carroll, *Annu. Rev. Biochem.*, 2014, **83**, 409–439.

47 Y. G. Kim, J. Cha and S. Chandrasegaran, *Proc. Natl. Acad. Sci.*, 1996, **93**, 1156–1160.

48 D. Carroll, *Genetics*, 2011, **188**, 773–782.

49 C. O. Pabo, E. Peisach and R. A. Grant, *Annu. Rev. Biochem.*, 2001, **70**, 313–340.

50 D. J. Segal, B. Dreier, R. R. Beerli and C. F. Barbas, *Proc. Natl. Acad. Sci.*, 1999, **96**, 2758–2763.

51 B. Dreier, R. P. Fuller, D. J. Segal, C. V. Lund, P. Blancafort, A. Huber, B. Koksch and C. F. Barbas, *J. Biol. Chem.*, 2005, **280**, 35588–35597.

52 B. Gonzalez, L. J. Schwimmer, R. P. Fuller, Y. Ye, L. Asawapornmongkol and C. F. Barbas, *Nat. Protoc.*, 2010, **5**, 791–810.

53 J. K. Joung and J. D. Sander, *Nat. Rev. Mol. Cell Biol.*, 2013, **14**, 49–55.

54 A. N.-S. Mak, P. Bradley, R. A. Cernadas, A. J. Bogdanove and B. L. Stoddard, *Science*, 2012, **335**, 716–719.

55 D. Deng, C. Yan, X. Pan, M. Mahfouz, J. Wang, J.-K. Zhu, Y. Shi and N. Yan, *Science*, 2012, **335**, 720–723.

56 Martin Jinek, Krzysztof Chylinski, Ines Fonfara, Michael Hauer, Jennifer A. Doudna and Emmanuelle Charpentier, *Science*, 2012, **337**, 816–821.

57 D. Carroll, *Mol. Ther.*, 2012, **20**, 1658–1660.

58 A. D. Weinberger and M. S. Gilmore, *Cell*, 2015, **161**, 964–966.

59 A. V. Wright, J. K. Nuñez and J. A. Doudna, *Cell*, 2016, **164**, 29–44.

60 F. Jiang and J. A. Doudna, *Annu. Rev. Biophys.*, 2017, **46**, 505–529.

61 G. J. Knott and J. A. Doudna, *Science*, 2018, **361**, 866–869.

62 D. S. Sigman, D. R. Graham, V. D'Aurora and A. M. Stern, *J. Biol. Chem.*, 1979, **254**, 12269–12272.

63 D. S. Sigman, *Acc. Chem. Res.*, 1986, **19**, 180–186.

64 G. Pratviel, J. Bernadou and B. Meunier, *Angew. Chemie Int. Ed.*, 1995, **34**, 746–769.

65 O. Zelenko, J. Gallagher and D. S. Sigman, *Angew. Chemie Int. Ed.*, 1997, **36**, 2776–2778.

66 T. Oyoshi and H. Sugiyama, *J. Am. Chem. Soc.*, 2000, **122**, 6313–6314.

67 D. S. Sigman, T. W. Bruice, A. Mazumder and C. L. Sutton, *Acc. Chem. Res.*, 1993, **26**, 98–104.

68 C. Wende, C. Lüdtke and N. Kulak, *Eur. J. Inorg. Chem.*, 2014, **2014**, 2597–2612.

69 M. Pitié, B. Donnadieu and B. Meunier, *Inorg. Chem.*, 1998, **37**, 3486–3489.

70 M. Pitié, C. J. Burrows and B. Meunier, *Nucleic Acids Res.*, 2000, **28**, 4856–4864.

71 M. Pitié, C. Boldron, H. Gornitzka, C. Hemmert, B. Donnadieu and B. Meunier, *Eur. J. Inorg. Chem.*, 2003, **2003**, 528–540.

72 B. C. F. Chu and L. E. Orgel, *Proc. Natl. Acad. Sci.*, 1985, **82**, 963–967.

73 M. Boidot-Forget, M. Chassignol, M. Takasugi, N. T. Thuong and C. Hélène, *Gene*, 1988, **72**, 361–371.

74 G. B. Dreyer and P. B. Dervan, *Proc. Natl. Acad. Sci.*, 1985, **82**, 968–972.

75 T. Le Doan, L. Perrouault, M. Chassignol, N. T. Thuong and C. Hélène, *Nucleic Acids Res.*, 1987, **15**, 8643–8659.

76 T. Le Doan, L. Perrouault, C. Helene, M. Chassignol and Nguyen Thanh Thuong, *Biochemistry*, 1986, **25**, 6736–6739.

77 J. C. François, T. Saison-Behmoaras, C. Barbier, M. Chassignol, N. T. Thuong and C. Helene, *Proc. Natl. Acad. Sci.*, 1989, **86**, 9702–9706.

78 C.-H. B. Chen and D. S. Sigman, *J. Am. Chem. Soc.*, 1988, **110**, 6570–6572.

79 C.-H. B. Chen and D. S. Sigman, *Proc. Natl. Acad. Sci.*, 1986, **83**, 7147–7151.

80 M. Pitié and B. Meunier, *Bioconjug. Chem.*, 1998, **9**, 604–611.

81 S. A. Ross, M. Pitié and B. Meunier, *Eur. J. Inorg. Chem.*, 1999, **1999**, 557–563.

82 M. Pitié, J. D. Van Horn, D. Brion, C. J. Burrows and B. Meunier, *Bioconjug. Chem.*, 2000, **11**, 892–900.

83 G. Felsenfeld, D. R. Davies and A. Rich, *J. Am. Chem. Soc.*, 1957, **79**, 2023–2024.

84 N. T. Thuong and C. Hélène, *Angew. Chemie Int. Ed.*, 1993, **32**, 666–690.

85 S. Buchini, *Curr. Opin. Chem. Biol.*, 2003, **7**, 717–726.

86 P. A. Beal and P. B. Dervan, *Nucleic Acids Res.*, 1992, **20**, 2773–2776.

87 G. Goldsmith, T. Rathinavelan and N. Yathindra, *PLoS One*, 2016, **11**, e0152102.

88 J. L. Mergny, J. S. Sun, M. Rougée, T. Montenay-Garestier, J. Chomilier, C. Hélène and F. Barcelo, *Biochemistry*, 1991, **30**, 9791–9798.

89 J. S. Sun, J. L. Mergny, R. Lavery, T. Montenay-Garestier and C. Hélène, *J. Biomol. Struct.*

Dyn., 1991, **9**, 411–424.

90 M. P. Knauert and P. M. Glazer, *Hum. Mol. Genet.*, 2001, **10**, 2243–2251.

91 M. Faria and C. Giovannangeli, *J. Gene Med.*, 2001, **3**, 299–310.

92 D. Praseuth, A. L. Guieysse and C. Hélène, *Biochim. Biophys. Acta - Gene Struct. Expr.*, 1999, **1489**, 181–206.

93 M. Grigoriev, D. Praseuth, A. L. Guieysse, P. Robin, N. T. Thuong, C. Helene and A. Harel-Bellan, *Proc. Natl. Acad. Sci.*, 1993, **90**, 3501–3505.

94 M. Duca, P. Vekhoff, K. Oussedik, L. Halby and P. B. Arimondo, *Nucleic Acids Res.*, 2008, **36**, 5123–5138.

95 V. N. Potaman, *Expert Rev. Mol. Diagn.*, 2003, **3**, 481–496.

96 L. A. Christensen, H. Wang, B. Van Houten and K. M. Vasquez, *Nucleic Acids Res.*, 2008, **36**, 7136–7145.

97 L. Perrouault, U. Asseline, C. Rivalle, N. T. Thuong, E. Bisagni, C. Giovannangeli, T. Le Doan and C. Hélène, *Nature*, 1990, **344**, 358–360.

98 A. S. Pavlova, P. E. Vorobyev and V. F. Zarytova, *Russ. J. Bioorganic Chem.*, 2009, **35**, 197–206.

99 D. Pei, D. R. Corey and P. G. Schultz, *Proc. Natl. Acad. Sci.*, 1990, **87**, 9858–9862.

100 R. Landgraf, C.-H. B. Chen and D. S. Sigman, *Biochemistry*, 1994, **33**, 10607–10615.

101 H. Moser and P. Dervan, *Science*, 1987, **238**, 645–650.

102 J. C. François, T. Saison-Behmoaras, M. Chassignol, N. T. Thuong and C. Helene, *J. Biol. Chem.*, 1989, **264**, 5891–5898.

103 M. Shimizu, H. Inoue and E. Ohtsuka, *Biochemistry*, 1994, **33**, 606–613.

104 A. Panattoni, A. H. El-Sagheer, T. Brown, A. Kellett and M. Hocek, *ChemBioChem*, 2020, **21**, 991–1000.

105 M. C. Uzagare, I. Claußnitzer, M. Gerrits and W. Bannwarth, *ChemBioChem*, 2012, **13**, 2204–2208.

106 E. D. Goddard-Borger and R. V Stick, *Org. Lett.*, 2007, **9**, 3797–3800.

107 C. Slator, Z. Molphy, V. McKee, C. Long, T. Brown and A. Kellett, *Nucleic Acids Res.*, 2018, **46**, 2733–2750.

108 N. Zuin Fantoni, Z. Molphy, C. Slator, G. Menounou, G. Toniolo, G. Mitrikas, V. McKee, C. Chatgilialoglu and A. Kellett, *Chem. - A Eur. J.*, 2019, **25**, 221–237.

109 S. Tabassum, W. M. Al-Asbahy, M. Afzal, F. Arjmand and V. Bagchi, *Dalt. Trans.*, 2012, **41**, 4955–4964.

110 Z. Molphy, A. Prisecaru, C. Slator, N. Barron, M. McCann, J. Colleran, D. Chandran, N. Gathergood and A. Kellett, *Inorg. Chem.*, 2014, **53**, 5392–5404.

111 Z. Molphy, C. Slator, C. Chatgilialoglu and A. Kellett, *Front. Chem.*, 2015, **3**, 1–9.

112 F. Seela, V. R. Sirivolu and P. Chittepu, *Bioconjug. Chem.*, 2008, **19**, 211–224.

113 A. B. Neef and N. W. Luedtke, *ChemBioChem*, 2014, **15**, 789–793.

114 J. Ludwig, *Acta Biochim. Biophys. Acad. Sci. Hungaricae*, 1981, 16, 131–133.

115 S. A. Ingale and F. Seela, *J. Org. Chem.*, 2013, **78**, 3394–3399.

116 S. S. Pujari, S. A. Ingale and F. Seela, *Bioconjug. Chem.*, 2014, **25**, 1855–1870.

117 F. Seela and V. R. Sirivolu, *Helv. Chim. Acta*, 2007, **90**, 535–552.

118 T. R. Chan, R. Hilgraf, K. B. Sharpless and V. V. Fokin, *Org. Lett.*, 2004, **6**, 2853–2855.

119 M. Uhlen, *Nature*, 1989, **340**, 733–734.

120 P. Ménová, V. Raindlová and M. Hocek, *Bioconjug. Chem.*, 2013, **24**, 1081–1093.

121 M. Citartan, T. H. Tang, S. C. Tan and S. C. B. Gopinath, *World J. Microbiol. Biotechnol.*, 2011, **27**, 1167–1173.

122 P. D. Kwong, R. Wyatt, J. Robinson, R. W. Sweet, J. Sodroski and W. A. Hendrickson, *Nature*, 1998, **393**, 648–659.

123 B. Faucon, J. L. Mergny and C. Helene, *Nucleic Acids Res.*, 1996, **24**, 3181–3188.

124 A. Panattoni, R. Pohl and M. Hocek, *Org. Lett.*, 2018, **20**, 3962–3965.

125 J. Diot, M. I. García-Moreno, S. G. Gouin, C. O. Mellet, K. Haupt and J. Kovensky, *Org. Biomol. Chem.*, 2009, **7**, 357–363.

126 J. Muñoz, J. L. Gelpí, M. Soler-López, J. A. Subirana, M. Orozco and F. J. Luque, *J. Phys. Chem. B*, 2002, **106**, 8849–8857.

127 Y. Aiba, J. Sumaoka and M. Komiyama, *Chem. Soc. Rev.*, 2011, **40**, 5657–5668.

128 P. Kielkowski, J. Fanfrlík and M. Hocek, *Angew. Chemie Int. Ed.*, 2014, **53**, 7552–7555.

129 P. Kielkowski, H. Macíčková-Cahová, R. Pohl and M. Hocek, *Angew. Chemie Int. Ed.*, 2011,

50, 8727–8730.

130 J. R. Kiefer, C. Mao, J. C. Braman and L. S. Beese, *Nature*, 1998, **391**, 304–307.

131 J. Fanfrlík, A. K. Bronowska, J. Řezáč, O. Přenosil, J. Konvalinka and P. Hobza, *J. Phys. Chem. B*, 2010, **114**, 12666–12678.

132 M. Lepšík, J. Řezáč, M. Kolář, A. Pecina, P. Hobza and J. Fanfrlík, *Chempluschem*, 2013, **78**, 921–931.

133 K. Fox, *Curr. Med. Chem.*, 2000, **7**, 17–37.

134 L. Lacroix and J.-L. Mergny, *Arch. Biochem. Biophys.*, 2000, **381**, 153–163.

135 S. Walsh, A. H. El-Sagheer and T. Brown, *Chem. Sci.*, 2018, **9**, 7681–7687.

136 M. Mills, P. B. Arimondo, L. Lacroix, T. Garestier, H. Klump and J. L. Mergny, *Biochemistry*, 2002, **41**, 357–366.

137 M. Böttger, B. Wiegmann, S. Schaumburg, P. G. Jones, W. Kowalsky and H.-H. Johannes, *Beilstein J. Org. Chem.*, 2012, **8**, 1037–1047.

138 V. Raindlová, M. Janoušková, M. Slavíčková, P. Perlíková, S. Boháčová, N. Milisavljevič, H. Šanderová, M. Benda, I. Barvík, L. Krásný and M. Hocek, *Nucleic Acids Res.*, 2016, **44**, 3000–3012.