Univerzita Karlova v Praze

Přírodovědecká fakulta

Studijní program: Organická chemie

Studijní obor: Organická chemie

Ing. Pavel Kielkowski

Nové funkcionalizované nukleové kyseliny pro aplikaci v chemické biologii

New functionalized nucleic acids for application in chemical biology

Dizertační práce

Vedoucí závěrečné práce/Školitel: Prof. Ing. Michal Hocek, CSc., DSc.

Praha, 2014

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, 15.8.2014

Podpis

Acknowledgement

I would like to thank my supervisor Prof. Michal Hocek for an interesting research project, his excellent guidance and support over the course of my research. My grand gratitude goes to Dr. Radek Pohl (IOCB) for measurements and interpretation of NMR spectra. I wish also to thank Dr. Nelson Brock and Prof. Jeroen S. Dickschat (Technische Univerität Braunschweig) for fruitful collaboration on *TdaD* gene cloning. Next, I want to thank Dr. Jindřich Fanfrlík for computational studies on competitive incorporations, Ivana Křížová for help with the metabolic labeling, Jan Tykvart for his advices on the surface plasmon resonance and Dr. Eva Kaletová for introduction into atomic force microscopy. My thanks goes also to Dr. Markéta Šmídková for helpful biochemical discussions. I must thank to all collegues from our research group for their advices, suggestions and pleasant social background. I am grateful to Dr. Pavla Perlíková, Dr. Hana Macíčková-Cahová, Anna Tokarenko, Dr. Tomáš Kubelka, Michal Tichý and Michael A. Downey for their friendship and a lot of fun. Finally, my great thanks belongs to my parents for their support and Michaela, Lia and Kira for having a patience and comprehension with my work.

All the synthetic work and biochemical experiments were performed by me. I have also performed most of the spectroscopic measurements, but the measurement and interpretation of NMR spectra of some intermediates and all final compounds were done by Dr. Radek Pohl. The parts that were done by our collaborating groups are always distinctly denoted in the thesis.

This work was supported by the Academy of Sciences of the Czech Republic (ASCR) (Z4 055 0506), by the Ministry of Education of the Czech Republic (LC512), by the Czech Science Foundation (203/09/0317 and P206/12/G151), and by Gilead Sciences, Inc. (Foster City, CA, U. S. A.).

Abstract

This work is focused on the synthesis of the modified 2'-deoxyribonucleoside triphosphates, their incorporation into DNA and use in chemical biology applications. The synthetic routes to the double-headed nucleosides and nucleotide triphosphates in which the two nucleobases were connected via ethynyl or propargyl linker has been developed. (Cytosin-5-yl)ethynyl, 3-(cytosin-1-yl)prop-1-yn-1-yl and 3-(5-fluorocytosin-1-yl)prop-1-yn-1-yl derivatives of pyrimidine and 7-deazaadenine 2'-deoxyribonucleosides and nucleoside triphosphates were prepared by aqueous palladium-catalyzed cross-coupling reactions. The double-headed modified nucleoside triphosphates were good substrates for DNA polymerases suitable for primer extension and PCR construction of DNA bearing linked cytosine or 5-fluorocytosine in the major groove mimicking the flliped-out nucleotide. The assay for the testing of the inhibition of DNA methyltransferases was developed. Next, the transient protection of DNA against cleavage by restriction endonucleases (REs) using (trialkylsilyl)ethynyl modified DNA was developed. A series of 7-(trialkylsilyl)ethynyl-7-deaza-2'-deoxyadenosine triphosphates was prepared and they were shown to be incorporated into DNA by primer extension and/or PCR using KOD XL polymerase. The deprotection conditions of the trialkylsilyl protecting groups were optimized on model nucleoside monophospahtes. The ability to protect the DNA against cleavage by REs by using (trialkylsilyl)ethynyl modifications was tested. It was found that the (triethylsilyl)ethynyl-protected DNA resists the cleavage by RE, but after it is treated with NH_3, the resulting deprotected ethynyl-modified DNA is fully cleavable by the REs. This 7-(triethylsilyl)ethynyl-7-deaza-2'-deoxyadenosine triphosphate was also used in a PCR-based synthesis of a gene internally protected against cleavage by restriction endonucleases. The unmodified flanking regions were cleaved for cloning into a plasmid which was replicated by *E. coli*, and used for protein production. Finally, a series of 7-substituted 7-deazaadenine and 5-substituted cytosine 2′-deoxyribonucleoside triphosphates were tested for their competitive incorporations (in the presence of their natural counterparts) into DNA by several DNA polymerases by using analysis based on cleavage by restriction endonucleases. 7-Aryl-7-deaza-2'-deoxyadenosine triphosphates were shown to be more efficient substrates than dATP because of their higher affinity for the active site of the enzyme, as proved by kinetic measurements and calculations.

Abstrakt

Tato práce je zaměřena na syntézu modifikovaných 2'-deoxyribonukleosid trifosfátů, jejich inkorporaci do DNA a aplikaci v chemické biologii. Byla vyvinuta syntéza "douhlavých" nukleosidů a nukleosid trifosfátů, které obsahují ethynylový nebo propargylový můstkem mezi dvěmi bázemi. (Cytosin-5-yl)ethynyl, 3-(cytosin-1-yl)prop-1-yn-1-yl a 3-(5-fluorcytosin-1-yl)prop-1-yn-yl substituované pyrimidiny a 7-deazaadeniny 2'-deoxyribonukleosidů a nukleosid trifosfátů byly připraveny palladiem katalyzovanou cross-coupling reakcí ve vodném prostředí. Tyto 2'-deoxyribonukleosid trifosfáty obsahující dvě nukleobáze jsou dobrými substráty pro DNA polymerasy v primer extension reakci a PCR a tudíž vhodné pro přípravu DNA s cytosinem nebo 5-fluorcytosinem ve velkém žlábku, která tak napodobuje nukleobázi vyjmutou ven z DNA dvoušroubovice. Byla také vyzkušena metoda pro testování inhibice DNA methyltransferáz. Dále, bylo vyvinuto přechodné chránění DNA proti štěpení restrikčními endonukleázami (RE) za pomocí (trialkylsilyl)ethynyl modifikované DNA. Bylo dokázáno, že všechny připravené (trialkylsilyl)ethnyl-modifikované 7-deaza-2'-deoxyadenosin trifosfáty se inkorporují do DNA KOD XL polymerásou v primer extension reakci a nebo pomocí PCR. Podmínky potřebné pro odchránění trialkylsilylové chránicí skupiny byly optimalizovány na nukleosid monofosfátech. Rovněž byla testována schopnost (trialkylsiyl)ethynyl modifikované DNA chránit tuto DNA proti štěpení RE. Zjistili jsme, že DNA chráněná (triethylsilyl)ethynylovou skupinou nepodléhá štěpení pomocí RE, po odchránění triethylsilylové skupiny amoniakem výsledná ethynyl modifikovaná DNA již štěpena RE je. Tento 7-(triethylsilyl)ethynyl-7-deaza-2'-deoxyadenosin trifosfát byl také použit při přípravě genu částečně chráněného proti štěpení RE. Pro přípravu byla použita vlastní nová metoda založená na PCR. V genu neobsažené nemodifikované sekvence byly dále použity pro klonování do plasmidu, jenž byl následně replikován v *E. coli* a použit pro produkci proteinu. A konečně, série 7-substituovaných 7-deazaadenin a 5-substituovaných cytosin 2'-deoxyribonukleosid trifosfátů byla testována v přímé kompetitivní inkorporaci (s jejich přirozenými analogy) do DNA pomocí celé řady DNA polymeras. K analýze byla použita metoda založená na štěpících vlastnostech RE. Ukázalo se, že 7-aryl-7-deaza-2'-deoxyadenosin trifosfáty jsou lepšími substráty pro DNA polymerasy než přirozená dATP, protože mají větší afinitu k aktivnímu místu polymerasy. Zmíněná zjištění byla dále potvrzena měřením kinetik a modelováním.

List of abbreviations

Ac	Acetyl
AdoHcy	*S*-Adenosyl-L-homocysteine
Bn	Benzyl
BrdU	5-Bromo-2'deoxyribouridine
Bz	Benzoyl
CuAAC	Copper-catalyzed azide–alkyne cycloaddition
DMF	*N*,*N*-Dimethylformamide
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
Et	Ethyl
HPLC	High performance liquid chromatography
IPTG	Isopropyl β-D-1-thiogalactopyranoside
MALDI	Matrix-assisted laser desorption/ionization
Me	Methyl
MS	Mass spectrometry
MT	Methyltransferase
NMR	Nuclear magnetic resonance
NTA	Nitriolotriacetic acid
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
PEX	Primer extension
Ph	Phenyl
Pol	Polymerase
RE	Restriction endonuclease
SAM	*S*-Adenosyl-L-methionine
TCA	Trichloroacetic acid
TEAB	Tetraethylammonium bicarbonate
TESE	(Triethylsilyl)ethynyl
THF	Tetrahydrofuran
TIPSE	(Triisopropylsilyl)ethynyl
TMS	Trimethylsilyl

TMSE	(Trimethylsilyl)ethynyl
TPPTS	Tris(3-sulfophenyl)phosphine trisodium salt

List of publications of the author related to the thesis

1. Kielkowski, P.; Pohl, R.; Hocek, M.: "Synthesis of Acetylene Linked Double-Nucleobase Nucleos(t)ide Building Blocks and Polymerase Construction of DNA Containing Cytosines in the Major Groove" *J. Org. Chem.* **2011**, *76*, 3457–3462.

2. Kielkowski, P.; Macíčková-Cahová, H.; Pohl, R.; Hocek, M.: "Transient and Switchable (Triethylsilyl)ethynyl Protection of DNA against Cleavage by Restriction Endonucleases" *Angew. Chem. Int. Ed.* **2011**, *50*, 8727 –8730.

3. Kielkowski, P.; Brock, N. L.; Dickschat, J. S.; Hocek, M.: "Nucleobase Protection Strategy for Gene Cloning and Expression" *ChemBioChem* **2013**, *14*, 801-804.

4. Kielkowski, P.; Fanfrlík, J.; Hocek, M.: "7-Aryl-7-deazaadenine dNTPs are better substrates for DNA polymerases than dATP in competitive incorporations" *Angew. Chem. Int. Ed.* **2014**, *53*, 7552−7555.

Contents

1 Introduction

1.1 Nucleic Acids

Nucleic acids were discovered in 1869 by Friedrich Miescher and in 1953 the double helix structure of deoxyribonucleic acid (DNA) was deciphered by James Watson and Francis Crick. Since then the acquired knowledge of structure, physical and chemical properties and function of DNA and later discovered ribonucleic acid (RNA) has had tremendous impact to the development of nowadays medicine, molecular biology, chemical biology and chemistry.

Double helical DNA has three major forms, the most common B-DNA, A-DNA and Z-DNA (**Figure 1**). All three DNA structures have in common the sugar-phosphate backbone and aromatic bases (A, T, G, and C), which occupy the core of the helix (**Figure 2**). Bases form complementary pairs A · T and G · C.

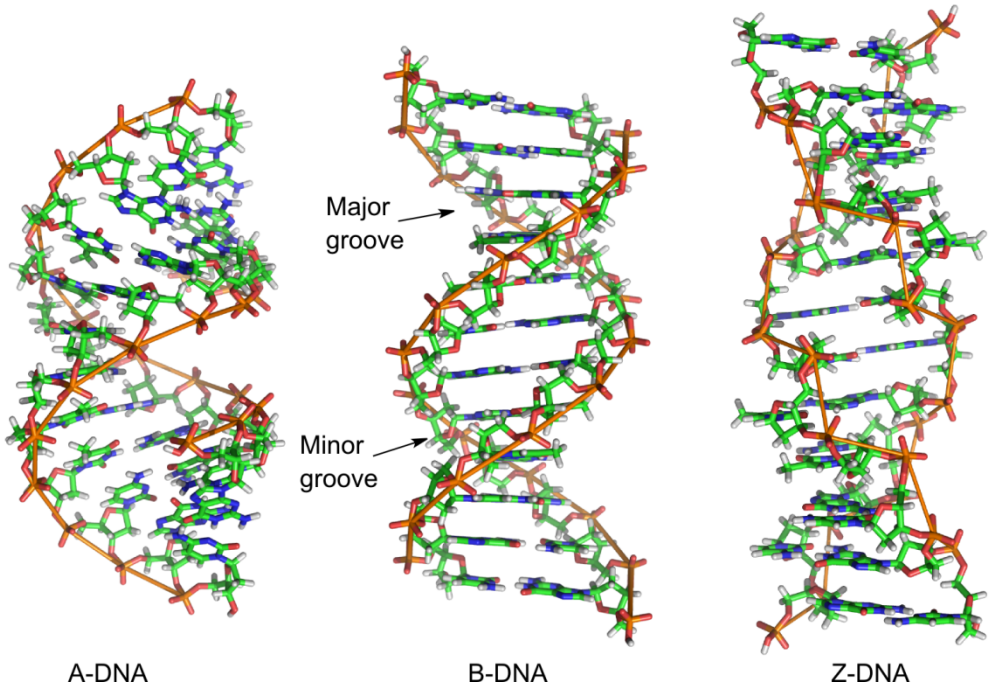


Figure 1. The major double helical forms of DNA.

A-DNA forms (at low humidity and high salt) a wider and flatter right handed helix than B-DNA does. It was among the first oligonucleotides to be crystallized and later on it has been observed in the active site of DNA polymerases and in spores of Gram-positive bacteria.[1a] Z-DNA that forms a left handed helix was discovered and crystallized by Rich and co-workers.[2] It is best typified by an alternating $(dG\text{-}dC)_n$ polymer in which the cytosines take the *anti* conformation and the guanines the *syn* conformation. The biologically predominant B-DNA is formed at high humidity (and low salt) into right handed double helix with the

relatively narrow minor groove and relatively wide major groove.[3] B-DNA in contrast to the A- and Z-forms is more flexible and capable to undergo small adjustments in local helix structures in response to particular base sequences in order to minimize non-covalent interactions between adjacent bases and maximize base stacking.

Lately, much attention was paid to the guanine rich sequences, which can fold into the non-canonical G-quadruplex structures formed through the stacking of Hoogsteen hydrogen bonded G-tetrads stabilized by a metal cation. The sequences possessing the potential to form stable G-quadruplexes were observed in both prokaryotic and eukaryotic genomes.[4]

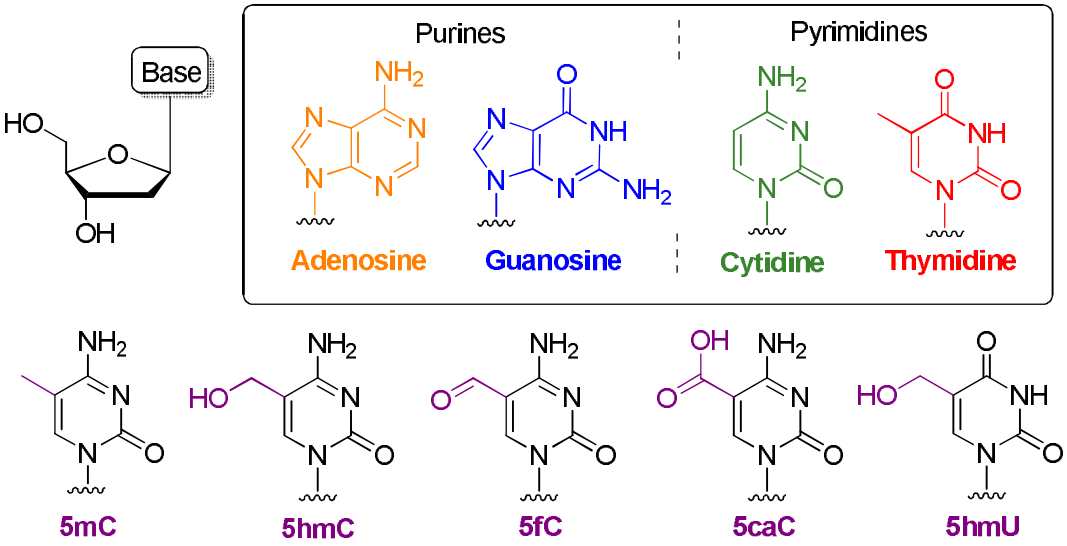
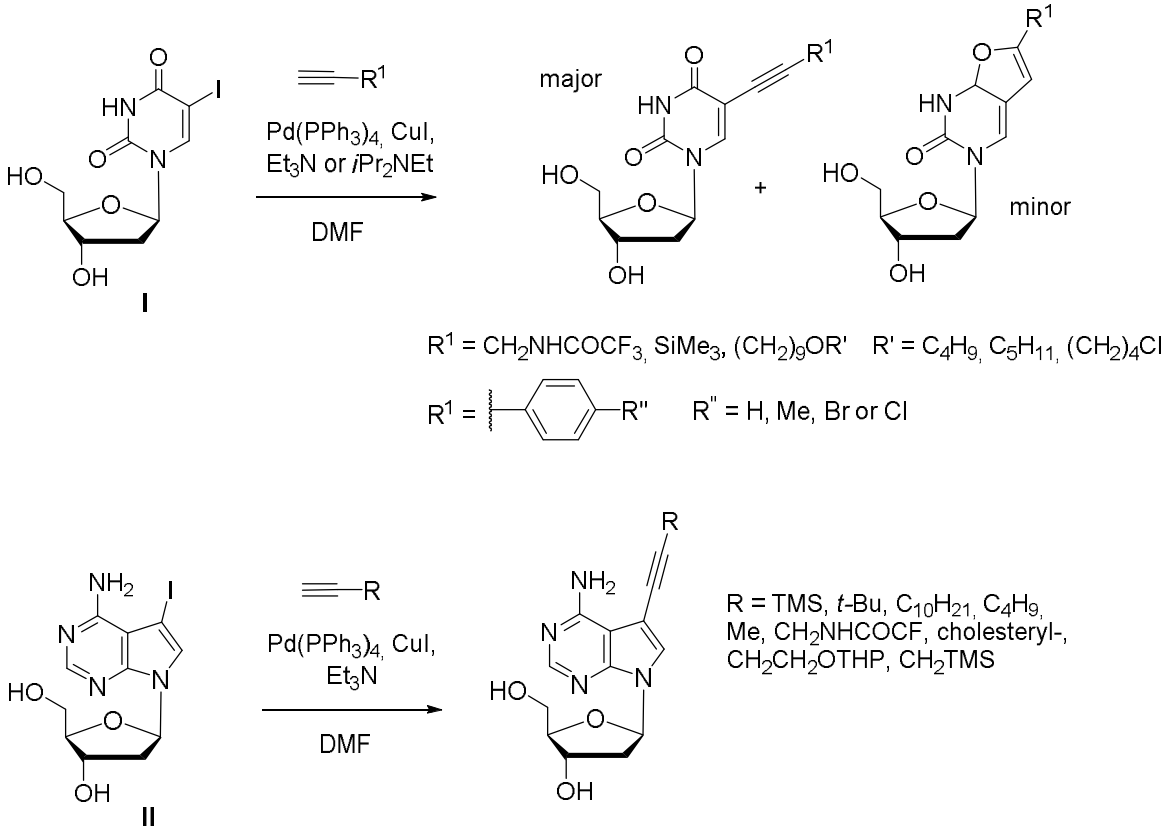


Figure 2. DNA building blocks A, G, C, T and pyrimidine natural analogues.

Since the organization of living organisms is complex the genetic information coded by four nucleotides is systematically modified in order to expand the information capacity, control the gene expression, protect the DNA but not involve a change in the nucleotide sequence itself. The most common modifications in prokaryotes are N^6-methylation of adenosine and C^5-methylation of cytidine. In eukaryotes the C^5-methylated cytidines (5mC)[5] in CpG sequences play a critical signalling role in biology. Recently, four additional modifications were found to exist in mammalian DNA 5-hydroxymethylcytosine (5hmC),[6] formylcytosine (5fC), 5-carboxycytosine (5caC)[7] and 5-hydroxymethyluridine (5hmU)[8] (**Figure 2**), these are involved in the demethylation pathway of the 5mC, but they might possess also additional functions in epigenetics, which are not yet fully understood.

1.2 Synthesis of Modified Nucleosides and Nucleotides

A vast number of nucleosides and nucleotides bearing modified bases have been designed and prepared. Many display significant biological activity, while others are used for applications in chemical biology, such as nucleic acid sequencing,[9] labeling,[10] investigation of nucleic-acid structure[11] and protein-nucleic acid interactions.[12] The functionality can be attached to the nucleobase for instance by palladium-catalyzed reaction[13] of corresponding halogenated nucleoside with terminal alkynes, alkenes or arylboronic acids. The most common is Sonogashira cross-coupling of 5-iodopyrimidines[14] (**I**) or 7-iodo-7-deazapurines[15] (**II**) with diverse terminal alkynes in DMF (**Scheme 1**).

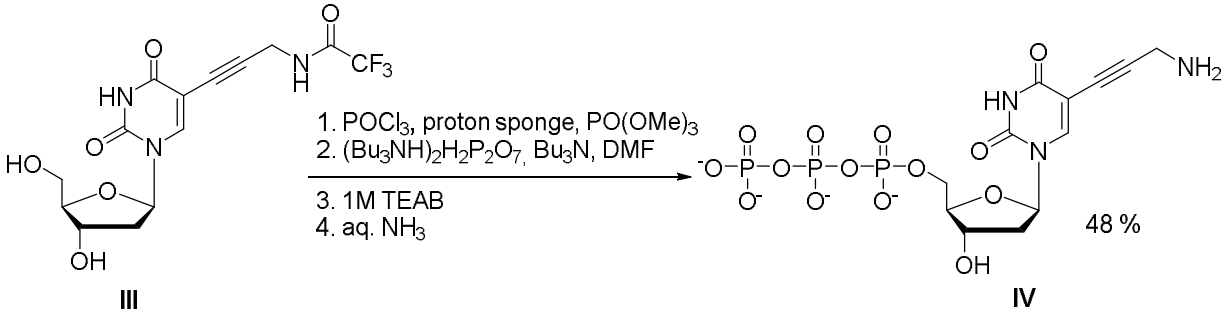


Scheme 1. Representative examples of palladium-catalyzed reactions for preparation of 5-substituted 2'-deoxyuridines and 7-substituted 7-deaza-2'-deoxyadenosines.

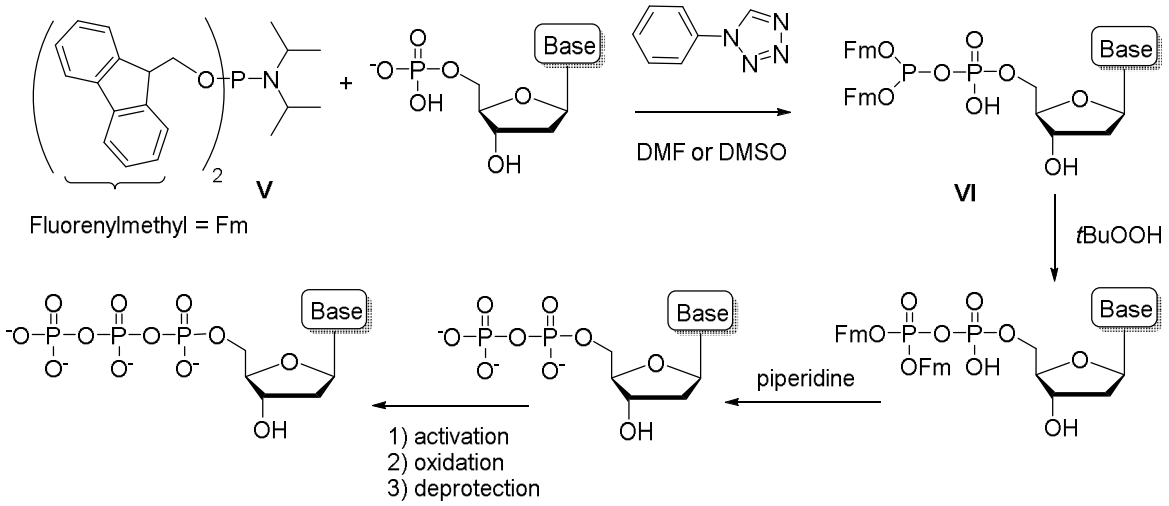
Resulting functionalized nucleoside **III** can be then used for synthesis of the corresponding nucleoside triphosphates **IV**, where the functional groups have to be protected and deprotected (**Scheme 2**).[16] Thus, the synthesis of the modified nucleoside triphosphates require laborious multistep sequences with rather low total yields and all steps in the synthesis need to be optimized for every triphosphate. The synthesis of triphosphates by standard

procedure[16] is done by phosphoryl chloride in trimethyl phosphate in the first step, followed by addition of pyrophosphate with tributylamine in DMF (**Scheme 2**).

Very recently, the alternative methodology for synthesis of triphosphates was developed (**Scheme 3**).[17] This method has proved that P-amidites (**V**) can be used in protecting-group-free iterative couplings to selectively give mixed P_{III}-P_V anhydrides (**VI**). These intermediates can be oxidized with following rapid removal of the two terminal fluorenylmethyl groups. The advantage of this approach is use of non-dry reagents and solvents and ambient conditions.



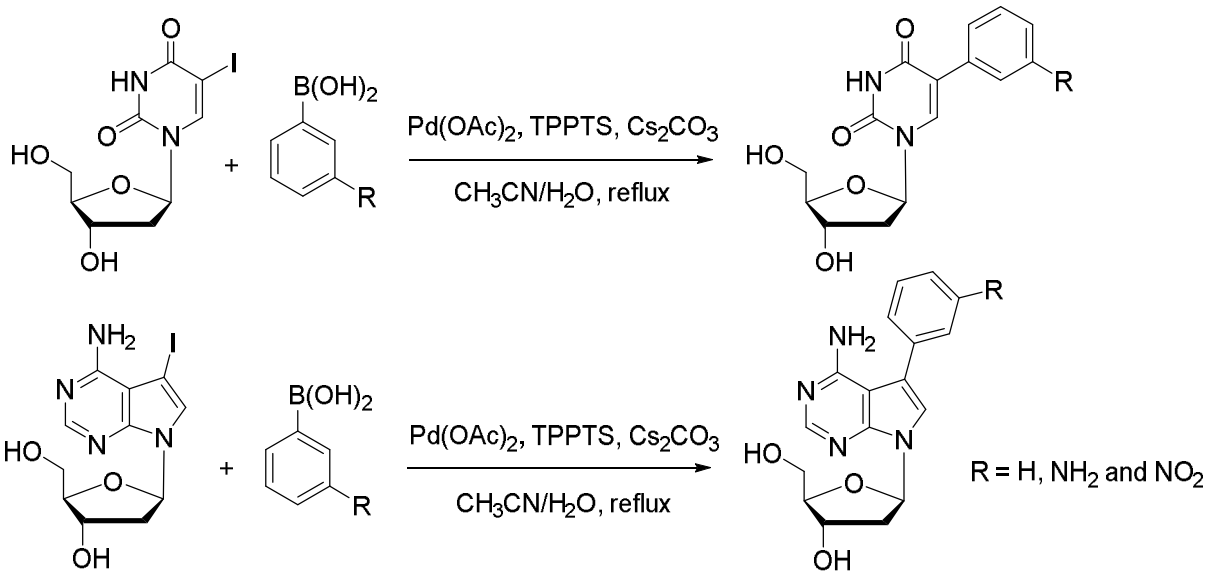
Scheme 2. Typical synthesis of functionalized 2'-deoxyribonucleoside triphosphate.



Scheme 3. Alternative synthesis of 2'-deoxyribonucleoside triphosphates.

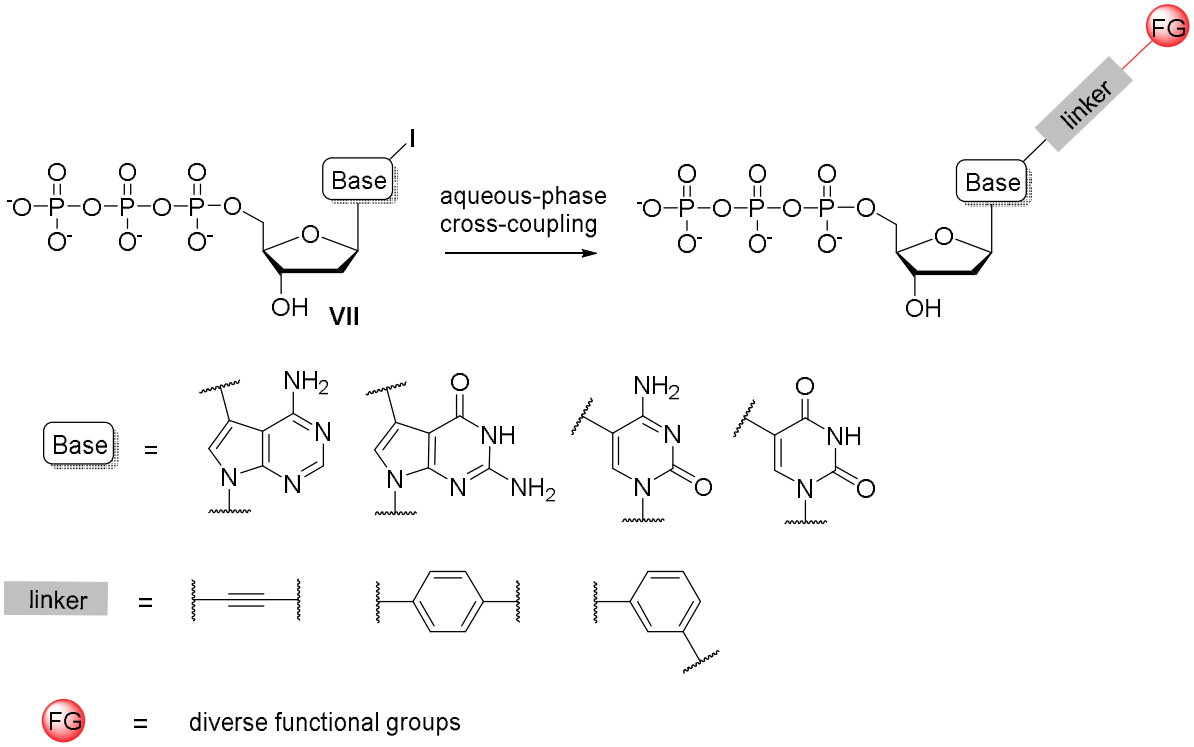
The development of the water soluble catalytic systems led to the advancement of aqueous-phase cross-coupling reactions. This was applied for the first time by Shaughnessy and co-workers[18] for arylation of unprotected nucleosides by Suzuki-Miyaura reaction with arylboronic acids in the presence of tris(3-sulfonylphenyl)phosphine (TPPTS), $Pd(OAc)_2$ and

inorganic base (*i.e.* Na_2CO_3) in a mixture of acetonitrile-water (**Scheme 4**).[19,20] Similar conditions were then employed for Sonogashira[20] and Heck[21] cross-couplings as well.



Scheme 4. Representative examples of aqueous Suzuki-Miyaura cross-coupling reactions for preparation of modified nucleosides.

Recently, the Hocek group developed a new very convenient method for the preparation of the functionalized nucleoside triphosphates based on the synthesis of iodinated nucleoside triphosphate **VII** (5-iodo-2'-deoxycytidine, 5-iodo-2'-deoxyuridine, 7-iodo-7-deaza-2'-deoxyadenosine, 7-iodo-7-deaza-2'-deoxyguanosine triphosphates) followed by aqueous cross-coupling reactions with diverse groups of terminal alkynes, alkenes and arylboronic acids bearing functional groups used for further applications (**Scheme 5**).[20,21] This approach represents very straightforward methodology, which avoids the time consuming optimization of the triphosphorylation and it is compatible with most of the functional groups. For diversity of the functional groups see **Figure 3**. Suzuki-Miyaura cross-coupling reaction of iodinated 2'-deoxyribonucleoside triphosphate with arylboronic acids or pinacol ester of biarylboronates were used in the syntheses of **VIII**–**IX**[22] and **X**–**XI**[12a] respectively. Syntheses of **XII** – **XV**[12,23] were carried out by Sonogashira cross-coupling and **XVI** – **XVII**[21c] were prepared by Heck reaction.



Scheme 5. General approach to synthesis of functionalized 2'-deoxyribonucleoside triphosphates with ethynyl or phenyl linker.

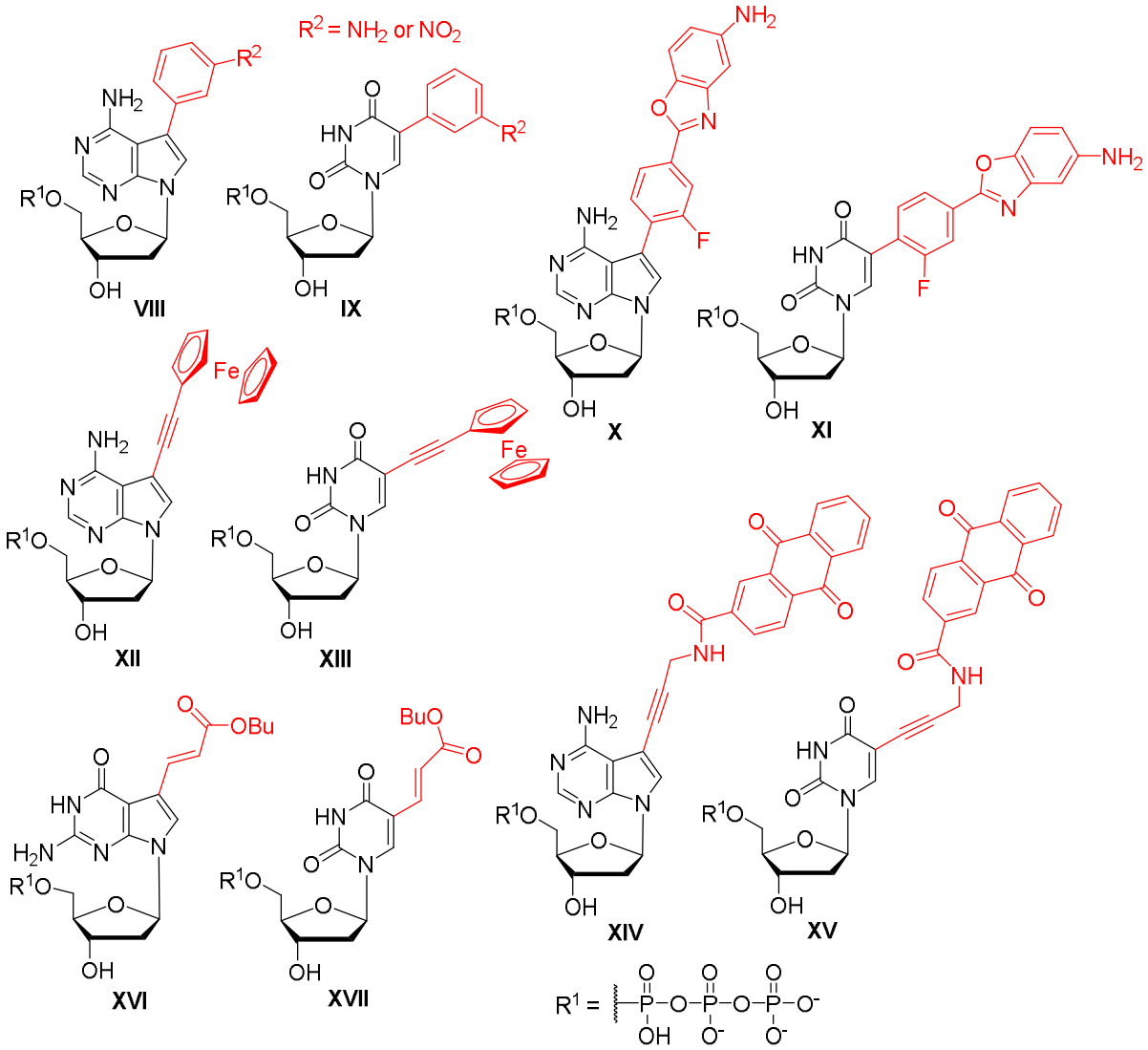


Figure 3. Representative examples of functionalized 2'-deoxyribonucloside triphosphates.

1.3 Synthesis of Oligonucleotides and Nucleic Acids

Generally in chemical oligonucleotide synthesis, phosphodiester bonds are formed between a 3'-hydroxyl group bearing a phosphate derivative and a 5'-hydroxyl group of another nucleoside. In contrary, DNA polymerases used in enzymatic synthesis of oligonucleotides and nucleic acids utilize a 3'-hydroxyl group nucleophilic attack on the α-phosphate of an incoming nucleoside triphosphate.

Synthesis of oligonucleotides of less than 100 residues is done preferably by chemical synthesis. However, both chemical and enzymatic synthesis are now well established methods for synthesis of these short oligonucleotides containing a modified nucleotides. For preparation of oligonucleotides (nucleic acids) of more than 100 bases (base-pairs) the enzymatic synthesis is usually used.

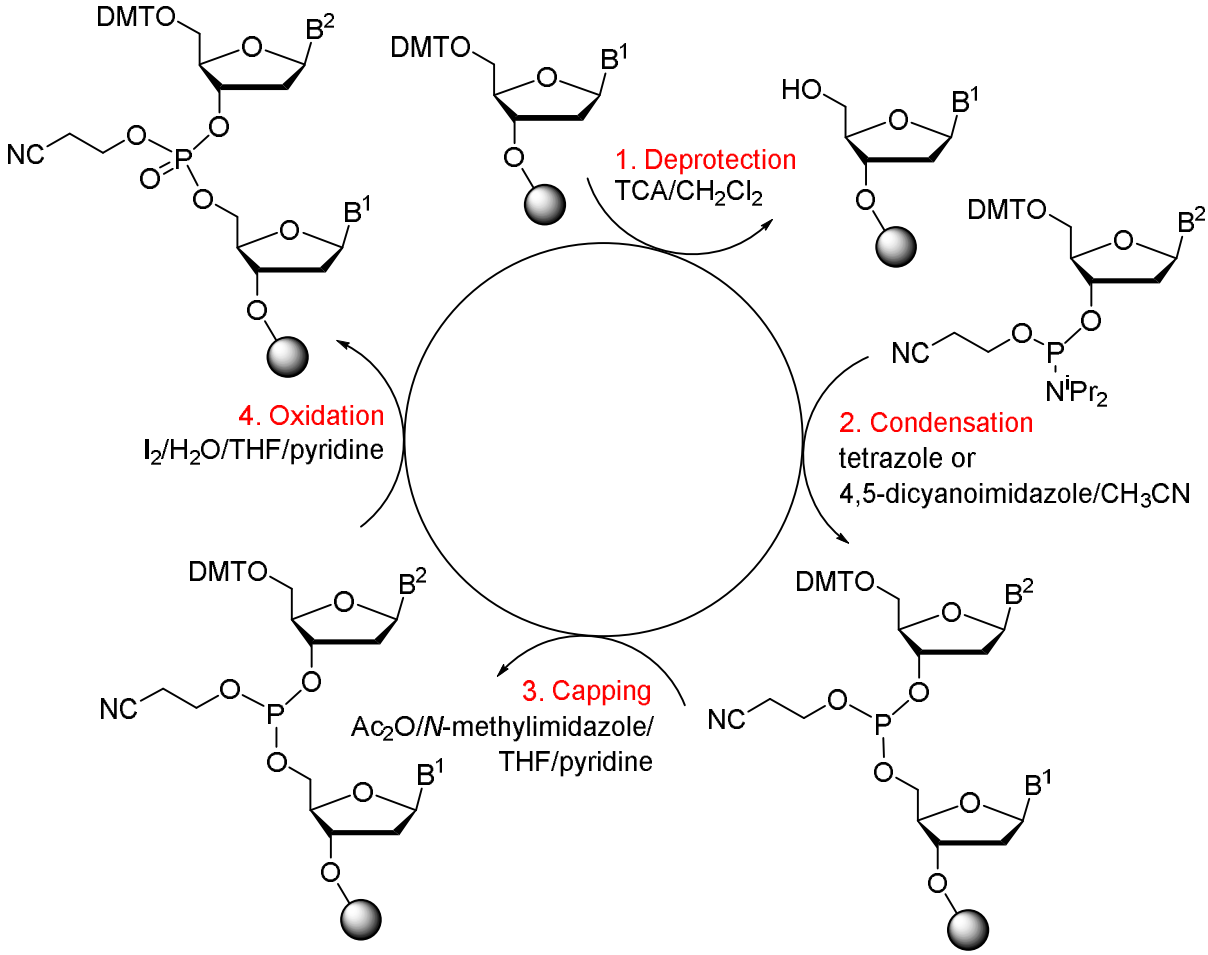
Preparations of functionalized DNAs could be done by both synthetic and enzymatic approaches. Another possibility for functionalization of DNA is a post-synthetic modification of oligonucleotides.[24]

1.3.1 Solid-Phase Chemical Synthesis

Formation of an internucleotide bond is the key step in oligonucleotide synthesis, however, only relatively mild conditions can be applied. The heterocyclic bases are prone to alkylation, oxidation and phosphorylation and the phosphodiester backbone is susceptible to hydrolysis. Formerly, the problem was solved by the development of the efficient methods of synthesis such as phosphodiester[25] and phosphotriester[26] method. Later, the solid-phase H-phosphonate[26] chemistry was developed, but suffered from serious side reactions.

Recently, phosphite triester (phosphoramidite) solid-phase method is used for the preparation of the most oligonucleotides in micrograms to many gram scale and up to 200 residues in length.[27] This process is automated, growing oligonucleotide is bound to polystyrene or controlled pore glass solid-support, the efficiency of the coupling is higher than 98 % and final oligonucleotide is removed and deprotected in single step by NH_3. The sequence of reactions is illustrated in **Scheme 6**. Synthesis of modified nucleotides is achieved by incorporation of functionalized nucleoside phosphoramidites[28] or by post-synthetic oligonucleotide modification.[24] Addition of the functional groups often leeds to necesity of using additional protecting groups and possibly lowers the yield of the coupling

step. In general, the solid-phase synthesis is not compatible with functional groups which are oxidizable, labile under acidic conditions or readily react with nucleophiles.



Scheme 6. Basic steps of oligonucleotide synthesis by the phosphoramidite method

1.3.2 Enzymatic Synthesis

1.3.2.1 DNA Polymerases

Genomic integrity is defined by efficiency and accuracy of the replication machinery in which DNA polymerases play a key role.[29] Based on primary sequence similarity, DNA polymerases from prokaryotes, archaea and eukaryotes can be categorized into families (**Table 1**).[30] Perhaps the most studied of these families is A polymerase family, which includes Klenow fragment of *Escherichia coli*, Taq pol of *Thermus aquaticus*, Bst pol of *Bacillus stearothermophilus* and T7 pol. The B family contain Human polymerase α, Vent pol of *Thermococcus literalis*, Pwo of *Pyrococcus woesii*, KOD of *Thermococcus kodakaraensis* and Pfu of *Pyrococcus furiosus*. Polymerases from families A and B are major polymerases

involved in replication machinery. The thermophilic polymerases from these two families (Taq, Vent, Pwo, KOD and Pfu) are crucial in PCR and often engineered[31] to process with even higher incorporation rate and fidelity. The B family polymerases are considered to proceed the replication with the highest fidelity.[32] Other well-known polymerases are pol β which is essential for short-patch base excision repair, Y family polymerases which replicate through damaged DNA and RT family which are heavily used in molecular biology.

Table 1. Representative members of families of DNA polymerases.

Family	Prokaryotic	Eukaryotic	Archaea	Viral
A	Pol I	Pol γ,θ		T3, T5, T7 pol
B	Pol II	Pol α,δ,ε,ζ	Pol BI, BII	Adenovirus, HSV, RB69, T4, T6
C	Pol III (α)			
D			Pol D	
X		Pol β,λ,μ, Tdt		
RT		Telomerase		Reverse transcriptase
Y	Pol IV, Pol V	Pol η,ι,κ		

The crystal structures and detailed kinetics studies of several DNA polymerases provide insight into structural and functional diversity during replication[33] and repair[34] (**Figure 4)**. In contrast to this diversity they all have in common overall shape which can be compared to the right hand and is described as consisting of thumb, palm and fingers domains, moreover they all utilize a two-metal-ioncatalyzed polymerase mechanism.[35] The process of nucleotide selection and faithful incorporation is carried out by fine synchronization of polymerases domains. The palm domains catalyze the phosphoryl transfer reaction and appear to be the only structurally and functionally homologous part through the A-, B- and Y- family whereas fingers and thumb domains are structurally completely different but the functions remain the same.[36] The fingers bind an incoming nucleoside triphosphate and interact with single-stranded template and thumb play a key role in positioning the DNA duplex and translocation. All together gave average 100 000-fold excess of the complementary to noncomplemetary incorporated nucleotide. This fidelity is increased approximately by 10-fold in some DNA polymerases by the exonucleolytic (3' - 5') proofreading domains.

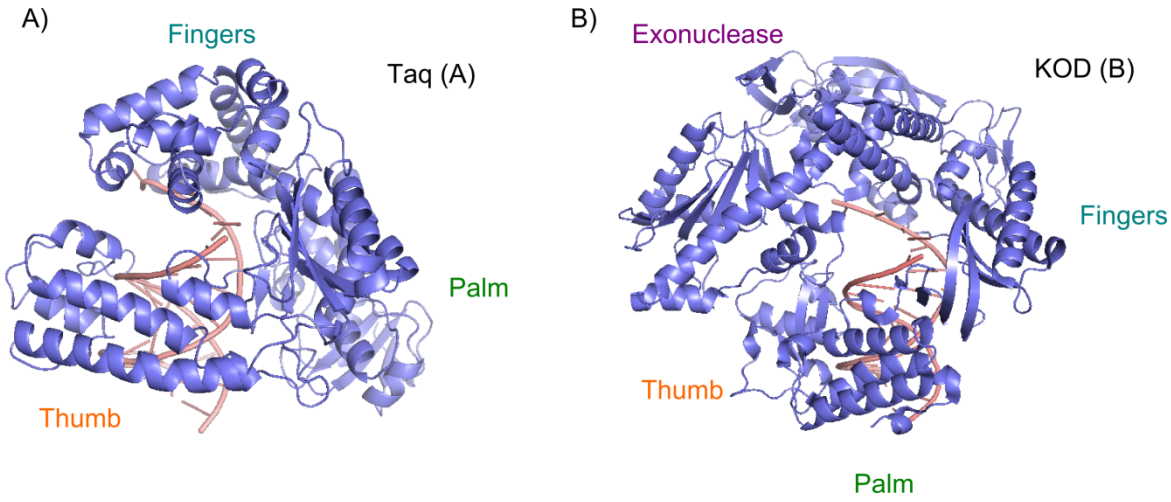


Figure 4. DNA polymerases from family A (Taq[37]) and family B (KOD[38]) domains organization.

1.3.2.2 Incorporation of Modified 2'-Deoxyribonucleoside Triphosphates by DNA Polymerases

Base-modified DNAs prepared by enzymatic synthesis are used in many applications: DNA sequencing,[9] structural and mechanistic investigation of polymerase catalysis,[11] extension of the genetic alphabet[39] and functionalization of DNA for further analysis[10] (e.g. microarrays and study of DNA-protein interactions). For all of these applications there is need for DNA polymerases which are able to incorporate modified 2'-deoxyribonucleoside triphosphates. Indeed, many DNA polymerases were tested and possess desirable properties, specifically from family A: Taq, Tth, Klenow fragment (exo-), Bst and from family B Pfu, Pwo, Vent (exo-), KOD.[20,49] In general, these tests revealed that family B perform better than family A polymerases. In addition, there are engineered polymerases[31] e.g. Phusion and KlenTaq which fulfill the requirements for high fidelity replication and incorporation of bulky modified dNTPs, respectively.

The scope of the modified nucleotides which are incorporated by several polymerases is very wide, selected structures are presented in **Figure 5**. In general, the modifications are coupled with the nucleobases via linker into the 7 position of the 7-deazapurines and 5 position of pyrimidines and after the incorporation to DNA duplex the modification is directed towards the major groove. In next-generation sequencing the 3'-*O*-azidomethyl dNTPs are readily incorporated with high fidelity, each dNTP is labeled with different removable fluorophore[9] (**XVIII**). 5-Ethynyl dCTP[40] (**XIX**), 5-bromo[41] (**XX**) and 5-azidomethyl[42] (**XI**) dUTPs and more functionalized nucleosides[43] are good substrates for

polymerases for labeling in living cells. Chromatin domains labeling during replication in living cells is done by bulky Cy3-dUTP[44] (**XXII**). Reactive vinylsulfonamide modified dCTP (**XXIII**) was used for preparation of DNA which was subsequently used for cross-linking with cysteine containing protein p53.[21] Solvatochromic 4-aminophatalimide (**XXIV**) fluorophore functionalized DNA for studying protein-DNA interactions was prepared using corresponding modified dATP. [45] Nitrophenyl modified dATP (**XXV**) is one representative of redox labeled dNTPs which are incorporated into DNA in PEX or PCR and then used for electrochemical detection of functionalized DNA.[49b] Even extremely bulky modifications of dNTPs such as oligonucleotides (**XXVI**) are substrates for KlenTaq polymerase and the resulting DNAs serve as barcoded DNAs. [46]

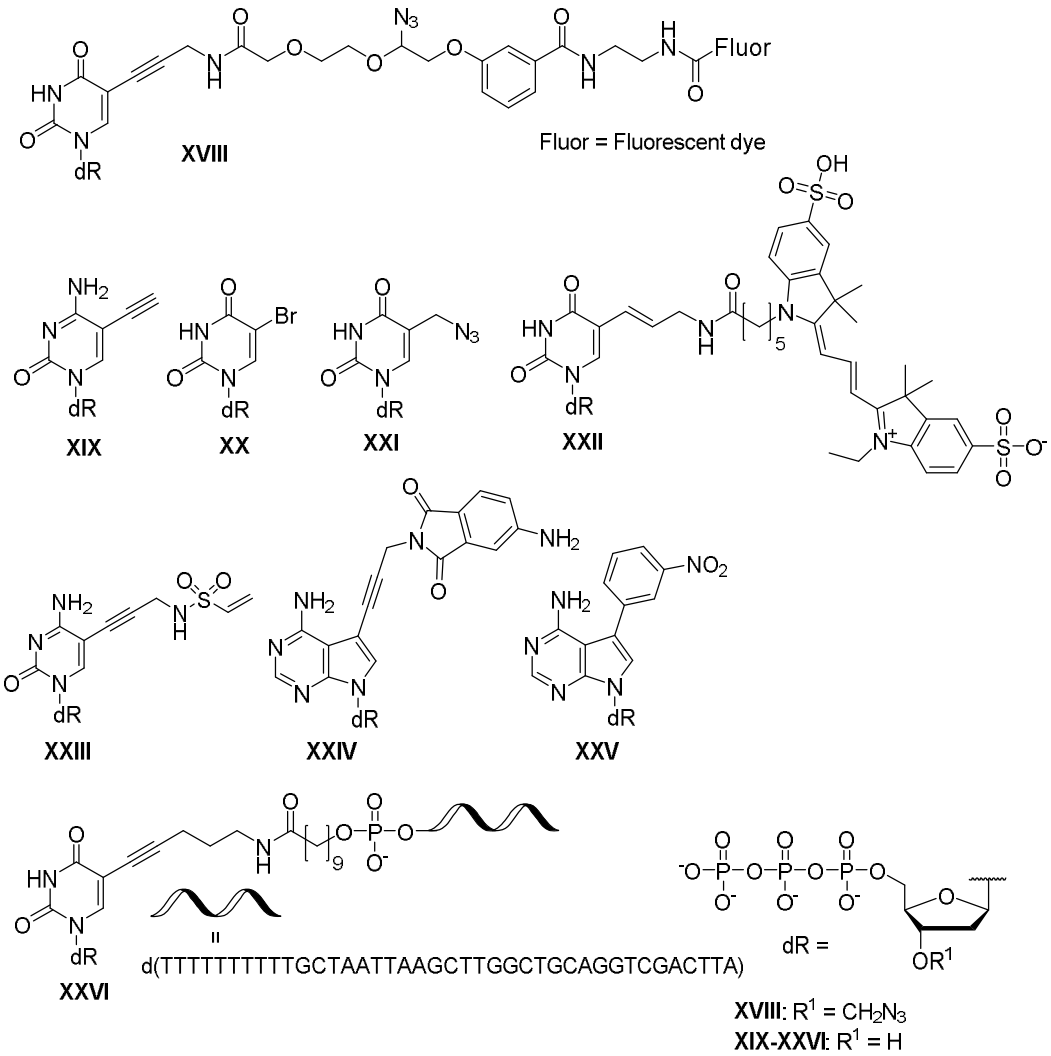


Figure 5. Diverse functionalized nucleotide triphosphates accepted by the DNA polymerases.

Several crystal structures of polymerases in complex with primer, template DNA and modified nucleoside triphosphates are available. They enable to elucidate some common features useful for understanding of incorporation of modified nucleotides.[47] However, the elongation of the DNA strand by polymerase is very complex and dynamic process, thus the simple prediction whether the modified nucleotide is a good substrate for the certain polymerase is now impossible and the screening of several polymerases needs to be done in order to find the suitable one. Also not all of the modified nucleotides which are successfully incorporated in PEX are good enough to be sufficient substrate for polymerase during PCR.

For many applications the competition for incorporation by DNA polymerase of modified triphosphates with natural counterparts is inevitable e.g. DNA *in vivo* labelling, chromatin labelling and microarrays. Marx and co-workers reported[48] kinetic and structural studies of competitive single nucleotide incorporations of modified triphosphates, results of these competitive incorporations were analysed by PAGE on which the oligonucleotides with modified nucleotide on the 3'- end has a different mobility than unmodified oligonucleotide. They also found out most of the modified nucleoside triphosphates tested were worse substrates for DNA polymerases than natural triphosphates with one exception for 7-[(hydroxydecanoyl)aminopentynyl]-7-deaza-dATP[47b] for KlenTaq DNA polymerase.

1.3.2.3 Enzymatic Synthesis of Nucleic Acids in Primer Extension Reaction and PCR

Enzymatic synthesis is used frequently to incorporate the triphosphate of a modified nucleoside onto the 3' -end of a primer (typically 12–25 nucleotide) annealed to a longer template (both chemically-synthesised) in a PEX.[20,22,49] Most commonly, this method is used for a synthesis of oligonucleotides which have less than 80 bp. The primer is usually labeled on the 5' -end by ^{32}P-phosphate or by fluorescent probe used to visualize the reaction. During the extension of a primer the incoming modified nucleoside triphosphate is tested whether is a suitable substrate for a DNA polymerase and if the addition of the following nucleotide is not restricted due to the modification (**Figure 6**).

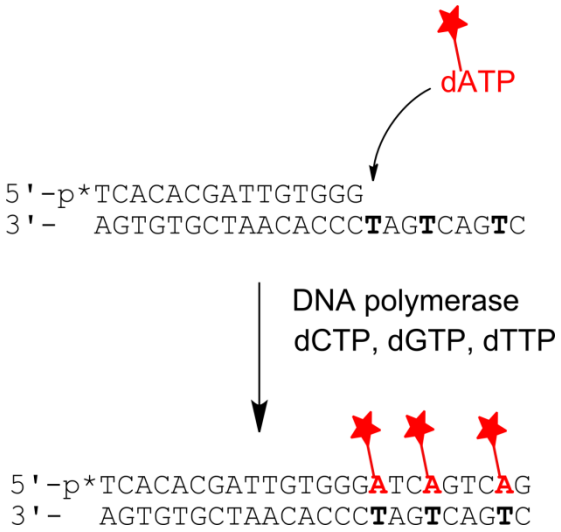


Figure 6. Primer extension experiment.

If DNA or gene of interest is not achievable by PEX or larger amount is required for further application then the synthesis is processed by use of polymerase chain reaction (PCR).[50] The target duplex DNA is denatured by heat and next annealed to the primers. The primers are then elongated in the first cycle by DNA polymerase yielding the twice amount of the starting DNA for a second round (**Figure 7**). The requirements for the successful amplification is a thermostable DNA polymerase, the optimized primer sequences and annealing temperature, sufficient extension time and number of cycles. The crucial point of the amplification with modified nucleotides are efficiency and fidelity of the incorporation by DNA polymerase.

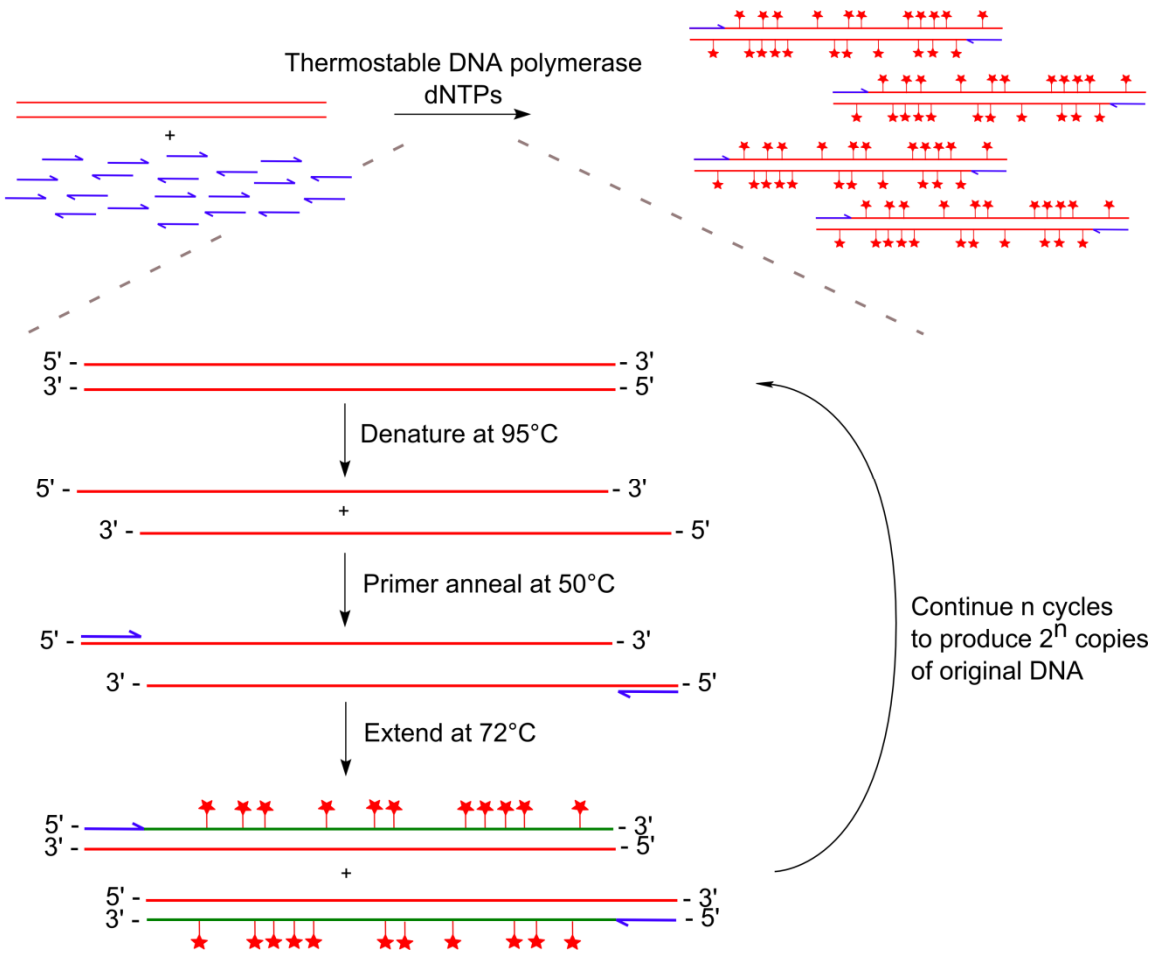


Figure 7. Steps typically involved in the polymerase chain reaction.

Recently, the Hocek group developed novel methodologies for preparation of short base modified oligonucleotides[51] and oligonucleotides with single modification at a specific position for redox labeling.[52] Enzymatic synthesis of short (10–22 nt) base-modified oligonucleotides was developed by nicking enzyme amplification reaction (NEAR) using Vent(exo-) polymerase, Nt.BstNBI nicking endonuclease, and a modified deoxyribonucleoside triphosphate derivative. This approach could be used for preparation of fluorescently labeled primers or ethynylated primers for labeling of PCR products. Preparation of single-nucleobase redox-labeled oligonucleotides was based on polymerase incorporation of a single modified nucleoside triphosphate followed by PEX with natural dNTPs or PEX with a biotinylated one-nucleotide overhang template, magnetoseparation and the second PEX with a full-length template. Both methods are useful for preparation of base modified DNAs which are not achievable with simple PEX or PCR.

1.4 Protein-Nucleic Acid Interactions

DNA duplex in B-form consists of the two anti-parallel polynucleotide strands that wind around each other with striking helical character. Base-pairs lie in roughly right angle to the helical axis and the sugar-phosphate units forming repetitive linkages. The major and minor grooves arise from the helical geometry and they are the main feature in the recognition by proteins.[1b] The major groove is wide and accessible to accommodate α-helix of proteins. Although the minor groove is narrower, it is suitable for insertion of the single peptide chains. The edges of the base-pairs in major and minor groove are accessible for hydrogen-bond donors and acceptors for direct recognition with proteins. On the other hand, indirect readout of DNA sequence is also important and is performed by shape change and deformability of DNA during complexation with protein.

1.4.1 Methyltransferases

DNA methylation of 5 position of cytosine bases occurs in both prokaryotic and eukaryotic cells but the biological functions are fundamentally different. In bacteria the DNA methylation has central role in host restriction of phage DNA while in eukaryotes the methylation of DNA is an epigenetic event. The methylated genes are silenced and thus not transcribed.

In eukaryotes the DNA is methylated at cytidines followed by guanidines also called as CpG sites. DNA modification typically correspond to long-term epigenetic memory: once methylated, genomic DNA remains methylated through generations. [53] There are two types of methyltransferases: de novo enzymes that would establish methylation patterns at CpG sites early in development and maintenance enzymes that would preserve methylation patterns during cell division by specific methylation of hemimethylated CpG dinucleotides produced by DNA replication (**Figure 8**).

The DNA methyltransferases access the target nucleobase by flipping it out of the DNA duplex into its active site[54] (**Figure 9**). This mechanism is common for both prokaryotic and eukaryotic methyltransferases and for certain DNA repair enzymes.[55]

The mechanism of cytosine methylation is illustrated in **Scheme 7**. DNA methyltransferases use a Michael addition in which the conserved cytosine residue is deprotonated to the thiolate anion and acts as a strong nucleophile. The cysteine thiolate attacks the C-6 atom of cytosine in a conjugate addition reaction, and a covalent bond is

formed between the cysteine sulfur atom and the cytosine C-6 atom. The negative charge on cytosine is stabilized by interaction with a glutamate residue. Nucleophilic attack then takes place on the methyl group of *S*-adenosyl-L-methionine (SAM), which is converted to *S*-adenosyl-L-homocysteine (AdoHcy). Finally, β-elimination occurs across the C-5 and C-6 bond, releasing the enzyme.[54] In the mechanism of methyltransferases-catalyzed methylation of cytosine a base is required to deprotonate the cysteine to form the thiolate. It is proposed that the base involved in this reaction is a DNA phosphate group via a bridging water molecule. Therefore, DNA acts as both a substrate and cofactor.[56]

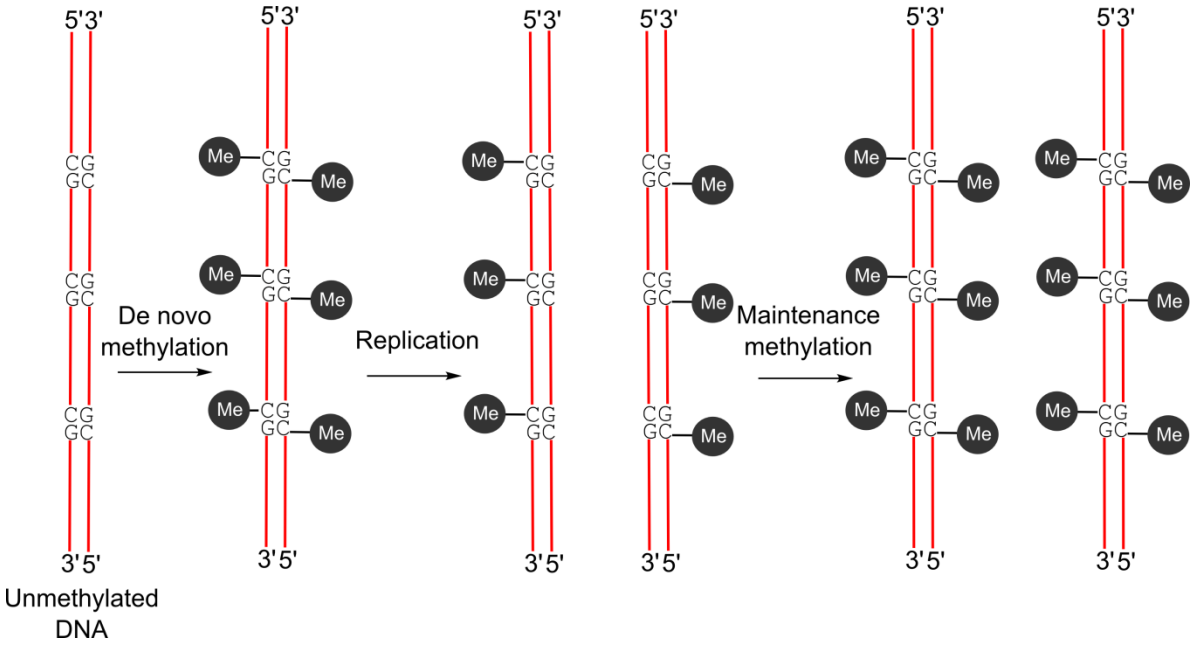


Figure 8. Schematic representation of the two types of DNA methylation: de novo and maintenance methylation of DNA.

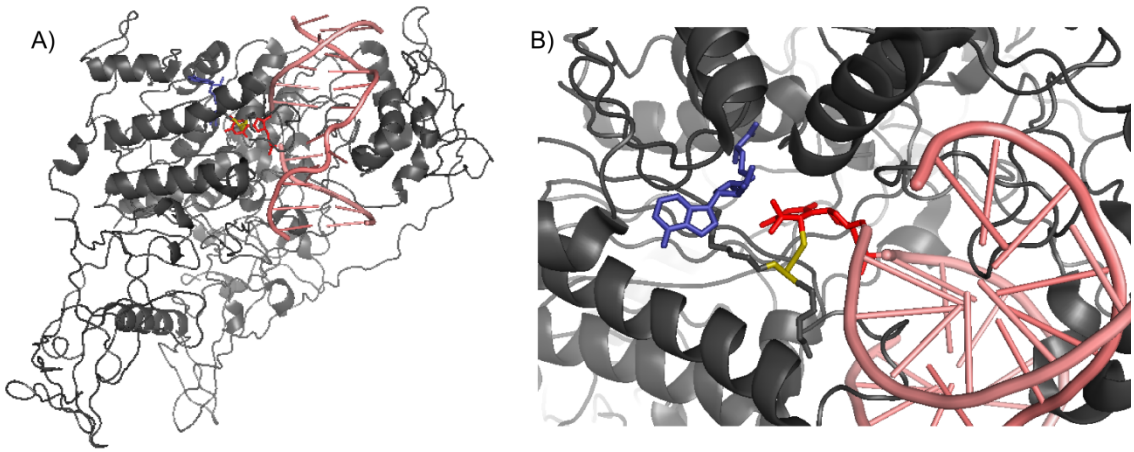
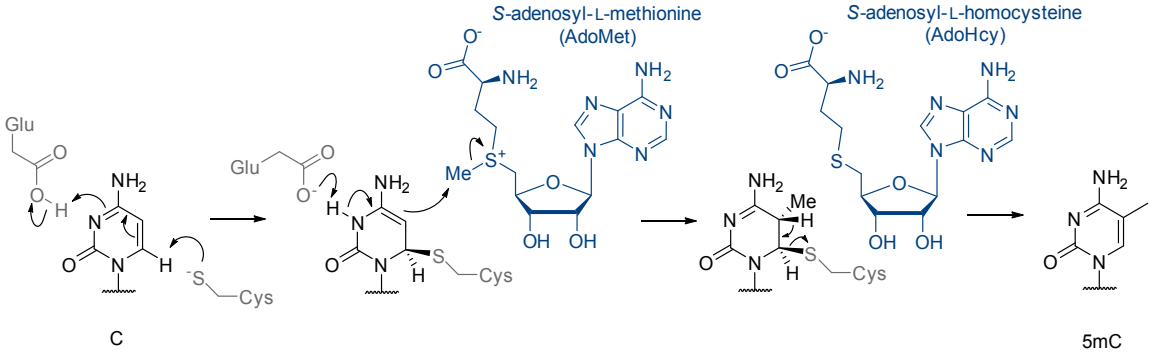


Figure 9. Crystal structure of Dnmt1 methyltransferase covalently bound to 5-fluorocytidine modified DNA in complex with *S*-adenosyl-L-homocysteine – the overall structure A) and zoom in active site B).[57]



Scheme 7. Mechanism of methylation of cytosine residue in DNA catalyzed by DNA methyltransferase.

Nucleoside inhibitors of DNA methylation[58] 5-aza-2'-deoxycytidine[59] (Decitabine), 5-azacytidine[59] (Vidaza), 1-(β-D-ribofuranosyl)-1,2-dihydropyrimidin-2-one[60] (Zebularine) and 5-fluoro-2'-deoxycytidine[61] (5FdC) (**Figure 10**) are first converted into the active triphosphates form by different kinases and then incorporated into the DNA (except Vidaza which is first converted by ribonucleotide reductase to it's 2'-deoxyanalogue). They all form, after being incorporated into DNA during replication, a covalent complex with the methyltransferase that cannot be released, since the β-elimination process is hampered. Vidaza is also incorporated into RNA and interferes with protein translation.

The crystal structure of complex of the 5FdC modified hemimethylated DNA with Dnmt1 gave us the mechanistic insight into the maintenance DNA methylation.[57] Similarly, it is possible to use oligonucleotides containing 5FdC as a probe for indentification of DNA

methyltransferases in whole-cell extract,[62] because the covalent complex of 5FdC-DNA methyltransferase withstands strongly denaturing conditions.[63]

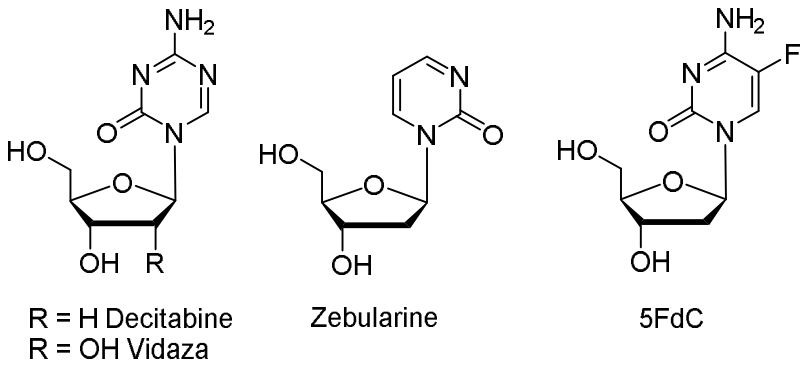


Figure 10. Methyltransferase inhibitors 5-fluoro-2'deoxycytidine (5FdC), zebularine and 5-azacytidine.

1.4.2 Restriction Endonucleases

Restricition endonucleases (RE) have a key function in protection of prokaryotes against the bacteriophage DNA. Function of these REs is tightly coupled with DNA methyltransferases which add methyl group into the 5 position of cytidine and N^6 position of adenosine. This modified DNA is then not digested by RE while DNA from host pathogen is digested.[64]

Currently, >19 000 REs are listed on REBASE.[65] REases are classified into four main groups, Type I, II, III and IV, with subdivision for convenience. Main features and requirements of REs are summarized in **Table 2.** Almost all require divalent metal cation such as Mg^{2+} for activity and cut the DNAs phosphodiester bound yielding 5'-phosphoryl and 3'-hydroxyl groups. Some REs require ATP and/or cofactor SAM.

Table 2. Characterization of the four types of restriction enzymes.

Type	Type I	Type II	Type III	Type IV
Feature	Oligomeric REs and MTs complex Require ATP hydrolysis for restriction Cleave variably, often far from recognition site	Separate REs and MTs or combined Cleave within or at fixed positions close to recognition site Many different subtypes	Combine REs + MTs complex ATP required for restriction Cleave at fixed position outside recognition sequence	Methylation-dependent REs Cleave at variable distance from recognition site Cleave m6A, m5C, hm5C and/or other modified DNA
Example	EcoKI	EcoRI	EcoP1I	No 'typical' example
REBASE	104 enzymes	3938 enzymes	21 enzymes	18 enzymes

Type II REs represent the largest group of well studied and characterized enzymes.[66] These are used in recombinant DNA technology and have advanced our understanding of protein-DNA interactions. The interactions of type II REs with their recognition sequence have many common features. These REs act mainly as homodimers or homotetramers and recognize usually short palindromic sequence which they cut to generate blunt or sticky ends (**Figure 11**). They interact symmetrically with a minimum of 10 nucleotide pairs, from which each

base pair offers a unique pattern in major and minor groove that enables the base-recognition by 'direct readout' , an additional factor is the distortion of DNA ('indirect readout').

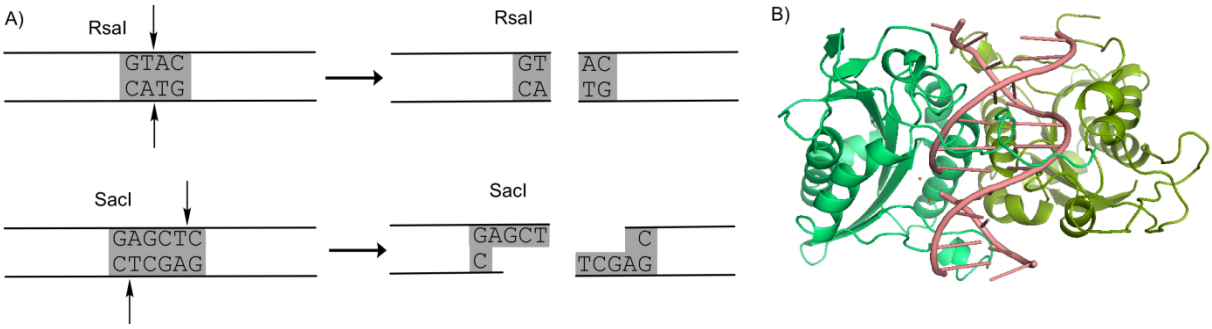


Figure 11. Type II restriction enzymes cleavage properties: A) RsaI cleavage produces blunt ends whereas SacI cleavage produces sticky ends. B) Crystal structure of BamHI homodimer with DNA after the cleavage of the one strand of dsDNA.

Digestion of modified DNA by type II REs has been studied only scarcely. There are some reports on 8-modified adenines and guanines,[67] 7-deazaadenine[68] and 7-deazaguanine[69] and unnatural base-pairs incorporated[70] in recognition sequence for REs. In all cases the modifications cause the inhibition of the cleavage reaction. Only the recent systematic study of interactions of REs with DNA modified with 7-substituted 7-deazaadenines and 5-substituted pyrimidines published by Hocek and co-workers revealed that any modification of cytosine residue results in the full inhibition of tested REs while some modifications of 7-deazaadenines and thymines are tolerated by certain REs (**Table 3**).[12,71]

Table 3. Cleavage efficiency of restriction endonucleases to cleave modified DNAs. (7D – 7-deaza, E – ethynyl, Ph – phenyl, NO2 – 3-nitrophenyl, Br – bromo, Me – methyl)

Restricition enzyme	A^{7D}	A^{E}	A^{Ph}	A^{NO2}	A^{Br}	A^{Me}	C^{E}	C^{Ph}	U^{E}	U^{Ph}
AfeI	+/-	-	-	-	+/-	+	-	-	+	+/-
PvuII	+	-	-	-	+	+	-	-	+/-	-
RsaI	+	+	-	-	+	+	-	-	+/-	-
ScaI	-	-	-	-	+/-	+/-	+	-	+	-
PspGI	+	+	+	+	+	+	-	-	+/-	+/-
EcoRI	-	-	-	-	+	+	-	-	-	-
KpnI	+	+	-	-	+	+	-	-	+/-	-
PstI	+	-	-	-	+	+	-	-	-	-
SacI	+	+	-	-	+/-	+/-	-	-	+/-	-
SphI	+	+/-	-	-	+/-	+/-	-	-	+/-	-

2 Specific Aims of the Thesis

1. Synthesis of modified 2'-deoxyribonucleosides and nucleotides substituted in 5-position of pyrimidines and 7-position of 7-deazaadenine with another nucleobase, study on their enzymatic incorporation into DNA and investigation of these modified DNAs interactions with DNA methyltransferases.
2. Synthesis of the series of (trialkylsilyl)ethynyl-modified 7-deaza-2'-deoxyadenosines and corresponding triphosphates.
3. Development of the methodology for transient protection of DNA against the cleavage by restriction endonucleases and application of this method in gene cloning and protein expression.
4. Study of competitive incorporations of modified 2'-deoxynucleosides triphosphates in presence of their natural counterparts.

2.1 Rationale of the Specific Aims

DNA editing enzymes such as DNA glycosylases and DNA methyltransferases use the nucleotide-flipping mechanism in which the whole nucleotide is flipped out of the DNA duplex into the active site of the enzyme. We had proposed that modified nucleobases bearing another nucleobase might mimic the flipped-out state and thus interact or inhibit these enzymes. My task was to synthesize double-headed nucleotides, incorporate them enzymaticaly into the DNA and develop the assay for testing of the interactions with DNA methyltransferases.

Building upon our previous studies on interactions of restriction endonucleases with modifications in DNA major groove we had designed the series of trialkylsilyl protecting groups of the ethynyl modification of 7-substituted 7-deaza-2'-deoxyadenosines. These bulky modification would protect the recognition sequence in DNA against the cleavage by restriction endonucleases. While after deprotection of the trialkylsilyl group the ethynyl modified DNA would be cleaved. This proof-of-principle experiment which use the chemical protection group and trigger to control biochemical reaction would be very interesting for further development of similar techniques and applications in gene cloning or transcription regulation. Particulary in the gene cloning the methodology for selective protection of restriction endonuclease target sequence in the clone with several copies of this sequence would be of great advantage and circumvent the consequences caused by codon mutations,

which is usually used. My goal was to develop such chemical trigger based on (trialkylsilyl)ethynyl modified 7-deaza-2'-deoxyadenosines and apply this approach in a cloning of gene with more than one resctriction site for same restriction endonuclease.

Finally, my objective for study of competitive incorporations was that in many applications of modified dNTPs their natural counterparts are inherently present and thus compete with modified ones (e.g. in *in vivo* labeling of DNA). In adition to this fact, the competitive incorporations were studied only scarcely and could possibly yield interesting results. My task was to developed the method for analysis of competitive incorporations and study competitive incorporations of different modified 2'-deoxyribonucleoside triphosphates with several DNA polymerases. It was presumed that the modified triphosphates would be worst substrates for DNA polymerases than natural dNTPs, because of the evolution of the natural dNTPs to perfect fit into the active site of the DNA polymerases.

3 Results and Discussion

3.1 Synthesis of Double-headed Nucleosides and Nucleotides, Their Enzymatic Incorporation, and Interactions with DNA Methyltransferases[72]

DNA methyltransferases[73] and DNA glycosylases[74] play an important role in transcription regulation and DNA repair. DNA editing enzymes employ the base flipping mechanism during which the whole nucleotide is flipped out of the DNA duplex into the active site of the enzyme.[54] We envisaged that DNA containing double-headed (**Figure 12**) nucleotides might mimic this flipped out state and thus interact or inhibit these enzymes. Three small series of the double-headed nucleosides (**Figure 13**) and nucleotides were prepared. The rationale for the 5-fluorocytosine derivative is that it might employ the same mechanism as 5-fluorocytidine,[61] thus covalently and irreversibly bind to the methyltransferase which would lead to inhibition of the methylation process. Therefore, I focused on the synthesis of the double-headed 2'-deoxynucleosides and nucleotides which after incorporation into DNA might interact with DNA methyltransferases.

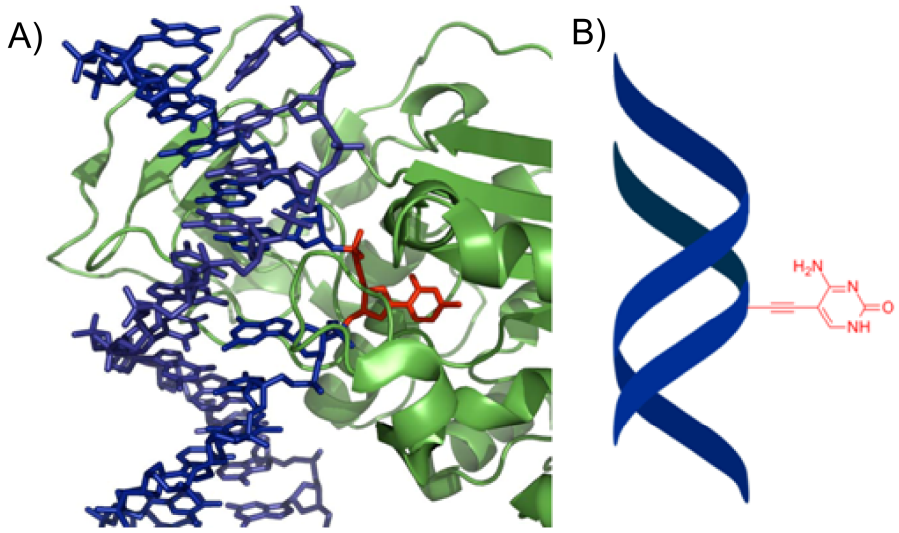


Figure 12 (A) Crystal structure of DNA methyltransferase with nucleotide flipped out of the DNA duplex.[73b] (B) Illustration of DNA bearing another cytosine base in the major groove for interactions with DNA methyltransferase.

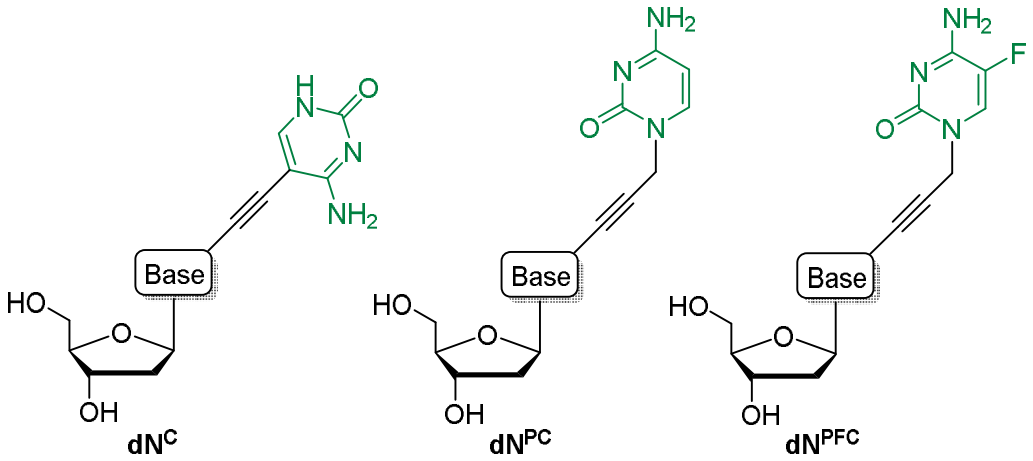
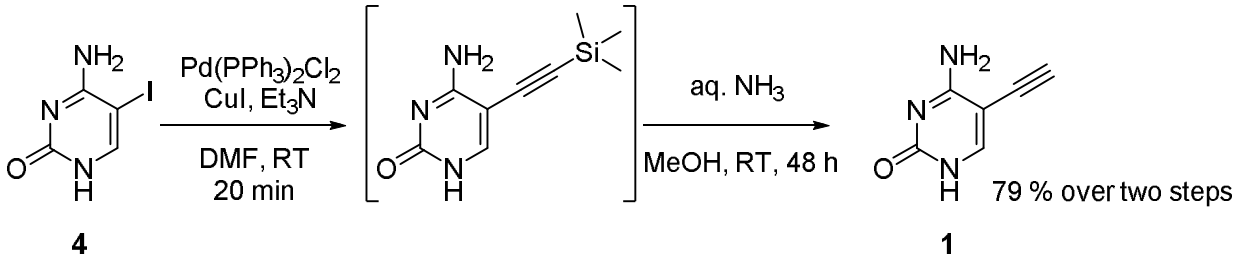


Figure 13 Double-headed 2'-deoxyribonucleosides.

3.1.1 Synthesis of Double-headed Nucleosides and Nucleotides

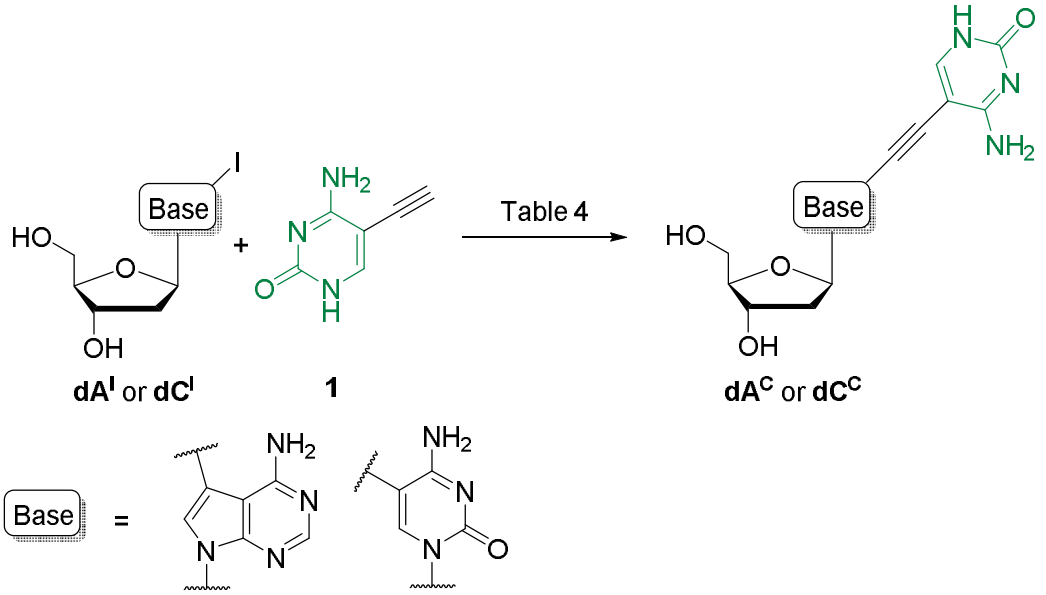
The synthetic methodology we have chosen consisted of Sonogashira cross-coupling of 5-ethynylcytosine (**1**) or 1-*N*-(prop-2-yn-1-yl)cytosine (**2**) or 5-fluoro-1-*N*-(prop-2-yn-1-yl)cytosine (**3**) with halogenated 2'-deoxyribonucleosides and nucleotides. First, the 5-ethynylcytosine series was prepared. The starting 5-ethynylcytosine[75] was synthesized by Sonogashira cross-coupling of 5-iodocytosine[76] (**4**) with trimethylsilylacetylene (TMSA) in the presence of the palladium catalyst $Pd(PPh_3)_2Cl_2$, CuI and Et_3N as a base in DMF at ambient temperature followed by the treatment with aqueous ammonia in methanol with good overall yield (79 %) (**Scheme 8**).



Scheme 8 Synthesis of 5-ethynylcytosine (**1**).

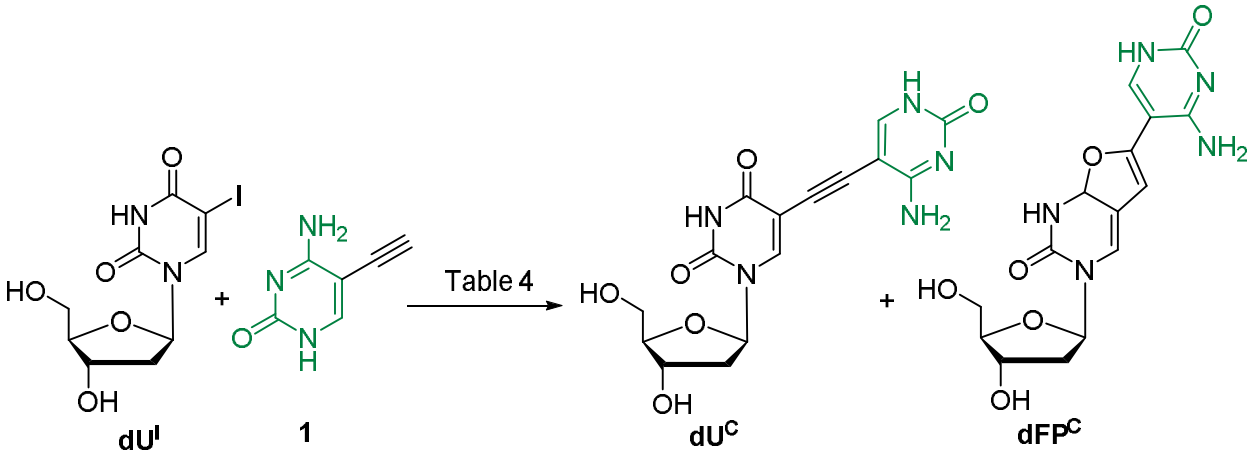
The Sonogashira cross-coupling reactions of **1** with 5-iodo-2'-deoxycytidine[14] (**dC^I**), 5-iodo-2'-deoxyuridine[14] (**dU^I**) or 7-iodo-2'-deoxy-7-deazaadenosine[15] (**dA^I**) were performed under previously optimized[20] aqueous conditions CH_3CN/H_2O 1:2 in presence of catalyst $Pd(OAc)_2$, CuI, TPPTS as a ligand and *i*Pr_2NEt as a base at 80°C for 1 h or in DMF in presence of $Pd(PPh_3)_2Cl_2$ (**Scheme 9**). The **dC^I** and **dA^I** reacted smoothly to give quantitative

conversions to the desired cytosinylethynyl derivatives **dC^C** and **dA^C**. The isolated yields 38–73 % were lowered by difficult isolation of the very polar and poorly soluble products, see **Table 4** entry 1,2,11 and 12.



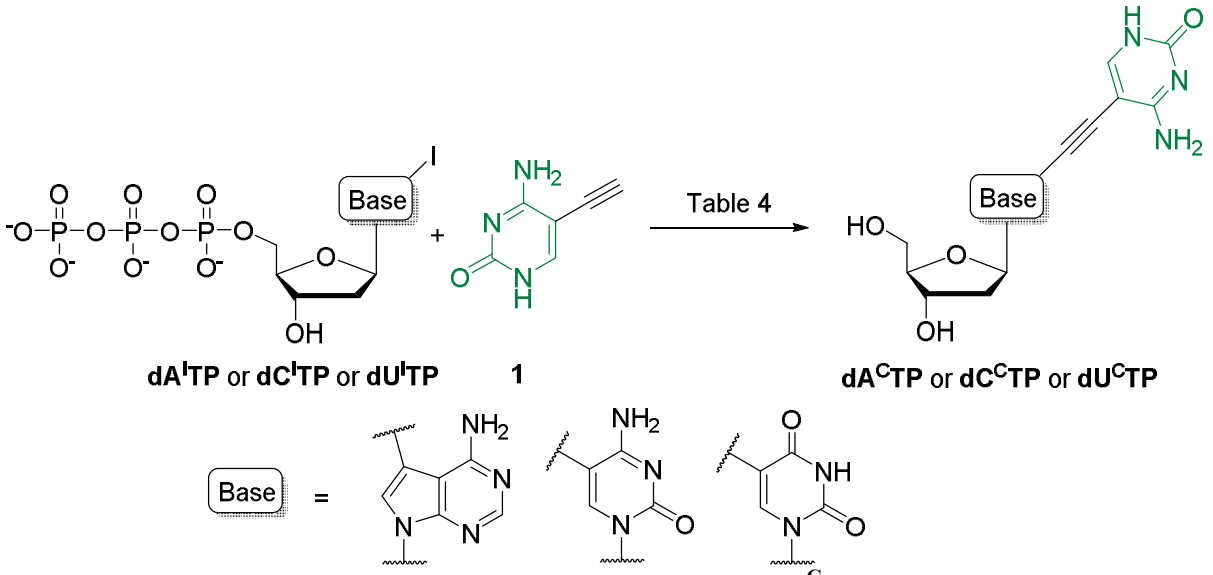
Scheme 9 Synthesis of 5-ethynylcytosine modified 7-deaza-2'-deoxyadenenine and 2'-deoxycytidine.

The attempted reaction of **dU^I** with **1** under the same conditions led to a mixture containing the undesired furopyrimidine[77] **dFP^C** as the main product formed by the Cu-mediated heterocyclization (**Scheme 10** and **Table 4**, entry 7 and 8). Only the use of Et_3N as a base and lower temperature 50 °C gave the non-cyclized desired acetylene linked nucleoside **dU^C** in 40 % in CH_3CN/H_2O or 60 % in DMF (**Table 4**, entry 5 and 6).



Scheme 10 Synthesis of 5-ethynylcytosine modified 2'-deoxyuridine with undesired furopyrimidine **dFP^C** product.

Analogous strategy of Sonogashira cross-coupling was used for preparation of the cytosinylethynyl modified **dN^{X}TP**s (**Scheme 11**). Generally, aqueous cross-coupling reactions of dNTPs suffer from lowering of the yields of modified **dN^{X}TP**s due to partial hydrolysis of the starting and final dNTPs to nucleoside diphosphates.[20] Therefore, the reactions must be optimized and performed in short times (less than 1 h). Typical isolated yields of modified **dN^{X}TP**s are 25-50 % which are sufficient, since for most of the biochemical applications only miligram amounts of **dN^{X}TP**s are needed. Thus the reactions of halogenated nucleotides **dC^{I}TP**,[49b] **dU^{I}TP**[78] or **dA^{I}TP**[79] with **2** were performed under the same conditions as for the nucleosides in aqueous phase. The pyrimidine nucleotides **dC^{I}TP** and **dU^{I}TP** reacted well at 80 °C to give the corresponding modified **dC^{C}TP** and **dU^{C}TP** in relatively good 35 and 31 % yields, respectively (**Table 4**, entry 15 and 18). No formation of furopyrimidine was observed in reaction of **dU^{I}TP**. Both **dN^{C}TPs** were isolated in sufficient amounts and purity by semi-preparative HPLC. On the other hand, reaction of **dA^{I}TP** was much slower and at 80 °C it did not lead to full conversion within 1 h. Therefore, the temperature was increased to 105 °C under which the cross-coupling (but unfortunately also the hydrolysis) was more efficient and the desired **dA^{C}TP** was isolated in moderate yield of 17 % (**Table 4**, entry 19).



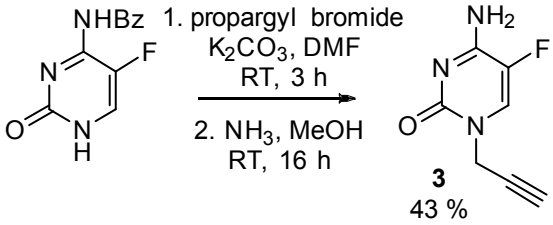
Scheme 11 Synthesis of the 5-ethynylcytosine modified **dN^{C}TP**s.

Table 4. Cross-coupling reaction conditions of iodonucleos(t)ides with alkynylcytosine derivatives **1**, **2** and **3**.

Entry	dN^I/dN^ITP	Catalyst	Ligand/base	Solvent	Conditions [°C, h]	Product	Yield (%)
1	dC^I	$Pd(OAc)_2$	TPPTS/iPr_2NEt	CH_3CN/H_2O	80 °C, 1 h	$\mathbf{dC^C}$	49
2	dC^I	$Pd(PPh_3)_2Cl_2$	Et_3N	DMF	80 °C, 1 h	$\mathbf{dC^C}$	73
3	dC^I	$Pd(OAc)_2$	TPPTS/iPr_2NEt	CH_3CN/H_2O	80 °C, 0.5 h	$\mathbf{dC^{PC}}$	30
4	dC^I	$Pd(OAc)_2$	TPPTS/iPr_2NEt	CH_3CN/H_2O	80 °C, 0.1 h	$\mathbf{dC^{PFC}}$	95
5	dU^I	$Pd(OAc)_2$	TPPTS/iPr_2NEt	CH_3CN/H_2O	50 °C, 1 h	$\mathbf{dU^C}$	40[a]
6	dU^I	$Pd(PPh_3)_2Cl_2$	Et_3N	DMF	50 °C, 1 h	$\mathbf{dU^C}$	60[b]
7	dU^I	$Pd(OAc)_2$	TPPTS/iPr_2NEt	CH_3CN/H_2O	80 °C, 1 h	$\mathbf{dFP^C}$	30
8	dU^I	$Pd(PPh_3)_2Cl_2$	Et_3N	DMF	80 °C, 1 h	$\mathbf{dFP^C}$	78
9	dU^I	$Pd(OAc)_2$	TPPTS/iPr_2NEt	CH_3CN/H_2O	50 °C, 1 h	$\mathbf{dU^{PC}}$	95
10	dU^I	$Pd(OAc)_2$	TPPTS/Et_3N	CH_3CN/H_2O	50 °C, 0.5 h	$\mathbf{dU^{PFC}}$	69
11	dA^I	$Pd(OAc)_2$	TPPTS/iPr_2NEt	CH_3CN/H_2O	80 °C, 1 h	$\mathbf{dA^C}$	38
12	dA^I	$Pd(PPh_3)_2Cl_2$	Et_3N	DMF	80 °C, 1 h	$\mathbf{dA^C}$	51
13	dA^I	$Pd(OAc)_2$	TPPTS/iPr_2NEt	CH_3CN/H_2O	80 °C, 0.5 h	$\mathbf{dA^{PC}}$	61
14	dA^I	$Pd(OAc)_2$	TPPTS/iPr_2NEt	CH_3CN/H_2O	80 °C, 1 h	$\mathbf{dA^{PFC}}$	58
15	dC^ITP	$Pd(OAc)_2$	TPPTS/iPr_2NEt	CH_3CN/H_2O	80 °C, 1 h	$\mathbf{dC^CTP}$	35
16	dC^ITP	$Pd(OAc)_2$	TPPTS/iPr_2NEt	CH_3CN/H_2O	80 °C, 1.5 h	$\mathbf{dC^{PC}TP}$	30
17	dC^ITP	$Pd(OAc)_2$	TPPTS/iPr_2NEt	CH_3CN/H_2O	80 °C, 1 h	$\mathbf{dC^{PFC}TP}$	80
18	dU^ITP	$Pd(OAc)_2$	TPPTS/iPr_2NEt	CH_3CN/H_2O	80 °C, 1 h	$\mathbf{dU^CTP}$	31
19	dA^ITP	$Pd(OAc)_2$	TPPTS/iPr_2NEt	CH_3CN/H_2O	105 °C, 1 h	$\mathbf{dA^CTP}$	17
20	dA^ITP	$Pd(OAc)_2$	TPPTS/iPr_2NEt	CH_3CN/H_2O	80 °C, 1 h	$\mathbf{dA^{PC}TP}$	80
21	dA^ITP	$Pd(OAc)_2$	TPPTS/iPr_2NEt	CH_3CN/H_2O	80 °C, 1 h	$\mathbf{dA^{PFC}TP}$	43

[a] 11 % of $\mathbf{dFP^C}$ was also isolated. [b] Ca. 40 % of unreacted $\mathbf{dU^I}$ remained.

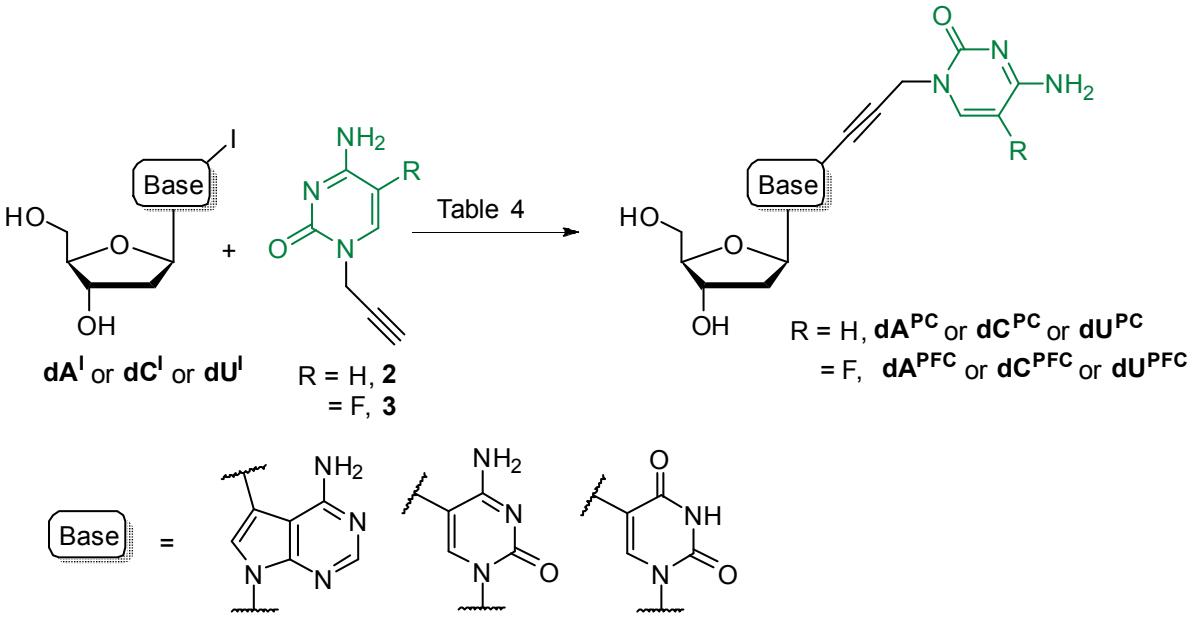
Synthesis of the propargyl linked double-headed 2'-deoxyribonucleosides and nucleotides was similar to the previous approach. The starting 1-*N*-(prop-2-yn-1-yl)cytosine (**2**) was prepared according the literature procedure[80] using acetyl protecting group to protect the amino group as was reported. In contrast, the acetyl group was not stable on 5-fluoro derivative **3**, thus more stable benzoyl protecting group was used.[81] The alkylation illustrated in **Scheme 12** was performed in the presence of K_2CO_3 as a base in DMF with propargyl bromide. Benzoyl protecting group was removed by methanolic ammonia to give 5-fluoro-1-*N*-(prop-2-yn-1-yl)cytosine (**3**) in 43 % yield.



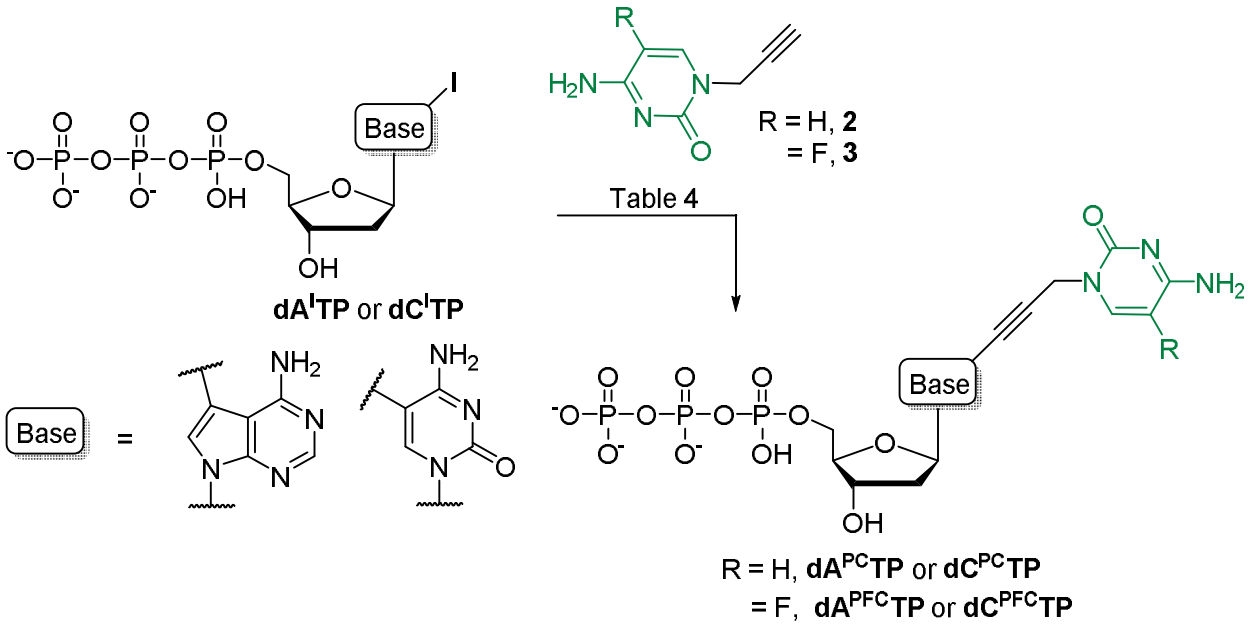
Scheme 12 Synthesis 5-fluoro-1-*N*-(prop-2-yn-1-yl)cytosine.

Synthesis of propargyl linked double-headed nucleosides was performed using an agueous Sonogashira cross-coupling reaction of **2** or **3** with a halogenated nucleoside in the presence of *i*Pr_2NEt or Et_3N as a base, $Pd(OAc)_2$ as a catalyst and TPPTS as a ligand in CH_3CN/H_2O 1:2 (**Scheme 13** and **Table 4**, entry 3, 4, 9, 10, 13 and 14). The resulting nucleosides were obtained in good to very good yields, in comparison with the series of double-nuclebase nucleosides with acetylene linker, because of their better solubility which make their isolation easier and yields higher.

Modified 2'-deoxyadenosine and 2'-deoxycytidine triphosphates were prepared under the same conditions as corresponding nucleosides (**Scheme 14**). The triphosphates were obtained in good yields (**Table 4**, entry 13, 14, 16, 17, 20 and 21). In conclusion, new series of double-headed nucleosides and nucleoside triphosphates containing propargyl linker and cytosine or 5-fluoro cytosine as a second base was prepared in good yields.



Scheme 13. Synthesis of 1-*N*-(prop-2-yn-1-yl)cytosine modified 2'-deoxyribonucleosides.



Scheme 14. Synthesis of 1-*N*-(prop-2-yn-1-yl)cytosine modified 2'-deoxyribonucleoside triphosphates.

3.1.2 Enzymatic Incorporation of Double-Headed Nucleosides Triphosphates

All the ethynylcytosine modified dNTPs (**dC^{C}TP**, **dU^{C}TP** and **dA^{C}TP**) were tested as substrates for *Pyrococcus kodakaraensis* (KOD XL) DNA polymerase in PEX. We have designed several different templates (**Table 5**) with different number of modifications in the resulting DNA. Each experiment was performed with one of the modified **dC^{C}TP**, **dU^{C}TP** or **dA^{C}TP** and three additional natural dNTPs. At first, we incorporated a single modification **dC^{C}**, **dU^{C}** or **dA^{C}** followed by three natural G using 19-mer template temp^{1C}, temp^{1T} and temp^{1A}, respectively (**Figure 14**). All the **dN^{C}TP**s were very good substrates giving full-length (19-nt) products. In the negative control experiment using temp^{1A} (in absence of **dA^{C}TP**), we have observed some misincorporation of dGTP at higher concentration (0.2mM) and therefore the concentration of dGTP needed to be lowered to 10μM. Next, we performed PEX using temp^{4base} as a template with four incorporations of one modification (**dC^{C}**, **dU^{C}** or **dA^{C}**) at separate positions. Again, all three modified **dN^{C}TP**s were successfully incorporated (**Figure 15A**). Finally, we challenged the polymerase with a PEX including four in-line incorporations of a modified **dN^{C}TP** (**dC^{C}TP**, **dU^{C}TP** or **dA^{C}TP**) using templates temp^{4N} (**Figure 15B**). Even this experiment was successful for all modifications showing that the

dN^{C}TP are excellent substrates for KOD polymerase and are incorporated even into difficult sequences at adjacent positions.

Table 5 Primers and templates used for PEX

primer	5'-CATGGGCGGCATGGG-3'
primer^{80mer}	5'-CATGGGGCTCATGGG-3'
temp^{1C}	5'-CCCGCCCATGCCGCCCATG-3'
temp^{1T}	5'-CCCACCCATGCCGCCCATG-3'
temp^{1A}	5'-CCCTCCCATGCCGCCCATG-3'
temp^{4base}	5'-CTAGCATGAGCTCAGTCCCATGCCGCCCATG-3'
temp^{4C}	5'-TCATCATCATAGGGGCCCATGCCGCCCATG-3'
temp^{4T}	5'-CGTCGTCGTCGAAAACCCATGCCGCCCATG-3'
temp^{4A}	5'-CAGCAGCAGCATTTTCCCATGCCGCCCATG-3'
temp^{80mer}	5'-ATACAGTATCCCATCACATGCTCATACAAGTCTACCAGC
	ATACCAGTCTATACCTACAGCAGCTACCCATGAGCCCCATG-3'

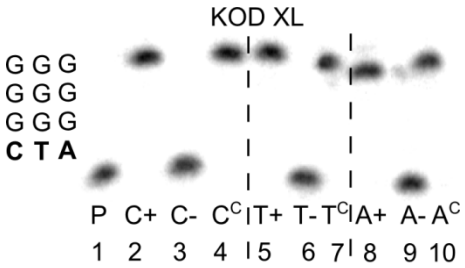


Figure 14. PAGE analysis of PEX with temp^{1C}, temp^{1T} and temp^{1A}. Lane 1: primer (5′-^{32}P-end labeled primer); lane 2, 5 and 8: natural dNTPs; lane 3: dTTP, dATP, dGTP; lane 4: **dC^{C}TP**, dATP, dGTP, dTTP; lane 6: dATP, dCTP, dGTP; lane 7: dCTP, **dU^{C}TP**, dATP, dGTP; lane 9: dCTP, dTTP, dGTP; lane 10: **dA^{C}TP**, dTTP, dCTP, dGTP.

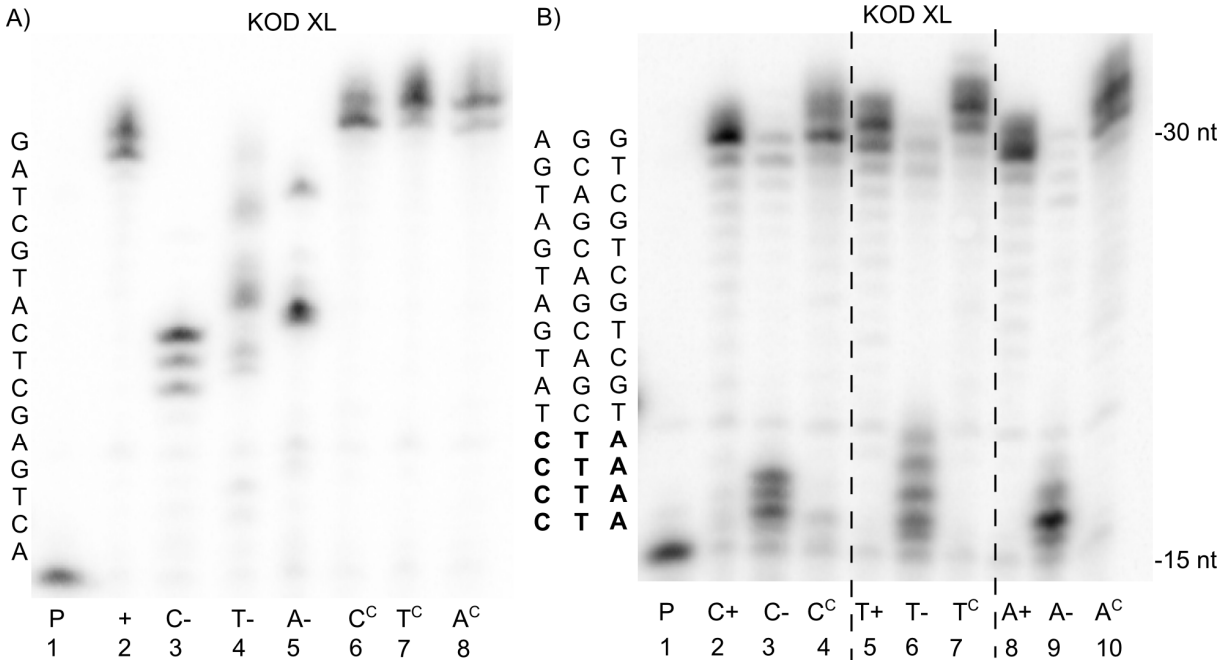


Figure 15. PAGE analysis of PEX with temp^{4base} (A) and temp^{4C}, temp^{4T}, temp^{4A} (B). (A) Lane 1: primer (5′-^{32}P-end labeled primer); lane 2: natural dNTPs; lane 3: dTTP, dATP, dGTP; lane 4: dATP, dCTP, dGTP; lane 5: dCTP, dTTP, dGTP; lane 6: **dC^{C}TP**, dATP, dGTP, dTTP; lane 7: dCTP, **dU^{C}TP**, dATP, dGTP; lane 8: **dA^{C}TP**, dTTP, dCTP, dGTP. (B) Lane 1: primer (5′-^{32}P-end labeled primer); lane 2, 5 and 8: natural dNTPs; lane 3: dTTP, dATP, dGTP; lane 4: **dC^{C}TP**, dATP, dGTP, dTTP; lane 6: dATP, dCTP, dGTP; lane 7: dCTP, **dU^{C}TP**, dATP, dGTP; lane 9: dCTP, dTTP, dGTP; lane 10: **dA^{C}TP**, dTTP, dCTP, dGTP.

In the case of 1-*N*-(prop-2-yn-1-yl)cytosine derivatives (**dN^{PC}TP** an **dN^{PFC}TP**) the PEX was performed with longer 80-mer template with six modifications in the resulting DNA. Similar 80-mer template was then used for testing of interactions of such modified DNA with DNA methyltransferase. We used three different DNA polymerases Pwo, Vent(exo-) and KOD XL. All these polymerases were shown to be able to incorporate both **dC^{PC}TP** and **dC^{FC}TP**. The KOD XL and Vent(exo-) we observed the primer in negative control experiment was partially elongated to the full length which is done by the misincorporation, for the Pwo under same reaction conditions no full length product in negative control was observed (**Figure 16**). Thus the Pwo polymerase was used in further experiments with 1-*N*-(prop-2-yn-1-yl)cytosine derivatives.

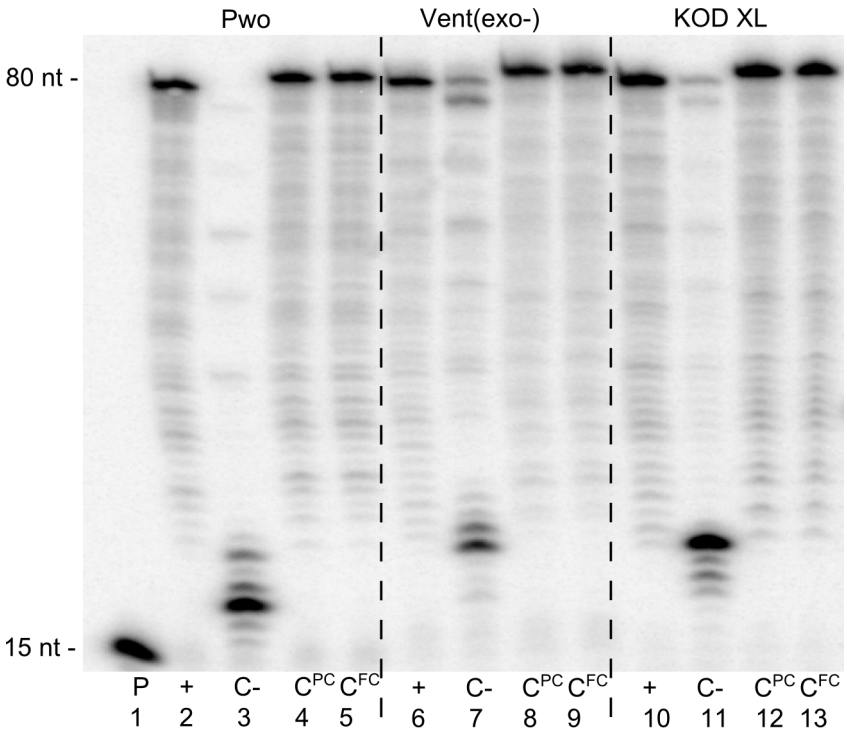


Figure 16. PAGE analysis of PEX with $temp^{80mer}$ with Pwo (lane 2–4), Vent(exo-) (lane 6–9) and KOD XL (lane 10–13). dNTPs composition: Lane 1: primer (5′-^{32}P-end labeled primer); lane 2, 6 and 10: natural dNTPs; lane 3, 7 and 11: dTTP, dATP, dGTP; lane 4, 8 and 12: **$dC^{PC}TP$**, dATP, dGTP, dTTP; lane 5, 9 and 13: **$dC^{FC}TP$**, dATP, dGTP, dTTP.

These results prompted us to attempt PCR experiments with ethynylcytosine modified **$dN^{C}TP$**s analogues with a 98-mer template and 20- and 25-mer primers in the presence of KOD XL polymerase. The formation of functionalized DNA after 30 PCR cycles in the presence of primers, a modified **$dN^{C}TP$** (**$dC^{C}TP$**, **$dU^{C}TP$** or **$dA^{C}TP$**) and three additional natural dNTPs was analyzed by agarose gel electrophoresis. PCR reaction proceeded smoothly to give full-length (97-nt) products. We also extended our study to longer DNA strand (297-mer template). In this experiment we found that **$dU^{C}TP$** was somewhat poorer substrate for the KOD polymerase than the other tested **$dN^{C}TP$**s (**$dC^{C}TP$** and **$dA^{C}TP$**) but all of them gave the desired product (**Figure 17**).

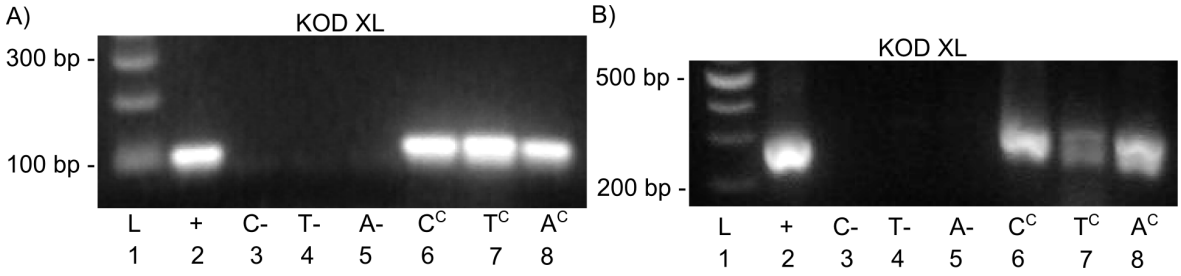


Figure 17. Agarose gel analysis of PCR experiments: A) 98-mer template, B) 297-mer template. Lane 1: 100 bp DNA ladder; lane 2: natural dNTPs; lane 3: dTTP, dATP, dGTP; lane 4: dATP, dCTP, dGTP; lane 5: dCTP, dTTP, dGTP; lane 6: **dC^{C}TP**, dATP, dGTP, dTTP; lane 7: dCTP, **dU^{C}TP**, dATP, dGTP; lane 8: **dA^{C}TP**, dTTP, dCTP, dGTP.

3.1.3 Interactions of the Double-Headed Nucleobases Modified DNA with DNA Methyltransferase M.SssI

Methyltransferase M.SssI[82] add the methyl group at 5 position of 2'-deoxycytidine within the CpG sites. Before the addition of the methyl group it flips out the cytosine 2'-deoxyribonucleotide from DNA duplex[54] which we propose could be mimicked by double-headed nucleobases incorporated into modified DNA. The mechanism of the methylation also employ the covalent M.SssI-DNA complex.[54] Because of the second nucleobase in the major groove it might be possible to prepare a covalent methyltransferase-DNA complex and thus inhibit the methylation of the CpG sites.

First, I developed an assay for testing of the inhibition of M.SssI by double-headed nucleobases in DNA. The **dC^{PC}** and **dC^{FC}** double-headed nucleosides were tested in this assay. The primer and template for PEX was designed to contain five CpG sites, two within the primer sequence (thus cannot be modified) and three which are only in the template sequence. The idea was that if the covalent complex of M.SssI and modified DNA is formed, it would decrease the amount of the methyltransferase which can methylate the CpG sequence within the primer. The CpG site within the primer is included in the recognition sequence for the restriction endonuclease AfeI. The AfeI is sensitive to DNA methylation and does not cleave DNA with 5mC in the CpG sites. The method is illustrated in **Figure 18**, primer and template are listed in **Table 6**.

The **dC^{PC}TP** and **dC^{PFC}TP** were incorporated into the DNA by Pwo DNA polymerase. Resulting modified DNA was treated with M.SssI in presence of the 3.2 μM SAM and then with AfeI. PAGE analysis revealed that DNA modified with **dC^{PC}** and **dC^{PFC}** do not significantly inhibit the methylation by M.SssI. On the other hand the assay was found to be

suitable for testing of the inhibition of M.SssI. Further screening of double-headed nucleobases (e.g. with different linkers and/or modification of the "second" bases) of inhibition of DNA methylation is out of the scope of this thesis.

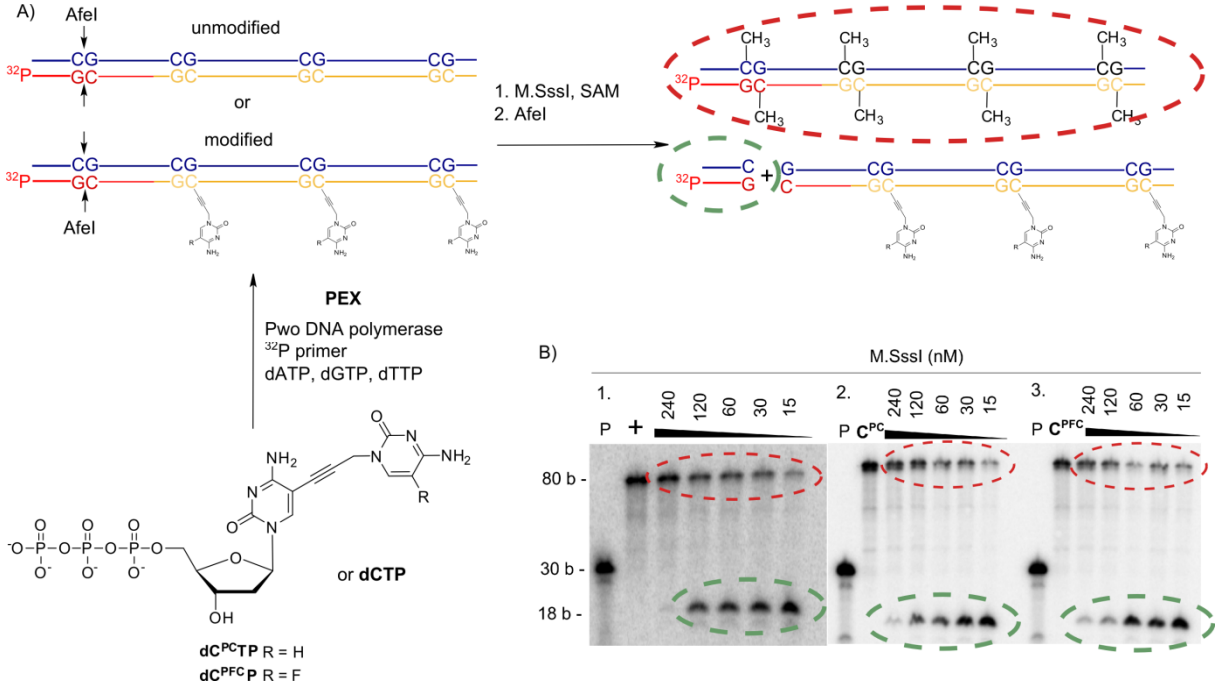


Figure 18. Inhibition of M.SssI DNA methylation assay. A) Outline of the assay for testing of inhibition of DNA methylation by M.SssI B) PAGE analysis of the inhibition assay performed with different concentrations of M.SssI. 1. PAGE: unmodified DNA; 2. PAGE: **dC^{PC}** modified DNA; 3. PAGE: **dC^{PFC}** modified DNA. P stands for primer.

Table 6. Primer and template used for inhibition assay of M.SssI. CpG sites are in bold. The AfeI site is underlined.

primer^{MTRE}	5'-CATGGGCGGCATGGGAGCGCTATGTAGGTA-3'
template^{MTRE}	5'-ATA**CG**ATATCCCATCACATACTCATA**CG**AATCTACCAACATA**CG**AAT CTATACCTACAT<u>AG**CG**CT</u>CCCATGC**CG**CCCATG-3'

Different physical and structural technicques such as atomic force microscopy and surface plasmon resonance were applied in order to find the M.SssI-DNA covalent complex and thus relevant candidate for DNA methylation inhibition. However, the biochemical approach was shown to be the most straightforward, cheap and robust.

In conclusion, above presented double-headed 2'-deoxyribonucleoside triphosphates are good substrates for DNA polymerases in PEX and PCR and thus they are suitable for construction of DNA bearing linked cytosine derivatives in the major groove mimicking the filliped-out nucleotide. In addition, the methodology for testing of the inhibition of DNA

methyltransferase was developed, but the tested **dC**$^{\mathbf{PC}}$ and **dC**$^{\mathbf{PFC}}$ derivatives did not show any significant inhibition of DNA methyltransferase M.SssI. Nevertheless, the research along this line will be continued in the Hocek group.

3.2 Transient and Switchable Protection of DNA Against Cleavage by Restriction Endonucleases by (Trialkylsilyl)ethynyl-Modified 7-Deaza-adenines[83]

Restriction endonucleases (REs) serve as DNA cutters and are widely used in molecular biology for gene manipulation and cloning. Type II REs recognize short 4–8 bp long recognition sequence in which they selectively at a specific position the phosphodiester bond[66,68e,84] cleave. Hundreds of type II REs are commercially available. Protection of DNA against cleavage by REs by incorporation of modified nucleotide within the recognition sequence could be useful in biotechnology especially for manipulation with large DNA sequences with more than one recognition sequence for one RE. REs search for the recognition sequence on DNA is done mainly by interacting with nucleobases in the major groove. DNA containing 7-deazaadenine[68] or 7-deazaguanine[69] in a recognition sequence was previously reported to inhibit restriction endonucleases, apparently because the N7 atom cannot form hydrogen bonds in the major groove. Recently, the systematic study of 5-substituted pyrimidines[71] and 7-substituted 7-deazaadenines[12b] was done in our group and it was found that some REs can tolerate 7-deazaadenine and 7-ethynyl analogs while more bulky 7-phenyl analog fully inhibit the cleavage. The major groove modified DNA is now accessible and could be prepared using modified **dN^{X}TP**s and DNA polymerases and thus this approach is well suited for preparation of such modified DNA and application for protection of DNA against cleavage by restriction endonucleases (**Figure 19**).

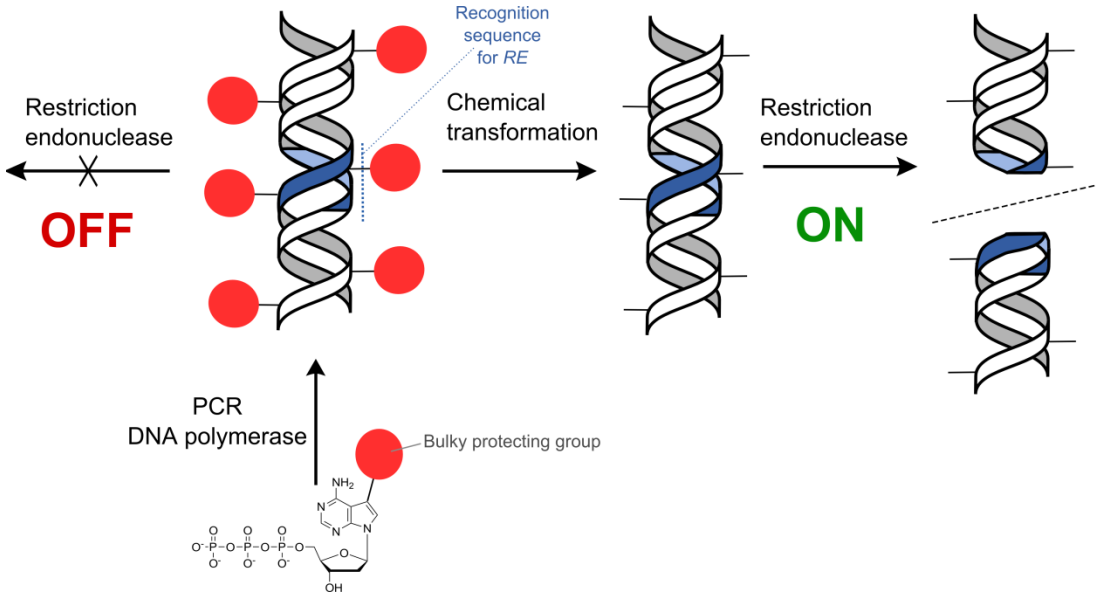


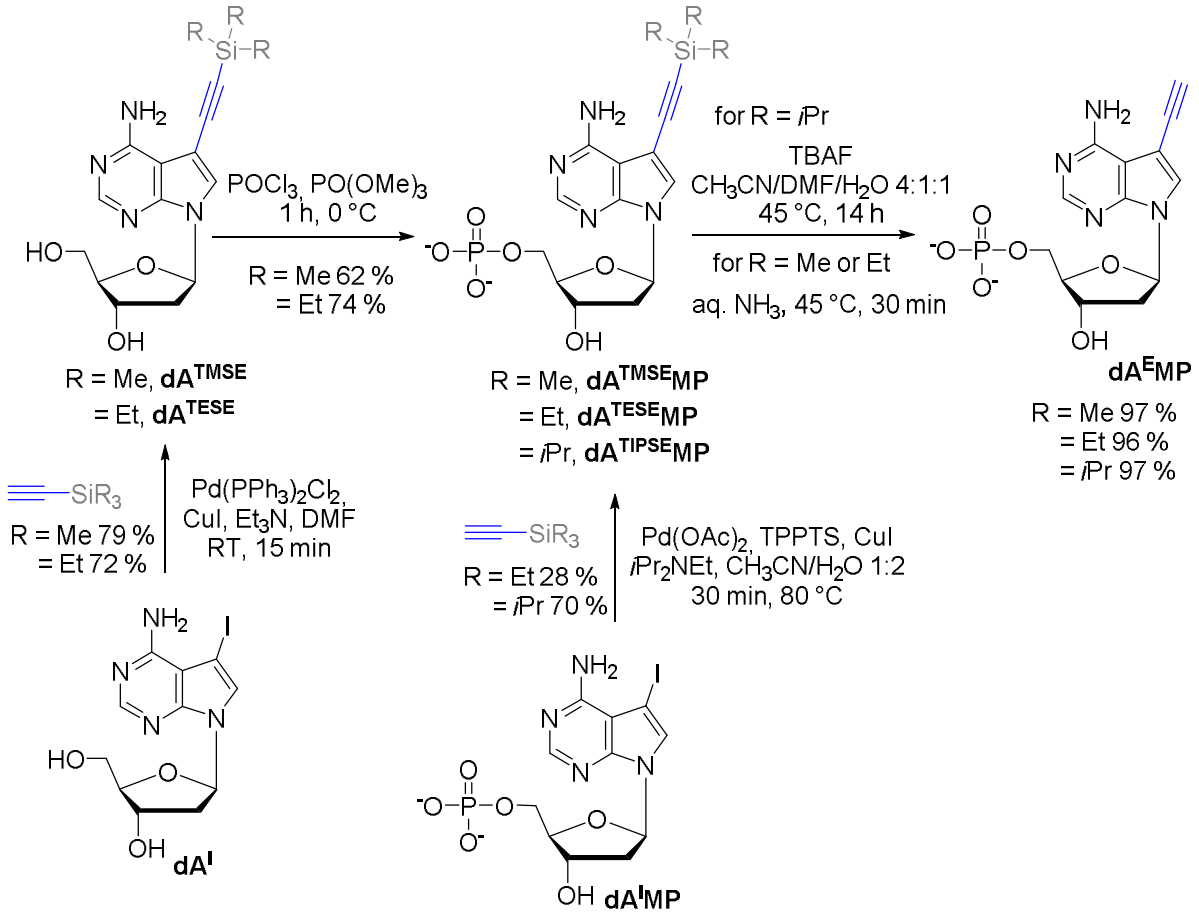
Figure 19. General approach to transient protection of DNA against cleavage by REs.

We proposed that 7-(trialkylsilyl)ethynyl-7-deaza-2'-deoxyadenines in the REs recognition sequences might be bulky enough to protect the DNA against the cleavage by REs (similary to phenyl derivative) while after chemical deprotection of the trialkylsilyl group the resulting ethynyl modified DNA would be accessible for REs and cleaved.[12b] This chemical trigger for biochemical reactions would be interesting for many applications in which the transient and switchable protection is needed (e.g. cloning techniques and transcription regulation).

3.2.1 Synthesis of the Series of 7-(Trialkylsilyl)ethynyl-7-deazaadenine 2'-Deoxyribonucleosides and Nucleotides

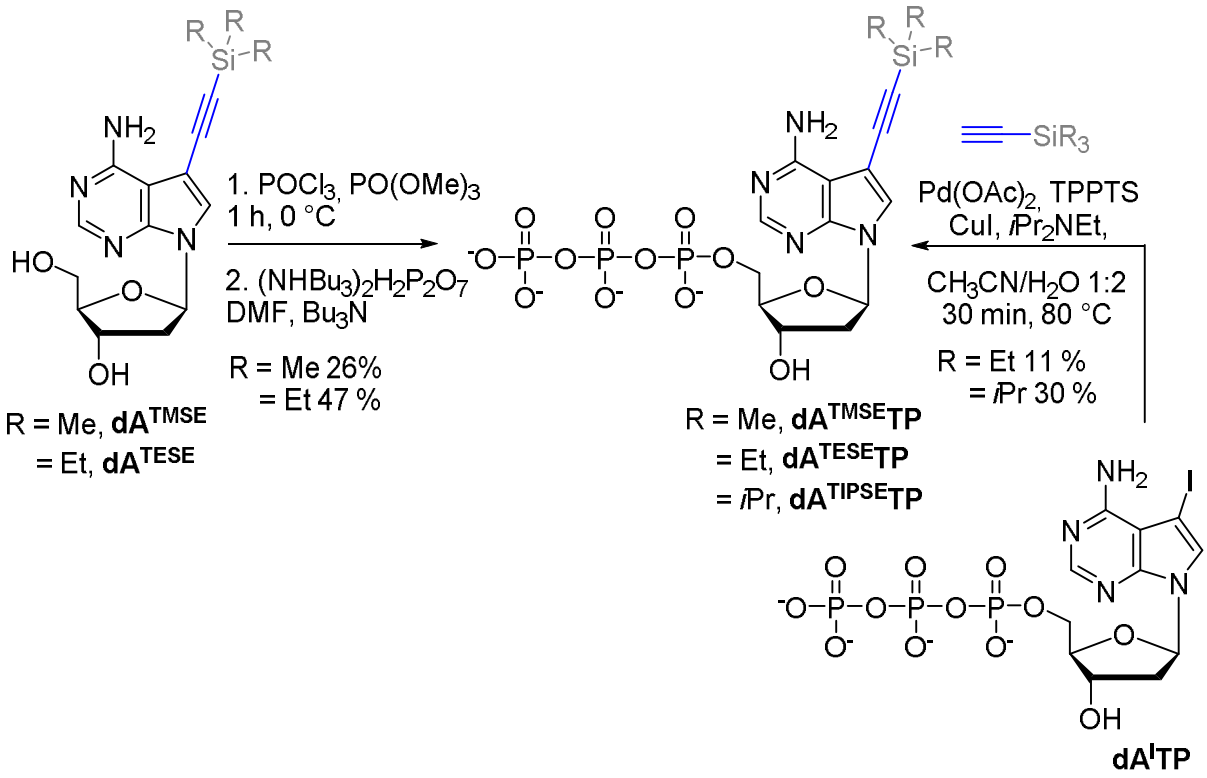
We have decided to synthesize three trialkylsilyl groups: trimethylsilyl (TMS), triethylsilyl (TES) and triisopropylsilyl (TIPS) as protection of the acetylene at the position 7 of 7-deazaadenosine[85]. Analogously to a method based on the (trialkylsilyl)alkyne protection[86] for the triple click labeling of DNA by triazole formation that was recently developed by the Carell and co-workers. The synthesis and deprotection was first studied on model nucleoside monophosphates. The **dA^{TMSE}MP** was not accessible by Sonogashira cross-coupling reaction of 7-iodo-7-deaza-2'-deoxyadenosine monophosphate with trimethylsilylacetylene under aqueous conditions because of the low solubility of the very volatile trimethylsilylacetylene in water which is not compatible with higher reaction temperature. Therefore, the corresponding nucleoside derivative was prepared in DMF at ambient temperature in 79 % yield and phosphorylated with $POCl_3$ in $PO(OMe)_3$ in moderate yield 26 %. The lower yield was caused by rather labile trimethylsilyl protection group which was partially cleaved during the HPLC separation of **dA^{TMSE}MP** in triethylammonium bicarbonate (TEAB) buffer. The **dA^{TESE}MP** was prepared by aqueous Sonogashira cross-coupling of **dA^{I}MP** with trimethylsilylacetylene in low yield 11 %. The Sonogashira reaction of **dA^{I}** with trimethylsilylacetylene in DMF gave us **dA^{TESE}** in good yield 72 % and was followed by phosphorylation to **dA^{TESE}MP** in moderate yield 46 %. Finally, the **dA^{TIPSE}MP** was synthesized directly from **dA^{I}MP** by Sonogashira reaction with triisopropylsilylacetylene in good yield 70 %, syntheses are summarized in **Scheme 15**.

All three **dA^{XE}MP**s were subjected for the deprotection study. The TMSE and TESE derivatives were quantitatively desilylated by treatment with aqueous ammonia, whereas the TIPSE derivative was deprotected by treatment with tetra-*n*-butylammonium fluoride (TBAF) in $CH_3CN/DMF/H_2O$ mixture (**Scheme 15**).



Scheme 15. Synthesis of (trialkylsilyl)ethynyl substituted 7-deaza-2'-deoxyadenosine monophosphates and cleavage conditions of the trialkylsilyl protecting group.

For the construction of the base-modified DNA using DNA polymerases[20] we needed to synthesize (trialkylsilyl)ethynyl modified deoxyadenosine triphosphates (**dA^{XE}TP**s). We first attempted to use straightforward aqueous Sonogashira cross-coupling reaction of **dA^{I}TP** with (trialkylsilyl)acetylenes (**Scheme 16**). Similarly to the deoxyribonucleoside monophosphates, the reaction of **dA^{I}TP** with (trimethylsilyl)acetylene did not proceed. However, the cross-coupling reaction of **dA^{I}TP** with (trimethylsilyl)- and (triisopropylsilyl)acetylene proceeded reasonably well to give desired **dA^{TESE}TP** and **dA^{TIPSE}TP** in moderate yields, 15 and 30 % respectively. The TMSE derivative was prepared from **dA^{TMSE}** by triphosphorylation in 26 % yield, also diminished by partial deprotection during the HPLC separation in TEAB buffer. The same approach was used for more efficient preparation of **dA^{TESE}TP** in 47 % yield.



Scheme 16. Synthesis of (trialkylsilyl)ethynyl substituted 7-deaza-2'-deoxyadenosine triphosphates.

3.2.2 Incorporation of Protected 7-Ethynyl-7-Deazaadenine 2'-Deoxyribonucleoside Triphosphates into DNA by DNA Polymerases and Cleavage of (Trialkylsilyl)ethynyl and Ethynyl-Modified DNAs by Restriction Endonucleases

We tested all three **dA^{XE}TP**s as substrates for the DNA polymerases in PEX using template temp^{4base} and prim^{A} (**Table 7**). The template temp^{4base} is designed to give 31 nt long DNA with four **dA^{XE}** separated by at least 2 nucleotides from each other. Both DNA polymerases KOD XL and Vent(exo-) efficiently incorporated all four **dA^{XE}** into the full-length product (**Figure 20**).

Table 7. Primers and templates used for PEX and subsequent cleavage with REs. Recognition sequence is in bold, cleavage site is marked with a slash.

Template primer	Restriction endonuclease	Sequence
temp^{4base}	-	5'- CTAGCATGAGCTCAGTCCCATGCCGCCCATG -3'
prim^{A}		3'- GGGTACGGCGGGTAC -5'
temp^{Kp}	KpnI	5'- AACGACGAC**GGTAC/C**CCCATGCCGCCCATG -3'
prim^{A}		3'- GGGTACGGCGGGTAC -5'
temp^{Rs}	RsaI	5'- AACGACGACG**GT/AC**GCCCATGCCGCCCATG -3'
prim^{A}		3'- GGGTACGGCGGGTAC -5'
temp^{Sa}	SacI	5'- AACGACGAC**GAGCT/C**CCCATGCCGCCCATG -3'
prim^{A}		3'- GGGTACGGCGGGTAC -5'

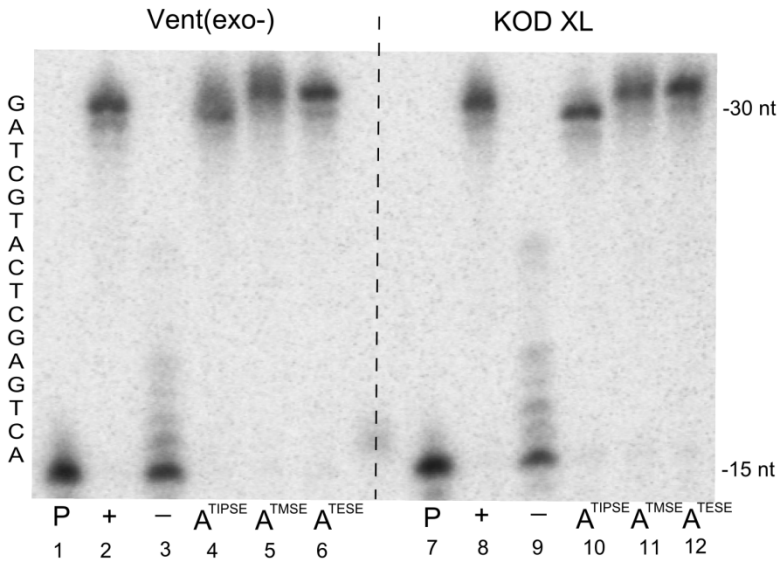


Figure 20. PAGE analysis of PEX with temp^{4base} in presence of Vent(exo-) or KOD XL. Lane 1: primer (5´-^{32}P-end labeled primer); lane 2 and 8: natural dNTPs; lane 3 and 9: dTTP, dCTP, dGTP; lane 4 and 10: **dA^{TIPSE}TP**, dTTP, dCTP, dGTP; lane 5 and 11: **dA^{TMSE}TP**, dTTP, dCTP, dGTP; lane 6 and 12: **dA^{TESE}TP**, dTTP, dCTP, dGTP.

Next, we incorporated each **dA^{XE}** into the dsDNA with recognition sequence for REs KpnI[87], RsaI[68a] and SacI[88] (all of these REs tolerate **dA^{E}** within the recognition sequence) and treated it with corresponding RE, in parallel the cleavage of the unmodified DNA was always used as a control of the cleavage reaction itself. We found out the TMSE modified DNA was partially cleaved by KpnI and RsaI (**Figure 21 A** and **B** respectively, line 4). The cleavage by SacI was inhibited. In contrast, the TESE and TIPSE modified DNAs were fully protected against cleavage by all tested REs (**Figure 22** and **23**, line 5).

The deprotection of (trimethylsilyl)- and (triethylsilyl)ethynyl modified DNA was carried out by aqueous ammonia at 45 °C for 2 h, followed by evaporation of the excess of ammonia under reduced pressure. Resulting ethynyl modified DNAs were treated with corresponding REs and indeed they were fully cleaved. Due to the low stability and insufficient protection of DNA against cleavage by REs the TMS protection group was excluded from further testing. For the deprotection of triisopropylsilyl groups we used TBAF thus the resulting ethynyl modified DNA was precipitated and then desalted on gel chromatography column, consequently the yield was lowered what is visible on the gel in **Figure 23** (line 6 and 7).

All ethynyl-DNAs after deprotection were cleanly cleaved by REs (**Figure 21–23**, line 7), except one prepared from TMSE-DNA and treated with RsaI (**Figure 21 b**, line 7). This clearly confirms the proof-of-principle of transient silyl protection against the restriction cleavage.

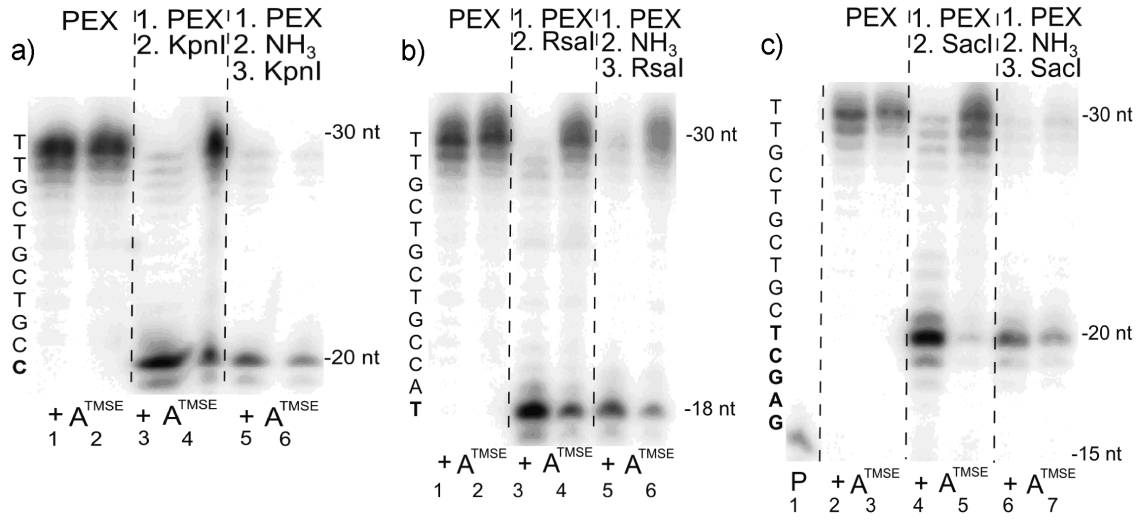


Figure 21. PAGE analysis of PEX, deprotection and cleavage products. a) PEX with temp^{Kp} , cleavage by KpnI; b) PEX with temp^{Rs} , cleavage by RsaI; c) PEX with temp^{Sa} , cleavage by SacI. Lane 1: primer; lane 2: product of PEX with natural dNTPs; lane 3: product of PEX with dTTP, dCTP, dGTP, **dA^{TMSE}TP**: lane 4: cleavage of the unmodified PEX product with a RE; lane 5: treatment of the TMSE-modified PEX product with a RE (partial cleavage); lane 6, treatment of the unmodified PEX product with NH_3 followed by cleavage with a RE; lane 7: treatment of the TMSE-modified PEX product with NH_3 followed by cleavage with a RE (full cleavage).

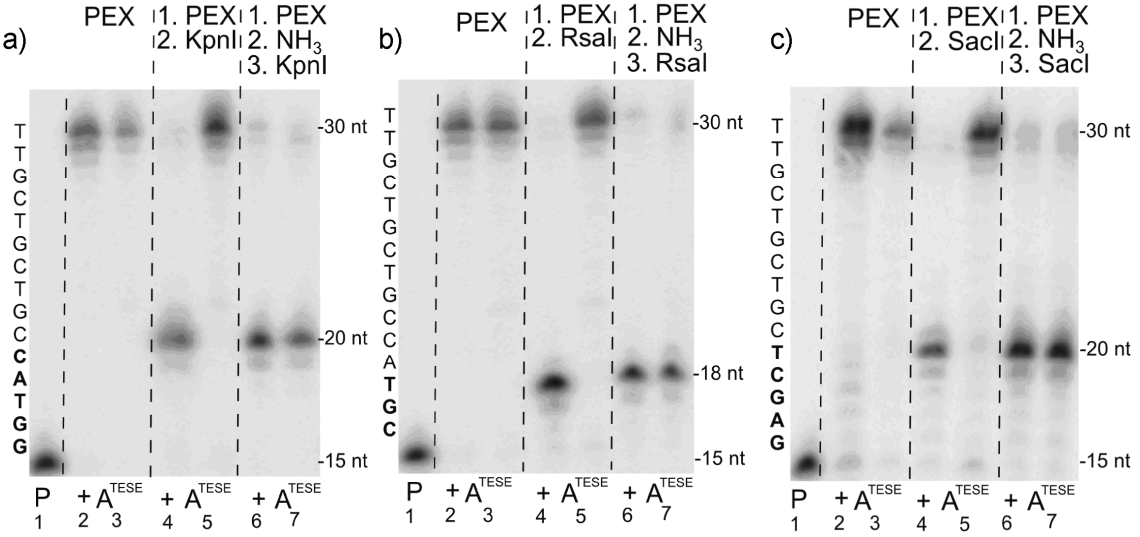


Figure 22. Page analyses of PEX experiments, deprotection and cleavage products: a) PEX with temp^{Kp} , cleavage by KpnI; b) PEX with temp^{Rs} , cleavage by RsaI; c) PEX with temp^{Sa} , cleavage by SacI. Lane 1: primer; lane 2: product of PEX with natural dNTPs; lane 3: product of PEX with dTTP, dCTP, dGTP, **dA^{TESE}TP**; lane 4: cleavage of the unmodified PEX product with a RE; lane 5: treatment of the TESE-modified PEX product with a RE (no cleavage); lane 6: treatment of the unmodified PEX product with NH_3 followed by cleavage

with a RE; lane 7: treatment of the TESE-modified PEX product with NH_3 followed by cleavage with a RE (full cleavage).

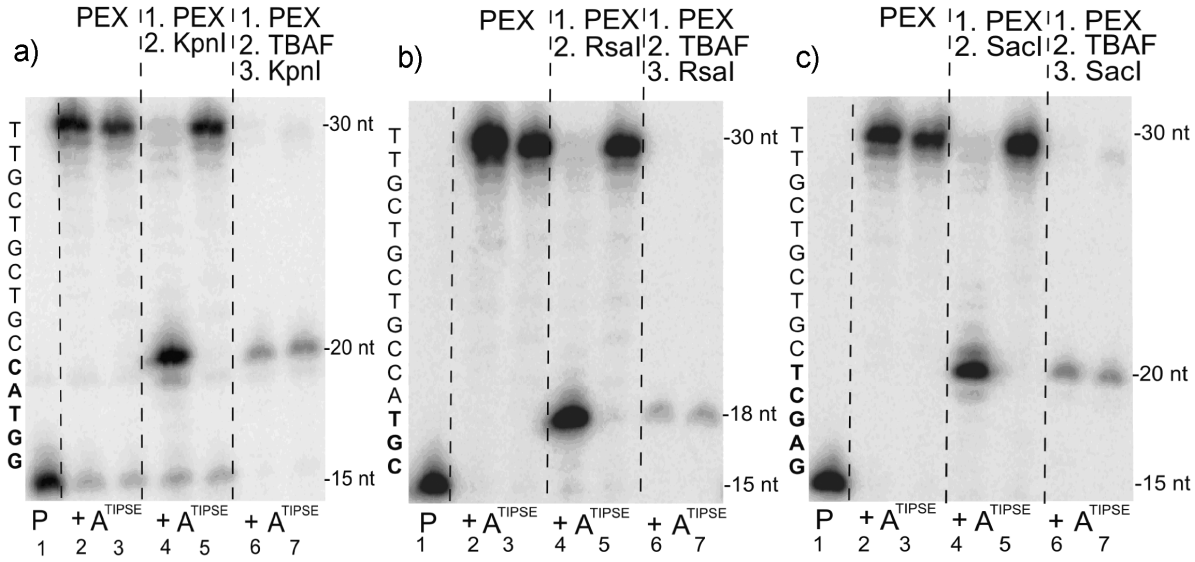


Figure 23. PAGE analysis of PEX, deprotection and cleavage products. a) PEX with temp^{Kp} , cleavage by KpnI; b) PEX with temp^{Rs} , cleavage by RsaI; c) PEX with temp^{Sa} , cleavage by SacI. Lane 1: primer; lane 2: product of PEX with natural dNTPs; lane 3: product of PEX with dTTP, dCTP, dGTP, **dA^{TIPSE}TP**; lane 4: cleavage of the unmodified PEX product with a RE; lane 5: treatment of the TIPSE-modified PEX product with a RE (no cleavage); lane 6: treatment of the unmodified PEX product with TBAF followed by cleavage with a RE; lane 7: treatment of the TIPSE-modified PEX product with TBAF followed by cleavage with a RE (full cleavage).

In order to apply this method in molecular biology we extended the study with testing the protection and cleavage on larger DNA. Hence, we needed first to test the **dA^{TESE}TP** and **dA^{TIPSE}TP** in PCR which is usually used for preparation of the larger DNA. Indeed, we performed the PCR using KOD XL and Vent(exo-). The incorporation of **dA^{TESE}TP** worked reasonably well with Vent(exo-) and very well with KOD XL polymerase to give the full length product (**Figure 24**, line 5 and 10), while incorporation of the more bulky **dA^{TIPSE}TP** did not work with any of the polymerases (**Figure 24**, line 4 and 9). Therefore we selected the **dA^{TESE}TP** for use in further applications.

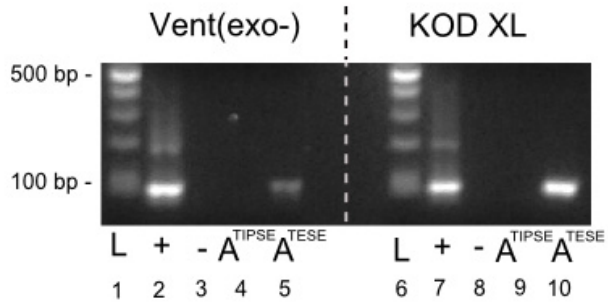


Figure 24. Agarose gel analysis of PCR experiments with 98-mer template and Vent(exo-) (lanes 2-5) or KOD XL (lanes 7-10) polymerases. Lanes 1 and 6: 100 bp DNA ladder; lanes 2 and 7: natural dNTPs; lanes 3 and 8: dCTP, dTTP, dGTP; lanes 4 and 9: **dA^{TIPSE}TP**, dTTP, dCTP, dGTP; lanes 5 and 10: **dA^{TESE}TP**, dCTP, dTTP, dGTP.

It was also proved that the transient protection of DNA against restriction works on larger DNA by preparation of 287 bp long DNA (coding DNA for HIV protease) using KOD XL polymerase in which all adenosines were replaced by **dA^{TESE}** analog (**Figure 25**). The DNA sequence contained the recognition site for RsaI, this RE was used for cleavage of the resulting TESE modified DNA but in agreement with our previous observations the modified DNA was not cleaved while the control unmodified DNA was fully cleaved (**Figure 25 B**, line 5 and 6). However, after the treatment of the TESE modified DNA with ammonia the resulting ethynyl modified DNA was spliced by RsaI as the unmodified DNA (**Figure 25 B**, line 7 and 8). We have clearly shown the **dA^{TESE}TP** is well incorporated into long DNA by KOD XL polymerase, the TESE modified DNA with recognition sequence for RsaI resists the treatment with this RE but after deprotection of TES group the recognition sequence becomes accessible for RsaI and is cleaved. Further interesting questions for future research on this topic might concern the interactions of the modified DNA with transcription factors, the possible regulation of gene expression by transient protection, and the possibility of using RE-resistant protected DNA to combat bacteria.

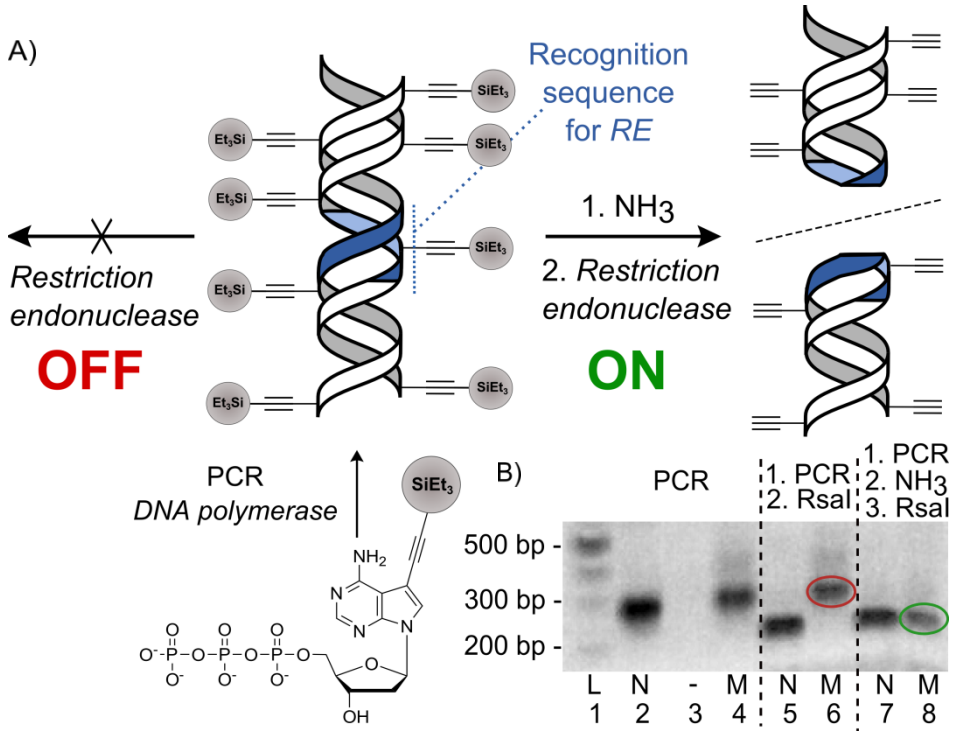


Figure 25. Transient protection of DNA against cleavage by restriction endonucleases using (trimethylsilyl)ethynyl modified DNA a) Scheme of the PCR incorporation of **dA^{TESE}TP**, deprotection and cleavage of the DNA. b) Agarose gel analysis of PCR, deprotection and cleavage experiments with 297-mer template. Lane 1: DNA ladder; lane 2: product of PCR with natural dNTPs; lane 3: product of PCR with dTTP, dCTP, dGTP; lane 4: product of PCR with dTTP, dCTP, dGTP, **dA^{TESE}TP**; lane 5: cleavage of the unmodified PCR product with RsaI; lane 6: treatment of the TESE-modified PCR product with RsaI (no cleavage, marked in red circle); lane 7: treatment of the unmodified PCR product with NH_3 followed by cleavage with RsaI; lane 8: treatment of the TESE-modified PCR product with NH_3 followed by cleavage with RsaI (full cleavage, marked in green circle).

3.2.3 Nucleobase Protection Strategy for Gene Cloning and Expression[89]

There are some techniques available for the gene cloning in which the gene of the interest contain target sequences for some REs in several copies. The common approach si codon mutation within the restriction site, but this may not circumvent the problem, for example if the expression system does not contain acomplementary transfer RNA.

In regard to our previous study of protection of DNA against cleavage by REs[83] using modified nucleobases we wanted to investigate whether it would be possible to use major-groove protection-group approach in gene cloning. Here, I present the first strategy for the polymerase synthesis of a chemically protected copy of a gene flanked by unmodified sequences, its site-specific cleavage by REases, cloning into a plasmid and use for protein expression. This work we did in collaboration with Prof. J. S. Dickschat and Dr. N. L. Brock from Braunschweig Universität in Germany who study the *tdaD* gene[90] in the thropodithietic acid biosynthetic gene cluster from *Phaeobacter gallaeciensis* DSM 17395.[91] The tdaD gene encodes the putative thioester hydrolase TdaD and contains the cleavage site for NcoI within the gene sequence. While they would use the same RE for the gene cloning (**Figure 26**).

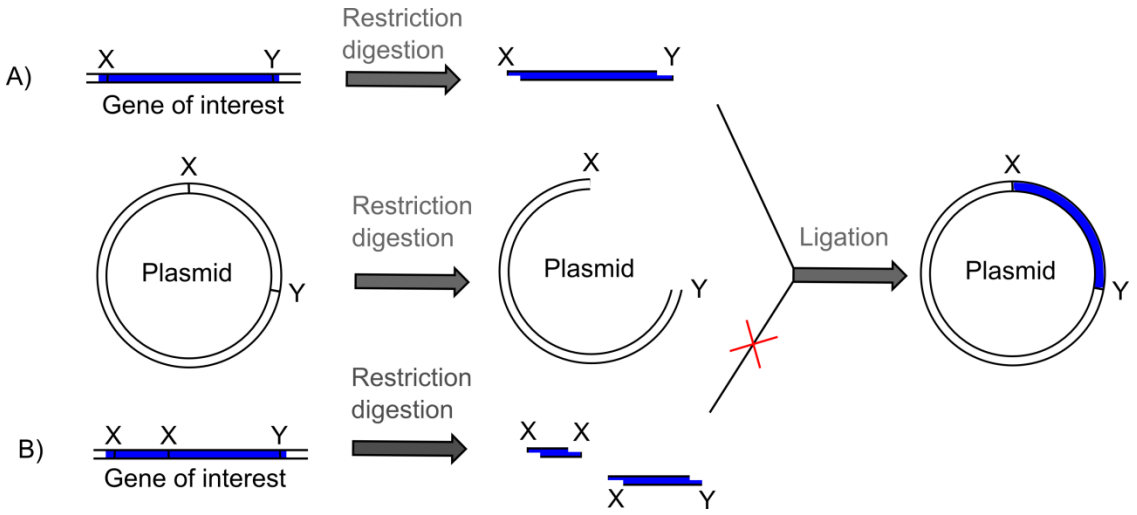


Figure 26. Gene cloning with A) one site of the each recognition sequence for RE used in the cloning and B) multiple site of the recognition sequence for RE used in the cloning.

In a standard procedure of gene cloning based on PCR the recognition sequences for REs are introduced by 5'- regions of forward and reverse primers. The PCR product is then treated with corresponding REs and the recognition sequences within the primers are cleaved (**Figure 27**). The widespread pET28c(+) expression system for isopropyl-β-D-thiogalactopyranosid (IPTG) inducible expressions of histidine-tag (His_6-tag) modified proteins in *Escherichia coli* takes advantage of the RE NcoI (recognition sequence CCATGG)[92]. This sequence is

displayed in the forward oligonucleotide primer in a way that the partial sequence ATG coincides with the start codon of the gene of interest. In the reverse oligonucleotide primer the stop codon (TGA, TAG, or TAA) of the original gene is together with the sequence of the following three base-pairs altered to the CTCGAG recognition sequence of the RE XhoI.[93] After digestion and cloning of the PCR product into the pET28c(+) vector the sequence of the cloned gene is fused to the *Xho*I recognition sequence followed by the His_6-tag sequence $(CAC)_6$ allowing protein purification by Ni^{2+}-nitriolotriacetic acid (NTA) chromatography and a stop codon (TGA). Another feature of the pET28c(+) vector is a strong IPTG inducable promoter together with the ribosome binding site upstream of the NcoI recognition sequence. Apart from the C-terminal His_6-tag the sequence of the expressed protein deviates only slightly from the sequence of the native protein, due to two additional amino acid residues encoded by the XhoI recognition sequence and the possible exchange of a C, T, or A against guanine (G) in the first position of the first codon following the start codon. This strategy of NcoI-XhoI cloning fails if any of these two recognition sites occurs within the sequence of the gene to be cloned. The multiple cloning site of the pET28c(+) vector contains alternative RE recognition sequences for cloning, but the modifications of the sequence of the expressed protein in comparison to the natural sequence are kept to a minimum, if the NcoI and XhoI recognition sites are used, making NcoI-XhoI cloning the preferred strategy. The concept using **dA^{TESE}TP** which would be incorporated into the PCR product by DNA polymerase and protect the NcoI site in the gene sequence against the restriction opens up new alternatives for NcoI-XhoI cloning strategy (**Figure 27**).

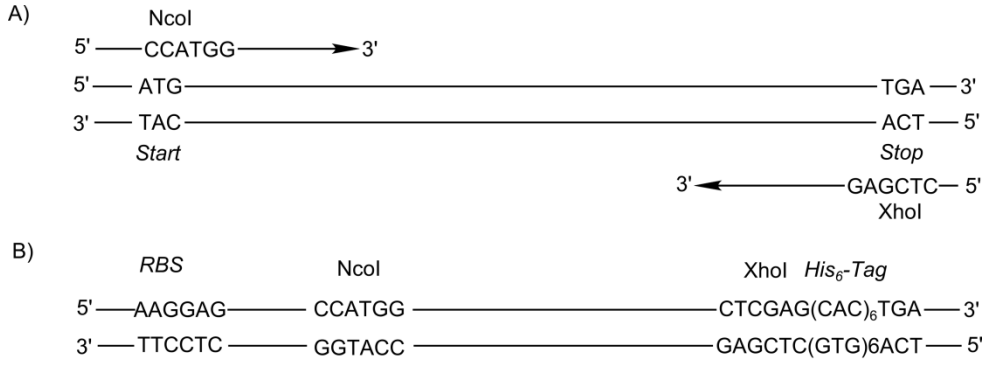


Figure 27. Cloning strategy for the pET-28c(+) expression system. A) Primer design for PCR amplification of gene of interest (from "start" to "stop"), B) multiple cloning site of pET-28c(+) vector showing ribosome binding site (RBS), NcoI and XhoI recognition sites, and sequence coding for the His_6 tag.

The new strategy for preparation of TESE modified PCR product needed to be developed because of the simple use of one set of primers cannot be applied, since the resulting DNA contains the TESE groups inherently in the complementary strand to the primers and thus protects the cleavage of the recognition sequences within the primer sequences. We envisaged the preparation of two partially overlapping (90 base-pairs) PCR products using two sets of primers and **dA^{TESE}TP**. In the next step, PCR 1 and PCR 2 products are used in the third PCR with natural 2'-deoxyribonucleoside triphosphates in which the PCR 1 and PCR 2 products served as the primers. This was allowed thanks to annealing of the coding strand of PCR 1 with its 3'-end as a megaprimer to the 5'-end of the noncoding strand of PCR 2 (and vice versa) and elongation to yield the full gene sequence, now with only the DNA sequence around the internal NcoI recognition site, but not the flanking regions that have to be cleaved by REases for cloning being chemically protected. The concept is clear from the **Figure 28**.

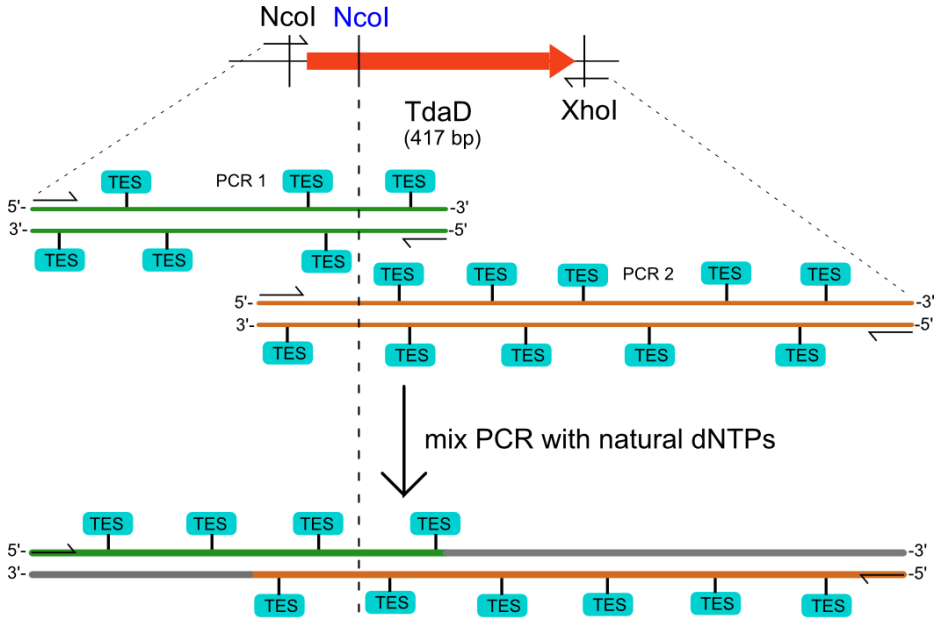


Figure 28. Strategy for PCR amplification with protection of NcoI site within the *tdaD* gene sequence.

Indeed, we have prepared PCR 1 (125 bp) and PCR 2 (399 bp) products using KOD XL polymerase and **dA^{TESE}TP** and used them in the third PCR with natural dNTPs which proceeded in ten PCR cycles and yielded the a copy of the full *tdaD* gene (343 bp) (**Figure 29 A**, line 3). This final product was in contrast to its unmodified amplificate of the *tdaD* gene not internally cleaved by NcoI-XhoI treatment (**Figure 29 A**, line 4 and 5). Treatment of the TESE-protected amplificate of *tdaD* with ammonia for silyl group deprotection followed by NcoI-XhoI digestion resulted in cleavage at the internal NcoI recognition sequence (**Figure**

29, lane 6). This deprotection is an option for reducing the bulkiness of the modification in case the TESE-modified DNA sequence is not replicated in *E. coli*.

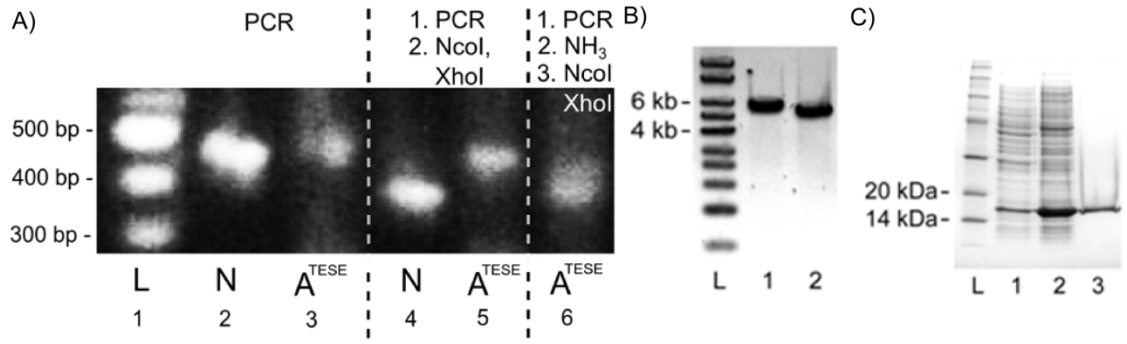


Figure 29. Gel electrophoretic analyses of PCR products and protein expression. Digestion experiments with the REases NcoI and XhoI. Lane 1: 100 bp DNA ladder; lane 2: PCR amplificate of *tdaD* obtained with natural dNTPs; lane 3: product of mix PCR obtained with **dA^{TESE}TP**; lane 4: unprotected copy of *tdaD* treated with NcoI and XhoI resulting in cleavage to yield a 363 bp product (second product of 61 bp not visible); lane 5: protected amplificate of *tdaD* obtained in mix PCR (lane 3) treated with NcoI and XhoI is not cleaved; lane 6: protected copy of *tdaD* first treated with NH_3 for removal of the TES protecting group, followed by treatment with NcoI and XhoI is cleaved to yield a 363 bp product. B) Linearized (EcoRV) pET-28c(+) expression vector containing the *tdaD* insert (lane 1, 5647 bp) and linearized pET-28c(+) expression vector without insert (lane 2, 5367 bp). E) Production of TdaD-His6. Lane 1: soluble proteins without IPTG induced gene expression; lane 2: soluble proteins after IPTG induced gene expression; lane 3: Ni^{2+} –NTA-chromatography-purified TdaD-His^6 (16.9 kDa); L: DNA or protein ladder. Gel A was visualized with DNA staining by GelRed, gel B was visualized with DNA staining by ethidium bromide, and gel C was visualized with protein staining by Coomassie Brilliant Blue. For gel B the camera was set to inverted mode, giving a white background and black DNA bands.

Dr. Nelson Brock carried out the construction of the plasmid, its amplification and production of the TdaD protein. The NcoI-XhoI end-digested and internally A^{TESE} protected *tdaD* gene was then cloned into the pET28c(+) vector using T4 DNA ligase and the resulting ligation product was transformed into *E. coli* DH5α (**Figure 30**). It was of upmost importance for the success of the cloning strategy with internally protected DNA that the A^{TESE} protected DNA sequence parts of the transformed plasmid DNA were replicated during cell division. This was demonstrated to be possible by isolation of plasmid DNA from *E. coli* clones and linearization by the single-cutting enzyme EcoRV. A gel electrophoretic analysis suggested insertion of the *tdaD* sequence into the multiple cloning site of pET28c(+) (**Figure 29 B**, line 1) as was corroborated by DNA sequencing. Therefore, the desilylation step was not needed for this cloning and amplification. The obtained expression construct was finally used for

expression of TdaD followed by protein purification via Ni^{2+}-NTA chromatography (**Figure 29 C**).

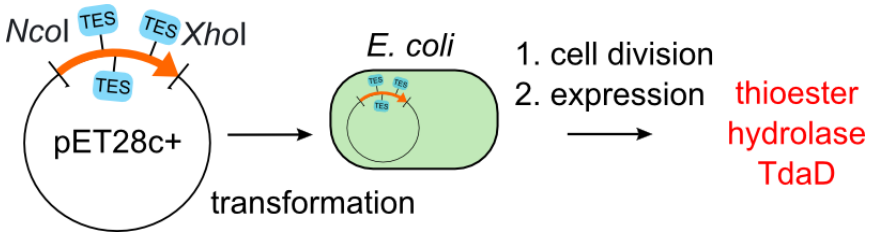


Figure 30. Cloning of the modified gene into the plasmid, transformation and production of thioester hydrolase TdaD.

In conclusion, we have demonstrated that a nucleobase modification can be successfully used for protection of a PCR-derived gene copy against cleavage by a REs in gene cloning. The original methodology of two PCR amplifications using a modified dNTP followed by a third PCR using the natural dNTPs was developed and shown to be suitable for construction of an internally protected gene copy flanked by unmodified sequences and subsequent protein production.

3.3 Study of Direct Competition Incorporations of 7-Substituted 7-Deazaadenine and 5-Substitute Cytosine 2'-Deoxyribonucleoside Triphosphates in the Presence of Their Natural Counterparts[94]

Base functionalized DNAs are used in many areas of the chemical biology and medicinal chemistry for different purposes such as DNA sequencing, protein-DNA conjugation or *in vivo* labeling. This is usually prepared by incorporation of base modified 2'-deoxyribonucleoside triphosphate in primer extension (PEX) experiment or PCR.[20] Reaction mixtures of both methods PEX and PCR contain modified **dN^{X}TP** in absence of the natural counterpart. In many applications the presence of the natural counterpart to the modified **dN^{X}TP** is inherent and cannot be avoided e.g. *in vivo* DNA labeling. It was preassumed that natural dNTPs are evolved to be the ideal substrates for DNA polymerases and also any modification of the natural dNTPs would lead to decrese of the incorporation rate. Thus only little attention was paid to competitive incorporations of modified **dN^{X}TP**s in the presence of the natural counterparts. Also there was not any suitable method for determination of the resulted mixture of modified and unmodified DNA. Recently, several structural and mechanistic studies of competitive single incorporations of modified **dN^{X}TP**s were performed by Marx and co-workers.[48] They analysed resulting mixture of modified and unmodified DNAs on PAGE, because of the different electroforetic mobilities of modified nucleotides attached to the 3' -end compared to the natural DNA. The general scheme of the competitive incorporation is illustrated in **Figure 31**.

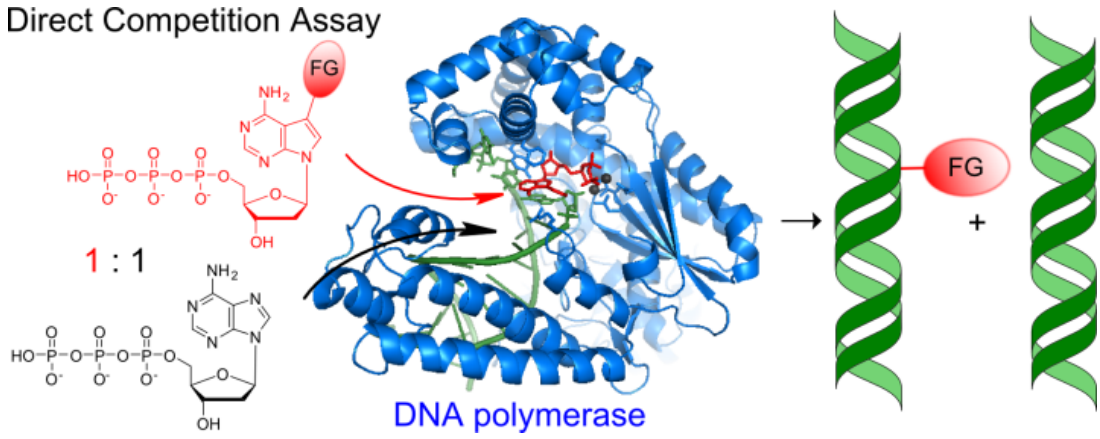


Figure 31. Direct competitive incorporation of functionalized **dA^{X}TP** versus natural dATP.

3.3.1 Method Validation

The absence of the any systematic study of competitive incorporations led us to develop the general method for analysis of the mixture of modified and unmodified DNA based on restriction endonucleases[95] which are inhibited by modified nucleobases within the recognition sequence. We took advantage of the knowledge acquired from previous studies of interactions of the REs with modified nucleobases performed in our group.[12b, 71] Thanks the fact that we were able to find suitable REs which are fully inhibited by tested modified nucleobases we were able to design the method shown in the **Figure 32**. First, we performed the proof-of-principle experiments with the **dA^{Ph}TP** and **dC^{Ph}TP** in the presence of Pwo polymerase followed by treatment with BamHI and ClaI respectively and separated the resulting DNA mixtures on the PAGE. Both **dA^{Ph}TP** and **dC^{Ph}TP** were tested in different ratios of modified **dN^{Ph}TP** to unmodified dNTP (1:1 and 10:1). The intensity of the two spots of cleaved natural and intact modified DNA were then quantified using image analysis. These experiments revealed that method is suitable for determination of the incorporation ratio of modified and unmodified nucleotides by DNA polymerases.

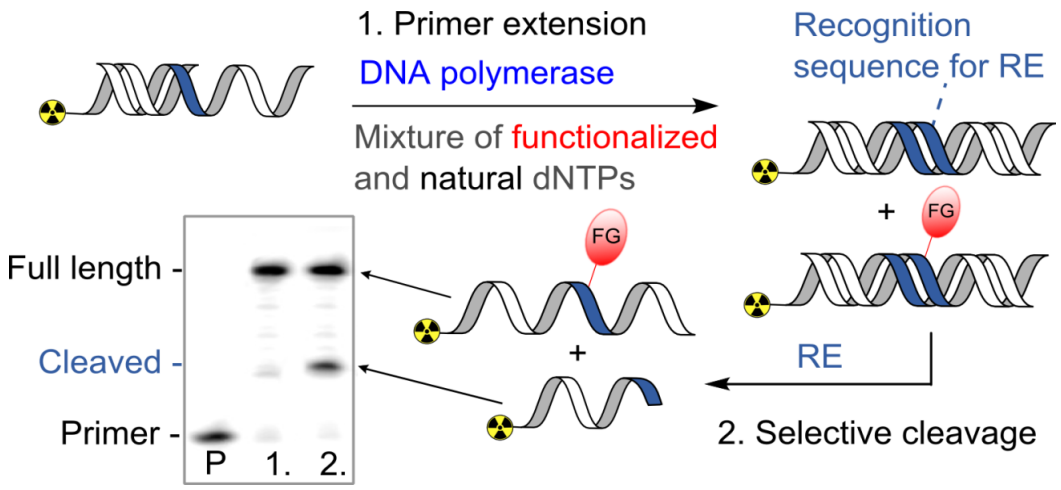
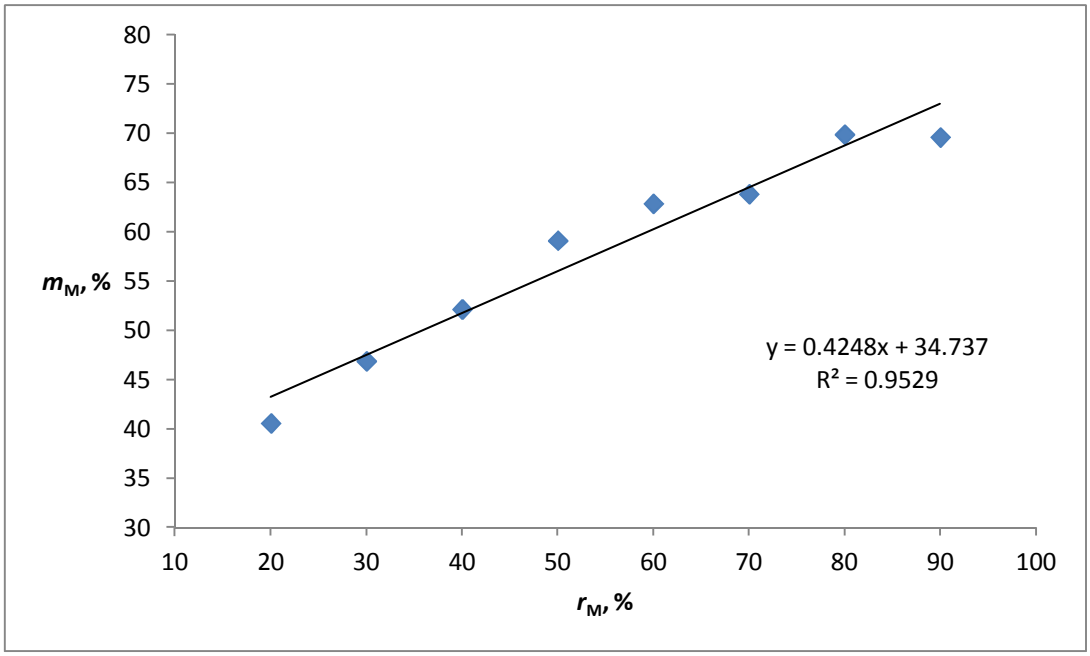


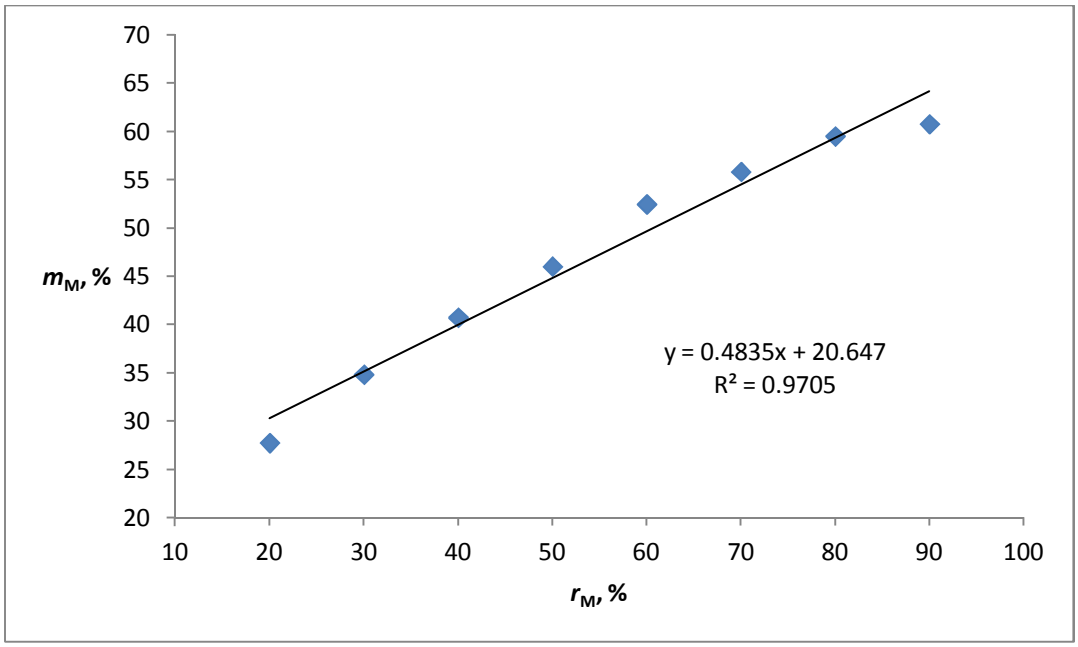
Figure 32. Method for analysis of competitive PEX using RE and PAGE.

In order to prove our method is correct and provides reliable data, we carried out competitive PEX using 5' –biotinylated template, so that we were then able isolate by magnetoseparation the elongated primer and measure MALDI of this single stranded DNA. However, it was not possible to directly deduct the ratio of modified and unmodified DNA from MALDI experiments even though the calibration curve was used because of the fact that MALDI is not a quantitative method. The ration could only be determined by using

another ssDNA as an internal standard in MALDI experiment. We compared the data from analysis based on REs with results from MALDI analysis with an internal standard and we found both were in clear correlation, see **Graph 1**, **Graph 2** and **Table 8**. Because of the REs are commercially available, cheap and their quality is standardized, the subsequent PAGE is convenient separation method in comparsion to the expensive and laborious isolation of ssDNA using streptavidin coated magnetic beads we used the REs based method in further studies of competitive incorporations.



Graph 1. The calibration curve for **dA**$^{\mathbf{Ph}}$ modified DNA. r_M is real abundance of **dA**$^{\mathbf{Ph}}$ modified DNA over the unmodified DNA. m_M is relative abundance of **dA**$^{\mathbf{Ph}}$ modified DNA over internal standard DNA calculated from measured MALDI spectra.



Graph 2. The calibration curve for **dC^{Ph}** modified DNA. . r_M is real abundance of **dC^{Ph}** modified DNA over the unmodified DNA. m_M is relative abundance of **dC^{Ph}** modified DNA over internal standard DNA calculated from measured MALDI spectra

Table 8. Competitive incorporations of **dC^{Ph}TP** and **dA^{Ph}TP** with Pwo analysed by MALDI.[a]

	dCTP/**dC^{Ph}TP**				dATP/**dA^{Ph}TP**			
N/M	1/1		1/10		1/1		1/10	
Analysis	MALDI	RE	MALDI	RE	MALDI	RE	MALDI	RE
% of modified DNA	-	9	60	59	44	43	79	79

[a]Values are in percent of resulting modified DNA.

3.3.2 Direct Competition Assays

With the suitable method in our hands we moved on to the batch determination of incorporation ratios of the series of 7-substituted 7-deaza-2'-deoxyadenosines and 5-substituted 2'-deoxycytidines bearing various bulky functional groups available in our group including ethynyl and phenyl derivatives,[12b, 71] redox labels (**dN^{NO2}TP** and **dN^{NH2}TP**),[49b] fluorescent labels (**dA^{BFU}TP**, **dA^{ABOX}TP**),[12a] reactive groups (**dC^{FT}TP**),[96] and very bulky groups (**dN^{STr}TP**)[97] (**Figure 33** and **Figure 34**). First, we chose the suitable RE for each modified **dN^{X}** based on our previous studies and designed corresponding primer and template sequences which are summarized in **Table 9**. Always only one modified **dN^{X}** could be incorporated into the recognition sequence for RE.

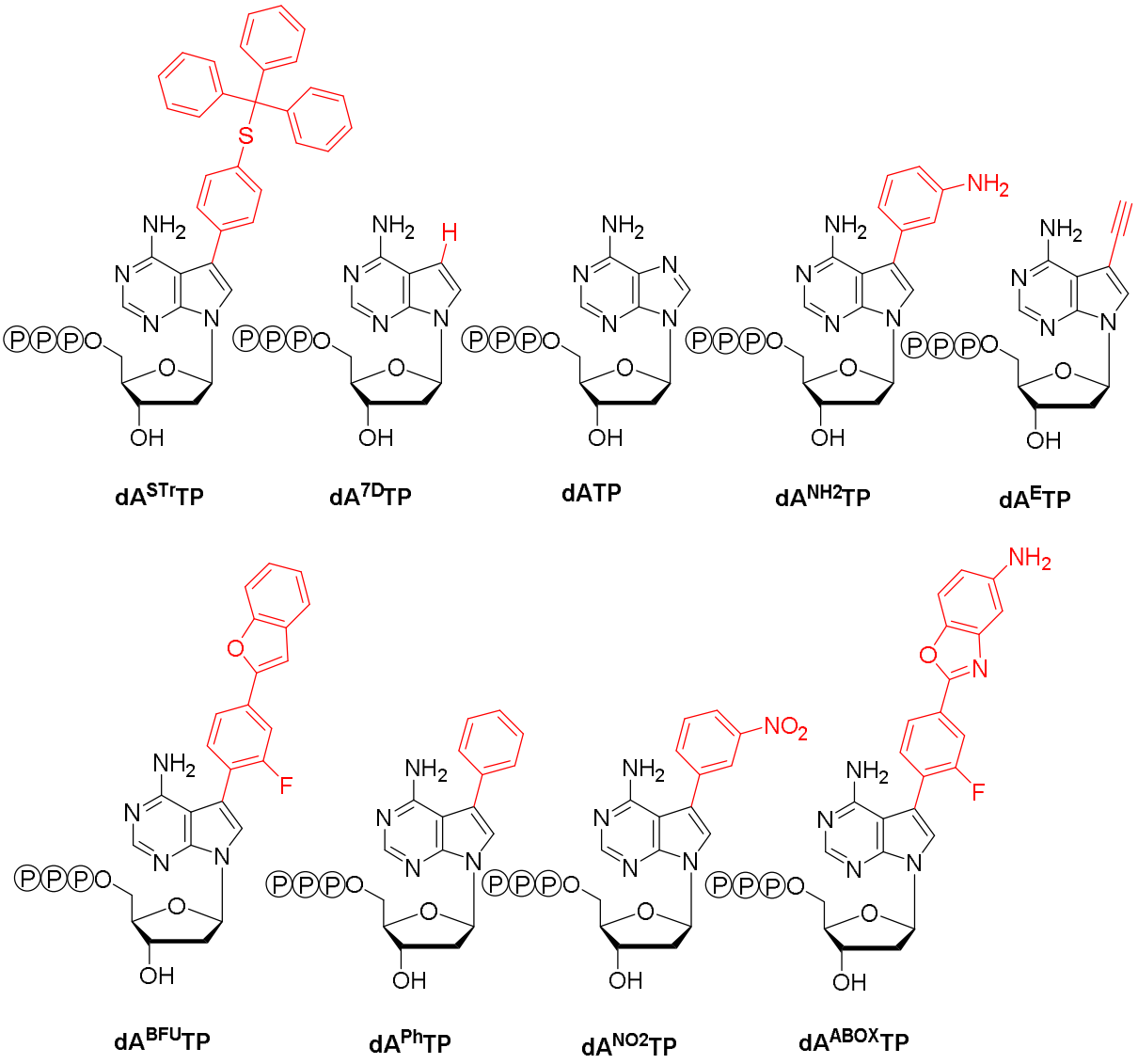


Figure 33. Functionalized **dA^{X}TP**s used in the study of competitive incorporation with K_m values and competitive incorporation ratios.

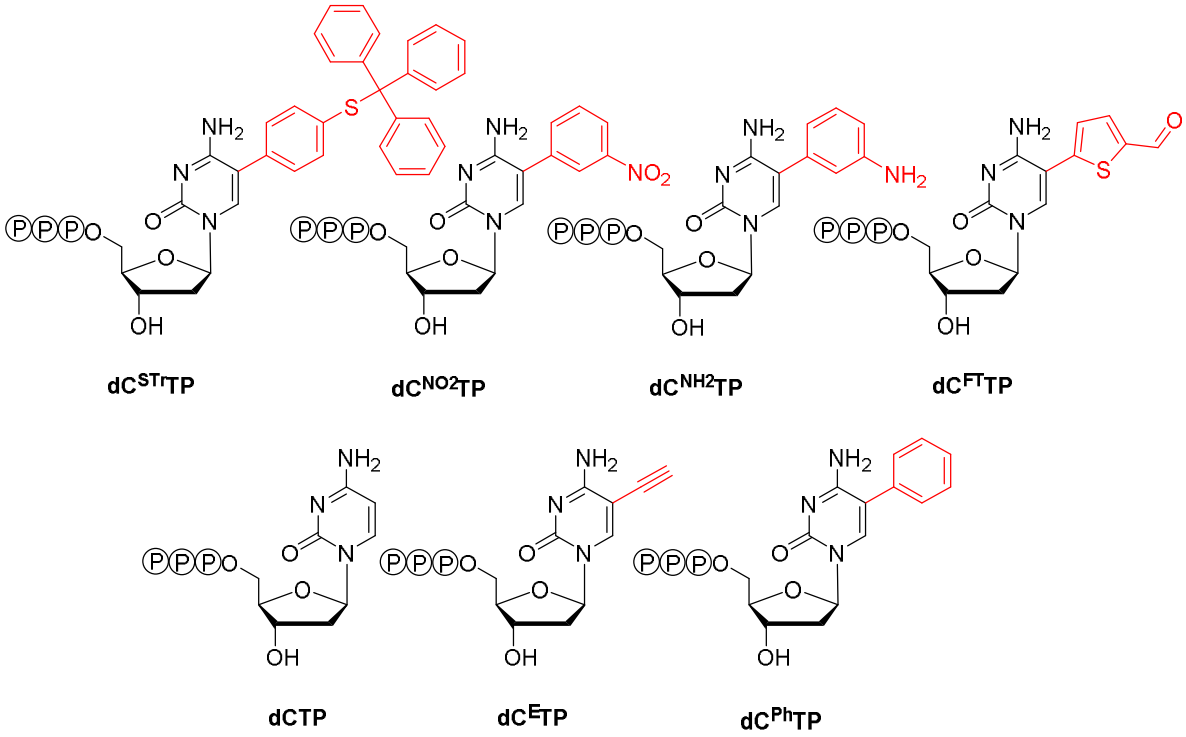


Figure 34. Functionalized **dC^{X}TP**s used in the study of competitive incorporation with K_m values and competitive incorporation ratios.

Table 9. Templates and primers used for PEX construction of DNA with modification within the recognition sequence.[a]

Name	Sequence (5' to 3')	dN^XTP used with the sequence
$Temp^{ClaIC}$	CACTACTCA**AT/CGAT**CCCATGCCG CCCATG	**$dC^{STr}TP$, $dC^{NO2}TP$, $dC^{NH2}TP$, $dC^{FT}TP$, $dC^{E}TP$, $dC^{Ph}TP$**
$Temp^{AluIC}$	CTCACTACTCA**AG/CT**CCCATGCCG CCCATG	
$Temp^{RsaIC}$	AACTACTACT**GT/AC**ACCCATGCCG CCCATG	
$Temp^{BamHIA}$	CACGACGCA**G/GATCC**CCCATGCC GCCCATG	**$dA^{STr}TP$, $dA^{Ph}TP$, $dA^{NH2}TP$, $dA^{NO2}TP$, $dA^{BFU}TP$, $dA^{ABOX}TP$**
$Temp^{PstIA}$	CACGACGCA**CTGCA/G**CCCATGCC GCCCATG	
$Temp^{KpnIA}$	AACGACGAC**GGTAC/C**CCCATGCC GCCCATG	
$Temp^{XhoIA}$	CACGACGCA**C/TCGAG**CCCATGCC GCCCATG	**$dA^{E}TP$**
$Temp^{ScaIA}$	AACGACGAC**AGT/ACT**CCCATGCCG CCCATG	**$dA^{7D}TP$**
$Temp^{MALDI-IS}$	CTAGCATGAGCTCAGTCCCATGCCG CCCATG	
$Prim^{ScaI}$	CATGGGCGGCATGGGAG	
$Temp^{1A}$	TCCCATGCCGCCCATG	
$Temp^{1C}$	CCCGCCCATGCCGCCCATG	
Primer	CATGGGCGGCATGGG	
PrimS1	GATCACTCTTTGGCAGCGACCCCTC GTCAC	
PrimS2	TTAAAGTGCAGCCAATCTGAGTCA ACAGAT	

[a] bold - target palindromic sequence for restriction endonuclease; / - cleavage site.

Next, we performed the measurements of competitive PEX of **$dC^{Ph}TP$** and **$dA^{Ph}TP$** in presence of natural counterparts in different ratios employing different DNA polymerases, results are presented in **Table 10**, **Figure 35** and **Figure 36**. Generally, the **$dC^{Ph}TP$**s was a worse substrate than dCTP for most polymerases, whereas **$dA^{Ph}TP$**s was, surprisingly, better substrate. Except for Pwo, all other polymerases preferentially incorporated **$dA^{Ph}TP$** (up to 67 % of **dA^{Ph}** for Klenow and Bst). On the other hand, **$dC^{Ph}TP$** was a worse substrate than dCTP for all enzymes, except for Bst.

Table 10. Competitive incorporations of **dC^{Ph}TP** and **dA^{Ph}TP** with different DNA polymerases.[a,b]

	dCTP/**dC^{Ph}TP**		dATP/**dA^{Ph}TP**	
N/M	1/1	1/10	1/1	1/10
Pwo	14 (7)	59 (2)	43 (2)	79 (3)
KOD XL	33 (5)	69 (2)	57 (1)	81 (1)
Klenow	33 (3)	59 (1)	67 (4)	90 (5)
Vent(exo-)	36 (2)	56 (2)	65 (1)	81 (2)
Taq	48 (4)	77 (5)	63 (1)	79 (2)
Bst	58 (5)	76 (3)	67 (2)	95 (1)
H Pol α	-	-	46	79

[a]Values are in percent of resulting modified DNA. [b] The standard errors are in brackets.

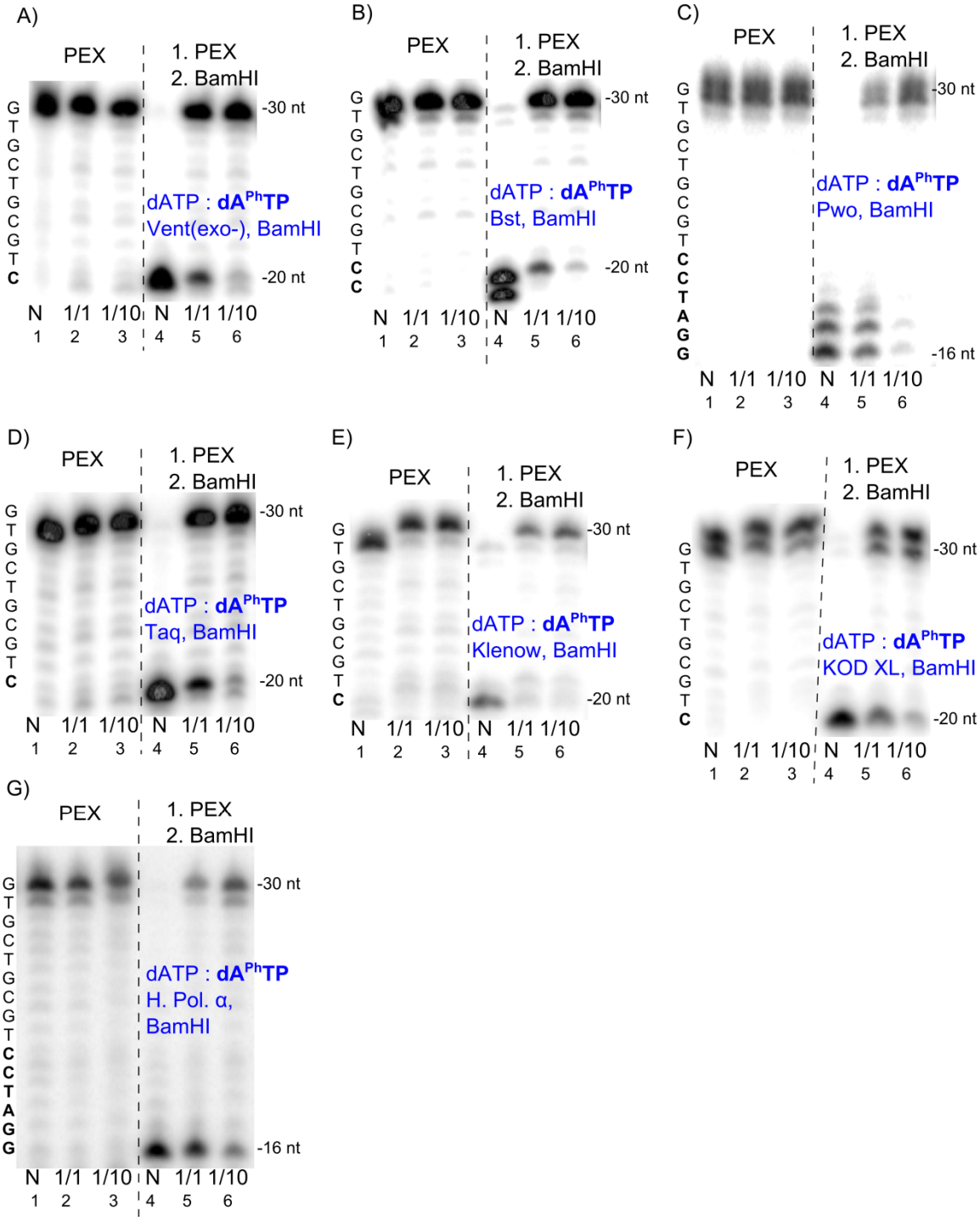


Figure 35. PAGE analyses of PEX experiments and cleavage products with **dA^{Ph}** using different DNA polymerases. Lanes 1, N: product of PEX with natural dNTPs; lanes 2 and 3, 1/1 and 1/10: product of PEX with dTTP, dCTP, dGTP and corresponding ratio of dATP/**dA^{Ph}TP**; lanes 4, N: products of cleavage of unmodified DNA by RE; lanes 5 and 6, 1/1 and 1/10: products of cleavage of modified DNAs by RE.

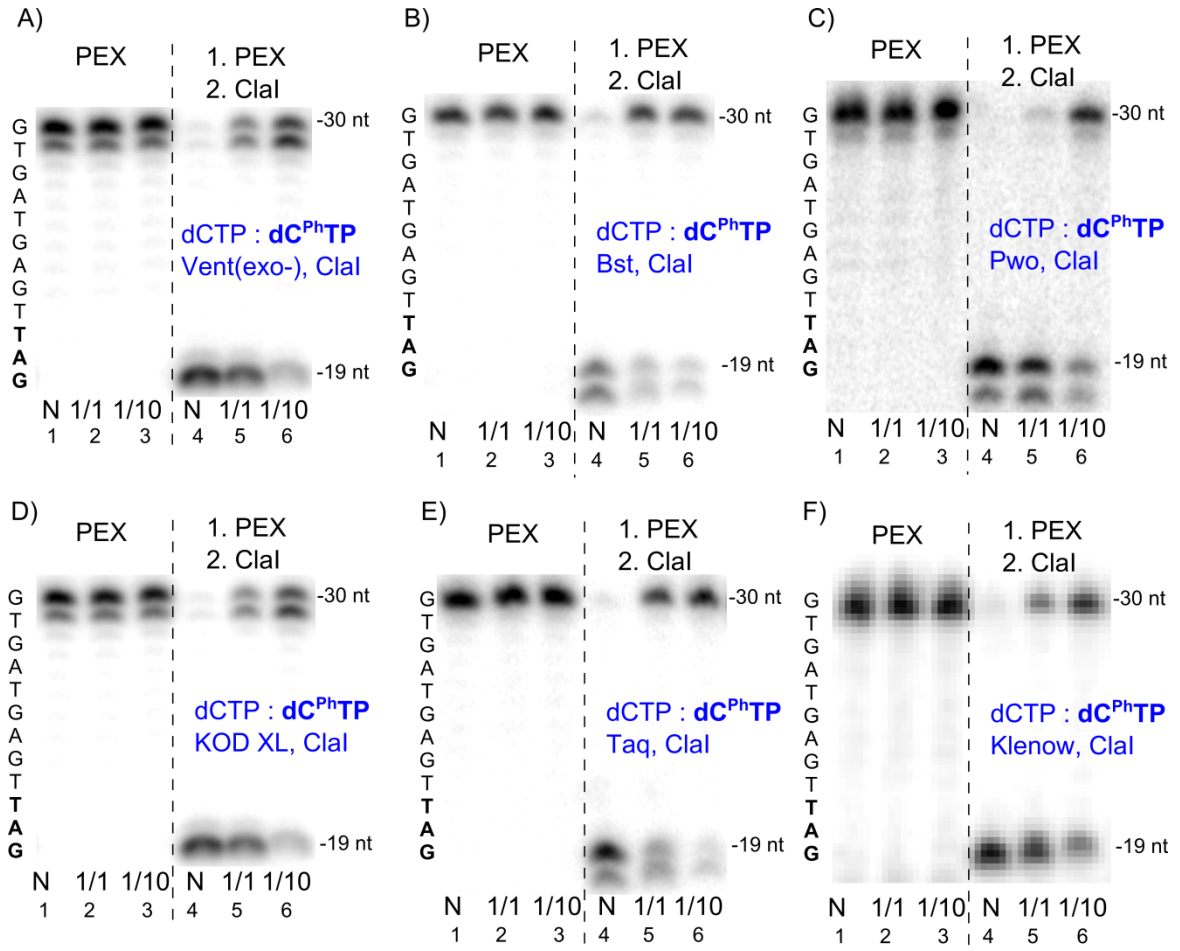


Figure 36. PAGE analyses of PEX experiments and cleavage products with **dC^{Ph}** using different DNA polymerases. Lanes 1, N: product of PEX with natural dNTPs; lanes 2 and 3, 1/1 and 1/10: product of PEX with dTTP, dATP, dGTP and corresponding ratio of dCTP/**dC^{Ph}TP**; lanes 4, N: products of cleavage of unmodified DNA by RE; lanes 5 and 6, 1/1 and 1/10: products of cleavage of modified DNAs by RE.

These promising results with **dA^{Ph}TP** promted us to test other 7-substituted 7-deaza-2'-deoxyadenosines and 5-substituted 2'-deoxycytidines listed in **Figure 33** and **Figure 34**. Competitive incorporations of all modified **dA^{X}TP**s with Bst DNA polymerase showed that **dA^{X}TP**s bearing aromatic groups (except for the very bulky **A^{STr}**) and ethynyl were more efficient substrates than dATP for Bst polymerase. On the other hand, the unsubstituted 7-deazaadenine derivative (**dA^{7D}TP**) was a poorer substrate than dATP probably because of loose of the interaction on N-7 without replacement by other functional group capable to substitute this interaction (**Table 11** and **Figure 37**).

Table 11. Competitive incorporations of **dA^{X}TP** by Bst polymerase[a,b].

N/ M	**dA^{STr}TP**	**dA^{7D}TP**	**dA^{NH2}TP**	**dA^{E}TP**	**dA^{BFU}TP**	**dA^{Ph}TP**	**dA^{NO2}TP**	**dA^{ABOX}TP**
1/1	16.9 (0.9)	32 (2)	61 (1)	67 (2)	67 (4)	67 (2)	70 (2)	73 (5)
1/10	27 (1)	85 (13)	89 (5)	87 (3)	88.3 (0.7)	95 (1)	93.5 (0.5)	89 (2)

[a]Values are in percent of resulting modified DNA. [b] The standard errors are in brackets.

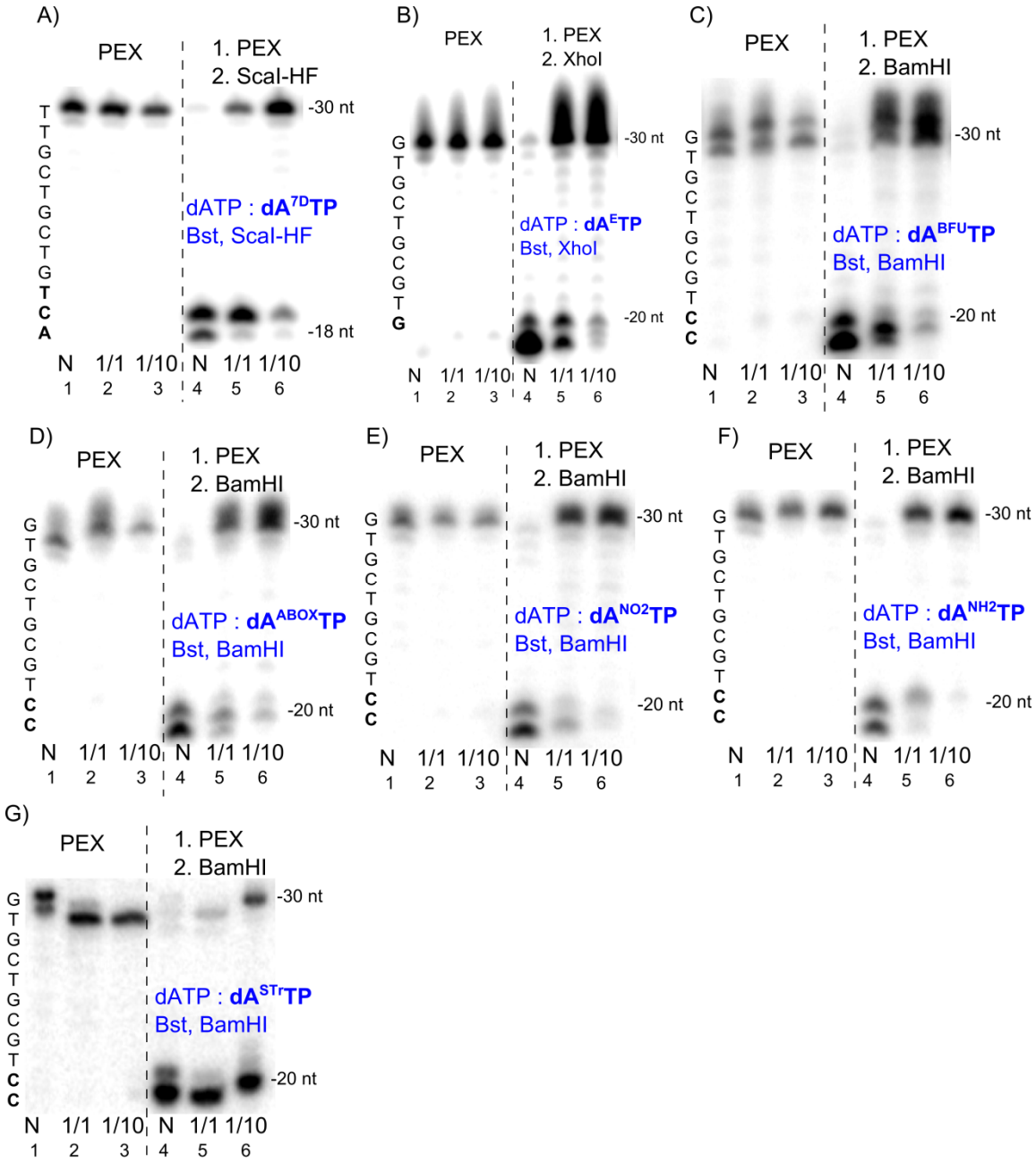


Figure 37. PAGE analyses of PEX experiments and cleavage products with **dA^{X}** using Bst polymerases. Lanes 1, N: product of PEX with natural dNTPs; lanes 2 and 3, 1/1 and 1/10: product of PEX with dTTP, dCTP, dGTP and corresponding ratio of dATP/**dA^{X}TP**; lanes 4, N: products of cleavage of unmodified DNA by RE; lanes 5 and 6, 1/1 and 1/10: products of cleavage of modified DNAs by RE.

In addition, we also determined the incorporation ratio of **dA^{E}TP** by human polymerase α to find that it is also a better substrate than natural dATP (**Table 12** and **Figure 38**) which might have an important implications for use of ethynyl modified adenosines for *in vivo* applications.

Table 12. Competitive incorporation of **dA^{E}TP** by human polymerase α.[a]

dATP/**dA^{E}TP**	1/1	1/10
H Pol α	60	85

[a]Values are in percent of resulting modified DNA.

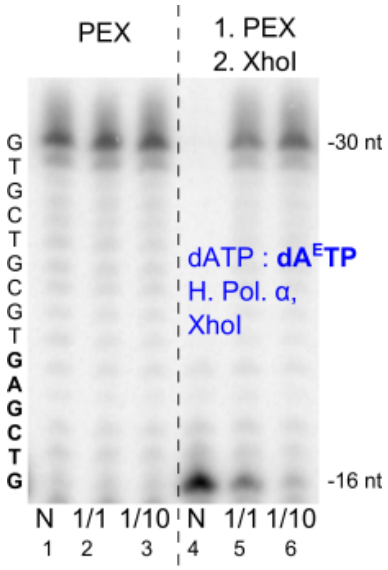


Figure 38. PAGE analyses of PEX experiments and cleavage product with dA^{E} using H. Pol. α. Lanes 1, N: product of PEX with natural dNTPs; lanes 2 and 3, 1/1 and 1/10: product of PEX with dTTP, dCTP, dGTP and corresponding ratio of dATP/dA^{E}TP; lanes 4, N: products of cleavage of unmodified DNA by XhoI; lanes 5 and 6, 1/1 and 1/10: products of cleavage of modified DNAs by RE.

In comparsion to the high incorporation rates of **dA^{X}TP**s in the presence of the dATP, different modified **dC^{X}TP** are worse substrates than natural dCTP for Bst polymerase except slightly better ratios for **dC^{E}TP** and **dC^{Ph}TP** (**Table 13** and **Figure 39**). This difference between incorporation rates of modified **dA^{X}TP**s and **dC^{X}TP**s was reported previously by Marx and co-workers[47]. The reason for this is yet unknown, one possibility would be control of the C-5 substitution at cytidine by DNA polymerase due to the importance of the epigenetic marks at this position.

Table 13. Competitive incorporations of **dC^{X}TP** by Bst polymerase.[a,b]

N/M	**dC^{STr}TP**	**dC^{NO2}TP**	**dC^{NH2}TP**	**dC^{FT}TP**	**dC^{E}TP**	**dC^{Ph}TP**
1/1	5 (2)	11.5 (0.7)	32 (6)	44 (2)	54 (3)	58 (5)
1/10	14 (3)	39 (4)	70 (0.5)	63 (6)	74 (5)	76 (3)

[a]Values are in percent of resulting modified DNA. [b] The standard errors are in brackets.

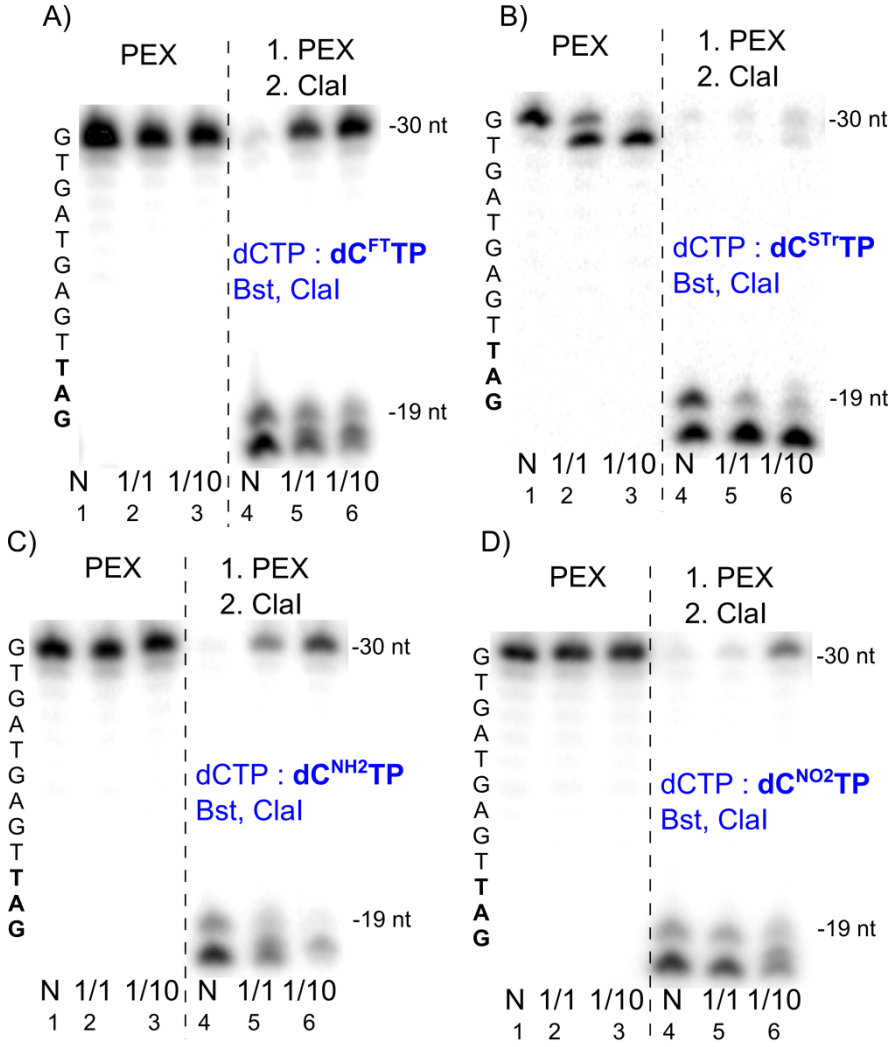


Figure 39. PAGE analyses of PEX experiments and cleavage products with **dC^{X}** using Bst polymerases. Lanes 1, N: product of PEX with natural dNTPs; lanes 2 and 3, 1/1 and 1/10: product of PEX with dTTP, dATP, dGTP and corresponding ratio of dCTP/**dC^{X}TP**; lanes 4, N: products of cleavage of unmodified DNA by ClaI; lanes 5 and 6, 1/1 and 1/10: products of cleavage of modified DNAs by RE.

The analysis of competitive PEX in different ratios of dNTPs with Bst polymerase showed that **dA^{Ph}TP** is efficiently incorporated (34%) even in the presence of a large excess of dATP

1:10. The **dC^{Ph}TP** analog followed the same trends as described above, but the incorporation ratios were lower (**Table 14** and **Figure 40**).

Table 14. Competitive incorporations by Bst.[a]

ratio dCTP/**dC^{Ph}TP**	10/1	5/1	2/1	1/1	1/2	1/5	1/10
dC^{Ph}	24	32	45	60	71	75	79
ratio dATP/**dA^{Ph}TP**	10/1	5/1	2/1	1/1	1/2	1/5	1/10
dA^{Ph}	34	44	56	69	77	90	94

[a]Values are in percent of resulting modified DNA.

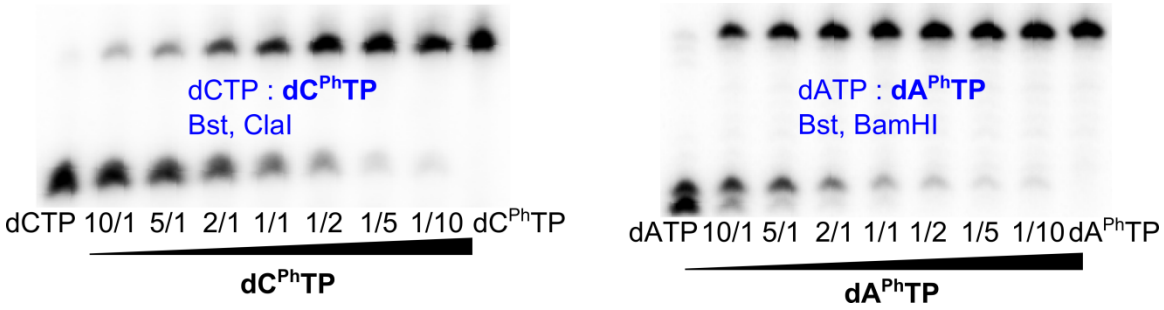


Figure 40. PAGE analyses of cleavage products using Bst polymerase. dCTP and dATP: products of PEX with natural dNTPs followed by cleavage by RE; 1/1 to 1/10: products of cleavage of modified DNA by RE with dTTP, dCTP or dATP, dGTP and corresponding ratio of dCTP/**dC^{Ph}TP** or dATP/**dA^{Ph}TP**, respectively.

We wanted to find out if the incorporation efficiency is sequence dependent and therefore three out of four possible combinations were tested, **dA^{Ph}** was incorporated after G, C and T and **dC^{Ph}** after T, G and A. Results are summarized in **Table 15** and **Table 16**. The efficiency of incorporation of **dA^{Ph}TP** and **dC^{Ph}TP** was sequence-dependent, but it always followed the same trend.

Table 15. Template dependence of competitive incorporations for **dA^{Ph}TP** incorporated by Pwo.[a,b]

Sequence	5'-…**GA^{Ph}**…-3'		5'-…**CA^{Ph}**…-3'		5'-…**TA^{Ph}**…-3'	
Ratio	1/1	1/10	1/1	1/10	1/1	1/10
% of modified DNA	43 (2)	79 (3)	35 (3)	68 (2)	53 (2)	75 (1)

[a]Values are in percent of resulting modified DNA. [b] The standard errors are in brackets.

Table 16. Template dependence of competitive incorporations for **dC^{Ph}TP** incorporated by Pwo.[a,b]

Sequence	5'-…**T**C^{Ph}…-3'		5'-…**G**C^{Ph}…-3'		5'-…**A**C^{Ph}…-3'	
Ratio	1/1	1/10	1/1	1/10	1/1	1/10
% of modified DNA	14 (7)	59 (2)	15 (3)	46 (3)	40 (2)	65 (4)

[a]Values are in percent of resulting modified DNA. [b] The standard errors are in brackets.

The same approach was also used for analysis of competitive PCR with **dA^{Ph}TP** or **dC^{Ph}TP** using KOD XL polymerase and RsaI for cleavage, which led to the same conclusion. The incorporation of **dC^{Ph}TP** in ratio 1:1 with natural dCTP resulted in 25 % of modified DNA while the incorporation of **dA^{Ph}TP** in ratio 1:1 with natural dATP resulted in 50 % of modified DNA (**Table 17** and **Figure 41**). The lower incorporation ratio of **dA^{Ph}TP** with KOD XL in PCR than in PEX could be caused by partial inhibition of cleavage reaction by modified nucleobases next to the restricition site.

Table 17. Competitive incorporations in PCR using KOD XL.[a]

N/ M	+	10/1	5/1	2/1	1/1	1/2	1/5	1/10
dC^{Ph}	0	0	0	9	25	44	66	81
dA^{Ph}	0	9	17	33	50	67	83	91

[a]Values are in percent of resulting modified DNA.

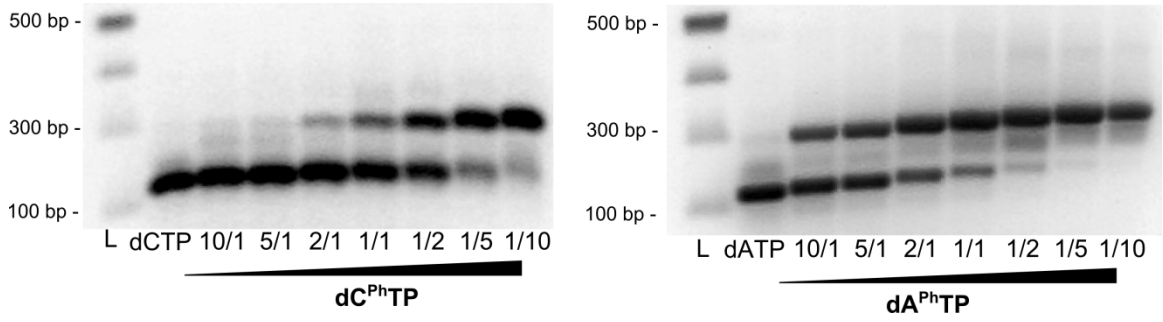


Figure 41. Analyses of competitive PCR using KOD XL followed by cleavage by RsaI. dCTP and dATP: cleaved unmodified products of PCR; 1/1 to 1/10: products of cleavage of modified DNA by RE.

3.3.3 Steady State Kinetic Assays

Because of the surprisingly high efficiency of 7-aryl **dA^{X}TP**s in competitive incorporations we were curious if the affinities of the incoming modified triphosphates to the active site of the Bst polymerase are higher than the affinity of the natural dATP. Thus we studied the kinetics of the single nucleotide incorporations.[98] Results for different **dN^{X}TP**s are summarized in the **Table 18**. The kinetics experiments revealed that the substrates which are better than natural substrates in competitive incorporations have lower K_m values thus higher affinity to the active site and also give higher values of the V_{max}/K_m in comparsion to the natural counterparts.

Table 18. Bst DNA polymerase.[a]

compound	K_m[b]	V_{max}[c]	V_{max} / K_m	discrimination[d]	competition[e]
dCTP	4.4 (0.8)	1.20 (0.07)	0.27	-	-
dATP	28 (4)	1.8 (0.1)	0.06	-	-
dC^{Ph}TP	2.4 (0.2)	1.32 (0.02)	0.54	2	58
dC^{NH2}TP	21 (2)	1.34 (0.04)	0.06	0.2	32
dC^{E}TP	3.1 (0.2)	1.29 (0.02)	0.41	1.5	54
dA^{NO2}TP	6.5 (0.5)	1.57 (0.04)	0.23	3.8	70
dA^{BFU}TP	5.7 (0.8)	1.27 (0.05)	0.22	3.7	67
dA^{Ph}TP	6.8 (0.8)	1.19 (0.04)	0.18	3	67
dA^{NH2}TP	7 (1)	1.02 (0.05)	0.14	2.3	61
dA^{E}TP	13 (2)	1.51 (0.06)	0.12	2	67
dA^{ABOX}TP	8 (2)	0.94 (0.09)	0.11	1.8	73
dA^{7D}TP	60 (16)	1.0 (0.1)	0.02	0.3	32
dA^{STr}TP	>1000	-	-	-	17

[a] The standard errors are in brackets. [b] K_m extension is μM. [c] V_{max} extension is 10 pmol.s^{-1}.U^{-1}. [d] Discrimination reflects the efficiency of functionalized **dA^{X}TP** incorporated compared to the efficiency of incorporation of the dATP, deduced as $(V_{max}/K_m)_{fnc}$ / $(V_{max}/K_m)_{natural}$. [e] Percent of functionalized DNA.

In order to compare the kinetics of **dA^{Ph}TP** for different DNA polymerases we tested all DNA polymerases used in competitive incorporations with this representative **dA^{Ph}TP** and compared the data (**Table 19**). In most cases, except for KOD XL and Vent(exo-), the V_{max}/K_m values were higher than for natural dATP which is in good correlation with results

from competitieve incorporations. A similar correlation was reported for analysis of misincorporations.[99] However, the PEX (and even more so PCR) is a much more complex process, where the efficiency depends not only on the rate of incorporation but also on the presence or absence of proofreading (no correlation was observed for this factor) and on further extension of the primer.

Table 19. Kinetics with different DNA polymerases.[a]

	dATP			**dA^{Ph}TP**			
	K_m[b]	V_{max}[c]	V_{max}/K_m	K_m[b]	V_{max}[c]	V_{max}/K_m	discr.[d]
Bst	28 (4)	1.8 (0.1)	0.06	6.8	1.19	0.179	3
Pwo	10 (1)	0.85 (0.03)	0.083	24 (5)	0.81 (0.07)	0.034	0.4
Taq	12 (1)	0.49 (0.02)	0.039	8.9 (0.4)	0.371 (0.005)	0.042	1.1
Vent(exo-)	8.7 (0.9)	0.67 (0.01)	0.077	24 (2)	0.73 (3)	0.031	0.4
Klenow	2.5 (0.6)	0.25 (0.01)	0.096	2.1 (0.2)	0.49 (0.1)	0.233	2.4
KOD XL	18 (2)	4.6 (0.2)	0.263	38 (6)	6.3 (0.4)	0.167	0.6
H Pol α	45 (8)	0.00095 (8.10^{-4})	2.10^{-4}	>100	-	-	-

[a] The standard errors are in brackets. [b] K_m extension is μM. [c] V_{max} extension is 10 pmol.s^{-1}.U^{-1}. [d] Discrimination reflects the efficiency of functionalized **dA^{X}TP** incorporated compared to the efficiency of incorporation of the dATP, deduced as $(V_{max}/K_m)_{fnc}$ / $(V_{max}/K_m)_{natural}$.

3.3.4 Computational Studies

In order to explain the higher affinity of the 7-aryl **dA^{X}TP**s to the active site of polymerases, molecular modeling studies were performed. The computational studies were done by Dr. Jindřich Fanfrlík. Both dATP and **dA^{Ph}TP** were docked to the complex of Bst polymerase with primer and template using the known crystal structure (pdb 4BDP).[100] The obtained tertiary complexes were studied using a semiempirical quantum mechanical scoring function. The resulting optimized structures are shown in **Figure 42**. The phenyl ring of **dA^{Ph}TP** filled the space between the aliphatic chain of Arg629 and the phenyl group of Phe710. The **dA^{Ph}TP** had a larger affinity (more negative score) than dATP (-72.8 and -65.0 kcal.mol^{-1}, respectively), which indicates a possible increase in π-π stacking as a result of the 7-phenyl group. A similar binding mode and increased stacking is expected for other 7-aryl-7-

deaza-dATP analogues and it is in accord with the high affinity of these **dA^{X}TP**s to this polymerase.

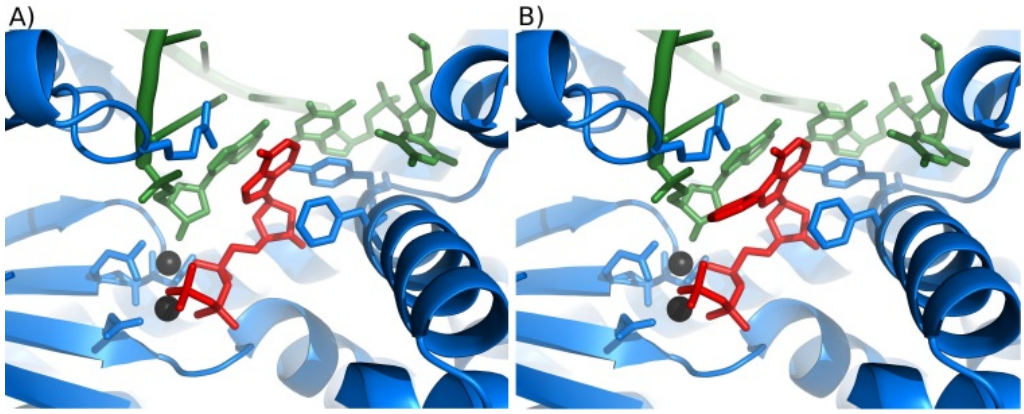


Figure 42. Bst polymerase active site with modeled A) dATP and B) **dA^{Ph}TP**.

In conclusion, we have developed a method that is based on cleavage of the DNA by RE and PAGE analysis for the analysis of products generated from competitive incorporations of a modified **dN^{X}TP** in the presence of the corresponding natural dNTP. Our systematic study of competitive PEX using different **dA^{X}TP**s and **dC^{X}TP**s in different ratios with the natural dATP and dCTP revealed a surprising and counterintuitive result that most 7-aryl **dA^{X}TP**s are better substrates for most polymerases than dATP, whereas most **dC^{X}TP**s are worse substrates. The kinetic study and semi-empirical calculations explain this by increased affinity of the 7-aryl **dA^{X}TP**s to the active site of the polymerase complex with the primer and template because of increased stacking. These findings are important for better understanding the mechanism of the incorporation of modified dNTPs by polymerase, as well as for the design of modified nucleosides and nucleotides for *in vivo* synthesis of base-modified DNAs.

3.4 Study of Possible Metabolic DNA Labeling Using Modified 2'-Deoxyribonucleosides and Pronucleotides

The study of nucleic acids in their native environment is becoming increasingly important in basic research and modern drug discovery. Metabolic labeling of DNA has traditionally been performed using [^{3}H]thymidine or 5-bromo-2'-deoxyuridine (BrdU).[41] These labels are limited in terms of their subsequent visualization, requiring either autoradiography, or DNA denaturation and antibody staining. BrdU immunostaining is currently the most commonly used method, but it requires harsh chemical denaturation of cellular DNA and is limited by the poor tissue penetration of the BrdU antibody. In addition, BrdU itself is both toxic and mutagenic when applied at high concentrations, and it can have a negative impact on DNA stability and the cell cycle. The recent emergence of bioorthogonal chemical reporter strategies has revolutionized the study of biological macromolecules in their native environments. The most commonly used chemoselective reactions for probe conjugation include the Staudinger ligation, copper(I)-catalyzed azide–alkyne cycloaddition (CuAAC), and strain-promoted azide–alkyne cycloaddition reactions. Luedtke and co-workers recently prepared and used nucleoside based probes with attached alkyne for the *in vivo* DNA labeling.[42,43] The successful nucleoside based probe must penetrate through the cell membrane, be phosphorylated to the corresponding **dN^{X}TP** analog by cellular kinases, get accepted by DNA polymerase as substrate and posses low cytotoxicity.

Based on our study of competitive incorporations[94] of modified **dN^{X}TP**s by DNA polymerase in which we found that many 7-aryl **dA^{X}TP**s are better substrates for DNA polymerases than a natural counterpart we have decided to try one step metabolic labeling of genomic and plasmid DNA using fluorescent **dA^{BFU}**,[12a] as a control label we have used **dA^E** and the study has been extended with **dA^{Ph}** probe as an example of the minimalistic 7-aryl modification. All probes used in the study are illustrated in **Figure 43**. In order to fulfill the criteria for successful labeling listed above we have decided to prepare also corresponding pronucleotides[101] which were reported to penetrate easily through the cell membrane and bypass the first phosphorylation by kinases which is considered to be the most difficult part of the overall process. The rationale for using **dA^{BFU}** for metabolic labeling is to avoid the necessity of the click-reaction chemistry used for addition of the fluorescent tag to alkyne based probes.

The nucleoside analogs were prepared modifying the literature procedures reported previously by McGuigan and co-workers.[102] Compounds **Pro-dA^E**, **Pro-dA^{Ph}** and **Pro-dA^{BFU}**

were prepared by reaction of modified 7-deaza-2'-deoxyadenosines with alanine-ester-phosphorochloridate[103] **5** in presence of the 1-methylimidazol in anhydrous THF. The mild conditions were used to avoid reaction at 2° alcohol and also led to the lower yields, on the other hand bypassed laborious the 2° alcohol protection based strategy (**Scheme 17**).

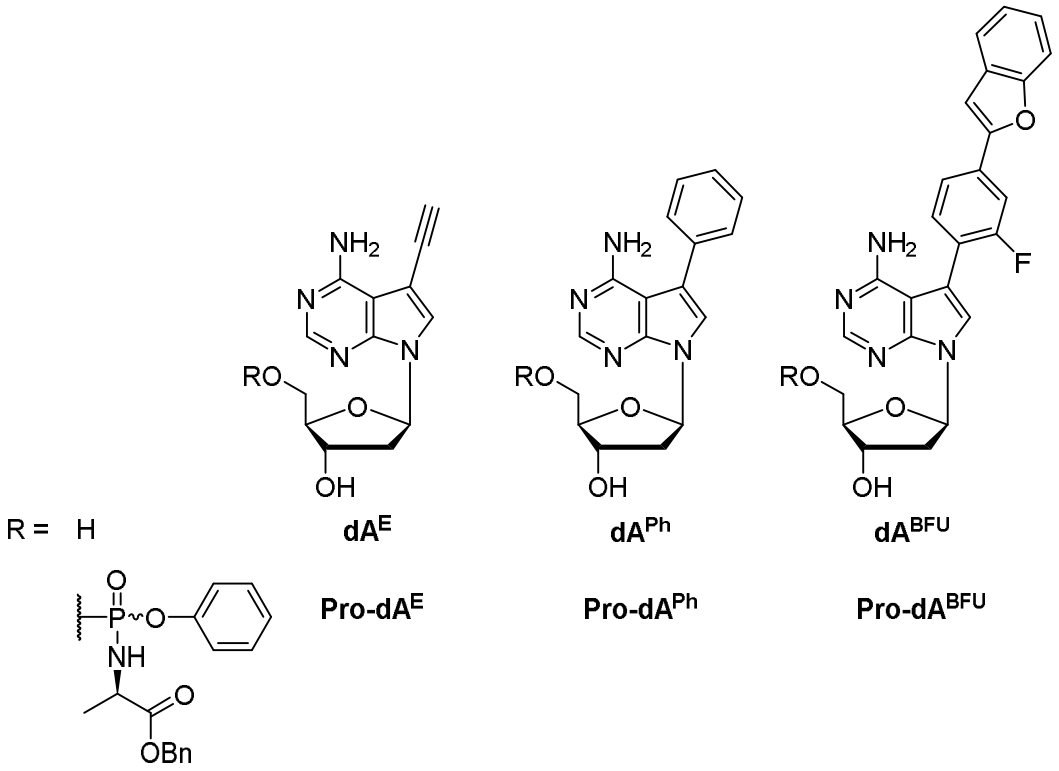
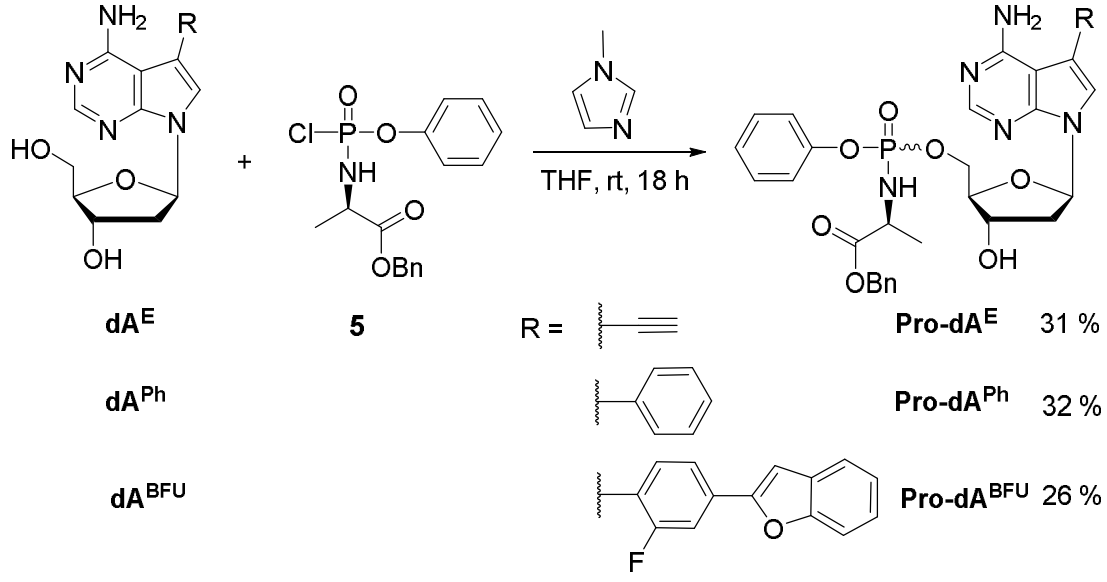


Figure 43. Nucleosides and pronucleotides used in the study.



Scheme 17. Design and synthesis of the pronucleotides for *in vivo* labeling.

The nucleosides **dA^{E}**, **dA^{BFU}** and pronucleotides **Pro-dA^{E}**, **Pro-dA^{BFU}** were tested for metabolic labeling of DNA in HeLa cells. The screening was carried out by Ivana Křížová. The cells were incubated with 10μM of each nucleoside or pronucleotide for 24 h and the **dA^{E}** and **Pro-dA^{E}** were subsequently stained with a fluorescent azide (AlexaFluor 488) using CuAAC reaction (**Figure 44** and **Figure 45**). The **dA^{E}** and **Pro-dA^{E}** gave intensive nuclear staining as they were expected (**Figure 45**). In contrast, both **dA^{BFU}** and **Pro-dA^{BFU}** unfortunately did not resulted in staining of cell nucleus, see **Figure 45**, they were rather caged in an unknown small circular structures around the nucleus. Because of these unsuccessful results in HeLa cells we moved on to the more simple system and tried the labeling of DNA in *E. coli* with **dA^{BFU}** and **Pro-dA^{BFU}**. Both probes penetrate into the cells but we have not observed any colocalization between the DAPI nuleus staining and our stain.

Finally, we have tried several methods for assessment if the modified nucleotides are present in plasmid isolated from E. coli treated with **dA^{E}**, **Pro-dA^{E}**, **dA^{Ph}**, **Pro-dA^{Ph}**, **dA^{BFU}**, **Pro-dA^{BFU}**. All methods we have proposed and tested were based on the REs and unfortunately the methods have shown to be unsuccessful so far.

The problems with metabolic incorporation of artificial 2'-deoxyribonucleosides into the cell DNA are general. Recently, similar problem was reported by Romesberg and co-workers in their work on extension of genetic alphabet in bacteria.[39d] They circumvented this problem by use of an enzymatic transport system for dNTPs through the cell membrane. Further development of modified 2'-deoxyribonucleoside and/or transport systems for metabolic labelling of DNA was out of the scope of this project, but the research line will be followed further in the Hocek group.

In conclusion, we found that both nucleosides and pronucleosides penetrate into the cell. The **dA^{E}** and **Pro-dA^{E}** are also metabolically phosphorylated and incorporated into the genomic DNA in HeLa cells. The more bulky **dA^{BFU}** and **Pro-dA^{BFU}** penetrate into the cell but were not incorporated in nucleus into the cellular DNA.

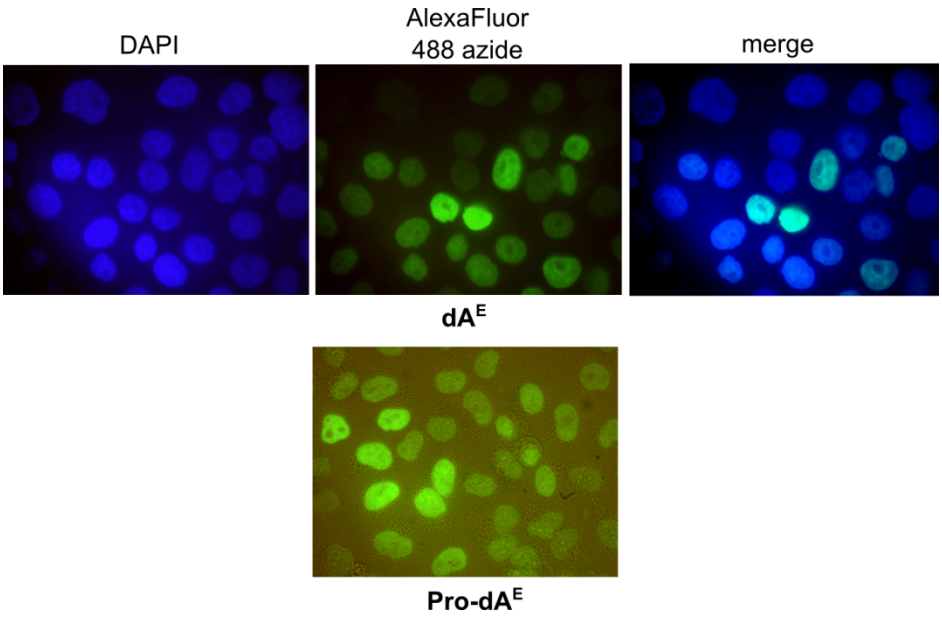


Figure 44. Metabolic labeling of genomic DNA in HeLa cells following a 24 h incubation with 10 μM of **dA**^{E} or **Pro-dA**^{E}. After washing and fixing the cells, ethynyl-modified DNA was stained using AlexaFluor 488 azide and Cu(I), and total cellular DNA was stained with DAPI. Negative controls received identical treatments, but were not exposed to a synthetic nucleoside.

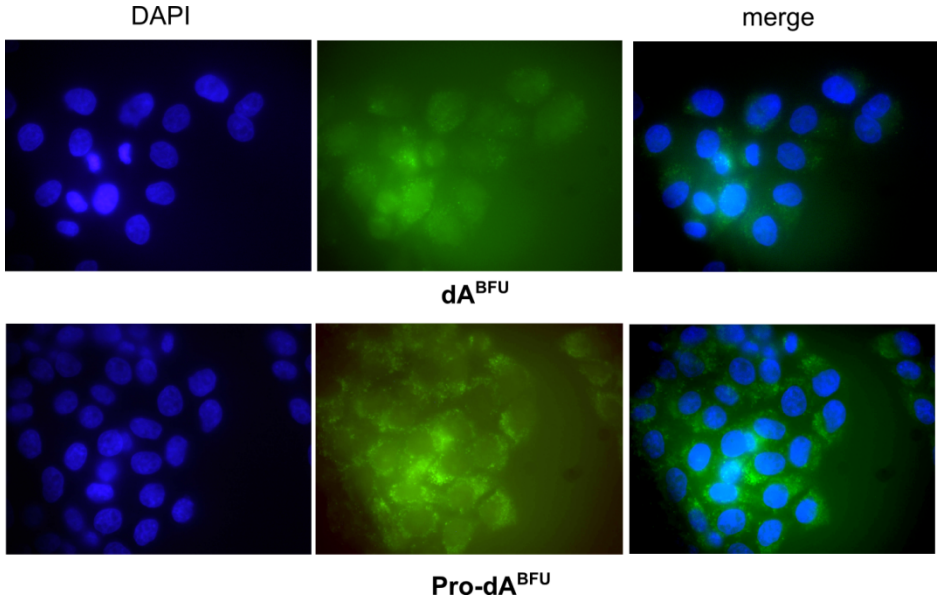


Figure 45. Metabolic labeling of genomic DNA in HeLa cells following a 24 h incubation with 10 μM of **dA**^{BFU} or **Pro-dA**^{BFU}. After washing and fixing the cells, the **dA**^{BFU} was measured and then cellular DNA was stained with DAPI. Negative controls received identical treatments, but were not exposed to a synthetic nucleoside.

4 Conclusion

The synthesis of the novel modified 2'-deoxyribonucleoside triphosphates which were incorporated into the DNA and used in several chemical biology applications were developed in this work.

Synthesis of double-headed nucleosides and nucleotides was performed. The synthetic approach was based on aqueous Sonogashira cross-coupling reaction of 5-iodopyrimidine and 7-iodo-7-deazaadenine 2'-deoxyribonucleosides or corresponding triphosphates with 5-ethynylcytosine, 1-*N*-(prop-2-yn-1-yl)cytosine or 5-fluoro-1-*N*-(prop-2-yn-1-yl)cytosine. The modified 2'-deoxyribonucleoside triphosphates were then tested in PEX and PCR. They were found to be good substrates for DNA polymerase in both PEX and PCR. Thus the double-headed nucleobases containing different linkers are suitable derivatives for further investigation of interactions between such modified DNA and DNA methyltransferases. Following this line, the assay for testing of inhibition of DNA methyltransferases with DNA containing double-headed nucleobases was developed. Taken together, the general design, preparation of DNA modified with double-headed nucleobases and assay for testing the inhibition of DNA methyltransferases was successfully studied but the tested $\mathbf{dC^{PC}}$ and $\mathbf{dC^{PFC}}$ derivatives did not show any significant inhibition of DNA methyltransferase M.SssI. Nevertheless, the research along this line will be continued in the Hocek group.

Transient protection of the DNA against cleavage by restriction endonucleases the series of three 7-(trialkylsilyl)ethynyl-7-deazaadenines was tested as the triggerable switch. All three 7-(trialkylsilyl)ethynyl-7-deaza-2'-deoxyadenosine triphosphates were synthesized by use of Sonogashira cross-coupling reaction and tested on incorporation into DNA in PEX or PCR. For the PEX all three $\mathbf{dA^{X}TP}$s worked well, while in PCR bulky triisopropyl derivative was an unsuitable substrate. Deprotection conditions of the trialkysilyl protecting groups were optimized first on corresponding nucleoside monophosphates and then on trialkylsilylethynyl modified DNA. The following study of the ability of (trialkylsilyl)ethynyl modified DNA to protect the DNA cleavage by REs (KpnI, RsaI, SacI) revealed that the trimethylsilyl group was somehow insufficient and DNA was partially cleaved by RsaI and KpnI while the triethylsilyl and triisopropyl groups fully inhibit the cleavage reaction with tested restriction endonucleases. After treatment of the (triisopropypl)ethynyl modified DNA with TBAF and (trimethyl)ethynyl and (trimethyl)ethynyl modified DNA with aqueous ammonia and subsequent purification the resulting ethynyl modified DNA was again cleaved with all tested restricition endonucleases. Finally, it was shown that 287-bp DNA with recognition sequence

for RsaI modified with (triethylsilyl)ethynyl on 7-deazaadenosines is fully protected against the cleavage by RsaI, but after deprotection of the triethylsilyl protecting group by aqueous ammonia, evaporation of the excess of the ammonia the recognition sequence is cleaved again. The last experiment clearly shows that the (triethylsilyl)ethynyl modified 7-deaza-2'-deoxyadenosine is suitable for transient protection of DNA against cleavage by REs. The reported proof-of-principle experiment also for the first time presents the protecting group traditionally used in organic synthesis applied in regulation of biochemical reaction.

This methodology was then applied for selective protection of the internal NcoI recognition site of the *tdaD* gene which contain couple of the NcoI recognition sites. The method consist of the preparation of two PCR products which overlaps and they are fully modified with 7-(triethylsilyl)ethynyl-7-deaza-2'-deoxyadenosine except the primer sequences. These two products are then used in the third PCR with natural dNTPs resulting in the desired partially modified DNA. This was cloned into plasmid and successfully amplified in *E. coli* by bacterial DNA polymerases after transformation and the thioester hydrolase tdaD was subsequently produced. Here reported method is an alternative to other methods for cloning of the genes containing multiple recognition sites for REs, but the main aspect of this work is in the general use of base modifications as bioorthogonal protection against specific interactions with proteins and therefore regulation of biological processes.

Since the competitive incorporations of the modified **dN^{X}TP**s in presence of their natural counterparts by DNA polymerases were investigated only scarcely it was assumed that the modified **dN^{X}TP**s are inherently worse substrates than natural dNTPs. In this work I found that several 7-aryl-7-deazaadenine 2'-deoxyribonucleoside triphosphates are better substrates for many DNA polymerases than their natural counterparts. First the new method was developed for analysis of competitive incorporations. Because of many REs cannot cleave modified bases in their recognition sequences it was possible to selectively cleave only unmodified DNA in mixture resulting from the direct competitive incorporation in PEX. Our systematic study of competitive PEX using different **dA^{X}TP**s and **dC^{X}TP**s in different ratios with the natural dATP and dCTP, respectively, revealed a surprising and counterintuitive result that most tested 7-aryl **dA^{X}TP**s are better substrates for DNA polymerases than dATP, whereas most **dC^{X}TP**s are worse substrates. The highest incorporation rates 70 and 73 % were achieved using **dA^{NO2}TP** and **dA^{BFU}TP**, respectively, incorporated by Bst polymerase in ratio 1:1 with natural dATP. The kinetic study and semi-empirical calculations explain this by increased affinity of the 7-aryl **dA^{X}TP**s to the active site of the polymerase complex with the

primer and template because of increased π-π stacking. This study gave us better insight into the mechanism of the incorporation of modified dNTPs by DNA polymerases, as well as opened a new window for the design of modified nucleosides and nucleotides for *in vivo* synthesis of base-modified DNA.

5 Experimental Section

5.1 General Remarks

Cross-coupling reactions were performed in evacuated flame-dried glassware with magnetic stirring under argon atmosphere. THF, toluene and diethylether were dried and distilled from sodium/benzophenone. Other reagents were purchased from commercial suppliers and used as received. NMR spectra were measured on Bruker Avance 400 MHz spectrometer (400.1 MHz for ^{1}H and 100.6 MHz for ^{13}C) or Bruker Avance 500 MHz spectrometer (499.8 MHz for ^{1}H, 125.7 MHz for ^{13}C, 202.3 MHz for ^{31}P) or Bruker Avance 600 MHz spectrometer (600.1 MHz for ^{1}H and 150.9 MHz for ^{13}C) in $CDCl_3$ (TMS was used as internal standard), MeOH-*d4* (referenced to the residual solvent signal), DMSO-*d6* (referenced to the residual solvent signal), D_2O (methanol as internal standard, referenced to C*H*$_3$OH singlet 3.34 ppm and to *C*H$_3$OH signal 49.5 ppm). NMR spectra of the triphosphates were measured in D_2O (dioxane and H_3PO_4 as internal standards, referenced to singlet 3.75 ppm and to 69.3 ppm). Chemical shifts are given in ppm (δ-scale), coupling constants (*J*) in Hz. Complete assignment of all NMR signals was performed using a combination of H,H-COSY, H,H-ROESY, H,C-HSQC and H,C-HMBC experiments. IR spectra were recorded on Bruker Alpha FT-IR spectrometer using attenuated total reflection (ATR). Both low resolution and high resolution mass spectra were measured using electrospray ionization. Melting points were determined on a Kofler block and are uncorrected. Semi-preparative HPLC separations were performed on column packed with 10 μm C18 reversed phase (Phenomenex, Luna C18(2)).

5.2 Double-headed Nucleosides and Nucleotides: Synthesis and Enzymatic Incorporations

4-Amino-5-ethynyl-1*H*-pyrimidin-2-one (1)

DMF (8 ml) was added through a septum to an argon-purged vial containing: 5-iodo-cytosine (**4)** (2.5 g, 10.5 mmol), $Pd(PPh_3)_2Cl_2$ (369 mg, 0.5 mmol), CuI (200 mg, 1 mmol), Et_3N (5.3 g, 52.5 mmol). Then a trimethylsilylacetylene (3.1 g, 31.5 mmol) was added to the solution. The mixture was stirred at ambient temperature for 20 min. Product was precipitated from the reaction mixture using MeOH, filtered off and washed with water (3 x 100 ml). Dried product was directly used in the following reaction with NH_3 aq. (40 ml) in methanol (10 ml) at

ambient temperature. After 48 h NH_3 and solvents were evaporated under reduced pressure. It give the product **2** as white solid in 79 % yield. ^{1}H NMR (500.0 MHz, DMSO-d_6): 4.28 (s, 1H, HC≡C); 6.66, 7.56 (2 × bs, 2 × 1H, NH_2); 7.74 (s, 1H, H-6); 10.98 (bs, 1H, NH). ^{13}C NMR (125.7 MHz, DMSO-d_6): 76.4 (-**C**≡CH); 85.9 (H**C**≡C-); 87.5 (C-5); 148.3 (CH-6); 155.6 (C-2); 165.3 (C-4). MS (ESI$^+$): *m/z (%)*: 158.0 (100) $[M + Na]^+$, 180.0 (38) $[M + 2Na]^+$. HRMS (ESI$^+$): *m/z* $[M + H]^+$ calculated for $C_6H_6ON_3$ = 136.05054; found 136.05041. IR: υ^{-cm} = 1649, 1452, 1232, 1148, 1062, 903 cm^{-1}.

Synthesis of modified 2'-deoxyribonucleosides using Sonogashira-coupling. General Procedure A:

A water/acetonitrile mixture 2:1 (2 ml) was added through a septum to an argon-purged vial containing halogenated **dN^I** (0.22 mmol), 5-ethynylcytosine (**1**) (61 mg, 0.45 mmol), CuI (4 mg, 0.02 mmol) and *i*Pr_2NEt (0.4 ml, 2.2 mmol). After the solids had dissolved, a solution of $Pd(OAc)_2$ (2 mg, 0.01 mmol) and TPPTS (32 mg, 0.05 mmol) in water/acetonitrile mixture 2:1 (1 ml) was injected into the reaction mixture and the resulting mixture was stirred with heating 80 °C for 1 h. The products were isolated from crude reaction mixture by column chromatography on silica gel, eluent gradient from 10 % MeOH in $CHCl_3$ to MeOH and reverse phase chromatography, eluent linear gradient from H_2O to MeOH to gave the products **dC^C**, **dFP^C** and **dA^C**.

Synthesis of modified 2'-deoxyribonucleosides using Sonogashira-coupling. General Procedure B:

A DMF (2.5 ml) was added through a septum to an argon-purged vial containing halogenated **dN^I** (0.28 mmol), 5-ethynylcytosine (**1**) (38.8 mg, 0.28 mmol), $Pd(PPh_3)_2Cl_2$ (10 mg, 0.01 mmol), CuI (5 mg, 0.02 mmol) and Et_3N (0.7 ml, 2.8 mmol) and the reaction mixture was stirred with heating 80 °C for 1 h. The products were isolated from crude reaction mixture by precipitation with MeOH (3 ml). The solid was filtered off and washed with water (3 x 30 ml), MeOH (2 x 30 ml) and EtOH (2 x 30 ml). The products (**dC^C**, **dFP^C** and **dA^C**) were dried under reduced pressure.

Synthesis of modified 2'-deoxyribonucleosides using Sonogashira-coupling. General Procedure C:

A CH_3CN/H_2O 1:2 (2 ml) was added through a septum to an argon-purged vial containing halogenated **dN^I** (1 eq.), 4-amino-1-(prop-2-yn-1-yl)pyrimidin-2-one (**2**) (2 eq.) or 4-amino-5-fluoro-1-(prop-2-yn-1-yl)pyrimidin-2-one (**3**) (2 eq.), $Pd(OAc)_2$ (0.01 eq.), TPPTS (0.04 eq.), CuI (0.01 eq.) and *i*Pr_2NEt (10 eq.) and the reaction mixture was stirred with heating 80 °C for 1 h.

4-Amino-5-(4-amino-1*H*-pyrimidin-2-on-5-yl)ethynyl-7-(2-deoxy-β-D-ribofuranosyl)-7*H*-pyrrolo[2,3-d]pyrimidine (dA^C)

Preparation according to the general procedure A gave yellowish solid in 38 % yield. For the preparation by general procedure B the yield was 51 %. ^{1}H NMR (500.0 MHz, DMSO-d_6): 2.20 (ddd, 1H, J_{gem} = 13.2, $J_{2'b,1'}$ = 6.0, $J_{2'b,3'}$ = 2.9, H-2'b); 2.44 (ddd, 1H, J_{gem} = 13.2, $J_{2'a,1'}$ = 8.1, $J_{2'a,3'}$ = 5.9, H-2'a); 3.51, 3.57 (2 × dd, 2 × 1H, J_{gem} = 12.0, $J_{5',4'}$ = 4.6, H-5'); 3.83 (td, 1H, $J_{4',5'}$ = 4.6, $J_{4',3'}$ = 2.6, H-4'); 4.34 (ddd, 1H, $J_{3',2'}$ = 5.9, 2.9, $J_{3',4'}$ = 2.6, H-3'); 5.08, 5.33 (2 × bs, 2 × 1H, OH-3',5'); 6.69, 7.23 (2 × bs, 2 × 2H, NH_2); 6.49 (dd, 1H, $J_{1',2'}$ = 8.1, 6.0, H-1'); 7.82 (s, 1H, H-8); 7.92 (s, 1H, H-6'''); 8.11 (s, 1H, H-2); 10.95 (bs, 1H, NH). ^{13}C NMR (125.7 MHz, DMSO-d_6): 39.9 (CH_2-2'); 62.2 (CH_2-5'); 71.2 (CH-3'); 83.3 (C5'''-**C**≡C-C7); 83.3 (CH-1'); 87.4 (C7-**C**≡C-C5'''); 87.7 (CH-4'); 88.8 (C-5'''); 95.5 (C-7); 102.2 (C-5); 127.1 (CH-8); 147.5 (CH-6'''); 149.6 (C-4); 152.9 (CH-2); 155.6 (C-2'''); 157.9 (C-6); 165.2 (C-4'''). MS (ESI$^+$): *m/z (%)*: 384.5 (22) $[M + H]^+$, 406.5 (100) $[M + Na]^+$, 428.6 (31) $[M + 2Na]^+$. HRMS (ESI$^+$): *m/z* $[M + H]^+$ calculated for $C_{17}H_{18}O_4N_7$ = 384.14148; found 384.14150. IR: υ^{-cm} = 3313, 3194, 1634, 1593, 1455, 1302, 1205, 1177, 1086, 1058, 926 cm^{-1}.

4-Amino-5-(4-amino-1*H*-pyrimidin-2-on-5-yl)ethynyl-1-(2-deoxy-β-D-ribofuranosyl)pyrimidin-2-one (dC^C)

Preparation according to the general procedure A gave yellow solid in 49 % yield. For the preparation by general procedure B the yield was 73 %. ^{1}H NMR (500.0 MHz, DMSO-d_6): 2.00 (ddd, 1H, J_{gem} = 13.5, $J_{2'b,1'}$ = 6.8, $J_{2'b,3'}$ = 6.1, H-2'b); 2.17 (ddd, 1H, J_{gem} = 13.5, $J_{2'a,1'}$ = 6.2, $J_{2'a,3'}$ = 3.9, H-2'a); 3.56 (dd, 1H, J_{gem} = 11.8, $J_{5'b,4'}$ = 3.9, H-5'b); 3.61 (dd, 1H, J_{gem} = 11.8, $J_{5'a,4'}$ = 3.9, H-5'b); 3.79 (q, 1H, $J_{4',5'} = J_{4',3'}$ = 3.9, H-4'); 4.21 (dt, 1H, $J_{3',2'}$ = 6.1, 3.9, $J_{3',4'}$ = 3.9, H-3'); 6.12 (dd, 1H, $J_{1',2'}$ = 6.8, 6.2, H-1'); 6.88 (bs, 1H, NH_a**H$_b$**-4'''); 7.05 (bs, 1H,

$NH_a\mathbf{H_b}$-4); 7.75 (bs, 1H, $N\mathbf{H_a}H_b$-4'''); 7.89 (bs, 1H, $N\mathbf{H_a}H_b$-4); 7.89 (s, 1H, H-6'''); 8.24 (s, 1H, H-6); 11.01 (bs, 1H, NH-1'''). ^{13}C NMR (125.7 MHz, DMSO-d_6): 40.7 (CH_2-2'); 61.2 (CH_2-5'); 70.2 (CH-3'); 85.6 (CH-1'); 85.8 (C5-**C**≡C-C5'''); 86.2 (C5-C≡**C**-C5'''); 87.6 (CH-4'); 88.6 (C-5'''); 90.2 (C-5); 144.9 (CH-6); 147.3 (CH-6'''); 153.6 (C-2); 155.3 (C-2'''); 163.8 (C-4); 164.7 (C-4'''). MS (ESI^+): *m/z (%)*: 383.3 (100) $[M + Na]^+$. HRMS (ESI^+): *m/z* [M + Na] $^+$ calculated for C_{15} H_{16} O_5 N_6 Na: 383.1074; found 383.1073. IR: υ^{-cm} = 3323, 3064, 1643, 1587, 1493, 1315, 1225, 1196, 1090, 1049, 980 cm^{-1}.

5-(4-Amino-1*H*-pyrimidin-2-on-5-yl)ethynyl-1-(2-deoxy-β-D-ribofuranosyl)pyrimidin-2,4-dione (dU^C)

a) A water/acetonitrile mixture 2:1 (2 ml) was added through a septum to an argon-purged vial containing halogenated **dN^I** (0.22 mmol), 5-ethynylcytosine (**1**) (61 mg, 0.45 mmol), CuI (4 mg, 0.02 mmol) and Et_3N (0.3 ml, 2.25 mmol). After the solids had dissolved, a solution of $Pd(OAc)_2$ (2 mg, 0.01 mmol) and TPPTS (32 mg, 0.05 mmol) in water/acetonitrile mixture 2:1 (1 ml) was injected into the reaction mixture and the resulting mixture was stirred with heating 50 °C for 1 h. The products were isolated from crude reaction mixture by column chromatography on silica gel, eluent gradient from 10 % MeOH in $CHCl_3$ to MeOH and reverse phase chromatography, eluent linear gradient from H_2O to MeOH to gave yellow solid in 40 % yield.

b) DMF (2.5 ml) was added through a septum to an argon-purged vial containing halogenated dN^I (0.28 mmol), 5-ethynylcytosine (**1**) (38.8 mg, 0.28 mmol), $Pd(PPh_3)_2Cl_2$ (10 mg, 0.01 mmol), CuI (5 mg, 0.02 mmol) and Et_3N (0.7 ml, 2.8 mmol) and reaction mixture was stirred with heating 50 °C for 1 h. The products were isolated from crude reaction mixture by precipitation with MeOH (3 ml). The solid was filtered off and washed with water (3 x 30 ml), MeOH (2 x 30 ml) and EtOH (2 x 30 ml). The resulting product was obtained as yellow solid in 60 % yield, slightly contaminated by starting 5-ethynylcytosine. ^{1}H NMR (600.1 MHz, DMSO-d_6): 1.97 (ddd, 1H, J_{gem} = 13.1, $J_{2'b,1'}$ = 7.9, $J_{2'b,3'}$ = 5.8, H-2'b); 2.02 (ddd, 1H, J_{gem} = 13.1, $J_{2'a,1'}$ = 6.0, $J_{2'a,3'}$ = 3.0, H-2'a); 3.52, 3.56 (2 × dd, 2 × 1H, J_{gem} = 11.7, $J_{5',4'}$ = 3.8, H-5'); 3.72 (td, 1H, $J_{4',5'}$ = 3.8, $J_{4',3'}$ = 3.0, H-4'); 4.19 (dt, 1H, $J_{3',2'}$ = 5.8, 3.0, $J_{3',4'}$ = 3.0, H-3'); 5.03, 5.18 (2 × bs, 2 × 1H, OH-3',5'); 6.21 (dd, 1H, $J_{1',2'}$ = 7.9, 6.0, H-1'); 7.57 (s, 1H, H-6'''); 7.60 (bs, 2H, NH_2-4'''); 7.78 (s, 1H, H-6); 11.00 (bs, 2H, NH-3,1'''). ^{13}C NMR (150.9 MHz, DMSO-d_6): 40.4 (CH_2-2'); 61.7 (CH_2-5'); 70.8 (CH-3'); 84.7 (CH-1'); 85.6 (C5-C≡**C**-C5'''); 87.1 (CH-4'); 89.6 (C-5'''); 93.8 (C5-**C**≡C-C5'''); 98.2 (C-5); 138.9 (CH-6); 143.2 (CH-6''');

155.7 (C-2); 156.0 (C-2'''); 165.7 (C-4'''); 170.8 (C-4). MS (ESI^+): *m/z (%)*: 384.1 (100) $[M + Na]^+$, 406.1 (55) $[M + 2Na]^+$. HRMS (ESI^+): *m/z* $[M + Na]^+$ calculated for $C_{15}H_{15}O_6N_5Na$ = 384.09145; found 384.09137. IR: υ^{-cm} = 3435, 1729, 1672, 1456, 1035 cm^{-1}. UV/VIS (MeOH) λ_{max} (ε) = 211 (12628), λ (ε) = 322 (3104).

6-(4-Amino-1*H*-pyrimidin-2-on-5-yl)-1-(2-deoxy-β-D-ribofuranosyl)-7a*H*-furo[2,3-*d*]pyrimidin-2-one (dFP^C)

Preparation according to the general procedure A gave yellowish solid in 30 % yield. For the preparation by general procedure B the yield was 78 %. ^1H NMR (600.1 MHz, DMSO-d_6): 2.05 (ddd, 1H, J_{gem} = 13.4, $J_{2'b,1'} = J_{2'b,3'}$ = 6.1, H-2'b); 2.40 (ddd, 1H, J_{gem} = 13.4, $J_{2'a,1'}$ = 6.1, $J_{2'a,3'}$ = 4.2, H-2'a); 3.62 (dd, 1H, J_{gem} = 12.1, $J_{5'b,OH}$ = 5.2, $J_{5'b,4'}$ = 3.9, H-5'b); 3.69 (dd, 1H, J_{gem} = 12.1, $J_{5'a,OH}$ = 5.2, $J_{5'a,4'}$ = 3.5, H-5'a); 3.92 (dt, 1H, $J_{4',5'}$ = 3.9, 3.5, $J_{4',3'}$ = 3.5, H-4'); 4.24 (dddd, 1H, $J_{3',2'}$ = 6.1, 4.2, $J_{3',OH}$ = 4.3, $J_{3',4'}$ = 3.5, H-3'); 5.20 (t, 1H, $J_{OH,5'}$ = 5.2, OH-5'); 5.33 (d, 1H, $J_{OH,3'}$ = 4.3, OH-3'); 6.19 (t, 1H, $J_{1',2'}$ = 6.1, H-1'); 6.77 (bs, 1H, $NH_a\mathbf{H_b}$-4'''); 6.85 (s, 1H, CH-furane); 7.57 (bs, 1H, $N\mathbf{H_a}H_b$-4'''); 7.80 (bs, 1H, H-6'''); 8.79 (s, 1H, H-6); 11.09 (bs, 1H, NH-1'''). ^{13}C NMR (125.7 MHz, DMSO-d_6): 41.6 (CH_2-2'); 60.9 (CH_2-5'); 69.8 (CH-3'); 87.7 (CH-1'); 88.4 (CH-4'); 94.4 (C-5'''); 101.4 (CH-furane); 107.1 (C-5); 137.5 (CH-6); 144.3 (CH-6'''); 149.3 (C-furane); 154.0 (C-2); 155.6 (C-2'''); 163.1 (C-4'''); 171.1 (C-4). MS (ESI^+): *m/z (%)*: 384.2 (100) $[M + Na]^+$. HRMS (ESI^+): *m/z* $[M + Na]^+$ calculated for $C_{15}H_{15}O_6N_5Na$ = 384.09145; found 384.09157. IR: υ^{-cm} = 3064, 1618, 1576, 1431, 1387, 1185, 1063 cm^{-1}. UV/VIS (MeOH) λ_{max} (ε) = 214 (14004), λ (ε) = 322 (6504).

4-Amino-5-fluoro-1-(prop-2-yn-1-yl)pyrimidin-2-one (3)

DMF (10 ml) was added through a septum to an argon-purged flask containing 4-*N*-benzoyl-5-fluorocytosine (400 mg, 1.7 mmol) and K_2CO_3 (284 mg, 2 mmol). Then a propargyl bromide (170 µl, 2.4 mmol) was added to the solution. The mixture was stirred at ambient temperature for 3 h. DMF was co-evaporated with toluene under reduced pressure. The crude reaction mixture was redissolved in $CHCl_3$ and filtrated through silica gel using $CHCl_3$/MeOH 80:1 mixture. Dried product was directly used in the following reaction with methanolic ammonia (15 ml) at ambient temperature. After 16 h NH_3 and MeOH were evaporated under reduced pressure and the product was isolated by column chromatography on silica gel, eluent gradient from $CHCl_3$ to $CHCl_3$/MeOH 20:1 to give the product **3** as white

solid in 43 % yield (120 mg, 0.7 mmol). ^{1}H NMR (499.8 MHz, DMSO-d_6): 3.35 (t, 1H, ^{4}J = 2.5, HC≡C); 4.43 (d, 2H, ^{4}J = 2.5, $J_{H,F}$ = 0.5, CH_2N); 7.50, 7.74 (2 × bs, 2 × 1H, NH_2); 7.98 (d, 1H, $J_{H,F}$ = 6.7, H-6). ^{13}C NMR (125.7 MHz, DMSO-d_6): 37.79 (CH_2N); 75.68 (H**C**≡C); 79.29 (**C**≡CH); 129.42 (d, $J_{C,F}$ = 31.2, CH-6); 135.96 (d, $J_{C,F}$ = 241.4, C-5); 153.55 (C-2); 157.87 (d, $J_{C,F}$ = 13.1, C-4). ^{19}F{^{1}H} NMR (470.3 MHz, DMSO-d_6): -164.58. MS (ESI$^+$): *m/z (%)*: 168.2 (48) $[M + H]^+$, 190.2 (100) $[M + Na]^+$. HRMS (ESI$^+$): *m/z* $[M + H]^+$ calculated for $C_7H_7ON_3F$ = 168.05669; found 168.5677. IR: υ^{-cm} = 2935, 1688, 1651, 1478, 1321, 1243, 1148, 1094, 897, 777, 678 cm^{-1}.

4-Amino-5-[3-(4-aminopyrimidin-2-on-1-yl)prop-1-yn-1-yl]-7-(2-deoxy-β-D-ribofuranosyl)-7*H*-pyrrolo[2,3-*d*]pyrimidine (dA^{PC})

Nucleoside **dA^{PC}** was prepared according to the general procedure C. Nucleoside **dA^{I}** (50 mg, 0.13 mmol) and **2** (39 mg, 0.27 mmol) were used. The crude reaction mixture was separated by semi-preparative HPLC on a C18 column with use of gradient H_2O to 40% MeOH in 2 h as eluent. Nucleoside **dA^{PC}** (34 mg, 61 %) was obtained as a white solid. ^{1}H NMR (600.1 MHz, DMSO-d_6): 2.17 (ddd, 1H, J_{gem} = 13.1, $J_{2'b,1'}$ = 6.0, $J_{2'b,3'}$ = 2.8, H-2'b); 2.44 (ddd, 1H, J_{gem} = 13.1, $J_{2'a,1'}$ = 8.0, $J_{2'a,3'}$ = 5.8, H-2'a); 3.50 (ddd, 1H, J_{gem} = 11.8, $J_{5'b,OH}$ = 5.8, $J_{5'b,4'}$ = 4.2, H-5'b); 3.56 (ddd, 1H, J_{gem} = 11.8, $J_{5'a,OH}$ = 5.2, $J_{5'a,4'}$ = 4.5, H-5'a); 3.81 (ddd, 1H, $J_{4',5'}$ = 4.5, 4.2, $J_{4',3'}$ = 2.6, H-4'); 4.33 (m, 1H, $J_{3',2'}$ = 5.8, 2.8, $J_{3',OH}$ = 4.1, $J_{3',4'}$ = 2.6, H-3'); 4.71 (s, 2H, CH_2N); 5.08 (dd, 1H, $J_{OH,5'}$ = 5.8, 5.2, OH-5'); 5.27 (d, 1H, $J_{OH,3'}$ = 4.1, OH-3'); 5.72 (d, 1H, $J_{5''',6'''}$ = 7.3, H-5'''); 6.46 (dd, 1H, $J_{1',2'}$ = 8.0, 6.0, H-1'); 6.86 (bs, 2H, NH_2-6); 7.05, 7.18 (2 × bs, 2 × 1H, NH_2-4'''); 7.72 (s, 1H, H-8); 7.73 (d, 1H, $J_{6''',5'''}$ = 7.3, H-6'''); 8.09 (s, 1H, H-2). ^{13}C NMR (150.9 MHz, DMSO-d_6): 39.42 (CH_2N); 40.10 (CH_2-2'); 62.05 (CH_2-5'); 71.14 (CH-3'); 77.37 (-**C**≡C-CH_2-); 83.44 (CH-1'); 87.66 (-C≡**C**-CH_2-); 87.72 (CH-4'); 94.20 (CH-5'''); 94.42 (C-7); 102.56 (C-5); 126.46 (CH-8); 145.41 (CH-6'''); 149.47 (C-4); 152.96 (CH-2); 155.81 (C-2'''); 157.66 (C-6); 166.36 (C-4'''). MS (ESI$^+$): *m/z (%)*: 282.2 (40) $[M - C_5H_8O_3 + H]^+$, 398.3 (100) $[M + H]^+$, 420.3 (25) $[M + Na]^+$. HRMS (ESI$^+$): *m/z* $[M + H]^+$ calculated for $C_{18}H_{20}O_4N_7$ = 398.15719; found 398.15713. IR: υ^{-cm} = 3187, 1641, 1598, 1479, 1270, 1192, 1093, 785, 677 cm^{-1}.

4-Amino-5-[3-(4-aminopyrimidin-2-on-1-yl)prop-1-yn-1-yl]-1-(2-deoxy-β-D-ribofuranosyl)pyrimidine-2-one (dC^{PC})

Nucleoside **dC^{PC}** was prepared according to the general procedure C. Nucleoside **dC^{I}** (50 mg, 0.14 mmol) and **2** (42 mg, 0.28 mmol) were used. The crude reaction mixture was separated by semi-preparative HPLC on a C18 column with use of gradient H_2O to 20% MeOH in 100 min as eluent. Nucleoside **dC^{PC}** (50 mg, 95 %) was obtained as a white solid. ^{1}H NMR (600.1 MHz, DMSO-d_6): 1.97 (ddd, 1H, J_{gem} = 13.2, $J_{2'b,1'}$ = 7.1, $J_{2'b,3'}$ = 6.1, H-2'b); 2.13 (ddd, 1H, J_{gem} = 13.2, $J_{2'a,1'}$ = 6.1, $J_{2'a,3'}$ = 3.5, H-2'a); 3.54, 3.60 (2 × ddd, 2 × 1H, J_{gem} = 12.0, $J_{5',OH}$ = 5.1, $J_{5',4'}$ = 3.6, H-5'); 3.78 (q, 1H, $J_{4',3'}$ = $J_{4',5'}$ = 3.6, H-4'); 4.19 (m, 1H, $J_{3',2'}$ = 6.1, 3.5, $J_{3',OH}$ = 4.3, $J_{3',4'}$ = 3.6, H-3'); 4.69 (s, 2H, CH_2N); 5.05 (t, 1H, $J_{OH,5'}$ = 5.1, OH-5'); 5.19 (d, 1H, $J_{OH,3'}$ = 4.3, OH-3'); 5.72 (d, 1H, $J_{5''',6'''}$ = 7.3, H-5'''); 6.10 (dd, 1H, $J_{1',2'}$ = 7.1, 6.1, H-1'); 6.94 (bs, 1H, $NH_a\mathbf{H_b}$-4); 7.06, 7.16 (2 × bs, 2 × 1H, NH_2-4'''); 7.73 (d, 1H, $J_{6''',5'''}$ = 7.3, H-6'''); 7.78 (bs, 1H, $N\mathbf{H_a}H_b$-4); 8.16 (s, 1H, H-6). ^{13}C NMR (150.9 MHz, DMSO-d_6): 38.93 (CH_2N); 40.95 (CH_2-2'); 61.14 (CH_2-5'); 70.25 (CH-3'); 76.66 (cyt-**C**≡C-CH_2-); 85.54 (CH-1'); 87.63 (CH-4'); 89.04 (C-5); 90.95 (cyt-C≡**C**-CH_2-); 94.17 (CH-5'''); 144.57 (CH-6); 145.14 (CH-6'''); 153.56 (C-2); 155.55 (C-2'''); 164.70 (C-4); 166.20 (C-4'''). MS (ESI$^+$): *m/z (%)*: 281.1 (30) [M – $C_5H_8O_3$ + Na]$^+$, 397.2 (100) [M + Na]$^+$. HRMS (ESI$^+$): *m/z* [M + H]$^+$ calculated for $C_{16}H_{19}O_5N_6$ = 375.14107; found 375.14114. IR: υ^{-cm} = 3188, 1651, 1599, 1479, 1321, 1242, 1187, 1094, 898, 781, 677 cm^{-1}.

5-[3-(4-Aminopyrimidin-2-on-1-yl)prop-1-yn-1-yl]-1-(2-deoxy-β-D-ribofuranosyl)pyrimidin-2,4-dione (dU^{PC})

Nucleoside **dU^{PC}** was prepared according to the general procedure C with modification, the reaction temperature was 50 °C. Nucleoside **dU^{I}** (60 mg, 0.17 mmol) and **2** (50 mg, 0.34 mmol) were used. The crude reaction mixture was separated by semi-preparative HPLC on a C18 column with use of gradient H_2O to 20% MeOH in 100 min as eluent. Nucleoside **dU^{PC}** (24 mg, 37 %) was obtained as a white solid. ^{1}H NMR (500.0 MHz, DMSO-d_6): 2.10 (ddd, 1H, J_{gem} = 13.3, $J_{2'b,1'}$ = 6.4, $J_{2'b,3'}$ = 4.0, H-2'b); 2.14 (ddd, 1H, J_{gem} = 13.3, $J_{2'a,1'}$ = 7.0, $J_{2'a,3'}$ = 5.5, H-2'a); 3.55, 3.61 (ddd, 2 × 1H, J_{gem} = 12.0, $J_{5',OH}$ = 5.1, $J_{5',4'}$ = 3.6, H-5'); 3.79 (q, 1H, $J_{4',3'}$ = $J_{4',5'}$ = 3.6, H-4'); 4.22 (m, 1H, $J_{3',2'}$ = 5.5, 4.0, $J_{3',OH}$ = 4.3, $J_{3',4'}$ = 3.6, H-3'); 4.69 (s, 2H, CH_2N); 5.10 (t, 1H, $J_{OH,5'}$ = 5.1, OH-5'); 5.24 (d, 1H, $J_{OH,3'}$ = 4.3, OH-3'); 5.73 (d, 1H, $J_{5''',6'''}$ = 7.3, H-5'''); 6.09 (dd, 1H, $J_{1',2'}$ = 7.0, 6.4, H-1'); 7.08, 7.18 (2 × bs, 2 × 1H, NH_2-4'''); 7.67 (d,

1H, $J_{6''',5'''}$ = 7.3, H-6'''); 8.22 (s, 1H, H-6); 11.64 (bs, 1H, NH-3). ^{13}C NMR (125.7 MHz, DMSO-d_6): 38.22 (CH_2N); 40.26 (CH_2-2'); 61.12 (CH_2-5'); 70.29 (CH-3'); 77.13 (ur-**C**≡C-CH_2-); 84.99 (CH-1'); 87.76 (CH-4'); 87.81 (ur-C≡**C**-CH_2-); 94.33 (CH-5'''); 97.80 (C-5); 144.35 (CH-6); 144.84 (CH-6'''); 149.59 (C-2); 155.30 (C-2'''); 161.74 (C-4); 166.12 (C-4'''). MS (ESI$^+$): *m/z (%)*: 282.0 (100) [M – $C_5H_8O_3$ + Na]$^+$, 398.0 (72) [M + Na]$^+$. HRMS (ESI$^+$): *m/z* [M + H]$^+$ calculated for $C_{16}H_{17}O_6N_5Na$ = 398.10712; found 398.10710. IR: υ^{-cm} = 3187, 1653, 1564, 1512, 1462, 1092, 786 cm^{-1}.

4-Amino-5-[3-(4-amino-5-fluoropyrimidin-2-on-1-yl)prop-1-yn-1-yl]-7-(2-deoxy-β-D-ribofuranosyl)-7*H*-pyrrolo[2,3-*d*]pyrimidine (dA^{PFC})

Nucleoside **dA^{PFC}** was prepared according to the general procedure C. Nucleoside **dA^{I}** (50 mg, 0.13 mmol) and **3** (46 mg, 0.27 mmol) were used. The crude reaction mixture was separated by semi-preparative HPLC on a C18 column with use of gradient H_2O to 40% MeOH in 2 h as eluent. Nucleoside **dA^{PFC}** (33 mg, 58 %) was obtained as a white solid. ^{1}H NMR (600.1 MHz, DMSO-d_6): 2.17 (ddd, 1H, J_{gem} = 13.1, $J_{2'b,1'}$ = 6.0, $J_{2'b,3'}$ = 2.8, H-2'b); 2.45 (ddd, 1H, J_{gem} = 13.1, $J_{2'a,1'}$ = 8.0, $J_{2'a,3'}$ = 5.8, H-2'a); 3.50 (ddd, 1H, J_{gem} = 11.8, $J_{5'b,OH}$ = 5.9, $J_{5'b,4'}$ = 4.3, H-5'b); 3.56 (ddd, 1H, J_{gem} = 11.8, $J_{5'a,OH}$ = 5.2, $J_{5'a,4'}$ = 4.3, H-5'a); 3.81 (td, 1H, $J_{4',5'}$ = 4.3, $J_{4',3'}$ = 2.6, H-4'); 4.33 (m, 1H, $J_{3',2'}$ = 5.8, 2.8, $J_{3',OH}$ = 4.1, $J_{3',4'}$ = 2.6, H-3'); 4.68 (s, 2H, CH_2N); 5.06 (dd, 1H, $J_{OH,5'}$ = 5.9, 5.2, OH-5'); 5.25 (d, 1H, $J_{OH,3'}$ = 4.1, OH-3'); 6.46 (dd, 1H, $J_{1',2'}$ = 8.0, 6.0, H-1'); 6.85 (bs, 2H, NH_2-6); 7.51, 7.72 (2 × bs, 2 × 1H, NH_2-4'''); 7.74 (s, 1H, H-8); 8.09 (s, 1H, H-2); 8.10 (d, 1H, $J_{H,F}$ = 6.6, H-6'''). ^{13}C NMR (150.9 MHz, DMSO-d_6): 39.70 (CH_2N); 40.09 (CH_2-2'); 62.01 (CH_2-5'); 71.09 (CH-3'); 77.56 (-**C**≡C-CH_2-); 83.39 (CH-1'); 87.23 (-C≡**C**-CH_2-); 87.69 (CH-4'); 94.27 (C-7); 102.48 (C-5); 126.58 (CH-8); 129.88 (d, $J_{C,F}$ = 31.4, CH-6'''); 135.94 (d, $J_{C,F}$ = 241.4, C-5'''); 149.45 (C-4); 152.92 (CH-2); 153.99 (C-2'''); 157.62 (C-6); 157.93 (d, $J_{C,F}$ = 13.0, C-4'''). ^{19}F{^{1}H} NMR (470.3 MHz, DMSO-d_6): -164.11. MS (ESI$^+$): *m/z (%)*: 300.1 (56) [M – $C_5H_8O_3$ + H]$^+$, 416.1 (100) [M + H]$^+$, 438.1 (30) [M + Na]$^+$. HRMS (ESI$^+$): *m/z* [M + H]$^+$ calculated for $C_{18}H_{19}O_4N_7F$ = 416.14768; found 416.14771. IR: υ^{-cm} = 3187, 1722, 1659, 1622, 1564, 1525, 1513, 1462, 1284, 1057, 954, 779, 446 cm^{-1}.

4-Amino-5-[3-(4-amino-5-fluoropyrimidin-2-on-1-yl)prop-1-yn-1-yl]-1-(2-deoxy-β-D-ribofuranosyl)pyrimidin-2-one (dC^{PFC})

Nucleoside **dC^{PFC}** was prepared according to the general procedure C. Nucleoside **dC^{I}** (40 mg, 0.11 mmol) and **3** (37 mg, 0.23 mmol) were used. The crude reaction mixture was separated by semi-preparative HPLC on a C18 column with use of gradient H_2O to 20 % MeOH in 100 min as eluent. Nucleoside **dC^{PFC}** (40 mg, 95 %) was obtained as a white solid. ^{1}H NMR (500.0 MHz, DMSO-d_6): 1.98 (ddd, 1H, J_{gem} = 13.2, $J_{2'b,1'}$ = 7.1, $J_{2'b,3'}$ = 6.2, H-2'b); 2.14 (ddd, 1H, J_{gem} = 13.2, $J_{2'a,1'}$ = 6.0, $J_{2'a,3'}$ = 3.5, H-2'a); 3.54, 3.61 (2 × ddd, 2 × 1H, J_{gem} = 11.9, $J_{5',OH}$ = 5.1, $J_{5',4'}$ = 3.7, H-5'); 3.78 (q, 1H, $J_{4',3'}$ = $J_{4',5'}$ = 3.7, H-4'); 4.19 (m, 1H, $J_{3',2'}$ = 6.2, 3.5, $J_{3',OH}$ = 4.2, $J_{3',4'}$ = 3.7, H-3'); 4.67 (s, 2H, CH_2N); 5.05 (t, 1H, $J_{OH,5'}$ = 5.1, OH-5'); 5.19 (d, 1H, $J_{OH,3'}$ = 4.2, OH-3'); 6.10 (dd, 1H, $J_{1',2'}$ = 7.1, 6.0, H-1'); 6.95 (bs, 1H, $NH_a\mathbf{H_b}$-4); 7.51, 7.72 (2 × bs, 2 × 1H, NH_2-4'''); 7.79 (bs, 1H, $N\mathbf{H_a}H_b$-4); 8.07 (d, 1H, $J_{H,F}$ = 6.6, H-6'''); 8.17 (s, 1H, H-6). ^{13}C NMR (125.7 MHz, DMSO-d_6): 38.99 (CH_2N); 40.96 (CH_2-2'); 61.16 (CH_2-5'); 70.28 (CH-3'); 76.28 (cyt-**C**≡C-CH_2-); 85.58 (CH-1'); 87.65 (CH-4'); 88.98 (C-5); 90.50 (cyt-C≡**C**-CH_2-); 129.71 (d, $J_{C,F}$ = 31.2, CH-6'''); 135.98 (d, $J_{C,F}$ = 242.2, C-5'''); 144.79 (CH-6); 153.56 (C-2); 153.81 (C-2'''); 157.85 (d, $J_{C,F}$ = 13.1, C-4'''); 164.67 (C-4). ^{19}F {1H} NMR (470.3 MHz, DMSO-d_6): -163.88. MS (ESI^+): *m/z (%)*: 277.1 (28) $[M - C_5H_8O_3 + H]^+$, 415.2 (100) $[M + Na]^+$. HRMS (ESI^+): *m/z* $[M + H]^+$ calculated for $C_{16}H_{18}O_5N_6F$ = 393.13172; found 393.13172. IR: υ^{-cm} = 3073, 1689, 1652, 1582, 1479, 1322, 1242, 1185, 1094, 976, 896, 778, 678 cm^{-1}.

5-[3-(4-Amino-5-fluoropyrimidin-2-on-1-yl)prop-1-yn-1-yl]-1-(2-deoxy-β-D-ribofuranosyl)pyrimidin-2,4-dione (dU^{PFC})

Nucleoside **dU^{PFC}** was prepared according to the general procedure C with modification, the reaction temperature was 50 °C. Nucleoside **dU^{I}** (60 mg, 0.17 mmol) and **3** (57 mg, 0.34 mmol) were used. The crude reaction mixture was separated by semi-preparative HPLC on a C18 column with use of gradient H_2O to 20 % MeOH in 100 min as eluent. Nucleoside **dU^{PFC}** (46 mg, 69 %) was obtained as a white solid. ^{1}H NMR (600.1 MHz, DMSO-d_6): 2.10 (ddd, 1H, J_{gem} = 13.3, $J_{2'b,1'}$ = 6.4, $J_{2'b,3'}$ = 3.9, H-2'b); 2.14 (ddd, 1H, J_{gem} = 13.3, $J_{2'a,1'}$ = 7.0, $J_{2'a,3'}$ = 5.6, H-2'a); 3.55, 3.61 (ddd, 2 × 1H, J_{gem} = 11.9, $J_{5',OH}$ = 5.1, $J_{5',4'}$ = 3.6, H-5'); 3.79 (q, 1H, $J_{4',3'}$ = $J_{4',5'}$ = 3.6, H-4'); 4.22 (m, 1H, $J_{3',2'}$ = 5.6, 3.9, $J_{3',OH}$ = 4.3, $J_{3',4'}$ = 3.6, H-3'); 4.66 (s, 2H, CH_2N); 5.10 (t, 1H, $J_{OH,5'}$ = 5.1, OH-5'); 5.25 (d, 1H, $J_{OH,3'}$ = 4.3, OH-3'); 6.09 (dd, 1H, $J_{1',2'}$ =

7.0, 6.4, H-1'); 7.51, 7.74 (2 × bs, 2 × 1H, NH_2-4'''); 8.01 (d, 1H, $J_{H,F}$ = 6.5, H-6'''); 8.22 (s, 1H, H-6); 11.65 (bs, 1H, NH-3). ^{13}C NMR (150.9 MHz, DMSO-d_6): 38.58 (CH_2N); 40.26 (CH_2-2'); 61.13 (CH_2-5'); 70.30 (CH-3'); 77.49 (ur-**C**≡C-CH_2-); 85.03 (CH-1'); 87.46 (ur-C≡**C**-CH_2-); 87.83 (CH-4'); 97.73 (C-5); 129.27 (d, $J_{C,F}$ = 31.2, CH-6'''); 136.13 (d, $J_{C,F}$ = 245.7, C-5'''); 144.37 (CH-6); 149.59 (C-2); 153.66 (C-2'''); 157.88 (d, $J_{C,F}$ = 13.6, C-4'''); 161.78 (C-4). $^{19}F\{^1H\}$ NMR (470.3 MHz, DMSO-d_6): -163.67. MS (ESI^+): *m/z (%)*: 278.1 (100) $[M - C_5H_8O_3 + H]^+$, 394.2 (67) $[M + H]^+$, 416.2 (30) $[M + Na]^+$. HRMS (ESI^+): *m/z* $[M + H]^+$ calculated for $C_{16}H_{17}O_6N_5F$ = 394.11571; found 394.11574. IR: υ^{-cm} = 3066, 1702, 1623, 1526, 1463, 1332, 1286, 1129, 1058, 956, 772, 575 cm^{-1}.

Synthesis of modified dNTPs Sonogashira-coupling. General Procedure D:

A CH_3CN/H_2O 1:2 (1 ml) was added through a septum to an argon-purged vial containing halogenated **dN^{I}TP** (0.05 mmol), 5-ethynylcytosine (**1**) (0.11 mmol, 15 mg), CuI (1 mg, 0.005mmol) and *i*Pr_2NEt (98 µl, 0.56 mmol). After the solids had dissolved, a solution of $Pd(OAc)_2$ (1 mg, 0.002 mmol) and TPPTS (8 mg, 0.01 mmol) in CH_3CN/H_2O 1:2 (1 ml) was injected into the reaction mixture and the resulting mixture was stirred with heating 80 °C for 1 h. The products were isolated from crude reaction mixture by semi-preparative HPLC on a C18 column with use of linear gradient of H_2O to 50% MeOH with 0.1M TEAB in 65 min. Several co-distillations with water and exchange of the counterion at Dowex50WX8 in Na^+ cycle followed by freeze drying from water, gave the products **dC^{C}TP**, **dU^{C}TP** and **dA^{C}TP**.

Synthesis of modified dNTPs Sonogashira-coupling. General Procedure E:

A CH_3CN/H_2O 1:2 (1 ml) was added through a septum to an argon-purged vial containing halogenated **dN^{I}TP** (0.06 mmol), **2** (0.18 mmol, 25 mg) or **3** (0.18 mmol, 30 mg), CuI (1 mg, 0.005 mmol) and *i*Pr_2NEt (98 µl, 0.56 mmol). After the solids had dissolved, a solution of $Pd(OAc)_2$ (1 mg, 0.006 mmol) and TPPTS (13 mg, 0.02 mmol) in CH_3CN/H_2O 1:2 (1 ml) was injected into the reaction mixture and the resulting mixture was stirred with heating 80 °C for 1 h. The products were isolated from crude reaction mixture by semi-preparative HPLC on a C18 column with use of gradient 5 % MeOH in H_2O to 12.5 %MeOH with 0.1M TEAB in 65 min. Several co-distillations with water and exchange of the counterion at Dowex50WX8 in Na^+ cycle followed by freeze drying from water, gave the products **dC^{PC}TP**, **dA^{PC}TP**, **dC^{PFC}TP**, **dA^{PFC}TP**.

5-(4-Amino-1*H*-pyrimidin-2-on-5-yl)ethynyl-1-(2-deoxy-β-D-ribofuranosyl)pyrimidin-2-on 5'-*O*-triphosphate (dC^{C}TP)

This compound was prepared according to the general procedure D from 5-iodo-2'-deoxycytidine triphosphate **dC^{I}TP** as yellow solid (35 %). ^{1}H NMR (499.8 MHz, D_2O, pD = 7.1, phosphate buffer, ref(dioxane) = 3.75 ppm): 2.32 (ddd, 1H, J_{gem} = 13.9, $J_{2'b,1'}$ = 7.0, $J_{2'b,3'}$ = 6.1, H-2'b); 2.47 (ddd, 1H, J_{gem} = 13.9, $J_{2'a,1'}$ = 6.0, $J_{2'a,3'}$ = 4.1, H-2'a); 4.24 (m, 3H, H-4',5'); 4.63 (dt, 1H, $J_{3',2'}$ = 6.1, 4.1, $J_{3',4'}$ = 4.1, H-3'); 6.25 (dd, 1H, $J_{1',2'}$ = 7.0, 6.0, H-1'); 7.96 (s, 1H, H-6'''); 8.33 (s, 1H, H-6). ^{13}C NMR (125.7 MHz, D_2O, pD = 7.1, phosphate buffer, ref(dioxane) = 69.3 ppm): 42.4 (CH_2-2'); 67.8 (d, $J_{C,P}$ = 6.3, CH_2-5'); 73.0 (CH-3'); 88.1 (C5-**C**≡C-C5'''); 88.5 (d, $J_{C,P}$ = 9.7, CH-4'); 88.6 (C5-C≡**C**-C5'''); 89.3 (CH-1'); 93.8 (C-5'''); 94.9 (C-5); 147.5 (CH-6); 150.1 (CH-6'''); 158.7 (C-2); 160.5 (C-2'''); 167.2 (C-4); 168.1 (C-4'''). ^{31}P{^{1}H} NMR (125.7 MHz, D_2O, pD = 7.1, phosphate buffer, ref(phosphate buffer) = 2.35 ppm): -21.32 (bdd, *J* = 19.3, 17.3, P_β); -10.83 (d, *J* = 19.3, P_α); -6.38 (d, *J* = 17.3, P_γ). MS (ESI$^-$): *m/z (%)*: 439.1 (62) [M – $P_2O_6H_3$]$^-$, 519.0 (100) [M - PO_3H_2]$^-$, 621.0 (18) [M – H + Na]$^-$,. HRMS (ESI$^-$): *m/z* [M - 2H + Na]$^-$ calculated for C_{15} H_{17} O_{14} N_6 Na P_3: 620.9919; found 620.9911.

5-(4-Amino-1*H*-pyrimidin-2-on-5-yl)ethynyl-1-(2-deoxy-β-D-ribofuranosyl)pyrimidin-2,4-dion 5'-*O*-triphosphate (dU^{C}TP)

This compound was prepared according to the general procedure D from 5-iodo-2'-deoxyuridine triphosphate **dU^{I}TP** as yellow solid (31 %). ^{1}H NMR (499.8 MHz, D_2O, pD = 7.1, phosphate buffer, ref(dioxane) = 3.75 ppm): 2.38 (dt, 1H, J_{gem} = 14.2, $J_{2'b,1'}$ = $J_{2'b,3'}$ = 6.8, H-2'b); 2.44 (ddd, 1H, J_{gem} = 14.2, $J_{2'a,1'}$ = 6.2, $J_{2'a,3'}$ = 3.9, H-2'a); 4.23 (m, 3H, H-4',5'); 4.67 (bm, 1H, H-3'); 6.30 (dd, 1H, $J_{1',2'}$ = 6.8, 6.2, H-1'); 7.99 (s, 1H, H-6'''); 8.32 (s, 1H, H-6). ^{13}C NMR (125.7 MHz, D_2O, pD = 7.1, phosphate buffer, ref(dioxane) = 69.3 ppm): 41.9 (CH_2-2'); 68.0 (d, $J_{C,P}$ = 4.8, CH_2-5'); 73.2 (CH-3'); 87.3 (C5-**C**≡**C**-C5'''); 88.6 (d, $J_{C,P}$ = 8.5, CH-4'); 88.6 (CH-1'); 89.0 (C-5'''); 102.1 (C-5); 146.8 (CH-6); 150.0 (CH-6'''); 153.6 (C-2); 160.5 (C-2'''); 167.6 (C-4); 168.2 (C-4'''). ^{31}P{^{1}H} NMR (125.7 MHz, D_2O, pD = 7.1, phosphate buffer, ref(phosphate buffer) = 2.35 ppm): -21.30 (br, P_β); -10.98 (br, P_α); -6.19 (br, P_γ). MS (ESI$^-$): *m/z (%)*: 440.1 (65) [M – $P_2O_6H_3$]$^-$, 520.0 (100) [M - PO_3H_2]$^-$, 542.0 (47) [M - PO_3H_2 + Na]$^-$,

622.0 (32) [M – H + Na]$^-$,. HRMS (ESI$^-$): *m/z* [M - 2H + Na]$^-$ calculated for C_{15} $H_{16}O_{15}$ N_5 Na P_3: 621.9759; found 621.9742.

4-Amino-5-(4-amino-1*H*-pyrimidin-2-on-5-yl)ethynyl-7-(2-deoxy-β-D-ribofuranosyl)-7*H*-pyrrolo[2,3-*d*]pyrimidin 5'-*O*-triphosphate (dA^{C}TP)

This compound was prepared according to the general procedure D from 7-iodo-7-deaza-2'-deoxyadenosine triphosphate **dA^{I}TP**. (17 %). ^{1}H NMR (499.8 MHz, D_2O, pD = 7.1, phosphate buffer, ref(dioxane) = 3.75 ppm): 2.52 (ddd, 1H, J_{gem} = 14.0, $J_{2'b,1'}$ = 6.3, $J_{2'b,3'}$ = 3.5, H-2'b); 2.67 (ddd, 1H, J_{gem} = 14.0, $J_{2'a,1'}$ = 7.4, $J_{2'a,3'}$ = 6.3, H-2'a); 4.15, 4.22 (2 × m, 2 × 1H, H-5'); 4.24 (m, 1H, H-4'); 4.79 (m, 1H, H-3'); 6.55 (dd, 1H, $J_{1',2'}$ = 7.4, 6.3, H-1'); 7.77 (s, 1H, H-6'''); 7.83 (s, 1H, H-8); 8.06 (s, 1H, H-2). ^{13}C NMR (125.7 MHz, D_2O, pD = 7.1, phosphate buffer, ref(dioxane) = 69.3 ppm): 41.7 (CH_2-2'); 68.2 (d, $J_{C,P}$ = 5.4, CH_2-5'); 73.8 (CH-3'); 84.9 (C5'''-**C**≡C-C7); 85.9 (CH-1'); 88.1 (d, $J_{C,P}$ = 8.7, CH-4'); 90.7 (C7-**C**≡C-C5'''); 94.1 (C-5'''); 99.0 (C-7); 105.2 (C-5); 129.4 (CH-8); 148.6 (CH-6'''); 151.2 (C-4); 154.6 (CH-2); 159.7 (C-6); 160.2 (C-2'''); 168.1 (C-4'''). ^{31}P{^{1}H} NMR (202.3 MHz, D_2O, pD = 7.1, phosphate buffer, ref(phosphate buffer) = 2.35 ppm): -21.25 (t, *J* = 19.5, P_β); -10.23 (d, *J* = 19.5, P_α); -6.59 (d, *J* = 19.5, P_γ). MS (ESI$^+$): *m/z (%)*: 668.0 (26) [M – 2H + 2Na]$^+$, 689.9 (51) [M – 3H + 3Na]$^+$, 711.9 (100) [M - 3H + 4Na]$^+$, 733.9 (49) [M - 3H + 5Na]$^+$, 755.9 (20) [M - 3H + 6Na]$^+$. HRMS (ESI$^+$): *m/z* [M - 3H + 4Na]$^+$ calculated for $C_{17}H_{17}O_{13}N_7Na_4P_3$: 711.96825; found 711.96846.

4-Amino-5-[3-(4-aminopyrimidin-2-on-1-yl)prop-1-yn-1-yl]-7-(2-deoxy-β-D-ribofuranosyl)-7*H*-pyrrolo[2,3-*d*]pyrimidin 5'-*O*-triphosphate (dA^{PC}TP)

This compound was prepared according to the general procedure E from 5-iodo-2'-deoxyadenosine triphosphate **dA^{I}TP** as yellow solid (80 %).^{1}H NMR (499.8 MHz, D_2O, pD = 7.1, phosphate buffer, ref(dioxane) = 3.75 ppm): 2.48 (ddd, 1H, J_{gem} = 13.9, $J_{2'b,1'}$ = 6.2, $J_{2'b,3'}$ = 3.1, H-2'b); 2.64 (ddd, 1H, J_{gem} = 13.9, $J_{2'a,1'}$ = 7.8, $J_{2'a,3'}$ = 6.3, H-2'a); 4.09 (ddd, 1H, J_{gem} = 11.4, $J_{H,P}$ = 5.5, $J_{5'b,4'}$ = 4.1, H-5'b); 4.17 (ddd, 1H, J_{gem} = 11.4, $J_{H,P}$ = 6.4, $J_{5'a,4'}$ = 4.1, H-5'a); 4.22 (qd, 1H, $J_{4',3'}$ = $J_{4',5'}$ = 4.1, $J_{H,P}$ = 1.4, H-4'); 4.73 (ddd, 1H, $J_{3',2'}$ = 6.3, 3.1, $J_{3',4'}$ = 4.1, H-3'); 4.79 (s, 2H, CH_2N, overlapped with HDO); 6.06 (d, 1H, $J_{5''',6'''}$ = 7.4, H-5'''); 6.53 (dd, 1H, $J_{1',2'}$ = 7.8, 6.2, H-1'); 7.63 (s, 1H, H-8); 7.80 (d, 1H, $J_{6''',5'''}$ = 7.4, H-6'''); 8.06 (s, 1H, H-2). ^{13}C NMR (150.9 MHz, D_2O, pD = 7.1, phosphate buffer, ref(dioxane) = 69.3 ppm): 41.22 (CH_2-

2'); 43.25 (CH_2N); 68.19 (d, $J_{C,P}$ = 5.7, CH_2-5'); 73.80 (CH-3'); 80.53 (-**C**≡C-CH_2-); 85.67 (CH-1'); 87.96 (d, $J_{C,P}$ = 8.8, CH-4'); 88.84 (-C≡**C**-CH_2-); 98.38 (C-7); 98.83 (CH-5'''); 105.67 (C-5); 129.17 (CH-8); 148.77 (CH-6'''); 151.31 (C-4); 154.88 (CH-2); 159.84 (C-6); 160.60 (C-2'''); 169.31 (C-4'''). $^{31}P\{^1H\}$ NMR (202.3 MHz, D_2O, pD = 7.1, phosphate buffer, ref(phosphate buffer) = 2.35 ppm): -21.39 (t, J = 19.6, P_β); -10.10 (d, J = 19.6, P_α); -7.19 (bd, J = 19.6, P_γ). MS (ESI^+): *m/z (%)*: 476.2 (50) [M – $H_3O_6P_2$ - H]$^-$, 556.2 (100) [M – H_2O_3P - H]$^-$, 578.2 (50) [M - H_2O_3P - 2H + Na]$^-$. HRMS (ESI^+): *m/z* [M - H] $^-$ calculated for $C_{18}H_{21}O_{13}N_7P_3$: 636.04081; found 636.04157.

4-Amino-5-[3-(4-aminopyrimidin-2-on-1-yl)prop-1-yn-1-yl]-1-(2-deoxy-β-D-ribofuranosyl)pyrimidin-2-on 5'-*O*-triphosphate (dC^{PC}TP)

This compound was prepared according to the general procedure E from 5-iodo-2'-deoxycytidine triphosphate **dC^{I}TP** as white solid (30 %). ^1H NMR (600.1 MHz, D_2O, pD = 7.1, phosphate buffer, ref(dioxane) = 3.75 ppm): 2.30 (ddd, 1H, J_{gem} = 14.1, $J_{2'b,1'}$ = 7.1, $J_{2'b,3'}$ = 6.0, H-2'b); 2.44 (ddd, 1H, J_{gem} = 14.1, $J_{2'a,1'}$ = 6.5, $J_{2'a,3'}$ = 4.4, H-2'a); 4.18-4.26 (bm, 3H, H-4',5'); 4.60 (bdt, 1H, $J_{3',2'}$ = 6.0, 4.4, $J_{3',4'}$ = 4.4, H-3'); 4.82 (s, 2H, CH_2N, overlapped with HDO); 6.07 (d, 1H, $J_{5''',6'''}$ = 7.3, H-5'''); 6.23 (dd, 1H, $J_{1',2'}$ = 7.1, 6.5, H-1'); 7.83 (d, 1H, $J_{6''',5'''}$ = 7.3, H-6'''); 8.20 (s, 1H, H-6). ^{13}C NMR (150.9 MHz, D_2O, pD = 7.1, phosphate buffer, ref(dioxane) = 69.3 ppm): 42.03 (CH_2-2'); 42.73 (CH_2N); 67.77 (d, $J_{C,P}$ = 5.3, CH_2-5'); 72.77 (CH-3'); 78.67 (cyt-**C**≡C-CH_2-); 88.27 (d, $J_{C,P}$ = 8.8, CH-4'); 89.04 (CH-1'); 92.40 (cyt-C≡**C**-CH_2-); 94.37 (C-5); 99.05 (CH-5'''); 148.18 (CH-6); 148.90 (CH-6'''); 158.75 (C-2); 160.68 (C-2'''); 167.79 (C-4); 169.34 (C-4'''). $^{31}P\{^1H\}$ NMR (202.3 MHz, D_2O, pD = 7.1, phosphate buffer, ref(phosphate buffer) = 2.35 ppm): -21.43 (dd, J = 19.3, 18.4, P_β); -10.64 (d, J = 19.3, P_α); -6.45 (d, J = 18.4, P_γ). MS (ESI^+): *m/z (%)*: 259.1 (35) [M – $C_5H_{13}O_{12}P_3$ + H]$^+$, 281.1 (60) [M – $C_5H_{13}O_{12}P_3$ + Na]$^+$, 397.2 (100) [M - $H_4O_9P_3$ + Na]$^+$. HRMS (ESI^+): *m/z* [M + H]$^+$ calculated for $C_{16}H_{22}N_6O_{14}P_3$: 614,03286; found 614,03546.

4-Amino-5-[3-(4-amino-5-fluoropyrimidin-2-on-1-yl)prop-1-yn-1-yl]-7-(2-deoxy-β-D-ribofuranosyl)-7*H*-pyrrolo[2,3-*d*]pyrimidin 5'-*O*-triphosphate (dA^{PFC}TP)

This compound was prepared according to the general procedure E from 5-iodo-2'-deoxyadenosine triphosphate **dA^{I}TP** as white solid (43 %).^1H NMR (499.8 MHz, D_2O, pD =

7.1, phosphate buffer, ref(dioxane) = 3.75 ppm): 2.48 (ddd, 1H, J_{gem} = 14.0, $J_{2'b,1'}$ = 6.3, $J_{2'b,3'}$ = 3.3, H-2'b); 2.66 (ddd, 1H, J_{gem} = 14.0, $J_{2'a,1'}$ = 7.8, $J_{2'a,3'}$ = 6.3, H-2'a); 4.11 (ddd, 1H, J_{gem} = 11.3, $J_{H,P}$ = 5.4, $J_{5'b,4'}$ = 4.1, H-5'b); 4.18 (ddd, 1H, J_{gem} = 11.3, $J_{H,P}$ = 6.4, $J_{5'a,4'}$ = 4.1, H-5'a); 4.23 (qd, 1H, $J_{4',3'}$ = $J_{4',5'}$ = 4.1, $J_{H,P}$ = 1.4, H-4'); 4.75 (ddd, 1H, $J_{3',2'}$ = 6.3, 3.3, $J_{3',4'}$ = 4.1, H-3'); 4.78 (s, 2H, CH_2N, overlapped with HDO); 6.56 (dd, 1H, $J_{1',2'}$ = 7.8, 6.3, H-1'); 7.70 (s, 1H, H-8); 7.97 (d, 1H, $J_{H,F}$ = 5.8, H-6'''); 8.10 (s, 1H, H-2).^{13}C NMR (150.9 MHz, D_2O, pD = 7.1, phosphate buffer, ref(dioxane) = 69.3 ppm): 41.26 (CH_2-2'); 43.52 (CH_2N); 68.17 (d, $J_{C,P}$ = 5.2, CH_2-5'); 73.74 (CH-3'); 80.95 (-**C**≡C-CH_2-); 85.69 (CH-1'); 88.00 (d, $J_{C,P}$ = 8.8, CH-4'); 88.42 (-C≡**C**-CH_2-); 98.27 (C-7); 105.75 (C-5); 129.46 (CH-8); 132.88 (d, $J_{C,F}$ = 32.0, CH-6'''); 139.89 (d, $J_{C,F}$ = 243.2, C-5'''); 151.44 (C-4); 155.08 (CH-2); 158.89 (C-2'''); 160.04 (C-6); 161.29 (d, $J_{C,F}$ = 14.2, C-4'''). ^{31}P{^{1}H} NMR (202.3 MHz, D_2O, pD = 7.1, phosphate buffer, ref(phosphate buffer) = 2.35 ppm): -21.57 (dd, J = 20.7, 19.2, P_β); -10.46 (d, J = 19.2, P_α); -6.76 (bd, J = 20.7, P_γ).^{19}F{^{1}H} NMR (470.3 MHz, D_2O, pD = 7.1, phosphate buffer): -162.19. MS (ESI$^+$): *m/z (%)*: 494.1 (35) [M – $H_3O_6P_2$ - H]$^-$, 574.1 (100) [M – H_2O_3P - H]$^-$, 596.1 (80) [M - H_2O_3P - 2H + Na]$^-$. HRMS (ESI$^+$): *m/z* [M - H] $^-$ calculated for $C_{18}H_{20}O_{13}N_7FP_3$: 654.03190; found 654.03214.

4-Amino-5-[3-(4-amino-5-fluoropyrimidin-2-on-1-yl)prop-1-yn-1-yl]-1-(2-deoxy-β-D-ribofuranosyl)pyrimidin-2-on 5'-*O*-triphosphate (dC^{PFC}TP)

This compound was prepared according to the general procedure E from 5-iodo-2'-deoxycytidine triphosphate **dC^{I}TP** as white solid (80 %). ^{1}H NMR (499.8 MHz, D_2O, pD = 7.1, phosphate buffer, ref(dioxane) = 3.75 ppm): 2.30 (dt, 1H, J_{gem} = 14.0, $J_{2'b,1'}$ = $J_{2'b,3'}$ = 6.7, H-2'b); 2.44 (ddd, 1H, J_{gem} = 14.0, $J_{2'a,1'}$ = 6.3, $J_{2'a,3'}$ = 4.3, H-2'a); 4.15-4.26 (bm, 3H, H-4',5'); 4.60 (ddd, 1H, $J_{3',2'}$ = 6.0, 4.4, $J_{3',4'}$ = 2.9, H-3'); 4.80 (s, 2H, CH_2N, overlapped with HDO); 6.24 (dd, 1H, $J_{1',2'}$ = 6.7, 6.3, H-1'); 8.01 (d, 1H, $J_{H,F}$ = 5.8, H-6'''); 8.22 (s, 1H, H-6).^{13}C NMR (150.9 MHz, D_2O, pD = 7.1, phosphate buffer, ref(dioxane) = 69.3 ppm): 42.09 (CH_2-2'); 43.06 (CH_2N); 67.81 (d, $J_{C,P}$ = 4.9, CH_2-5'); 72.89 (CH-3'); 79.02 (cyt-**C**≡C-CH_2-); 88.30 (d, $J_{C,P}$ = 8.7, CH-4'); 89.10 (CH-1'); 92.06 (cyt-C≡**C**-CH_2-); 94.29 (C-5); 133.06 (d, $J_{C,F}$ = 31.7, CH-6'''); 139.95 (d, $J_{C,F}$ = 222.4, C-5'''); 148.32 (CH-6); 158.75 (C-2); 158.91 (C-2'''); 161.35 (d, $J_{C,F}$ = 14.3, C-4'''); 167.83 (C-4); ^{31}P{^{1}H} NMR (202.3 MHz, D_2O, pD = 7.1, phosphate buffer, ref(phosphate buffer) = 2.35 ppm): -21.45 (dd, J = 20.1, 19.0, P_β); -10.17 (d, J = 19.0, P_α); -7.42 (d, J = 20.1, P_γ). ^{19}F{^{1}H} NMR (470.3 MHz, D_2O, pD = 7.1, phosphate buffer): -

161.89. MS (ESI^+): *m/z (%)*: 471.1 (48) [M – $H_3O_6P_2$ - H]$^-$, 551.1 (100) [M – H_2O_3P - H]$^-$, 573.1 (70) [M - H_2O_3P - 2H + Na]$^-$. HRMS (ESI^+): *m/z* [M - H] $^-$ calculated for $C_{16}H_{19}O_{14}N_6FP_3$: 631.01489; found 631.01616.

Primer extension experiments:

Monoincorporations:

The reaction mixture (20 µl) contained KOD XL DNA Polymerase (Novagen®, 2.5 U/µl, 0.02 µl), dGTP (4 mM, 0.05 µl), dC^CTP, dU^CTP or dA^CTP (4 mM, 1 µl), primer (3 µM, 1 µl, Primer: 3'-GGGTACGGCGGGTAC-5'), and 19-mer template (3 µM, 1.5 µl) in KOD XL DNA polymerase reaction buffer (2 µl) supplied by the manufacturer. Primer was labeled by use of [γ^{32}P]-ATP according to standard techniques. Reaction mixtures were incubated for 10 min at 60 °C in a thermal cycler and were stopped by addition of stop solution (40 µl, 80 % [v/v] formamide, 20 mM EDTA, 0.025 % [w/v] bromphenole blue, 0.025 % [w/v] xylene cyanol) and heated 5 min at 95 °C. Reaction mixtures were separated by use of a 12 % denaturizing PAGE. Visualization was performed by phosphoimaging.

Multi-incorporations:

The reaction mixture (20 µl) contained KOD XL DNA Polymerase (Novagen®, 2.5 U/µl, 0.1 µl), dNTP (either natural or functionalized, 4 mM, 1 µl), primer (3 µM, 1 µl, Primer: 3'-GGGTACGGCGGGTAC-5'), and 31-mer template (3 µM, 1.5 µl, $temp^{4base}$ or $temp^{4N}$) in KOD XL DNA polymerase reaction buffer (2 µl) supplied by the manufacturer. Primer was labeled by use of [γ^{32}P]-ATP according to standard techniques. Reaction mixtures were incubated for 30 min at 60 °C in a thermal cycler and were stopped by addition of stop solution (40 µl, 80 % [v/v] formamide, 20 mM EDTA, 0.025 % [w/v] bromphenole blue, 0.025 % [w/v] xylene cyanol) and heated 5 min at 95 °C. Reaction mixtures were separated by use of a 12 % denaturizing PAGE. Visualization was performed by phosphoimaging.

Polymerase chain reactions:

98-mer

The PCR reaction mixture (20 µl) contained KOD XL DNA Polymerase (Novagen®, 2.5 U/µl, 0.02 µl), DMSO (100 %, 1 µl), formamide (5 %, 1 µl), betaine (0.75 M, 1 µl), TMAC

(50 mM, 1 µl), natural dNTP (4 mM, 0.5 µl), functionalized dNTP (4 mM, 2 µl), primers (10 µM, 2 µl, LT25TH: 5‘-CAAGGACAAAATACCTGTATTCCTT-3‘ and 10 µM, 2 µl, L20-: 5‘-GACATCATGAGAGACATCGC-3‘), and a 98-mer template (1 µM, 0.5 µl, FVL-A: 5‘-GACATCATGAGAGACATCGCCTCTGGGCTAATAGGACTACTTCTAATCTGTAAGA GCAGATCCCTGGACAGGCAAGGAATACAGGTATTTTGTCCTTG-3‘) in KOD XL reactions buffer (2 µlb) supplied by the manufacturer. 30 PCR cycles were run under the following conditions: denaturation for 1 min at 94 °C, annealing for 1 min at 55 °C, extension for 1 min at 72 °C, followed by final extension step of 5 min at 72 °C. PCR products were analyzed on a 2 % agarose gel in 0.5x TBE buffer, followed by staining with Ethidium Bromide (Sigma Aldrich, 500 µg/ml).

287-mer

The PCR reaction mixture (20 µl) contained KOD XL DNA Polymerase (Novagen®, 2.5 U/µl, 0.5 µl), DMSO (100 %, 1 µl), formamide (5 %, 1 µl), betaine (0.75 M, 1 µl), TMAC (50 mM, 1 µl), natural dNTP (4 mM, 0.5 µl) functionalized dNTP (4 mM, 2 µl), primers (20 µM, 2 µl, Prim S1-HIV1: 5‘-GAT CAC TCT TTG GCA GCG ACC CCT CGT CAC -3‘ and 20 µM, 2 µl, Prim S2-HIV1: 5‘-TTA AAG TGC AGC CAA TCT GAG TCA ACA GAT-3‘), and a 297-mer template (74.02 ng/µl, 0.3 µl, wt HIV-1 PR: 5‘-CCTCAGATCACTCTTTGGCAGCGACCCCTCGTCACAATAAAGATAGGGGGGCAAT TAAAGGAAGCTCTATTAGATACAGGAGCAGATGATACAGTATTAGAAGAAATGA ATTTGCCAGGAAGATGGAAACCAAAAATGATAGGGGGAATTGGAGGTTTTATCA AAGTAAGACAGTATGATCAGATACTCATAGAAATCTGCGGACATAAAGCTATAG GTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAATCTGTTGAC TCAGATTGGCTGCACTTTAAATTTT-3‘) in KOD XL reactions buffer (2 µl) 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 60 °C, extension for 30 s at 75 °C, followed by final extension step of 2 min at 75 °C. PCR products were analyzed on a 1.3 % agarose gel in 0.5x TBE buffer, followed by staining with Ethidium Bromide (Sigma Aldrich, 500 µg/ml).

Inhibition assay

Primer extension

The reaction mixture (152 µl) contained Pwo DNA Polymerase (Peqlab, 2 U/µl, 2.4 µl), dGTP, dATP, dTTP (each 4 mM, 8 µl), **dC^{PC}TP** or **dC^{PFC}TP** (4 mM, 8 µl), primer^{MTRE} (3 µM, 8 µl), and 80-mer template (3 µM, 8 µl) in Pwo DNA polymerase reaction buffer (16 µl) supplied by the manufacturer. Primer was labeled by use of [γ^{32}P]-ATP according to standard techniques. Reaction mixtures were incubated for 30 min at 60 °C in a thermal cycler and the resulting dsDNA was isolated using NucleoSpin-Extract II columns (Macherey-Nagel).

Treatment with M.SssI and cleavage by AfeI

The **dC^{PC}** and **dC^{PFC}** modified DNAs were treated with M.SssI (New England Biolabs) (different concentrations 15−240 nM, (0.4−8 U)) in the presence of NEB 2 buffer (New England Biolabs) and 3.2 µM SAM in total volume of 20 µl for 1 h at 37 °C and then the AfeI (5 U) was added for 1 h at 37 °C.

5.3 Transient and Switchable Protection of DNA Against Cleavage by Restriction Endonucleases by (Trialkylsilyl)ethynyl-Modified 7-Deaza-adenines

4-Amino-5-(trimethylsilyl)ethynyl-7-(2-deoxy-β-D-ribofuranosyl)-7*H*-pyrrolo[2,3-*d*]pyrimidine (dA^{TMSE})

DMF (5 ml) and **TMSA** (19 µl, 1.3 mmol) were added through a septum to an argon-purged vial containing **dA^{I}** (0.26 mmol, 100 mg), $Pd(PPh_3)_2Cl_2$ (0.01 mmol, 9 mg), CuI (1 mg, 0.005mmol) and Et_3N (370 µl, 2.6 mmol). The resulting mixture was stirred at r.t. for 15 min. The product was isolated from crude reaction mixture by column chromatography in $CHCl_3$:MeOH 20/1 followed by precipitation from *i*PrOH and H_2O to gave white solid (110 mg), 79 % yield. Analytical and spectral data were identical with the literature[85a].

4-Amino-5-(triethylsilyl)ethynyl-7-(2-deoxy-β-D-ribofuranosyl)-7*H*-pyrrolo[2,3-*d*]pyrimidine (dA^{TESE})

DMF (8 ml) and **TESA** (59 µl, 3.3 mmol) were added through a septum to an argon-purged vial containing **dA^{I}** (0.66 mmol, 250 mg), $Pd(PPh_3)_2Cl_2$ (0.03 mmol, 23 mg), CuI (12 mg, 0.006 mmol) and Et_3N (932 µl, 2.6 mmol). The product was isolated from crude reaction mixture by column chromatography in $CHCl_3$:MeOH 20/1 followed by precipitation from MeOH and H_2O to gave white solid (185 mg, 72 %). ^{1}H NMR (499.8 MHz, DMSO-d_6): 0.67 (q, 6H, J_{vic} = 7.9, $CH_3C\mathbf{H_2}Si$); 1.01 (t, 9H, J_{vic} = 7.9, $C\mathbf{H_3}CH_2Si$); 2.17 (ddd, 1H, J_{gem} = 13.2, $J_{2'b,1'}$ = 6.0, $J_{2'b,3'}$ = 2.8, H-2'b); 2.47 (ddd, 1H, J_{gem} = 13.2, $J_{2'a,1'}$ = 8.2, $J_{2'a,3'}$ = 5.6, H-2'a); 3.51 (ddd, 1H, J_{gem} = 11.7, $J_{5'b,OH}$ = 5.8, $J_{5'b,4'}$ = 4.4, H-5'b); 3.57 (ddd, 1H, J_{gem} = 11.7, $J_{5'a,OH}$ = 5.5, $J_{5'a,4'}$ = 4.4, H-5'a); 3.82 (td, 1H, $J_{4',5'}$ = 4.4, $J_{4',3'}$ = 2.5, H-4'); 4.33 (m, 1H, $J_{3',2'}$ = 5.6, 2.8, $J_{3',OH}$ = 4.1, $J_{3',4'}$ = 2.5, H-3'); 5.05 (dd, 1H, $J_{OH,5'}$ = 5.8, 5.5, OH-5'); 5.26 (dd, 1H, $J_{OH,3'}$ = 4.1, OH-3'); 6.47 (dd, 1H, $J_{1',2'}$ = 8.2, 6.0, H-1'); 7.84 (s, 1H, H-8); 8.13 (s, 1H, H-2). ^{13}C NMR (125.7 MHz, DMSO-d_6): 4.06 ($CH_3\mathbf{C}H_2Si$); 7.62 ($\mathbf{C}H_3CH_2Si$); 40.00 (CH_2-2'); 62.01 (CH_2-5'); 71.09 (CH-3'); 83.36 (CH-1'); 87.76 (CH-4'); 94.06 (Si-**C**≡C-C7); 94.76 (C-7); 100.61 (C7-**C**≡C-Si); 102.56 (C-5); 127.34 (CH-8); 149.26 (C-4); 153.12 (CH-2); 157.80 (C-6). MS (ESI$^+$): *m/z (%)*: 411.1 (100) $[M - H + Na]^+$. HRMS (ESI$^+$): *m/z* $[M + H]^+$ calculated for $C_{19}H_{29}O_3N_4Si$: 389.2003; found 389.2003. IR: υ = 3137, 2953, 2873, 2147, 1635, 1572, 1455, 1312, 1090, 1017 cm^{-1}.

4-Amino-5-iodo-7-(2-deoxy-β-D-ribofuranosyl)-7*H*-pyrrolo[2,3-*d*]pyrimidin 5'-*O*-monophosphate (dA^{I}MP)

Nucleoside **dA^{I}** (212 mg, 0.5 mmol) was placed to a vial and dried for 2 h at 80°C. Then it was suspended in $PO(OMe)_3$ (1.5 ml, 12 mmol) at 0°C and $POCl_3$ (150 µl, 1.6 mmol) was added. The mixture was stirred at 0 °C for 1 h and then quenched by addition of 2M TEAB (2 ml). The solvents were evaporated in vacuo and the residue was co-distilled several times with water. The product was isolated on a DEAD Sephadex column (300 ml) with elution with gradient of 0 to 1.2M TEAB, evaporated, co-distilled with water (5x) and lyophilized to yield **dA^{I}MP** as a white powder (117 mg, 42 %). ^{1}H NMR (499.8 MHz, D_2O, ref(dioxane) = 3.75 ppm): 2.50 (ddd, 1H, J_{gem} = 13.9, $J_{2'b,1'}$ = 6.2, $J_{2'b,3'}$ = 3.2, H-2'b); 2.63 (ddd, 1H, J_{gem} = 13.9, $J_{2'a,1'}$ = 7.5, $J_{2'a,3'}$ = 6.3, H-2'a); 3.94 (m, 2H, H-5'b); 4.18 (td, 1H, $J_{4',5'}$ = 4.8, $J_{4',3'}$ = 3.2, H-4'); 4.65 (dt, 1H, $J_{3',2'}$ = 6.3, 3.2, $J_{3',4'}$ = 3.2, H-3'); 6.50 (dd, 1H, $J_{1',2'}$ = 7.5, 6.2, H-1'); 7.56 (s, 1H, H-8); 8.00 (s, 1H, H-2). ^{13}C NMR (125.7 MHz, D_2O, ref(dioxane) = 69.3 ppm): 41.32

(CH_2-2'); 54.53 (C-7); 67.18 (d, $J_{C,P}$ = 4.8, CH_2-5'); 74.19 (CH-3'); 85.55 (CH-1'); 88.08 (d, $J_{C,P}$ = 8.5, CH-4'); 106.26 (C-5); 129.65 (CH-8); 151.40 (C-4); 153.57 (CH-2); 159.05 (C-6). $^{31}P\{^1H\}$ NMR (202.3 MHz, D_2O): 2.45. MS (ESI^-): *m/z (%)*: 454.9 (100) $[M - H]^-$. HRMS (ESI^-): *m/z* $[M - H]^-$ calculated for $C_{11}H_{13}O_6N_4I$ P: 454.9622; found 454.9622.

4-Amino-5-(triisopropylsilyl)ethynyl-7-(2-deoxy-β-D-ribofuranosyl)-7*H*-pyrrolo[2,3-*d*]pyrimidin 5'-*O*-monophosphate (dA^{TIPSE}MP)

A CH_3CN/H_2O 1:2 (1 ml) and **TIPSA** (45 µl, 0.2 mmol) were added through a septum to an argon-purged vial containing **dA^{I}MP** (0.1 mmol, 50 mg), CuI (2 mg, 0.01 mmol) and *i*Pr_2NEt (174 µl, 1 mmol). After the solids had dissolved, a solution of $Pd(OAc)_2$ (1.2 mg, 0.005 mmol) and TPPTS (14 mg, 0.02 mmol) in CH_3CN/H_2O 1:2 (1 ml) was injected into the reaction mixture and the resulting mixture was stirred at 80 °C for 30 min. The product was isolated from crude reaction mixture by semi-preparative HPLC on a C18 column with use of linear gradient of 0.1M TEAB in H_2O to 0.1M TEAB in H_2O/MeOH 1:1 as eluent. Several co-distillations with water, followed by freeze drying from water, gave **dA^{TIPSE}MP** as white solid (39 mg, 70 %). ^1H NMR (499.8 MHz, D_2O, $ref_{dioxane}$ = 3.75 ppm): 0.98 (bm, 21H, $(CH_3)_2$CHSi); 2.38 and 2.53 (2 × bm, 2 × 1H, H-2'); 3.90 (bm, 2H, H-5'); 4.13 (bm, 1H, H-4'); 4.48 (bm, 1H, H-3'); 6.34 (bm, 1H, H-1'); 7.43 (bs, 1H, H-8); 8.05 (bs, 1H, H-2). ^{13}C NMR (125.7 MHz, D_2O, $ref_{dioxane}$ = 69.3 ppm): 13.76 (Si**C**H$(CH_3)_2$); 21.02 ((**C**$H_3)_2$CHSi); 40.92 (CH_2-2'); 67.38 (CH_2-5'); 73.91 (CH-3'); 85.60 (CH-1'); 87.92 (CH-4'); 96.98 (Si**C**≡C); 99.12 (**C**≡CSi); 102.72 (C-7); 105.79 (C-5); 129.00 (CH-8); 151.05 (C-4); 155.06 (CH-2); 159.82 (C-6). $^{31}P\{^1H\}$ NMR (202.3 MHz, D_2O,): 2.71. MS (ESI^-): *m/z (%)*: 509.1 (100) $[M - H]^-$. HRMS (ESI^-): *m/z* $[M - H]^-$ calculated for $C_{22}H_{34}N_4O_6PSi$: 509.20; found 509.1990.

4-Amino-5-(triethylsilyl)ethynyl-7-(2-deoxy-β-D-ribofuranosyl)-7*H*-pyrrolo[2,3-*d*]pyrimidin 5'-*O*-monophosphate (dA^{TESE}MP)

A CH_3CN/H_2O 1:2 (1 ml) and **TESA** (64 µl, 0.3 mmol) were added through a septum to an argon-purged vial containing **dA^{I}MP** (0.06 mmol, 30 mg), CuI (1 mg, 0.01 mmol) and *i*Pr_2NEt (100 µl, 0,6 mmol). After the solids had dissolved, a solution of $Pd(OAc)_2$ (0.7 mg, 0.003 mmol) and TPPTS (9 mg, 0.01 mmol) in CH_3CN/H_2O 1:2 (1 ml) was injected into the

reaction mixture and the resulting mixture was stirred with heating 80 °C for 30 min. The product was isolated from crude reaction mixture by semi-preparative HPLC on a C18 column with use of linear gradient of 0.1M TEAB in H_2O to 0.1M TEAB in H_2O/MeOH 1:1 and then MeOH as eluent. Concentration of fractions by rotating evaporator, followed by freeze drying from water, gave **dA^{TESE}MP** as white solid (10 mg, 28 %). ^{1}H NMR (600.1 MHz, CD_3OD): 0.715 (q, 6H, J_{vic} = 7.9, $CH_3C\mathbf{H_2}Si$); 1.07 (t, 9H, J_{vic} = 7.9, $C\mathbf{H_3}CH_2Si$); 1.27 (t, 9H, J_{vic} = 7.3, $C\mathbf{H_3}CH_2N$); 2.32 (ddd, 1H, J_{gem} = 13.5, $J_{2'b,1'}$ = 5.8, $J_{2'b,3'}$ = 2.4, H-2'b); 2.61 (ddd, 1H, J_{gem} = 13.5, $J_{2'a,1'}$ = 8.6, $J_{2'a,3'}$ = 5.9, H-2'a); 3.13 (q, 6H, J_{vic} = 7.3, $CH_3C\mathbf{H_2}N$); 3.98 (dt, 1H, J_{gem} = 11.0, $J_{H,P}$ = $J_{5'b,4'}$ = 4.9, H-5'b); 4.02 (ddd, 1H, J_{gem} = 11.0, $J_{H,P}$ = 5.9, $J_{5'a,4'}$ = 4.1, H-5'a); 4.09 (ddd, 1H, $J_{4',5'}$ = 4.9, 4.1, $J_{4',3'}$ = 2.4, H-4'); 4.59 (dt, 1H, $J_{3',2'}$ = 5.9, 2.4, $J_{3',4'}$ = 2.4, H-3'); 6.64 (dd, 1H, $J_{1'2'}$ = 8.6, 5.8, H-1'); 7.76 (s, 1H, H-8); 8.13 (bs, 1H, H-2). ^{13}C NMR (150.9 MHz, CD_3OD): 5.28 ($CH_3\mathbf{C}H_2Si$); 7.86 ($\mathbf{C}H_3CH_2Si$); 9.18 ($\mathbf{C}H_3CH_2N$); 41.02 (CH_2-2'); 47.51 ($CH_3\mathbf{C}H_2N$); 66.29 (d, $J_{C,P}$ = 5.1, CH_2-5'); 73.21 (CH-3'); 84.70 (CH-1'); 87.50 (d, $J_{C,P}$ = 8.9, CH-4'); 95.64 (Si-**C**≡C-C7); 97.73 (C-7); 100.99 (C7-**C**≡C-Si); 104.30 (C-5); 127.78 (CH-8); 150.46 (C-4); 153.69 (CH-2); 159.21 (C-6). ^{31}P {^{1}H} NMR (202.3 MHz, CD_3OD): 2.42. MS (ESI$^-$): *m/z (%)*: 467.4 (100) $[M - H]^-$.

4-Amino-5-(trimethylsilyl)ethynyl-7-(2-deoxy-β-D-ribofuranosyl)-7*H*-pyrrolo[2,3-*d*]pyrimidin 5'-*O*-monophosphate (dA^{TMSE}MP)

Nucleoside **dA^{TMSE}** (50 mg, 0.1 mmol) was placed to a vial and dried for 2 h at 80 °C. Then it was suspended in $PO(OMe)_3$ (300 µl, 2.5 mmol) at 0°C and $POCl_3$ (14 µl, 0.1 mmol) was added. The mixture was stirred at 0 °C for 1 h and then quenched by addition of 2M TEAB (0.5 ml). The product was isolated from crude reaction mixture by semi-preparative HPLC on a C18 column with use of linear gradient of H_2O and MeOH as eluent. Concentration of fractions by rotating evaporator, followed by freeze drying from water, gave **dA^{TMSE}MP** (42 mg, 62 %) as a white powder. ^{1}H NMR (500.0 MHz, CD_3OD): 0.25 (s, 9H, CH_3Si); 1.21 (t, 9H, J_{vic} = 7.3, $C\mathbf{H_3}CH_2N$); 2.32 (ddd, 1H, J_{gem} = 13.5, $J_{2'b,1'}$ = 5.9, $J_{2'b,3'}$ = 2.6, H-2'b); 2.59 (ddd, 1H, J_{gem} = 13.5, $J_{2'a,1'}$ = 8.4, $J_{2'a,3'}$ = 5.8, H-2'a); 2.97 (q, 6H, J_{vic} = 7.3, $CH_3C\mathbf{H_2}N$); 3.97 (dt, 1H, J_{gem} = 11.0, $J_{H,P}$ = $J_{5'b,4'}$ = 5.0, H-5'b); 4.03 (ddd, 1H, J_{gem} = 11.0, $J_{H,P}$ = 5.9, $J_{5'a,4'}$ = 4.0, H-5'a); 4.09 (m, 1H, H-4'); 4.60 (dt, 1H, $J_{3',2'}$ = 5.8, 2.6, $J_{3',4'}$ = 2.6, H-3'); 6.64 (dd, 1H, $J_{1'2'}$ = 8.4, 5.9, H-1'); 7.76 (s, 1H, H-8); 8.12 (bs, 1H, H-2). ^{13}C NMR (125.7 MHz, CD_3OD): -0.13 (CH_3Si); 9.74 ($\mathbf{C}H_3CH_2N$); 41.14 (CH_2-2'); 47.33 ($CH_3\mathbf{C}H_2N$); 66.17 (d, $J_{C,P}$ = 5.0, CH_2-5'); 73.28 (CH-3'); 84.71 (CH-1'); 87.65 (d, $J_{C,P}$ = 8.8, CH-4'); 97.70 (C-7); 98.03 (C7-**C**≡C-Si);

99.79 (Si-**C**≡C-C7); 104.29 (C-5); 127.84 (CH-8); 150.47 (C-4); 153.65 (CH-2); 159.20 (C-6). ^{31}P {^{1}H} NMR (202.3 MHz, CD_3OD): 2.84. MS (ESI$^-$): *m/z (%)*: 425.1 (100) [M - H]$^-$. HRMS (ESI$^-$): *m/z* [M - H]$^-$ calculated for $C_{16}H_{22}O_6N_4PSi$: 425.1051; found 425.1052.

Deprotection of dA^{XE}MP to dA^{E}MP:

Method A) The mixture of **dA^{TIPSE}MP** (35 mg, 0.04 mmol) and TBAF (77 mg, 0.29 mmol) in CH_3CN:DMF:H_2O 4/1/1 (3 ml) was stirred at 45 °C for 14 h. The products were isolated from crude reaction mixture by semi-preparative HPLC on a C18 column with use of linear gradient of 0.1M TEAB in H_2O to 0.1M TEAB in H_2O/MeOH 1:1 as eluent. Several co-distillations with water, followed by freeze drying from water, gave **dA^{E}MP** as white solid (21 mg, 97 %).

Method B) The solution of **dA^{TESE}MP** (4.6 mg, 0.01 mmol) in NH_3 aq. (2 ml) was stirred with heating 45 °C for 30 min. After evaporation of NH_3 and water followed by freeze drying we obtained **dA^{E}MP** as a white solid (3 mg), 96 % yield.

Method C) The solution of **dA^{TMSE}MP** (20 mg, 0.03 mmol) in NH_3 aq. (8 ml) was stirred with heating 45 °C for 30 min. After evaporation of NH_3 and water followed by freeze drying we obtained **dA^{E}MP** as a white solid (13 mg), 96 % yield.

4-Amino-5-ethynyl-7-(2-deoxy-β-D-ribofuranosyl)-7*H*-pyrrolo[2,3-*d*]pyrimidin 5'-*O*-monophosphate (dA^{E}MP)

^{1}H NMR (600.1 MHz, D_2O, $ref_{dioxane}$ = 3.75 ppm): 2.49 (ddd, 1H, J_{gem} = 14.0, $J_{2'b,1'}$ = 6.2, $J_{2'b,3'}$ = 3.1, H-2'b); 2.68 (ddd, 1H, J_{gem} = 14.0, $J_{2'a,1'}$ = 7.7, $J_{2'a,3'}$ = 6.1, H-2'a); 3.69 (s, 1H, HC≡C-); 3.88, 3.91 (2 × dt, 2 × 1H, J_{gem} = 11.4, $J_{H,P}$ = $J_{5',4'}$ = 5.0, H-5'); 4.18 (td, 1H, $J_{4',5'}$ = 5.0, $J_{4',3'}$ = 3.1, H-4'); 4.67 (dt, 1H, $J_{3',2'}$ = 6.1, 3.1, $J_{3',4'}$ = 3.1, H-3'); 6.57 (dd, 1H, $J_{1'2'}$ = 7.7, 6.2, H-1'); 7.75 (s, 1H, H-8); 8.09 (s, 1H, H-2). ^{13}C NMR (150.9 MHz, D_2O, $ref_{dioxane}$ = 69.3 ppm): 41.26 (CH_2-2'); 66.72 (d, $J_{C,P}$ = 4.7, CH_2-5'); 74.38 (CH-3'); 79.14 (-**C**≡CH); 84.37 (-C≡**C**H); 85.69 (CH-1'); 88.55 (d, $J_{C,P}$ = 8.5, CH-4'); 98.22 (C-7); 105.70 (C-5); 129.81 (CH-8); 151.22 (C-4); 155.04 (CH-2); 160.06 (C-6). ^{31}P (^{1}H dec.) NMR (202.3 MHz, D_2O): 3.89. MS (ESI$^-$): *m/z (%)*: 353.1 (100) [M - H]$^-$.

4-Amino-5-(triisopropysilyl)ethynyl-7-(2-deoxy-β-D-ribofuranosyl)-7*H*-pyrrolo[2,3-*d*]pyrimidin 5'-*O*-triphosphate (dA^{TIPSE}TP)

A CH_3CN/H_2O 1:2 (1 ml) and **TIPSA** (22.5 µl, 0.1 mmol) were added through a septum to an argon-purged vial containing **dA^{I}TP** (0.05 mmol, 50 mg), CuI (1 mg, 0.005mmol) and *i*Pr_2NEt (87 µl, 0.5 mmol). After the solids had dissolved, a solution of $Pd(OAc)_2$ (1 mg, 0.002 mmol) and TPPTS (8 mg, 0.01 mmol) in CH_3CN/H_2O 1:2 (1 ml) was injected into the reaction mixture which was stirred at 80 °C for 30 min. The product was isolated from crude reaction mixture by semi-preparative HPLC on a C18 column with use of linear gradient of 0.1 MTEAB in H_2O to 0.1M TEAB in H_2O/MeOH 1:1 as eluent. Several co-distillations with water, followed by freeze drying from water, gave **dA^{TIPSE}TP** as white solid (11 mg, 30 %). ^{1}H NMR (600.1 MHz, CD_3OD+D_2O): 1.14 (m, 21H, $(CH_3)_2CH$); 2.38 (ddd, 1H, J_{gem} = 13.6, $J_{2'b,1'}$ = 6.0, $J_{2'b,3'}$ = 2.4, H-2'b); 2.61 (ddd, 1H, J_{gem} = 13.6, $J_{2'a,1'}$ = 8.5, $J_{2'a,3'}$ = 6.4, H-2'a); 4.11, 4.17 (2 × m, 2 × 1H, H-5'); 4.21 (m, 1H, H-4'); 4.70 (m, 1H, H-3'); 6.58 (dd, 1H, $J_{1'2'}$ = 8.5, 6.0, H-1'); 7.68 (s, 1H, H-8); 8.15 (s, 1H, H-2). ^{13}C NMR (150.9 MHz, CD_3OD+D_2O): 12.24 (**C**H$(CH_3)_2$); 18.99 ((**C**$H_3)_2$CH); 39.98 (CH_2-2'); 66.76 (d, $J_{C,P}$ = 5.8, CH_2-5'); 72.44 (CH-3'); 84.23 (CH-1'); 86.72 (d, $J_{C,P}$ = 9.3, CH-4'); 95.41 (Si-**C**≡C-); 97.82 (C-7); 101.04 (-**C**≡C-Si); 104.13 (C-5); 127.28 (CH-8); 149.90 (C-4); 153.63 (CH-2); 158.87 (C-4). ^{31}P{^{1}H} NMR (162.0 MHz, CD_3OD+D_2O): -21.18 (t, *J* = 18.8, P_β); -10.31 (d, *J* = 18.8, P_α); -8.33 (bd, *J* = 18.8, P_γ). MS (ESI$^-$): *m/z (%)*: 589.2 (100) [M - PO_3H_2] $^-$, 691 (30) [M – H + Na] $^-$. HRMS (ESI$^-$): *m/z* [M - H] $^-$ calculated for $C_{22}H_{34}O_{12}N_4Na_2P_3Si$: 713.0956; found 713.0966.

4-Amino-5-(triethylsilyl)ethynyl-7-(2-deoxy-β-D-ribofuranosyl)-7*H*-pyrrolo[2,3-*d*]pyrimidin 5'-*O*-triphosphate (dA^{TESE}TP)

Method A) A CH_3CN/H_2O 1:2 (1 ml) and **TESA** (63 µl, 0.3 mmol) were added through a septum to an argon-purged vial containing **dA^{I}TP** (0.07 mmol, 50 mg), CuI (1 mg, 0.007 mmol) and *i*Pr_2NEt (87 µl, 0.71 mmol). After the solids had dissolved, a solution of $Pd(OAc)_2$ (1 mg, 0.003 mmol) and TPPTS (10 mg, 0.01 mmol) in water CH_3CN/H_2O 1:2 (1 mlb) was injected into the reaction mixture which was then stirred at 80 °C for 30 min. The product was isolated from crude reaction mixture by semi-preparative HPLC on a C18 column with use of linear gradient of 0.1M TEAB in H_2O to 0.1M TEAB in H_2O/MeOH 1:1 and then MeOH as eluent. Concentration of fractions by rotating evaporator, followed by freeze drying from water, gave **dA^{TESE}TP** as white solid (6 mg, 11 %).

Method B) dA^{TESE} (52 mg, 0.1 mmol) was placed to a vial and dried for 2 h at 80 °C. Then it was suspended in $PO(OMe)_3$ (0.31 ml, 2.6 mmol) at 0°C and $POCl_3$ (18 µl,0.2mmol) was added. The mixture was then stirred at 0 °C for 1 h followed by addition of $(Bu_3NH)_2H_2P_2O_7$ (368 mg, 0.6 mmol) and Bu_3N (127 µl, 0.5 mmol) in DMF (1.5 ml) and stirring at 0 °C for another 1 h. The reaction was quenched by addition of 2M TEAB (0.5 ml) and water (3 ml). The crude reaction mixture was separated by semi-preparative HPLC on a C18 column with use of linear gradient of 0.1M TEAB in H_2O to 0.1M TEAB in H_2O/MeOH 1:1 and then MeOH as eluent. Concentration of fractions by rotating evaporator, followed by freeze drying from water, gave **dA^{TESE}TP** as white solid (45 mg, 47 %). ^{1}H NMR (499.8 MHz, CD_3OD): 0.72 (q, 6H, J_{vic} = 7.9, $CH_3C\mathbf{H_2}Si$); 1.07 (t, 9H, J_{vic} = 7.9, $C\mathbf{H_3}CH_2Si$); 1.29 (bt, 27H, J_{vic} = 7.0, $C\mathbf{H_3}CH_2N$); 2.32 (bm, 1H, H-2'b); 2.61 (bm, 1H, H-2'a); 3.17 (bq, 18H, J_{vic} = 7.0, $CH_3C\mathbf{H_2}N$); 4.10-4.28 (bm, 3H, H-4',5'); 4.68 (bm, 1H, H-3'); 6.62 (dd, 1H, $J_{1'2'}$ = 7.8, 6.1, H-1'); 7.74 (s, 1H, H-8); 8.14 (bs, 1H, H-2). ^{13}C NMR (125.7 MHz, CD_3OD): 5.24 ($CH_3\mathbf{C}H_2Si$); 7.87 ($\mathbf{C}H_3CH_2Si$); 9.16 ($\mathbf{C}H_3CH_2N$); 40.50 (CH_2-2'); 47.48 ($CH_3\mathbf{C}H_2N$); 67.00 (b, CH_2-5'); 72.66 (CH-3'); 84.44 (CH-1'); 87.20 (b, CH-4'); 95.95 (Si-**C**≡C-C7); 97.84 (C-7); 100.85 (C7-**C**≡C-Si); 103.59 (C-5); 127.69 (CH-8); 150.34 (C-4); 153.66 (CH-2); 159.17 (C-6). ^{31}P {^{1}H} NMR (202.3 MHz, CD_3OD): -22.30, -10.17, -9.45 (3 × bs). MS (ESI$^-$): *m/z (%)*: 547.1 (100) [M - PO_3H_2] $^-$, 627 (95) [M - H] $^-$. HRMS (ESI$^-$): *m/z* [M - H] $^-$ calculated for $C_{19}H_{30}O_{12}N_4P_3Si$: 627.0847; found 627.0832.

4-Amino-5-(trimethylsilyl)ethynyl-7-(2-deoxy-β-D-ribofuranosyl)-7*H*-pyrrolo[2,3-*d*]pyrimidine 5'-*O*-triphosphate (dA^{TMSE}TP)

Nucleoside **dA^{TMSE}** (40 mg, 0.1 mmol) was placed to a vial and dried for 2 hours at 80 °C. Then it was suspended in $PO(OMe)_3$ (0.26 ml, 2.3 mmol) at 0 °C and $POCl_3$ (16 µl,0.1mmol) was added. Then the reaction mixture was stirred at 0 °C for 1 h followed by the addition of a solution of $(Bu_3NH)_2H_2P_2O_7$ (316 mg, 0.5 mmol) and Bu_3N (110 µl, 0.4 mmol) in DMF (1 ml) and stirring at 0 °C for another 1 h. The reaction was quenched by addition of 2M TEAB (0.5 ml) and water (3 ml). The crude reaction mixture was separated by semi-preparative HPLC on a C18 column with use of linear gradient H_2O to MeOH as eluent. Concentration of fractions by rotating evaporator, followed by freeze drying from water, gave **dA^{TMSE}TP** as white solid (21 mg, 26 %). ^{1}H NMR (600.1 MHz, CD_3OD): 0.27 (s, 9H, CH_3Si); 0.985 (t, 27H, J_{vic} = 7.3, $C\mathbf{H_3}CH_2CH_2CH_2N$); 1.41 (bs, 18H, J_{vic} = 7.3, $CH_3C\mathbf{H_2}CH_2CH_2N$); 1.69 (bm,

18H, $CH_3CH_2C\mathbf{H_2}CH_2N$); 2.34 (ddd, 1H, J_{gem} = 13.5, $J_{2'b,1'}$ = 6.0, $J_{2'b,3'}$ = 2.9, H-2'b); 2.56 (ddd, 1H, J_{gem} = 13.5, $J_{2'a,1'}$ = 7.9, $J_{2'a,3'}$ = 6.0, H-2'a); 3.10 (bm, 18H, $CH_3CH_2CH_2C\mathbf{H_2}N$); 4.15 (m, 1H, H-4'); 4.20-4.28 (m, 2H, H-5'); 4.65 (dt, 1H, $J_{3',2'}$ = 6.0, 2.9, $J_{3',4'}$ = 2.9, H-3'); 6.62 (dd, 1H, $J_{1'2'}$ = 7.9, 6.0, H-1'); 7.79 (s, 1H, H-8); 8.26 (bs, 1H, H-2). ^{13}C NMR (150.9 MHz, CD_3OD): -0.02 (CH_3Si); 14.04 ($\mathbf{C}H_3CH_2CH_2CH_2N$); 21.01 ($CH_3\mathbf{C}H_2CH_2CH_2N$); 26.50 ($CH_3CH_2\mathbf{C}H_2CH_2N$); 41.43 ($CH_2$-2'); 53.35 ($CH_3CH_2CH_2\mathbf{C}H_2N$); 67.02 (d, $J_{C,P}$ = 4.1, CH_2-5'); 72.64 (CH-3'); 84.91 (CH-1'); 87.63 (d, $J_{C,P}$ = 8.2, CH-4'); 98.07 (C-7); 98.61 (C7-**C**≡C-Si); 99.29 (Si-**C**≡C-C7); 103.31 (C-5); 128.82 (CH-8); 149.15 (C-4); 149.35 (CH-2); 155.65 (C-6). ^{31}P {^{1}H} NMR (202.3 MHz, CD_3OD): -22.64 (bdd, J = 20.8, 19.7, P_β); -10.24 (d, J = 20.8, P_α); -9.40 (d, J = 19.7, P_γ).

Primer extension experiments:

The reaction mixture (20 µl) contained Vento(exo-) DNA Polymerase (New England Biolabs, 2 U/µl, 2 µl), dNTP (either natural or functionalized, 4 mM, 1 µl), primer (3 µM, 1 µl, Primer: 3'-GGGTACGGCGGGTAC-5'), and 30-mer template (3 µM, 1.5 µl, temp^{KpnI}, temp^{RsaI} or temp^{SacI}) in thermopol reaction buffer (2 µl) supplied by the manufacturer. Primer was labeled by use of [γ^{32}P]-ATP according to standard techniques. Reaction mixtures were incubated for 30 min at 60 °C in a thermal cycler.

Sample preparation before gel loading and separation:

Samples (20 µl) were denatured by addition of stop solution (40 µl, 80 % [v/v] formamide, 20 mM EDTA, 0.025 % [w/v] bromphenole blue, 0.025 % [w/v] xylene cyanol) and heated 5 min at 95 °C. Reaction mixtures were separated by use of a 12% denaturizing PAGE. Visualization was performed by phosphorimaging.

TIPSE modified DNA deprotection:

A mixture of CH_3CN:DMF 4/1/1 (150 µl) and TBAF (5.6 mg, 0.01 mmol) was added to 30 µl of reaction mixture after primer extension and the resulting mixture was stirred (1000 rpm) with heating 45 °C for 16 h. Deprotected DNA was precipitated by adding of 3M AcONa buffer pH 5.2 (18 µl) and *i*PrOH (594 µl) and treated at 0 °C for 24 h. After centrifugation 13 400 rpm for 30 min solvent was decanted and washed by 70% EtOH (150 µl). After suspension in H_2O (50 µl) sample was purged on GE Healthcare column.

TMSE and TESE modified DNA deprotection:

A NH_3 aq. (200 µl) was added to 30 µl of reaction mixture after primer extension and stirred (1000 rpm) at 45 °C for 2 h then NH_3 and water was evaporated under reduced pressure. **For PCR products:** A NH_3 aq. (140 µl) was added to 20 µl of reaction mixture after PCR and stirred (1000 rpm) at 45 °C for 2 h then NH_3 and water was evaporated under reduced pressure.

Conditions of cleavage by *RE* after primer extension:

For RsaI: To a solution of unmodified or modified DNA in water (10 µl), NEBuffer 4 (1 µl) and RsaI (0.4 µl) were added and the mixture was incubated for 1 h at 37 °C in a thermal cycler.

For KpnI: To a solution of unmodified or modified DNA in water (10 µl), NEBuffer 1 (1 µl), BSA (0.1 µl, 100 µg/ml) and KpnI (0.4 µl) were added and the mixture was incubated for 1 h at 37 °C in a thermal cycler.

For SacI: To a solution of unmodified or modified DNA in water (10 µl), NEBuffer 1 (1 µl), BSA (0.1 µl, 100 µg/ml) and SacI (0.4 µl) were added and the mixture was incubated for 1 h at 37 °C in a thermal cycler..

Conditions of cleavage by *RE* after deprotection:

For RsaI: To a solution of deprotected modified DNA in water (50 µl), NEBuffer 4 (5 µl) and RsaI (2.5 µl) were added and the mixture was incubated for 1 h at 37 °C in a thermal cycler.

For KpnI: To a solution of deprotected modified DNA in water (50 µl), NEBuffer 1 (5 µl), BSA (0.5 µl, 100 µg/ml) and KpnI (2 µl) were added and the mixture was incubated for 1 h at 37 °C in a thermal cycler.

For SacI: To a solution of deprotected modified DNA in water (50 µl), NEBuffer 1 (5 µl), BSA (0.5 µl, 100 µg/ml) and SacI (2.5 µl) were added and the mixture was incubated for 1 h at 37 °C in a thermal cycler.

Polymerase chain reactions:

98-mer

The PCR reaction mixture (20 µl) contained KOD XL (Novagen, 2.5 U/µl, 0.5 µl) or Vent(exo-) (New England Biolabs, 2.5 U/µl, 0.5 µl) DNA Polymerase, natural dNTP (4 mM, 0.5 µl), functionalized dNTP (4 mM, 2 µl), primers (10 µM, 2 µl, LT25TH: 5'-CAAGGACAAAATACCTGTATTCCTT-3' and 10 µM, 2 µl, L20-: L20-: 5'-GACATCATGAGAGACATCGC-3'), and a 98-mer template (1 µM, 0.5 µl, FVL-A: 5'-GACATCATGAGAGACATCGCCTCTGGGCTAATAGGACTACTTCTAATCTGTAAGA GCAGATCCCTGGACAGGCAAGGAATACAGGTATTTTGTCCTTG-3') in KOD XL reactions buffer (2 µl) or thermopol buffer (2 µl) supplied by the manufacturer. Thirty PCR cycles were run under the following conditions: denaturation for 1 min at 94 °C, annealing for 1 min at 55 °C, extension for 1 min at 72 °C, followed by final extension step of 5 min at 72 °C. PCR products were analyzed on a 2% agarose gel in 0.5x TBE buffer, followed by staining with Ethidium Bromide (Sigma Aldrich, 500 µg/ml).

287-mer

The PCR reaction mixture (20 µl) contained KOD XL DNA Polymerase (Novagen®, 2.5 U/µl, 1.5 µl), natural dNTP (4 mM, 0.5 µl) functionalized dNTP (4 mM, 2 µl), primers (20 µM, 2 µl, Prim S1-HIV1: 5'-GAT CAC TCT TTG GCA GCG ACC CCT CGT CAC -3' and 20 µM, 2 µl, Prim S2-HIV1: 5'-TTA AAG TGC AGC CAA TCT GAG TCA ACA GAT-3'), and a 297-mer template (74.02 ng/µl, 0.3 µl, wt HIV-1 PR: 5'-CCTCAGATCACTCTTTGGCAGCGACCCCTCGTCACAATAAAGATAGGGGG GCAATTAAAGGAAGCTCTATTAGATACAGGAGCAGATGATACAGTATTAGAAGA AATGAATTTGCCAGGAAGATGGAAACCAAAAATGATAGGGGGAATTGGAGGTTT TATCAAAGTAAGACAGTATGATCAGATACTCATAGAAATCTGCGGACATAAAGCT ATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAATCTGT TGACTCAGATTGGCTGCACTTTAAATTTT-3') in KOD XL reactions buffer (2 µl) 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 60 °C, extension for 4 min at 75 °C, followed by final extension step of 2 min at 75 °C. PCR products were analyzed on a 1.3 % agarose gel in 0.5x TBE buffer, followed by staining with Ethidium Bromide (Sigma Aldrich, 500 µg/ml).

Conditions of cleavage by *RE* for PCR products:

To a solution of modified DNA in water (20 µl), NEBuffer 4 (2 µl) and RsaI (2 µl) were added and the mixture was incubated for 1 h at 37 °C in a thermal cycler.

5.4 Construction of TdaD Gene Using (Trialkylsilyl)ethynyl-Modified DNA

Strains and culture conditions.

Escherichia coli strains were grown in LB medium (10 g/l tryptone, 5 g/l yeast extract, 5 g/l NaCl, pH = 7.2) at 37°C with shaking (200 rpm) or on corresponding solid agar medium (20 g/l agar) unless otherwise stated. The medium was spiked after autoclavation with kanamycin sulfate (50 µg/ml).

Plasmid construction.

Primers used in this study are listed in **Table 20**.

PCR 1

The PCR reaction mixture (20 µl) contained KOD XL DNA polymerase (Novagen, 2.5 U/ µl, 0.2 µl), natural dCTP, dGTP and dTTP (4 mM, 1 µl), natural dATP (4mM, 1 µl) or functionalized $dA^{TESE}TP$ (4 mM, 1 µl), primers (20 µM, 1 µl, PCR1f and 20 µM, 1 µl, PCR1r) and 434-bp template (24.0 ng/µl, 0.2 µl, 5'-GAGACACCATGTCCAACAAAGCCTTCGAAACCCCGGTCGAAATCCGTTACCGCGATACCGATTCCATGGGCCATGTCAGCAGCCCCGTCTATTACGACTACATGCAGTCTGCCTATCTGGAATACATGCATGACCTTCTTGAGCTGCCAAAGTCCGAGAAGCTGCCCCATATCATGGTCAAAACCTCCTGCGATTACATTTCTCAGGCGTTGTATGGCGACAACCTCGTCGTTTGCAGCCGGGTGATCAGCTTCGGCGGTAAGAGCTTCGAGATGGAGCATGTGATGCATATCGACGATGACGATCGCCGCGTCGTAGCCAATGCCAAATCGGTGCATGTGATGTTCGACTACGACAAACAGGCCACCTATCCGGTGCCGGATGCCTTCAAAGAAAGCGTTGCAGCTTTTCAGGAAACCGCCTGATCCCTGTTT-3') in KOD XL reaction buffer (2 µl) supplied by manufacturer. Forty PCR cycles were run under the following conditions: denaturation 95 °C, 1 min, anneling 52 °C 1 min, extension 72 °C 1 min, followed by final extension step 72 °C, 2 min. PCR products were purified on

NucleoSpin ExtractII column (Macherey-Nagel) and analyzed on 1.3% agarose gel stained with GelRed nucleid acid stain in 0.5xTBE buffer.

PCR 2

The PCR reaction mixture (20 µl) contained KOD XL DNA polymerase (Novagen, 2.5 U/ µl, 1 µl), natural dCTP, dGTP and dTTP (4 mM, 1 µl), natural dATP (4mM, 1 µl) or functionalized $dA^{TESE}TP$ (4 mM, 1 µl), primers (20 µM, 1 µl, PCR2f and 20 µM, 1 µl, PCR2r) and 434-bp template (24.0 ng/µl, 1 µl, 5'-GAGACACCATGTCCAACAAAGCCTTCGAAACCCCGGTCGAAATCCGTTACCGCGA TACCGATTCCATGGGCCATGTCAGCAGCCCCGTCTATTACGACTACATGCAGTCT GCCTATCTGGAATACATGCATGACCTTCTTGAGCTGCCAAAGTCCGAGAAGCTGC CCCATATCATGGTCAAAACCTCCTGCGATTACATTTCTCAGGCGTTGTATGGCGAC AACCTCGTCGTTTGCAGCCGGGTGATCAGCTTCGGCGGTAAGAGCTTCGAGATGG AGCATGTGATGCATATCGACGATGACGATCGCCGCGTCGTAGCCAATGCCAAATC GGTGCATGTGATGTTCGACTACGACAAACAGGCCACCTATCCGGTGCCGGATGCC TTCAAAGAAAGCGTTGCAGCTTTTCAGGAAACCGCCTGATCCCTGTTT-3') in KOD XL reaction buffer (2 µl) supplied by manufacturer. Forty PCR cycles were run under the following conditions: denaturation 95 °C, 1 min, anneling 52 °C 1 min, extension 72 °C 1 min, followed by final extension step 72 °C, 2 min. PCR products were purified on NucleoSpin ExtractII (Macherey-Nagel) column and analyzed on 1.3% agarose gel stained with GelRed nucleid acid stain in 0.5 x TBE buffer.

PCR 3

The PCR reaction mixture (20 µl) contained KOD XL DNA polymerase (Novagen, 2.5 U/µl, 1 µl), natural dNTPs (10 mM, 1 µl), **PCR 1** (39.7 ng/µl, 5 µl) and **PCR 2** (87.8 $ng.\mu l^{-1}$, 5 µl) products in KOD XL reaction buffer (2 µl) supplied by manufacturer. Ten PCR cycles were run under the following conditions: denaturation 95 °C, 1 min, anneling 50 °C 1 min, extension 72 °C 1 min, followed by final extension step 72 °C, 2 min. PCR products were purified on Macherey-Nagel NucleoSpin ExtractII column and analyzed on 1.3% agarose gel stained with GelRed nucleid acid stain in 0.5xTBE buffer.

Deprotection of TES

Into the **PCR 3** product (42.35 ng/µl, 10 µl) was added 70 µl of aq. NH_3 and sttired for 1.5 h at 45 °C. The NH_3 was evaporated under reduced pressure.

Conditions of cleavage by RE

The **PCR 3** product (42.35 ng/µl, 10 µl) in reaction mixture (20 µl) was cleaved by NcoI (1 µl, New England Biolabs, Ipswich, MA, USA, 10000 U/ml) and XhoI (1 µl, New England Biolabs, 10000 U/ml) in NEBuffer 4 (2 µl) in the presence of the BSA (0.2 µl, 10 mg/ml, New England Biolabs). The reaction mixture was treated for 1 h at 37 °C and analyzed on 1.3% agarose gel stained with GelRed nucleid acid stain in 0.5xTBE buffer.

The protected PCR amplicon was digested with *Nco*I and *Xho*I (New England Biolabs) and ligated with T4 DNA ligase (New England Biolabs) into pET-28c(+), cut with *Nco*I and *Xho*I, to give the pET28c(+) expression vector containing the *tdaD* insert. The plasmid was sequenced (**Figure 46**) and transformed into *E. coli* DH-5α by electroporation. Verification of clones was performed by DNA sequencing (Seqlab GmbH, Göttingen). For protein expression, the plasmid was transformed into *E. coli* BL-21 (DE3) by electroporation.

Table 20. Oligonucleotide primers used in this study.

Primer	Sequence
PCR1f	GAGACACCATGGCCAACAAAGC
PCR1r	GTATTCCAGATAGGCAGACTGC
PCR2f	GTCGAAATCCGTTACCGC
PCR2r	AAACAGCTCGAGGGCGGTTTCCTG

Expression, protein extraction and purification.

An culture of *E. coli* BL-21 (DE3) transformed with pET28c(+) containing the *tdaD* insert (250 ml) was grown to OD_{600} ~0.4 at 37 °C and equilibrated at 16 °C for 30 min. Expression was induced with IPTG (final concentration 100 µM) at 16 °C for 20 h. The cells were centrifuged (4 °C, 5000 rpm, 10 min), washed with STE buffer (25 ml, 100 mM NaCl, 10 mM Tris-HCl, 1 mM EDTA, pH 7.5) and stored in five aliquots at –20 °C. One aliquot was suspended in binding buffer (5 ml, 500 mM NaCl, 20 mM Na_2HPO_4, 20 mM imidazole, pH

7.4) and sonicated on ice (10 x 15 s). The cell debris was spun down (4 °C, 5000 rpm, 30 min) and the supernatant was sterile-filtered. The soluable proteins were purified by Ni^{2+}-affinity chromatography (1 ml Ni-NTA superflow, QIAGEN, Hilden, Germany). Loading onto the column, washing with binding buffer (10 ml) and elution with elution buffer (5 x 500 µL, 500 mM NaCl, 20 mM Na_2HPO_4, 500 mM imidazole, pH 7.4) gave purified TdaD-HIS.

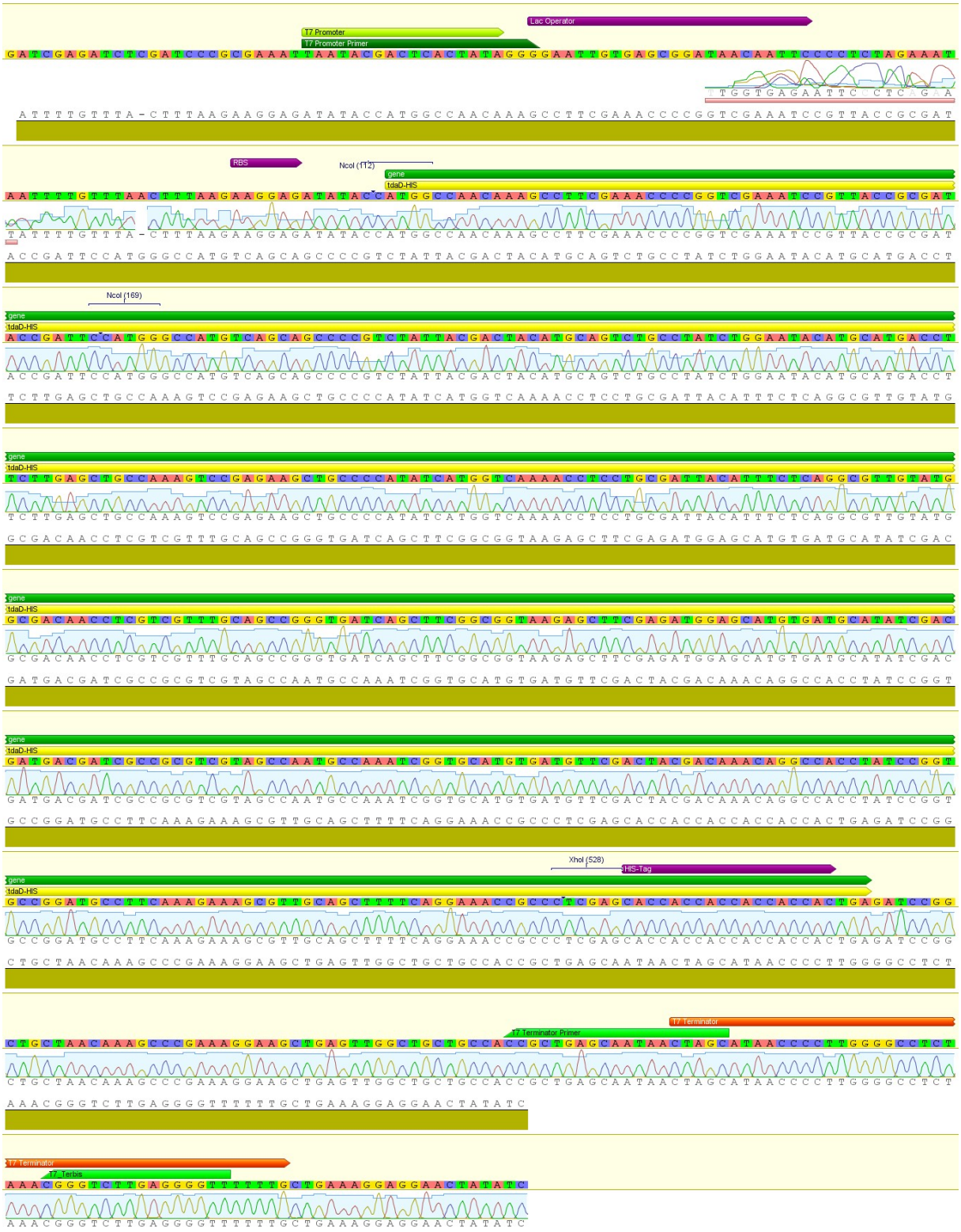


Figure 46. Sequence of the *tdaD* gene cloned into the pET28c(+) expression vector.

5.5 Study of Direct Competition Incorporations of 7-Substituted 7-Deazaadenine 2'-Deoxyribonucleoside Triphosphates in Presence of Natural Counterparts

Compounds **dA^{7D}TP**[104], **dA^{E}TP**[85a,12b], **dA^{Ph}TP**[12b], **dA^{NH2}TP**[49b], **dA^{NO2}TP**[49b], **dA^{BFU}TP**[12a], **dA^{ABOX}TP**[12a], **dA^{STr}TP**[97], **dC^{FT}TP**[96], **dC^{E}TP**[71], **dC^{NH2}TP**[49b], **dC^{NO2}TP**[49b], **dC^{STr}TP**[97] were available in our laboratory. Mass spectra of short DNAs were measured by MALDI-TOF, Reflex IV (Bruker) with nitrogen laser. Concentrations of DNA were measured on Nanodrop 1000 Spectrophotometer (Thermo Scientific). Synthetic oligonucleotides were purchased from Generi Biotech. Dynabeads M-270 Streptavidin (DB_{Stv}) were obtained from Novagen. DNA polymerases: Vent (exo-), Taq, DNA polymerase I (Klenow) and Bst large fragment were purchased from New England Biolabs, KOD XL DNA polymerase from Novagen, Pwo from Peqlab, Human DNA polymerase α from Chimerx. All restriction endonucleases were purchased from New England Biolabs. Unmodified nucleoside triphosphates (dATP, dTTP, dCTP, and dGTP) from Fermentas, and γ^{32}P-ATP from MP Empowered Discovery. Other chemicals were of analytical grade. Primers were labeled by use of [γ ^{32}P]-ATP (if it is not stated otherwise) according to standard techniques. 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 dehydrated after the electrophoresis and exposed to a phosphor screen, and products were analysed with a phosphorimager (Typhoon 9410, Amersham Biosciences). All solutions for biochemistry experiments were prepared in Milli-Q water. V_{max} extension is 10 pmol.s^{-1}.U^{-1}, K_m is in μM.

5.5.1 Method Validation

Direct Competition Assay. Incorporation of dA^{Ph}TP and dC^{Ph}TP Employing Pwo Polymerase in Proof-of-Principle Experiment.

Competitive incorporation of dATP vs **dA^{Ph}TP**: The reaction mixture (30 μl) contained Pwo (0.3 U), natural dNTPs (dGTP, dTTP and dCTP, 0.2 mM), for ratio 1/1 dATP (0.1 mM) and **dA^{Ph}TP** (0.1 mM), for ratio 1/10 dATP (0.02 mM) and **dA^{Ph}TP** (0.2 mM), primer (0.15 μM), temp^{BamH1A} (0.225 μM) in reaction buffer for Pwo from supplier. Reaction mixtures were incubated for 30 min at 60 °C. The reaction mixtures were then divided into two portions (12 and 18 μl). The stop solution (20 μl) was added to the 12 μl portions which was then

denatured at 95 °C for 5 min and analysed using 12.5% denaturing PAGE. The second portion was used in following cleavage reaction. All experiments were done in triplicate.

Competitive incorporation of dCTP vs **dC^{Ph}TP**: The procedure was same as written above with the following changes in reaction mixture – natural dNTPs (dGTP, dATP and dTTP, 0.2 mM), for ratio 1/1 dCTP (0.1 mM) and **dC^{Ph}TP** (0.1 mM), for ratio 1/10 dCTP (0.02 mM) and **dC^{Ph}TP** (0.2 mM), primer (0.15 μM), temp^{Cla1C} (0.225 μM).

Cleavage by Restriction Endonucleases:

The 18 μl portion of product of PEX experiments were mixed with 1 x reaction buffer supplied by manufacturer relevant to RE, BSA (100 μg/ml, 0.2 μl) and corresponding RE (10 U). Reaction mixtures were incubated at 37 °C for 60 min and then the stop solution was added (40 μl). Products of cleavage by RE were denatured at 95 °C for 5 min and analysed using 12.5% denaturing PAGE.

MALDI General:

All samples were prepared in triplicates and from each sample four mass spectra were acquired. Calculations of amount of the modified DNA in the samples were calculated according to the equation $p = \frac{v \times 100}{(v+s)}$ in which p is percentage of relative intensity of modified DNA over internal standard DNA ,v is relative intensity of modified DNA peak, s is relative intensity of the internal standard peak. The calibration curves were made using linear regression.

Magnetic Separation Procedure:

Reaction mixture (100 μl) and NaCl (2.5M, 0.6 μl) was added to suspension of Dynabeads® M-270 Streptavidin (DBStv) [50 μl of stock solution washed three times by wash buffer (0.3M NaCl, 10 mM Tris, pH 7.4, 200 μl)]. The suspensions were shaken (1200 rpm) at 20 °C for 30 min. The DBStv beads were washed three times with phosphate-buffer saline (PBS) (0.14M NaCl, 3mM KCl, 4 mM sodium phosphate, pH 7.4, 200 μl), then three times with wash buffer (200 μl) and twice with H_2O (200 μl). Single-strand oligonucleotides were

released into 50 µl of H_2O by shaking (900 rpm) at 60 °C for 2 min. Buffer exchanges were performed using a magnetoseparator (Dynal).

MALDI Calibration Curves:

The reaction mixture for preparation of the ssDNA samples (100 µl) contained Pwo DNA polymerase (5 U), natural dNTPs (dGTP, dTTP and dCTP, 0.2 mM), dATP (0.2 mM) or **dA^{Ph}TP** (0.2 mM), primer (1.35 µM), 5´-biotinylated temp$^{BamH\text{I}A}$ (1.35 µM) in 1 x reaction buffer for Pwo. Reaction mixtures were incubated for 30 min at 60 °C in a thermal cycler. Single strand DNAs were isolated as is described above. The reaction mixture for preparation of an internal standard ssDNA (190 µl) contained KOD XL DNA polymerase (12.5 U), dNTPs (0.2 mM), primer (1.42 µM), temp$^{MALDI\text{-}IS}$ (1.42 µM) in 1 x reaction buffer for KOD XL. Reaction mixture was incubated for 30 min at 60 °C in thermal cycler. Single strand DNA was isolated as is described above.
Concentrations of both the natural ssDNA and modified ssDNA were adjusted to 10.4 ng/µl and mixed together into the standards containing 10-90 % of modified DNA and the total volume of 10 µl, the internal standard (3 µl, 24.3 ng/µl) was added. All calibration curves were measured in triplicates.

The procedure for calibration curve with **dC^{Ph}** was same as written above with the following changes in reaction mixture – natural dNTPs (dGTP, dATP and dTTP, 0.2 mM), dCTP (0.2 mM) or **dC^{Ph}TP** (0.2 mM), 5´-biotinylated temp$^{Cla\text{I}C}$ (1.35 µM).

MALDI Samples:

Competitive incorporation of dATP vs **dA^{Ph}TP**: The reaction mixture for preparation of the ssDNA samples (100 µl) contained Pwo DNA polymerase (5 U), natural dNTPs (dGTP, dTTP and dCTP, 0.2 mM), for ratio 1/1 dATP (0.1 mM) and **dA^{Ph}TP** (0.1 mM), for ratio 1/10 dATP (0.02 mM) and **dA^{Ph}TP** (0.2 mM), primer (1.35 µM), 5´-biotinylated temp$^{BamH\text{I}A}$ (1.35 µM) in 1 x reaction buffer for Pwo. Reaction mixtures were incubated for 30 min at 60 °C in a thermal cycler. Single strand DNA was isolated as is described in magnetic separation procedure. Concentrations and volumes of the samples were adjusted to 10.4 ng/µl and 10 µl then the internal standard (3 µl, 24.3 ng/µl) was added.

Competitive incorporation of dCTP vs **dC^{Ph}TP**: The procedure was same as written above with the following changes in reaction mixture – natural dNTPs (dGTP, dATP and dTTP, 0.2 mM), for ratio 1/1 dCTP (0.1 mM) and **dC^{Ph}TP** (0.1 mM), for ratio 1/10 dCTP (0.02 mM) and **dC^{Ph}TP** (0.2 mM), 5′-biotinylated temp^{ClaIC} (1.35 μM).

5.5.2 Direct Competition Assay

Cleavage by Restriction Endonucleases:

The 18 μl portion of product of PEX experiments were mixed with 1 x reaction buffer supplied by manufacturer relevant to RE [for cleavage by BamHI, ClaI, PstI and KpnI the BSA (100 μg/ml, 0.2 μl) was added] and corresponding RE (10 U). Reaction mixtures were incubated at 37 °C for 60 min and then the stop solution was added (40 μl). Products of cleavage by RE were denatured at 95 °C for 5 min and analysed using 12.5% denaturing PAGE.

dA^{Ph}TP and dC^{Ph}TP Employing Different DNA Polymerases:

Competitive incorporation of dATP vs **dA^{Ph}TP**: The reaction mixture (30 μl) contained DNA polymerase (Table S3), natural dNTPs (dGTP, dTTP and dCTP, 0.2 mM), for ratio 1/1 dATP (0.1 mM) and **dA^{Ph}TP** (0.1 mM), for ratio 1/10 dATP (0.02 mM) and **dA^{Ph}TP** (0.2 mM), primer (0.15 μM), temp^{BamHIA} (0.225 μM) in reaction buffer. Reaction mixtures were incubated for 30 min, for incubation temperatures see Table S3. The reaction mixtures were then divided into two portions (12 and 18 μl). The stop solution (20 μl) was added to the 12 μl portions which was then denatured at 95 °C for 5 min and analysed using 12.5% denaturing PAGE. The second portion was used in following cleavage reaction. All experiments were done at least in duplicate.

Competitive incorporation of dCTP vs **dC^{Ph}TP**: The procedure was same as written above with the following changes in reaction mixture – natural dNTPs (dGTP, dATP and dTTP, 0.2 mM), for ratio 1/1 dCTP (0.1 mM) and **dC^{Ph}TP** (0.1 mM), for ratio 1/10 dCTP (0.02 mM) and **dC^{Ph}TP** (0.2 mM), primer (0.15 μM), temp^{ClaIC} (0.225 μM).

dA^{Ph}TP Employing Human polymerase α:

The reaction mixture (15 µL) contained primed template (0.1 µM primer, 0.15 µM temp^{BamHIA}), Human DNA polymerase α (0.75 U), BSA (0.1 mg ml^{-1}), glycerol (10 %), Tris·HCl (pH 7.5, 50 mM), DTT (5 mM), $MgCl_2$ (5 mM), natural dNTPs (dGTP, dTTP and dCTP, 0.1 mM) for ratio 1/1 dATP (50 µM) and **dA^{Ph}TP** (50 µM), for ratio 1/10 dATP (10 µM) and **dA^{Ph}TP** (100 µM). The reactions were carried out at 37 °C for 3 h. The reaction mixtures were then divided into two portions (6 and 9 µl). The stop solution (10 µl) was added to the 6 µl portion. The 9 µl portions of product of PEX experiment were mixed with NaCl (1 M, 1 µl) and RE (10 U). Reaction mixtures were incubated at 37 °C for 60 min and then the stop solution (20 µl) was added. All samples were then denatured at 95 °C for 5 min and analysed using 12.5% denaturing PAGE.

Incorporations of Multiple Ratios of Natural to Modified Nucleotides.

Competitive incorporation of dATP vs **dA^{Ph}TP**: The reaction mixture (20 µl) contained Bst DNA polymerase, Large Fragment (0.2 U), natural dNTPs (dGTP, dTTP and dCTP, 0.2 mM), the ratio of dATP to **dA^{Ph}TP** varied from 10/1 to 1/10 (10/1, 5/1, 2/1, 1/1, 1/2, 1/5, 1/10), primer (0.15 µM), temp^{BamHIA} (0.225 µM) in 1 x ThermoPol reaction buffer. Reaction mixtures were incubated for 30 min at 65 °C in a thermal cycler. The stop solution (40 µl) was added to the 12 µl portion which was then denatured at 95 °C for 5 min and analysed using 12.5% denaturing PAGE.

Competitive incorporation of dCTP vs **dC^{Ph}TP**: The procedure was same as written above with the following changes in reaction mixture – natural dNTPs (dGTP, dATP and dTTP, 0.2 mM), the ratio of dCTP to **dC^{Ph}TP** varied from 10/1 to 1/10 (10/1, 5/1, 2/1, 1/1, 1/2, 1/5, 1/10), primer (0.15 µM), temp^{ClaIC} (0.225 µM).

Incorporation of dA^{Ph}TP and dC^{Ph}TP Employing KOD XL Polymerase in Proof-of-Principle PCR Experiment.

The PCR reaction mixture (20 µl) contained KOD XL DNA Polymerase (Novagen®, 2.5 U/µl, 1.5 µl), natural dNTP (4 mM, 0.5 µl) modified **dN^{X}TP** (4 mM, 2 µl), primS1 (20 µM, 2 µl), primS2 (20 µM, 2 µl) and a 297-mer template (74.02 ng/µl, 0.3 µl, wt HIV-1 PR: 5'-CCTCAGATCACTCTTTGGCAGCGACCCCTCGTCACAATAAAGATAGGGGG

GCAATTAAAGGAAGCTCTATTAGATACAGGAGCAGATGATACAGTATTAGAAGA
AATGAATTTGCCAGGAAGATGGAAACCAAAAATGATAGGGGGAATTGGAGGTTT
TATCAAAGTAAGACAGTATGATCAGATACTCATAGAAATCTGCGGACATAAAGCT
ATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAATCTGT
TGACTCAGATTGGCTGCACTTTAAATTTT-3') in KOD XL reactions buffer (2 µl) 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 60 °C, extension for 4 min at 75 °C, followed by final extension step of 2 min at 75 °C. PCR products were cleaved by *Rsa*I under cleavage conditions described below and analyzed on a 1.3 % agarose gel in 0.5x TBE buffer, followed by staining with Ethidium Bromide (Sigma Aldrich, 500 µg/ml).

5.5.3 Steady State Kinetic Assays

Steady State Kinetic Assay with Bacterial DNA Polymerases:

Assays typically contained 5'-6-FAM labelled primer (1 µM) and template (1 µM) (temp^{1A} for **dA^{X}TP** incorporation or temp^{1C} for **dC^{X}TP** incorporation), ThermoPol buffer (1 x) and DNA polymerase, in a total volume of 10 µl. Reactions were initiated via addition of varying concentrations of natural or modified dNTPs, mixtures incubated at 60 °C (except Bst at 65 °C and Klenow at 37 °C) for 5 min (except Bst 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 of incorporation for the modified (V_{max}/K_M) to the efficiency of incorporation for the corresponding natural nucleotide (V_{max}/K_M normalized to 1).

Steady State Kinetic Assay with Human Polymerase α:

Assays contained annealed primer (0,1 µM) and template (0,1 µM) (temp^{1A} for **dA^{X}TP** incorporation or temp^{1C} for **dC^{X}TP** incorporation), BSA (0.1 mg ml^{-1}), glycerol (10 %), Tris·HCl (pH 7.5, 50 mM), DTT (5 mM), $MgCl_2$ (5 mM) and Human polymerase α, in a total volume of 10 µl. Reactions were initiated via addition of varying concentrations of natural or modified dNTPs, mixtures incubated at 37 °C for 30 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.

5.5.4 Computational Studies

In silico models were prepared using the X-ray structure of DNA polymerase complexed to DNA duplex (pdb code 4BDP)[100].The second Mg^{2+} ion was modeled in the position of water A2771, and adenine residue dA3 was modified to thymine dT3. dATP and **dA^{Ph}TP** were docked to the active site by using DOCK 6[105]. Ligand conformations for docking were prepared by Multiconf-DOCK program[106]. The obtained complexes were optimized by using PM6-D3H4 method[107] combined with the COSMO solvent model[108] and LBFGS algorithm. We only optimized docked triphosphates, Mg^{2+} ions, dT3, dT4, dA29, Arg629, Asp653, Asn709, Phe710, Gly711, Ile712, Val713, Tyr714, Asp830 and Glu831. The rest of the complex was frozen. The binding energy (score), was calculated as the sum of the gas-phase interaction energy, the interaction solvation free energy and the change of conformation 'free' energy[109]. The gas phase energies are calculated using PM6-D3H4 method on geometries optimized in water environment. The solvation free energy is calculated using the COSMO solvent model.

5.6 Study of Possible Metabolic DNA Labeling Using Modified 2'-Deoxyribonucleosides and Pronucleotides

4-Amino-5-ethynyl-7-(2-deoxy-β-D-ribofuranosyl)-7*H*-pyrrolo[2,3-*d*]pyrimidin 5'-*O*-[phenyl(benzyloxy-L-alaninyl)]phosphate (Pro-dA^{E})

Nucleoside **dA^{E}** (50 mg, 0.18 mmol) was dissolved in dry THF (5 ml). 1-Methylimidazol (58 µl, 0.73 mmol) was added drop-wise and the reaction mixture was stirred for 15 minutes. Then solution of a phosphorochloridate **5** (129 mg, 0.36 mmol) in THF (2 ml) was added and the reaction mixture was stirred overnight at room temperature. Saturated NH_4Cl solution (2 ml) was added and reaction mixture was extracted with EtOAc and water. Organic layer was dried over $MgSO_4$ and evaporated under reduced pressure. Crude product was purified using semi-preparative HPLC on a C18 column with use of gradient H_2O to MeOH in 2 h. Pronucleotide **Pro-dA^{E}** (34 mg, 31 %) was obtained as white solid (mixture of diastereomers 2 : 1). ^{1}H NMR (500.0 MHz, CD_3OD): 1.26 (dd, 3H, J_{vic} = 7.2, $J_{H,P}$ = 1.2, H-3-Ala-A); 1.31

(dd, 3H, J_{vic} = 7.2, $J_{H,P}$ = 1.1, H-3-Ala-B); 2.29-2.48 (m, 4H, H-2'-A,B); 3.728 (s, 1H, HC≡C-A); 3.731 (s, 1H, HC≡C-B); 3.92 (dq, 1H, $J_{H,P}$ = 9.3, J_{vic} = 7.2, H-2-Ala-A); 3.99 (dq, 1H, $J_{H,P}$ = 10.1, J_{vic} = 7.2, H-2-Ala-A); 4.08 (m, 2H, H-4'-A,B); 4.19-4.30 (m, 4H, H-5'-A,B); 4.48 (dt, 1H, $J_{3',2'}$ = 6.0, 3.5, $J_{3',4'}$ = 3.5, H-3'-B); 4.51 (dt, 1H, $J_{3',2'}$ = 6.0, 3.4, $J_{3',4'}$ = 3.4, H-3'-A); 5.06, 5.11 (2 × d, 2 × 1H, J_{gem} = 12.2, CH_2Ph-B); 5.11, 5.14 (2 × d, 2 × 1H, J_{gem} = 12.3, CH_2Ph-A); 6.56 (dd, 1H, $J_{1',2'}$ = 7.6, 6.3, H-1'-B); 6.58 (dd, 1H, $J_{1',2'}$ = 7.3, 6.4, H-1'-A); 7.17 (m, 2H, H-*p*-Ph-A,B); 7.21 (m, 4H, H-*o*-Ph-A,B); 7.27-7.34 (m, 14H, H-*m*-Ph-A,B, H-*o,m,p*-Bn-A,B); 7.567 (s, 1H, H-6-A); 7.573 (s, 1H, H-6-B); 8.118 (s, 1H, H-2-B); 8.124 (s, 1H, H-2-A). ^{13}C NMR (125.7 MHz, CD_3OD): 20.22 (d, $J_{C,P}$ = 7.4, CH_3-3-Ala-A); 20.38 (d, $J_{C,P}$ = 6.7, CH_3-3-Ala-B); 41.18 (CH_2-2'-A); 41.24 (CH_2-2'-B); 51.58 (CH-2-Ala-B); 51.67 (CH-2-Ala-A); 67.35 (d, $J_{C,P}$ = 5.3, CH_2-5'-A); 67.80 (d, $J_{C,P}$ = 5.5, CH_2-5'-B); 67.97 (**C**H_2-Ph-A,B); 72.26 (CH-3'-A,B); 77.66 (HC≡**C**-B); 77.70 (HC≡**C**-A); 82.29 (H**C**≡C-B); 82.33 (H**C**≡C-A); 84.88 (CH-1'-A); 84.93 (CH-1'-B); 86.32 (d, $J_{C,P}$ = 8.0, CH-4'-B); 86.40 (d, $J_{C,P}$ = 8.6, CH-4'-A); 96.97 (C-5-B); 97.00 (C-5-A); 104.28 (C-4a-A); 104.32 (C-4a-B); 121.44 (d, $J_{C,P}$ = 5.1, CH-*o*-Ph-B); 121.45 (d, $J_{C,P}$ = 4.8, CH-*o*-Ph-A); 126.19 (CH-*p*-Ph-A,B); 127.86 (CH-6-A); 127.92 (CH-6-B); 129.30, 129.33, 129.55, 129.56 (CH-*o,m,p*-Bn-A,B); 130.78 (CH-*m*-Ph-B); 130.81 (CH-*m*-Ph-A); 137.17 (C-*i*-Bn-B); 137.28 (C-*i*-Bn-A); 150.28 (C-7a-B); 150.36 (C-7a-A); 152.07 (d, $J_{C,P}$ = 6.9, C-*i*-Ph-B); 152.13 (d, $J_{C,P}$ = 6.8, C-*i*-Ph-A); 153.33 (CH-2-B); 153.39 (CH-2-A); 158.88 (C-4-B); 158.91 (C-4-A); 174.63 (d, $J_{C,P}$ = 4.8, C-1-Ala-B); 174.86 (d, $J_{C,P}$ = 4.2, C-1-Ala-A). ^{31}P{^{1}H} NMR (162.0 MHz, CD_3OD): 3.63 (B); 3.97 (A). MS (ESI$^-$): *m/z (%)*: 592.1 (60.97) [M + H]$^+$, 614.1 (100) [M + Na]$^+$. HRMS (ESI$^-$): *m/z* [M + H]$^+$ calculated for $C_{29}H_{31}N_5O_7P$: 592.19541; found 592.19556. IR: υ^{-cm} = 3316, 2976, 1744, 1629, 1592, 1457, 1213, 1150, 1093, 1010, 936, 752, 693 cm^{-1}.

4-Amino-5-phenyl-7-(2-deoxy-β-D-ribofuranosyl)-7*H*-pyrrolo[2,3-*d*]pyrimidin 5'-*O*-[phenyl(benzyloxy-L-alaninyl)]phosphate (Pro-dA^{Ph})

Nucleoside **dA^{Ph}** (120 mg, 0.37 mmol) was dissolved in dry THF (17 ml). 1-Methylimidazol (0.12 ml, 1.5 mmol) was added drop-wise and the reaction mixture was stirred for 15 minutes. Then solution of a phosphorochloridate **5** (260 mg, 0.74 mmol) in THF (3 ml) was added and the reaction mixture was stirred overnight at room temperature. Saturated NH_4Cl solution (1 ml) was added and reaction mixture was extracted with EtOAc and water. Organic layer was dried over $MgSO_4$ and evaporated under reduced pressure. Crude product was purified using semi-preparative HPLC on a C18 column with use of gradient H_2O to MeOH in 2 h.

Pronucleotide **Pro-dA^{Ph}** (77 mg, 32 %) was obtained as colorless solid (mixture of diastereomers 10 : 9). ^{1}H NMR (500.0 MHz, $CDCl_3$): 1.25 (bd, 6H, J_{vic} = 6.9, H-3-Ala-A,B); 2.41-2.55 (m, 4H, H-2'-A,B); 3.93-4.02 (m, 2H, H-2-Ala-A,B); 4.10 (m, 2H, H-4'-A,B); 4.22-4.35 (m, 4H, H-5'-A,B); 4.57 (dt, 1H, $J_{3',2'}$ = 6.3, 4.6, $J_{3',4'}$ = 4.6, H-3'-B); 4.62 (td, 1H, $J_{3',2'}$ = 5.7, 4.3, $J_{3',4'}$ = 5.7, H-3'-A); 4.97 (d, 1H, J_{gem} = 12.2, $CH_a\mathbf{H_b}$Ph-B); 5.02 (d, 1H, J_{gem} = 12.2, $CH_a\mathbf{H_b}$Ph-A); 5.04 (d, 1H, J_{gem} = 12.2, $C\mathbf{H_a}H_b$Ph-B); 5.06 (d, 1H, J_{gem} = 12.2, $C\mathbf{H_a}H_b$Ph-A); 6.72 (t, 1H, $J_{1',2'}$ = 6.7, H-1'-B); 6.75 (t, 1H, $J_{1',2'}$ = 6.5, H-1'-A); 7.01-7.34 (m, 24H, H-6-A,B, H-*p*-Ph-A,B, H-*o,m,p*-PhO-A,B, H-*o,m,p*-Bn-A,B); 7.37 (m, 4H, H-*m*-Ph-A,B); 7.41 (m, 4H, H-*o*-Ph-A,B); 8.249 (s, 1H, H-2-A); 8.252 (s, 1H, H-2-B).^{13}C NMR (125.7 MHz, $CDCl_3$): 20.60 (d, $J_{C,P}$ = 5.6, CH_3-3-Ala-B); 20.65 (d, $J_{C,P}$ = 5.2, CH_3-3-Ala-A); 39.89 (CH_2-2'-A); 39.96 (CH_2-2'-B); 50.20 (CH-2-Ala-B); 50.30 (CH-2-Ala-A); 65.82 (d, $J_{C,P}$ = 4.9, CH_2-5'-A); 66.36 (d, $J_{C,P}$ = 5.4, CH_2-5'-B); 67.19 (**C**H_2-Ph-A,B); 70.92 (CH-3'-A); 71.29 (CH-3'-B); 82.76 (CH-1'-A); 83.04 (CH-1'-B); 84.22 (d, $J_{C,P}$ = 7.1, CH-4'-B); 84.37 (d, $J_{C,P}$ = 6.8, CH-4'-A); 101.28 (C-4a-A); 101.32 (C-4a-B); 117.94 (C-5-B); 118.00 (C-5-A); 119.65 (CH-6-A); 119.82 (CH-6-B); 120.02 (d, $J_{C,P}$ = 4.8, CH-*o*-PhO-A); 120.03 (d, $J_{C,P}$ = 4.8, CH-*o*-PhO-B); 125.00 (CH-*p*-PhO-A,B); 127.45 (CH-*p*-Ph-B); 127.49 (CH-*p*-Ph-A); 128.15 (CH-*o*-Bn-A); 128.16 (CH-*o*-Bn-B); 128.43 (CH-*p*-Bn-A); 128.46 (CH-*p*-Bn-B); 128.56 (CH-*m*-Bn-A,B); 128.76 (CH-*o*-B); 128.78 (CH-*o*-A); 129.04 (CH-*m*-A,B); 129.62 (CH-*m*-PhO-A); 129.65 (CH-*m*-PhO-B); 134.08 (C-*i*-Ph-B); 134.12 (C-*i*-Ph-A); 135.03 (C-*i*-Bn-B); 135.10 (C-*i*-Bn-A); 150.27 (C-7a-B); 150.33 (d, $J_{C,P}$ = 6.8, C-*i*-Ph-B); 150.35 (d, $J_{C,P}$ = 6.4, C-*i*-Ph-A); 150.36 (C-7a-A); 150.59 (CH-2-B); 150.62 (CH-2-A); 156.23 (C-4-A,B); 173.26 (d, $J_{C,P}$ = 6.8, C-1-Ala-A); 173.32 (d, $J_{C,P}$ = 6.4, C-1-Ala-B). ^{31}P{^{1}H} NMR (162.0 MHz, $CDCl_3$): 2.93 (B); 3.31 (A). MS (ESI$^-$): *m/z (%)*: 644.6 (91.31) $[M + H]^+$, 666,6 (100) $[M + Na]^+$. HRMS (ESI$^-$): *m/z* $[M + H]^+$ calculated for $C_{33}H_{34}N_5O_7NaP$: 666.21; found 666.20881. IR: υ^{-cm} = 3318, 2948, 1743, 1624, 1588, 1568, 1461, 1213, 11,48, 1024, 931, 760, 694 cm^{-1}.

4-Amino-5-[4-(1-benzofuran-2-yl)-2-fluorophenyl]-7-(2-deoxy-β-D-ribofuranosyl)-7*H*-pyrrolo[2,3-*d*]pyrimidine 5'-*O*-[phenyl(benzyloxy-L-alaninyl)]phosphat (Pro-dA^{BFU})

Nucleoside **dA^{BFU}** (11 mg, 0.024 mmol) was dissolved in dry THF (3 ml). 1-Methylimidazol (8 µl, 0.096 mmol) was added drop-wise and the reaction mixture was stirred for 15 minutes. Then solution of a phosphorochloridate **5** (17 mg, 0.048 mmol) in THF (0.5 ml) was added and the reaction mixture was stirred overnight at room temperature. Saturated NH_4Cl solution (1 ml) was added and reaction mixture was extracted with EtOAc and water. Organic layer

was dried over $MgSO_4$ and evaporated under reduced pressure. Crude product was purified using semi-preparative HPLC on a C18 column with use of gradient H_2O to MeOH in 2 h. Pronucleotide **Pro-dA^{BFU}** (5 mg, 26 %) was obtained as white solid (mixture of diastereomers 5 : 3). ^{1}H NMR (499.8 MHz, CD_3OD): 1.205 (dd, 3H, J_{vic} = 7.1, $J_{H,P}$ = 1.3, H-3-Ala-A); 1.215 (dd, 3H, J_{vic} = 7.1, $J_{H,P}$ = 1.1, H-3-Ala-B); 2.39 (m, 2H, H-2'b-A,B); 2.50 (ddd, 1H, J_{gem} = 14.0, $J_{2'a,1'}$ = 7.4, $J_{2'a,3'}$ = 6.2, H-2'a-A); 2.53 (ddd, 1H, J_{gem} = 14.0, $J_{2'a,1'}$ = 7.7, $J_{2'a,3'}$ = 6.2, H-2'a-B); 3.83-3.92 (m, 2H, H-2-Ala-A,B); 4.12 (m, 2H, H-4'-A,B); 4.21-4.33 (m, 4H, H-5'-A,B); 4.53 (dt, 1H, $J_{3',2'}$ = 6.2, 3.3, $J_{3',4'}$ = 3.3, H-3'-B); 4.58 (dt, 1H, $J_{3',2'}$ = 6.2, 3.3, $J_{3',4'}$ = 3.3, H-3'-A); 4.97 (d, 1H, J_{gem} = 12.2, $CH_a\mathbf{H_b}$Ph-B); 4.98 (d, 1H, J_{gem} = 12.3, $CH_a\mathbf{H_b}$Ph-A); 5.02 (d, 1H, J_{gem} = 12.3, $C\mathbf{H_a}H_b$Ph-A); 5.04 (d, 1H, J_{gem} = 12.2, $C\mathbf{H_a}H_b$Ph-B); 6.72 (dd, 1H, $J_{1',2'}$ = 7.7, 6.2, H-1'-B); 6.74 (dd, 1H, $J_{1',2'}$ = 7.4, 6.1, H-1'-A); 7.05 (m, 4H, H-*o*-Ph-A,B); 7.09 (m, 2H, H-*p*-Ph-A,B); 7.17-7.25 (m, 14H, H-*m*-Ph-A,B, H-*o,m,p*-Bn-A,B); 7.24-7.29 (m, 4H, H-3,5-BZF-A,B); 7.34 (m, 2H, H-6-BZF-A,B); 7.46 (dd, 2H, $J_{H,F}$ = 10.6, $J_{6,5}$ = 7.9, H-6-C_6H_3F-A,B); 7.51 (s, 1H, H-6-A); 7.52 (s, 1H, H-6-B); 7.55 (m, 2H, H-7-BZF-A,B); 7.64 (m, 2H, H-4-BZF-A,B); 7.68 (dd, 2H, $J_{H,F}$ = 10.5, $J_{3,5}$ = 1.6, H-3-C_6H_3F-A,B); 7.70 (ddd, 2H, $J_{5,6}$ = 7.9, $J_{H,F}$ = 2.0, $J_{5,3}$ = 1.6, H-5-C_6H_3F-A,B); 8.18 (bs, 2H, H-2-A,B).^{13}C NMR (125.7 MHz, CD_3OD): 20.20 (d, $J_{C,P}$ = 7.2, CH_3-3-Ala-A); 20.30 (d, $J_{C,P}$ = 6.9, CH_3-3-Ala-B); 41.29 (CH_2-2'-A); 41.48 (CH_2-2'-B); 51.49 (CH-2-Ala-B); 51.66 (CH-2-Ala-A); 67.32 (d, $J_{C,P}$ = 5.1, CH_2-5'-A); 67.81 (d, $J_{C,P}$ = 5.2, CH_2-5'-B); 67.89 (**C**H_2-Ph-A,B); 72.37 (CH-3'-A); 72.44 (CH-3'-B); 84.77 (CH-1'-A); 84.84 (CH-1'-B); 86.33 (d, $J_{C,P}$ = 8.9, CH-4'-B); 86.35 (d, $J_{C,P}$ = 8.8, CH-4'-A); 102.98 (C-4a-A); 103.03 (C-4a-B); 104.00 (CH-3-BZF-B); 104.02 (CH-3-BZF-A); 111.43 (C-5-A,B); 112.08 (CH-7-BZF-A,B); 113.26 (d, $J_{C,F}$ = 25.0, CH-3-C_6H_3F-B); 113.30 (d, $J_{C,F}$ = 24.9, CH-3-C_6H_3F-A); 121.33 (d, $J_{C,P}$ = 4.3, CH-*o*-Ph-B); 121.36 (d, $J_{C,P}$ = 4.3, CH-*o*-Ph-A); 122.13 (CH-5-C_6H_3F-A,B); 122.37 (CH-4-BZF-A,B); 122.87 (d, $J_{C,F}$ = 15.9, C-1-C_6H_3F-B); 122.90 (d, $J_{C,F}$ = 15.9, C-1-C_6H_3F-A); 123.22 (CH-6-A); 123.28 (CH-6-B); 124.40 (CH-5-BZF-A,B); 126.07, 126.10 (CH-*p*-Ph-A,B, CH-6-BZF-A,B); 129.22, 129.25, 129.28, 129.51 (CH-*o,m,p*-Bn-A,B); 130.41 (C-3a-BZF-A,B); 130.68 (CH-*m*-Ph-A,B); 133.29 (C-4-C_6H_3F-B); 133.36 (C-4-C_6H_3F-A); 133.57 (d, $J_{C,F}$ = 2.6, CH-6-C_6H_3F-A,B); 137.08 (C-*i*-Bn-B); 137.12 (C-*i*-Bn-A); 151.40 (C-7a-B); 151.47 (C-7a-A); 151.949 (CH-2-B); 151.952 (d, $J_{C,P}$ = 6.8, C-*i*-Ph-A,B); 152.02 (CH-2-A); 155.39 (d, $J_{C,F}$ = 2.7, C-2-BZF-A); 155.44 (d, $J_{C,F}$ = 2.6, C-2-BZF-B); 156.41 (C-7a-BZF-A,B); 158.31 (C-4-B); 158.35 (C-4-A); 161.49 (d, $J_{C,F}$ = 245.5, C-2-C_6H_3F-B); 161.52 (d, $J_{C,F}$ = 245.3, C-2-C_6H_3F-A); 174.59 (d, $J_{C,P}$ = 4.7, C-1-Ala-B); 174.79 (d, $J_{C,P}$ = 4.3, C-1-Ala-A).^{19}F{^{1}H} NMR (470.3 MHz, CD_3OD): -112.24 (B); -

112.23 (A).$^{31}P\{^{1}H\}$ NMR (202.3 MHz, CD_3OD): 4.28 (B); 4.63 (A). MS (ESI^-): *m/z (%)*: 800.1 (100) $[M + Na]^+$. HRMS (ESI^-): *m/z* $[M + H]^+$ calculated for $C_{41}H_{38}O_8N_5FP$: 778.24393; found 778.24365. IR: υ^{-cm} = 3317, 2930, 1743, 1590, 1457, 1213, 1149, 1093, 1026, 937, 752, 694 cm^{-1}.

Cell culture medium and handling

For all experiments, HeLa cells were cultivated at 37 °C / 5% CO_2 in DMEM (D-5796, Sigma-Aldrich) containing 10% FBS (F-9665, Sigma-Aldrich), 50'000 units Penicillin and 50 mg Streptomycin per l (Sigma-Aldrich). Cells were grown to confluency and passaged every 2–4 days using Trypsin-EDTA solution (Sigma-Aldrich). Cells were counted using trypan blue staining prior to the determination of seeding densities.

Metabolic labeling of cellular DNA using synthetic nucleosides.

Cells were seeded in 24-well plates containing glass coverslips (VWR; thickness 1.5, diameter 13 mm) at 50'000–100'000 cells per well and incubated overnight. The supernatant was removed, and fresh media solutions containing 10 μM nucleoside or pronucleotide (diluted from 10 mM stock solutions in DMSO) were added. After incubating for 24, the cells were fixed in paraformaldehyde (3.7%) for 15 min at room temperature, quenched with PBS containing 50 mM glycine and 50 mM NH_4Cl for 5 min, and washed with PBS once.

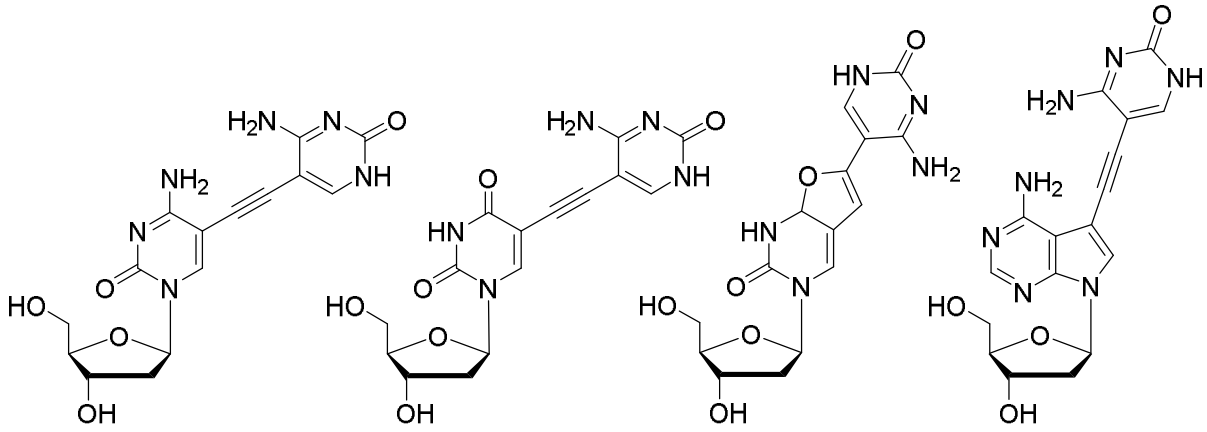
Cu(I)-catalyzed azide-alkyne cycloaddition (CuAAC) staining

Cells were grown and fixed on glass coverslips as described above and incubated upsidedown on 50 μl drops of freshly prepared staining mix (10 μM AlexaFluor 488 azide, 1 mM $CuSO_4$, and 10 mM sodium ascorbate in PBS) for 2 h at room temperature in the dark. The cells were washed with PBS (1 x 2 min), 0.1% Triton X-100 in PBS (1 x 2 min), and PBS (2 x 2 min). After non-covalent staining (600 nM DAPI in PBS, 15 min, room temperature, dark) and washing in PBS (3x) and H_2O (1x), coverslips were glued upside-down on microscopy slides using Glycergel (Dako).

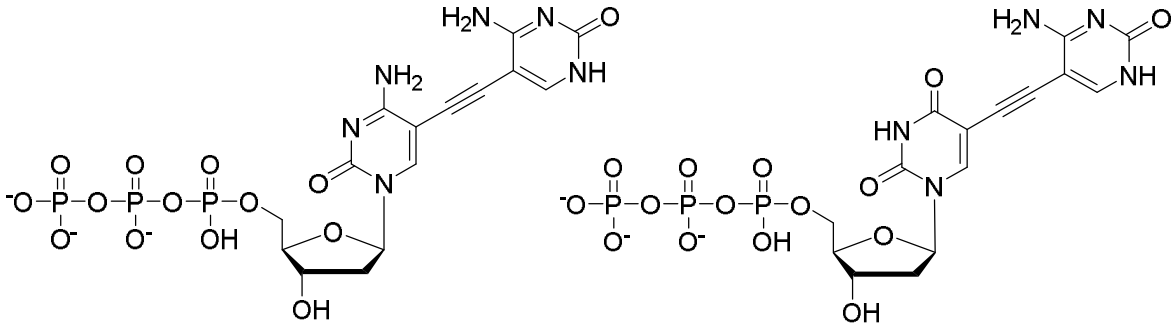
Microscopy

Images were acquired on a Olympus IX 81 microscope. DAPI was excited at 405 nm, and emission was sampled between 415 and 500 nm; AlexaFluor 488 was excited at 488 nm, and emission was sampled between 500 and 560 nm; $\mathbf{dA}^{\mathbf{BFU}}$ was excited at 340 nm, and emission was amplified between 400 and 480 nm. Image analysis was performed using ImageJ (National Institutes of Health, USA).

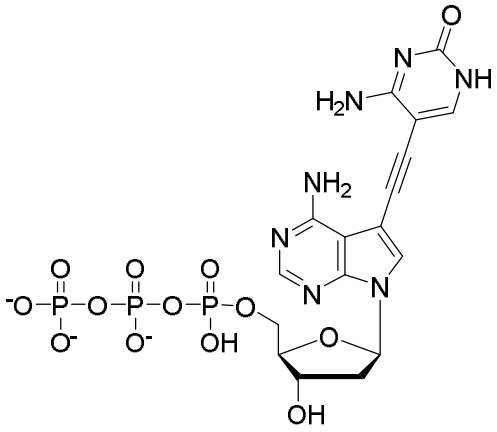
List of all prepared compounds



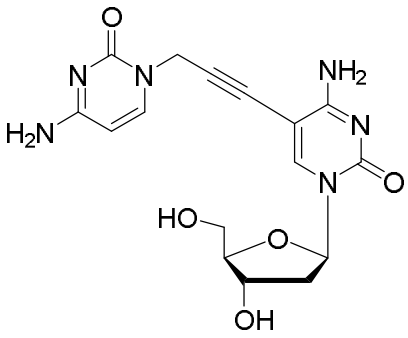
dC^C **dU^C** **dFP^C** **dA^C**



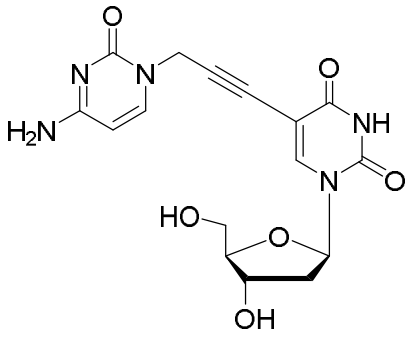
dC^{C}TP **dU^{C}TP**



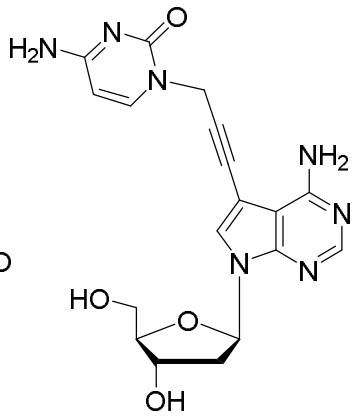
dA^{C}TP



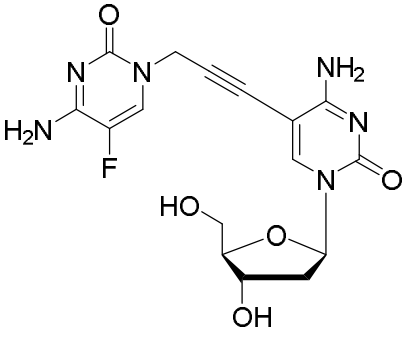
dC^{PC}



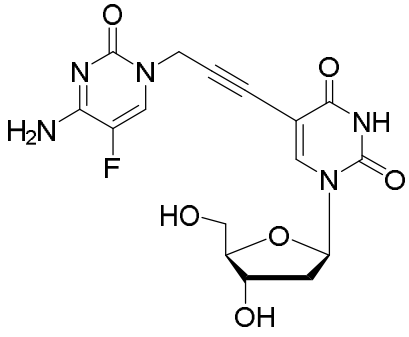
dU^{PC}



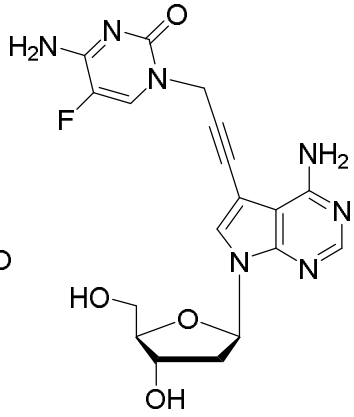
dA^{PC}



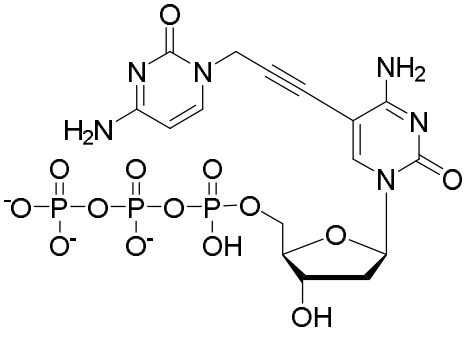
dC^{PFC}



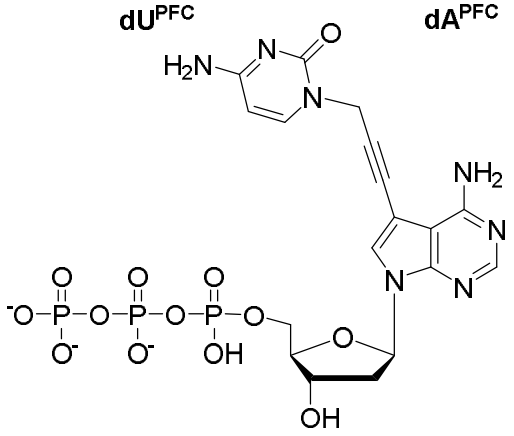
dU^{PFC}



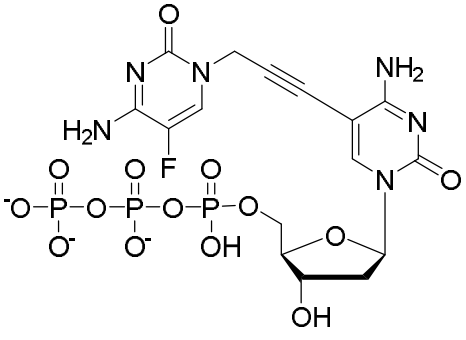
dA^{PFC}



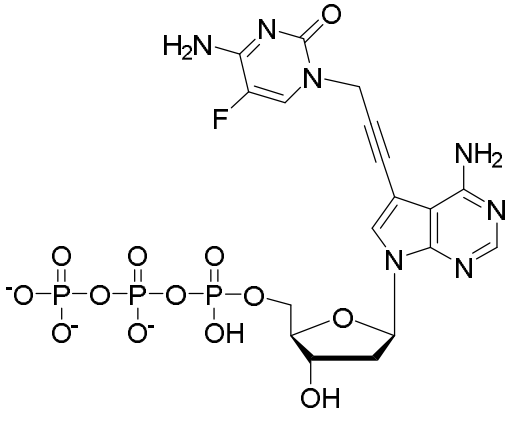
dC^{PC}TP



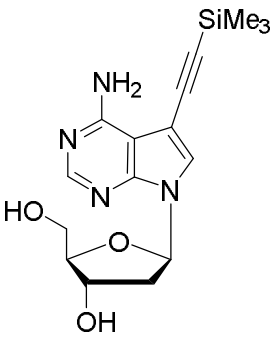
dA^{PC}TP



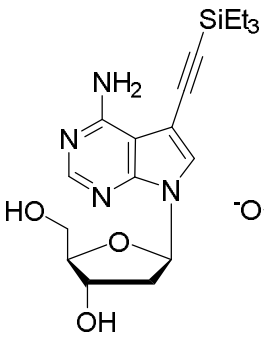
dC^{PFC}TP



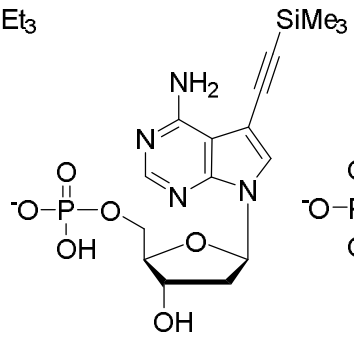
dA^{PFC}TP



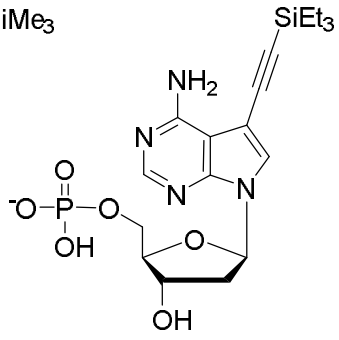
dA^{TMSE}



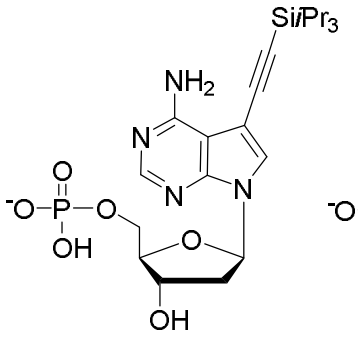
dA^{TESE}



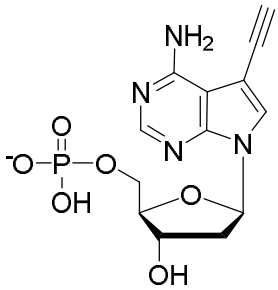
dA^{TMSE}MP



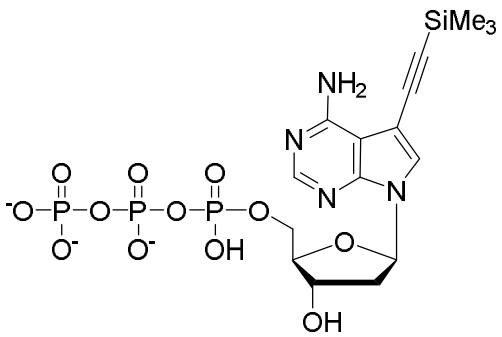
dA^{TESE}MP



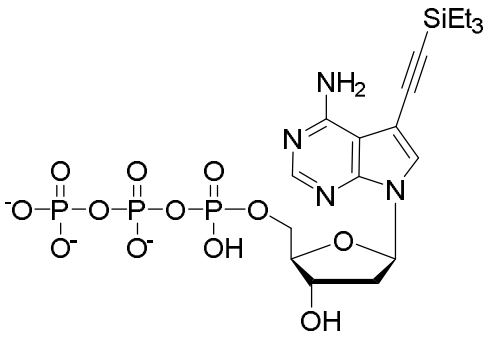
dA^{TIPSE}MP



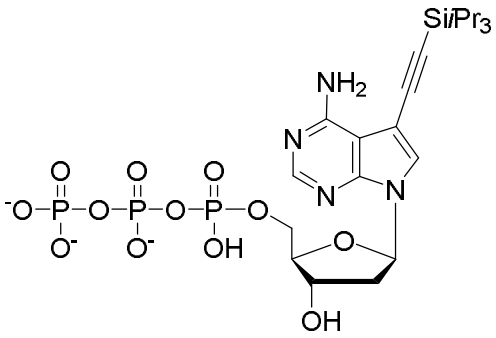
dA^{E}MP



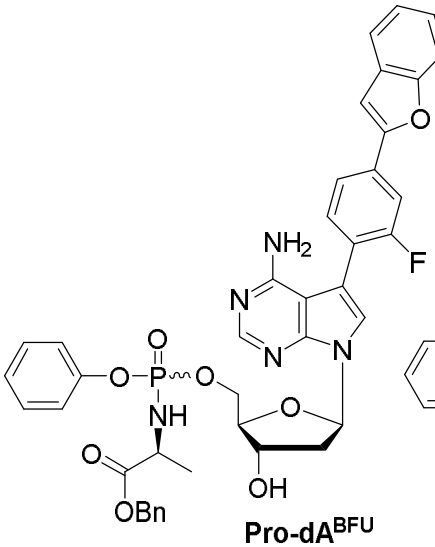
dA^{TMSE}TP



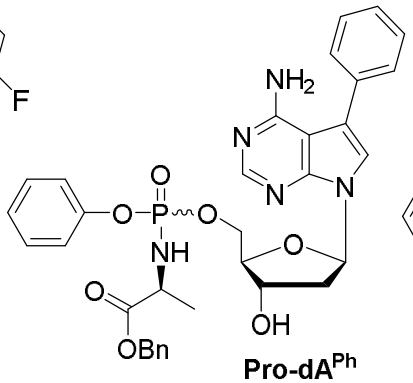
dA^{TESE}TP



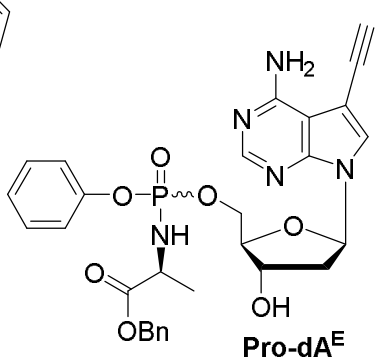
dA^{TIPSE}TP



Pro-dA^{BFU}



Pro-dA^{Ph}



Pro-dA^{E}

References

1 Blackburn, G. M.; Gait, M. J.; Loakes, D.; Williams, D. M.: *Nucleic Acids in Chemistry and Biology* 3rd Edition, The Royal Society of Chemistry **2006** (a) 27 (b) 384−391.

2 (a) Rich, A.; Zhang, S.: *Nat. Rev. Genet.* **2003**, *4*, 566–572. (b) Kim, Y.-G.; Muralinath, M.; Brandt, T.; Pearcy, M.; Hauns, K.; Lowenhaupt, K.; Jacobs B. L.; Rich A.: *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 6974–6979. (c) Oh, D.-B.; Kim, Y.-G.; Rich, A.: *Proc. Natl. Acad. Sci. USA* **2002**, 99, 16666–16671. (d) Herbert, A.; Alfken, J.; Kim, Y.-G.; Mian, I. S.; Nishikura, K.; Rich, A.: *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 8421–8426;

3 Minaso, G.; Tereshko, V.; Egli. M.: *J. Mol. Biol.* **1999**, *291*, 83–99.

4 (a) Lipps, H. J.; Rhodes, D.: *Trends Cell Biol.* **2009**, *19*, 414–422. (b) Huppert, J. L.; Balasubramanian, S.: *Nucleic Acids Res.* **2007**, *35*, 406–413.

5 Smith, Z. D.; Meissner, A.: *Nat. Rev. Genet.* **2013**, *14*, 204–220.

6 Tahiliani, M.; Koh, K. P.; Shen, Y.; Pastor, W. A.; Bandukwala, H.; Brudno, Y.; Agarwal, S.; Iyer, L. M.; Liu, D. R.; Aravind, L.; Rao, A.: *Science* **2009***, 324*, 930–935.

7 (a) Ito, S.; Shen, L.; Dai, Q.; Wu, S. C.; Collins, L. B.; Swenberg, J. A.; He, C.; Zhang, Y.: *Science* **2011**, *333*, 1300–1303. (b) He, Y.-F.; Li, B.-Z.; Li, Z.; Liu, P.; Wang, Y.; Tang, Q.; Ding, J.; Jia, Y.; Chen, Z.; Li, L.; Sun, Y.; Li, X.; Dai, Q.; Song, C.-X.; Zhang, K.; He, C.; Xu, G.-L.: *Science* **2011***, 333*, 1303–1307.

8 Pfaffeneder, T.; Spada, F.; Wagner, M; Brandmayr, C.; Laube, S. K.; Eisen, D.; Truss, M.; Steinbacher, J.; Hackner, B.; Kotljarova, O.; Schuermann, D.; Michalakis, S.; Kosmatchev, O.; Schiesser, S.; Steigenberger, B.; Raddaoui, N.; Kashiwazaki, G.; Müller, U.; Spruijt, C. G.; Vermeulen, M.; Leonhardt, H.; Schär, P.; Müller, M.; Carell, T.: *Nat. Chem. Biol.* **2014**, *10*, 574–581.

9 (a) Bentley, D. R. *et al.*: *Nature* **2008**, *456*, 53–59. (b) Wheeler, D. A.; Srinivasan, M.; Egholm, M.; Shen, Y.; Chen, L.; McGuire, L.; He, W.; Chen, Y.-J.; Makhijani, V.; Roth, G. T.; Gomes, X.; Tartaro, K.; Niazi, F.; Turcotte, C. L.; Irzyk, G. P.; Lupski, J. R.; Chinault, C.; Song, X.-Z.; Liu, Y.; Yuan, Y.; Nazareth, L.; Qin, X.; Muzny, D. M.; Margulies, M.; Weinstock, G. M.; Gibbs, R. A.; Rothberg, J. M.: *Nature* **2008**, *452*, 872 – 876. (c) Balintová, J.; Plucnara, M.; Vidláková, P.; Pohl, R.; Havran, L.; Fojta, M.; Hocek, M.: *Chem. - Eur. J.* **2013**, *19*, 12720–12731.

10 (a) Hocek, M.; Fojta, M.: *Chem. Soc. Rev.* **2011**, *40*, 5802–5814. (b) Paleček, E.; Bartosik, M. *Chem. Rev.* **2012**, *112*, 3427 –3481. (c) Salic, A.; Mitchison, T. J.: *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 2415–2420.

11 (a) McLaughlin, L. W.; Wilson, M.; Ha, S. B.: *Compr. Nat. Prod.Chem.* **1999**, *7*, 251. (b) Kool, E. T.; Morales, J. C.; Guckian, K. M.: *Angew. Chem. Int. Ed.* **2000**, *39*, 990–1009.

12 (a) Riedl, J.; Ménová, P.; Pohl, R.; Orság, P.; Fojta, M.; Hocek, M.: *J. Org. Chem.* **2012**, *77*, 8287–8293. (b) Macíčková-Cahová, H.; Hocek, M.: *Nucleic Acids Res*. **2009**, *37*, 7612–22. (c) Okamoto, A.; Tainaka, K.; Saito, I.: *Bioconjugate Chem.* **2005**, *16*, 1105–1111.

13 Agrofoglio, L. A.; Gillaizeau, I.; Saito, Y.: *Chem. Rev.* **2003**, *103*, 1875–1916.

14 Sakthivel, K.; Barbas III, C. F.: *Angew. Chem. Int. Ed.* **1998**, *37*, 2872 –2875.

15 Seela, F.: *Collect. Czech. Chem. Commun.Special Issue* **2002**, *5*, 1–15.

16 Jager, S.; Famulok, M.: *Angew. Chem., Int. Ed.* **2004**, *43*, 3337–3340.

17 Cremosnik, G.; Hofer, A.; Jessen, H.: *Angew. Chem., Int. Ed.* **2014**, *53*, 286–289.

18 Shaughnessy, K. H.: *Eur. J. Org. Chem.* **2006**, 1827–1835.

19 Hervé, G.; Sartori, G.; Enderlin, G.; Mackenzie, G.; Len, C.: *RSC Adv.* **2014**, *4*, 18558–18594.

20 Hocek, M.; Fojta, M.: *Org. Biomol. Chem.* **2008**, *6*, 2233–2241.

21 (a) Tobrman, T.; Dvořák, D.: *Eur. J. Org. Chem.* **2008**, 2923−2928. (b) Cho, J. H.; Shaughnessy, K. H.: *Synlett* **2011**, 2963−2966. (c) Dadová, J.; Vidláková, P.; Pohl, R.; Havran, L.; Fojta, M.; and Hocek, M.: *J. Org. Chem.* **2013**, *78*, 9627–9637.

22 Cahová, H.; Havran, L.; Brázdilová, P.; Pivoňková, H.; Pohl, R.; Fojta, M.; Hocek, M.: *Angew. Chem., Int. Ed.* **2008**, *47*, 2059 –2062.

23 Brázdilová, P.; Vrábel, M.; Pohl, R.; Pivoňková, H.; Havran, L.; Hocek, M.; Fojta, M.: *Chem.—Eur. J.* **2007**, *13*, 9527-9533.

24 (a) Hammond, D. M.; Manetto, A.; Gierlich, J.; Azov, V. A.; Gramlich, P. M. E.; Burley, G. A.; Maul, M.; Carell, T. *Angew. Chem., Int. Ed* **2007**, *46*, 4184–4187. (b) Fojta, M.; Kostecka, P.; Trefulka, M.; Havran, L.; Palecek, E.: *Anal. Chem.* **2007**, *79*, 1022–1029. (c) El-Sagheer, A. H.; Brown, T.: *Chem. Soc. Rev.* **2010**, *39*, 1388–1405.

25 Khorana, G. H.: *Science* **1979**, *203*, 614–625.

26 Reese, C. B.: *Tetrahedron* **2002**, *58*, 8893–8920.

27 (a) Caruthers, M. H.: *Science* **1985**, *4723*, 281–285. (b) Caruthers, M. H.: *Acc. Chem. Res.* **1991**, *24*, 278–284.

28 (a) Rosemeyer, H.; Ramzaeva, N.; Becker, E.-M.; Feiling, E.; Seela, F.: *Bioconjugate Chem.* **2002**, *13*, 1274–1285. (b) Rozners, E.; Smicius, R.; Uchiyama, C.: *Chem. Commun.* **2005**, 5778– 5780. (c) Kočalka, P.; Andersen, N. K.; Jensen, F.; Nielsen, P.: *ChemBioChem* **2007**, *8*, 2106–2116.

29 Kornberg, A.; Baker, T. A.: *DNA Replication,* 2nd ed., W. H. Freeman and Company: New York, **1991.**

30 (a) Steitz, T. A.: *J. Biol. Chem.* **1999**, *18,* 17395–17398. (b) Patel, P. H.; Loeb, L. A.: *Nat. Struct. Biol.* **2001**, *8*, 656–659.

31 Kranaster, R.; Marx, A.: *ChemBioChem* **2010**, *11*, 2077–2084.

32 Kool, E. T.: Annu. Rev. Biochem. 2002, *71*, 191–219.

33 (a) Kunkel, T. A.; Bebenek, K.: *Annu. Rev. Biochem.* **2000**, *69*, 497. (b) Kool, E. T.; Morales, J. C.; Guckian, K. M.: *Angew. Chem. Int. Ed.* **2000**, *39*, 991. (c) Kunkel, T. A.; Wilson, S. H.: *Nat. Struct. Biol.* **1998**, *5,* 95.

34 Beard, W. A.; Wilson, S. H.: *Chem. Rev.* **2006**, *106*, 361.

35 Ollis, D. L.; Brick, P.; Hamlin, R.; Xuong, N. G.; Steitz, T. A.: *Nature* **1985**, *313*, 762.

36 Wu, S.; Beard, W. A.; Pedersen, L. G.; Wilson, S. H.: *Chem. Rev.* **2014**, *114*, 2759–2774.

37 Bergen, K.; Betz, K.; Welte, W.; Diederichs, K; Marx, A.: *ChemBioChem* 2013, *14*, 1058–1062.

38 Li, Y.; Korolev, S.; Waksman, G.: *EMBO J.* **1998**, *17*, 7514.

39 (a) Hirao, I.; Kimoto, M.; Mitsui, T.; Fujiwara, T.; Kawai, R.; Sato, A.; Harada, Y.; Yokoyama, S.: *Nat. Methods* **2006**, *3*, 729–735. (b) Betz, K.; Malyshev, D. A.; Lavergne, T.; Welte, W.; Diederichs, K.; Dwyer, T. J.; Ordoukhanian, P.; Romesberg, F. E.; Marx, A.: *Nat. Chem. Biol.* **2012**, *8*, 612–614. (c) Li, L.; Degardin, M.; Lavergne, T.; Malyshev, D. A.; Dhami, K.; Ordoukhanian, P.; Romesberg, F. E.: *J. Am. Chem. Soc.* **2014**, *136*, 826–829. (d) Malyshev, D. A.; Dhami, K.; Lavergne, T.; Chen, T.; Dai, N.; Foster, J. M.; Corrêa, I. R.; Romesberg, F. E.: *Nature* **2014**, *509*, 385−388.

40 a) Qu, D.; Wang, G.; Wang, Z.; Zhou, L.; Chi, W.; Cong, S.; Ren, X.; Liang, P.; Zhang, B.: *Anal. Biochem.* **2011**, *417*, 112–121. b) Guan, L.; van der Heijden, G.; Bortvin, A.; Greenberg, M. M.: *ChemBioChem* **2011**, *12*, 2184–2190.

41 a) Bessman, M. J.; Lehman, I. R.; Adler, J.; Zimmerman, S. B.; Simms, E. S.; Kornberg, A.: *Proc. Natl. Acad. Sci. USA* **1958**, *44*, 633–641; b) Fakan, S: *J. Microsc.* **1976**, *106*, 159–171.

42 Neef, A. B.; Luedtke, N. W.: *ChemBioChem* **2014**, *15*, 789–793.

43 (a) Neef, A. B.; Samain, F.; Luedtke, N. W.: *ChemBioChem* **2012**, *13*, 1750–1753. (b) Neef, A. B.; Luedtke, N. W.: *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 20404–20409.

44 Bornfleth, H.; Edelmann, P.; Zink, D.; Cremer, T.; Cremer, C.: *Biophys. J.* **1999**, *77*, 2871–2886.

45 Riedl, J.; Pohl, R.; Ernsting, N.; Orság, P.; Fojta, M.; Hocek, M.: *Chem. Sci.* **2012**, *3*, 2797–2806.

46 Baccaro, A.; Steck, A.-L. L.; Marx, A.: *Angew. Chem., Int. Ed.* **2012**, *51*, 254–257.

47 (a) Bergen, K.; Betz, K.; Welte, W.; Diederichs, K.; Marx, A.: *ChemBioChem* **2013**, *14*, 1058–1062. (b) Obeid, S.; Bußkamp, H.; Welte, W.; Diederichs, K.; Marx, A.: *J. Am. Chem. Soc.* **2013**, *135*, 15667–15669. (c) Bergen, K.; Steck, A.-L. L.; Strütt, S.; Baccaro, A.; Welte, W.; Diederichs, K.; Marx, A.: *J. Am. Chem. Soc.* **2012**, *134*, 11840–11843.

48 a) Obeid, S.; Baccaro, A.; Welte, W.; Diederichs, K.; Marx, A.: *Proc. Natl. Acad. Sci. USA* **2010**, 107, 21327 – 21331. b) Betz, K.; Streckenbach, F.; Schnur, A.; Exner, T.; Welte, W.; Diederichs, K.; Marx, A.: *Angew. Chem. Int. Ed.* **2010**, *49*, 5181 – 5184. c) Baccaro, A.; Steck, A.-L.; Marx, A.: *Angew. Chem. Int. Ed.* **2012**, *51*, 254 – 257. d) Obeid, S.; Busskamp, H.; Welte, W.; Diederichs, K.; Marx, A.: *Chem. Commun.* **2012**, *48*, 8320 – 8322. e) Obeid, S.; Bußkamp, H.; Welte, W.; Diederichs, K.; Marx, A.: *J. Am. Chem. Soc.* **2013**, *135*, 15667 – 15669. f) Bergen, K.; Betz, K.; Welte, W.; Diederichs, K.; Marx, A.: *ChemBioChem* **2013**, *14*, 1058 –1062.

49 (a) Jäger, S.; Rasched, G.; Kornreich-Leshem, H.; Engeser, M.; Thum, O.; Famulok, M.: *J. Am. Chem. Soc.* **2005**, *127*, 15071–82. (b) Marx, A.; Detmer, I.; Gaster, J.; Summerer, D.: *Synthesis* **2004**, *1*, 1–14. (c) Creighton, S.; Bloom, L. B.; Goodman, M. F.: *Methods Enzymol.* **1995**, *262*, 232.

50 Mullis, K. B.; Faloona, F. A.: *Methods Enzymol.***1987**, *155*, 335–350.

51 (a) Ménová, P.; Hocek, M.: *Chem. Commun.* **2012**, *48*, 6921− 6923. (b) Ménová, P.; Raindlová, V.; Hocek, M.: *Bioconjugate Chem.***2013**, *24*, 1081−1093.

52 Ménová, P.; Cahová, H.; Plucnara, M.; Havran, L.; Fojta, M.; Hocek, M.: *Chem. Commun.* **2013**, *49*, 4652−4653.

53 Jurkowska, R. Z.; Jurkowski, T. P.; Jeltsch, A.: *ChemBioChem* **2011**, *12*, 206–222.

54 (a) Cheng, X.; Roberts, R.: *Nucleic Acids Res.* **2001**, *18*, 3784–3795. (b) Chen, L.; MacMillan, A. M.; Chang, W.; Ezez-Nikpay, K.; Lane, W. S.; Verdine, G. L.: *Biochemistry* **1991**, *30*, 11018–11026. (c) Chen, L.; MacMillan, A. M.; Verdine, G. L.: *J. Am. Chem. Soc.* **1993**, *115*, 5318–5319. (d) Santi, D. V.; Garrett, C. E.; Barr, P. J.: *Cell*

1983, *33*, 9–10. (e) Randerath, K.; Tseng, W. C.; Harris, J. S.; Lu, L. J.: *Recent Results Cancer Res.* **1983**, *84*, 283–297. (f) Bestor, T. H.; Verdine, G. L.: *Curr. Opin. Cell Biol.* **1994**, *6*, 380–389.

55 Friedman, J. I.; Stivers, J. T.: *Biochemistry* **2010**, *49*, 4957–4967.

56 Zangi, R.; Arrieta, A.; Cossío, F.: *J. Mol. Biol.* **2010**, *400*, 632–644.

57 Song, J.; Teplova, M.; Ishibe-Murakami, S.; Patel, D. J.: *Science* **2012**, *335*, 709–712.

58 Martinet, N.; Michel, B. Y.; Bertrand, P.; Benhida, R.: *Med. Chem. Commun.* **2012**, *3*, 263–273.

59 (a) Sorm, F.; Piskala, A.; Cihak, A.; Vesely, J.: *Experientia* **1964**, *20*, 85–93. (b) Cihak, A.: *Oncology* **1974**, *30*, 505–522. (c) Jones, P. A.; Taylor, S. M.: *Cell* **1980**, *20*, 85–93. (d) Kaminskas, E.; Farrell, A.; Abraham, S.; Baird, A.; Hsieh, L. S.; Lee, S. L.; Leighton, J. K.; Patel, H.; Rahman, A.; Sridhara, R.; Wang, Y. C.; Pazdur, R.; *Clin. Cancer Res.* **2005**, *11*, 3604–3608. (e) Kantarjian, H.; Issa, J. P.; Rosenfeld, C. S.; Bennett, J. M.; Albitar, M.; DiPersio, J.; Klimkev, V.; Slack, J.; De Castro, C.; Ravandi, F.; Helmer, R.; Shen, L.; Nimer, S. D.; Leavitt, R.; Raza, A.; Saba, H.: *Cancer* **2006**, *106*, 1794–803.

60 Champion, C.; Guianvarc'h, D.; Senamaud-Beaufort, S.; Jurkowska, R. Z.; Jeltsch, A.; Ponger, L.; Arimondo, P. B.; Guieysse-Peugeot, A. L.: *PLoS One* **2010**, *5*, 12388.

61 (a) Gowher, H.; Jeltsch, A.; *Cancer Biol Ther.* **2004**, *11*,1062–1068. (b) Ellis, L.; Atadja, P. W.; Johnstone, R. W.: *Mol Cancer Ther.* **2009**, *8*, 1409–1420. (c) Ren, J.; Singh, B. N.; Huang, Q.; Li, Z.; Gao, Y.; Mishra, P.; Hwa, Y. L.; Li, J.; Dowdy, S. C.; Jiang, S. W.: *Cell Signal.* **2011**, *7*, 1082–1093.

62 Yoder, J. A.; Soman, N. S.; Verdine, G. L.; Bestor, T. H.: *J. Mol. Biol.* **1997**, *270*, 385–395.

63 (a) Klimasauskas, S.; Kumar, S.; Roberts, R. J.; Cheng, X.: *Cell* **1994**, *76*, 357–369. (b) Osterman, D. G.; DePillis, G. D.; Wu, J. C.; Matsuda, A.; Santi, D. V.: *Biochemistry* **1988**, *27*, 5204–5210. (c) Smith, S. S.; Kaplan, B. E.; Sowers, L. C.; Newman, E. M.: *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 4744– 4748.

64 Loenen, W. A.; Dryden, D. T.; Raleigh, E. A.; Wilson, G. G.; Murray, N. E.: *Nucleic Acids Res*. **2014**, *42*, 3–19.

65 Roberts, R. J., Vincze, T., Posfai, J., Macelis, D.: *Nucleic Acids Res.* **2010**, *38*, 234–236.

66 Pingoud, A.; Wilson, G.; Wende, W.: *Nucleic Acids Res*. **2014**, *447*.

67 Komatsu, H.; Kim, S.-G.; Sakabe, I.; Ichikawa, T.; Nakai, M.; Takaku, H.: *Bioorg. Med. Chem. Lett.* **1992**, *2*, 565–570

68 (a) Bodnar J.W.; Zempsky W.; Warder D.; Bergson C.; Ward D.C.: *J. Biol. Chem.* **1983**, *258*, 15206–15213 (b) Zebala J.A.; Choi J.; Trainor G.L.; Barany F.: *J. Biol. Chem.* **1992**, *267*, 8106–8116 (c) Seela F.; Driller H.: *Nucleic Acid Res.* **1986**, *14*, 2319–2332 (d) Seela F.; Kehne A.: *Biochemistry* **1987**, *26*, 2232–2238 (e) Fliess A.; Wolfes H.; Seela F.; Pingoud A.: *Nucleic Acid Res.* **1988**, *16*, 11781–11793 (f) Seela F.; Roling A.: *Nucleic Acid Res.* **1992**, *20*, 55–61.

69 Jiricny J.; Wood S.G.; Martin D.; Ubasawa A.: *Nucleic Acid Res.*, **1986**, *14*, 6579–6590.

70 Chen, F.; Yang, Z.; Yan, M.; Alvarado, J. B.; Wang, G.; Benner, S. A.: *Nucleic Acids Res.* **2011**, *39*, 3949–3961.

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

72 Kielkowski, P.; Pohl, R.; Hocek, M.: *J. Org. Chem.* **2011**, *76*, 3457–3462.

73 (a) Estabrook, R. A.; Lipson, R.; Hopkins, B.; Reich, N.: *J. Biol. Chem.* **2004**, *279*, 31419–31428. (b) Shieh, F. K.; Reich, N. O.: *J. Mol. Biol.* **2007**, *373*, 1157–1168.

74 (a) Slupphaug, G.; Mol, C. D.; Kavli, B.; Arvai, A. S.; Krokan, H. E.; Tainer, J. A.: *Nature* **1996**, *384*, 87–92. (b) Bruner, S. D.; Norman, D. P. G.; Verdine, G. L.: *Nature* **2000**, *403*, 859–866. (c) Parick, S. S.; Walcher, G.; Jones, G. D.; Slupphaug, G.; Krokan, H. E.; Blackburn, G. M.; Tainer, J. A.: *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 5083–5088.

75 Jones, A. S.; Serafinowski, P.; Walker, R. T.: *Tetrahedron Lett.* **1977**, *18*, 2459–2460.

76 Doi, Y.; Chiba, J.; Morikava, T.; Inouye, M.: *J. Am. Chem. Soc.* **2008**, *130*, 8762–8768.

77 (a) Robins, M. J.; Barr, P. J.: *J. Org. Chem.* **1983**, *48*, 1854–1862. (b) Yu, C. J.; Yowanto, H.; Wan, Y. J.; Meade, T. J.; Chong, Y.; Strong, M.; Donilon, L. H.; Kayyem, J. F.; Gozin, M.; Blackburn, G. F.: *J. Am. Chem. Soc.* **2000**, *122*, 6767–6768. (c) Pike, A. R.; Ryder, L. C.; Horrocks, B. R.; Clegg, W.; Connoly, B. A.; Houlton, A.: *Chem.—Eur. J.* **2005**, *11*, 344–353.

78 Kovacs, T.; Ötvös, L.: *Tetrahedron Lett.* **1988**, *29*, 4525–4528.

79 Čapek, P.; Cahová, H.; Pohl, R.; Hocek, M.; Gloeckner, C.; Marx, A.: *Chem.—Eur. J.* **2007**, *13*, 6196–6203.

80 Lindesll, E. W.; Murray, C.; Preston, P. N.; Woodman, T. A. J.: *Tetrahedron* **2000**, *56*, 1233–1256.

81 Duschinsky, R.; Gabriel, T.; Hoffer, M.; Berger, J.; Tistworth, E.; Grunberg, E.; Burchenal, J. H.; Fox, J. J.: *J. Med. Chem.* **1966**, *9*, 566−572.

82 Renbaum, P.; Abrahamove, D.; Fainsod, A.; Wilson, G. G.; Rottem, S.; Razin, A.: *Nucleic Acid Res.*, **1990**, *18*, 1145–1152.

83 Kielkowski, P.; Macíčková-Cahová, H.; Pohl, R.; Hocek, M.: *Angew. Chem. Int. Ed.* **2011**, *50*, 8727–8730.

84 Perona, J. J.: *Methods* **2002**, *28*, 353–364.

85 (a) Seela, F.; Zulauf, M.: *Synthesis* **1996**, *13*, 726–730; b) Seela, F.; Zulauf, M: *Helv. Chim. Acta* **1999**, *82*, 1878–1898; c) Ding, P.; Wunnicke, D.; Steinhoff, H.-J.; Seela, F.: *Chem. Eur. J.* **2010**, *16*, 14385–14396.

86 Gramlich, P. M. E.; Warncke, S.; Gierlich, J.; Carell, T: *Angew. Chem. Int. Ed.* **2008**, *47*, 3442–3444.

87 Tomassini, J.; Roychoudhury, R.; Wu, R.; Roberts, R. J.: *Nucleic Acids Res.* **1978**, *5*, 4055–4064.

88 Yang, R. C. A.; Wu, R.: *J. Virol.* **1978**, *28*, 851–864.

89 Kielkowski, P.; Brock, N. L.; Dickschat, J. S.; Hocek, M.: *ChemBioChem* **2013**, *14*, 801-804.

90 a) Thiel, V.; Brinkhoff, T.; Dickschat, J. S.; Wickel, S.; Grunenberg, J.; Wagner-Dçbler, I; Simon, M.; Schulz, S.: *Org. Biomol. Chem.* **2010**, *8*, 234–246; b) Berger, M.; Brock, N. L.; Liesegang, H.; Dogs, M.; Preuth, I.; Simon, M.; Dickschat, J. S.; Brinkhoff, T.: *Appl. Environ. Microbiol.* **2012**, *78*, 3539–3551.

91 a) Geng, H.; Bruhn, J. B.; Nielsen, K. F.; Gram, L.; Belas, R.: *Appl. Environ. Microbiol.* **2008**, *74*, 1535–1545. b) Thole, S.; Kalhoefer, D.; Voget, S.; Berger, M.; Engelhardt, T.; Liesegang, H.; Wollherr, A.; Kjelleberg, S.; Daniel, R.; Simon, M.; Thomas, T.; Brinkhoff, T.: *ISME J.* **2012**, *6*, 2229–2244.

92 van Cott, E. M.; Wilson, G. G.: *Gene* **1988**, *74*, 55–59.

93 Gingeras, T. R.; Myers, P. A.; Olson, J. A.; Hanberg, F. A.; Roberts, R. J.: *J. Mol. Biol.* **1978**, *118*, 113–122.

94 Kielkowski, P.; Fanfrlík, J.; Hocek, M.: *Angew. Chem. Int. Ed.* **2014**, *53*, 7552−7555.

95 a) Chen, F.; Yang, Z.; Yan, M.; Alvarado, J. B.; Wang, G.; Benner, S. A.: *Nucleic Acids Res.* **2011**, *39*, 3949–3961. b) Yang, Z.; Chen, F.; Alvarado, J. B.; Benner, S. A.: *J. Am. Chem. Soc.* **2011**, *133*, 15105–15112.

96 Raindlová, V.; Pohl, R.; Šanda, M.; Hocek, M.: *Angew. Chem. Int. Ed.* **2010**, *49*, 1064–1066.

97 Macíčková-Cahová, H.; Pohl, R.; Horáková, P.; Havran, L.; Špaček, J.; Fojta, M.; Hocek, M.: *Chem. Eur. J.* **2011**, *17*, 5833–5841.

98 a) Polesky, A. H.; Steitz, T. A.; Grindley, N. D. F.; Joyce, C. M.: *J. Biol. Chem.* **1990**, *265*, 14579–14591. b) Kong, H.; Kucera, R. B.; Jack, W. E.: *J. Biol. Chem.* **1993**, *268*, 1965–1975. c) Patro, J. N.; Urban, M.; Kuchta, R. D.: *Biochemistry* **2009**, *48*, 8271–8278.

99 Bertram, J. G.; Oertell, K.; Petruska, J.; Goodman, M. F.: *Biochemistry* **2010**, *49*, 20–28.

100 Kiefer, J. R.; Mao, C.; Braman, J. C.; Beese, L. S.: *Nature* **1998**, *391*, 304–307.

101 (a) Perrone, P.; Daverio, F.; Valente, R.; Rajyaguru, S.; Martin, J. A.; Lévêque, V.; Le Pogam, S.; Najera, I.; Klumpp, K.; Smith, D. B.;McGuigan, C.: *J. Med. Chem.* **2007**, *50*, 5463–5470. (b) Perrone, P.; Luoni, G. M.; Kelleher, M. R.; Daverio, F.; Angell, A.; Mulready, S.; Congiatu, C.; Rajyaguru, S.; Martin, J. A.; Levêque, V.; Le Pogam, S.; Najera, I.; Klumpp, K.; Smith, D. B.; McGuigan, C.: *J. Med. Chem.* **2007**, *50*, 1840–1849. (c) Mehellou, Y.; Balzarini, J.; McGuigan, C.: *ChemMedChem* **2009**, *4*, 1779–1791. (d) Quintiliani, M.; Persoons, L.; Solaroli, N.; Karlsson, A.; Andrei, G.; Snoeck, R.; Balzarini, J.; McGuigan, C.: *Bioorg. Med. Chem.* **2011**, *19*, 4338–4345.

102 McGuigan, C.; Cahard, D.; Salgado, A.; De Clercq, E.; Baizarinl, J.: *Antiviral Chem. Chemother.***1996**, *7*, 31−36.

103 a) Derudas, M.; Carta, D.; Brancale, A.; McGuigan, C.; Vanpouille, C.; Lisco, A.; Margolis, L.; Balzarini, J.: *J. Med. Chem.* **2009**, *52*, 5520–5530. (b) McGuigan, C.; Thiery, J.-C.; Daverio, F.; Jiang, W.G.; Davies, G.; Mason, M.: *Bioorg. Med. Chem.* **2005**, *13*, 3219–3227. c) Quintiliani, M.; McGuigan, C.; Persoons, L.; Andrei, G.; Snoeck, R.; Balzarini, J.; Solaroli, N.; Karlsson, A.: *Bioorg. Med. Chem.* **2011**, *19*,4338–4345.

104 Newman, P. C.; Nwosu, V. U.; Williams, D. M.; Cosstick, R.; Seela, F.; Connolly. B. A.: *Biochemistry* **1990**, *29*, 9891–9901.

105 Lang, P. T.; Brozell, S. R.; Mukherjee, S.; Pettersen, E. F.; Meng, E. C.; Thomas, V.; Rizzo, R.C.; Case, D. A.; James, T. L.; Kuntz, I. D.: *RNA* **2009**, *15*, 1219–1230.

106 Sauton, N.; Lagorce, D.; Villoutreix, B. O.; Miteva, M. A.: *BMC Bioinformatics* **2008**, *9*, 184.

107 a) Řezáč, J.; Fanfrlík, J.; Salahub, D.; Hobza, P.: *J. Chem. Theory Comput*. **2009**, *5*, 1749-1760. b) Řezáč, J.; Hobza, P.: *J. Chem. Theory Comput.* **2012**, *8*, 141–151.

108 Klamt, A.; Schuurmann, G.: *J. Chem. Soc., Perkin Trans. 2* **1993**, *5*, 799–805.

109 a) Fanfrlík, J.; Bronowská, A. K.; Řezáč, J.; Přenosil, O.; Konvalinka, J.; Hobza, P.: *J.Phys. Chem. B* **2010**, *114*, 12666–12678. b) Lepšík, M.; Řezáč, J.; Kolař, M.; Pecina, A.; Hobza, P.; Fanfrlík, J.: *ChemPlusChem* **2013**, *78*, 921–929.