Univerzita Karlova v Praze

Přírodovědecká fakulta / Faculty of Science

Studijní program: Organická chemie / Organic chemistry

Studijní obor: Organická chemie / Organic chemistry



AMBRA SPAMPINATO

Modified Nucleotides and DNA bearing reactive groups for bioconjugation and cross-linking with aminoacids

Modifikované nukleotidy a DNA nesoucí reaktivní skupiny pro bioconjugace a cross-linkování s aminokyselinami

DISSERTATION

Disertační práce

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

Praha, 2024

The thesis was drawn up at the Institute of organic chemistry and biochemistry, Academy of Sciences (IOCB) in the period September 2018 - September 2024.

Disertační práce byla vypracována na Ústavu organické chemie a biochemie, Akademie věd České Republiky v Praze v období září 2018 - září 2024

Declaration

I declare that I wrote the thesis independently and I cited all the information sources and literature. This work will not be presented to obtain another academic degree or equivalent title.

Prohlášení

Prohlašuji, že jsem diplomovou práci vypracovala samostatně a že jsem uvedla všechny použité informační zdroje a literaturu. Tato práce nebude předložena k získání jiného akademického titulu nebo rovnocenného titulu.

Prague, 2024

V Praze, 2024

M.Sc. Ambra Spampinato



SUMMARY

This PhD thesis is divided in four parts describing the development and the synthesis of nucleotides and DNA-reactive probes bearing different reactive groups for bio-imaging, bioconjugation and cross-linking with aminoacids and peptides. The first part regarding the design and synthesis of a series of 2'-deoxyribonucleoside triphosphates (dNTPs) and DNA-reactive probes bearing 2- or 4-linked *trans*-cyclooctene (TCO) or bicyclononyne (BCN) tethered through a shorter propargylcarbamate or longer triethyleneglycol-based spacer for *in vitro* and *in cellulo* labeling with various tetrazines by IEDDA reaction. The general strategy was based on the functionalization of the position 5 of pyrimidines via acqueous Sonogashira coupling with propargylamino linker and then addition of reactive building block bearing TCO or BCN by nucleophilic substitution to synthetize modified 2'-deoxycitidine monophosphate (dCMP) and triphosphates (dCTPs). The following step was enzymatic incorporation of such functionalized dCTPs into DNA via primer extension (PEX). The 4- linked TCO monophosphate derivative was used, as model compound, for reaction with 3,6-dipyridyl tetrazine to prove the formation of the conjugate by NMR analysis. In parallel, the reactivity of all functionalized triphosphates with excess of various tetrazines was tested in acqueous solution and the fluorescent light-up of formed conjugates was measured. As next step, reactions of functionalized DNA with excess of tetrazines were successful and conjugates were observed by gel analysis and characterized by LC-MS. The last step was the postsynthetic labeling of DNA in cells which consisted in delivering of TCO- and BCN- functionalized dCTPs to cells by SNTT1 transporter and afterwards their incubation with tetrazines to label the nascent genomic DNA. Various combination of modified dNTPs with tetrazines were efficiently used for staining of DNA and imaging of DNA synthesis in live cells, showing the characteristic punctuate pattern of nuclei visualized by confocal microscopy. Also, the BCN-linked nucleotide in combination with TAMRA-linked (TAMRA = carboxytetramethylrhodamine) tetrazine was useful for staining of DNA for flow cytometry.

The second and third part of the thesis describes the development and synthesis of nucleotides and DNA-reactive probes bearing β-sultam/3-oxo- β-sultam or azirine moieties which can be used for cross-linking with Ser- or Glu/Asp- containing peptides and proteins, respectively. The general approach was to insert the reactive building blocks either via Sonogashira cross coupling or click reaction to synthetize the modified 2'-deoxyribonucleoside monophosphate (dNMP) and

triphosphates (dNTPs). In parallel, also the alternative strategy of click reaction on azido-functionalized DNA with β-sultam/3-oxo-β-sultam was attempted. Unfortunately, in both approaches these candidates revealed to be unstable and they were hydrolysed. Therefore, these moieties were not suitable for further conjugation reactions with aforementioned aminoacids.

The fourth and the last part of the thesis was focused on design and synthesis of modified 2′-deoxyuridine nucleotides (dUMP and dUTP) bearing the ABNO−H attached at position 5 via CuAAC click reaction for covalent conjugation with tryptophan and tryptophan containing peptides/proteins. The reactive form of ABNO−H is the radical N-oxyl or the oxoammonium cation, which can generated either by $NaNO_2$ oxidation or through disproportionation of the radical. Therefore, model reactions of keto-ABNO with a large excess of Ac-Trp or tripeptide Ac-AlaTrpAla were performed in presence or absence of $NaNO_2$ to determine the structures of the conjugates and if the presence of the only O_2 as oxidant was enough to generate the reactive form. In parallel, the reactivity of model ABNOH- linked dUMP and DNA was tested under similar conditions. In all model reactions the covalent adducts were successfully formed and characterized even in absence of $NaNO_2$ by partial NMR assignment and mass analysis. Additionally, the ABNOH- linked dUTP was good substrate for KOD XL DNA polymerase in PEX and the full-length product was observed. Moreover, studies with modified DNA probes and Trp- containing peptides were successful and the modified DNA-peptides covalent conjugates were characterized by LC-MS. On the other hand, peptides containing all aminoacids except tryptophan were used as negative controls to prove the selectivity towards tryptophan. As final step, cross-linking reactions with Trp-containing *bs*GntR transcription factor was performed and covalent conjugates were observed by gel analysis and proved by native mass analysis. Moreover, cross-linking reactions with proteins not containing tryptophan (such as histones) or non-binding proteins containing tryptophan (such as lysozyme) did not show any covalent adducts. The keto-ABNOH moiety is a useful reactive group to be inserted in DNA for bioconjugation and cross-linking with Trp-containing peptides or Trp-containing DNA binding proteins.

SOUHRN

Tato disertační práce je rozdělena do čtyř částí popisující vývoj a syntézu nukleotidů a DNA nesoucích různé reaktivní skupiny pro spojování s aminokyselinami a peptidy.

První část zahrnuje návrh a syntézu několika 2'-deoxyribonukleosid trifosfátů (dNTP) a DNA-reaktivních sond nesoucích deriváty 2-/4-*trans*-cyklooktenu (TCO) či bicyklononynu (BCN) vázané buď přes kratší propargylkarbamát nebo delší spojku na bázi triethyleneglykolu pro značení tetraziny pomocí IEDDA reakce jak *in vitro*, tak i v buňkách. Obecná strategie byla založena na derivatizaci polohy 5 u pyrimidinů pomocí vodné Sonogashira reakce s propargylaminovým linkerem a následnou nukleofilní substitucí s reaktivními deriváty TCO nebo BCN poskytující příslušný modifikovaný 2'-deoxycytidin monofosfát (dCMPs) a trifosfáty (dCTPs). Následujícím krokem bylo enzymové začlenění takto funkcionalizovaných dCTPs do DNA prostřednictvím metody prodlužování primeru (PEX). Derivát 4-*trans*-TCO-dCMP byl použit pro modelovou reakci s 3,6-dipyridyltetrazinem za účelem potvrzení vzniku příslušného konjugátu pomocí NMR analýzy. Současně byla snímána interakce všech modifikovaných dCTPs s nadbytkem různých tetrazinů ve vodném roztoku vedoucí ke zvýšení intenzity fluorescence vzniklých konjugátů. V dalším kroku byly úspěšně provedeny reakce funkcionalizované DNA s nadbytkem tetrazinů a vzniklé konjugáty byly charakterizovány gelovou analýzou (PAGE). Posledním krokem bylo „post-syntetické" značení DNA v buňkách, které spočívalo v dodání modifikovaných TCO-/BCN-dCTPs do buněk pomocí transportéru SNTT1 a jejich následné inkubaci s tetraziny, které vedlo k úspěšnému fluorescenčnímu zobrazení nově vznikající genomové DNA v reálném čase pomocí konfokální mikroskopie.

Druhá a třetí část práce popisuje vývoj a syntézu nukleotidů a DNA-reaktivních sond nesoucích β-sultamové nebo azirinové skupiny, vhodné pro připojování s peptidy a proteiny obsahujícími Serin (Ser), resp. kyselinu glutamovou (Glu) a asparágovou (Asp). Obecný přístup spočíval v přípravě reaktivních stavebních bloků 2'-deoxyribonukleosid monofosfátu (dNMPs) a trifosfátů (dNTPs) pomocí Sonogashira reakce nebo mědí katalyzované azid-alkynové cykloadice (CuAAC) a v dalším kroku použít dNTPs ke konstrukci příslušně modifikované DNA. Jako alternativní cesta byla současně vyzkoušena také CuAAC reakce s β-sultamem nebo azirinovým derivátem přímo na DNA již nesoucí azido skupinu.

Bohužel, v obou testovaných případech se ukázalo, že β-sultamové ani azirinové deriváty nejsou dostatečně stabilní, a tudíž konjugace s výše uvedenými aminokyselinami neprobíhala.

Poslední část disertační práce byla zaměřena na návrh a syntézu modifikovaných 2′-deoxyuridinových nukleotidů (dUMP a dUTP) nesoucích ABNO-H navázaný v poloze 5 pomocí CuAAC reakce pro kovalentní konjugace s peptidy či proteiny obsahující Tryptofan (Trp). Reaktivní formou ABNO-H je radikál N-oxyl nebo oxo-amonný kationt, který může vzniknout buď oxidací $NaNO_2$ nebo disproporcionací radikálu. Ke zjištění struktury konjugátů a určení, zda ke vzniku reaktivní formy postačí přítomnost pouze O_2 jako oxidantu, byla provedena modelová reakce keto-ABNO s velkým přebytkem Ac-Trp nebo tripeptidu Ac-AlaTrpAla v přítomnosti nebo nepřítomnosti $NaNO_2$. Podobně byla testována i konjugace s Trp na ABNO-H-modifikovaném monofosfátu, ale i ABNO-H-DNA, protože dUTP nesoucí ABNO-H se ukázal při PEX reakci jako velmi dobrý substrát pro KOD XL DNA polymerázu, a poskytoval odpovídající DNA o správné délce. Ve všech případech byly kovalentní adukty úspěšně syntetizovány v přítomnosti i nepřítomnosti $NaNO_2$ a plně charakterizovány. Jako důkaz selektivity vůči tryptofanu byla také provedena negativní kontrola testující konjugaci ABNO-H-DNA s peptidy obsahujícími všechny ostatní aminokyseliny kromě Trp. Dále byla studována konjugační reakce s transkripčním faktorem *bs*GntR obsahujícím Trp. Tvorba kovalentních konjugátů byla pozorována pomocí gelové analýzy a řádně prokázána nativní hmotnostní analýzou. Posledním krokem pak byla negativní kontrola konjugací s proteiny neobsahujícími Trp (histony) či nevazebnými proteiny (lysozym), která, dle předpokladu, neprokázala žádnou tvorbu kovalentních aduktů.

ACKNOWLEDGEMENT

I would like to thank my supervisor Prof. Michal Hocek for his support and constant guidance in overcoming challenging projects. I would give special thanks to Tomáš Kraus and Milan Vrábel for their help and precise and professional scientific advices. I thank also Dr Radek Pohl for measurements and interpretation of complex NMR spectra. I thank my collegue Dr. Denise-Liu' Leone for her friendly and practice advices. Among them, a big thanks goes to Dr. Veronika Sýkorová, who was always present with her ideas on difficult matters and in reading all raw drafts. I thank all my collegues for motivating me to always do better. Also, I thank Markéta Šoltysová for her help in designing specific DNA sequences to target *bs*GntR protein and for having provided the protein. I thank Dr. Ján Tarábek for his friendly approach and EPR measurements and interpretation. I also thank my special friends and collegues Bianca, Valentina and Flaminia for mutual support in every important step, laughs and all growing experiences we had together. I thank my friends Roberta and Margherita to have been present and supportive during these years of PhD. Also, I thank my partner Domenico to have always been present during whole PhD adventure and motivating me for not giving up. I thank my sister Selene for her constant support in standing all difficult moments during PhD. Also, I thank my brother Luigi to have inspired me in making creative and colourful pictures for publications and thesis. I would like to thank my mum for her lovely support and last but not least, I warmly thank my dad to have always supported me giving his wise advices to keep going and having taught me the patience to know how to wait results which was really important for whole PhD experience.

My PhD work was a part of multidisciplinary project performed in Prof. Hocek research group at the Institute of Organic Chemistry and Biochemistry AS CR (IOCB, Prague). All the synthetic work, enzymatic incorporations and conjugation studies given in this thesis were performed by me and part of them, concerning the fourth part of the thesis, in collaboration with Denise-Liu`Leone. All the measurements and interpretation of complex NMR spectra were done by Dr. Radek Pohl. All the measurements and interpretation of EPR spectra were done by Dr. Ján Tarábek. All cells experiments and confocal microscopy images were performed and acquired by Dr. Tomáš Kraus. Reactive diene (coumarin tetrazines derivatives) used in the first part of the thesis were synthetized and provided by Milan Vrabel`s group. Cell cycle progression assay was done by Erika Kužmová. Designing of specific DNA sequences to target covalently tryptophan in bsGntR protein was done

by Markéta Šoltysová from Pavlína Řezáčová's group. Mass analyses of DNA-protein covalent adducts were performed by Alan Kádek and Marek Polák from Biocev Institute, Prague.

This work was funded by the Czech Science Foundation (20-00885X) and by the Ministry of Education, Youth and Sports of the Czech Republic grant RNA for therapy (CZ.02.01.01/00/22_008/0004575).

List of publications of the autor related to the thesis

1. Spampinato A., Kužmová E., Pohl R., Sýkorová V., Kraus T., Hocek M.: **"***trans*-Cyclooctene- and Bicyclononyne-Linked Nucleotides for Click Modification of DNA with Fluorogenic Tetrazines and Live Cell Metabolic Labeling and Imaging" *Bioconjugate Chem.* **2023**, 34, 4, 772–780.

2. Spampinato A., Leone D-L., Pohl R., Ján Tarábek J., Šoltysová M., Polák M., Kádek A., Sýkorová V., Řezáčová P., and Hocek M. "ABNOH-Linked Nucleotides and DNA for Bioconjugation and Cross-linking with Tryptophan-Containing Peptides and Proteins" *Chem. Eur. J.* **2024**, e202402151.

Table of Contents

1 INTRODUCTION 15

1.1 DNA properties, structure 15

1.2 Enzymatic and chemical synthesis of natural DNA 17

1.2.1 Enzymatic synthesis 17

1.2.2 Chemical synthesis of DNA 19

1.2.3 Enzymatic synthesis of base-modified DNA 20

1.3 Synthesis of modified 2´-deoxyribonucleoside-5´-*O*-triphosphates 21

1.4 Triphosphorylation of base-modified nucleosides 22

1.4.1 Sonogashira/Heck/Suzuki cross-coupling 23

1.5 Click reactions 24

1.5.1 Introduction to bioorthogonal chemistry 24

1.5.2 CuAAC reaction - Azide-alkyne cycloaddition 25

1.5.3 SPAAC reaction - Strain-promoted azide-alkyne cycloaddition 27

1.5.4 IEDDA reaction - inverse electron demands Diels-Alder cycloaddition 28

1.5.4.1.1 Strained alkyne-tetrazines dienophiles cycloaddition 29

1.5.4.1.2 Features of diene (tetrazines) in IEDDA reaction 31

1.5.4.1.3 Features of dienophile in IEDDA reaction 33

1.6 Modified nucleotides and their delivering *in cellulo* 35

1.7 DNA-protein interactions and bioconjugation 36

1.7.1 DNA-protein cross-linking 36

1.8 Reagents for modification of serine 40

1.9 Reagents for modification of glutamic and aspartic acids 43

1.10 Reagents for modification of tryptophan 44

2 SPECIFIC AIMS OF THE WORK: 46

2.1 Rationale of the specific aims 47

3 RESULTS AND DISCUSSION 49

3.1 PART I 49

3.1.1 Design and synthesis of modified nucleotides bearing 2- or 4-linked *trans*-cyclooctene (TCO) or bicyclononyne (BCN) tethered through a shorter propargylcarbamate or longer triethyleneglycol-based spacer, their enzymatic incorporation and testing of reactivity with tetrazines 49

3.1.1.1 Synthesis of precursors **dC^{PAF3}MP, dC^{NH2}MP, dC^{PAF3}TP, dC^{NH2}TP..... 49**

3.1.1.2 Synthesis of 4-linked trans-cyclooctene (4-TCO) modified monophosphate (**dC^{4TCO}MP**)..... **50**

3.1.1.3 Synthesis of 2- or 4-linked trans-cyclooctene (TCO) or bicyclononyne (BCN) triphosphates **(dC^{2TCO}TP, dC^{4TCO}TP, dC^{p2TCO}TP, dC^{p4TCO}TP, dC^{pBCN}TP) 51**

3.1.1.4 Inverse electron demands Diels-Alder reactions (IEDDA) **54**

3.1.1.4.1 Structures of tetrazines used in this study 54

3.1.1.4.2 Model IEDDA reaction between **dC^{4TCO}MP** and 3,6-di-2-pyridyl-1,2,4,5-tetrazine (T5)……..55

3.1.1.4.3 Reaction between modified 2- or 4-linked trans-cyclooctene (TCO) or bicyclononyne (BCN) triphosphates (**dC^{2TCO}TP, dC^{4TCO}TP, dC^{p2TCO}TP, dC^{p4TCO}TP, dC^{pBCN}TP**) and tetrazines, fluorescence turn-on measurements, LC/MS characterizations and study of kinetics formation of **dC^{XTY}TP** click products 56

3.1.1.4.4 Reactions between **dC^{4TCO}TP** or **dC^{2TCO}TP** with **T1** or **T2** and between **dC^{p4TCO}TP** or **dC^{p2TCO}TP** or **dC^{pBCN}TP (dC^{X}TP)** with **T1**, **T2** or **T3** tetrazines (**dC^{XTY}TP**) by IEDDA cycloaddition..... 57

3.1.1.5 Fluorescence "turn-on" measurements and LC-MS characterizations of **dC^{XTY}TP** click products**……..... 62**

3.1.1.6 Kinetic formation of click products **dC^{XTY}TP..... 64**

3.1.1.7 Enzymatic incorporation of **dC^{2TCO}TP, dC^{4TCO}TP, dC^{p2TCO}TP, dC^{p4TCO}TP, dC^{pBCN}TP** into DNA by PEX and MALDI analysis **66**

3.1.1.8 IEDDA reactions of **19DNA_C^{4TCO} or 19DNA_C^{2TCO} or 19DNA_C^{p4TCO} or 19DNA_C^{p2TCO} or 19DNA_C^{pBCN}** with tetrazines, MALDI or LC-MS analysis of ON-click products (**19ON_C^{XTY}**)..... **68**

3.1.1.8.1 Post-synthetic labelling of **19DNA_C^{X}** (**X = 4TCO or 2TCO or p4TCO or p2TCO or pBCN**) with **T1** or **T2** tetrazines and of **19DNA_C^{X}** (**X = p4TCO or pBCN**) with **T3** or **T4** tetrazines ……68

3.1.1.9 MALDI-TOF analysis of ON-click products **(19ON_C^{XTY})** **80**

3.1.1.10 Oligonucleotide-click products (**19ON_C^{XTY}**) fluorescence turn-on measurements of **19ON_C^{p4TCO} or 19ON_C^{pBCN}** with **T1** or **T3** tetrazine **(19ON_C^{XTY})**..... **82**

3.1.1.11 Labelling of DNA using **dC^{p4TCO}TP or dC^{p2TCO}TP or dC^{pBCN}TP** derivatives delivered to U-2 OS cells by SNTT1..... **85**

3.1.1.11.1 Cell cycle progression assay..... 96

3.2 PART II..... 100

3.2.1 Design and synthesis of modified nucleotides bearing thiazetidine (β-sultam) for cross-linking with serine, their enzymatic incorporation and testing of reactivity..... 100

3.2.1.1 Synthesis of β-sultam modified nucleoside monophosphate (**dC^{SU1}MP**) and triphosphates (**dC^{SU1}TP, dC^{SU2}TP, dU^{SU1}TP, dU^{SU2}TP**)..... **100**

3.2.1.2 Copper-free cross-coupling of dC^{I}TP with SU1..... **106**

3.2.2 Click reaction on N_3- modified DNA with SU1 or SU2..... 107

3.2.3 Design and synthesis of modified nucleotides bearing 1,1-dioxido-3-oxo-4,4-dimethyl-thiazetidine (4,4-dimethyl-3-oxo-β-sultam) for cross-linking with serine, their enzymatic incorporation and testing of reactivity..... 111

3.2.3.1 Synthesis of modified monophosphates **dC^{3OXOSU}MP** and **dU^{3OXOSU}MP**..... **111**

3.2.3.2 Click reaction on DNA^{N3} using OXOSU building block..... **114**

3.3 PART III..... 115

3.3.1 Design and synthesis of modified nucleotides bearing azirine for cross-linking with glutamic and aspartic acids, enzymatic incorporation and testing of reactivity..... 115

3.3.1.1 Synthesis of azirine-modified nucleoside (**dU^{AZ}**)..... **116**

3.4 PART IV..... 120

3.4.1 Design and synthesis of modified nucleotides bearing 9-hydroxy-9-azabicyclo[3.3.1]nonane (keto-ABNOH) for cross-linking with tryptophan, enzymatic incorporation and testing of reactivity..... 120

3.4.1.1 Synthesis of keto-ABNOH modified monophosphate (**dU^{ABNOH}MP**) and triphosphate (**dU^{ABNOH}TP**)..... **121**

3.4.1.2 Investigation on the origin of signal in NMR spectra of **dU^{ABNOH}TP**..... **122**

3.4.1.3 EPR measurements of N-O radical formation on keto-ABNOH..... **124**

3.4.1.4 EPR measurements of N-O radical formation on **dU^{ABNOH}MP** and **dU^{ABNOH}TP**..... **127**

3.4.1.5 Bio-conjugation reactions..... **128**

3.4.1.5.1 Structure of peptides and amino acids used in keto-ABNO study..... 128

3.4.1.5.2 Cross-linking model reactions of 9-Azabicyclo[3,3,1]nonan-3-one-9-oxyl (keto-ABNO) moiety and Acetyl-Tryptophan (**Ac-Trp**) or Acetyl-AlaTrpAla tripeptide (**Pept1**), testing of reactivity with amino group 128
3.4.1.5.3 NMR characterization of adducts (**keto-ABNO_Trp, keto-ABNO_Pept1)** 129
3.4.1.5.4 Cross-linking model reactions of **dU^{ABNOH}MP** and Acetyl-Tryptophan (**Ac-Trp**) or Acetyl-AlaTrpAla tripeptide **(Pept1**)..... 134
3.4.1.5.5 LC-MS characterization of conjugate products **dU^{ABNO_Trp}MP, dU^{ABNO_Pept1}MP** .. 135
3.4.1.5.6 Stability studies of conjugated products **dU^{ABNO_Trp}MP, dU^{ABNO_Pept1}MP**..... 136
3.4.1.6 Incorporation of **dU^{ABNOH}TP** into DNA by PEX and MALDI analysis **138**
3.4.1.7 Model reaction of **19DNA^{ABNOH}** with Ac-Trp in presence or absence of $NaNO_2$ at different incubation temperatures **143**
3.4.1.7.1 Cross-linking reactions of **19DNA^{ABNOH}** with Trp-containing peptides and UPLC-MS characterization of adducts..... 146
3.4.1.8 Cross-linking reaction of and **31a/31bDNA^{ABNOH}** with Ac-Trp and Trp-containing peptides comparison and UPLC-MS characterization of adducts..... **149**
3.4.1.8.1 Stability studies of conjugates **31a/31bDNA^{ABNO_Pept1}** 153
3.4.1.9 Preparation of GntR from Bacillus subtilis (**bsGntR**)..... **156**
3.4.1.10 Design of DNA sequences **31aDNA^{Nat}/31aDNA^{ABNOH}** or **31bDNA^{Nat}/31bDNA^{ABNOH}** to target selectively W13 of **bsGntR** protein..... **157**
3.4.1.10.1 Reactions of **31a/31bDNA^{ABNOH}** with *bs*GntR 159
3.4.1.11 Reaction of **19DNA^{Nat}** or **19DNA^{ABNOH}** with DNA-binding proteins not containing Trp (Histones H2A and H4) and with non-DNA-binding protein containing Trp (Lysozyme) – SDS PAGE analysis **166**
3.4.1.12 Native mass analysis of **31DNA^{ABNO_bsGntR}** adduct..... **168**

4 CONCLUSIONS 170

5 EXPERIMENTAL PART 174

5.1 Design and synthesis of modified nucleotides bearing 2- or 4-linked *trans*-cyclooctene (TCO) or bicyclononyne (BCN) tethered through a shorter propargylcarbamate or longer triethyleneglycol-based spacer, their enzymatic incorporation and testing of reactivity with tetrazines..... 176

5.1.1 Synthesis of modified monophosphates (**dC^{PAF3}MP, dC^{NH2}MP, dC^{4TCO}MP**) and modified triphosphates **(dC^{4TCO}TP, dC^{2TCO}TP, dC^{p4TCO}TP, dC^{p2TCO}TP, dC^{pBCN}TP)**..... 176

5.1.2 IEDDA model reaction of **dC^{4TCO}MP** with **T5** (3,6-di-2-pyridyl)-tetrazine (16) 186
5.1.3 IEDDA click reactions of **dC^{4TCO}TP** or **dC^{2TCO}TP** with **T1** or **T2** tetrazines and of **dC^{p4TCO}TP** or **dC^{p2TCO}TP** or **dC^{pBCN}TP** with **T1**, **T2** or **T3** tetrazines (**dC^{XTY}TP**) 186
5.1.4 Preparation of nucleotide-tetrazine conjugates **dC^{XTY}TP** for LC/MS characterization 188
5.1.5 Enzymatic synthesis of TCO-, BCN-modified DNA by PEX 188
5.1.6 DB Streptavidin magneto-separation procedure: 190
5.1.7 Reaction of **19DNA_C^{X}** (**X = 4TCO or 2TCO or p4TCO or p2TCO or pBCN**) with tetrazines 190
5.1.8 Formation of tetrazine-ONs click products (**19ON_C^{XTY}**) for MALDI-TOF analysis. 191
5.1.9 Oligonucleotide fluorescence turn-on measurement of **19ON_C^{p4TCO}** or **19ON_C^{pBCN}** with **T1** or **T3** tetrazine (**19ON_C^{XTY}**) 192
5.1.10 Labelling of DNA using **dC^{p4TCO} or dC^{p2TCO}TP or dC^{pBCN}TP** derivatives delivered to U-2 OS cells by SNTT1 193
5.1.10.1 Protocol for the cell cycle progression experiment **196**

5.2 Design and synthesis of modified nucleotides bearing thiazetidine (β-sultam) for cross-linking with serine, their enzymatic incorporation and testing of reactivity 197
5.2.1 Synthesis of modified monophosphates (**dC^{SU1}MP, dC^{SU2}MP, dU^{SU2}MP**) and triphosphates (**dC^{SU1}TP, dC^{SU2}TP**) 198

5.3 Design and synthesis of modified nucleotides bearing 1,1-dioxido-3-oxo-4,4-dimethyl-thiazetidine (4,4-dimethyl-3-oxo-β-sultam) for cross-linking with serine, their enzymatic incorporation and testing of reactivity 204
5.3.1 Synthesis of modified monophosphate **(dC^{OXOSU}MP, dU^{OXOSU}MP)** 205

5.4 Design and synthesis of modified nucleotides bearing azirine for cross-linking with glutamic and aspartic acids, enzymatic incorporation and testing of reactivity 207
5.4.1 Synthesis of modified dU^{AZ} nucleoside 208

5.5 Design and synthesis of modified nucleotides bearing 9-hydroxy-9-azabicyclo[3.3.1]nonane (keto-ABNOH) for cross-linking with tryptophan, enzymatic incorporation and testing of reactivity 210
5.5.1 Synthesis of modified monophosphate (**dU^{ABNOH}MP**) and triphosphate (**dU^{ABNOH}TP**). 213
5.5.2 Model reactions 216
5.5.3 Enzymatic synthesis of ABNO- modified DNA 219
5.5.4 Bio-conjugation reactions 221

6 LIST OF ABBREVIATIONS AND SYMBOLS 228

7 APPENDIX 231

7.1 Fluorescence and LC-MS spectra 231

7.2 Cell imaging 247

7.3 NMR spectra 248

8 REFERENCES 252

1 INTRODUCTION

1.1 DNA properties, structure

All known living organisms inherit a molecule that carries most of the genetic instructions used in the development, functioning, and reproduction, the DNA or deoxyribonucleic acid. It is made up of repeating units called nucleotides (deoxyribonucleoside triphosphates, dNTPs) which are linked together by phosphodiester bonds, building the backbone of DNA (**Figure 1**). Each nucleotide consists of a sugar molecule (deoxyribose in DNA), a phosphate group, and one of four nitrogenous bases: the purines (adenine [A] and guanine [G]), and the pyrimidines (cytosine [C], thymine [T]). The structure of DNA is described as a double helix, where two strands of nucleotides are oriented antiparallel to each other. The two DNA strands are held together in double helix structure by specific so-called Watson-Crick base pairing (A with T and C with G) through hydrogen bonds (**Figure 1**). The double helix is stabilized by hydrophobic interactions between hydrophobic surfaces of adjacent base pairs and by Van der Vaals attractive forces that exist between atoms and molecules in close proximity. These interactions contribute to the structural stability of the DNA molecule. The stability of the DNA double helix is crucial to protect the genetic information and ensuring accurate transmission of the instructions necessary for the growth, development, functioning, and reproduction of living organisms from one generation to the next one. X-Ray diffraction studies showed that double helical DNA can occur in three different conformations depending on environmental conditions and interactions with other molecules (**Figure 2**). **B-form** DNA is the right-handed double helix structure which was described by James Watson and Francis Crick and the biologically relevant form under normal physiological conditions. Under dehydration condition, in presence of certain salts or in complex with proteins **A-form** is often observed, which is a right-handed helical structure with a shorter and wider helix compared to B-DNA. Less common in biological contexts is **Z-form** DNA which is involved in a role of gene regulation and DNA repair. It was proved to be a left-handed double helix with a zig-zag pattern, which can be formed under high salt concentrations or in sequences with alternating purine-pyrimidine repeats.[1] The DNA molecule can adopt also non-canonical structures, including three-stranded (*i.e.* triplex) structures, which is a third strand binding to the major groove of a DNA double helix through Hoogsteen base pairing. An example of triplex is the **H-DNA form**, naturally occurring, which is involved in genetic recombination during exchange

of genetic material between the two DNA strands in the homologous recombination process. Triplex DNA has been studied for its potential applications in gene targeting and regulation.[2] Certain DNA sequences, particularly those rich in guanine (G) bases, can form **G-quadruplex** structures, where four guanine bases interact through Hoogsteen hydrogen bonding. G-quadruplex structures are implicated in various cellular processes, including telomere maintenance and gene regulation.[3] Different conformations of DNA determine different functions, but primarily, one of the more important function of DNA, in physiological conditions, is to be used as a template for the mRNA synthesis. In summary, the sequence of DNA nucleotides is the genetic code that determines, by translation, the amino acid sequence of proteins and regulation of all cellular processes.

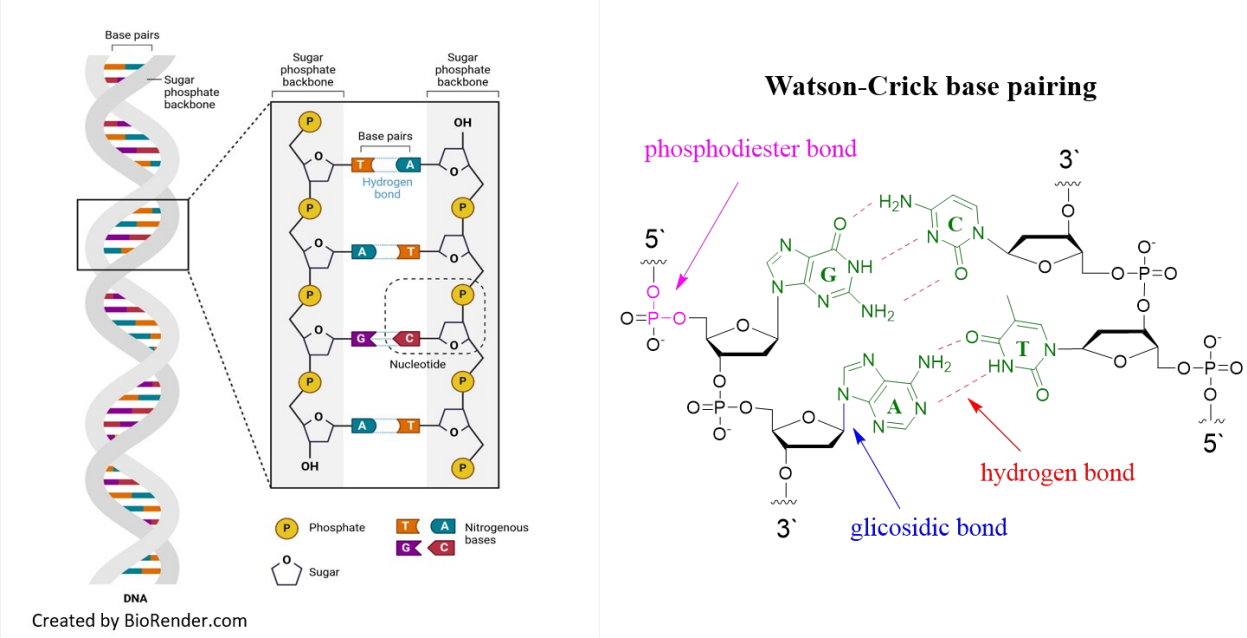


Figure 1. Structure of DNA and Watson-Crick base pairing of canonical nucleotides via hydrogen bonds.

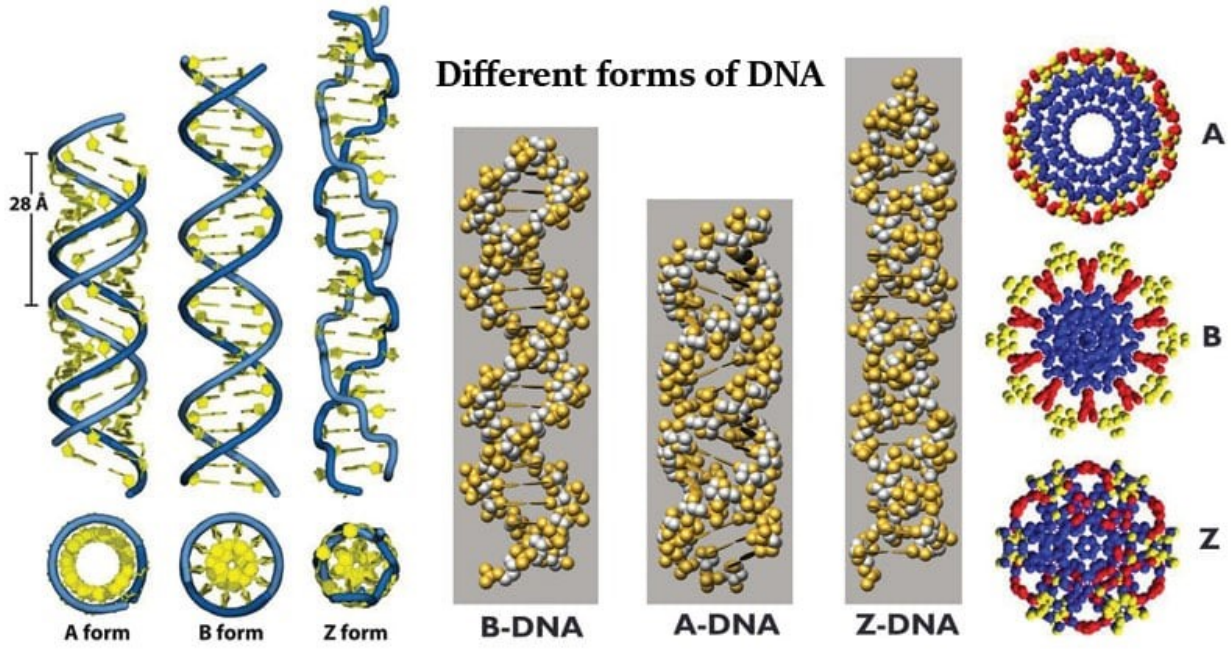


Figure 2. Different forms of DNA, a) A form, b) B form, c) Z form.[4]

1.2 Enzymatic and chemical synthesis of natural DNA

1.2.1 Enzymatic synthesis

Enzymatic synthesis of DNA involves the use of enzymes to build a complementary strand of DNA based on a template strand. The key enzyme involved in this process is DNA polymerase. DNA polymerase catalyzes the addition of dNTPs to the 3`-end of the growing DNA chain, extending the new DNA strand in the 5' to 3' direction. This enzymatic reaction to produce dsDNA at 5'- end, using a short sequence (15-25 nt), is called PEX or primer extension[5] (**Figure 3**). The primer is usually labelled with ^{32}P or with fluorophores, such as derivative of fluorescein or cyanine (FAM, Cy5, Cy3) The PEX is initiated when DNA polymerase recognizes the primer as a starting point to begin the addition of complementary deoxyribonucleotides to the 3' end of the primer, catalyzing the formation of phosphodiester bonds between the nucleotides. The synthesis is continued until the end of the template strand is reached by DNA polymerase.

The proper selection of DNA polymerase can lead to the incorporation of just one or multiple modified dNTPs[6].

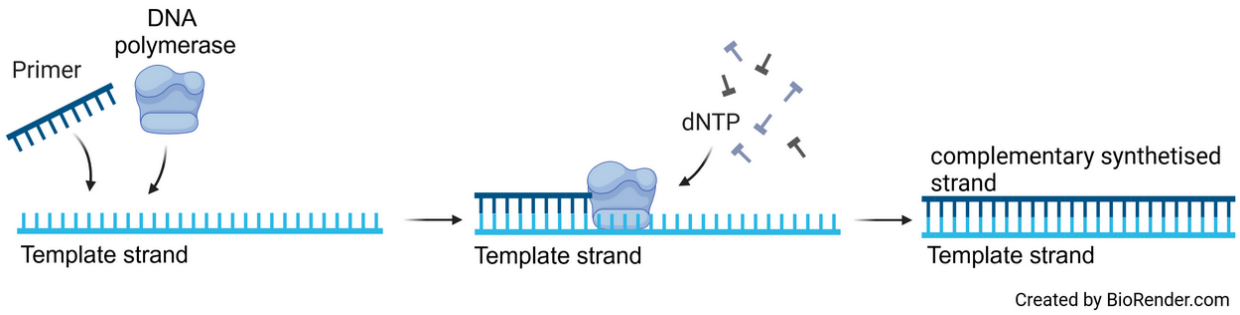


Figure 3. Primer extension (PEX) reaction scheme

Enzymatic DNA synthesis also plays a significant role in various molecular biology techniques, such as PCR (Polymerase Chain Reaction), which allows the amplification of specific DNA sequences. PCR was invented by Kary Mullis[7] in 1983 and has revolutionized molecular biology by enabling researchers to generate large amounts of specific DNA sequences from small amounts of starting material. The PCR process involves the use of a DNA template, DNA primers, a DNA polymerase, and nucleotide building blocks to generate multiple copies of an interested DNA segment. The reaction takes place in a thermal cycler, which controls the temperature of the reaction mixture through a series of cycles. The three main steps of a PCR cycle are denaturation, annealing, and extension (**Figure 4**). The reaction begins by heating the DNA template to a high temperature (94-98°C) which breaks the hydrogen bonds between complementary bases separating the DNA into two single strands. The reaction temperature is lowered (to 50-65°C), allowing short DNA primers to anneal to the complementary sequences on each single-stranded DNA template. The primers provide a starting point for the DNA polymerase to synthesize new DNA strands. Subsequently, the temperature is raised again (usually to 72°C), and DNA polymerase synthesizes a new DNA strand by adding nucleotides to the 3' end of each primer. This step extends the DNA sequence and produces a copy of the targeted segment and is called extension. These three steps constitute one PCR cycle, but usually the process is repeated for 20 to 40 cycles, leading to the exponential amplification of the targeted DNA sequence. Each cycle doubles the amount of DNA, resulting in a substantial increase in the number of copies.

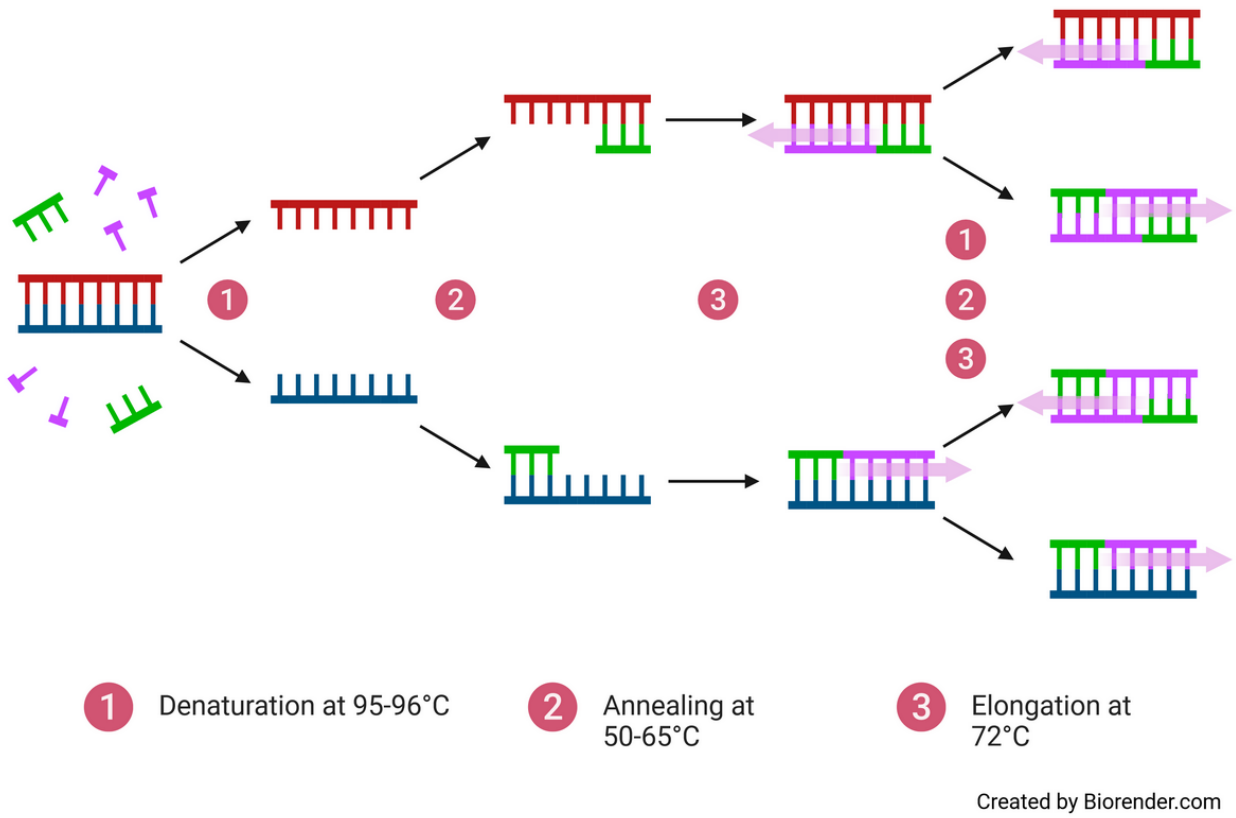


Figure 4. Polymerase chain reaction (PCR) scheme

1.2.2 Chemical synthesis of DNA

DNA can be prepared not only enzymatically but also chemically and the most used method is the solid-phase synthesis. It involves a step-by-step assembly of nucleotides to create customized DNA sequences. The main advantages of this approach is the efficient purification and manipulation of DNA sequences. The most common solid-phase DNA synthesis is *via* phosphoramidite chemistry (**Figure 5**), involving the sequential addition of nucleotide building blocks onto a growing DNA chain. The synthesis starts from the first nucleotide building block, protected at its reactive groups, which is covalently attached to the solid support. Afterwards, by removal of the protected group and, sequentially, by addition of other activated and protected nucleotide phosphorammidites, the DNA chain is growing via phosphodiester bond formation in direction 3' to 5'. Every step of new added nucleotide is called coupling reaction involving the activation of the nucleotide at 5' hydroxyl group, via tetrazole activation, and subsequent reaction with the 3' hydroxyl group of the growing chain. After each coupling step, unreacted hydroxyl groups on the solid support are capped to prevent nonspecific reactions in subsequent cycles. The

phosphite linkages, formed after coupling, is oxidized into phosphodiester bonds ensuring a stable linkage between nucleotides. As final step, the desired DNA sequence, by cleavage from the solid support and by deprotection of any remaining protecting groups, is released.

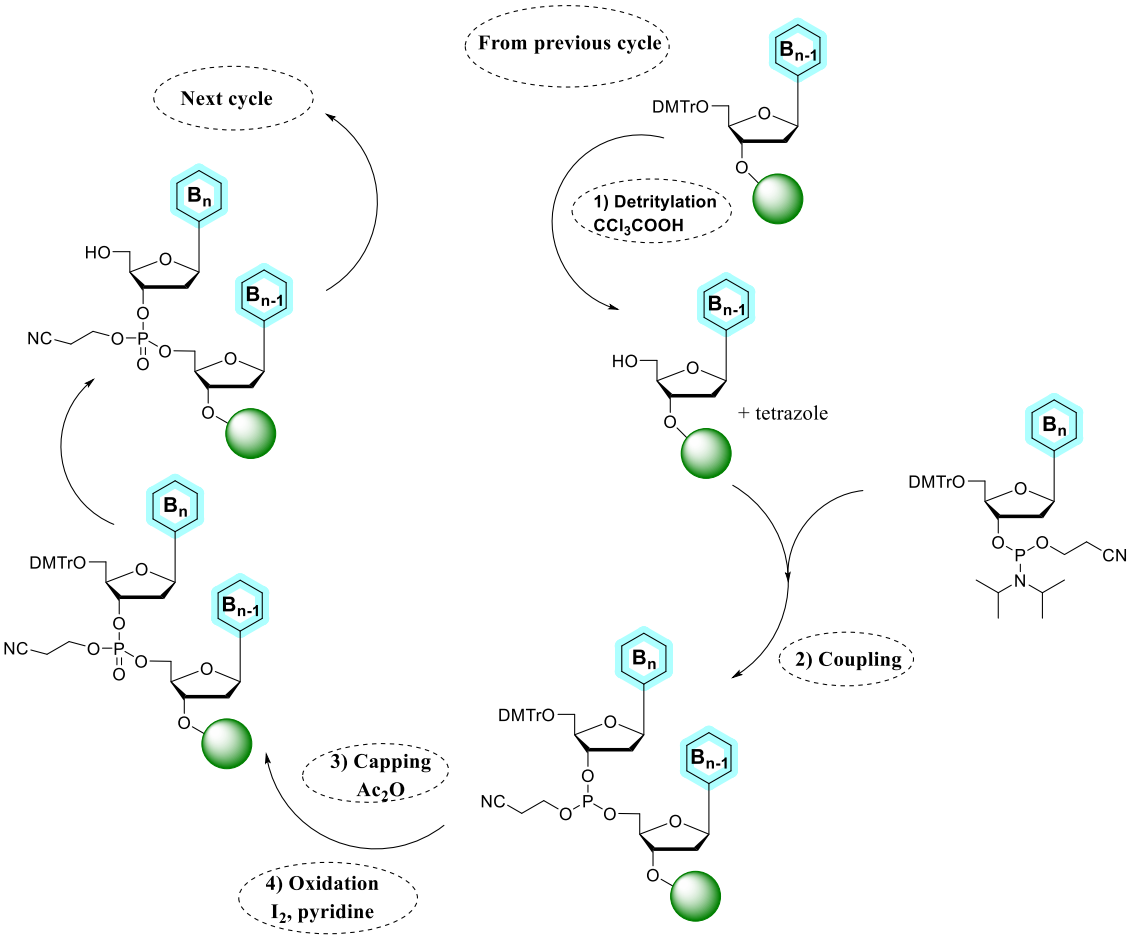


Figure 5. DNA solid phase chemical synthesis *via* Phosphoramidite.

1.2.3 Enzymatic synthesis of base-modified DNA

2'-Deoxyribonucleoside triphosphates are critical components in DNA replication and other DNA synthesis processes, providing the necessary substrates for the construction of new DNA strands. Modified nucleotides are nucleotide analogs with alterations to their structure, typically involving changes to the base, sugar, or phosphate components. These modified nucleotides can be used as building blocks for the modification of nucleic acids (DNA and RNA) in various applications.

Introduction of chemical modification at C-5 position of pyrimidine or C-7 position of 7-deazapurines were observed to be well tolerated by polymerase (**Figure 6**).[8]

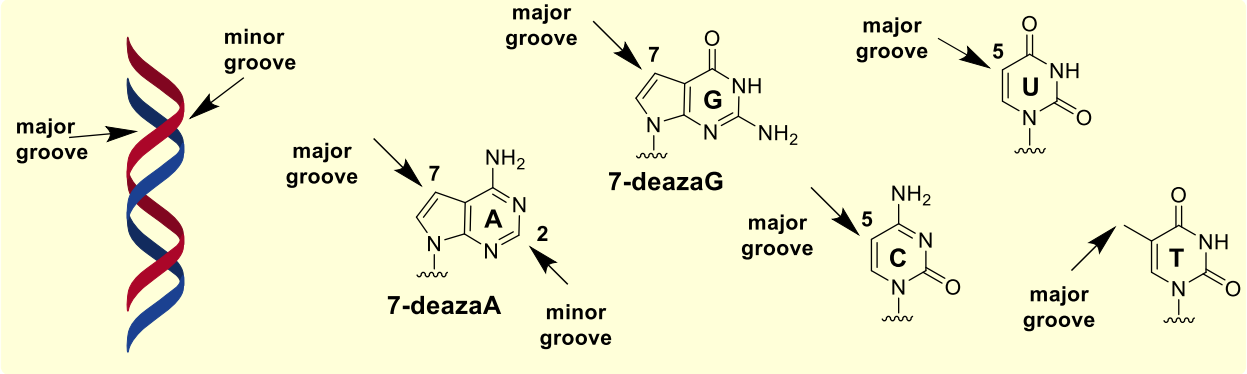


Figure 6. DNA modifications tolerated in minor (position 2) and major groove (positions 5 and 7).

1.3 Synthesis of modified 2´-deoxyribonucleoside-5´-*O*-triphosphates

Different kind of modification can be inserted, from small groups such as alkyl, ethynyl[9] or steric hindered groups such as phenyl[9] (**Figure 7**).

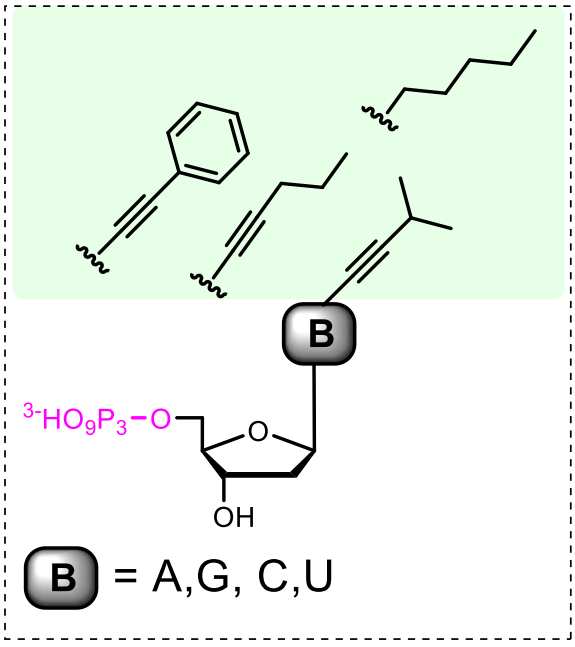


Figure 7. Examples of modified triphosphates with different groups.

Modified dNTPs can be synthesized, chemically, following two different pathways: the first one consists in the insertion of modification by Sonogashira/Heck/Suzuki cross coupling or click reaction on nucleoside followed by its triphosphorylation, the second one is the reverse way, triphosphorylation on prior functionalized nucleosides including the protection step if the reactive group is a nucleophile and then insertion of the modification (**Figure 8**).

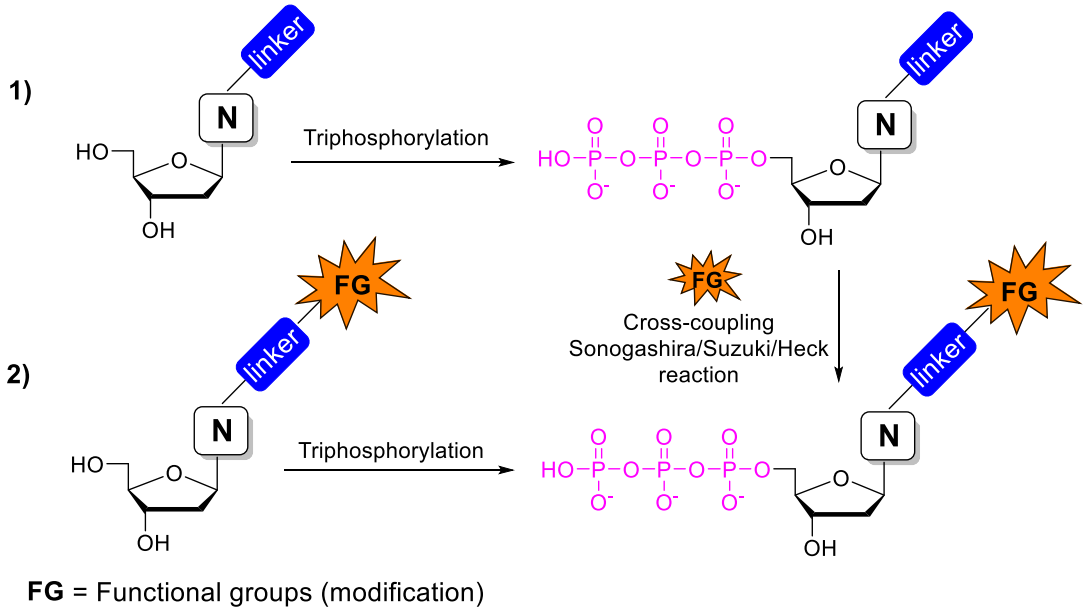


Figure 8. Different ways to synthesize modified triphosphate

1.4 Triphosphorylation of base-modified nucleosides

Nucleotides (phosphates esters of nucleosides) are prepared by chemical or enzymatic phosphorylation of nucleosides. Enzymatic phosphorylation can be performed by nucleoside (or nucleotide) kinases and proceeds chemoselectively[10]. Chemical phosphorylations are usually more efficient and convenient[11].

The primary OH (at position 5' of 2'-deoxyribose) is usually more reactive compared to the secondary one at position 2` or 3`, therefore, there is no need of protection.

The most efficient current method consists in two steps which was described by Yoshikawa[12], in the first step the reaction proceeds in presence of $POCl_3$ as phosphorylating agent in trimethyl phosphate as solvent. In the second step the 5'-dichlorophosphate is treated with pyrophosphate in

presence of tributylamine in DMF to form a cyclic triphosphate intermediate that is hydrolyzed to linear triphosphate by aqueous triethylammonium bicarbonate (**Figure 9**).

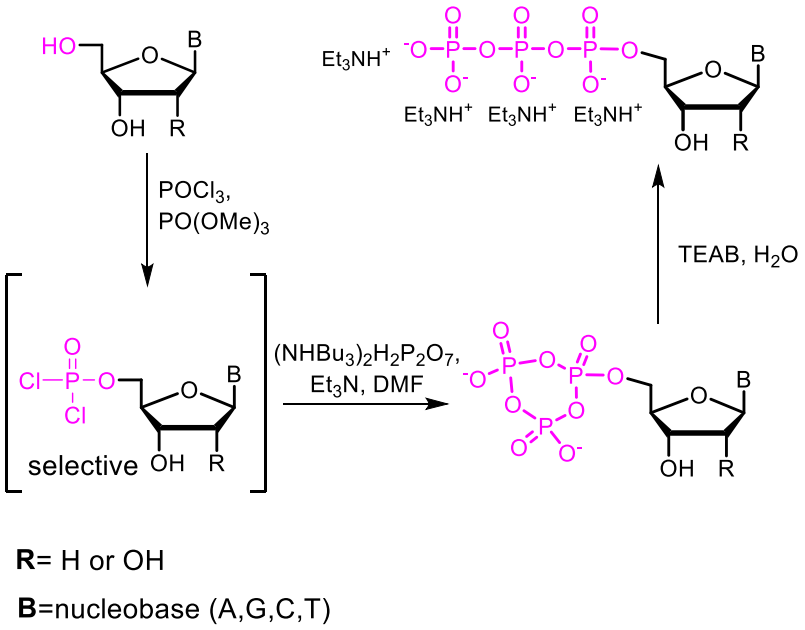


Figure 9. Yoshikawa synthesis of 2`-deoxyribonucleoside 5`-*O*-triphosphate.[12]

1.4.1 Sonogashira/Heck/Suzuki cross-coupling

In some cases, the modification is unstable under triphosphorylation conditions, therefore they are inserted directly on nucleoside triphosphate. The two step procedure was developed and optimized in Hocek group[13,14], which consists in performing on iodo-functionalised nucleoside triphosphates Sonogashira coupling with alkyne-modified moiety, with boronic acid- or ester-modified moiety[15], in case of Suzuki-Miyaura, or with alkene-modified moiety for Heck coupling (**Figure 10**).[16] The reaction is usually performed in mild condition, using water/acetonitrile as solvent in presence of palladium as catalyst and water-soluble co-catalyst (e.g TPPTS). Short time usually is required for the reaction, otherwise the triphosphate could be hydrolysed to diphosphate. Also presence of base is necessary, K_2CO_3 for Suzuki and trimethylamine for Sonogashira coupling. The coupled nucleoside triphosphates are, subsequently purified by reverse-phase HPLC in order to remove hydrolytic side-products.

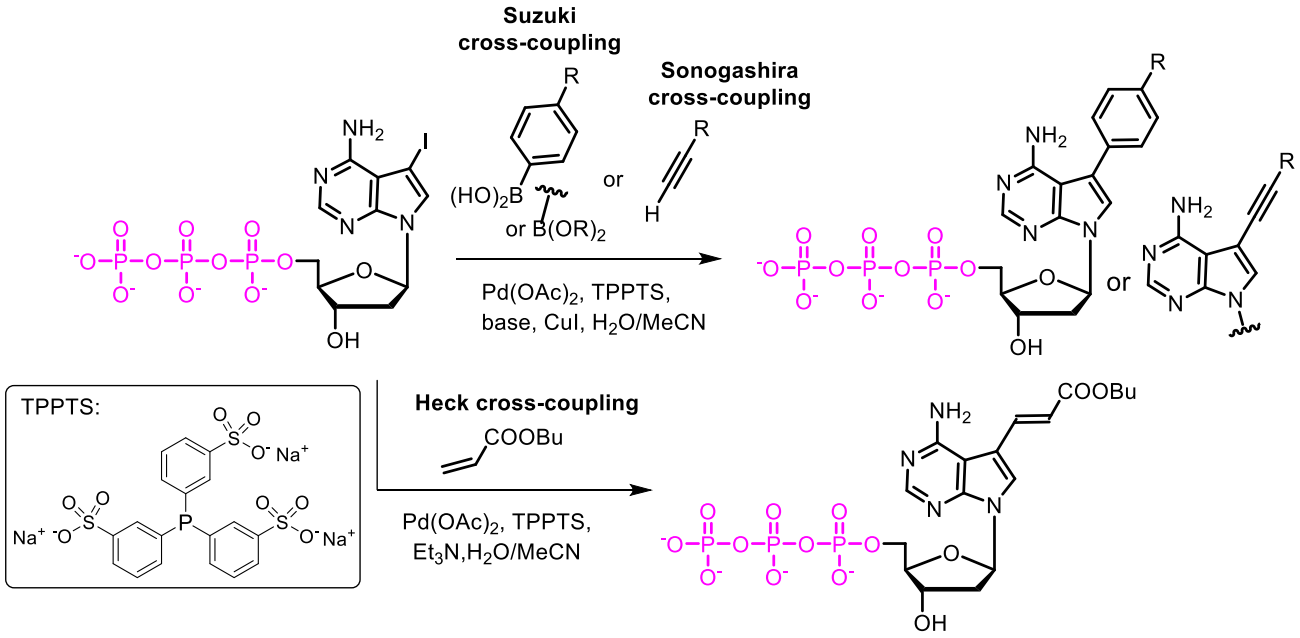


Figure 10. Examples of functionalized nucleoside triphosphates by cross-coupling reactions.

1.5 Click reactions

1.5.1 Introduction to bioorthogonal chemistry

The term bioorthogonal refers to chemical transformations that can occur selectively in biological systems without interfering with native biochemical processes, enabling the visualization, tracking, and manipulation of specific biomolecules in complex biological environments (bio-imaging). For instance, they can be tools to insert functionalities (affinity probes, fluorophores, reactive tags, post-translation modification etc.) in proteins without any significant alteration of their native functions.[17] Ideally, requirements of such bioorthogonal reactions are 1) the selectivity, meaning they should react only with specific functional groups over other potential reactive ones on biomolecules, 2) solvents, the biorthogonal reactions should be proceeded in aqueous media, 3) the pH, which should be close to physiological value and 4) fast kinetics at room temperature (or up to 37 °C) with low concentration of reactant used. However, it is difficult to fulfill simultaneously all these requirements, and the specific labelling of biomolecules using chemical reactions is still a very challenging task[18]. For instance, in 1,3-dipolar cycloaddition between azides and alkynes (CuAAC) is involved the use of copper which is toxic for cells[19,20]. Therefore,

to avoid the use of copper, the strain-promoted azide-alkyne cycloaddition (SPAAC) was developed[21], but in this case the specificity of the reaction is not guaranteed due to other side reactions on strained alkyne with nucleophiles present in the biological environment. Among all biorthogonal reactions, the inverse electron demands Diels-Alder cycloaddition (IEDDA) is the one which meets most of the criteria of this class of reaction in terms of selectivity, biocompatibility and no need of copper (**Figure 11**).

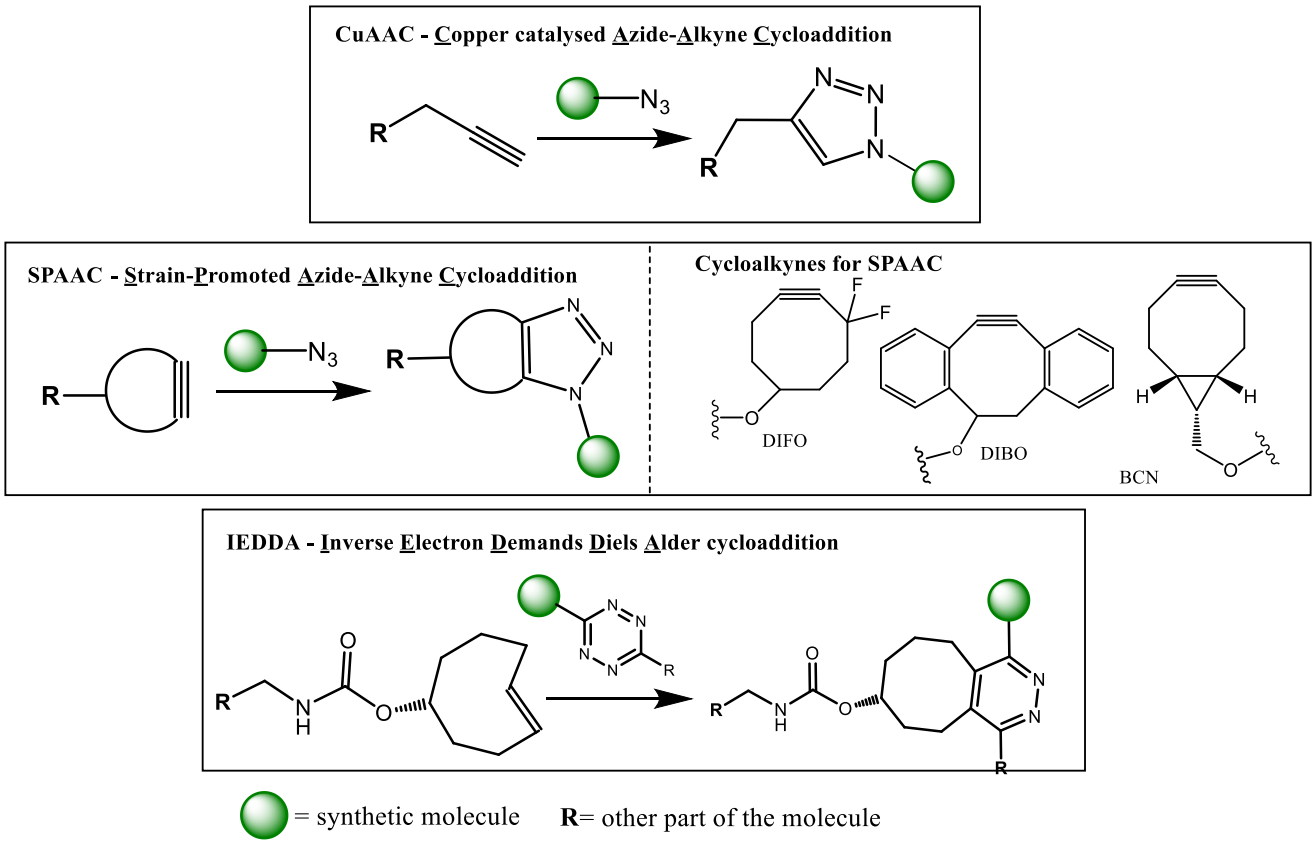


Figure 11. Representative examples of bioorthogonal reactions.

1.5.2 CuAAC reaction - Azide-alkyne cycloaddition

CuAAC stands for **C**opper catalyzed **A**zide-**A**lkyne **C**ycloaddition, often referred to as the "click chemistry" reaction, which is a 1,3-dipolar cycloaddition between an azido group (-N_3) and an alkyne group (-C≡C-) to form a triazole ring in the presence of a copper catalyst, Cu(I) which facilitates the formation of a copper acetylide intermediate from the alkyne accelerating the reaction rate, and subsequently, acetylide reacts with the azide to form the triazole product.

The reaction is usually performed under room temperature in aqueous or organic solvents. Advantages of CuAAC, like excellent regioselectivity, good selectivity which minimizes the formation of side-products, high efficiency, makes this reaction widely used in various fields of chemistry, especially in bioconjugation and chemical biology. In 2022, the Nobel Prize in Chemistry was awarded to Carolyn R. Bertozzi, Morten P. Meldal and K. Barry Sharpless, "for the development of click chemistry and bioorthogonal chemistry" [22]. Another type of azide -alkyne cycloaddition uses ruthenium as catalyst was developed by Fokin and co-workers, which is also regioselective but for the 1,5-disubstituted 1,2,3-triazoles[23]. In the context of CuAAC, functionalization of nucleosides or nucleotides with azide or alkyne group[24,25] is useful for different aims[26], such as the use of alkyne-functionalized DNA as click handle to detect it with fluorescent azide[27] (**Figure 12**).

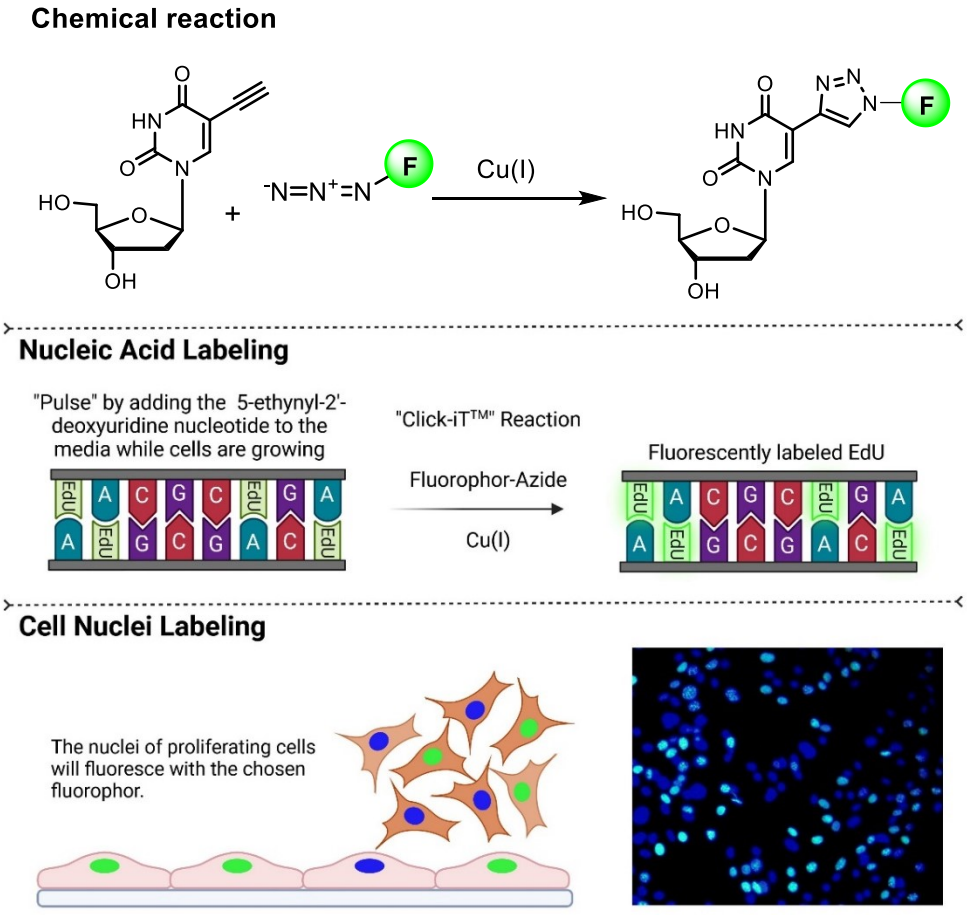


Figure 12. Use of alkyne functionalized DNA as click handle with fluorescent azide for bioimaging application

In particular, the introduction of azido group can be achieved through various synthetic methods such as nucleophilic substitution, diazotization, or azide displacement reactions. On the other hand, the alkyne can be introduced by acylation or Sonogashira cross-coupling[28] reactions.

1.5.3 SPAAC reaction - Strain-promoted azide-alkyne cycloaddition

There are more biocompatible click reactions which avoid the use of transition metals as catalysts, like the **S**train-**P**romoted **A**zide-**A**lkyne **C**ycloaddition (SPAAC). Strained alkynes or alkenes (cyclooctenes, bicyclononynes) and azide are involved in the reaction giving a fast kinetic comparable to the Cu-catalysed click reactions. However, SPAAC reactions are not regiospecific and give mixture of 1:1 triazole regioisomers. Among SPAAC reactions class there is the use of azides as chemical reporters explored by the Staudinger ligation with phosphines[29]. In contrast with CuAAC and SPAAC reactions where azides produce a stable triazole linkage, the Staudinger ligation generates a stable amide bond. Afterwards, Raines *et al.*[30] and Bertozzi *et al.*[31] simultaneously have reported the so-called traceless Staudinger ligation (**Figure 13**). Although this click reaction has sufficient biocompatibility to be performed in living systems, the main disadvantage involving the phosphine oxidation is the very slow kinetics.[32] Other chemical transformations are included in biorthogonal reactions such as aldehyde/ketone-amine ligations[33], photoclick 1,3-dipolar cycloadditions between tetrazoles and substituted alkenes[34,35] and, strain-promoted cycloadditions involving nitrones and alkynes[36] (**SPANC**=**S**train **P**romoted **A**lkyne **N**itrone **C**yloaddition), **Figure 13**. These bioorthogonal reactions, however, have been used mainly for *in vitro* bioconjugation because of the limited reaction conditions, for example the need of an acidic pH or a light source in the case of ketones/aldehydes and tetrazoles, respectively.

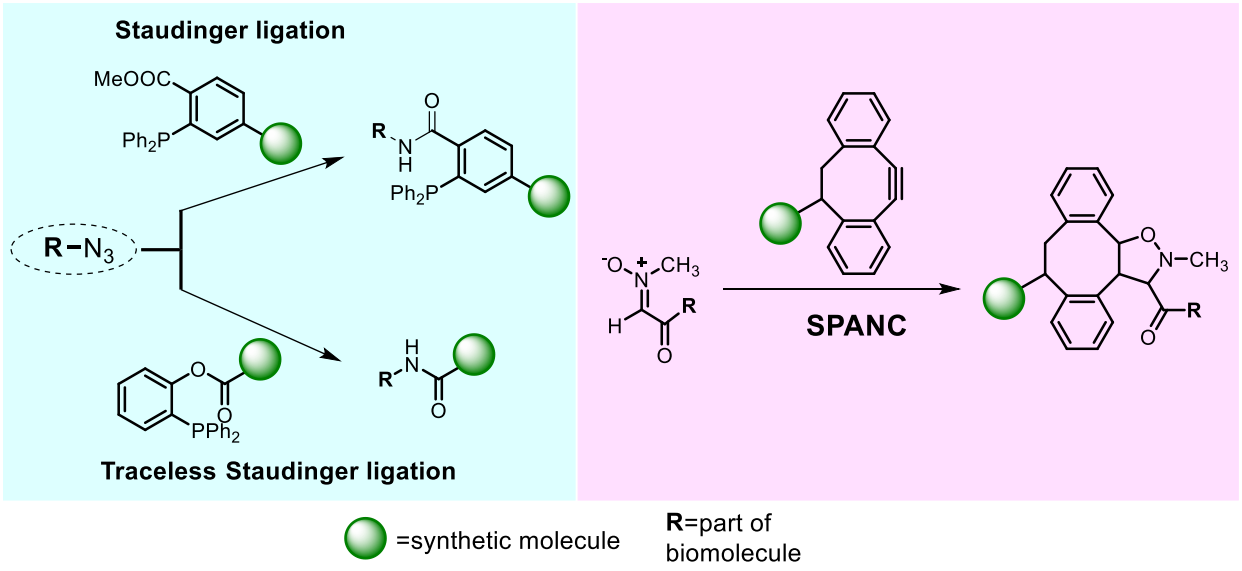


Figure 13. Schematic representation of traceless or canonical Staudinger ligations (left in blue) and SPANC=Strain Promoted Alkyne Nitrone Cycloaddition (right in pink).

Finally, SPAAC reactions can be included in a selective and efficient bioconjugation and also bioorthogonal method that enable the rapid and specific attachment of molecules to biomolecules.[37]

1.5.4 IEDDA reaction - inverse electron demands Diels-Alder cycloaddition

Even more biocompatible reaction conditions, which have been revolutionized the biorthogonal chemistry were achieved with **I**nverse-**E**lectron **D**emand **D**iels-**A**lder Cycloaddition reaction (IEDDA). The typical electron demand Diels Alder reaction is a concerted [4π e^- + 2π e^-] cycloaddition in which an electron-rich diene (the 4π e^- component) combines with an electron-deficient dienophile (the 2π e- component) to produce a 6-membered ring. In particular, two new C-C sigma bonds are formed by overlapping of π molecular orbitals of diene with π molecular orbital of dienophile. From the frontier molecular orbitals (FMO) and in terms of energy, the electron rich diene molecular orbitals are higher than the orbitals of the electron poor dienophile, the HOMO of the diene ($HOMO_{diene}$) and the LUMO of the dienophile ($LUMO_{dienophile}$) are more similar in energy than the $HOMO_{dienophile}$ and the $LUMO_{diene}$. Due to that, the strongest orbital interaction results from the combination $HOMO_{diene}$ and $LUMO_{dienophile}$ (**Figure 14-a**).

In other words, the constructive overlap is possible between the highest-occupied molecular orbital (HOMO) of the diene with the lowest-unoccupied molecular orbital (LUMO) of the dienophile**.** The reaction is very useful because is stereospecific which means the stereochemistry of both starting materials is conserved in the product. In contrast IEDDA reaction is a concerted [4π e^- + 2π e^-] cycloaddition in which an electron-poor diene (the 4π e^- component) combines with an electron-rich dienophile (the 2π e- component) to produce a 6-membered ring with formation of new C-C sigma bonds by overlapping of π molecular orbitals of diene with π molecular orbital of dienophile. From the frontier molecular orbitals (FMO) and in terms of energy, the electron poor diene molecular orbitals are higher than the orbitals of the electron poor dienophile, the LUMO of the diene ($LUMO_{diene}$) and the HOMO of the dienophile ($HOMO_{dienophile}$) are more similar in energy than the $LUMO_{dienophile}$ and the $HOMO_{diene}$. Due to that, the strongest orbital interaction results from the combination $LUMO_{diene}$ and $HOMO_{dienophile}$ (**Figure 14-b**). The IEDDA reaction is important in organic chemistry thanks to its main beneficial aspects like biocompatible mild reaction conditions, versatility, high selectivity, its applications in bioorthogonal chemistry and fast kinetics in labelling biomolecules in living systems without interfering with biological processes.

1.5.4.1.1 Strained alkyne-tetrazines dienophiles cycloaddition

The main characteristic of the IEDDA reaction is the use of strained cyclic alkynes or alkenes as dienophiles such as *trans*-cyclooctene or bicyclononyne which can be inserted also on dNTPs or DNA[38]. These strained compounds have high energy and are very reactive. The other component involved in the reaction is a diene which contains electron-rich groups like azides, tetrazines, or enamines. The ability of tetrazines to react with unsaturated compounds was first revealed in 1959 by Lindsey *et al.*[39] The reaction proceeds *via* the 1,4-addition of the –C=N–N=C– diene system of the tetrazine to an appropriate alkene, yielding a highly strained bicyclic intermediate (**Figure 14-c).** The adduct undergoes a retro-Diels–Alder reaction and by elimination of N_2 is formed the corresponding 4,5-dihydropyridazine, which either isomerizes to the corresponding 1,4-dihydro-isomers or is oxidized to give a pyridazine product. According to the FMO, the IEDDA reaction kinetics is governed by the energy gap between the corresponding HOMO and LUMO of the reactants. Smaller is the $HOMO_{dienophile}$–$LUMO_{diene}$ energy difference faster is the IEDDA reaction. Electron-donating groups (EDG) raise both HOMO and LUMO energy of the reactants,

whereas electron-withdrawing groups (EWG) cause the opposite. Therefore, by electronically tuning the two species reacting together it is possible to manipulate the $LUMO_{diene}$–$HOMO_{dienophile}$ energy gap and enhance their reactivity (**Figure 14-b**).

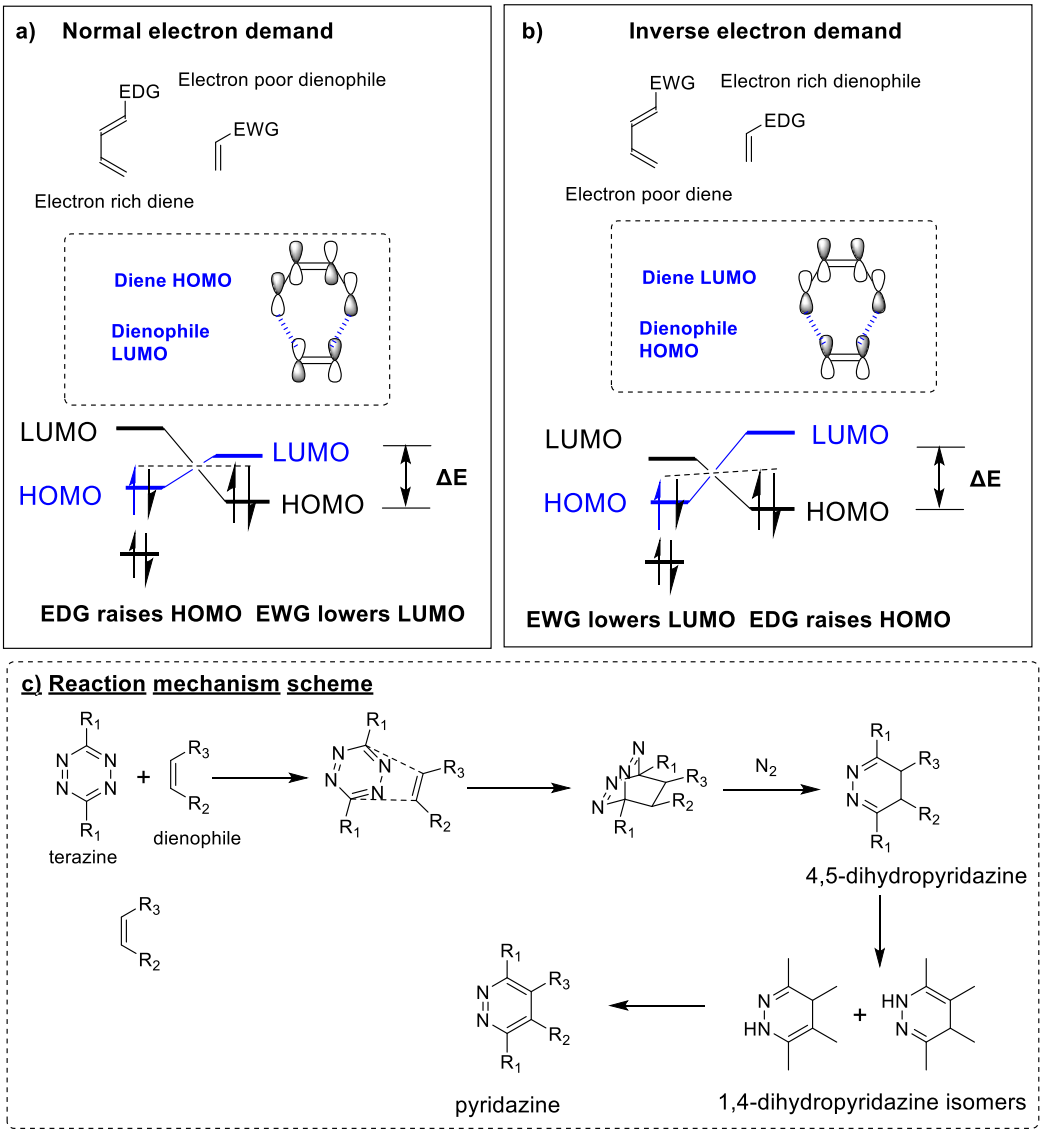


Figure 14. (a) Frontier molecular orbital (FMO) of normal and of (b) inverse electron demand Diels–Alder reaction. EDG = electron-donating group, EWG = electron-withdrawing group. (c) Reaction mechanism scheme between a dienophile and a tetrazine[40].

1.5.4.1.2 Features of diene (tetrazines) in IEDDA reaction

Boger *et al*[41] have shown that tetrazines attached to carboxylate EWGs present faster kinetics than tetrazines bearing methoxy groups due to lower $LUMO_{diene}$ energy. Similarly, EWGs substituents groups such as aromatic heterocycles at C3 and C6 decreases the electron density of the tetrazine ring increases the reaction rate against dienophiles[42] (**Figure 15-a**). Additionally, the electronic effect of the number of N-heteroatoms in the heterocyclic ring was also studied by Sauer *et al*[43], (**Figure 15-b**). Based on calculations of rate constants (k), it was shown that tetrazines reacts significantly faster than, for example, triazine, according to this order tetrazine > triazine > pyridazine > pyridine.[44] Boger *et al*[41] also showed that the electronic properties of the substituent not only influence tetrazine reactivity but also affect the regioselectivity of the cycloadditions (**Figure 15-c**). Accordingly, the authors demonstrated experimentally and theoretically that it is possible to predict the product formation by analysing the partial net charge on C3 and C6 positions.[41,45]

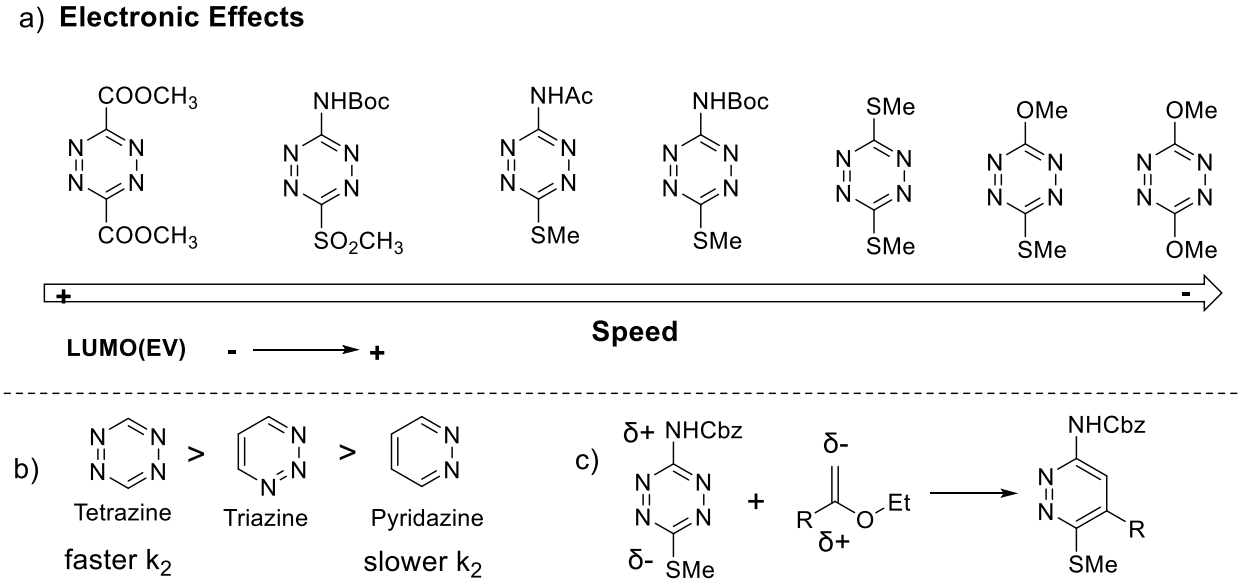


Figure 15. (a) Electronic effects of EDGs substituents on LUMO energy of tetrazine. (b) Reactivity of heterocyclic ring bearing various numbers of nitrogen atoms. (c) Electronic effects on regioselectivity.

A really important beneficial aspect of using tetrazines as coupling partners is the increase of fluorescence observed as a consequent effect of a bioorthogonal reaction, resulting in a "turn-on

system" useful for example for cell imaging applications. Fluorescence is a phenomenon where a substance absorbs light at a particular wavelength and then emits light at a longer wavelength. Fluorogenic compounds typically do not produce significant fluorescence until they undergo a specific chemical reaction or environmental change that triggers their fluorescence. Recently, Vrabel and co-workers have developed an important class of coumarin–tetrazine[46] probes with exceptional turn-on ratios when reacted with axial *trans*-cyclooctene (TCO) and bicyclononyne (BCN) dienophiles (**Figure 16**). By screening a series of tetrazines, the photochemical properties of the resulting 1,4-dihydropyridazines could be tuned towards 91-fold fluorescence intensity increase and emissions ranging from 480 to 605 nm. Their design was performed in the way to aligned transition dipoles of the coumarin fluorophore and the tetrazine in order to maximize contributions from FRET (Fluorescence Resonance Energy Transfer) in addition to TBET (Through Bond Energy Transfer). FRET phenomenon is based on energy transfer from a donor, when it is excited, to an acceptor molecule which emits light at its characteristic wavelength. The donor molecule has to emit light in a wavelength range that overlaps with the absorption spectrum of the acceptor molecule. TBET typically occurs when two fluorophores are covalently linked together within a molecule or a molecular complex. The energy transfer in TBET is mediated by the electronic coupling between the molecular orbitals of the fluorophores, which allows the transfer of excitation energy through σ and π bonds. As last, the fast reaction kinetics and low background signal make fluorogenic tetrazines probes ideal for live-cell imaging where it is possible to track and visualize cellular processes like nascent DNA synthesis, in real-time, with no washing steps to remove excess of unreacted probes.

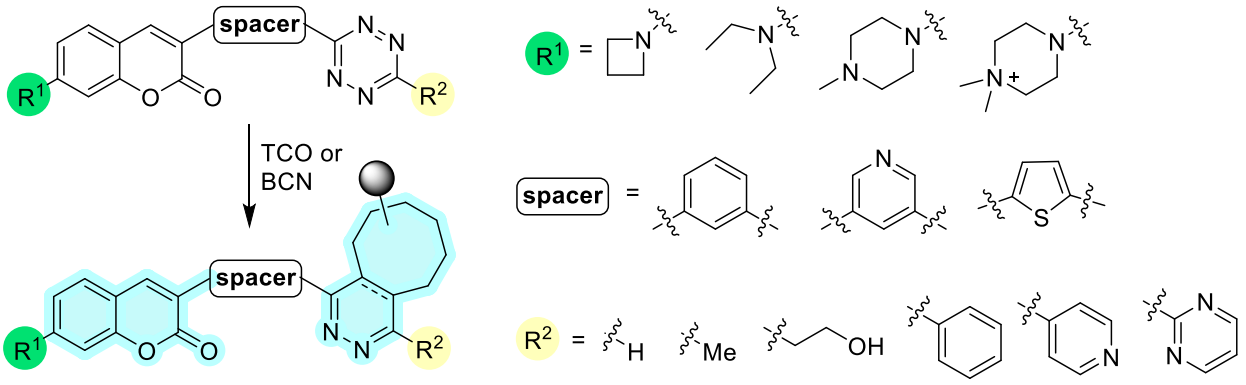


Figure 16. General structure of coumarin-tetrazine probes and schematic presentation of their fluorogenic reactions with TCO or BCN dienophiles.[46]

1.5.4.1.3 Features of dienophile in IEDDA reaction

On the other hand, even the type of dienophile affects the IEDDA kinetics. High dependence between the degree of ring strain and reactivity is observed for dienophile. In particular, dienophiles with electron-rich substituents (EDGs) enhance the reactivity against the diene. Concerning unstrained dienophiles Sauer and co-workers[47] in many papers have shown that olefinic dienophiles react faster than acetylenic dienophiles due to the increased electron withdrawing character of the triple bond, therefore lowering the energy of the HOMO.[48,49] Another important aspect which plays a very important role in IEDDA reaction rate is the ring strain of the dienophile, independent from the EDGs substituents. The ring strain raises the energy of $HOMO_{dienophile}$ increasing the rate of the reaction according to the following order: cyclopropene > cyclobutene > cyclopentene > cyclohexene > cyclooctene (**Figure 17-a**). Furthermore, it was observed that the use of *trans*-cyclooctene (TCO) has a huge energy difference compared to *cis*-cyclooctene according to the different conformation. In particular, Houk *et al.*[50] have studied that strained dienophiles are more reactive due to their pre-distorted conformation. Specifically, the 'crown' conformation adopted by *trans*-TCO is lower in energy compared to the 'half-chair' conformation of *cis*-cyclooctene (**Figure 17-b**).

As last, also the stereochemistry of strained dienophiles is a crucial aspect in terms of the reaction rate and the axial isomer (**Figure 17-c**) of a functionalized TCO has shown to be more reactive than the corresponding equatorial isomer by 1.1 kcal mol^{-1} (~4×).[51]

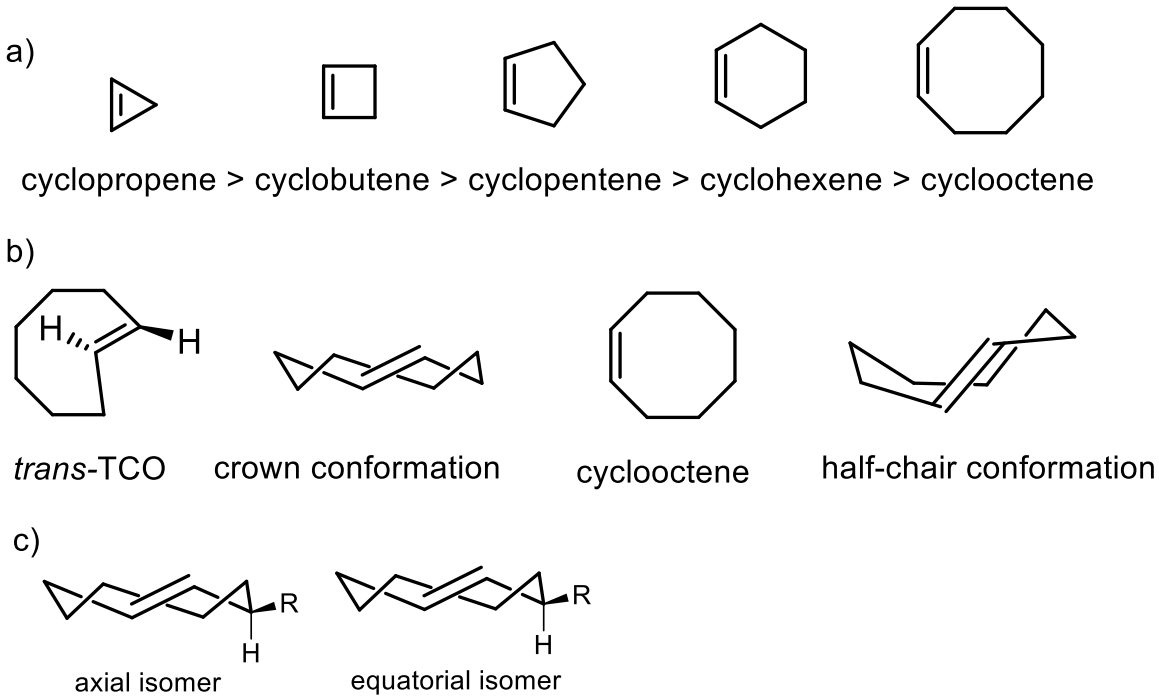


Figure 17. Influence of the strain effect on the IEDDA reaction. (a) Degree of ring strain along a series of cyclic alkenes. (b) The crown conformation of TCO makes the *trans* isomer orders of magnitude more reactive than *cis*-isomer. (c) axial and equatorial isomers of *trans*-TCO.

This biorthogonal IEDDA is fast and highly selective, making it a valuable tool in the field of chemical biology and bioimaging, offering a non-invasive and selective approach for visualizing specific biomolecules and cellular processes in complex biological environments. Additional application of IEDDA is about pre-targeting approaches used in radiochemistry[52,53]. Whereas in the past two decades the main focus of bioorthogonal chemistry was applied on 'bond formation' reactions for protein modification, very recently the focus has been moved on reactions that can cleave specific bonds under bioorthogonal conditions. This new approach, the so called ‘click and release’ method is applied also in therapy, in specifically delivering of drugs in case of tumor for example[54].

These bioconjugation strategies enable the integration of modified DNA into diverse platforms, facilitating applications in diagnostics, imaging, drug delivery, and materials science. The versatility of modified DNA allows researchers to design and tailor bioconjugates for specific purposes.

1.6 Modified nucleotides and their delivering *in cellulo*

In order to deliver modified DNA *in cellulo* is usually required the transfection which is the process of introducing nucleic acids (such as DNA, RNA, or oligonucleotides) into cells by different approaches such as electroporation, use of transfecting agent like cationic lipids[55] or polymers[56] to form complexes with nucleic acids, which can then be taken up by cells etc. This is the most common method used in molecular biology and biotechnology. About nucleotides, instead, they cannot cross the membrane without any presence of some carrier due to their anionic characteristic as the cell membrane. In particular, synthetic nucleoside triphosphate transporters (SNTTs)[57] is needed to transport them inside cells. The transporter is composed of a receptor, consisting in a cyclodextrin with Arg-rich tail (**Figure 18**), which forms a non-covalent complex with a triphosphate anion, and a cell-penetrating agent, which translocate the complex across the plasma membrane (**Figure 18**).[57] From this moment the DNA polymerase can incorporate the modified triphosphate in the nascent DNA during S phase in nuclei.

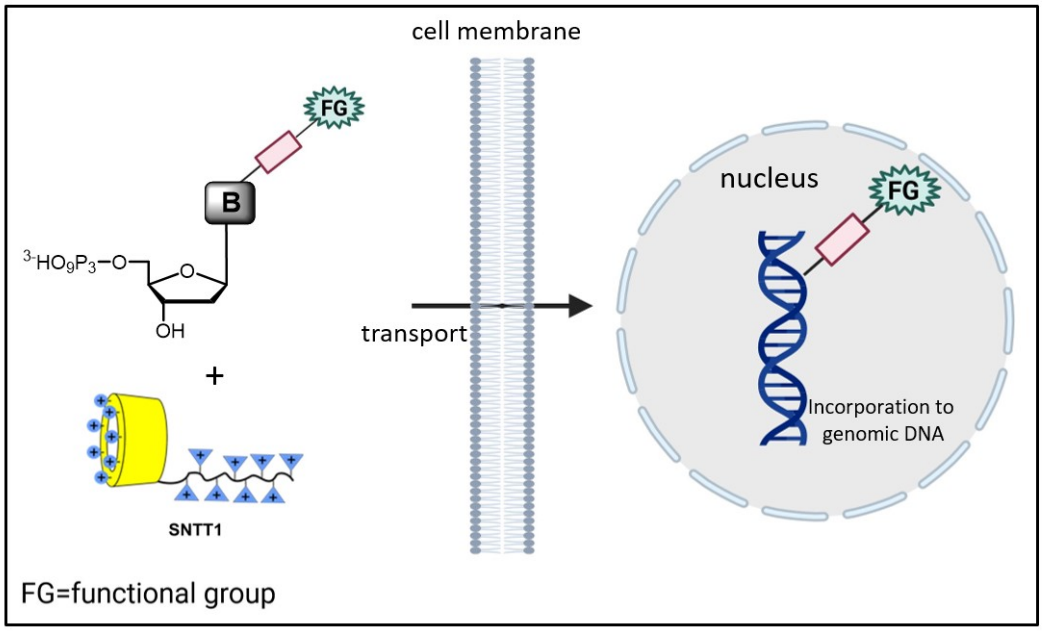


Figure 18. Representation of synthetic nucleoside triphosphate transporter (**SNTT1**)

1.7 DNA-protein interactions and bioconjugation

DNA-protein interactions and bioconjugations play crucial roles in various biological processes and biotechnological applications, such as cellular functions and signal transduction, structural biology and functionality and bioconjugations. In the group of cellular functions is worth to mention the gene regulation and DNA repair. DNA-binding proteins, such as transcription factors, regulate gene expression by binding to specific DNA sequences influencing cellular functions and development[58]. An other important role in DNA-protein interaction is the DNA repair in which many enymes such as helicases, polymerases, and ligases, are involved interacting with DNA in order to correct damage and maintain genomic integrity. Shana and co-workers discovered a correlation between carboxymethyl modification in DNA and consume of meat. In particular O^6-CMdG is removed by synergic cooperation of O^6-alkylguanine-DNA alkyltransferase (AGT) and nucleotide excision repair (NER) obtaining the DNA repaired.[59] Also, DNA-protein interactions are useful in keeping the chromatin structure and histones are the proteins interacting with DNA to form nucleosomes, the structural units of chromatin. This interaction is crucial for DNA packaging within the nucleus and regulation of gene accessibility[60]. Moreover, there are structural studies of protein-DNA complexes which help to understand the mechanisms of DNA recognition, binding, and manipulation by proteins, providing insights into their functional roles. Generally, bioconjugation refers to the covalent linking of two or more biomolecules or biomolecules with other molecules, such as drugs or labels, to form a single complex. Among bioconjugation reactions which are commonly used it is important to mention chemical cross-linking reactions which involve the formation of covalent bonds between functional groups on biomolecules.

1.7.1 DNA-protein cross-linking

DNA-protein cross-linking (DPCs) is one example of formation of covalent bonds between biomolecules. Many factors can induce these cross-links including chemical agents, radiation, and endogenous cellular processes. Certain chemicals such as cisplatin used in chemotherapy for cancer treatment brings to cross-links between DNA.[61] In the same way, exposure to ionizing radiation, such as X-rays or gamma rays, can cause DNA damage by DNA-protein cross-linking. Overall, these cross-links can compromise cellular processes like the progression of transcription and replication contributing to mutagenesis, carcinogenesis or cell death. There are different types of DPCs, according to DNA structure involved and the nature of the covalent cross-links itself,

distinguished in two main categories: 1) Non-enzymatic DPCs are caused by reactive agents, which nonspecifically cross-link chromatin proteins to DNA. One example of this group is the cross-linking caused by ROS (Reactive Oxygen Species) which generate high radical reactive species from DNA or amino acids[62]. In contrast, 2) enzymatic DPCs involve specific proteins that form covalent intermediates with DNA as part of their evolved enzymatic function. In this group it is important to mention the mode of action of chemotherapeutics resulting in the formation of topoisomerase DPCs. Enzymatic cross-linking involves, for example, Topoisomerase I (TOP I enzymes which relieve torsional stress, introducing a single-strand break, within the DNA molecule during transcription and replication[63]. In particular, during relaxation TOP I is covalently attached to the 3′ end of the single-strand break through tyrosine, resulting in the formation of TOP I covalent complex (TOP Icc). Nevertheless, site-specific DNA substrates containing transient hydrolytically stable DNA–protein conjugates are useful in chemical biology in order to understand the biological consequences of DPCs formation and their repair mechanisms. Another example of enzymatic methods involve glycosylase/β-lyases or DNA methyltransferase.[64] In the first case, during base excision repair (BER) the enzyme uses a basic amino acid side chain to displace the damaged base, forming a covalent DNA–protein intermediate by reducing the transient imino bond between BER proteins and their DNA substrates. In the second case, DNA methyltransferases (DNMT) are conjugated to DNA strands containing 5-aza-2'-deoxycitidine (dC^{N3}) or 5-fluoro-2'-deoxycitidine (dC^{F}). The normal biological function of DNMT is to generate 5-methylcytosines by transferring a methyl group from *S*-adenosylmethionine (SAM) to cytosine bases within CG dinucleotides of DNA, via a transient DNA-protein intermediate.[65] In contrast, when DNA substrates contains 5-fluoro-dC or 5-aza-dC instead of natural dC, DNMT traps, irreversibly, the DNA leading to important changes in the biological process. Concerning the synthetic method, DNA strands bearing unnatural nucleobases, which bind specifically a protein-reactive moiety, can be prepared by solid-phase synthesis. For example, Sanchez and Lloyd have prepared DNA strands incorporating site-specific oxopropyl-2′-deoxyguanosine[66], via phosphoramidite intermediate. When incubated with target peptides and proteins, forms reversible Schiff base adducts[67], which can be stabilized in the presence of $NaBH_3CN$ to generate stable conjugates. An alternative method in which both biomolecules, DNA and protein, are site-specific employs copper-catalyzed [3 + 2] Huisgen cycloaddition (click reaction) between azide-functionalized proteins and alkyne-containing DNA[68]. Copper-free alternative to CuAAC is the

strain-promoted alkyne-azide cycloadditions (SPAAC) previously described in Section 1.5.3. These methods are widely used in medicinal chemistry, molecular and chemical biology, diagnostic and imaging. The most investigated field by Hocek and co-workers include the amino acid-specific cross-linking reactions (Figure 10) where a nucleobase in DNA bears a reactive modification which targets natural side chains of amino acids.[69] Modified 2`-deoxyribonucleoside triphosphates are prepared and used for enzymatic synthesis of DNA, which is post-synthetically used for cross-linking with peptides and proteins. There are many examples of reactive probes to target specifically different amino acids, such as vinylsulfonammide[70] and acrylsulfonammide groups for cysteine-containing peptides and proteins by Michael addition reaction or DNA and most recently RNA bearing chloroacetamide[71,72] to cross-link cysteine or histidine- containing peptides and proteins. Moreover, DNA containing aldehyde both in major[73] and minor groove[74] was developed to target lysine containing peptides and proteins by reductive amination. As alternative, to target lysine containing peptides and proteins, in absence of other reagents, squaramate-dNTP and –DNA were prepared[75] (**Figure 19**). Most recently, Hocek and co-workers have developed 2`-deoxyribonucleotides and DNA bearing diketone[76] or glyoxal[77] moiety to cross-link arginine containing peptides and proteins by formation of stable pyrimidine or imizadolone-linked conjugates. Other aminoacids, such as tyrosine or tryptophan should be investigated for cross-linking with DNA to further extend the portfolio of aminoacids to target selectively.

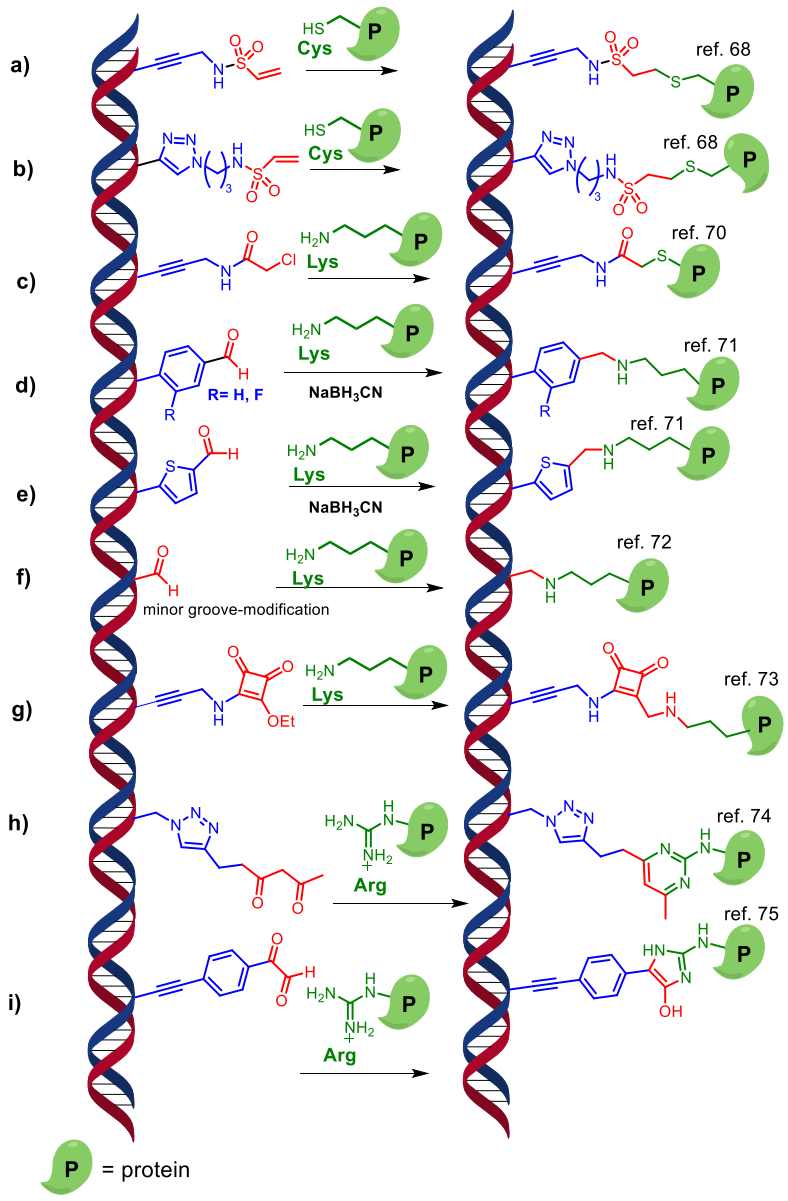


Figure 19. Reactive modified DNA previously developed in Hocek group. a), b) vinylsulfonammide modified-DNA for DPC with cysteine, c) chloroacetamide modified-DNA for DPC with cysteine and histidine, d), e), f) aldehyde modified-DNA for DPC with lysine through reductive amination, g) squaramate modified-DNA for DPC with lysine in absence of any reductive agents, h), i) 1,3-diketone and glyoxal modified-DNA for DPC with arginine.

1.8 Reagents for modification of serine

Serine (Ser, S) has a primary hydroxyl group and it is abundant at the surface of proteins with the function of site for post-translational modifications such as glycosylations and phosphorylations. The low nucleophilicity of the Ser hydroxyl group compared to the amino group of Lys and the thiol group of Cys, makes chemoselective reactions with electrophiles a challenging prospect.

It should be noted that serine has very low nucleophilicity with pKa at ~ 13, but in some proteins it is enhanced through interaction with residues such as the histidine and aspartate/glutamate present in the catalytic site of enzyme.[78] In particular, the imidazole ring of histidine behaves as proton shuttle interacting with carboxylate of the aspartic/glutamic acid and the OH- group of serine. In addition, a highly conserved oxyanion-binding site, commonly referred to as the oxyanion hole, stabilizes the negative charge on the carbonyl oxygen of the tetrahedral intermediate (**Figure 20-A**). In this way the pKa of serine is decreased at ~ 7, improving its nucleophilicity, which can be covalently modified by electrophilic species, as reported in most of the previous works targeted serine in active sites.[78] However, targeting less nucleophilic serines remains a challenge.

The first method for chemoselective serine conjugation on proteins was developed by Cravatt B. using fluorophosphonate/fluorophosphate (FP) derivatives like diisopropyl fluorophosphates as electrophiles to inhibit serine hydrolase (**Figure 20-A**).[79]

The second method was reported by Page M.I., Sieber A.S. and co-workers.[80] They used the four-membered ring β-sultams, structurally similar to β-lactones, to covalently conjugate to serine hydrolases and proteasome subunits by opening of four-membered ring (**Figure 20-B**). β-Sultams exhibit a smaller intramolecular resonance stabilization compared to the β-lactam analogues that in combination with ring strain contribute to their elevated reactivity.[81] Moreover, Page and co-workers have designed and synthetized a small collection of diverse *N*-acylated-β-sultams with different alkyne side chains based on previous work on analogous compound used as protease inhibitor.[82] An other similar approach involving the use of 3-oxo-β-sultams[83] was developed.

3-oxo-β-sultams behave as strong electrophiles, as *N*-acylated-β-sultams, which react with nucleophiles either at the carbonyl carbon or at the sulfonyl sulfur atoms.[84] They have been to inhibit serine hydrolases via acylation of the active-site serine residue (**Figure 20-C**). By crystallographic data was proved that 3-oxo-β-sultams probes are regioselective sulfonylating

electrophiles, covalently binding to the catalytic serine of human and porcine elastases through the sulfur atom. Also, the density functional theory (DFT) calculations shows that the sulfonyl sulfur atoms exhibit substantial positive charges, in contrast with the negative charges determined for the carbonyl carbon atoms, suggesting the sulfonyl center as the most electrophilic position.[85] Moreover, 3-oxo-β-sultam displayed to inhibit porcine pancreatic elastase (PPE) with a second-order rate $\sim 10^3$ higher than that of its *N*-acyl-β-sultam counterpart. Application of 3-oxo-β-sultam-derived activity-based probes in a human proteome revealed their potential to label disease-related serine hydrolases and proteasome subunits.

An other method for chemoselective serine conjugation on proteins was reported by Phil S. Baran and co-workers.[86] Based on inspiration of primary alcohols phosphorylation in protein performed by kinases and a new class of phosphorus reagents, they envisaged the reagents based on phosphorus(V) oxidation state reagents to target primary alcohols in peptides and proteins. The P(V) based bioconjugation reaction showed great selectivity for Ser. The conjugation reaction was performed at 25 °C in a 4 : 1 DMF/water mixture using 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU) as base. These conditions are far from physiological conditions and might not be applicable for many proteins (**Figure 20-D**). They used using these probes as proteomic tools for the identification of irreversibly bound protein targets in living bacterial cells. The reactive probes were incubated with different proteins and analysed by gel electrophoresis and mass spectroscopy.

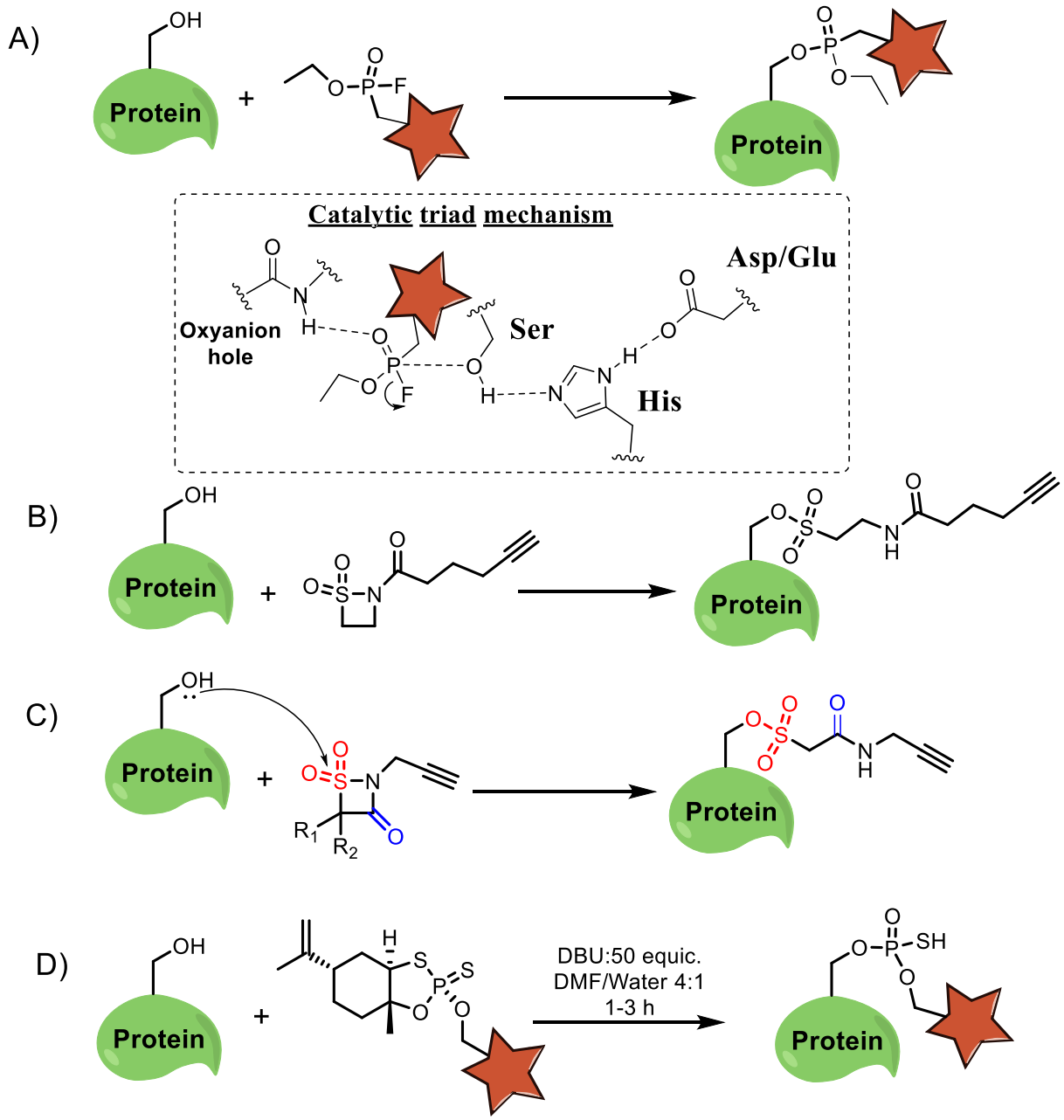


Figure 20. Examples of reactions used for conjugation to serine residue. A) Conjugation using phosphorus(V) oxidation state reagents; B) Conjugation using N-acylated-β-sultams- based reagents; C) Conjugation using 3-oxo-β-sultams- based reagents, involving sulfonylation over acylation; D) Conjugation using a new class of phosphorus reagents.

1.9 Reagents for modification of glutamic and aspartic acids

The side chains of aspartic acid (Asp, D) and glutamic acid (Glu, E) contain a carboxylic acid group with a pK_a value of approximately 4. Therefore, they have a negative charge at physiological pH and the carboxylate reactivity is limited in water, making Asp- and Glu-selective reactions a challenging scenario. The nucleophilicity of the carboxylate is low, requiring the probe to be highly reactive which increases the risk of cross reactions with more nucleophilic amino acids. For this reason, reagent with a directing group are needed to be used. An example of this approach is to use affinity directed tetrazole reagents which react with carboxylic residues by a photo induced reaction[87] (**Figure 21-A**). Upon irradiation with UV light, the tetrazole is converted to a reactive carboxy-nitrile imine that can react with various nucleophiles, including carboxylic group residue (**Figure 21-A**). An other example for chemoselective protein conjugation at carboxyl residues was shown in the study by Zhengqiu Li and co-workers.[88] 2H-azirine-based reagents was used by them and the reaction was performed at room temperature at pH 7.3 to label proteins both in vitro and in live cells (**Figure 21-B**). Glutamate is preferentially labeled than aspartate by 2H-azirine because the steric hindrance is lower. Only a few off-target reactions were observed when the probe is used for protein profiling in living cells with >95 % selectivity towards Asp/Glu. Moreover, selectivity was shown towards buried residues in active sites compared to surface exposed residues making 2H-azirine probe ideal for reactivity profiling and proteomics. The main limitation of this strategy is the impossibility to be applied in protein constructs where the activity of the protein must be intact. A third approach was demonstrated by Chaubet and co-workers who studied multicomponent reactions for Asp/Glu conjugation on native proteins.[89] The model protein trastuzumab was reacted with a combination of suitable aldehyde and isocyanide reagents, they found that protein conjugation could occur at Asp or Glu in vicinity of a Lys with the formation of dual residue modification product or single residue modification products (**Figure 21-C**). Bioconjugation reactions were performed in PBS buffer, pH 7.4 at 25 °C for 16–20 h. Conjugation was observed with varying conversions between 0 and 97 % depending on the reagent structure, the amount and the ratio between the reagents. Analysis of the conjugates by LC-MS/MS validated that the reaction is specific for Asp/Glu.

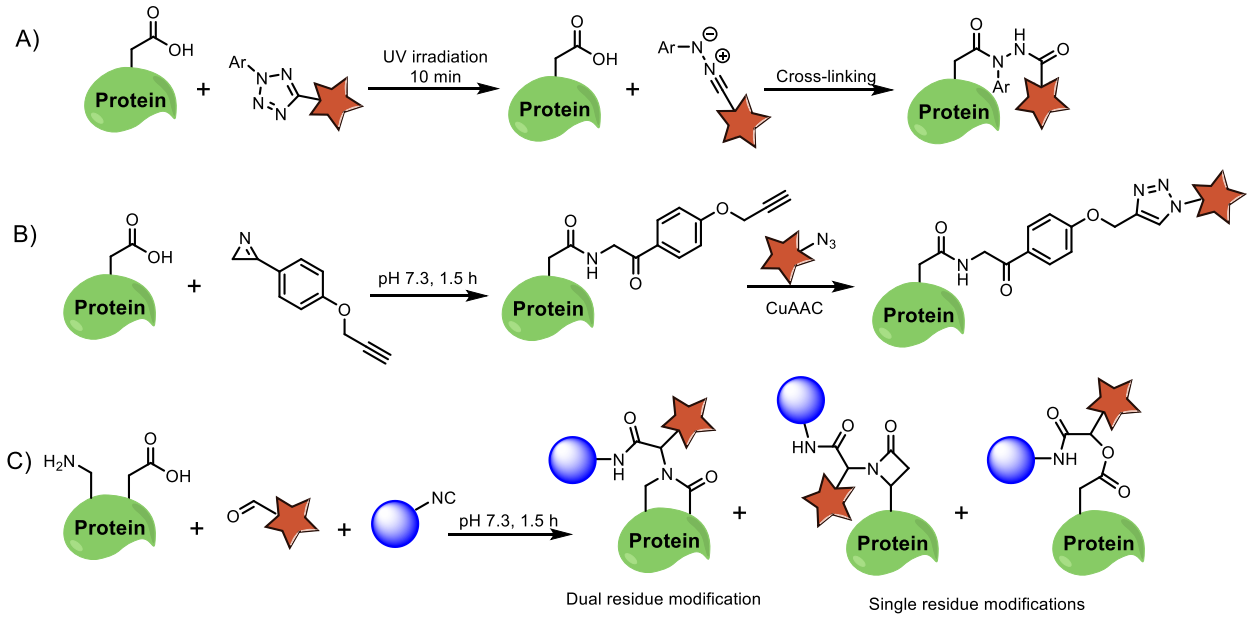


Figure 21. Examples of reactions used for conjugation to carboxylic acids. A) Conjugation using tetrazole based probes; B) Conjugation using 2*H*-azirine-based reagents; C) Three-component reaction, resulting in either dual or single residue modification.

1.10 Reagents for modification of tryptophan

Tryptophan (Trp) is the most rarely encountered in proteins[90] and due to its highly hydrophobic nature, it is often buried deeply within globular proteins.

In its backbone there is a five-membered electron-rich aromatic heterocyclic ring containing a nitrogen atom and is fused to a six-membered aromatic ring which makes the tryptophan different from other amino acids (**Figure 22**). In other words, like tyrosine, tryptophan is π-nucleophilic, with electrophilic substitution readily taking place at the C2 position of the indole ring.

What makes, tryptophan as amino acid, interesting to study, according to its rare presence (~1%), is the possibility to achieve high control over protein conjugates.

Previously, John M. Antos and Matthew B. Francis reported some methods for the conjugation of Trp which require toxic heavy metals, rhodium, or biochemically incompatible conditions[91] (**Figure 22-a**). The mechanism behind, involving the formation of metallocarbenoids, bring toward the alkylation at the N or C_2 position of Trp. The main disadvantage of these methods was to exhibit cross reactivity to other amino acids (particularly to tyrosine), thus limiting the range of applications (**Figure 22-c**). An other method reported, requires the presence of Pd [92] (**Figure 22-**

b). Later on, an alternative mild method and transition metal-free was described by Seki Y. *et al.*[93] to selectively modifying Trp for the bioconjugation of peptides and proteins.

This method tolerates many different functional groups and involves an oxidation-mediated functionalization in the presence of 9-azabicyclo[3.3.1]nonane-3-one-*N*-oxyl (keto-ABNO). The mechanism behind involves the oxidation of keto-ABNO in presence of oxidant like sodium nitrite $NaNO_2$ (**Figure 22-d**). The resulted reactive electrophilic specie, oxoammonium cation, further reacts with tryptophan by making a C–O covalent bond at the 3-position of the indole moiety of a Trp residue.[93] The hydrated end-product was proved to be stable by the crystal structure of the Trp-modified lysozyme. Moreover, tryptophan has one chiral center, so it exists in two enantiomeric forms, L-tryptophan (the naturally occurring form found in proteins) and D-tryptophan.

From biological point of view, Trp plays a crucial role in the synthesis of proteins and is the precursor of some important molecules in the body, such as serotonin and melatonin. Serotonin is a neurotransmitter that regulates behavior, appetite, and sleep, while melatonin regulates the sleep-wake cycle.[94]

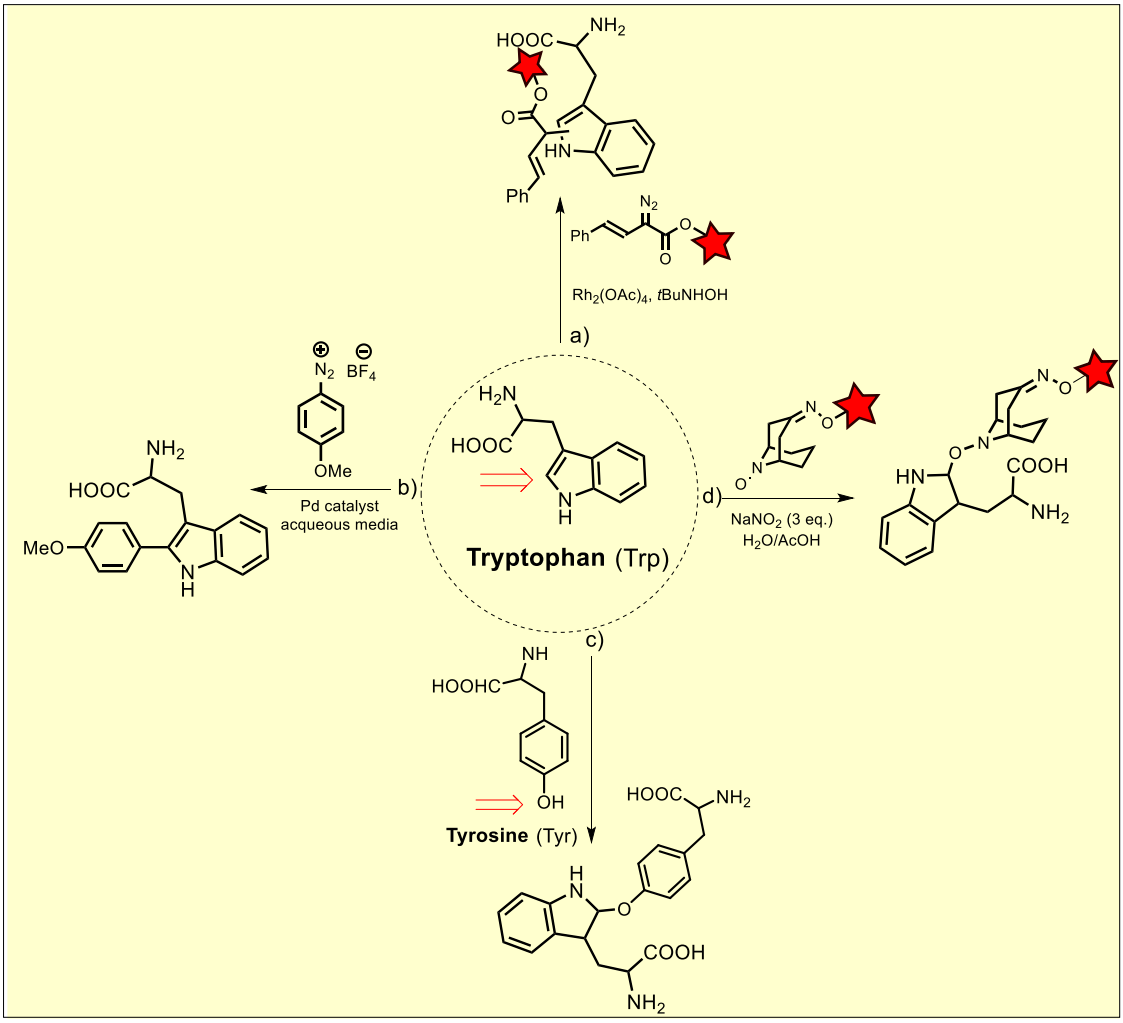


Figure 22. Chemical structure of tryptophane (in the centre) and different methods how to target it selectively. Red arrows indicate the reactive centers of molecules.

2 SPECIFIC AIMS OF THE WORK:

Design and chemical synthesis of 2′-deoxycitidine triphosphates (dCTPs) bearing 2- or 4-linked *trans*-cyclooctene (TCO) or bicyclononyne (BCN) tethered through a shorter propargylcarbamate or longer triethyleneglycol-based spacer. Enzymatic incorporation of TCO- and BCN-dCTPs in DNA and testing their reactivity *in vitro* or *in cellulo* after treatment with fluorophore-containing tetrazines by inverse electron-demand Diels–Alder (IEDDA) click reaction for metabolic labeling and imaging of DNA.

Design and chemical synthesis of 2'-deoxyribonucleotides bearing β-sultam or 3-oxo- β-sultam reactive group selective for serine, enzymatic incorporation and testing of cross-linking with Ser- containing peptides and proteins.

Design and chemical synthesis of 2'-deoxyribonucleotides bearing 3-membered azirine ring reactive group selective for glutamic or aspartic aminoacids, enzymatic incorporation and testing of cross-linking with Glu- or Asp- containing peptides and proteins.

Design and chemical synthesis of 2'-deoxyuridine-5'-*O*-mono- and triphosphate bearing ABNOH reactive group selective for tryptophan, enzymatic incorporation of ABNOH-dUTP in DNA and testing of cross-linking with tryptophan-containing peptides and proteins.

2.1 Rationale of the specific aims

Methods commonly used to detect and quantify DNA replication during S phase of cell cycle are based on incorporation of modified nucleotide analogues bearing halogen (dU^{Br}, dU^{I})[95] or alkyne (dU^{E}) and subsequently staining by immunocytochemical (dU^{Br}, dU^{I}) or by chemical reaction (dU^{E}).[96] Monitoring of cell cycle progression by metabolic labelling of DNA is used in experimental cell biology, cancer research, and clinical medicine.[97] In contrast to previous methods, employing fluorescent probes which require permealization of the cell membrane or long period of UV light irradiation, T. Kraus and co-workers have developed a new method[57]. This method is based on the use of synthetic nucleoside triphosphate transporter (**SNTT1**)[57], allowing fluorescently labeled deoxynucleoside triphosphates (FL-dNTPs), polar and negatively charged, to be delivered inside cells. At this point, FL-dNTPs are incorporated into genomic DNA in live cells. There are many examples[98,99] of clickable nucleosides and nucleotides used in the chemical[100,101] or enzymatic[102] synthesis of clickable nucleic acid probes, including several dNTPs (mostly 2′-deoxyuridine nucleotides) bearing *trans*-cyclooctene (TCO) or bicyclononyne (BCN)[38] undergoing IEDDA reactions with various tetrazines. However, these studies were so far limited to *in vitro* DNA labeling. The ability to perform the labeling reactions in the complex environment of the nuclei in living cells was not yet investigated. Therefore, one of the aims of the thesis was to systematically study and synthetize several new dCTPs bearing TCO or BCN moieties linked through different tethers in order to use by **SNTT1**[57]-assisted delivery for live cell metabolic labeling with different tetrazines through incorporation of the reactive nucleotides into genomic DNA followed by the IEDDA click reactions.

Concerning the DNA-protein cross-linking, it is often studied to understand cellular processes in which are involved transcription factors, DNA repair proteins, or chromatin-associated proteins. Previously, DNA functionalized with chemical[103] or photochemical[104,105] reactive groups can be used in molecular biology to study DNA-proteins conjugates and to target unknown DNA-binding proteins.[106,107] One step to fulfill was to target specifically DNA-binding proteins, therefore, several amino acids were explored in terms of cross-linking with reactive groups linked to DNA. For this reason, Hocek and co-workers have designed and synthetized several modified DNAs, such as dNTPs and DNAs bearing vinylsulfonammide[69] or chloroacetamide[71] to cross-link with Cys or His containing peptides and proteins, dNTPs and DNAs bearing aldehyde[108] and squaramate[74] have been developed to target Lys containing peptides and proteins with or without the need of reductive agents. Most recently 1,3-diketone[75] and glyoxal[76] linked to dNTPs and DNAs were designed and synthetized for cross-linking reactions with arginine containing peptides and proteins. Many aminoacids were investigated previously for cross-linking with modified DNA, but no specific reactive group for targeting aminoacids such as serine, glutamic/aspartic acids and tryptophan have been reported in DNA so far. Therefore, an other aim of this thesis was to study and use specific reactive groups for this purpose.

For DNA-probes selective for serine, the β-sultam[80] moiety was chosen to design and synthetize new modified nucleotides and fashioned DNA-probes for cross-linking with proteins.

For DNA-probes selective for glutamic/aspartic bearing, the azirine[88] reactive mojety was chosen to design and synthetize new modified nucleotides and fashioned DNA-probes for cross-linking with proteins.

Tryptophan is an important amino acid rarely present in proteins, for this reason, it is interesting to study thanks to the possibility to achieve high control over protein conjugates. Also, DNA-probes selective for this amino acid were not yet investigated, therefore, the last aim of the thesis was to use ABNOH[93] mojety in order to synthetize modified nucleotides to build Trp- selective fashioned DNA-probes for cross-linking with proteins.

3 RESULTS AND DISCUSSION

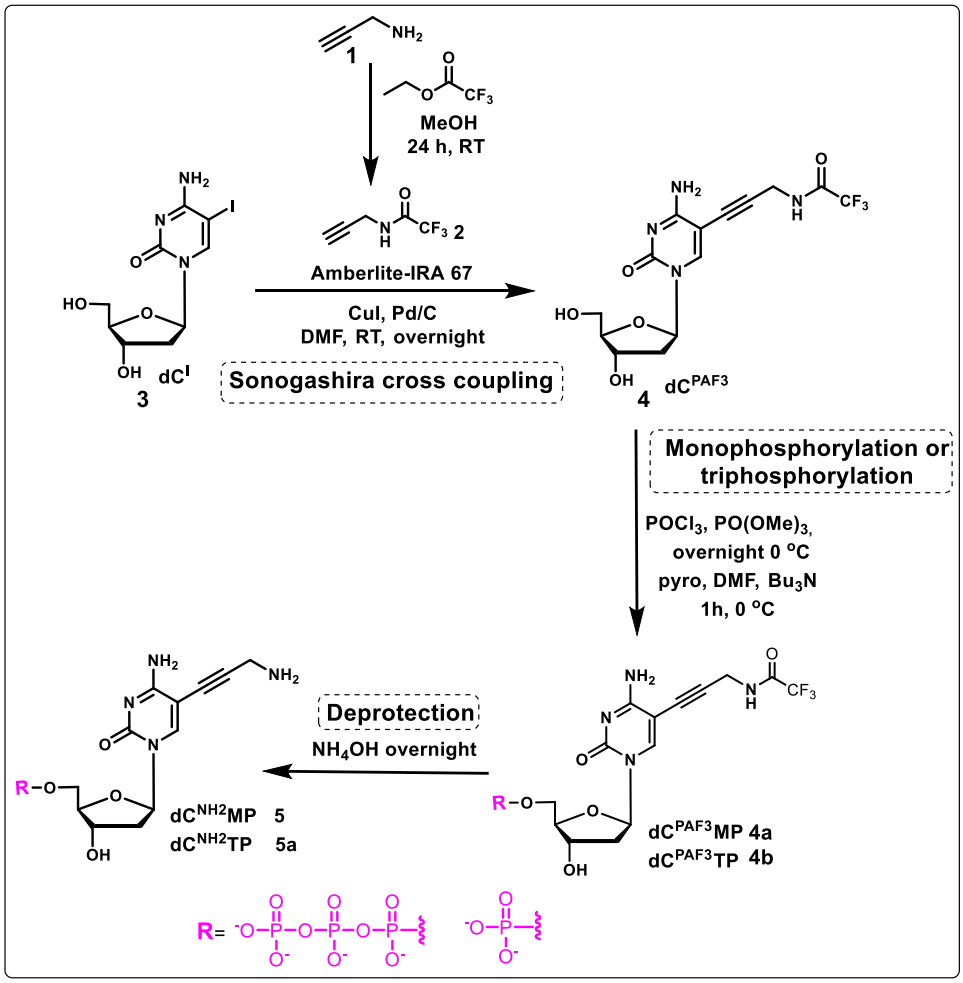
3.1 PART I

3.1.1 Design and synthesis of modified nucleotides bearing 2- or 4-linked *trans*-cyclooctene (TCO) or bicyclononyne (BCN) tethered through a shorter propargylcarbamate or longer triethyleneglycol-based spacer, their enzymatic incorporation and testing of reactivity with tetrazines

2- or 4- *trans*-cyclooctene (TCO) or bicyclononyne (BCN) were chosen as potential reactive groups to be inserted in DNA for fluorogenic IEDDA click labeling with tetrazines as tool for bioanalytical applications. Therefore, TCO- or BCN modified triphosphates were systematically studied by comparing their substrate activity for polymerase incorporation, reactivity with different tetrazines, efficiency of transport across the cell membranes using **SNTT1**[57] and capacity for live cell metabolic labeling through incorporation of the reactive nucleotides into genomic DNA followed by the IEDDA click reactions.

3.1.1.1 Synthesis of precursors $dC^{PAF3}MP$, $dC^{NH2}MP$, $dC^{PAF3}TP$, $dC^{NH2}TP$

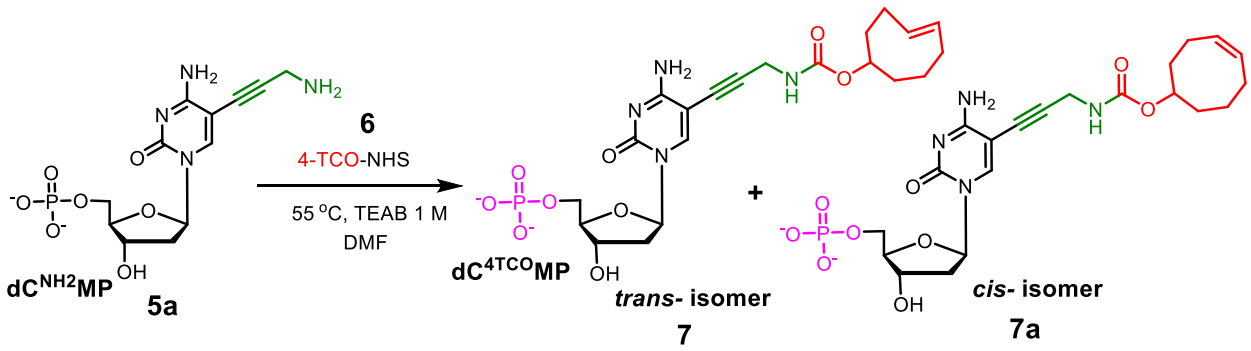
Several new dCTPs bearing *trans*-cyclooctene (TCO) or bicyclononyne (BCN) moieties tethered through a shorter propargylcarbamate or longer triethyleneglycol-based spacer were synthetized. The synthesis of new modified 2′-deoxycytidine monophosphate or triphosphate started from propargyl trifluoroacetamide 2′-deoxycytidine (**dC^{PAF3}**)[109] which was inserted at position 5 by Sonogashira cross-coupling. The reaction was performed in presence of CuI and $Pd(PPh_3)_4$ in DMF. Propargylamine moiety was prior protected by trifluoroacetyl group, subsequently the nucleoside **dC^{PAF3}** was converted in monophosphate **$dC^{PAF3}MP$** or triphosphate **$dC^{PAF3}TP$** and then deprotected by aqueous NaOH for 24h.[110] Therefore, the desired products were isolated after HPLC purification in acceptable yields (**$dC^{NH2}MP$** 81%) or triphosphate (**$dC^{NH2}TP$**, 20%) (**Scheme 1**). 5-Propargylamino-2'-deoxycytidine **$dC^{NH2}TP$**110 was used as the key intermediate for the synthesis of all functionalized nucleoside triphosphates.



Scheme 1. Synthetic scheme of 5-propargylamino 2′-deoxycytidine monophosphate or triphosphate (**dC^{NH2}MP** or **dC^{NH2}TP**).

3.1.1.2 Synthesis of 4-linked trans-cyclooctene (4-TCO) modified monophosphate (dC^{4TCO}MP)

In order to have a model nucleotide to test for IEDDA reaction with tetrazines, 4-TCO modified 2′-deoxycytidine monophosphate was synthesized. The reactive TCO dienophile group was installed to **5** (**dC^{NH2}MP**) through carbamate moiety by nucleophilic substitution. The reaction of **dC^{NH2}MP** with 4-TCO-*N*-succinimidyl ester was conducted at 55 °C for 4 h and the final product (**dC^{4TCO}MP, 7**), *trans*- isomer, was isolated after HPLC purification in 48% yield (**Scheme 2**). Also, formation of cis isomer **7a** was observed most probably due to high temperature.

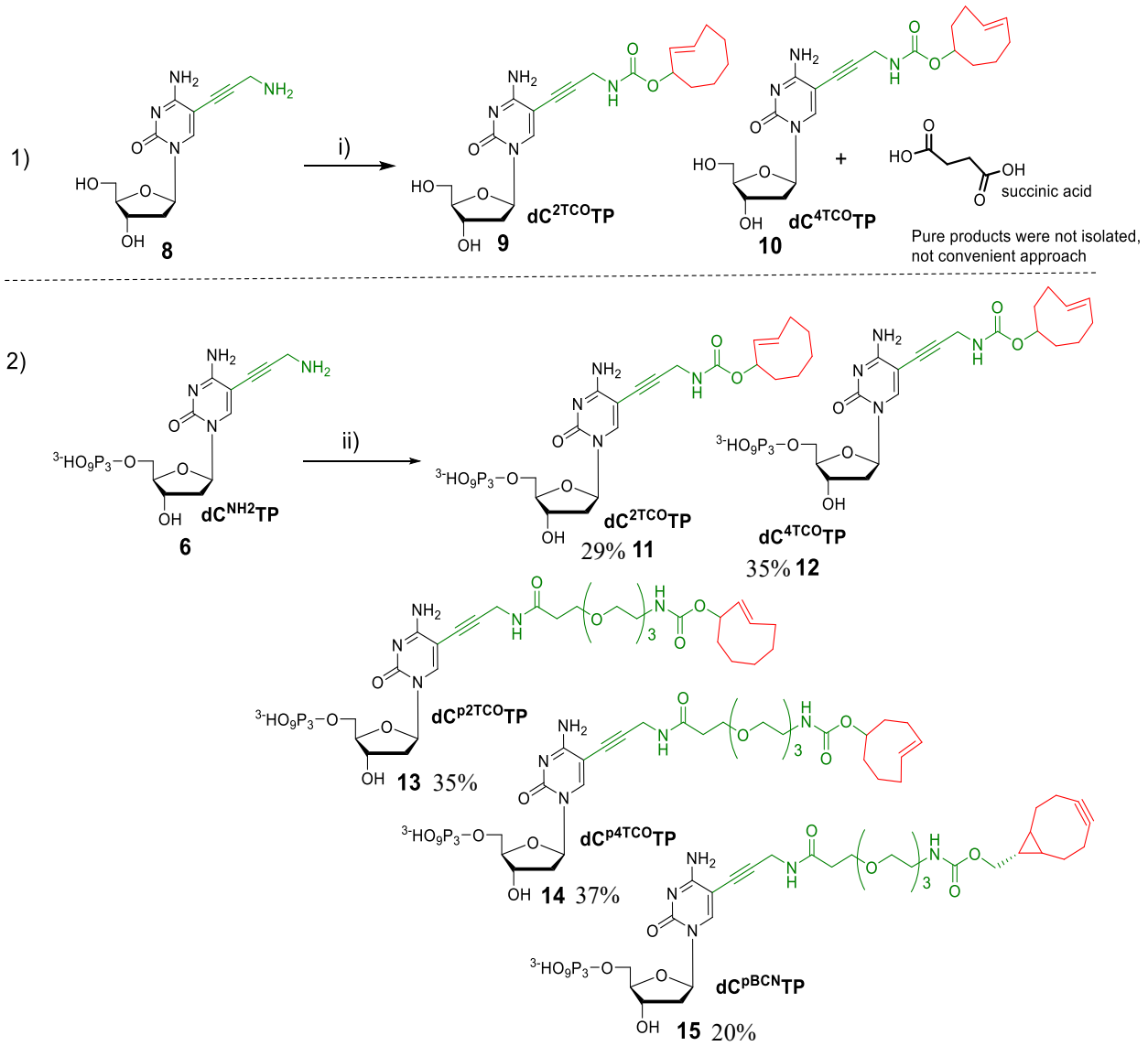


Scheme 2. Synthesis of 4-TCO modified 2′-deoxycytidine monophosphate.

3.1.1.3 Synthesis of 2- or 4-linked trans-cyclooctene (TCO) or bicyclononyne (BCN) triphosphates (dC^{2TCO}TP, dC^{4TCO}TP, dC^{p2TCO}TP, dC^{p4TCO}TP, dC^{pBCN}TP)

The design of the dNTP conjugates bearing reactive TCO and BCN groups was partly inspired by prior studies.[38] In order to prepare four dCTP derivatives with appended *trans*-cyclooct-2-en-1-oxy- (2TCO) and *trans*-cyclooct-4-en-1-oxy- (4TCO) reactive groups tethered to the nucleobase either via a short linker **11** (**dC^{2TCO}TP**) or **12** (**dC^{4TCO}TP**) or via a triethyleneglycol (PEG3) spacer **13** (**dC^{p2TCO}TP**) or **14** (**dC^{p4TCO}TP**) were tried two different approaches. In the first approach the synthesis started from 5-propargylamino-2′-deoxycitidine nucleoside **8** (**dC^{NH2}**), which was synthetized according to published protocol, and by nucleophilic substitution 2TCO-NHS-carbonate (**9a**) or 4TCO-NHS-carbonate (**10a**) (commercially available) were attached and the reaction was conducted at 55 °C for 4 h in DMF (**Scheme 3**). Unfortunately, succinic acid formed as by-product and as well DMF present in the reaction mixture were not possible to remove completely from modified nucleosides. Therefore, purified products were not isolated. The second approach was more convenient and the synthesis started from 5-propargylamino-2′-deoxycitidine triphosphate **dC^{NH2}TP**110 as intermediate for the synthesis of all functionalized nucleotides. Thus, the reactive ester bearing 2-TCO or 4-TCO reactive moieties were connected to **dC^{NH2}TP**110 via nucleophilic substitution reaction. Slow gradient was required to separate pure products **11** and **12** which were isolated after two HPLC separations in 29% and 35% yields, respectively. Similarly, the reactions of **dC^{NH2}TP** with 2TCO-PEG3 (**13a**) or 4TCO-PEG3-NHS (**14a**) esters (as well commercially available) gave PEG3 spacer-containing **13** and **14** (**dC^{p2TCO}TP** and **dC^{p4TCO}TP**) nucleotides in 35% and 37% yields, respectively. As last, the BCN derivative **15** (**dC^{pBCN}TP**) was

obtained in 20% yield by the reaction of the nucleotide **dC^{NH2}TP** with *endo*-BCN-PEG3-NHS ester (**Scheme 3**). In all cases of modified triphosphates the nucleophilic substitution reaction was conducted at 55 °C for 4 h in TEAB 1M pH 8 buffer. For nucleotides bearing triethyleneglycol (PEG3) spacer only one HPLC run was enough to isolate the pure desired products **13-15** (**dC^{p2TCO}TP**, **dC^{p4TCO}TP**, **dC^{pBCN}TP**) with no presence of additional impurity.



Scheme 3. Synthesis of 4TCO-, 2TCO-, p4TCO-, p2TCO- and pBCN- modified-2'-deoxycitidine triphosphates. Conditions i) 2TCO-NHS-carbonate (**9a**) or 4TCO-NHS-carbonate (**10a**) at 55 °C for 4 h in DMF or ii) 4TCO-NHS-carbonate (**9a**) or 2TCO-NHS-carbonate (**10a**) 2TCO-PEG3-NHS-carbonate (**13a**), 4TCO-PEG3-NHS-carbonate (**14a**) or endo-BCN-PEG3-NHS-ester (**15a**) at 55 °C for 4 h in TEAB 1 M pH 8.

In case of compound **11** mixture of diastereoisomers ~3:2 was formed, observable from ^{1}H and ^{13}C NMR on signal of H-1′, C-1′,2′ (**Figure 23-A,B**), due to new stereogenic center on the cyclooctene ring.

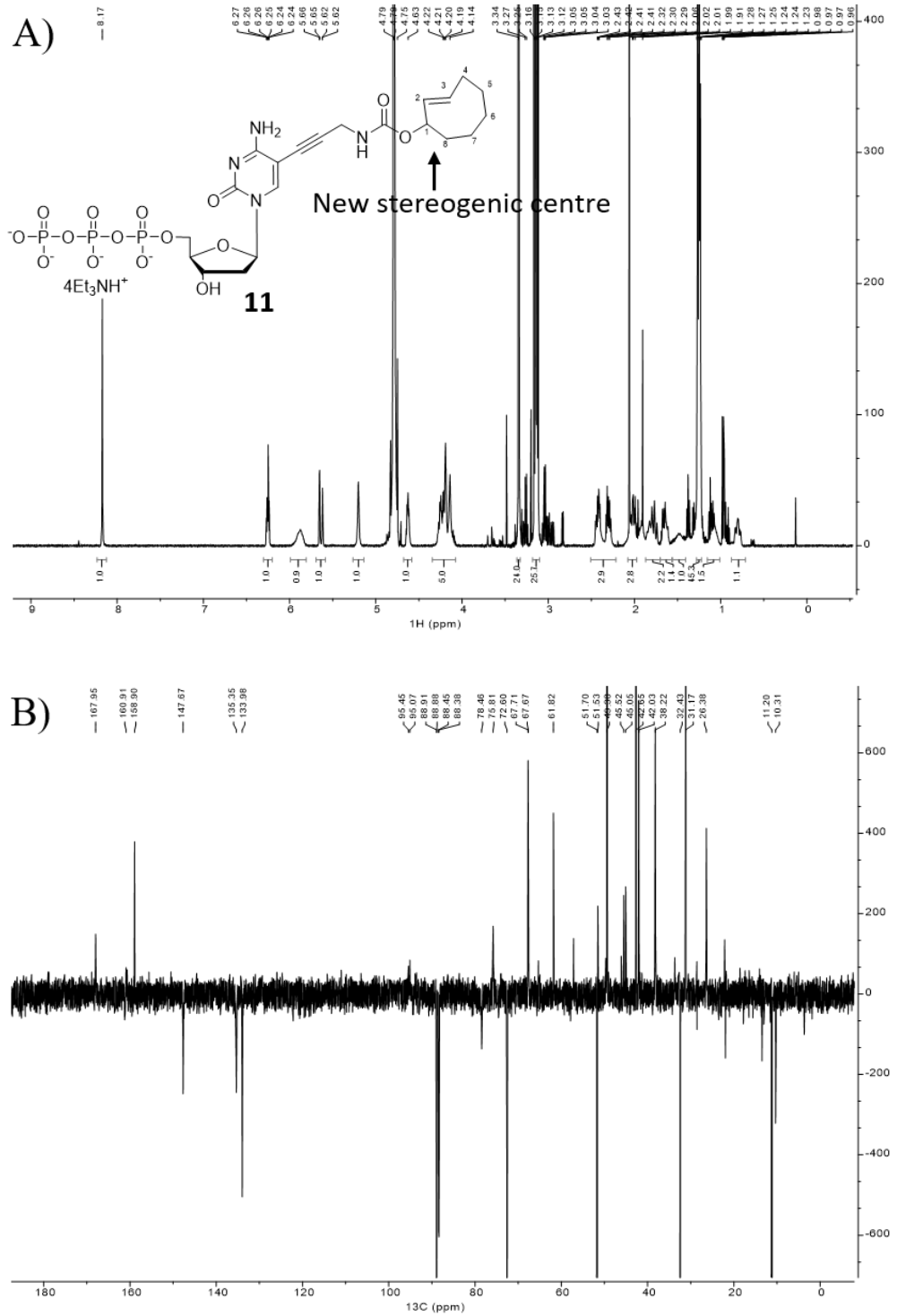


Figure 23. NMR spectra of A) ^{1}H NMR and B) ^{13}C NMR of **11**

3.1.1.4 Inverse electron demands Diels-Alder reactions (IEDDA)

3.1.1.4.1 Structures of tetrazines used in this study

In order to study the reactivity of TCO- (*trans*-cyclooctene) and BCN- (bicyclononyne) in IEDDA reactions, different kind of tetrazines were taken in consideration in this study.
7-(Azetidin-1-yl)-4-methyl-3-(5-(6-pyrimidin-2-yl)-1,2,4,5-tetrazin-3-yl)pyridin-3-yl)-2*H*-chromen-2-one (**T1**)[46], 7-(Azetidin-1-yl)-4-methyl-3-(3-(6-(pyridin-4-yl)-1,2,4,5-tetrazin-3-yl)phenyl)-2*H*-chromen-2-one (**T2**)[46], N,N-dimethyl-6-(6-methyl-1,2,4,5-tetrazin-3-yl)acridin-3-amine (**T3**)[111], 2-(6-(dimethylamino)-3-(dimethyliminio)-3H-xanthen-9-yl)-5-((3-(4-(6-(pyrimidin-2-yl)-1,2,4,5-tetrazin-3 yl)benzamido)propyl)carbamoyl)benzoate or pyrimidyl-tetrazine-Cy3 (**T4**) and 3,6-di-2-pyridyl-1,2,4,5-tetrazine (**T5**) (**Figure 24**). **T1**, **T2** and **T3** tetrazines were synthetized according to published literature[46] and provided by Milan Vrabel`s IOCB group, instead **T4** and **T5** tetrazines are commercially available.

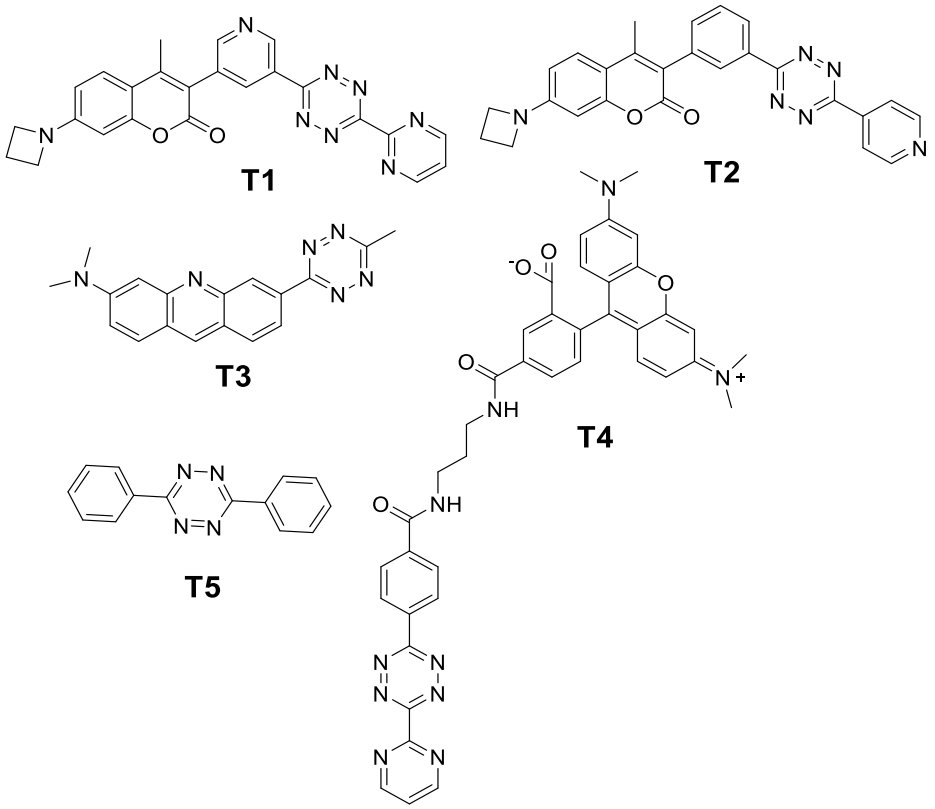
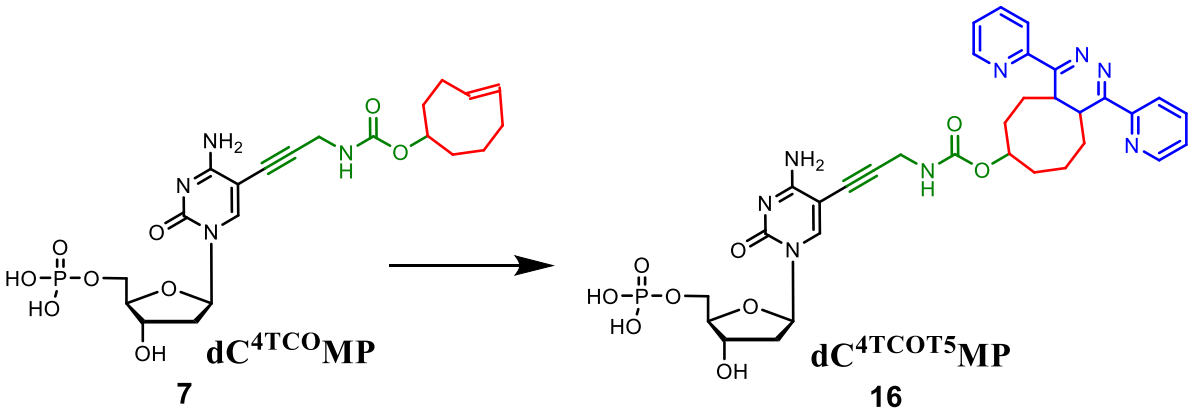


Figure 24. Chemical structures of tetrazines used in this study.

3.1.1.4.2 Model IEDDA reaction between dC^{4TCO}MP and 3,6-di-2-pyridyl-1,2,4,5-tetrazine (T5)

Subsequently, the reactivity of **dC^{4TCO}MP** was tested in model reaction with 3,6-di-2-pyridyl-1,2,4,5-tetrazine, **T5** tetrazine (**Scheme 4**). The reaction was conducted in MeCN/H_2O (3:1) at 37 ºC for 2 h, reaction conditions as close as natural environment for cells allowing them to be alive. The formation of the desired product **dC^{4TCOT5}MP** was confirmed by MS and nuclear magnetic resonance (NMR). However, the complete assignment of ^{1}H and ^{13}C NMR spectra was impossible for two reasons. The first one, the concentration of the sample was not sufficient for direct observation of ^{13}C NMR spectrum even though the reaction was repeated several times. Second, based on reaction mechanism[40], click reaction of substituted trans-cyclooctene with tetrazine provides a mixture of several isomers that contributes to complexity of NMR spectra as well, which were not possible to separate by HPLC (**Figure 25**). Instead, 2D correlation experiments such as H,C-HSQC and H,C-HMBC were used to identify structural fragments of the product in ^{1}H NMR spectrum, such as 2`-deoxyribose, nucleobase, linker, cyclooctane ring and pyridyl moiety (**Figure 25**). Therefore, the determination of the yield was not possible but the formation of conjugate was proved by NMR.



Scheme 4. Reaction of **dC^{4TCO}MP** with **T5** tetrazine. Conditions: MeCN/H_2O (3:1) with 3,6-di-2-pyridyl-1,2,4,5-tetrazine (**T5**) 3.2 equiv., 37 ºC, 2 h.

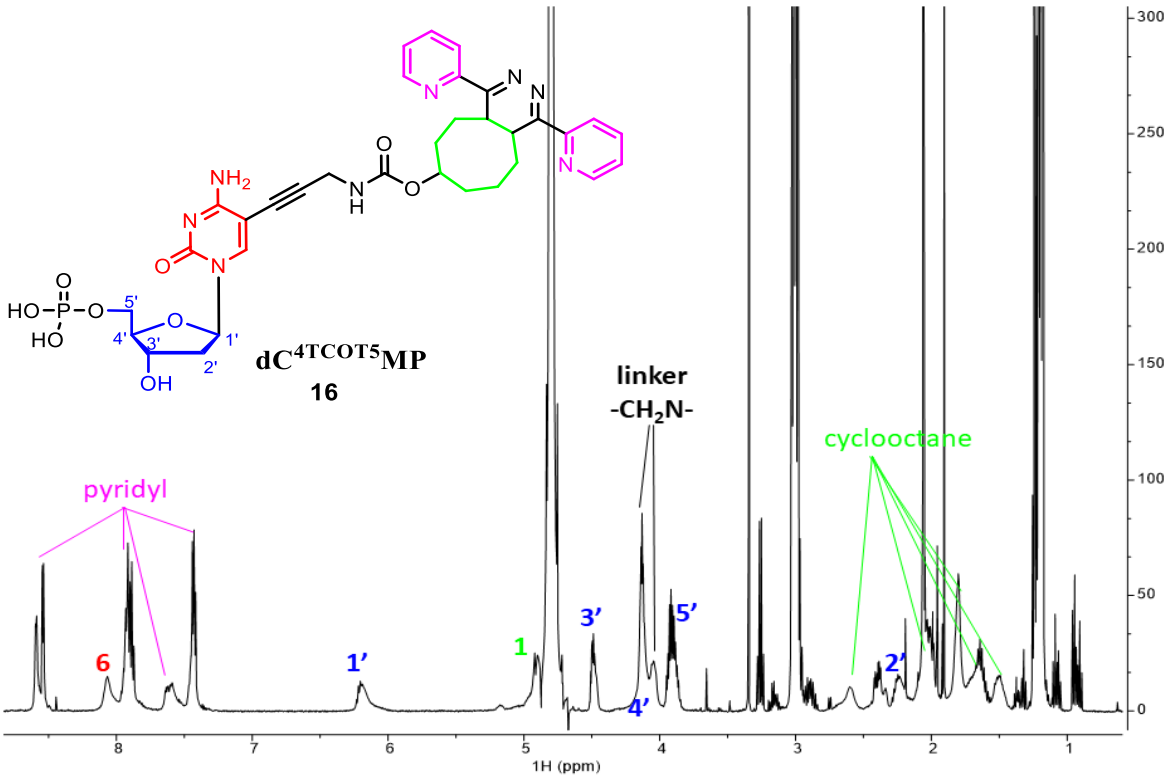


Figure 25. Assignment of structural fragments of **dC^{4TCOT5}MP** in ^{1}H NMR spectrum.

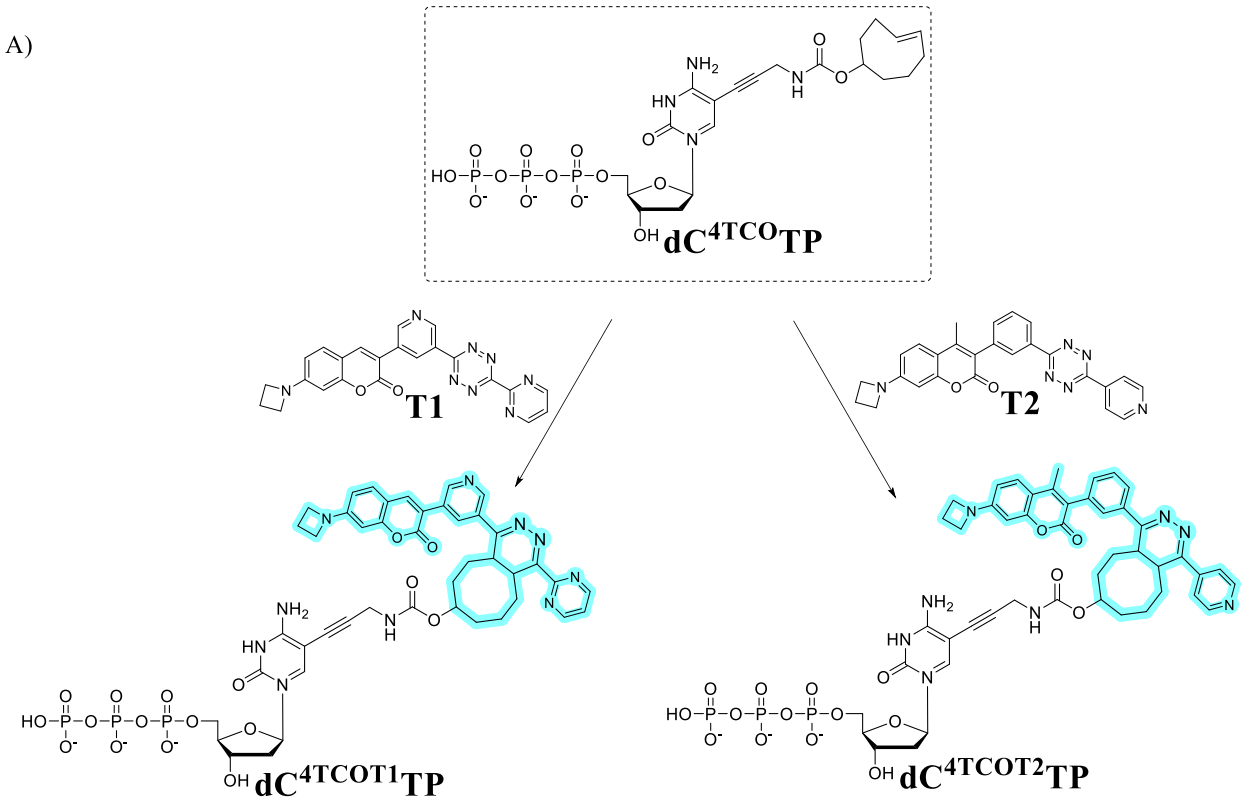
3.1.1.4.3 Reaction between modified 2- or 4-linked trans-cyclooctene (TCO) or bicyclononyne (BCN) triphosphates (dC^{2TCO}TP, dC^{4TCO}TP, dC^{p2TCO}TP, dC^{p4TCO}TP, dC^{pBCN}TP) and tetrazines, fluorescence turn-on measurements, LC/MS characterizations and study of kinetics formation of dC^{XTY}TP click products

After having demonstrated the proof of concept of click product **dC^{4TCOT5}MP** formation by IEDDA model reaction between **dC^{4TCO}MP** and symmetric tetrazine **T5** by NMR, the reactivity of modified triphosphates bearing 2TCO, 4TCO and BCN reactive moieties was tested towards the other tetrazines, previously shown, in order to measure the fluorescence derived from tetrazine-TPs conjugates (**dC^{XTY}TP**), studying their kinetic formations and characterize them by LC-MS. The purpose of these analysis is to have reliable *in vitro* studies before moving to cells which are more complex system.

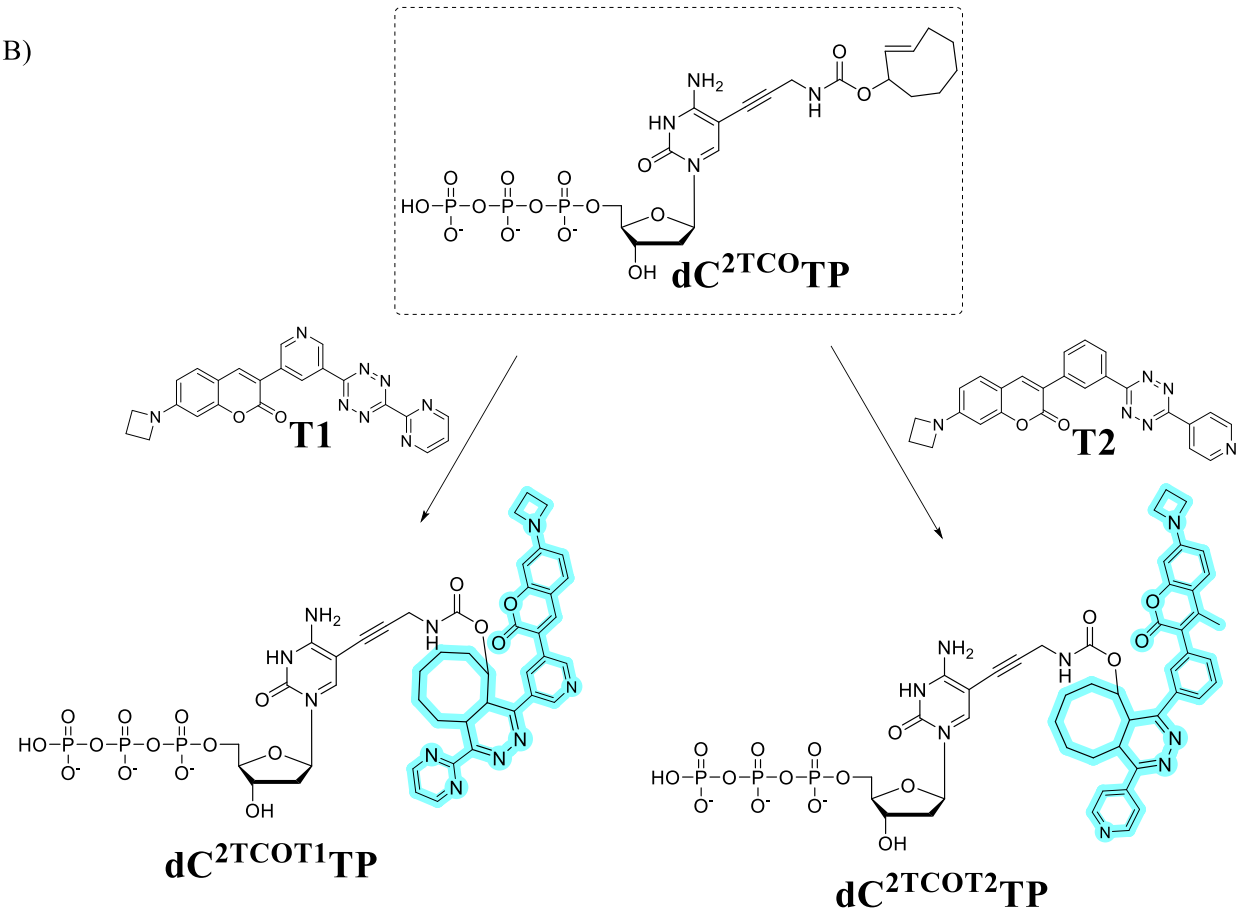
3.1.1.4.4 Reactions between dC^{4TCO}TP or dC^{2TCO}TP with T1 or T2 and between dC^{p4TCO}TP or dC^{p2TCO}TP or dC^{pBCN}TP (dC^{X}TP) with T1, T2 or T3 tetrazines (dC^{XTY}TP) by IEDDA cycloaddition.

The reactivity of all functionalized nucleoside triphosphates (**dC^{2TCO}TP**, **dC^{4TCO}TP**, **dC^{p2TCO}TP**, **dC^{p4TCO}TP**, and **dC^{pBCN}TP**) with an excess of tetrazine **T1**–**T3** (5 equiv.) was tested *in vitro*. (**Scheme 5**). Specifically 5 µL of 10 µM **dC^{X}TPs** aqueous solution were combined with 5 µL of 100 µM tetrazines (**T1**–**T3**) dissolved in DMSO. The reactions with **T1** or **T2** proceeded at 37 °C for 30 min, while the reaction with **T3** was conducted for 18 h. The identity of the **dC^{X}TP** controls and **dC^{XTY}TP** click products was verified only by analytical LC-MS, they were not synthetized or isolated in semipreparative scale. Additionally, from MS spectra addition of water to the click product (**dC^{XTY}TP**) was observed (an example shown in **Figure 26**, other spectra are shown in **Figures 86-92** (Appendix, section 7.1).

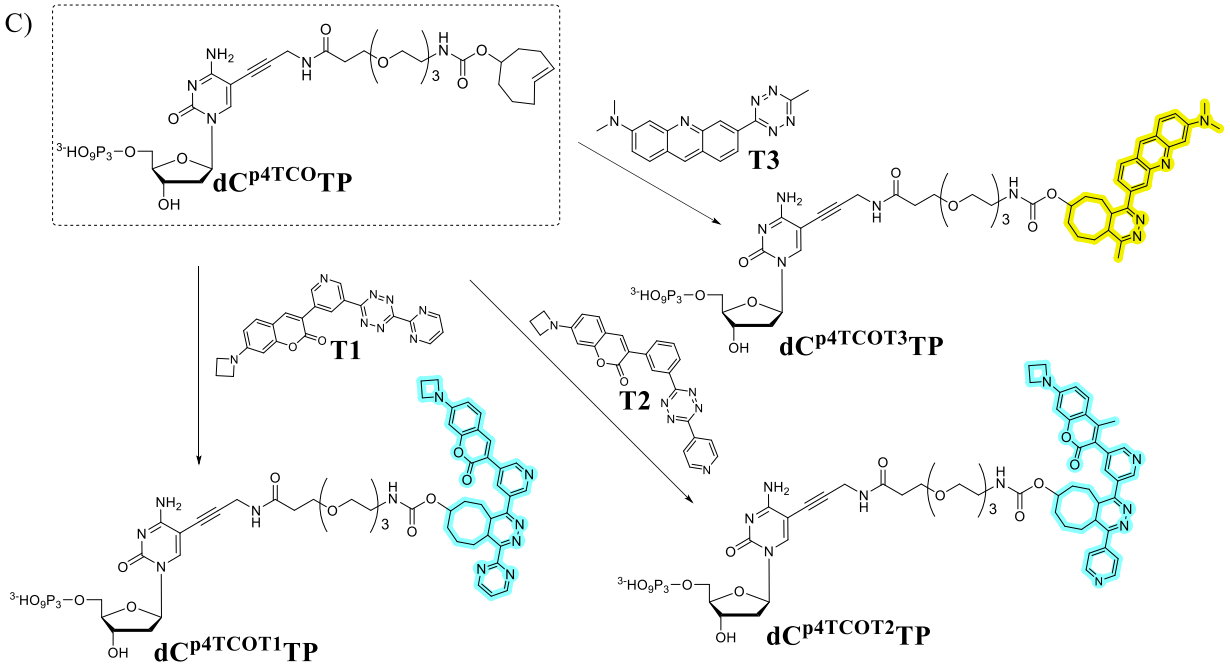
A)



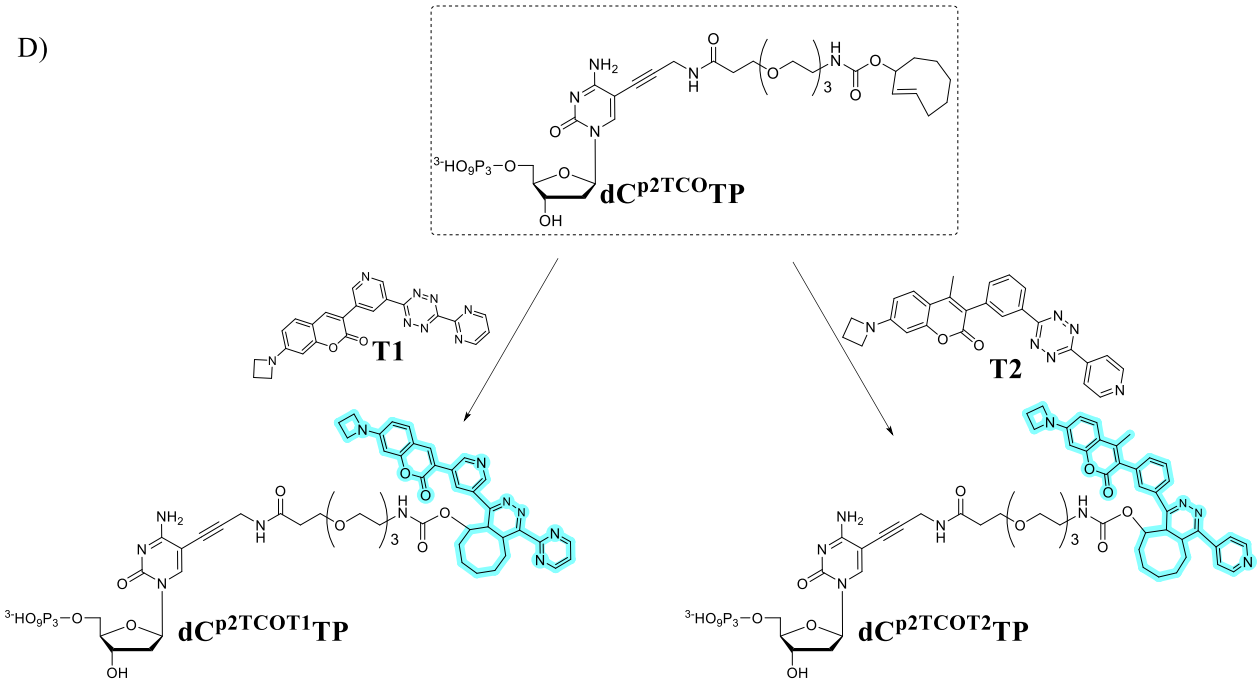
B)

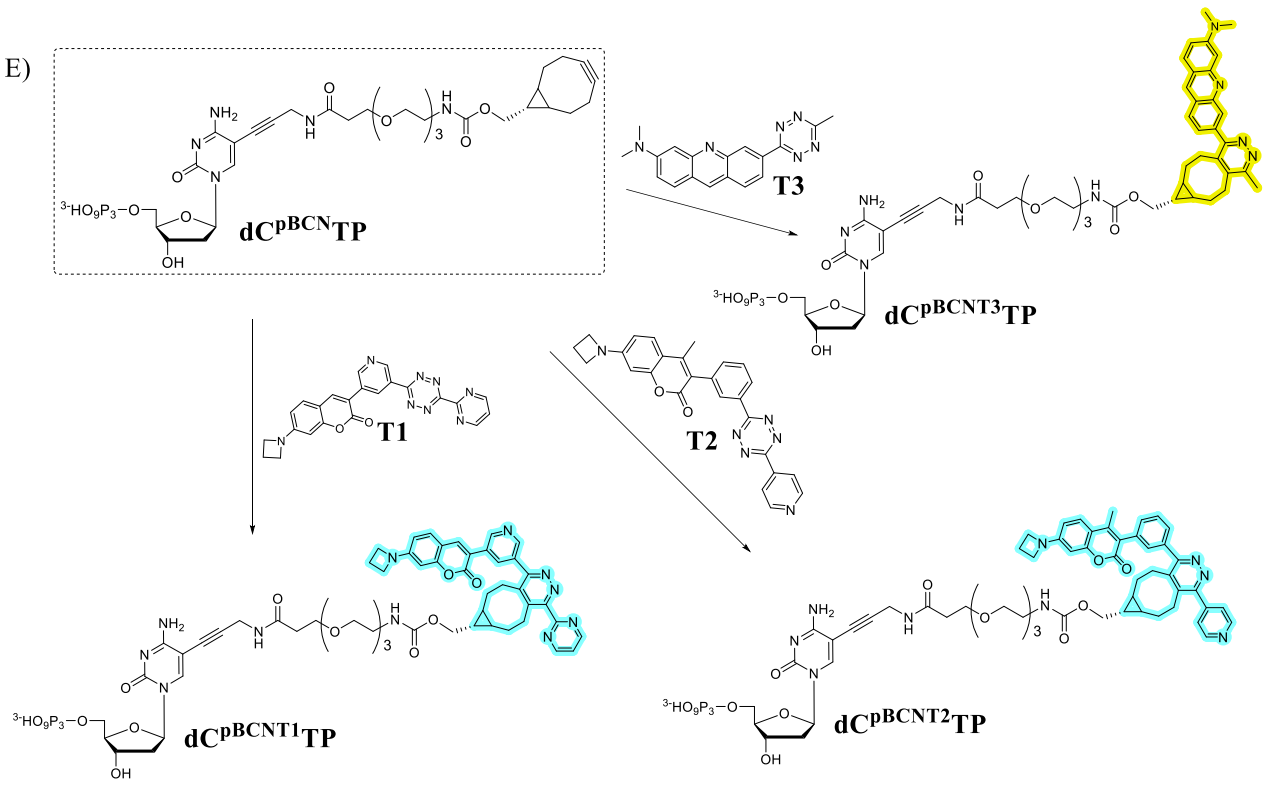


C)


dC^p4TCOTP
T3
dC^p4TCOT3TP
T1
T2
dC^p4TCOT1TP
dC^p4TCOT2TP

D)


dC^p2TCOTP
T1
T2
dC^p2TCOT1TP
dC^p2TCOT2TP



Scheme 5. Reaction scheme of the click product formed in the reaction of A) **dC^{4TCO}TP** or B) **dC^{2TCO}TP** with **T1** or **T2**, C) **dC^{p4TCO}TP** or D) **dC^{p2TCO}TP** or E) **dC^{pBCN}TP** with **T1**, **T2**, **T3** tetrazines. Only one isomer of the click product is shown for clarity.

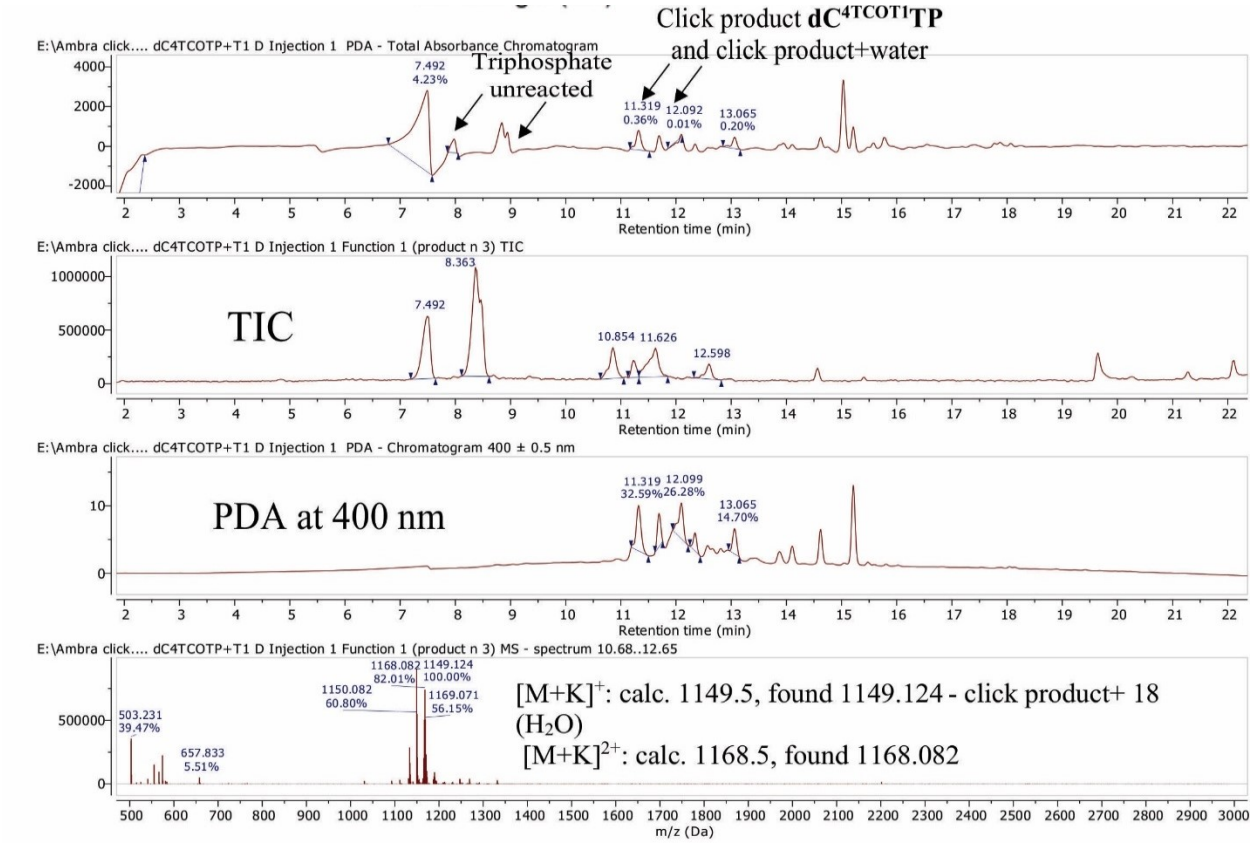


Figure 26. LC-MS characterization of **dC^{4TCOT1}TP** click product.

The outcomes of reactivity were lower conversions for dNTPs containing the shorter propargyl-carbamate linker (27–78%) and better for the compounds bearing the longer PEG3-based linker (55–90%). Moreover, the reactivity of tetrazine **T2** was lower than that of tetrazines **T1** and **T3**. Estimated formation (%) of **dC^{XTY}TP** click products are shown in (**Table 1)** and calculated from the peak areas by Mestrenova software.

Table 1. Estimated conversions of **dC^{XTY}TP** click products.

Click product code (dC^{XTY}TP)	Conversion (%)
dC^{4TCOT1}TP	58%
dC^{4TCOT2}TP	27%
dC^{2TCOT1}TP	78%
dC^{2TCOT2}TP	49%
dC^{p4TCOT1}TP	94%
dC^{p4TCOT2}TP	95%
dC^{p4TCOT3}TP	88%
dC^{p2TCOT1}TP	90%
dC^{p2TCOT2}TP	55%
dC^{pBCNT1}TP	75%
dC^{pBCNT2}TP	74%
dC^{pBCNT3}TP	73%

3.1.1.5 Fluorescence "turn-on" measurements and LC-MS characterizations of dC^{XTY}TP click products

Subsequently, after characterization of click conjugates **dC^{XTY}TP**, their kinetics formation and fluorescence "turn-on" were measured by fluorescence spectroscopy. "Turn on" is the increase of fluorescence intensity resulting from the difference of the fluorescence intensities exhibited by the products and the starting materials. The reactions of **T1** with TCO- and BCN-linked nucleotides were performed by mixing equimolar amounts of reactants in quartz cuvettes to achieve their final 1 µM concentration in phosphate-buffered saline (PBS). In particular 1 µL of the tetrazine stock solution (1 mM in DMSO) was diluted in 1 mL PBS buffer (1 X, pH 7.4, giving 1 µM final concentration) and the fluorescence was measured using 370 nm excitation (slit 3) and 400-650 nm for emission after ca. 30 minutes.

To this solution was added solution of the triphosphates (**dC^{4TCO}TP** or **dC^{2TCO}TP** or **dC^{p4TCO}TP** or **dC^{p2TCO}TP** or **dC^{pBCN}TP)**: 1 µL from 1 mM fresh stock in water, giving 1 µM final concentration, 1 equiv. to tetrazine).

The turn-on values were calculated from the observed fluorescence intensities of the click products (**dC^{XTY}TP**) at the emission maximum (485 nm) after ca. 15 min of reaction divided by the highest residual fluorescence of the quenched coumarin tetrazine dye. Additionally, all spectra were subtracted from the baseline (PBS buffer as the blank).

The largest turn-on effects, 210- and 220-fold, were observed with **dC^{4TCO}TP** and **dC^{p4TCO}TP**, respectively. The analogous 2-TCO isomers, **dC^{2TCO}TP** and **dC^{p2TCO}TP**, showed somewhat diminished light-up effect: 130- and 125-fold, respectively (**Table 2**). These results indicate that the apparent light-up effect measured at 30 min arbitrary time point was significantly higher for the 4-TCO nucleotides and that the length of the linker did not affect the observed light-up effect. The analogous IEDsection

DA reactions with tetrazine **T2** gave only a 3–5-fold increase of fluorescence, a 5-fold light-up effect was obtained for the reaction of **dC^{4TCO}TP**, whereas a 3-fold increase of fluorescence was observed for **dC^{2TCO}TP** and **dC^{pBCN}TP** (**Table 2**). Due to the lower reactivity of the **T2** tetrazine the reaction mixtures were diluted in a smaller volume of 8 µL PBS (instead of previously used 1 µL) and incubated in the dark at room temperature for 1 h. In case of triphosphates with short linker **dC^{2TCO}TP** and **dC^{4TCO}TP** was observed slightly higher fluorescence increase with 40 and 100 fold, respectively.

Additionally, acridine orange-tetrazine "PINK" (**T3**) tetrazine111 was tested and according to its intercalating behavior on DNA should show higher reactivity towards TCO- modified DNA but in our study was the opposite.

The conjugates of **dC^{p4TCO}TP** and **dC^{pBCN}TP** with **T3** showed measurable light-up with. The reaction was conducted at a concentration of 1 mM (for both components) for 2 h at 37 °C, due to lower reactivity of **T3**, giving low fluorescence enhancement with light-up factors of 15 and 50 for **dC^{p4TCO}TP** and **dC^{pBCN}TP**, respectively (**Table 2**). In this latest case the fluorescence was measured using 500 nm excitation wavelength and 525-750 nm emission window (slit 3) due to acridine dye presence taking part to **T3** chemical structure. All fluorescence spectra are shown in the **Figures 86-92** (Appedix, Section 7.1).

Table 2. Summary of the turn-on values for the click products (**dC^{XTY}TP**)

Tetrazine	click products	Fluorescence enhancement[a]
T1	**dC^{4TCOT1}TP**	210-fold
T2	**dC^{4TCOT2}TP**	100-fold
T1	**dC^{2TCOT1}TP**	130-fold
T2	**dC^{2TCOT2}TP**	40-fold
T1	**dC^{p4TCOT1}TP**	220-fold
T2	**dC^{p4TCOT2}TP**	5-fold
T3	**dC^{p4TCOT3}TP**	15-fold
T1	**dC^{p2TCOT1}TP**	125-fold
T2	**dC^{p2TCOT2}TP**	3-fold
T1	**dC^{pBCNT1}TP**	15-fold
T3	**dC^{pBCNT3}TP**	50-fold

[a]determined in PBS buffer (pH 7.4) using 370 nm excitation and calculated by dividing the fluorescence value at 485 nm for click product and that of the quenched **T1** or **T2** tetrazines. For **T3** tetrazine was used 500 nm excitation wavelength.

3.1.1.6 Kinetic formation of click products dC^{XTY}TP

Kinetic analysis of **dC^{p2TCO}TP**, **dC^{p4TCO}TP**, and **dC^{pBCN}TP** with **T1** was performed to reveal the differences in rate constants, especially, for 4-TCO versus 2-TCO derivatives.

Based on previous experimental observation the outcome was significantly different in rate constants between two derivatives series; **dC^{p4TCO}TP** displays a reaction rate (k_{obs} = 0.00736 s^{-1}) one order of magnitude higher than that of **dC^{p2TCO}TP** (0.00071 s^{-1}) (**Figure 27**). In order to have a comparison how the fluorescence intensity increase according to the products at the same concentration, fluorescence intensities at half-life of reactions were calculated (**Table 3**). From their comparison was revealed that both products of reactions of **T1** with **dC^{p2TCO}TP** and **dC^{p4TCO}TP** have comparable fluorescence intensities, indicating that the lower turn-on value observed for **dC^{p2TCO}TP** at the 30 min time point is excusively due to the lower reaction rate.

Furthermore, the reaction of **dC^{pBCN}TP** with **T1** exhibits a lower reaction rate (0.0021 s^{-1}), and the product displays a significantly lower brightness than the TCO products at half-life.

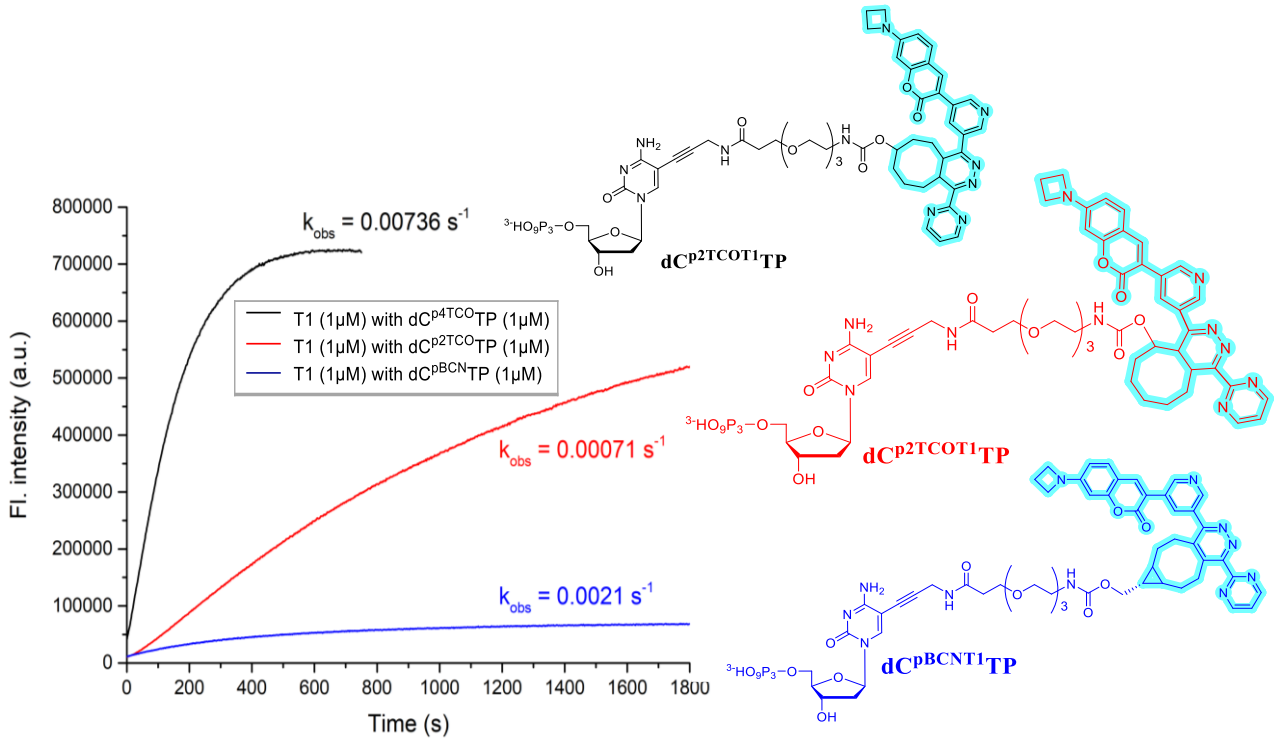


Figure 27. Fluorescence time-lapse measurements of **dC^{p4TCO}TP** or **dC^{p2TCO}TP** or **dC^{pBCN}TP** with **T1** tetrazine showing changes in the fluorescence signal of the click products in time. The observed rate constants were calculated by fitting the curves with single exponential equation using the Origin Pro software.

	Fl intensity dC^{p4TCOT1}TP (a.u.)	**Fl intensity dC^{p2TCOT1}TP** (a.u.)	**Fl intensity dC^{pBCNT1}TP** (a.u.)
$t_{1/2}$ T1 with **dC^{p4TCO}TP** = 135 s	422000	-	-
$t_{1/2}$ T1 with **dC^{p2TCO}TP** = 1408 s	-	460000	-
$t_{1/2}$ T1 with **dC^{pBCN}TP** = 475 s	-	-	48600

Table 3. Fluorescence intensities at half-life of reactions between **dC^{p4TCO}TP** or **dC^{p2TCO}TP** or **dC^{pBCN}TP** and **T1**.

3.1.1.7 Enzymatic incorporation of dC^{2TCO}TP, dC^{4TCO}TP, dC^{p2TCO}TP, dC^{p4TCO}TP, dC^{pBCN}TP into DNA by PEX and MALDI analysis

All five modified dNTPs (**dC^{2TCO}TP**, **dC^{4TCO}TP**, **dC^{p2TCO}TP**, **dC^{p4TCO}TP**, and **dC^{pBCN}TP**) were tested as substrates for KOD XL DNA polymerase in primer extension (PEX) reaction.

Among the aims of the study was to understand either the polymerase used (KOD XL) would accept these modified triphosphates, used sequences in **Table 4**. In all cases the reaction mixtures were incubated for 20 min at 60 °C in a thermomixer. Efficient incorporation of each of the modified nucleotides into full-length 19-mer double-stranded DNA products containing one modification was observed as shown by PAGE analysis (**Figures 28-29**).

Table 4. List of primer and templates used and nucleic acids and ONs synthetised in this study

Name	Sequence (5'→3')	Length
Prim A	5'-CATGGGCGGCATGGG-3'	15 nt
Prim A_FAM [a]	5'-CATGGGCGGCATGGG-3'	15 nt
Temp19_1C	5'-CCCGCCCATGCCGCCCATG-3'	19 nt
Temp19_1C_bio [b]	5'-CCCGCCCATGCCGCCCATG-3'	19 nt
Temp19_1C_2xbio [c]	5'-CCCGCCCATGCCGCCCATG-3'	19 nt
19DNA_C^{natural} [a]	5'-CATGGGCGGCATGGG**C**GGG-3'	19 nt
19DNA_C^{4TCO} [a]	5'-CATGGGCGGCATGGG**C^{4TCO}**GGG-3'	19 nt
19DNA_C^{2TCO} [a]	5'-CATGGGCGGCATGGG**C^{2TCO}**GGG-3'	19 nt
19DNA_C^{p4TCO} [a]	5'-CATGGGCGGCATGGG**C^{p4TCO}**GGG-3'	19 nt
19DNA_C^{p2TCO} [a]	5'-CATGGGCGGCATGGG**C^{p2TCO}**GGG-3'	19 nt
19DNA_C^{pBCN} [a]	5'-CATGGGCGGCATGGG**C^{pBCN}**GGG-3'	19 nt
19ON_C^{4TCO}	5'-CATGGGCGGCATGGG**C^{4TCO}**GGG-3'	19 nt
19ON_C^{2TCO}	5'-CATGGGCGGCATGGG**C^{2TCO}**GGG-3'	19 nt
19ON_C^{p4TCO}	5'-CATGGGCGGCATGGG**C^{p4TCO}**GGG-3'	19 nt
19ON_C^{p2TCO}	5'-CATGGGCGGCATGGG**C^{p2TCO}**GGG-3'	19 nt
19ON_C^{p2TCO} [a]	5'-CATGGGCGGCATGGG**C^{p2TCO}**GGG-3'	19 nt
19ON_C^{pBCN}	5'-CATGGGCGGCATGGG**C^{pBCN}**GGG-3'	19 nt
19ON_C^{pBCN} [a]	5'-CATGGGCGGCATGGG**C^{pBCN}**GGG-3'	19 nt

[a] 5`-(6-FAM)-labelled, [b] 5'-biotinylated; [c] 5'double-biotinylated; ONx-single-stranded DNA; DNAx-double-stranded DNA, the complementary strand is unmodified; primer regions underlined.

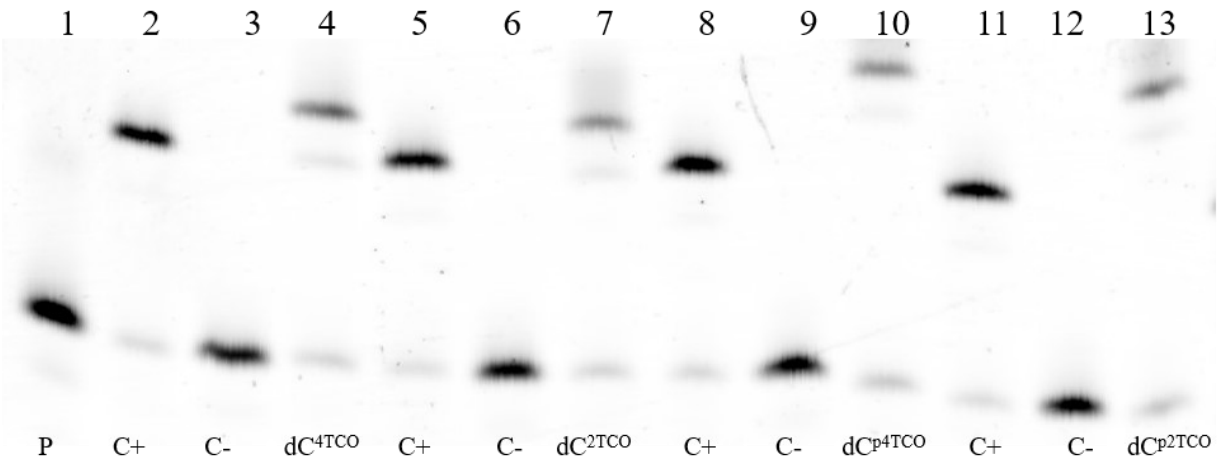


Figure 28. Denaturing PAGE analysis of PEX using KOD XL DNA polymerase and template **Temp19_1C** (lane 2-13). (P): primer **Prim A_FAM**, lane 1; (C+): natural dGTP, dCTP, lanes 2,5,8,11; (C-): negative control without dCTP, lanes 3,6,9,12; (dC^{4TCO}): modified DNA containing **dC^{4TCO}TP**, dGTP, lane 4; (dC^{2TCO}TP): modified DNA containing **dC^{2TCO}TP**, dGTP, lane 7; (dC^{p4TCO}TP): modified DNA containing **dC^{p4TCO}TP**, dGTP, lane 10; (dC^{p2TCO}TP): modified DNA containing **dC^{p2TCO}TP**, dGTP, lane 13.

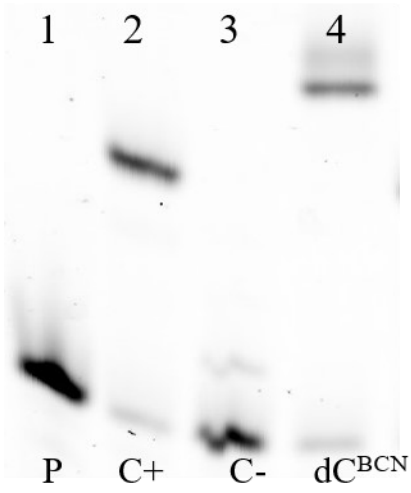


Figure 29. Denaturing PAGE analysis of PEX using KOD XL DNA polymerase and template **Temp19_1C** (lane 2-4). (P): primer **Prim A_FAM**, lane 1; (C+): natural dGTP, dCTP lane 2; (C-): negative control without dCTP, lane 3; (dC^{pBCN}): modified DNA containing **dC^{pBCN}TP**, dGTP lane 4.

The magnetoseparation (for procedure see section 5.1.6) was applied to isolate the modified oligonucleotides **19ON_C^{X}** which were characterized by MALDI-TOF (**Table 5**).

Table 5. Overview of dC^{4TCO} or dC^{2TCO} or dC^{p4TCO} or dC^{p2TCO} or dC^{pBCN}-modified nucleic acids or oligonucleotides and their masses

Nucleic acids	Mw calculated [Da]	Mw found [Da]	Δ [Da]
19ON_ C^{4TCO}	6156.0	6157.2 [b]	1.2
19ON_C^{2TCO}	6156.0	6159.3 [b]	3.3
19ON_C^{p4TCO}	6359.1	6360.1 [b]	1.0
19ON_C^{p2TCO}	6359.1	6360.6 [b]	1.5
19ON_C^{pBCN}	6383.1	6384.6 [b]	1.5
19ON_C^{pBCN} [a]	6920.1	6918.4	1.7
19DNA_C^{pBCN} [a]	6920.1	6921.3 [b]	1.2
19ON_C^{p4TCO} [a]	6896.1	6896.1	0
19DNA_C^{p4TCO} [a]	6896.1	6897.6 [b]	1.5

[a] 5`-(6-FAM)-labelled, ON^{X}-single stranded; DNA^{X}-double-stranded DNA
[b] Nucleic acids measured in positive mode.

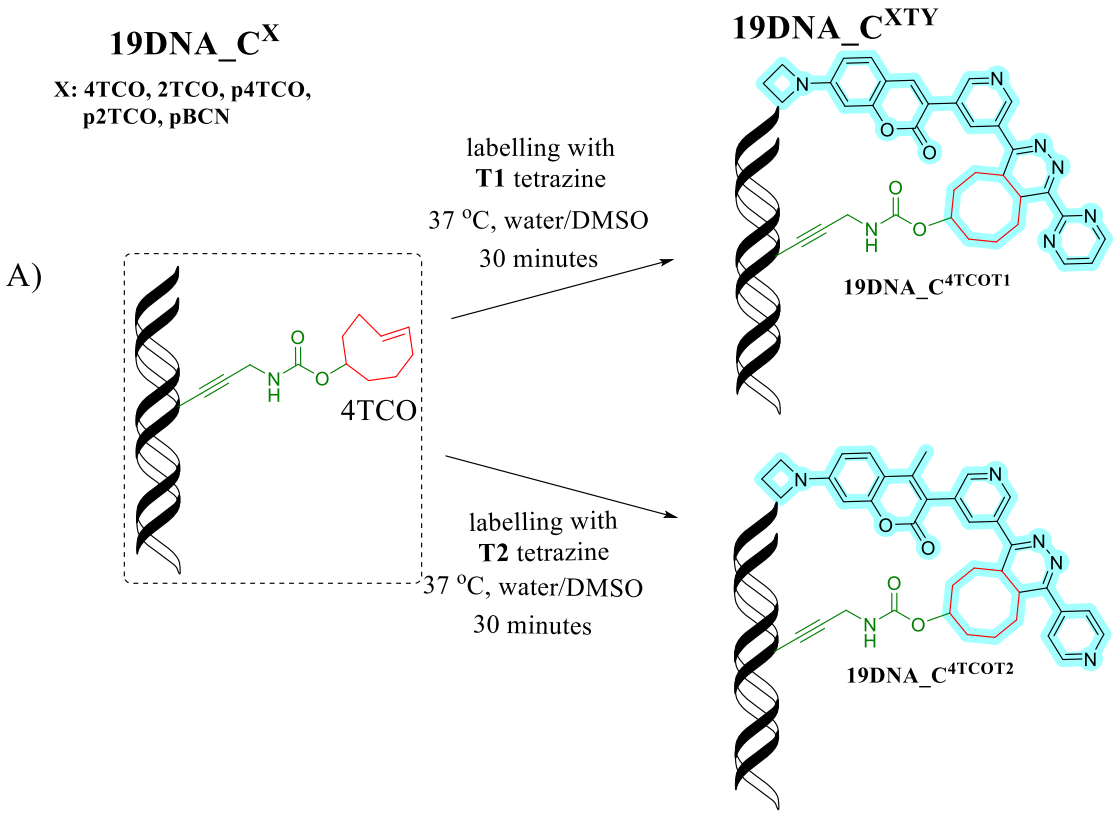
3.1.1.8 IEDDA reactions of 19DNA_C^{4TCO} or 19DNA_C^{2TCO} or 19DNA_C^{p4TCO} or 19DNA_C^{p2TCO} or 19DNA_C^{pBCN} with tetrazines, MALDI or LC-MS analysis of ON-click products (19ON_C^{XTY})

3.1.1.8.1 Post-synthetic labelling of 19DNA_C^{X} (X = 4TCO or 2TCO or p4TCO or p2TCO or pBCN) with T1 or T2 tetrazines and of 19DNA_C^{X} (X = p4TCO or pBCN) with T3 or T4 tetrazines

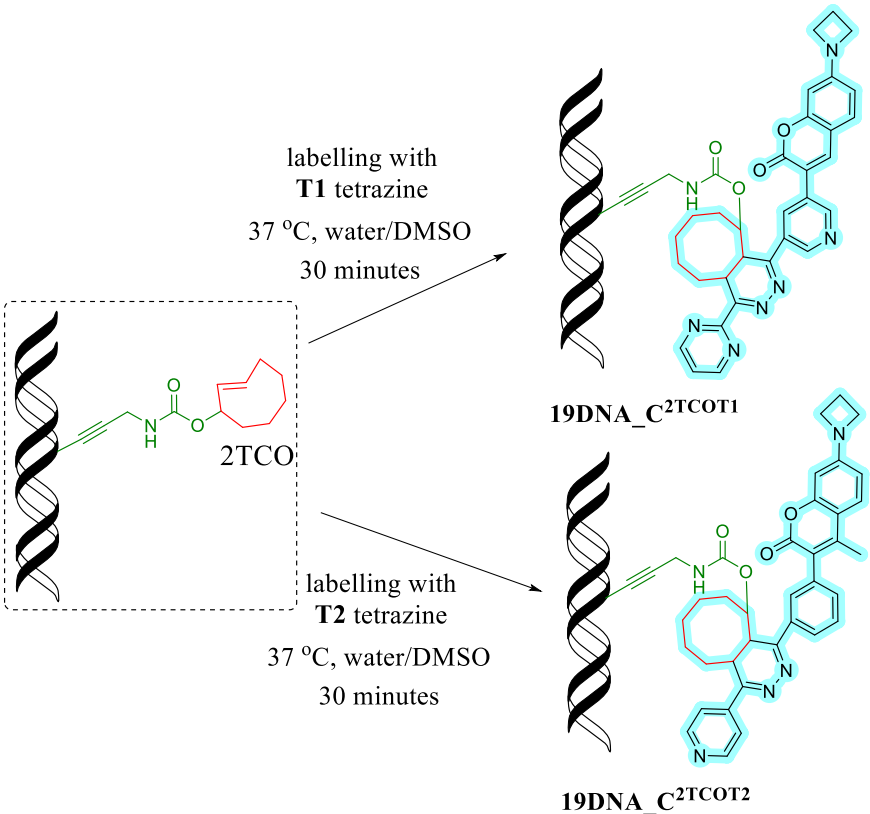
After successfully incorporation of modified triphosphates **dC^{X}TPs** in DNA were synthesized the correspondent modified **19DNA_C^{X}**. As next step, the reactivity of the reactive functionalized 19- mer dsDNA towards tetrazines was tested in order to find the best combination in terms of fluorogenicity, *in vitro* firstly (**Scheme 6**), and then in cellulo. Therefore, the PEX products

(0.2 μM) were incubated with excess (5 or 50 μM) of tetrazines **T1**–**T4** in the H_2O–DMSO mixture for 30 min at 37 °C, and the reaction mixtures were analyzed by PAGE analysis. Excess of tetrazines was used in order to allow the reaction with reactive strained dienophiles before their degradation, due to the instability of the latter in solution. In order to compare the reactivity of modified triphosphates (**dC^{X}TPs**) or DNA (**19DNA_C^{X}**) with tetrazines, the same combinations attempted for **dC^{X}TPs** were tried as well for modified **19DNA_C^{X}**, which means **19DNA_C^{X}** (X = **4TCO** or **2TCO** or **p4TCO** or **p2TCO** or **pBCN**) with **T1** or **T2** tetrazines and **19DNA_C^{X}** (X = **p4TCO** or **pBCN**) with **T3** tetrazine. Based on previously obtained data about triphosphates with short linker **dC^{4TCO}TP** and **dC^{2TCO}TP** showed lower reactivity with **T1** and **T2** tetrazines compared to triphosphates serie containing PEG3 spacer, consequently was evalued not worthy to proceed neither in case of modified **19DNA_C^{4TCO}** or **19DNA_C^{2TCO}** with **T3**.

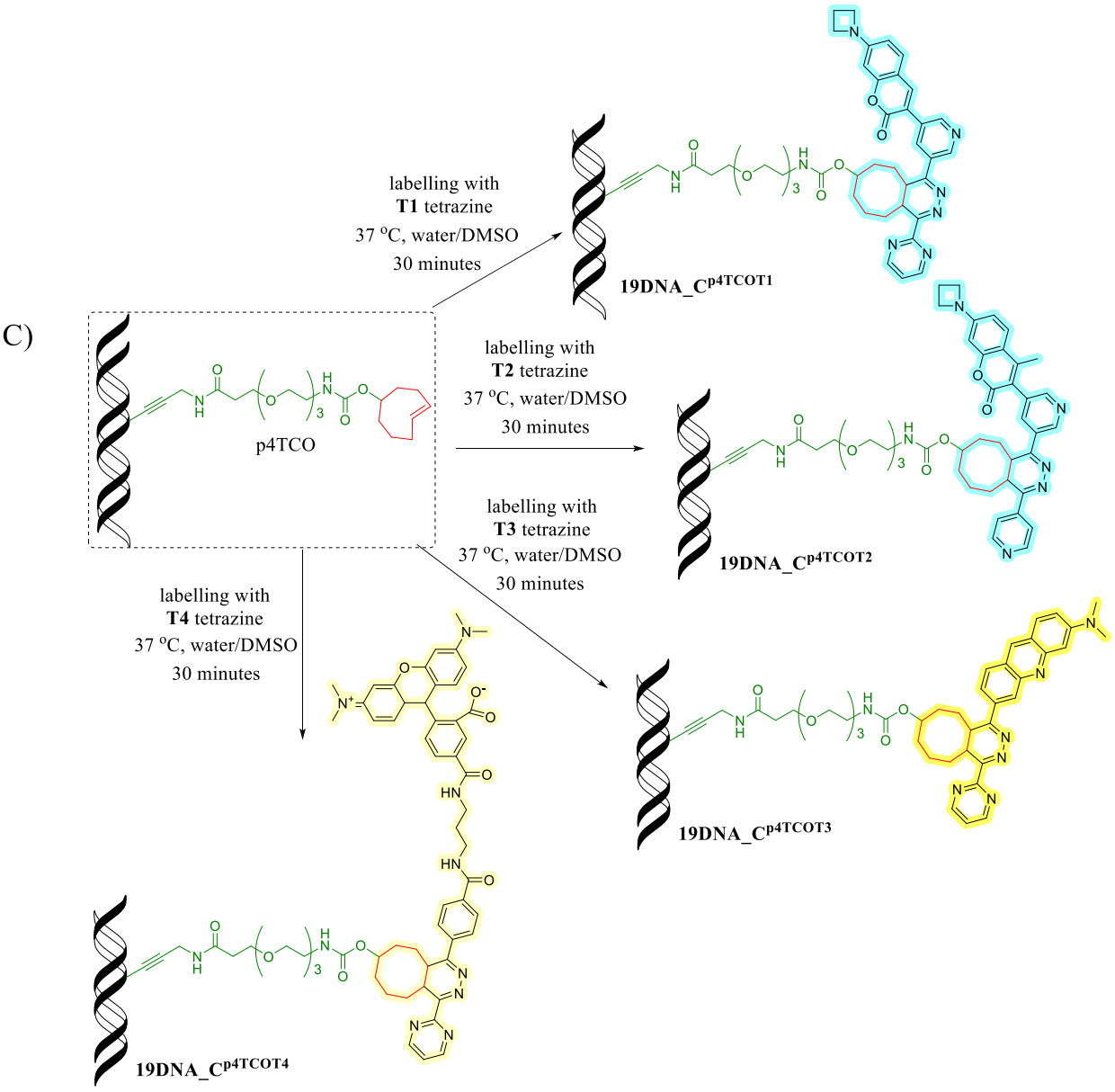
The reactivity of **19ON_C^{p4TCO}** or **19ON_C^{pBCN}** was tested with **T4** but is not crucial in terms of fluorogenicity due to rhodamine dye present in its structure, as conquence was not further used in combination with other modified oligonucleotide (**19ON_C^{X}**).

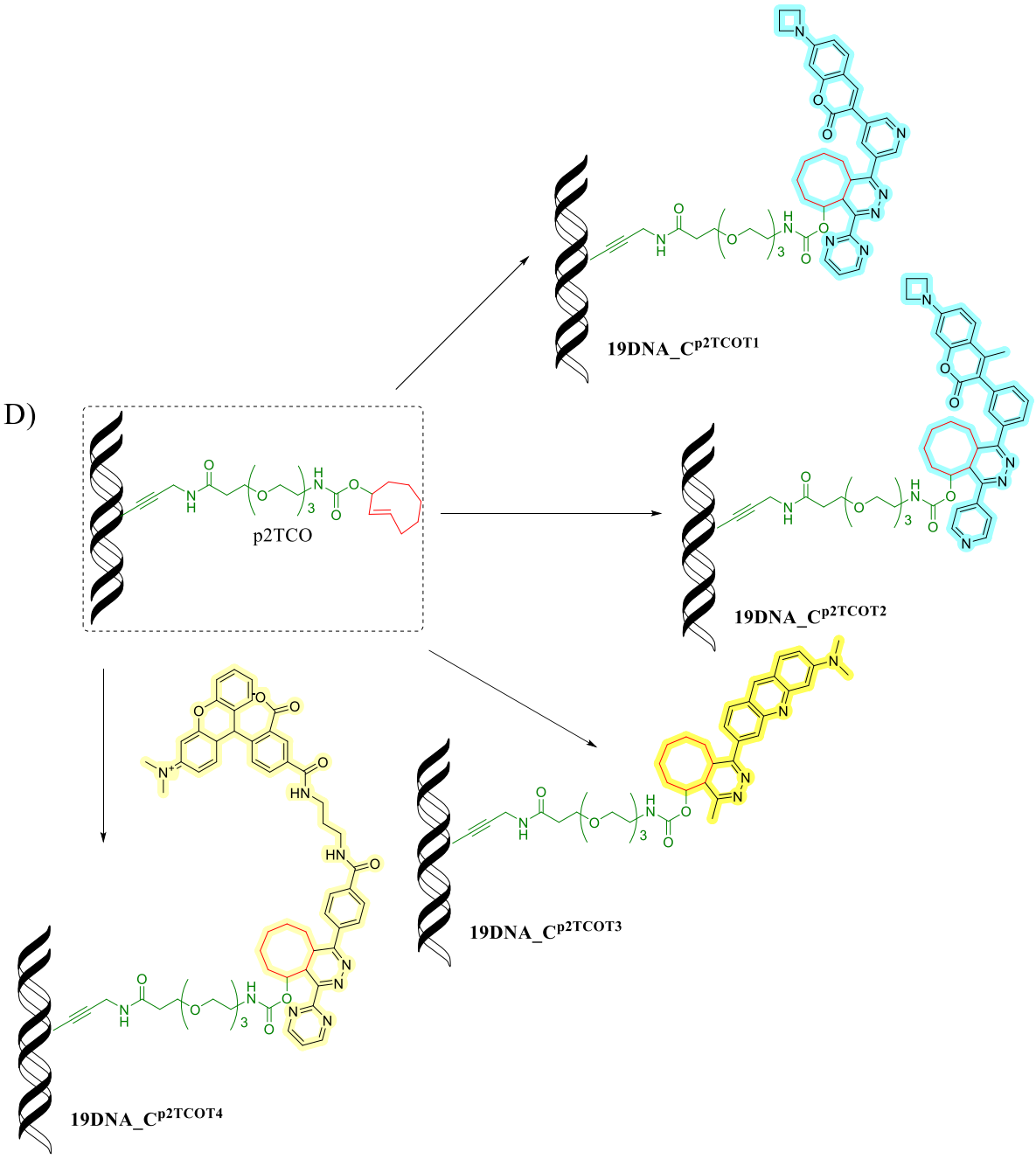


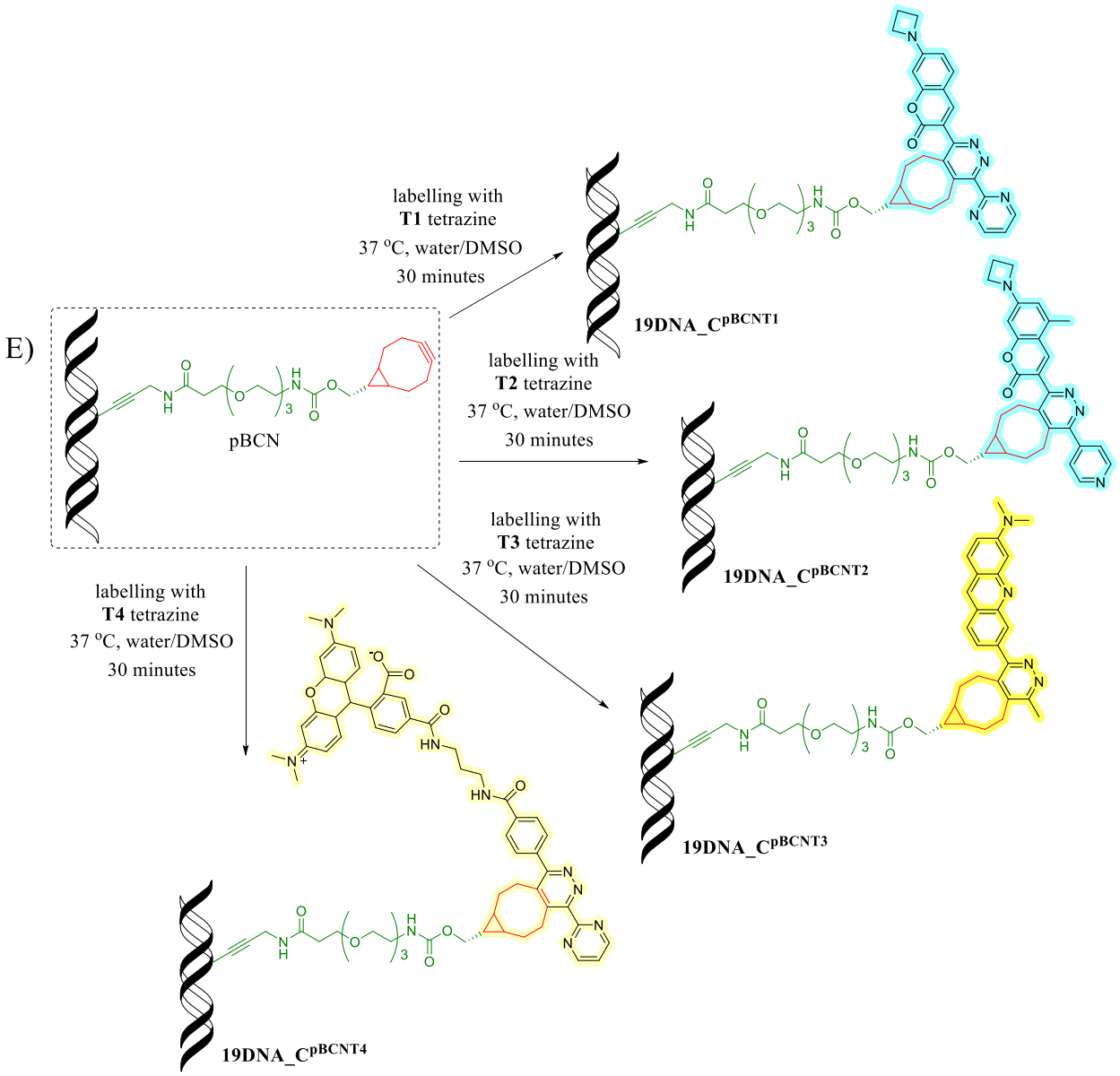
B)


labelling with
T1 tetrazine
37 °C, water/DMSO
30 minutes
2TCO
19DNA_C2TCOT1
labelling with
T2 tetrazine
37 °C, water/DMSO
30 minutes
19DNA_C2TCOT2

C)


labelling with
T1 tetrazine
37 °C, water/DMSO
30 minutes
19DNA_C^p4TCOT1
labelling with
T2 tetrazine
37 °C, water/DMSO
30 minutes
p4TCO
19DNA_C^p4TCOT2
labelling with
T3 tetrazine
37 °C, water/DMSO
30 minutes
labelling with
T4 tetrazine
37 °C, water/DMSO
30 minutes
19DNA_C^p4TCOT3
19DNA_C^p4TCOT4


D)
p2TCO
19DNA_C^p2TCOT1
19DNA_C^p2TCOT2
19DNA_C^p2TCOT3
19DNA_C^p2TCOT4



Scheme 6. Reaction of modified DNA A) **19DNA_C^{4TCO}** or B) **19DNA_C^{2TCO}** or C) **19DNA_C^{p4TCO}** or D) **19DNA_C^{p2TCO}** or E) **19DNA_C^{pBCN}** with **T1**, **T2**, **T3**, **T4** tetrazines. Only one isomer of the click product is shown for clarity.

All TCO-modified DNAs reacted with **T1**, forming the isomeric mixtures of dihydropyridazine conjugates. Compared to **19DNA_C^{4TCO}** and **19DNA_C^{2TCO}** with short linkers which gave 42–64% conversions, the PEG3-linked TCO-derivatives **19DNA_C^{p4TCO}** and **19DNA_C^{p2TCO}** have given higher conversions (81–90%, **Figure 30** lanes 9,10,12,13) which were calculated according to intensity bands by ImageJ software. In case of **19DNA_C^{4TCO}** and **19DNA_C^{2TCO}** (**Figure 30** lanes

3,4,6,7) increasing the labeling reaction time (1 hour) did not improve the conversions (%) due to possible degradation or isomerization of strained dienophile moiety on DNA to cis form, unreactive towards tetrazines. This result indicates the positive role of a longer linker in reactivity towards tetrazines and their higher stability to not isomerize. Also, the high reactivity of the strained dienophiles and lower stability are related[112], which could explain why in case of **19DNA_C^{2TCO}** with short linker was observed slightly better conversion than for **19DNA_C^{4TCO}**, more reactive but at the same time more unstable. Moreover, the BCN-linked **19DNA_C^{pBCN}** was efficiently labeled with **T1** (83-86%, **Figure 30**).

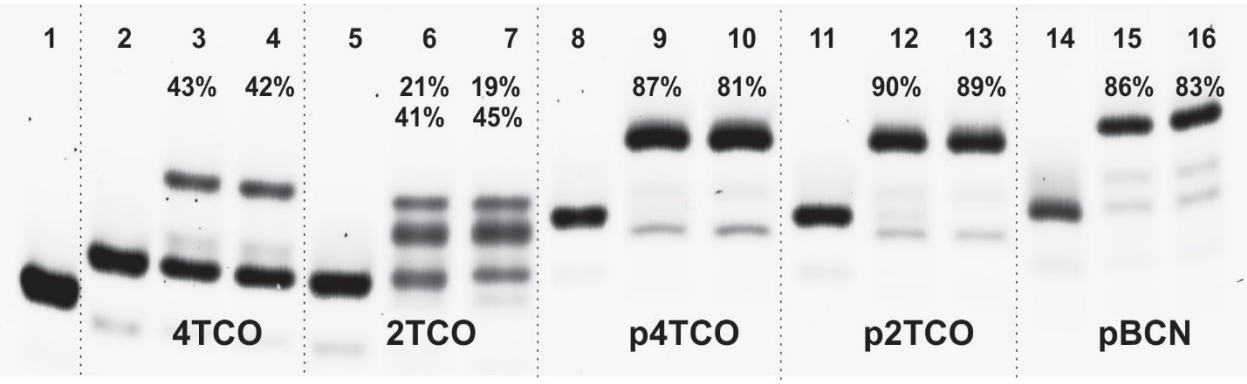
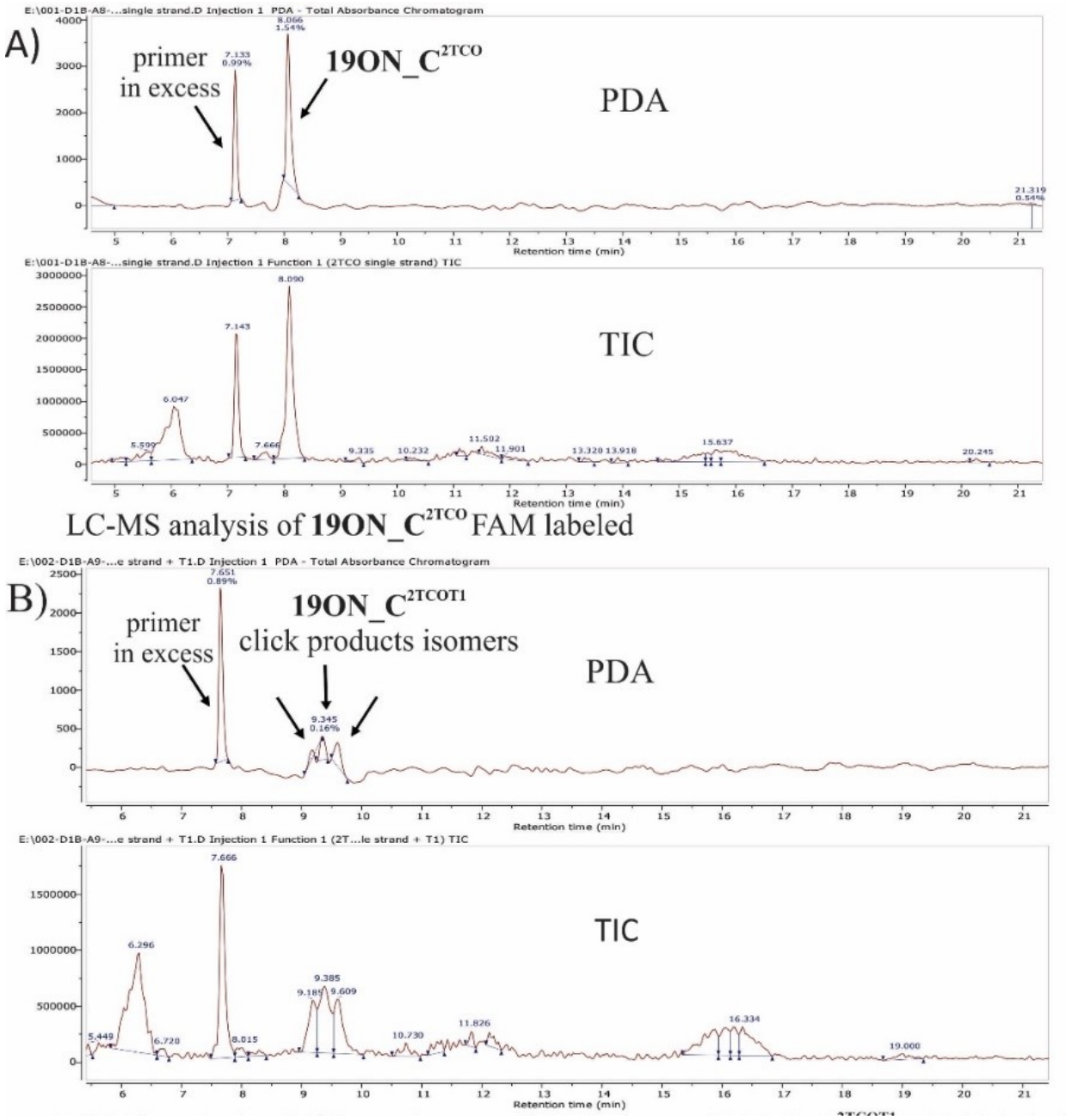


Figure 30. Denaturing PAGE analysis of tetrazine-DNA click reaction of **19DNA_C^{4TCO}** with **T1** (5 μM) lane 3 or **T1** (50 μM) lane 4; of **19DNA_C^{2TCO}** with 5 μM **T1** (lane 6) or 50 μM **T1** (lane 7); of **19DNA_C^{p4TCO}** with 5 μM **T1** (lane 9) or 50 μM **T1** (lane 10); of **19DNA_C^{p2TCO}** with 5 μM **T1** (lane 12) or 50 μM **T1** (lane 13); of **19DNA_C^{pBCN}** with 5 μM **T1** (lane 15) or 50 μM **T1** (lane 16). Reaction conditions: **19DNA_C^{X}** in water and **T1** tetrazine in DMSO incubated at 37 °C, 30 minutes. Negative controls (-): **19DNA^{natural}** lane 1, **19DNA_C^{4TCO}** lane 2, **19DNA_C^{2TCO}** lane 5, **19DNA_C^{p4TCO}** lane 8, **19DNA_C^{p2TCO}** lane 11, **19DNA_C^{pBCN}** lane 14.

After click reaction of **19DNA_C^{2TCO}** with **T1**, multiple bands were observed in **Figure 30** (lanes 6-7). To explain these multiple bands, LC-MS analysis was performed on crude mixture and several isomers of click product were observed which were confirmed by the same mass (**Figure 31**). Next, the reaction mixture (**19ON_C^{2TCO}** after click reaction with **T1)** was purified on semipreparative scale by analytical HPLC (**Figure 32-A**). Several compounds were isolated and analysed on denaturing PAGE gel (**Figure 32-B**).


A)
primer in excess
19ON_C2TCO
PDA
TIC
Retention time (min)
LC-MS analysis of 19ON_C2TCO FAM labeled
B)
primer in excess
19ON_C2TCOT1
click products isomers
PDA
TIC
Retention time (min)

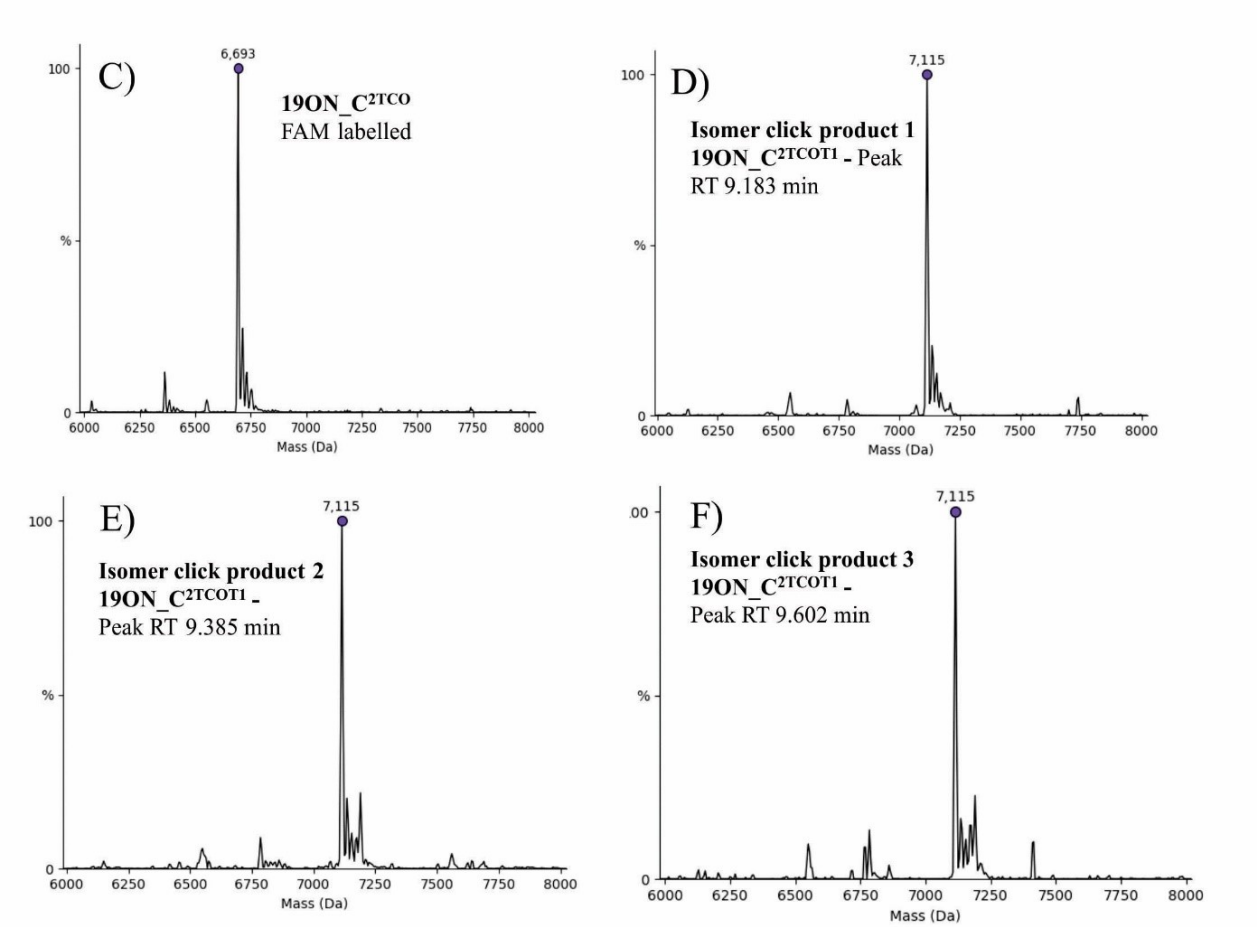


Figure 31. A) LC-MS analysis of **19ON_C^{2TCO}** FAM labelled, B) LC-MS analysis of isomer click products **19ON_C^{2TCOT1}** FAM labelled, C) ESI-MS(-) spectrum of **19ON_C^{2TCO}** FAM labelled, calculated [M] 6693.03 Da; found 6693 Da, D), E), F) ESI-MS(-) spectrum of isomer click products. **19ON_C^{2TCOT1}** FAM labelled, calculated [M] 7115.18 Da; found 7115 Da.

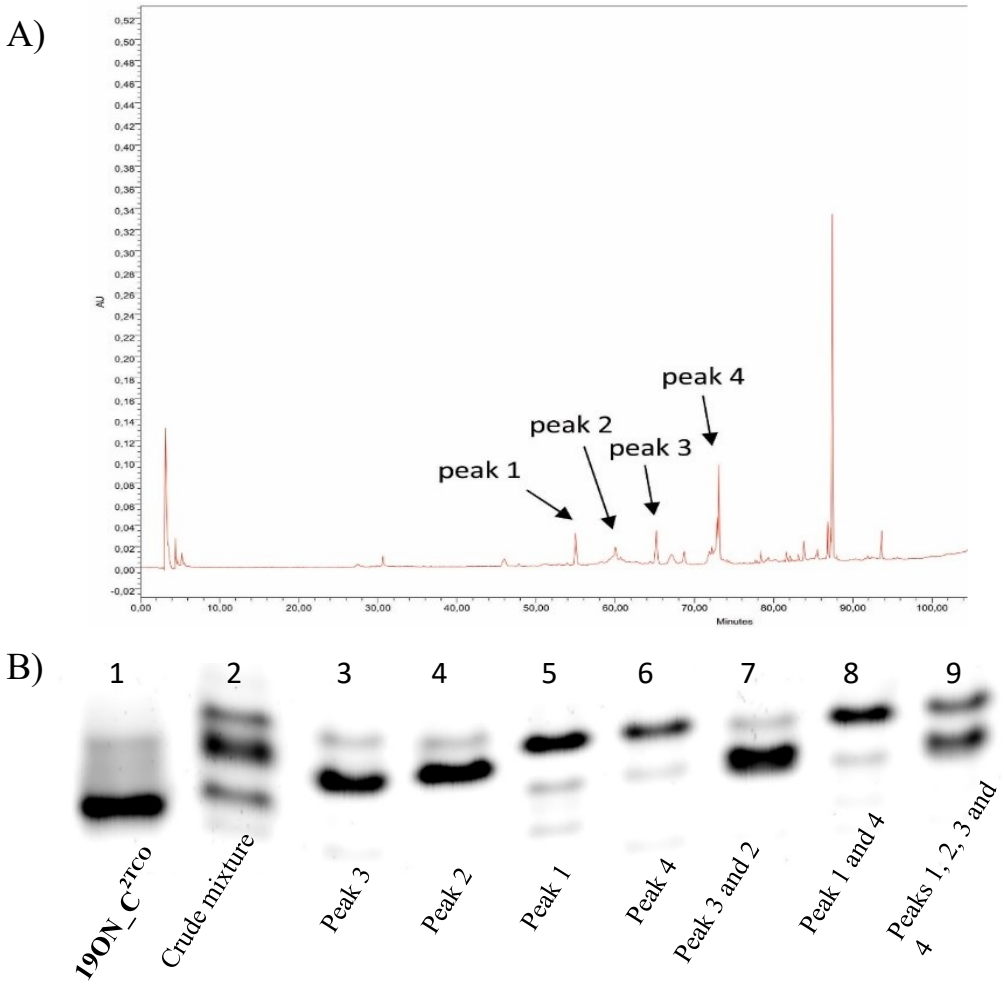
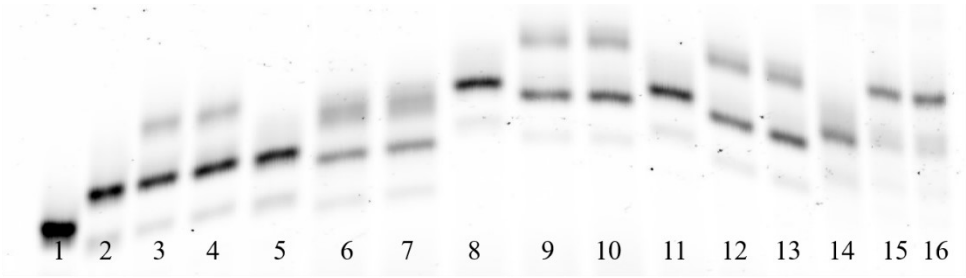


Figure 32. Chromatogram of click reaction between **19ON_C^{2TCO}** semipreparative scale and **T1** after analytical HPLC purification and denaturing PAGE analysis of three isomer click products of **19ON_C^{2TCOT1}**. In line 7 peak 3 and 2 were loaded together and they have shown slightly different mobility shift on gel (not 100 % overlapped). In line 8 peak 1 and 4 have shown the identical mobility on gel. In line 9 all peaks were mixed together.

On the other side, tetrazine **T2** reacted poorly with incomplete conversions in combination with TCO-linked DNA and readily and completely only with BCN-linked **19DNA_C^{pBCN}** (**Figure 33**).



	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
19DNA^{natural}	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
19DNA_C^{4TCO}	-	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-
19DNA_C^{2TCO}	-	-	-	-	+	+	+	-	-	-	-	-	-	-	-	-
19DNA_C^{p4TCO}	-	-	-	-	-	-	-	+	+	+	-	-	-	-	-	-
19DNA_C^{p2TCO}	-	-	-	-	-	-	-	-	-	-	+	+	+	-	-	-
19DNA_C^{pBCN}	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+
T2 (5 µM)	-	-	+	-	-	+	-	-	+	-	-	+	-	-	+	-
T2 (50 µM)	-	-	-	+	-	-	+	-	-	+	-	-	+	-	-	+

Figure 33. Denaturing PAGE analysis of tetrazine-DNA click reaction of **19DNA_C^{4TCO}** with **T2** (10 µM) lane 3 or **T2** (100 µM) lane 4; of **19DNA_C^{2TCO}** with **T2** (10 µM) lane 6 or **T2** (100 µM) lane 7; of **19DNA_C^{p4TCO}** with **T2** (10 µM) lane 9 or **T2** (100 µM) lane 10; of **19DNA_C^{p2TCO}** with **T2** (10 µM) lane 12 or **T2** (100 µM) lane 13; of **19DNA_C^{pBCN}** with **T2** (10 µM) lane 15 or **T2** (100 µM) lane 16. Reaction conditions: **19DNA_C^{X}** in water and **T2** tetrazine in DMSO incubated at 37 °C, 30 minutes. Negative controls (-): **19DNA^{natural}** lane 1, **19DNA_C^{4TCO}** lane 2, **19DNA_C^{2TCO}** lane 5, **19DNA_C^{p4TCO}** lane 8, **19DNA_C^{p2TCO}** lane 11, **19DNA_C^{pBCN}** lane 14.

Concerning **T3** tetrazine, the main disadvantage was the need to have higher concentration (500 µM) and 18 h of incubation at 37 °C to allow the reaction with **19DNA_C^{p4TCO}** and **19DNA_C^{pBCN}** to proceed. Additional bands were also observed in PAGE gel analysis, indicating the presence of some decomposition (**Figure 34**).

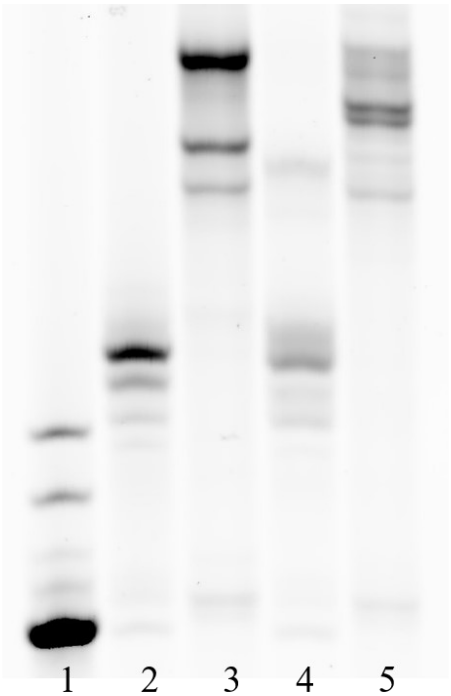


Figure 34. 20% NATIVE PAGE analysis about the formation of tetrazine-DNA click products between **19DNA_C^{p4TCO}** with **T3** tested at 1 mM (lane 3); **19DNA_C^{pBCN}** with **T3** tetrazine tested at 1 mM (lane 5). Negative controls: **19DNA_C^{natural}** (lane 1), **19DNA_C^{p4TCO}** (lane 2), **19ON_C^{pBCN}** (lane 4).

As last, in case of reaction between **T4** tetrazine and **19DNA_C^{p4TCO}** (50 μM of **T4**; 30 min at 37 °C) was observed nearly complete conversion in reaction with **19DNA_C^{p4TCO}** (50 μM of **T4**; 30 min at 37 °C), whereas in presence of **19DNA_C^{p2TCO}** and **19DNA_C^{pBCN}** was shown only partial conversion under these conditions (**Figure 35**).

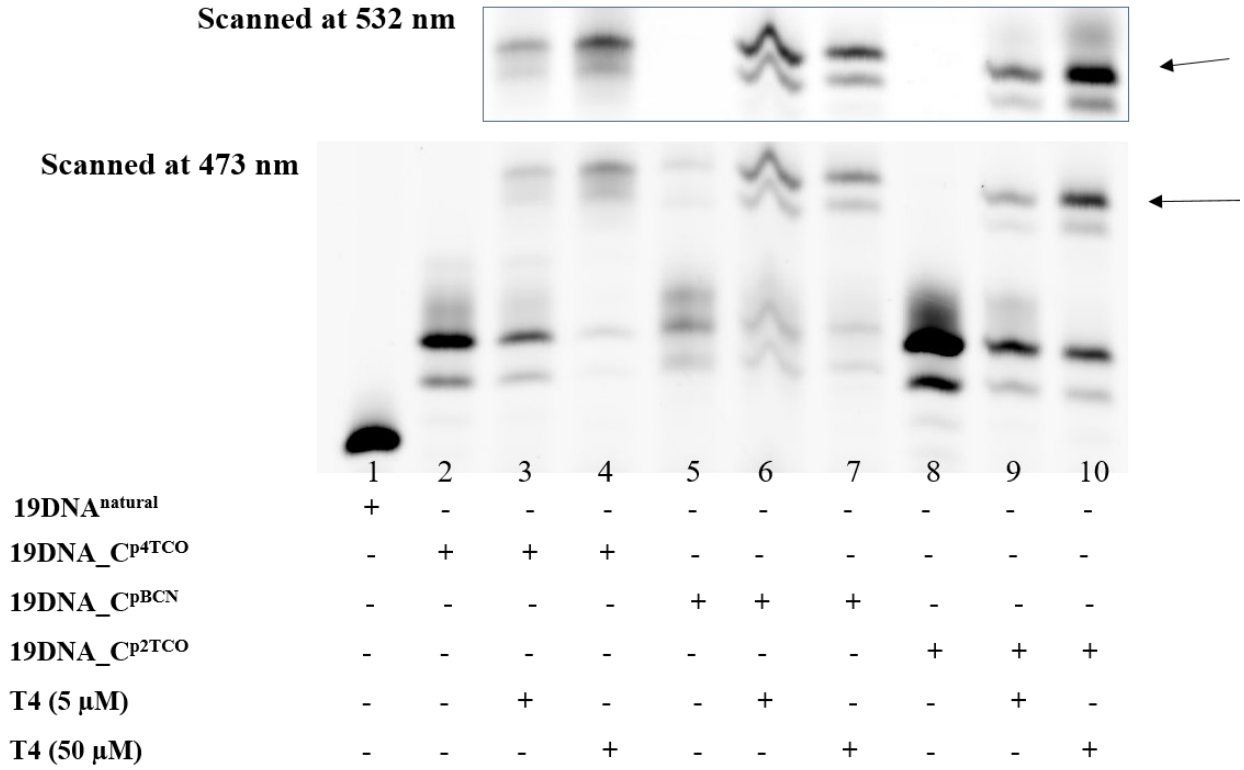


Figure 35. Denaturing PAGE analysis of tetrazine-DNA click reaction of **19DNA_C^{p4TCO}** with **T4** (10 µM) lane 3 or **T4** (100 µM) lane 4; **19DNA_C^{pBCN}** with **T4** (10 µM) lane 6 or **T4** (100 µM) lane 7; of **19DNA_C^{p2TCO}** with **T4** (10 µM) lane 9 or T4 (100 µM) lane 10. Reaction conditions: **19DNA_C^{X}** in water and **T4** tetrazine in DMSO incubated at 37 °C, 30 minutes. Negative controls: **19DNA^{natural}** lane 1, section **19DNA_C^{p4TCO}** lane 2, **19DNA_C^{pBCN}** lane 5, **19DNA_C^{p2TCO}** lane 8. Upper part: scanning at 532 nm=ʎex TAMRA (excitation wavelength) to visualize only click products **19DNA_C^{XTY}**, lower part: scanning at 473 nsem=ʎex FAM, natural and modified DNA are FAM labelled, to visualize click products (**19DNA_C^{XTY}**) and starting materials (**19DNA^{natural}** and **19ON_C^{X}**).

3.1.1.9 MALDI-TOF analysis of ON-click products (19ON_C^{XTY})

The PEX products (**see Table 5**) **19DNA_C^{4TCO}** or **19DNA_C^{2TCO}** or **19DNA_C^{p4TCO}** or **19DNA_C^{p2TCO}** or **19DNA_C^{pBCN}** (0.2 µM) were incubated with excess (5 or 50 µM) of tetrazines **T1**–**T4** in the H_2O–DMSO mixture for 30 min, except in case of **T3** tetrazine for 18 h, at 37 °C.

The formation of ON-click products **19ON_C^{XTY}** was proved by MALDI-TOF analysis (**Table 6**). The conjugate products **19DNA_C^{XTY}** were purified by magneto-separation[113], method

described in Section 5.1.6 using 5'-biotinylated template, in order to get the single strand **19ON_C^{4TCOT1}** or **19ON_C^{4TCOT2}**, **19ON_C^{2TCOT1}** or **19ON_C^{2TCOT2}**, **19ON_C^{p4TCOT1}** or **19ON_C^{p4TCOT2}** or **19ON_C^{p4TCOT3}** or **19ON_C^{p4TCOT4}**, **19ON_C^{p2TCOT1}** or **19ON_C^{p2TCOT2}**, **19ON_C^{pBCNT1}** or **19ON_C^{pBCNT2}** or **19ON_C^{pBCNT3}** or **19ON_C^{pBCNT4}** tetrazine-linked products. In order to visualize both strands of DNA (natural and modified), via MALDI analysis, is necessary higher concentration of sample compared to single strand ON, due to easier ionization of the latter. Therefore, the isolation of single strand ON, through magnetoseparation, was necessary.

Table 6. MALDI data of tetrazine-linked modified nucleic acids adducts

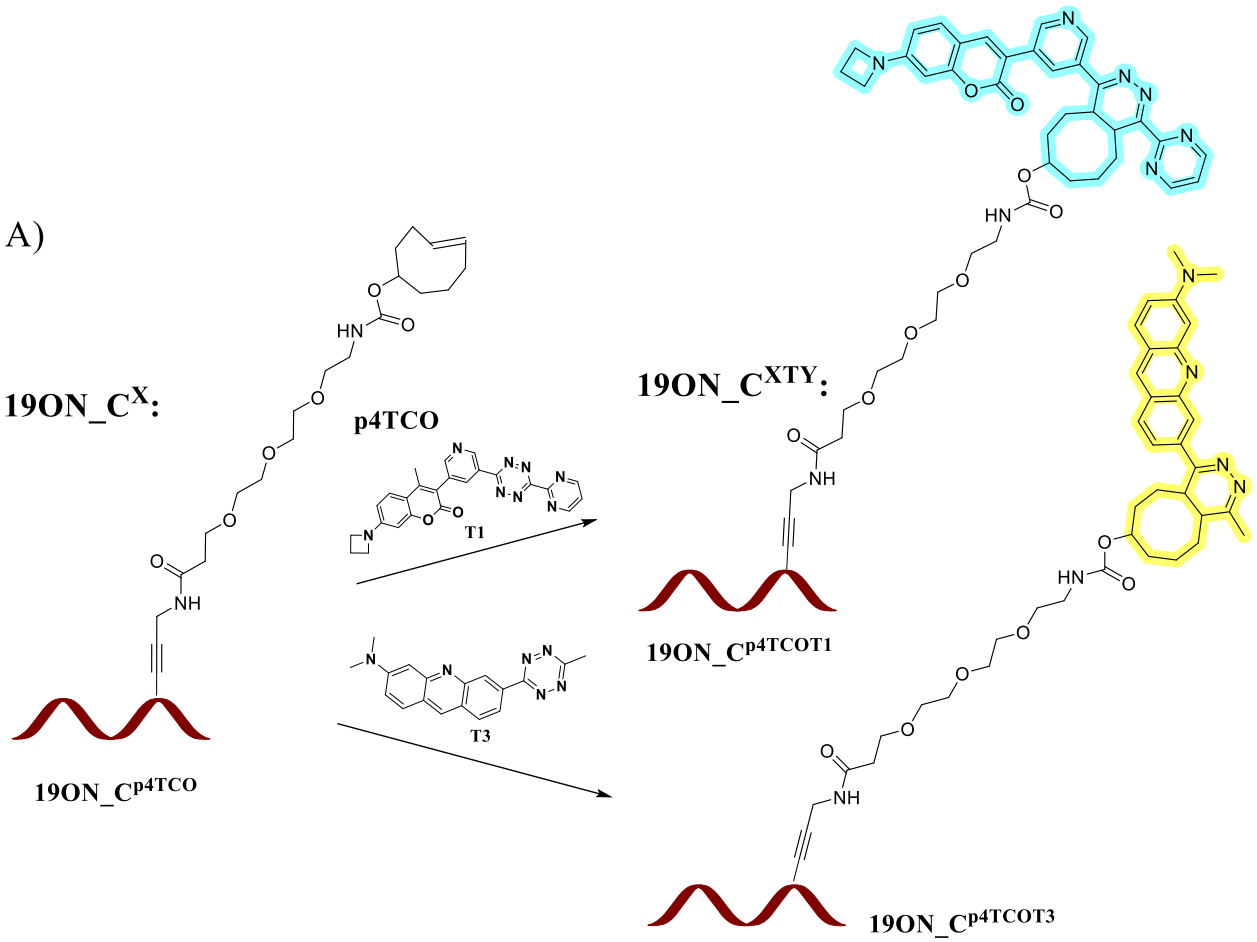
Nucleic acids	Mw calculated [Da]	Mw found [Da]	Δ [Da]
19ON_C^{4TCOT1}	6578.1	6578.2	0.02
19ON_C^{4TCOT2}	6576.1	6577.1[b]	1.0
19ON_C^{2TCOT1}	6578.1	6580.4	2.3
19ON_C^{2TCOT2}	6576.1	6579.5	3.4
19ON_C^{p4TCOT1}	6781.2	6782.4 [b]	1.2
19ON_C^{p4TCOT2}	6779.3	6780.4 [b]	1.1
19ON_C^{p2TCOT1}	6781.2	6782.8 [b]	1.6
19ON_C^{p2TCOT2}	6779.3	6780.6 [b]	1.3
19ON_C^{pBCNT1}	6805.3	6804.8	0.5
19ON_C^{pBCNT2}	6803.3	6803.2	0.1
19ON_C^{p4TCOT4}	7077.4	7078.9	1.5
19ON_C^{pBCNT4}	7103.4	7103.2	0.2
19ON_C^{pBCNT3}	6671.2	6670.8 [b]	0.5
19ON_C^{p4TCOT3}	6647.2	6646.7	0.6
19DNA_C^{p4TCOT3} [a]	7184.2	7184.4 [b]	0.2
19ON_C^{p2TCOT3} [a]	7184.4	7185.6 [b]	1.2
19DNA_C^{pBCNT3} [a]	7208.2	7209.9 [b]	1.7
19ON_C^{pBCNT3} [a]	7208.2	7209.4 [b]	1.2

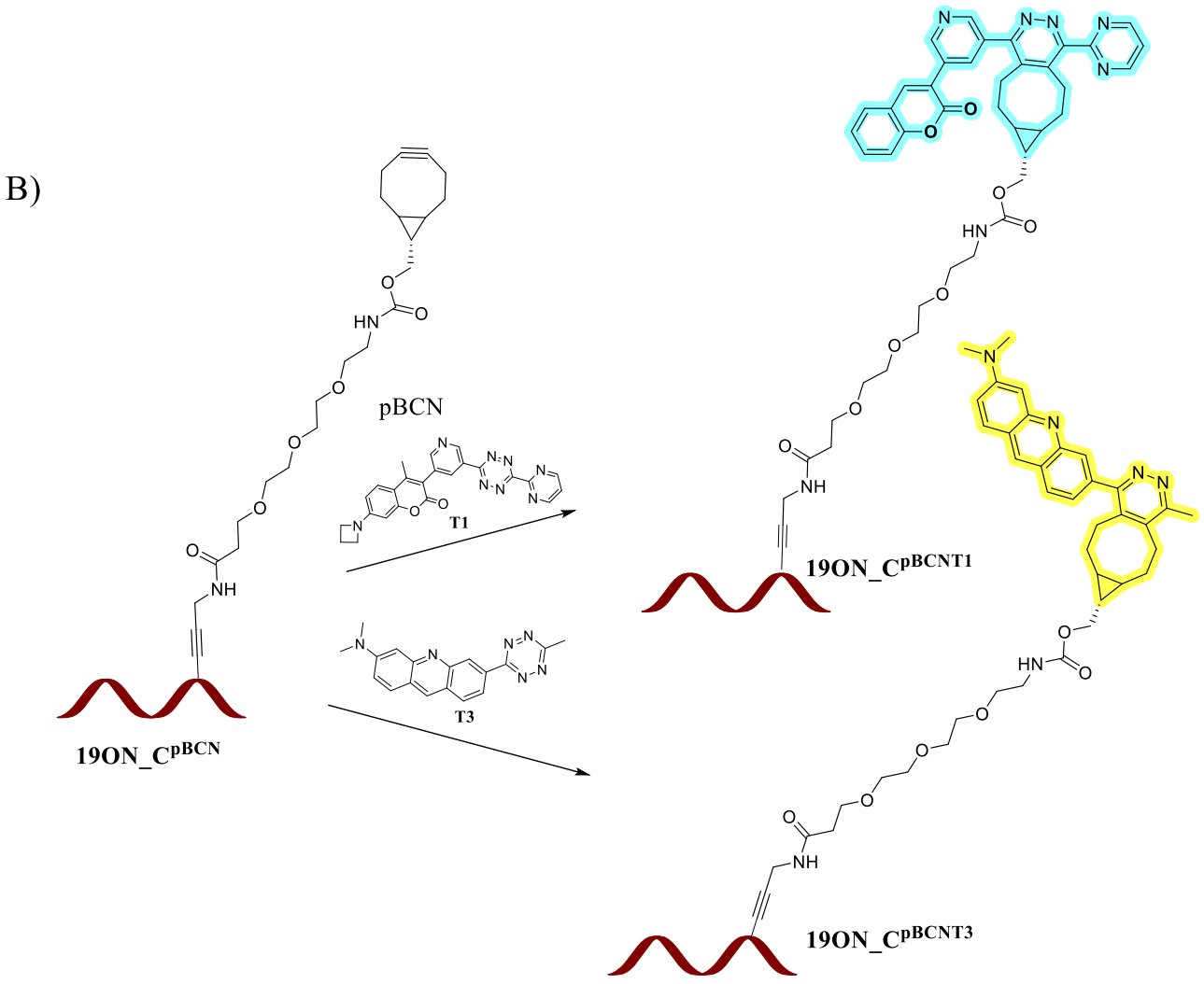
[a] 5`-(6-FAsectiM)-labelled, ON^{X}-single stranded DNA, DNA^{X}-double strand DNA.

[b] Nucleic acids were measured in positive mode

3.1.1.10 Oligonucleotide-click products (19ON_C^{XTY}) fluorescence turn-on measurements of 19ON_C^{p4TCO} or 19ON_C^{pBCN} with T1 or T3 tetrazine (19ON_C^{XTY})

Also for modified ONs **19ON_C^{X}**, the fluorescence light up of conjugates was measured after reaction with tetrazines before testing directly in cells. **T1** and **T3** tetrazines were both tested and analysed in combination with **19ON_C^{p4TCO}** and **19ON_C^{pBCN}** by PAGE gel analysis (**Figure 16,** above). Therefore we use them to measure their fluorescence, as examples of ON labelling, only **19ON_C^{p4TCO}** and **19ON_C^{pBCN}** were selected for model reactions with **T1** or **T3** (**Scheme 7**).
For the purposes of our research was not necessary to measure the fluorescence of all combination ON-tetrazines **19ON_C^{XTY}** but some of them were enough to have a comparison with labelled **dC^{XTY}TP**.
19ON_C^{p4TCO} and **19ON_C^{pBCN}** (1 μM final concentration) PEX products, obtained after magnetoseparation, were mixed with 1 μL of **T1** tetrazine (1 μM final concentration) and the fluorescence turn-on measurements were performed using 370 nm excitation and 400-650 nm for emission which are shown in **Figures 93-94** (Appendix, Section 7.1)
Due to the lower reactivity of the **T3** tetrazine, 0.18 μM solution of **19ON_C^{p4TCO}** or 0.14 μM solution of **19ON_C^{pBCN}** (final concentration) was combined with 1 μM (final concentration) of **T3** tetrazine and the reaction mixtures were incubated in the dark at room temperature for 3 hours. Then, the fluorescence "turn-on" was measured using 500 nm excitation wavelength and 525-750 nm emission window (slit 3) which is shown in **Figure 95** (Appendix, Section 7.1).
The spectra were recorded at several time points (final measurements after 30 minutes) and all measurements were typically repeated two-three times.


A)
19ON_C^X:
p4TCO
T1
T3
19ON_C^p4TCO
19ON_C^XTY:
19ON_C^p4TCOT1
19ON_C^p4TCOT3



Scheme 7. Reaction scheme of the click product of A) **19ON_C^{p4TCO}** or B) **19ON_C^{pBCN}** with **T1** or **T3** tetrazine. Only one isomer of the click product is shown for clarity.

The fluorescence turn-on values of the click products formed in the reaction of **19ON_C^{p4TCO}** and **19ON_C^{pBCN}** with **T1** or **19ON_C^{pBCN}** with **T3** tetrazine are shown in the (**Table 7**). In case of **19ON_C^{p4TCO}** with **T3** was not observed any turn on in the fluorescence, either due to degration of strained dienophile *trans*-cyclooctene and low reactivity of **T3**.

Table 7. Summary of the turn-on values for the click products (**19ON_C^{XTY}**)

Tetrazine	ON-click product	Fluorescence Measurement[a]
T1	**19ON_C^{p4TCO}**	110 fold
T1	**19ON_C^{pBCN}**	8.5 fold
T3	**19ON_C^{p4TCO}**	-
T3	**19ON_C^{pBCN}**	14 fold

[a]determined in PBS buffer (pH 7.4) using 370 nm excitation and calculated by dividing the fluorescence value at 485 nm for click product and that of the quenched **T1** or **T2** tetrazines. For **T3** tetrazine 500 nm excitation wavelength was used.

3.1.1.11 Labelling of DNA using dC^{p4TCO}TP or dC^{p2TCO}TP or dC^{pBCN}TP derivatives delivered to U-2 OS cells by SNTT1

After *in vitro* labelling experiments testing several combination of different modified nucleotides, bearing reactive group (TCO, BCN), and as well of modified oligonucleotides or DNA **19ON_C^{X}/19DNA_C^{X}** with tetrazines **T1**–**T4**, the final goal of the study was to test tetrazines **T1**–**T4** in cells for labeling of genomic DNA in order to visualize the nuclei by fluorescence confocal microscopy. All cell experiments and cell imaging were performed by Tomáš Kraus. For these experiments U-2OS cells were selected as the most suitable for measurements in confocal microscopy.[114]

The treatment consists in incubation of cells with an equimolar mixture (10 μM) of a specific **dN^{X}TP** (**dC^{2TCO}TP**, **dC^{4TCO}TP**, **dC^{p2TCO}TP**, **dC^{p4TCO}TP**, or **dC^{pBCN}TP**) and the **SNTT1** transporter in tricine buffer for 5 min at 37 °C to ensure triphosphates delivery to cells as previously described in literature.[57]

As next step, U-2 OS cells were incubated in a cultivation medium for other 60 min, and then tetrazine (**T1**–**T4**) in DMSO solution were added into it (**Figure 36**).

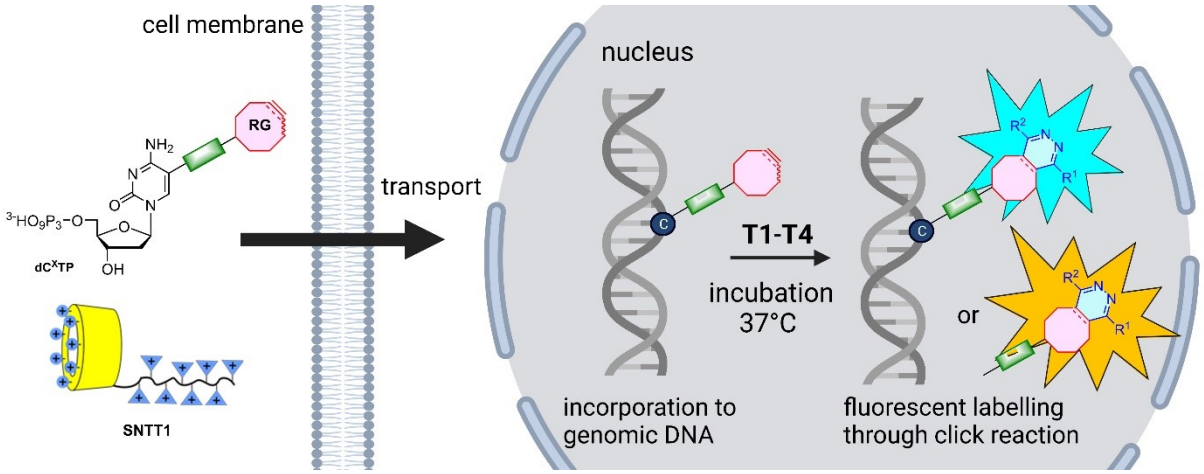


Figure 36. Delivery of modified dNTPs (**dC^{2TCO}TP**, **dC^{4TCO}TP**, **dC^{p2TCO}TP**, **dC^{p4TCO}TP**, or **dC^{pBCN}TP**) to U-2 OS cells in complex with **SNNT1** transporter and further incubation with tetrazines **T1**–**T4** to visualize labeled genomic DNA in nuclei by confocal microscopy.

As final step, the cells were incubated at 37 °C on a confocal microscope114 stage and observed at regular intervals of 10 min to follow the progress of DNA labeling, in which a clear punctuate pattern of nuclei is supposed to be shown, typical for DNA foci. In case of cells treated with **dC^{2TCO}TP** or **dC^{4TCO}TP** was not shown DNA labeling in combination of tetrazines **T1** or **T2** (1–10 μM) with unspecific staining of cytosolic structures, most probably mitochondria, whereas nuclei remained relatively dark (**Figure 37**).

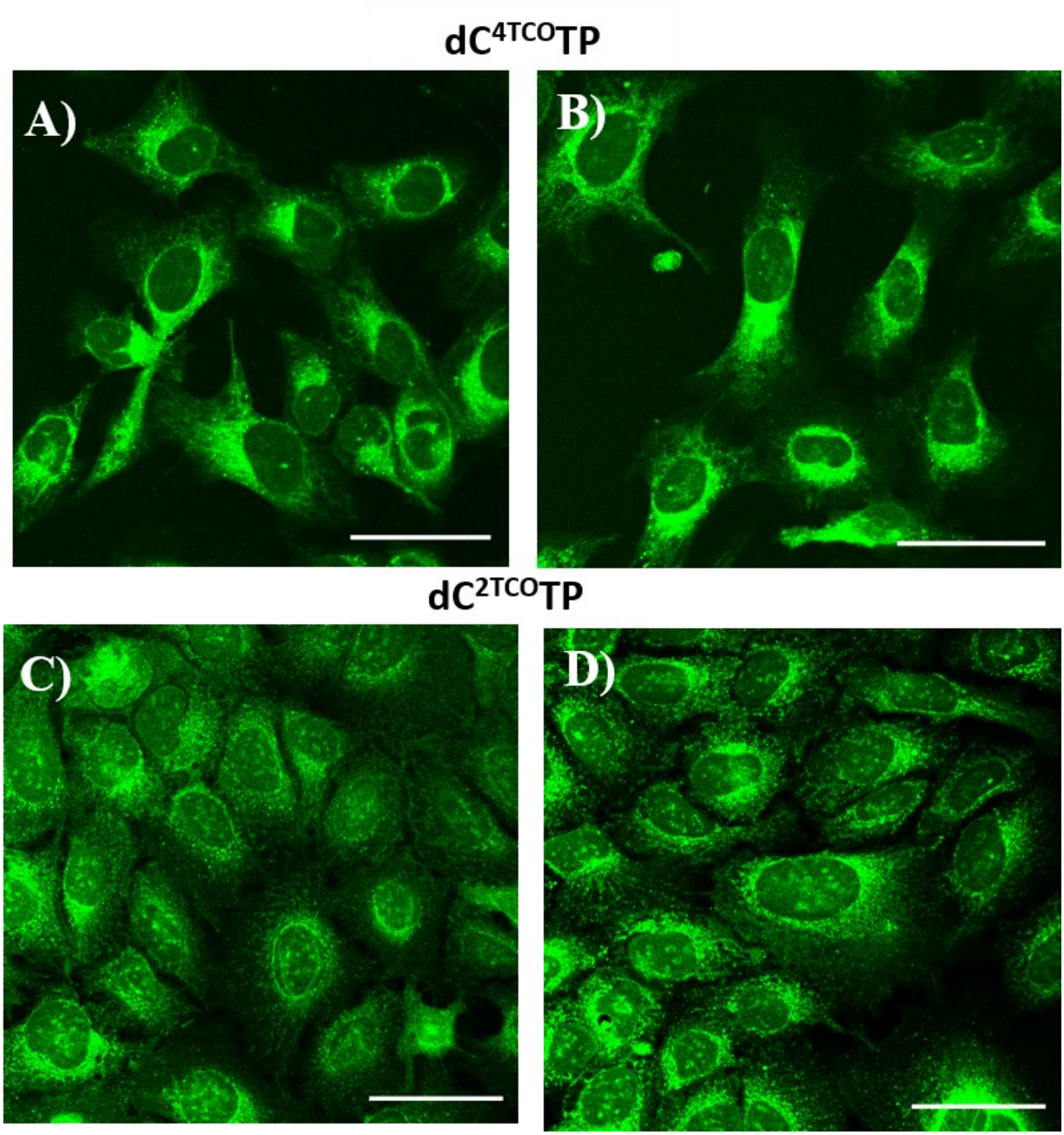


Figure 37. Confocal microscope image of U-2OS cells treated with A) **dC^{4TCO}TP** + **T1**; B) **dC^{4TCO}TP** + **T2**. Conditions: Transport **dC^{4TCO}TP/SNTT1**: 5 min/37° C. Incubation in medium: 60 min/37° C;

C) **dC^{2TCO}TP** + **T1**; D) **dC^{4TCO}TP** + **T2**. Conditions: Transport **dC^{2TCO}TP/SNTT1**: 5 min / 37° C. Incubation in medium: 60 min/37° C; Labelling with **T1** or **T2**: 15 min/37° C. Scale bar: 50 μm.

Incubation time with tetrazines **T1** or **T2** was 15 min and prolongation of cell treatment increases cytosolic fluorescence. As negative controls, cells only exposed to tetrazines **T1** or **T2** were used and they have shown a similar staining pattern, indicating noisy background fluorescence signal (**Figure 38**). The reason of this can be assigned either to **dC^{2TCO}TP** or **dC^{4TCO}TP** which are probably not good substrate for genomic DNA polymerase or to their low reactivity towards tetrazines.

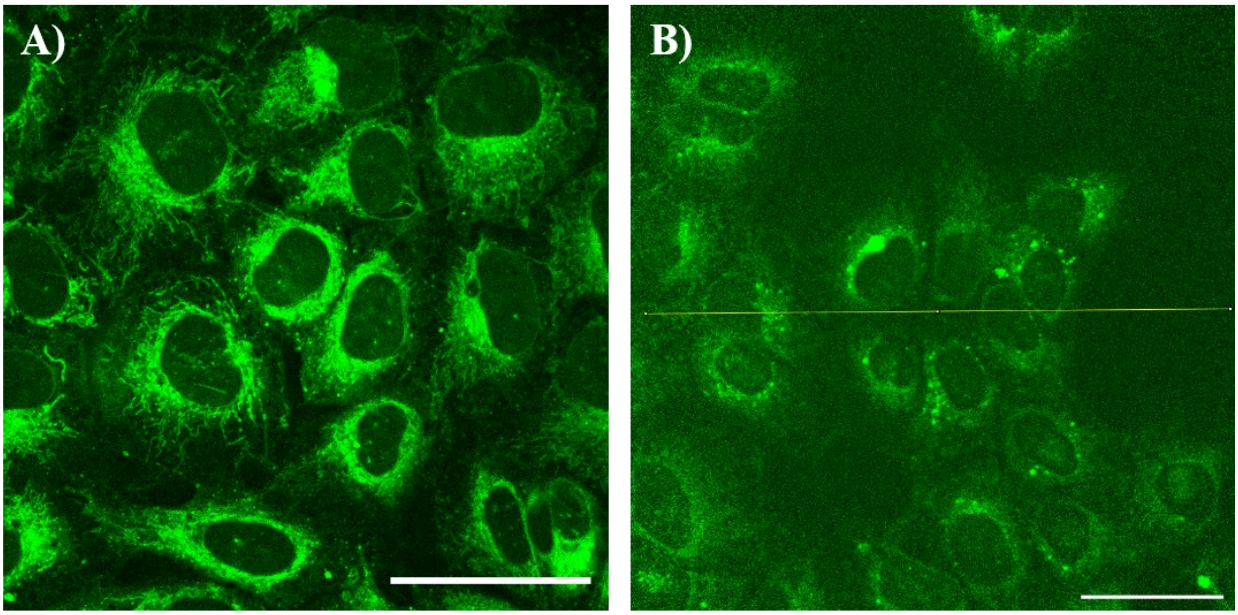


Figure 38. Confocal microscope images of control experiments **T1** and **T2**. A) Treatment of U2-OS cells with **T1** only; B) Treatment of U2-OS cells with **T2** only. Conditions: 15 min / 37° C; Scale bar: 50 μm

On the other hand, cells treated with **dC^{p2TCO}TP** or **dC^{p4TCO}TP**, in which the TCO groups were more distant from the nucleobase with the PEG3 linker, and exposed to **T1**, shown specific staining of nuclei (**Figure 39-A,B**). By comparison of cells images treated in the same conditions with **dC^{p2TCO}TP** or **dC^{p4TCO}TP** with **T1**, acquired by confocal microscope, have shown higher contrast of fluorescence signal between background and DNA in case of **dC^{p4TCO}TP** (**Figure 39-C**). Based on literature[115] we hypothesized that triphosphates bearing the longer linker were well accepted and incorporated by genomic DNA polymerase as well as **T1** to be the best in terms of cell permeability and reactivity towards strained diene, TCO which was confirmed by microscopy images.

Labeling of DNA with both tetrazines *in cellulo* has required a short period of 15 min which was sufficient to obtain clear images. The background fluorescence was increased in case of longer reaction times.

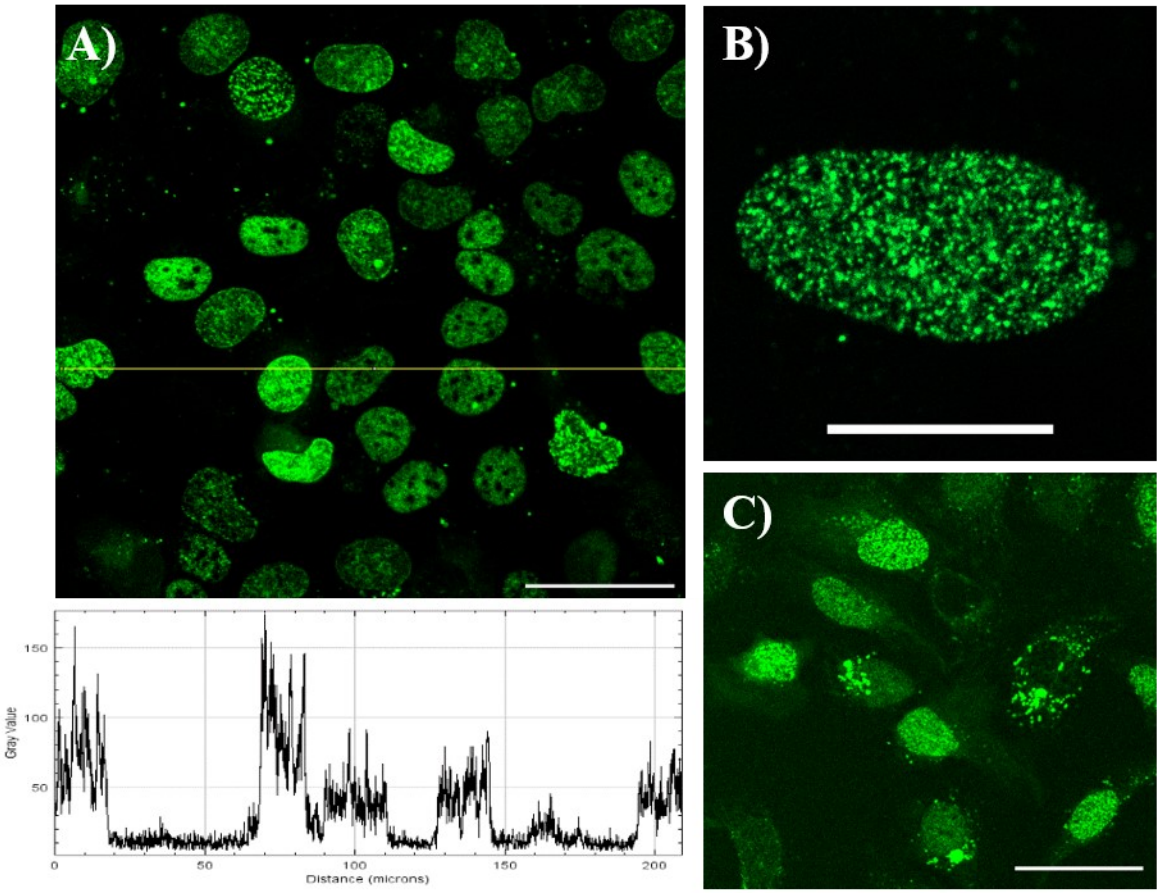


Figure 39. Confocal microscope images of A) U-2OS; **dC^{p4TCO}TP** + **T1**; B) Zoom of U-2OS cell showing DNA foci with incorporated **dC^{p4TCO}** and labeled with **T1**; C) U-2OS; **dC^{p2TCO}TP** + **T1**. Conditions: Transport **dC^{p4TCO}TP** or **dC^{p2TCO}TP/SNTT1,** 5 min / 37° C followed incubation in medium for 60 min / 37° C, then labelling with **T1**: 15 min / 37° C; Scale bar: 50 µm.

Treatment with **T2** tetrazine have shown clear DNA labeling in case of combination with **dC^{p4TCO}TP**, whereas in **dC^{p2TCO}TP** treated cells was observed unspecific staining (**Figure 40**). Tetrazine **T2,** regardless of the nucleotide used (**dC^{p2TCO}TP** or **dC^{p4TCO}TP**), has displayed considerably lower contrast than **T1.**

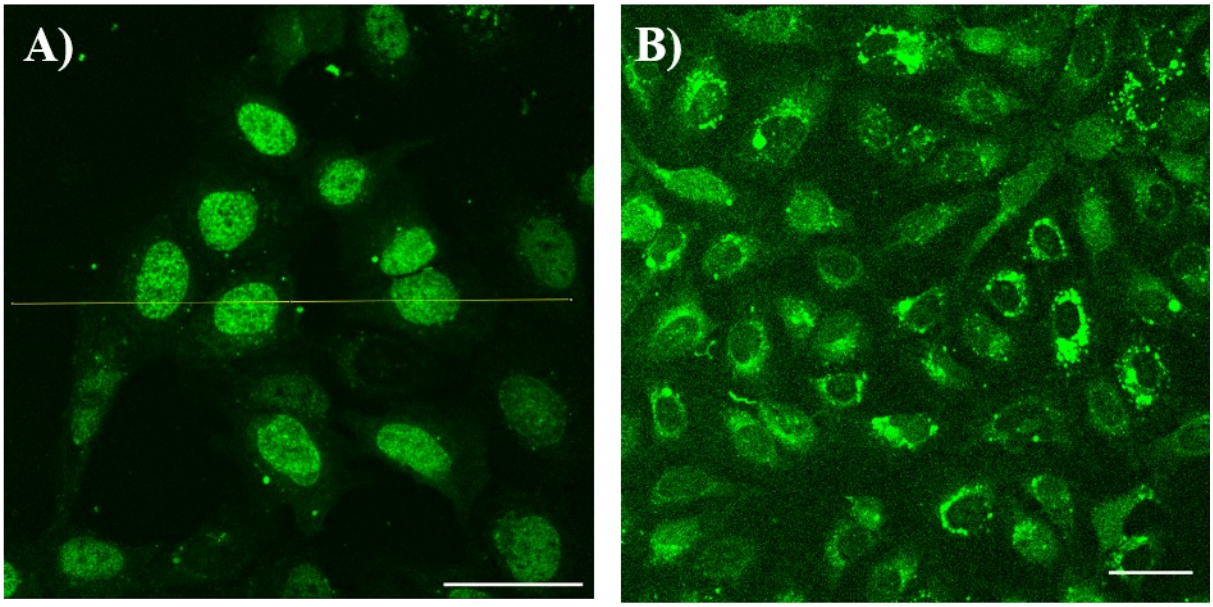


Figure 40. Confocal microscope images of A) U-2OS; **dC^{p4TCO}TP** + **T2**; B) U-2OS; **dC^{p2TCO}TP** + **T2**. Conditions: Transport **dC^{p4TCO}TP** or **dC^{p2TCO}TP**/**SNTT1**: 5 min / 37° C. Incubation in medium: 60 min / 37° C; Labelling with **T2**: 15 min / 37° C. Scale bar: 50 µm

Additionally, **dC^{pBCN}TP** was incubated with tetrazines **T1** or **T2** providing visible DNA foci, but the contrast between the labeled DNA and background was lower than for **dC^{p4TCO}TP** (**Figure 41**).

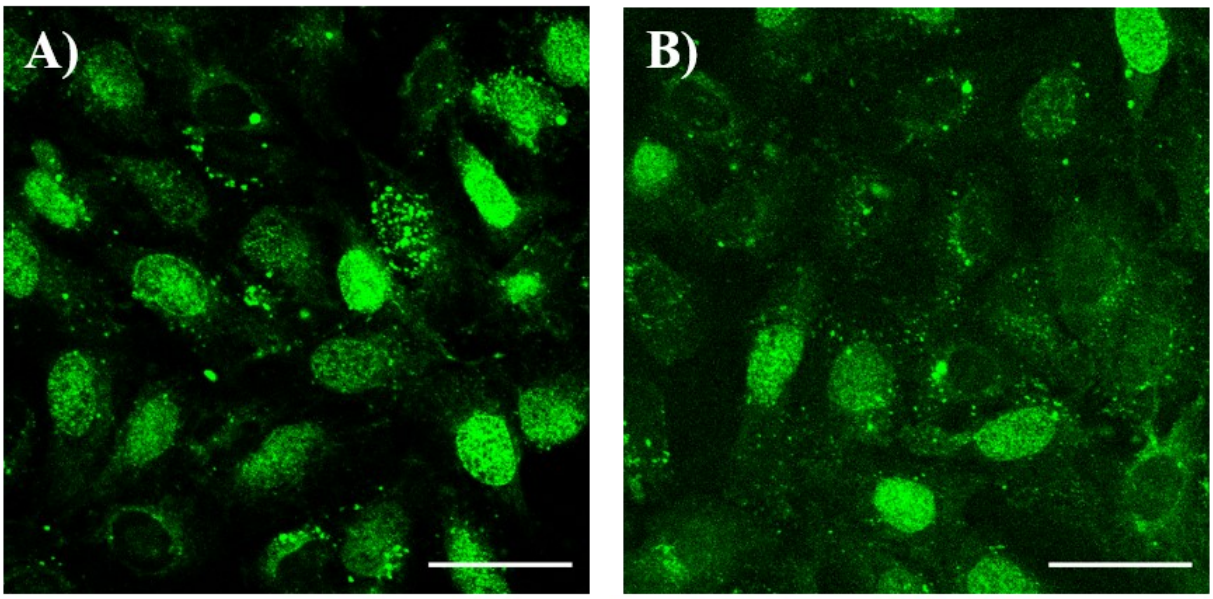


Figure 41. Confocal microscope images of U-2OS treated with A) **dC^{pBCN}TP** + **T1**; B) **dC^{pBCN}TP** + **T2**. Conditions: Transport **dC^{pBCN}TP** /**SNTT1**: 5 min / 37° C. Incubation in medium: 60 min / 37° C; Labelling with **T1** or **T2**: 15 min / 37° C. Scale bar: 50 µm.

Next, tetrazine **T3** was tested and incubated for longer time of 4 h, due to lower reactivity, with **dC^{p2TCO}TP**, **dC^{p4TCO}TP**, or **dC^{pBCN}TP**, delivered by **SNTT1**[57] and clear labeling of DNA was observed in case of **dC^{pBCN}TP** (**Figure 42-A,B,C**). Nevertheless, non specific staining was observed in the cytosol in case of only treatment of cells with tetrazine **T3** (**Figure 42-D**) and longer treatment time has increased the fluorescence signal background.

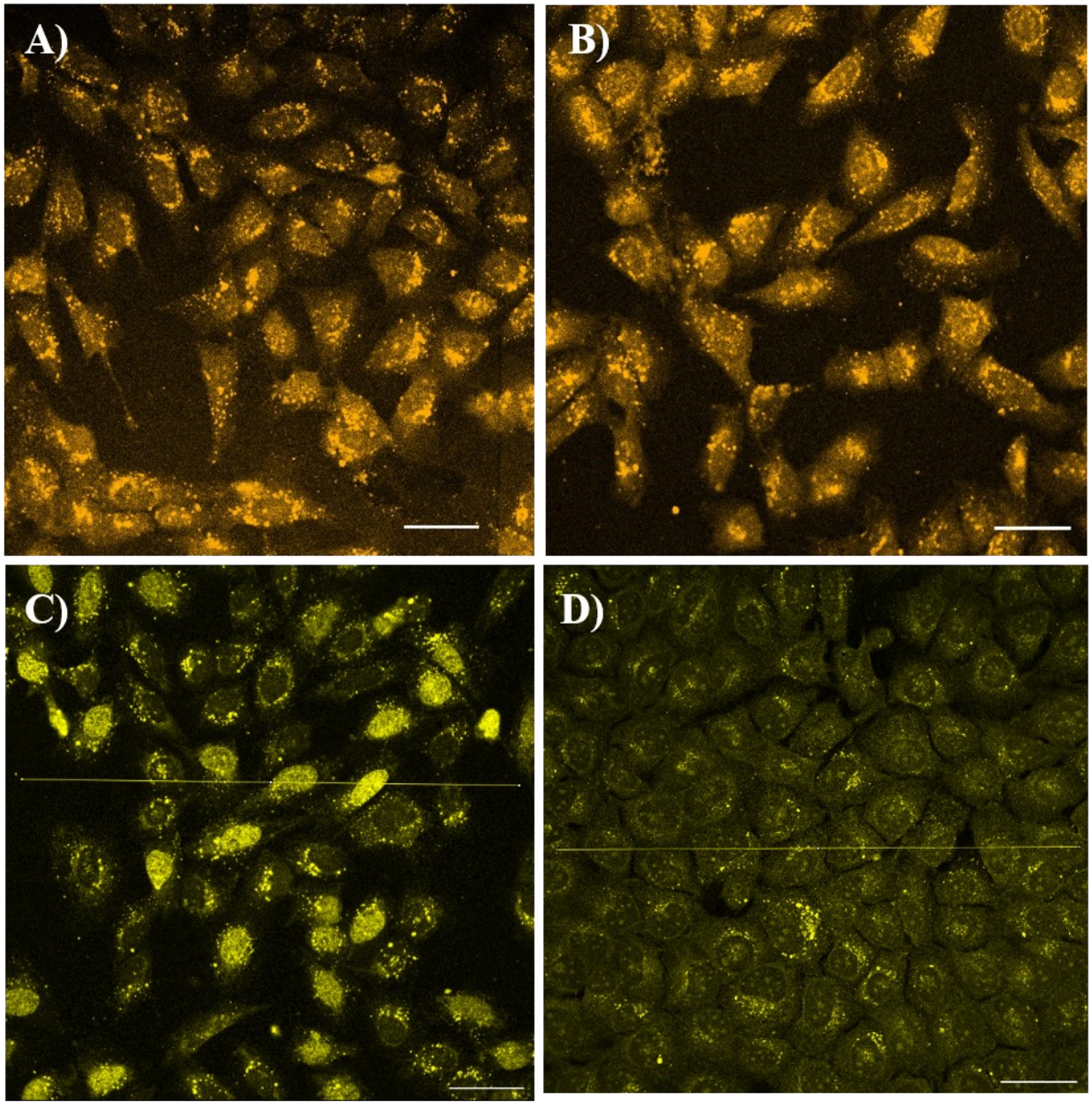


Figure 42. Confocal microscope images of U-2OS treated with A) **dC^{p4TCO}TP** + **T3**; B) **dC^{p2TCO}TP** + **T3**; C) **dC^{pBCN}TP** + **T3**; D) **T3** only. Conditions: Transport **dC^{p4TCO}TP** or **dC^{p2TCO}TP** or **dC^{pBCN}TP** /**SNTT1**: 5 min / 37° C. Incubation in medium: 60 min / 37° C; Labelling with **T3**: 4 hours / 37° C. Scale bar: 50 µm

Negative controls of cell incubated with **dC^{p2TCO}TP** or **dC^{p4TCO}TP** have shown the same non specific fluorescence background in cytosol compartments (**Figure 43**).

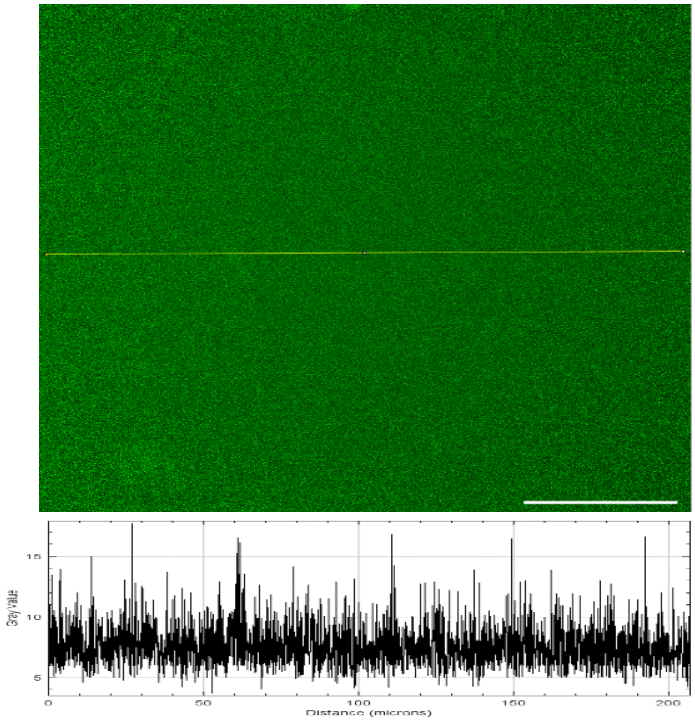


Figure 43. Confocal microscope image of U2-OS cells treated only with **dC^{p4TCO}TP**. Control experiment: Autofluorescence of U2OS cells; Scale bar: 50 µm

As last, the reactivity of commercially available 5-TAMRA-pyrimidyl-tetrazine **T4** was tested in cells incubated with **dC^{p4TCO}TP** or **dC^{pBCN}TP**. Live U-2OS cells and fixed ones were used in this experiment. In particular the first were used to maintain their natural physiological state, including active metabolism, growth, and division. Therefore, live U-2OS cells were used to follow the different phase of cell cycle by fluorescence. Instead, fixed cells were used for high-resolution imaging of cellular structures, such as nuclei. In live cells was shown less efficient labeling due to low cell membrane permeability, whereas in fixed cells clear images of nuclei with DNA foci were shown (**Figure 44**, **A-D**). This tetrazine has shown several disadvantages, beyond that slugghish permeation through cell membrane it requires also high concentration (2 µM) and longer incubation time (60 min) compared to other tetrazines in order to clearly visualize DNA foci.

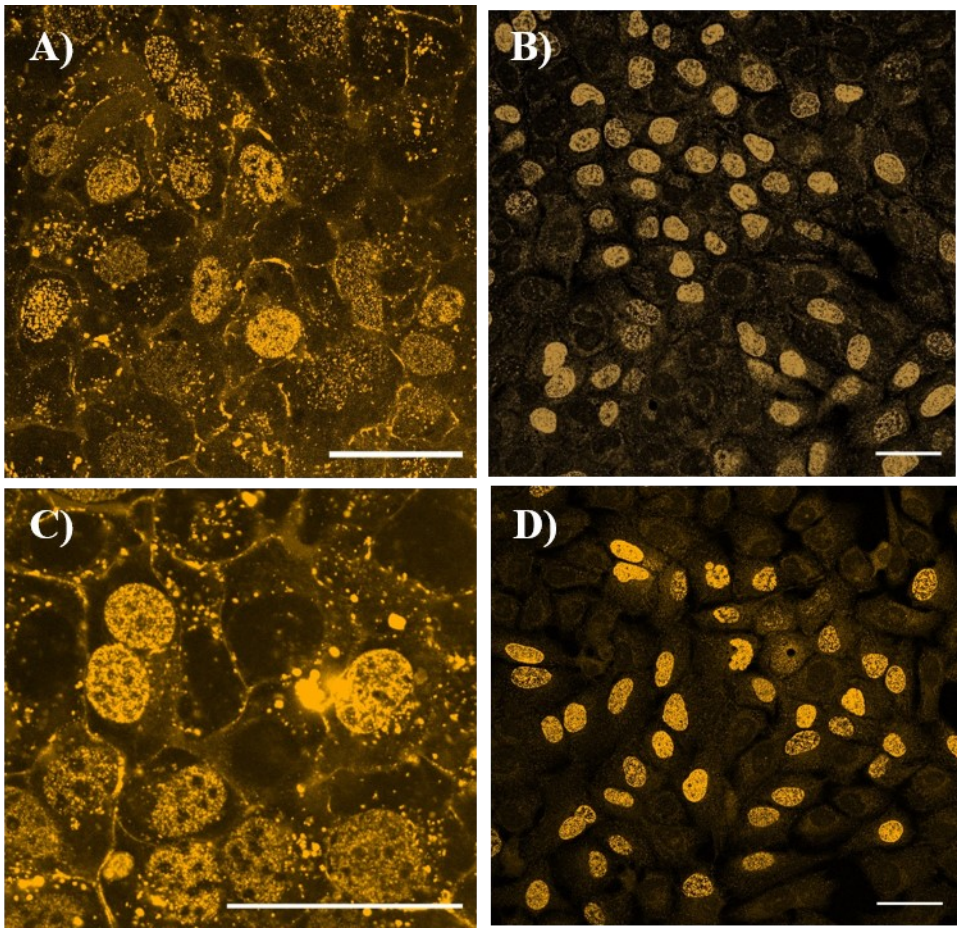


Figure 44. A) U-2OS (live); **dC^{p4TCO}TP** + **T4**, B) U-2OS (fixed); **dC^{p4TCO}TP** + **T4**; C) U-2OS (live); **dC^{pBCN}TP** + **T4**; D) U-2OS (fixed); **dC^{pBCN}TP** + **T4**. Conditions for live cells: transport **dC^{p4TCO}TP** /**SNTT1**: 5 min / 37° C. Incubation in medium: 60 min / 37° C; Labelling with **T4**: 45 min / 37° C. Scale bar: 50 µm. Conditions for fixed cells: transport **dC^{p4TCO}TP** /**SNTT1**: 5 min / 37° C. Incubation in medium: 60 min / 37° C; Fixation with methanol: 30 min / 0 ° C. Labelling with **T4**: 30 min (in combination with **dC^{p4TCO}TP**) or 45 min (in combination with **dC^{pBCN}TP**/ 37° C. Scale bar: 50 µm.

In terms of microscopy images, in case of tetrazine **T4**, strong background fluorescence was shown in cytosol and plasma membrane. By comparison of cells images, whose DNA was modified by p4TCO versus pBCN, were displayed a higher contrast for pBCN (**Figure 45**).

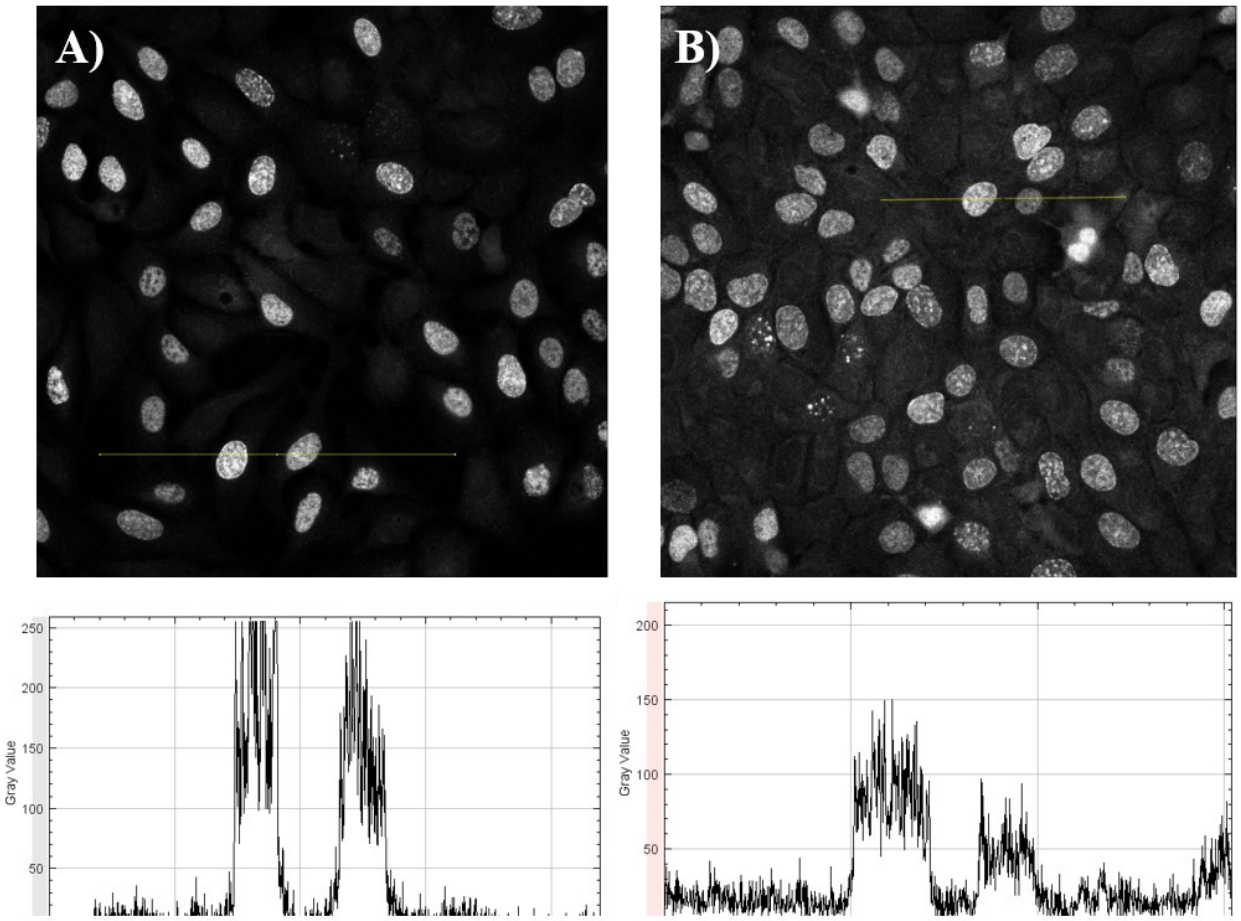


Figure 45. Comparison of contrast in images of U2OS cells labeled with A) **dC^{pBCN}TP** + **T4** and B) **dC^{p4TCO}TP** + **T4.** Images were acquired under identical conditions, acquisition parameters were adjusted according to the brightest nuclei in Image A. Image B) exhibits the lower intensity of labeled DNA foci and higher background fluorescence than image A). Condition: Transport **dC^{p4TCO}TP** or **dC^{pBCN}TP** /**SNTT1**: 5 min / 37° C; Incubation in medium: 60 min / 37° C; Fixation with methanol: 30 min / 0° C; Labelling with **T4**: 45 min / 37° C.

In short, to achieve efficient labeling in living cells is crucial the right combination between dienophile and tetrazine. In particular PEG3-2-TCO, PEG3-4-TCO, and PEG3-BCN reactive groups are more suitable for DNA labeling in live cells; and high reactivity was shown for PEG3-4-TCO and PEG3-BCN nucleotides incorporated into DNA with coumarin–tetrazine **T1** and 5-TAMRA-pyrimidyl-tetrazine **T4** conjugates. Instead, BCN-labeled DNA has shown low efficacy with PINK tetrazine **T3** showing a high unspecific background fluorescence signal. All these combinatorial experiments revealed that the presence of longer linker PEG3 is crucial for efficient incorporation into genomic DNA and good reactivity in the IEDDA click reaction with tetrazines to allow staining of DNA and imaging of DNA synthesis in live cells within time periods as short as 15 min.

3.1.1.11.1 Cell cycle progression assay

A cell progression assay[116] is a laboratory technique used to study the various stages of cell cycle progression in a population of cells. This assay, performed by Erika Kužmová, is useful to follow the nascent labelled genomic DNA in S phase and distinguish from not labelled DNA in G1 and G2-M mithosis phases, by fluorescence. Therefore, for this purpose it is important to use the best combinations of modified triphosphates **dC^{p4TCO}TP** or **dC^{pBCN}TP** with **T1** or **T4** tetrazines, observed in confocal microscopy, also for cell cycle progression assay. Firstly, the coumarin–tetrazine conjugate **T1** was taken in consideration. As first step, the cells were treated with either **dC^{p4TCO}TP** or **dC^{pBCN}TP** and **SNTT1** in tricine buffer and then incubated in a complete cultivation medium for 60 min in order to deliver the modified triphosphate inside cells. Then, cells were treated with tetrazine **T1** for 15 min and then divided in two fractions, whose the first one was used to check by cell cytometry the labeling efficiency (i.e., the difference between fluorescence intensities of the labeled and non-labeled cell populations, (**Figure 46**, **A-F**). The second fraction was, additionally, treated with a cell permeable far-red fluorescent DNA-staining, DRAQ5 after treatment with **T1**, and then analysed by cell cytometry (**Figure 46**, **A-C**).
In case of first fraction **dC^{pBCN}TP** transport to cells and treatment with tetrazine **T4**, the second fraction was stained with DAPI and then analysed in the same way by cell cytometry (**Figure 47**, **E-F**). Different staining were used in order to avoid the overlapping of click labeled DNA emission

wavelength (λ_{em}) and the one stained with DRAQ5 or DAPI, even though, as shown further, was observed partial overlap in case of DRAQ5.

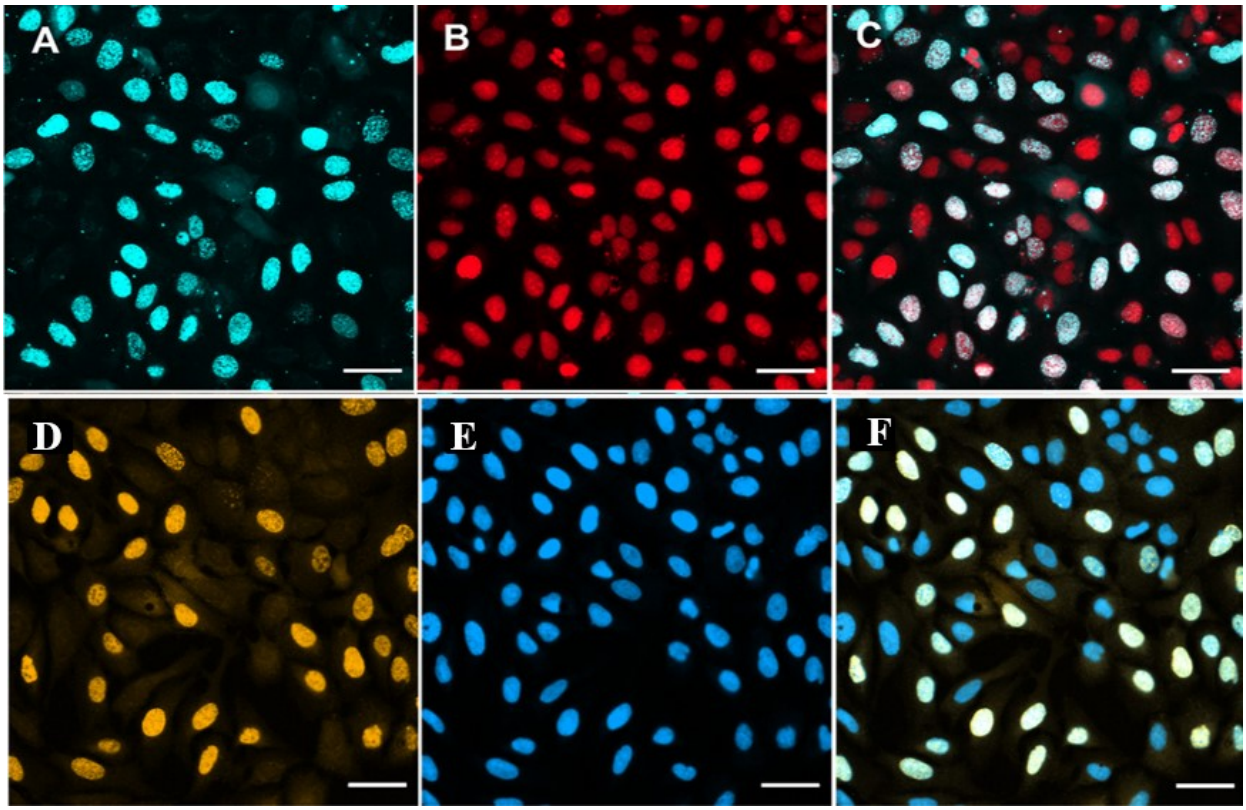


Figure 47. Confocal microscopy imaging of U-2 OS cells with labeled DNA. Cells were treated with **dC^{p4TCO}TP/SNTT1** (A–C; 10 μM) or **dC^{pBCN}TP/SNTT1** (D–F; 10 μM) for 5 min, then incubated in a medium for 60 min. Subsequently, tetrazine **T1** (A–C; 1 μM; 15 min) was added. Cells incubated with **dC^{pBCN}TP** (D–F) were fixed with methanol before tetrazine **T4** (D–F; 2 μM; 30 min) was added. Prior to microscopy, DRAQ5 (B–C; 2.5 μM) or DAPI (E–F; 0.5 μM) was added. **T1** channel (A), DRAQ5 channel (B), and merged (C); **T4** channel (D), DAPI (E), and merged (F). Bar size 50 μm.

The outcome of the assay, using the combination of **dC^{p4TCO}TP** with **T4**, was the first fraction revealed that the two (G1 and G2/M phase non-labeled vs S phase labeled) populations of cells were expected to be differentiated with separated peaks of fluorescence intensities but were less separated compared to published previous work.[57]

In case of second fraction of cells was shown decreasing of fluorescence intensity of the S-phase cells, after staining with DRAQ5 and the separation of the two populations was lost.

It was challenging, in this case, to appreciate the separations between cell populations in different phases, the reason of that is probably associated with quenching of fluorescence of coumarin fluorophores caused by DRAQ5 which is intercalated into DNA. In other words, as previously mentioned, there is a partial overlap of DRAQ5 absorption and coumarin emission bands between 400 and 500 nm.

In contrast, once DNA into fixed cells is labeled with **dC^{pBCN}TP**/**T4** reagents and treated with the addition of DAPI (**Figure 48-A**), acceptable separation of all cell cycle phases is shown (**Figure 48-B**). In case of control experiment of treatment with **T4** only, any separation of peaks was shown (**Figure 48-C**).

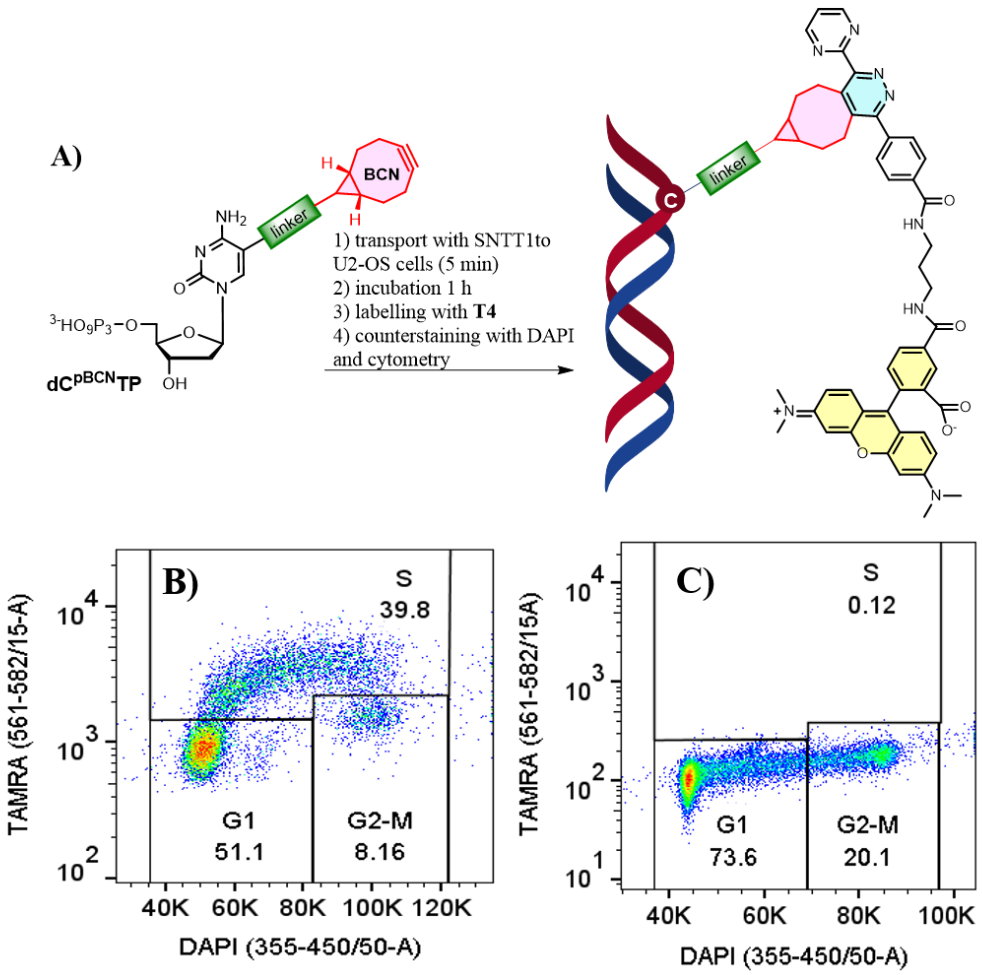


Figure 48. Scheme of A) DNA incorporation-based cell cycle analysis of U-2 OS cells; Flow cytometry images of B) DNA incorporation-based cell cycle analysis of U-2 OS cells. The cells were pulse-treated with a mixture of **dC^{pBCN}TP / SNTT1** (10 µM) in tricine buffer for 5 min and further incubated in a conditioned medium for 1h. After this period, the cells were fixed with methanol, incubated with 2 µM TAMRA-tetrazine (**T4**) for 30 minutes at 37 °C, counterstained with DAPI and analyzed by flow cytometry. C) Control experiment: cells were fixed and treated with tetrazine **T4** only.

To summarize, the modified **dC^{X}TP**s were delivered to cells by nucleoside triphosphate transporter **SNTT1** which were incubated at 37 °C for 60 min. Subsequent treatment of cells with one of the tetrazines used (**T1**–**T4**) revealed that the DNA with incorporated **dC^{p2TCO}**, **dC^{p4TCO}**, and **dC^{pBCN}** were labeled efficiently with **T1** or **T2**, instead unspecific labeling was observed for cells treated with **dC^{2TCO}TP** or **dC^{4TCO}TP**. Tetrazine **T3** labeled only the DNA with incorporated

dC^{pBCN} but required a longer reaction time (4 h). Commercially available **T4** was able to label the DNA with incorporated **dC^{p4TCO}** or **dC^{pBCN}** in live cells but less efficiently, compared to fixed cells, due to low cell membrane permeability. Fixed cells after the incorporation of **dC^{p4TCO}TP** or **dC^{pBCN}TP**, the labeling with **T4** showed clear images of nuclei with DNA foci (**Figure 96**, section 7.2, Appendix).

Additionally, the **dC^{pBCN}TP**/**T4** system was successfully used in a cell cycle cytometric assay, which allowed the quantification of cells at different stages of the cell cycle.

Unfortunately, for combination **dC^{pBCN}TP**/**T4** system was not optimal but the development of this two-step procedure for *in cellulo* labeling of the nascent DNA could be useful in cases where the direct incorporation of fluorescent tagged dNTPs is not possible due to the incompatibility of the fluorophore with cellular polymerases.

In this first part of the thesis the design and synthesis of a series of 2′-deoxyribonucleoside triphosphates (dNTPs) bearing 2- or 4-linked *trans*-cyclooctene (TCO) or bicyclononyne (BCN) tethered through a shorter propargylcarbamate or longer triethyleneglycol-based spacer was successfully performed. Moreover, several combination of dNTPs and tetrazines were efficiently used for labeling of DNA and imaging of DNA synthesis allowing to develop a new two-step procedure method which is operationally simple and overcoming several problems of previous used methods.

3.2 PART II

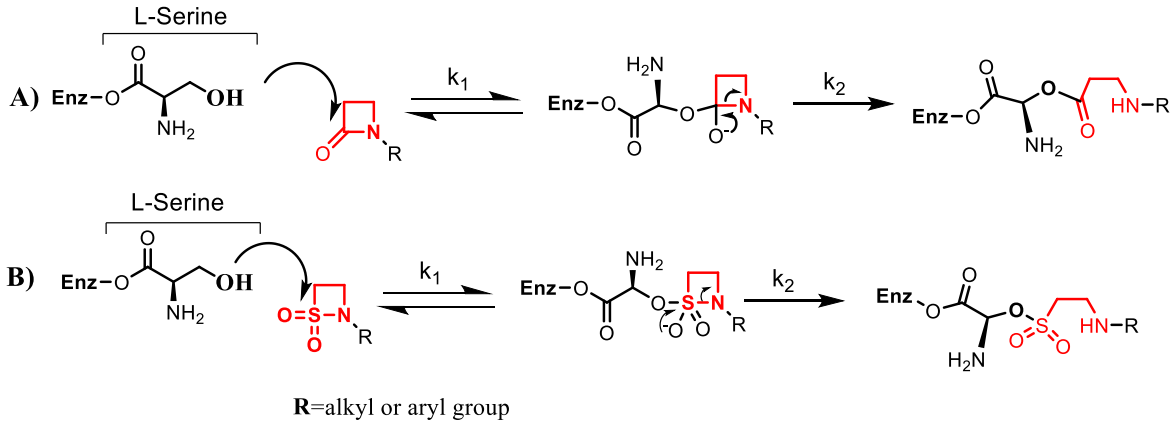
3.2.1 Design and synthesis of modified nucleotides bearing thiazetidine (β-sultam) for cross-linking with serine, their enzymatic incorporation and testing of reactivity

3.2.1.1 Synthesis of β-sultam modified nucleoside monophosphate (dC^{SU1}MP) and triphosphates (dC^{SU1}TP, dC^{SU2}TP, dU^{SU1}TP, dU^{SU2}TP)

Traditionally, β-lactams are used as anti-bacterial agents by inhibiting serine transpeptidases, but sulfonation of serine proteases provides an alternative strategy to the previous mechanism-based acylation process reported in literature[80] (**Scheme 8**, **A**). Sulfonation of serine enzymes is not a

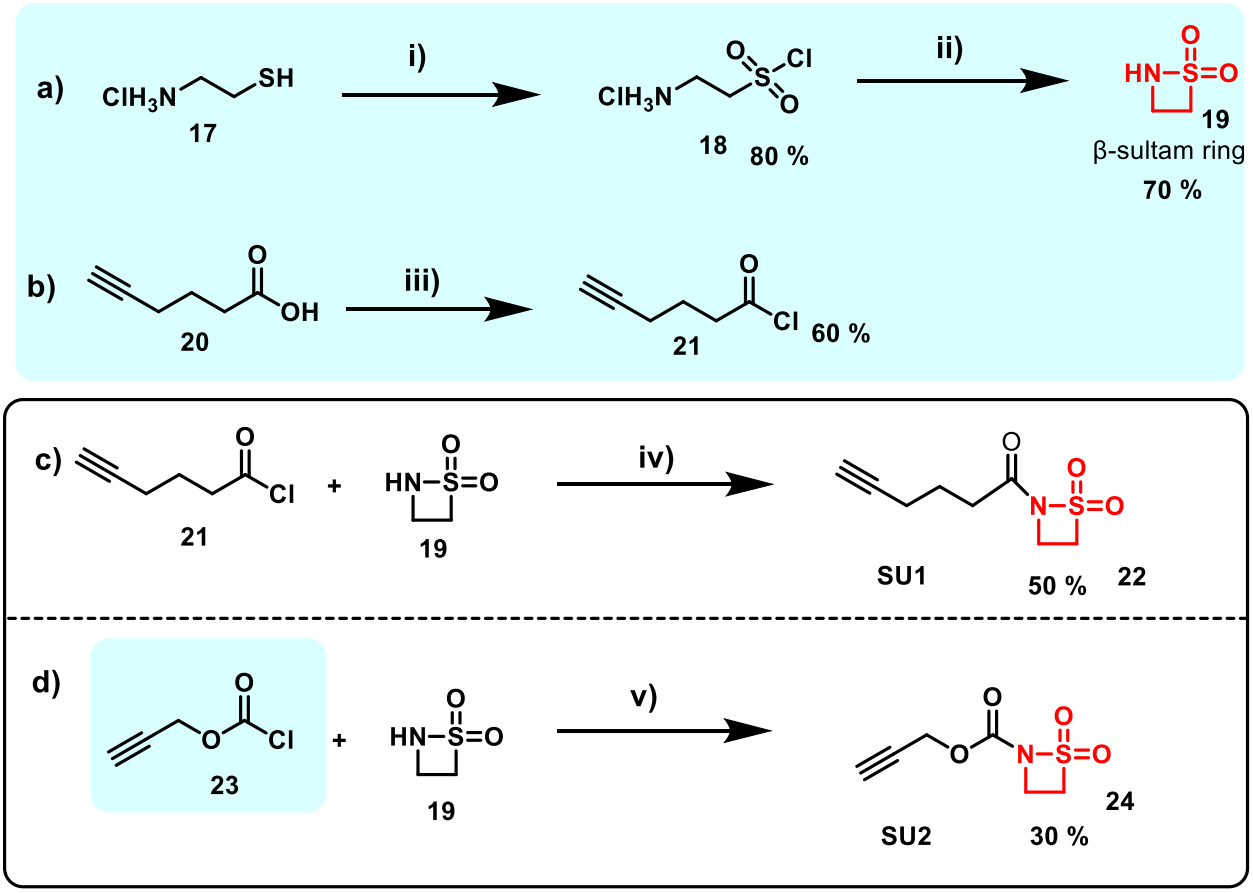
well studied process because sulfonyl derivatives are much less reactive than their acyl analogs (102 to 104 folder slower). Additionally, acyclic sulfonamides are extremely resistant to alkaline and acidic hydrolysis, but β-sultams (cyclic sulfonamides) showed extraordinary rate enhancements, with the acid and base catalysed hydrolysis compared to the acyclic ones.[80]

Therefore, β-sultams moieties were chosen as reactive groups to target serine aminoacid or serine containing peptides/proteins as reported in literature[80] as inhibitors of serine active enzymes such as elastase, protease etc. The mechanism of the reaction consists of nucleophilic addition of serine –OH group to the sulfonyl centre with consequent ring opening of the β-sultam giving the sulfonate ester (**Scheme 18, B**).



Scheme 8. Schematic representation of reaction mechanism of serine containing proteins with A) β-lactams and B) β-sultams.

Before the synthesis of the modified nucleoside triphosphates, the building block was designed having two reactive parts: the 1,1-dioxido-thiazetidin (β-sultam) ring and the alkyne handle. The goal was to link the reactive moiety with 5-iodo-2'-deoxycytidine-5'-*O*-triphosphate via Sonogashira cross coupling or with 5-azidomethyl-2'-deoxyuridine-5'-*O*-triphosphate via CuAAC reaction. As alkyne handle, hexanoyl acylchloride (synthetized from hex-5-ynoic acid) and propargyl chloroformate (commercially available) were taken in consideration in order to synthetize two different reactive moieties **22** (**SU1**) and **24** (**SU2**) (**Scheme 9**) to test their reactivity towards serine and then their tolerance to Sonogashira cross-coupling and CuAAC reaction conditions. The synthesis of both building blocks was performed according to published literature[80] starting from the 1,1-dioxido-thiazetidin ring and connected to the alkyne part by replacing the chloride group *via* nucleophilic substitution of NH group (**Scheme 9**).



Scheme 9. Synthesis of a) thiazetidin (β-sultam) ring; synthesis of b) alkyne handle, hexanoyl acylchloride; Synthesis of c) 1,1-dioxido-thiazetidin hexynone (**SU1**, compound **22**) or d) 1,1-dioxido-thiazetidin propargylformate (**SU2**, compound **24**). Conditions: i) CCl_4/EtOH (5:1), Cl_2 (g) at 0 °C. ii) Na_2CO_3 (2 equiv.) in EtOAc, 2 days at room temperature (RT). iii) oxalyl chloride (2 equiv.) in DCM, 2h at RT. iv) DMAP (2 equiv.) in DCM, 4h at 0 °C. v) BuLi (1 equiv.) in THF at -78 °C.

The confirmation of the β-sultam ring, in closed form, on **22** and **24** compounds can be mainly proved by the presence of characteristic S=O group stretch, between 1670 cm^{-1} and 1730 cm^{-1}, by IR analysis (**Figure 49**).

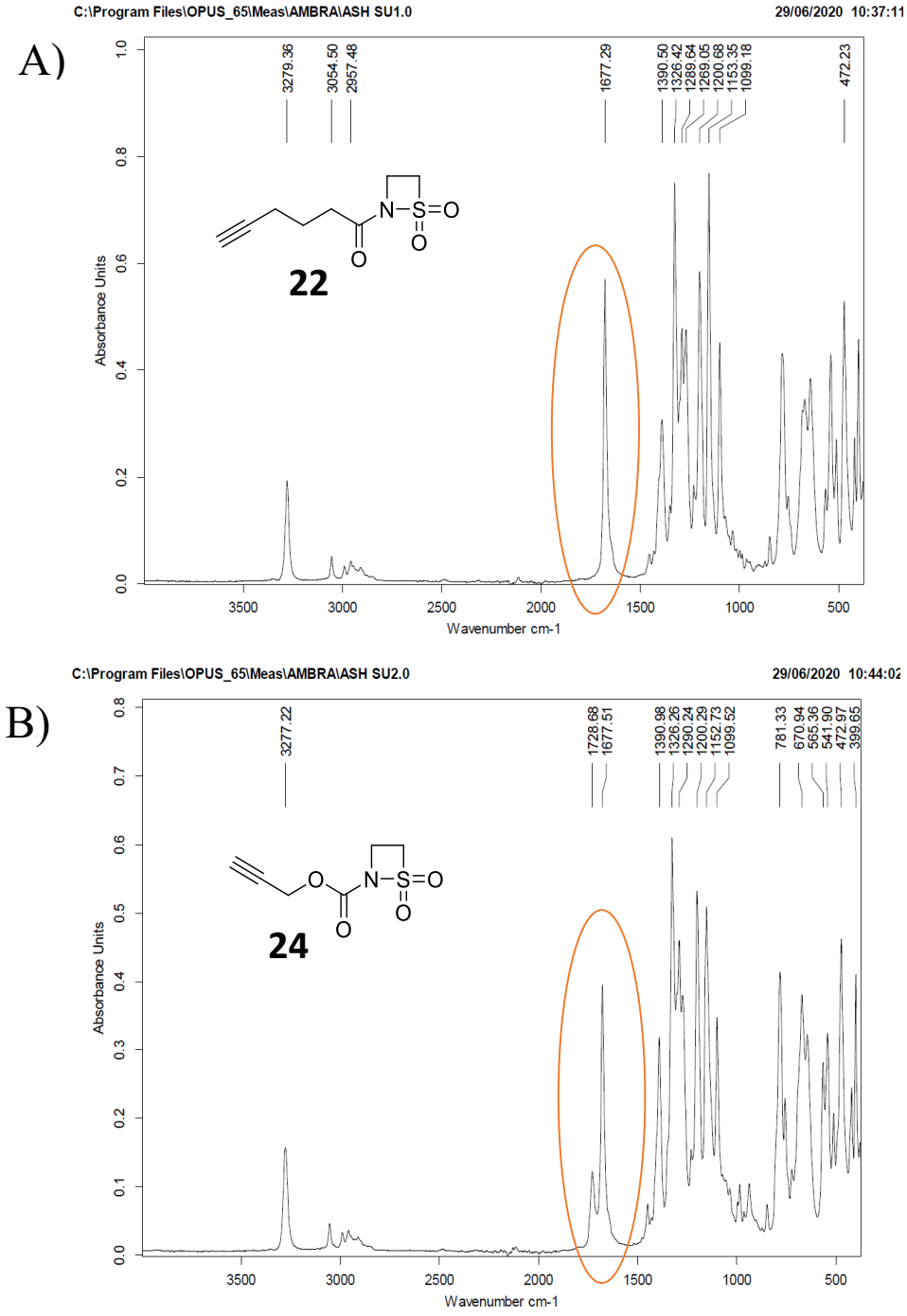


Figure 49. IR analysis of A) **SU1 linker**, compound **22** and of B) **SU2 linker**, compound **24**.

Moreover, the stability test on β-sultam ring **19** was investigated by incubation at room temperature in water at 30 min, 1 h, 5 h, 18 h and 24 h. Mass analysis at the different time was measured and the ring was confirmed to be stable in water even after 24 h by the same mass value (**Figure 50**).

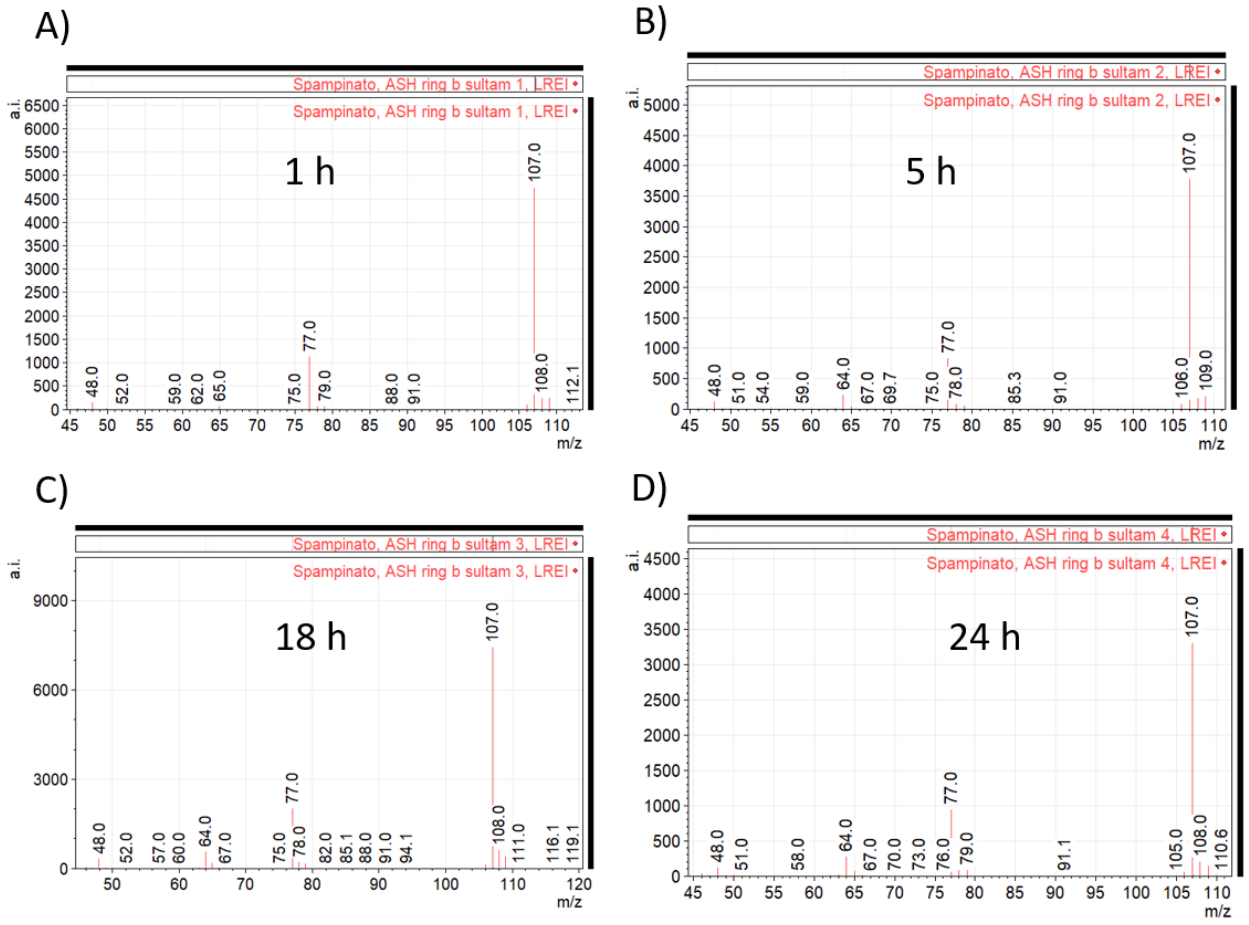
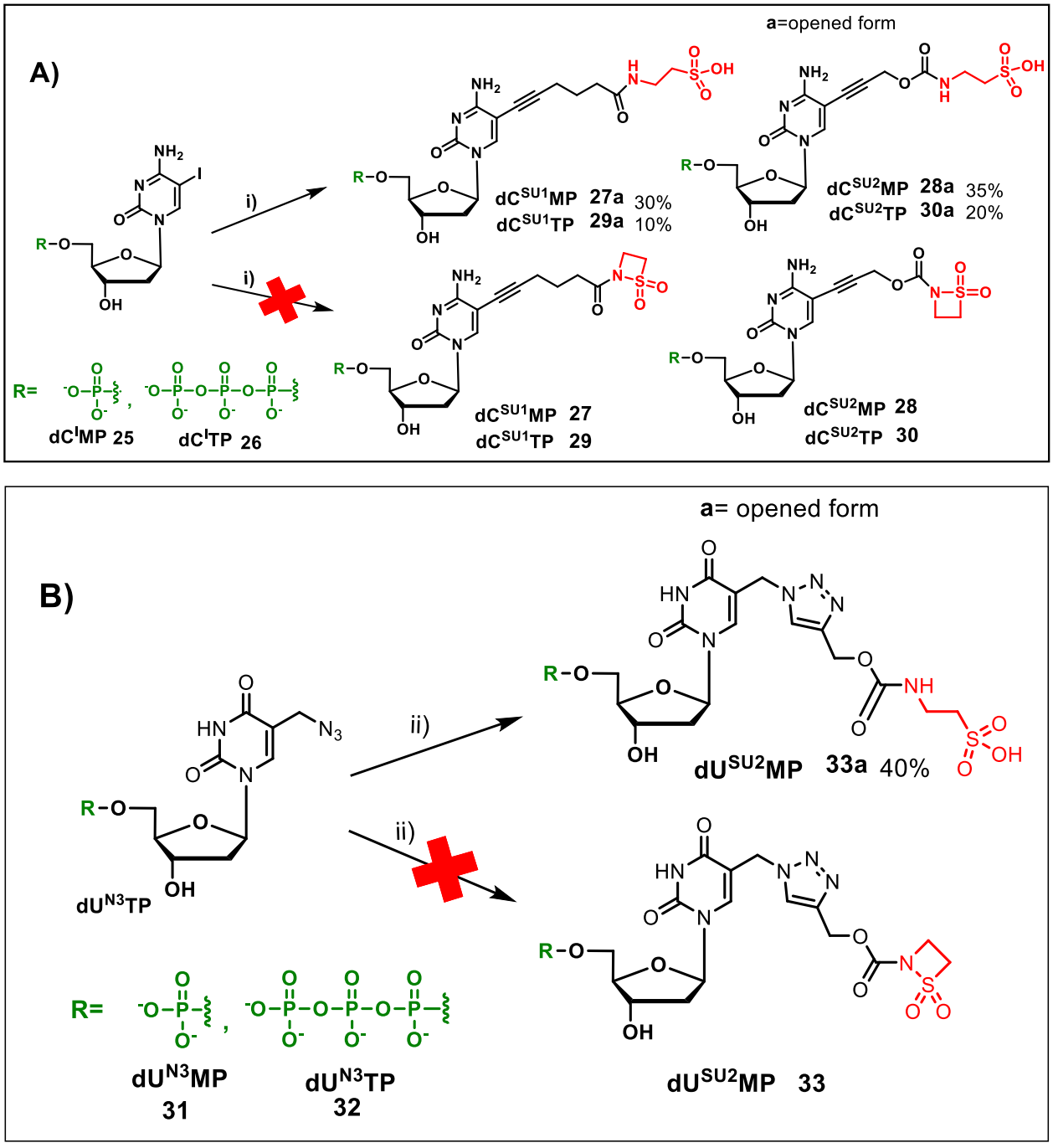


Figure 50. Mass analysis of β-sultam ring **19** after A) 1 h; B) 5 h; C) 18 h and D) 24 h.

Next, Sonogashira cross-coupling was applied on 5-iodo-2'-deoxycytidine triphosphate (**Scheme 10-A**) in presence of CuI, $Pd(OAC)_2$ and TPPTS or CuAAC click reaction on 5-azidomethyl-2'-deoxyuridine monophosphate[117] with functionalized 1,1-dioxido-thiazetidin moiety (**SU2**) in presence of $CuSO_4*5H_2O$ and sodium ascorbate in H_2O/DMSO/tBuOH reaction mixture (9:3:1) were performed at room temperature for 18 h (**Scheme 10-B**). As expected, NMR analysis did not give any information about the state of the ring (either opened or closed). Instead, by mass analysis was observed the addition of water (+18) on modified monophosphates and triphosphates, which means the opening of the ring of 1,1-dioxidothiazetidin moiety (**Scheme 10**). The addition of water

could be due to copper presence in reaction conditions which coordinates to the ring allowing the water addition and consequently its opening.

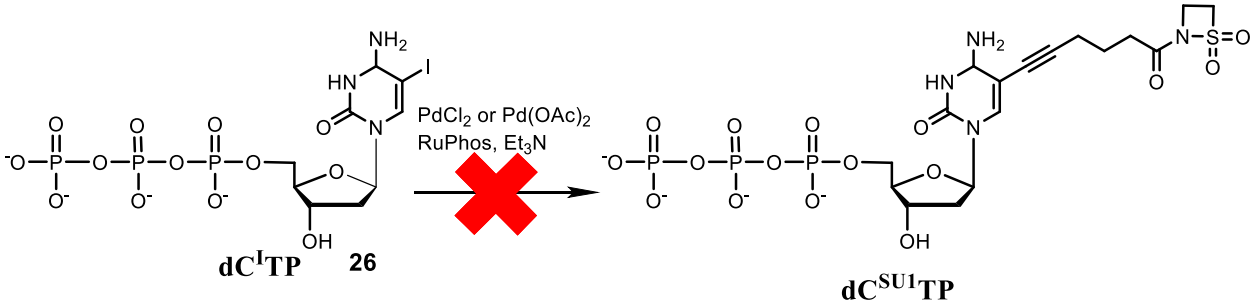


Scheme 10. Synthesis of β-sultam modified nucleotides (MP and TPs). Conditions: Sonogashira cross coupling: i) **SU1** or **SU2** (1.3 equiv.) $Pd(OAc)_2$ (0.2 equiv.), CuI (10%), TPPTS (10%), DIPEA (10 equiv.), H_2O/MeCN (2:1), 18 h at RT; ii) CuAAC click reaction: **SU1** or **SU2** (1.3

equiv.), $CuSO_4*5H_2O$ (0.1 equiv.), Na ascorbate (0.4 equiv.), H_2O/DMSO/tBuOH (9:3:1), 18 h at RT.

3.2.1.2 Copper-free cross-coupling of dC^{I}TP with SU1

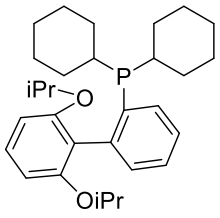
In order to avoid the presence of copper and keeping the ring closed, the copper-free cross coupling reaction was performed (**Scheme 11**). As reagents of copper-free cross-coupling were tested two different kind of Pd catalyst, such as $PdCl_2$ or $Pd(OAc)_2$ and two different types of additives, such as pyrrolidine or Et_3N and additionally RuPhos as ligand. In all cases, the reaction was conducted at RT for 18 h or 60 °C for 4 h, in H_2O/MeCN (1:2) or H_2O/THF (1:2) but in all cases no product formation was observed (**Table 8**).



Scheme 11. Trial of copper-free cross-coupling reaction of **dC^{I}TP** with **SU1** linker.

Table 8. Copper free cross-coupling conditions

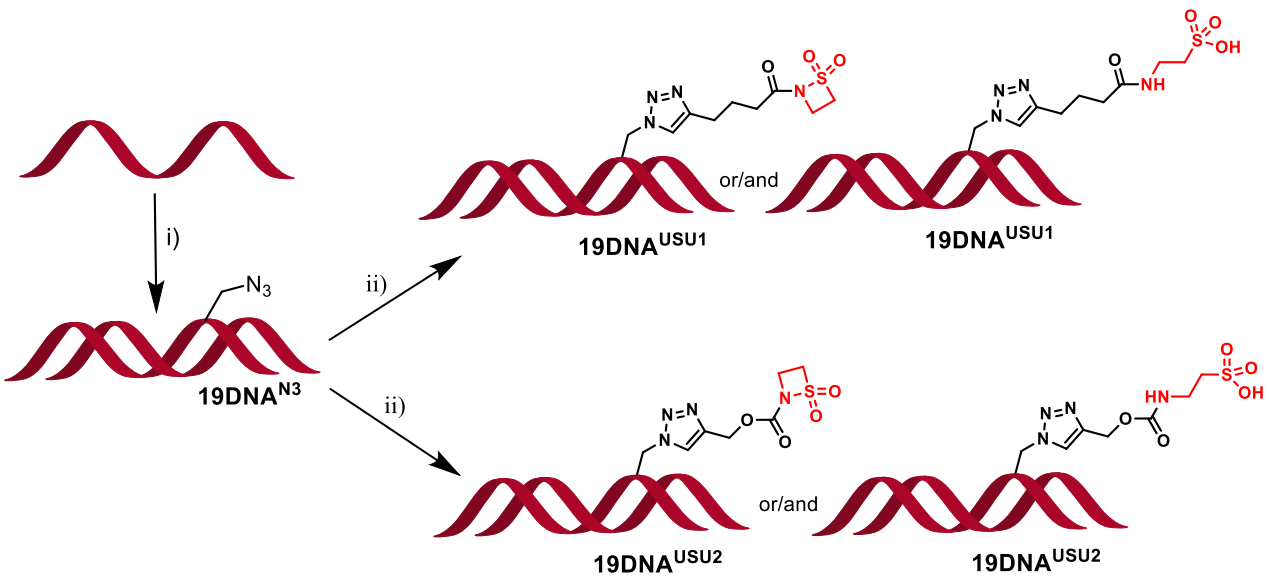
Starting compound	SU1 linker	Catalyst	Ligand	Temperature	Time	Additives	Solvent
dC^{I}TP (1 equiv.)	1.2 equiv.	$PdCl_2$ (0.01 equiv.)	-	RT 60 °C	18 h 4 h	Pyrrolidine (5 equiv.)	H_2O/MeCN (1:2)
dC^{I}TP (1 equiv.)	1.2 equiv.	$Pd(OAc)_2$ (0.025 equiv.)	RuPhos (0.075 equiv.)	RT 60 °C	18 h 4 h	Et_3N (3 equiv.)	H_2O/THF (1:2)



RuPhos ligand

3.2.2 Click reaction on N_3- modified DNA with SU1 or SU2

Because of the ring opening under click reaction conditions on nucleotides, the click reaction on DNA, to attach postsynthetically **SU1** or **SU2** linkers (**Scheme 12**), was tried as milder approach by reducement of copper amount used. The presence of azido group, as click handle, is needed, therefore DNA was functionalized with N_3- (azido group) by PEX synthesis using **dU^{N3}TP**, sequences and primer shown in **Table 9**.



Scheme 12. Synthetic scheme of **19DNA^{UN3}** and click reaction between **19DNA^{UN3}** and **SU1** or **SU2** linker. Conditions: i) PEX reaction: template **Temp^{19}_^{1T}** (3 µM), primer **Prim A_FAM** (3 µM), dGTP (4 mM), **dU^{N3}TP** (4 mM) at 60 °C for 30 min, ii) CuBr (2 mM or 4 mM), BTTAA ligand (50 mM), Na ascorbate (10 µM) and **SU1** or **SU2** linker (2 mM) in DMSO/tBuOH (3:1) at 37 °C for 18 h.

Table 9. List of primers and templates used and ONs synthesized in the study.

Name	Sequence (5'→3')	Length
Prim^{A}	5`-CATGGGCGGCATGGG-3`	15 nt
Prim^{A}_FAM [a]	5`-CATGGGCGGCATGGG-3`	15 nt
Temp^{19}_^{1T}	5`-CCCACCCATGCCGCCCATG-3`	19 nt
Temp^{19}_^{1C}	5`-CCCGCCCATGCCGCCCATG-3`	19 nt

Temp^{19_1T}_bio [b]	5`-CCCACCCATGCCGCCCATG-3`	19 nt
Temp^{19_1C}_bio [b]	5`-CCCGCCCATGCCGCCCATG-3`	19 nt
19ON^{N3}	5`-GGG**U$^{N3}$**GGGTACGGCGGGTAC-3`	19 nt
19DNA^{N3}	5`-GGG**U$^{N3}$**GGGTACGGCGGGTAC-3`	19 nt

[a] 5´-6-FAM-labelled products, [b] 5`-biothinylated.

For PEX reaction was used **Temp^{19_1T}** and **Prim^{A}_FAM** and after incubation at 60 °C for 30 min **19DNA^{N3}** was obtained and visualized by PAGE gel (**Figure 54**). In order to get the oligonucleotide **19ON^{N3}** (for sequence synthetized see **Table 9**) was used **Temp^{19_1T}_bio** and **Prim^{A}** for PEX and then was applied the magnetoseparation and MALDI-TOF analysis (**Table 10**).

Table 10. MALDI data of modified ONs.

ON	M (calc.) [Da]	M (found) [Da]
19ON^{UN3}	6005.9	6007.1

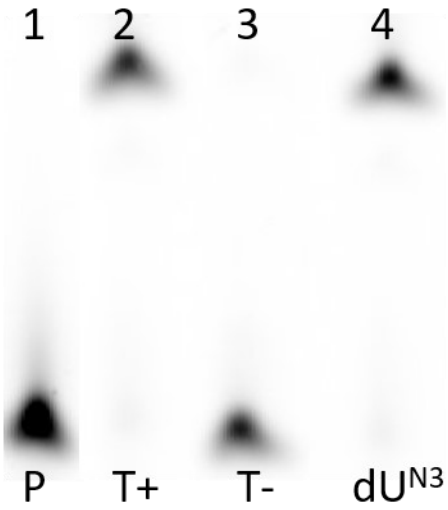


Figure 54. Denaturing PAGE analysis of PEX using KOD XL DNA polymerase and template **Temp^{19_1T}** (lane 2-4). (P): primer **Prim A_FAM**, lane 1; (T+): natural dGTP, dTTP, lane 2; (T-): negative control without dCTP, lane 3; (dU^{N3}): modified DNA containing **dU^{N3}TP**, dGTP, lane 4.

As next step, click reaction on **19DNA^{N3}** (prior synthesized by PEX), was performed with **SU1**. For click reaction on **19DNA^{N3}** (5 μM) were used two different concentration of CuBr (2 mM or 4 mM) to improve the yield, 2-(4-((Bis((1-(*tert*-butyl)-1*H*-1,2,3-triazol-4-yl)methyl)amino)methyl)-1*H*-1,2,3-triazol-1-yl)acetic acid ligand (BTTAA, 50 mM) and Na ascorbate (10 μM) dissolved in DMSO/tBuOH (3:1) prior prepared as stock solutions. Then to the reaction **SU1** or **SU2** linker (2 mM) was added and left stirring for 18 h at 37 °C. Click reaction on DNA was attempted in order to avoid the PEX temperature of 60 °C and analysed by PAGE gel analysis (**Figure 55**, only shown for **SU2** linker used). From PAGE gel analysis the formation of a second band was observed, by increased time, which supposed to be the click product.

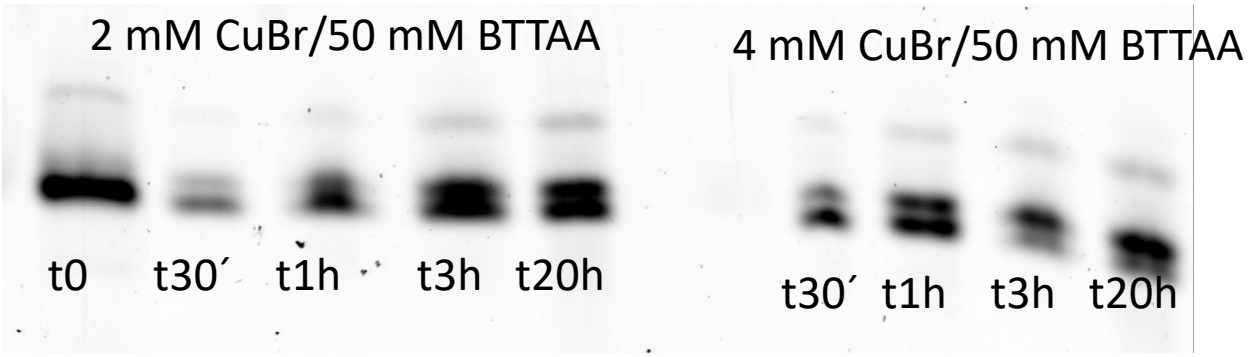


Figure 55. Denaturing PAGE gel analysis of click reaction between **19DNA^{N3}** and **SU1** linker using 2 mM or 4 mM CuBr/BTTA.

Therefore, after click reaction on DNA, the reaction mixture was purified to remove the excess of copper and the magnetoseparation was applied on **19DNA^{USU2}** to isolate the single strand **19ON^{USU2}** and then analysed by MALDI-TOF. Unfortunately, only traces of **19ON^{USU2}** in closed and in its hydrolysed form were observed as shown by mass spectrum (**Figure 56**), whereas none of **19ON^{SU1}** was found and only unreacted starting material **19ON^{N3}** was observed.

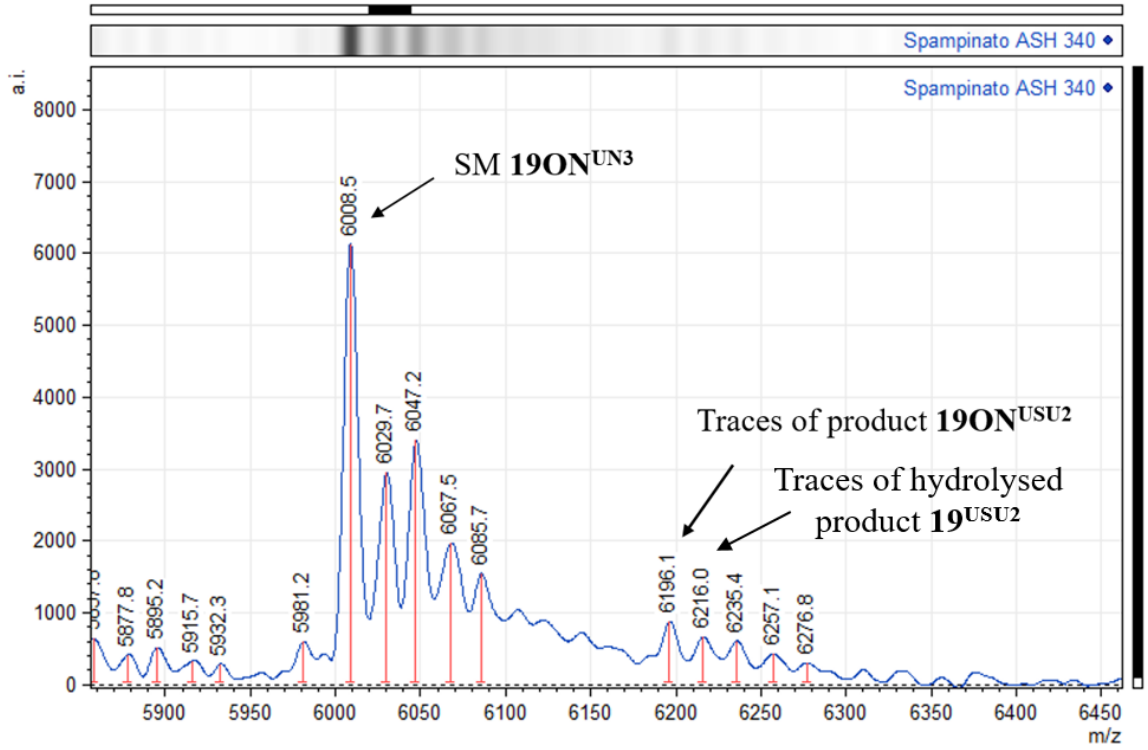


Figure 56. MALDI-TOF analysis of **19DNA^{USU2}**.

Based on unsuccessful experiments on β-sultam modified nucleotides and click reaction performed on DNA, we concluded that the reactive moiety was not stable and was hydrolysed. Therefore, further studies with β-sultam modified nucleotides and DNA probes to cross-link serine containing peptides and proteins were not proceeded. Therefore, targeting serine remains a challenge for future research.

3.2.3 Design and synthesis of modified nucleotides bearing 1,1-dioxido-3-oxo-4,4-dimethyl-thiazetidine (4,4-dimethyl-3-oxo-β-sultam) for cross-linking with serine, their enzymatic incorporation and testing of reactivity

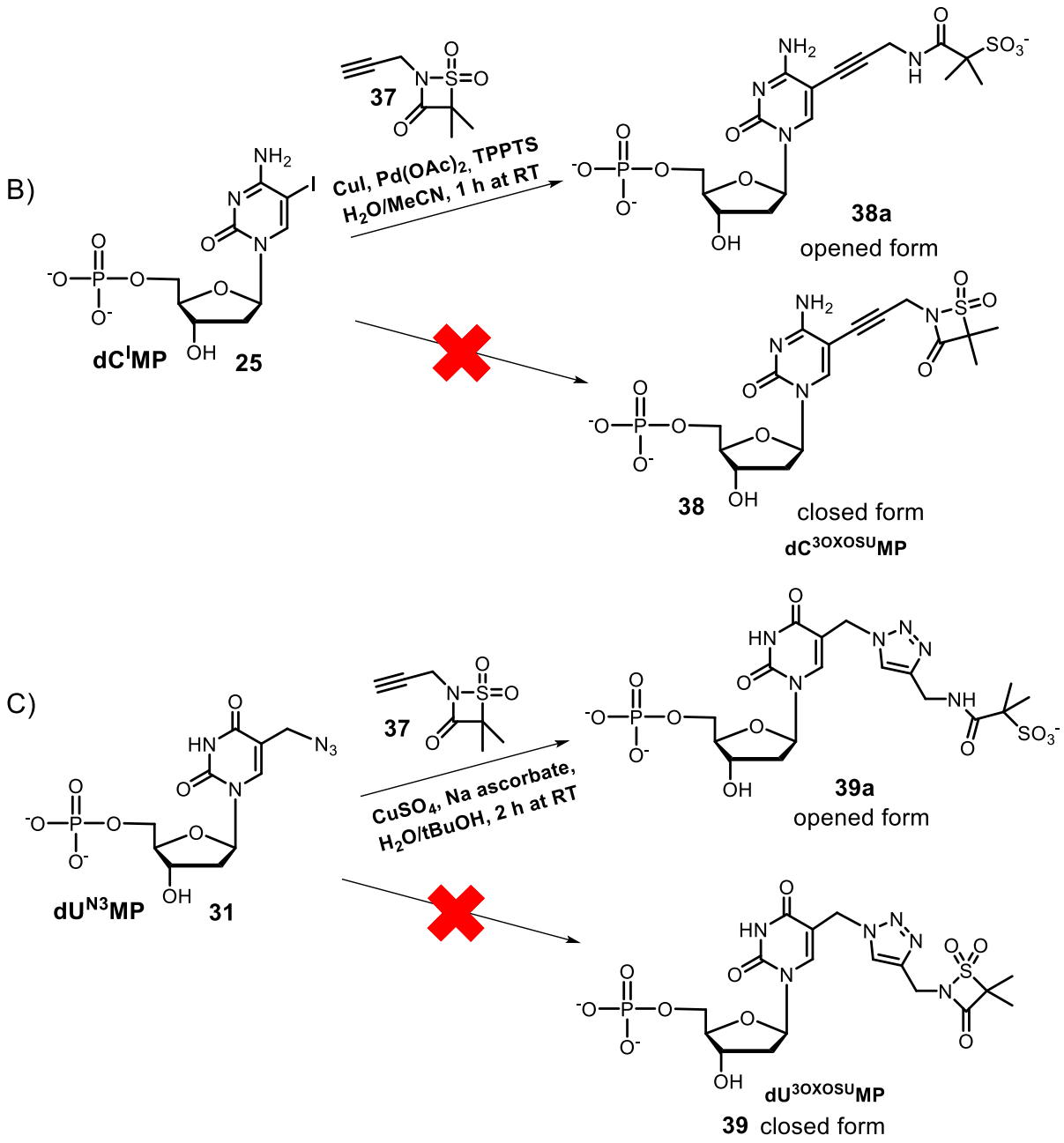
3.2.3.1 Synthesis of modified monophosphates dC^{3OXOSU}MP and dU^{3OXOSU}MP

An other attempt to target serine aminoacid was to test the reactivity of 1,1-dioxido-3-oxo-4,4-dimethyl-thiazetidine (4,4-dimethyl-**3-oxo-β-sultam**) and to be linked to nucleotides or DNA for further studies.

According to published literature[83,84,85], 4,4-dimethyl-3-oxo-β-sultam (**OXOSU**) exhibits high reactivity towards serine, therefore, was used to be inserted in nucleotides and subsequently in DNA.

Based on already optimized conditions[83,84,85], 4,4-dimethyl-3-oxo-β-sultam (**OXOSU**) was prepared according to published synthetic pathway involving the ring closure of the corresponding (chlorosulfonyl)acetyl chloride with propargyl amine (**Scheme 13-A**). The alkyne handle of linker was used to attach the moiety at position 5 to 5-iodo-2'-deoxycitidine monophosphate (**25**, **dC^{I}MP**) through Sonogashira cross-coupling or to 5-azido-2'-deoxyuridine monophosphate (**31**, **dU^{N3}MP**) through CuAAC click reaction (**Scheme 13-B,C**). After the reactions, hydrolysis of the ring was observed with its consequent opening and formation of compound **38a** instead of desired product **38** (**dC^{3OXOSU}MP**) or **39a** instead of desired product **39** (**dU^{3OXOSU}MP**), (**Scheme 13-B, C**).



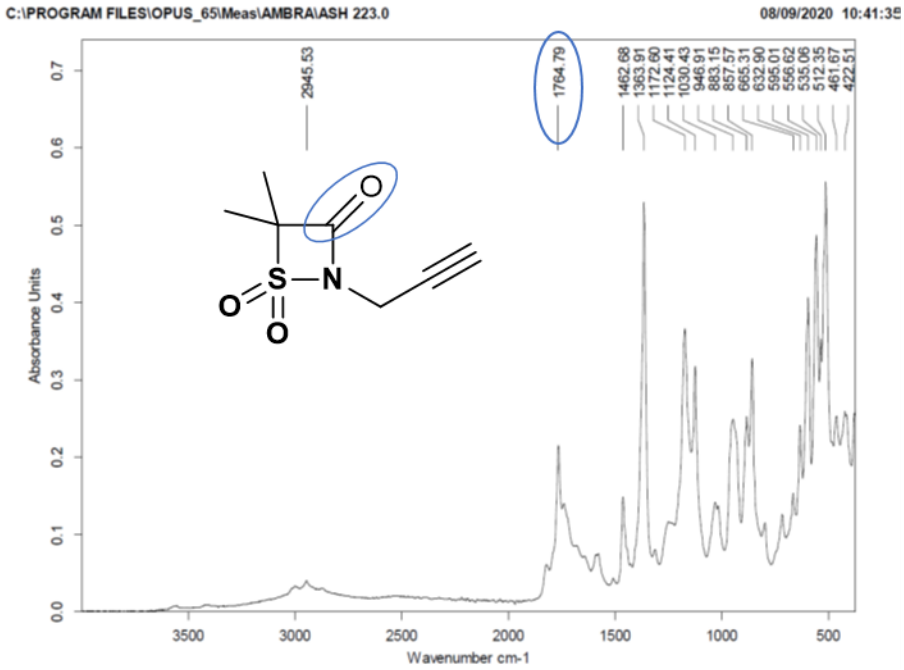


Scheme 13. A) Synthetic scheme of 4,4-dimethyl-3-oxo-β-sultam linker **37** (**OXOSU**); B) Sonogashira cross-coupling synthetic scheme of **dC^{3OXOSU}MP**; C) CuAAC click reaction synthetic scheme of **dU^{3OXOSU}MP**.

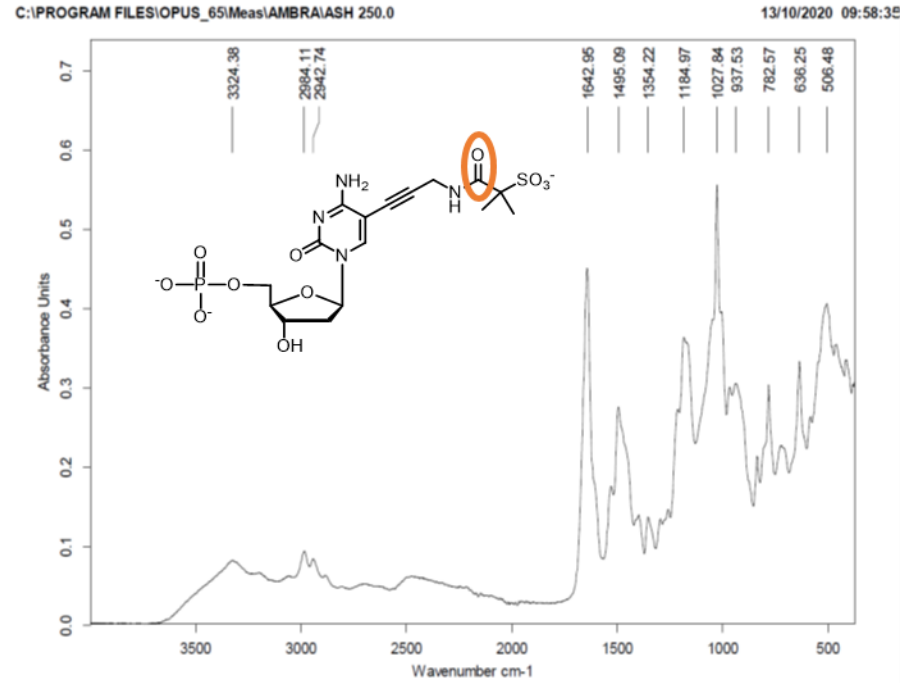
The ring-closure was checked by IR spectroscopy, via characteristic C=O IR stretching frequencies of 3-oxo-β-sultam, occurring around 1790 cm^{-1} which was in accordance with 4,4-dimethyl-3-oxo-β-sultam (**Figure 57-A**). Unfortunately this characteristic peak was not detected for modified

monophosphate **dC^{3OXOSU}MP** or **dU^{3OXOSU}MP** but exhibited instead 1642.95 cm^{-1} referring to C=O stretching of opened ring only in case of **dC^{3OXOSU}MP** (**Figure 57-B**). For **dU^{3OXOSU}MP** the peak of water was really broaden hiding the C=O stretching (**Figure 57-C**).

A)



B)



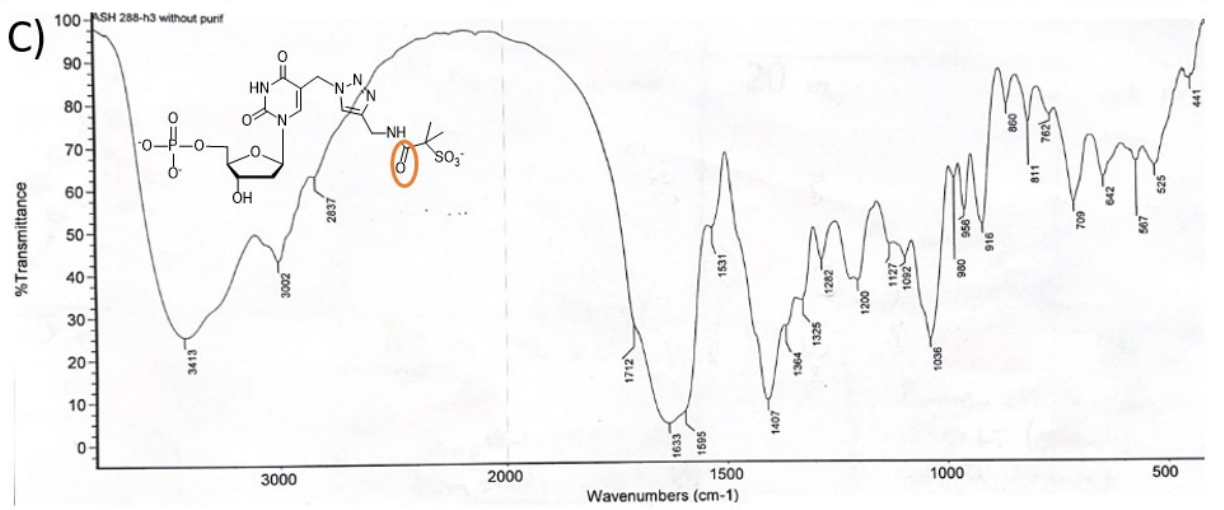
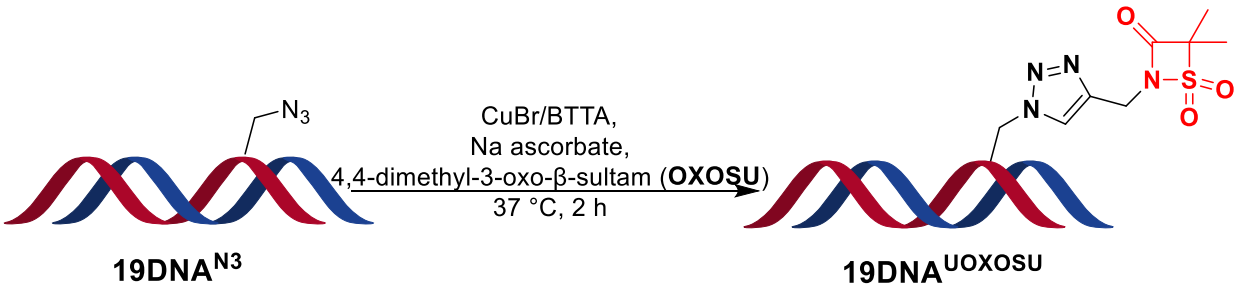


Figure 57. A) IR spectrum of 4,4-dimethyl-3-oxo-β-sultam; B) IR spectrum of **38a** compound; C) IR spectrum of **39a** compound.

3.2.3.2 Click reaction on DNA using OXOSU building block

Nevertheless, click reaction on DNA, functionalized with N_3 group, was still performed with 4,4-dimethyl-3-oxo-β-sultam linker in order to avoid high temperature of PEX (60 °C) and also high amount of copper used on nucleotides, which could probably affect the stability of the ring. For click reaction on DNA was used CuBr (2 mM), BTTAA ligand (50 mM) dissolved in H_2O/DMSO/tBuOH (9:3:1).

Then N_3- modified DNA (5 μM) was mixed with CuBr/BTTAA, sodium ascorbate (10 μM) and **OXOSU** linker (10 mM). The reaction was kept for 2 h at 37 °C in a thermal cycler (**Scheme 14**).



Scheme 14. Synthetic scheme of click reaction between **19DNA^{N3}** and 4,4-dimethyl-3-oxo-β-sultam (**OXOSU**) linker.

The click reaction on **19DNA^{N3}** with **OXOSU** linker was monitored at different time, using 2 mM or 4 mM concentration of CuBr and 50 mM of BTTAA. Afterwards, the reaction mixture was purified to remove the excess of copper and magnetoseparation[113] was applied to isolate the single strand **19ON^{3OXOSU}** and analysed by MALDI-TOF. Only traces of **19ON^{3OXOSU}** was observed and the rest was the unreacted **19DNA^{N3}** (**Table 11**).

Table 11. MALDI data of modified ON after click reaction with **OXOSU**

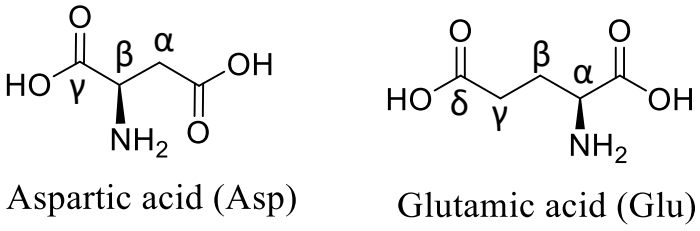
ON	M (calc.) [Da]	M (found) [Da]
19ON^{3OXOSU}	6193.9	6195.8
19ON^{N3} (unreacted)	6006.9	6008.4

In opposite of what expected, **OXOSU**- modified nucleotides and DNA reveals not to be more stable than previously tested **SU1-** or **SU2-** modified nucleotides and DNA and the opening of the ring was proved by IR and mass analysis. Therefore 3-oxo-β-sultam project was not proceeded for further studies to cross-link serine containing peptides and proteins.

3.3 PART III

3.3.1 Design and synthesis of modified nucleotides bearing azirine for cross-linking with glutamic and aspartic acids, enzymatic incorporation and testing of reactivity

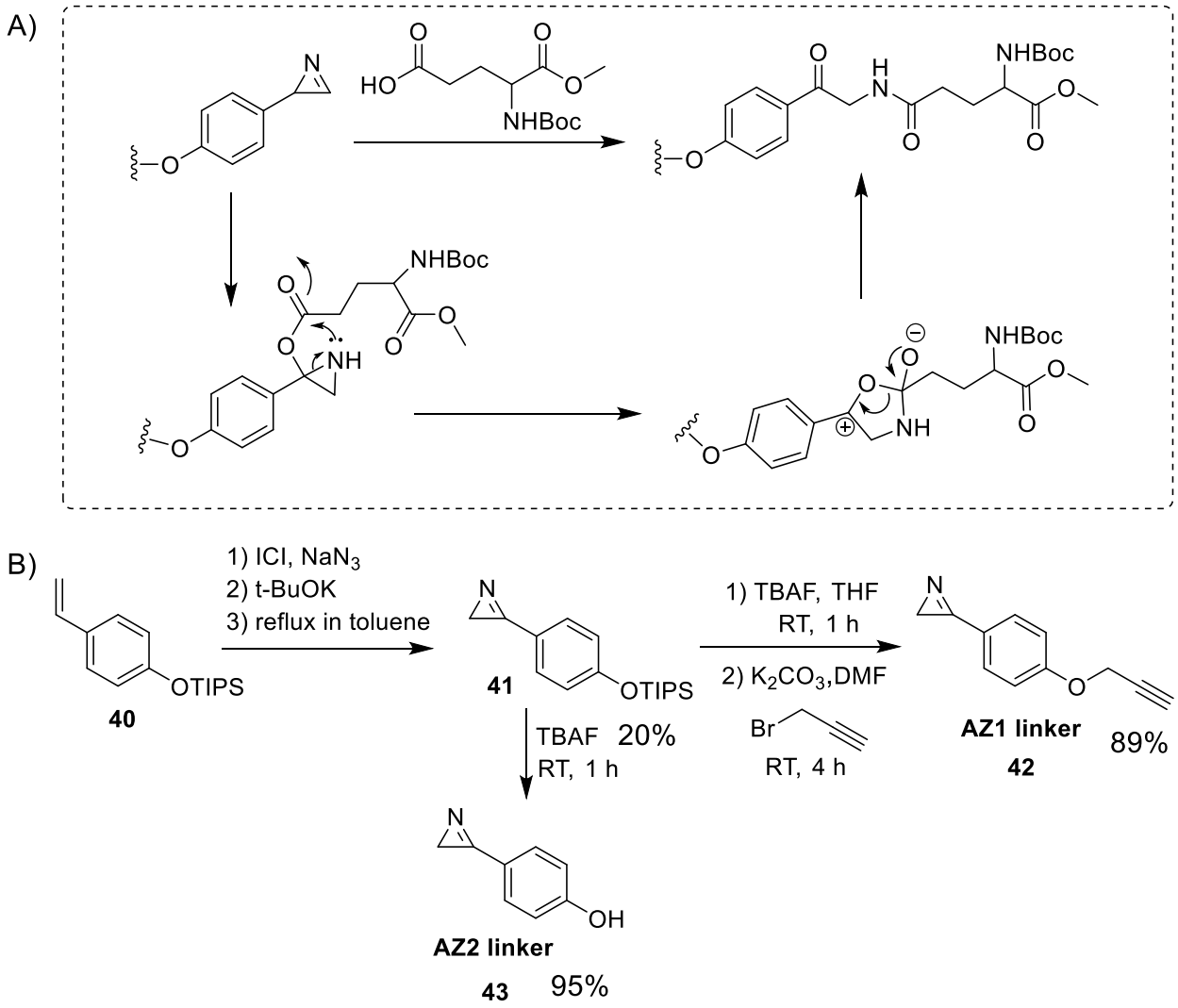
To extend the portfolio of aminoacids to target, the study was focused on glutamic and aspartic acids (Glu/Asp). In particular, an other reactive moiety was found in literature[88] to be specific to bind covalently Glu/Asp, which was the azirine ring. The azirine moiety was chosen as reactive group to be linked to nucleotides or DNA for targeting carboxyl aminoacids, glutamic (Glu) and aspartic acids (Asp) (**Scheme 15**) or Glu/Asp-containing peptides/proteins.



Scheme 15. Chemical structure of aspartic (Asp) and glutamic (Glu) acids.

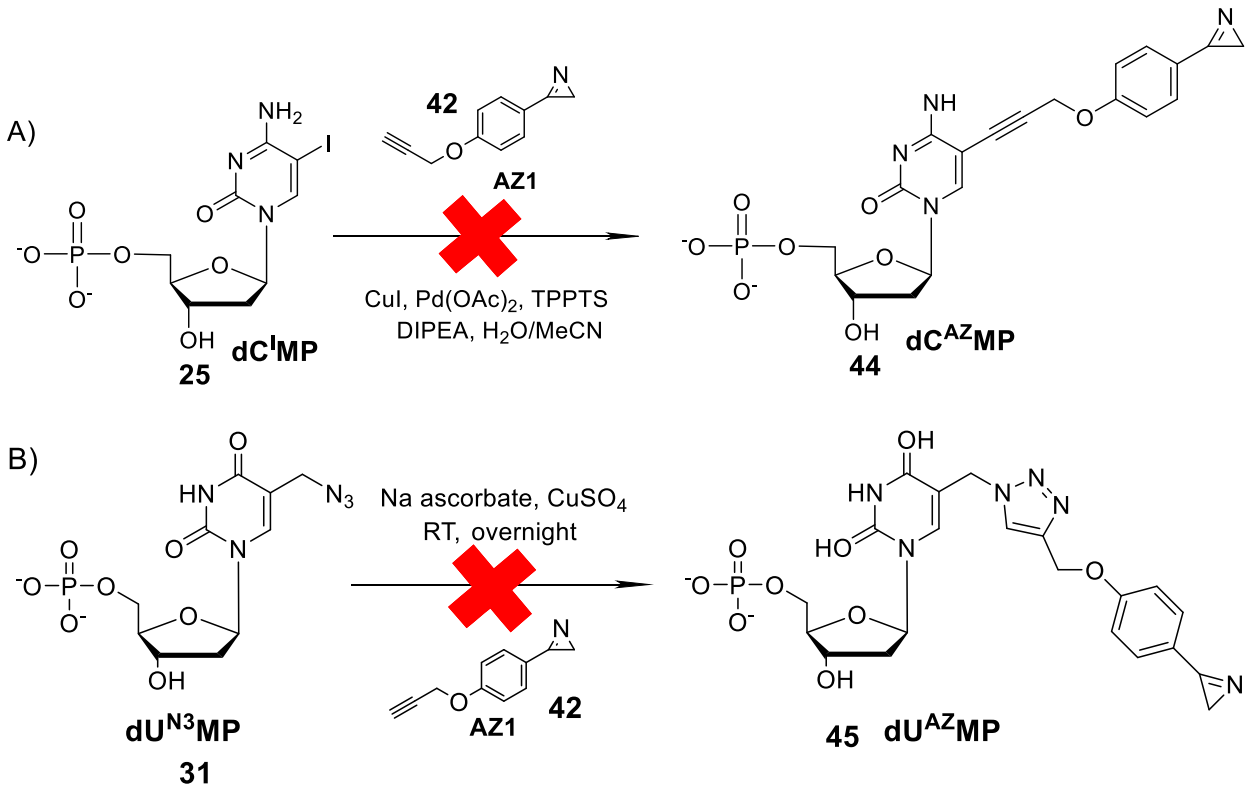
3.3.1.1 Synthesis of azirine-modified nucleoside (dU^{AZ})

In the proposed bioconjugation mechanism, as shown in **Scheme 16-A** the carboxylic acid reacts firstly with 3-phenyl-2*H*-azirine and then is followed by intramolecular nucleophilic addition and ring expansion to form the *N*-phenacylacetamide adduct. The synthesis of reactive linker to insert into nucleotide or DNA for further studies was done according to published literature.[88]
(**Scheme 16-B**). The synthesis started from triisopropyl(4-vinylphenoxy)silane (**40**) in which the azirine ring was built from double bond of p-tosylstirene to give the compound **41** and then after deprotection of OH group was obtained the **AZ2** linker (**43**) in almost quantitative yield of 95% which was used directly in an alternative synthetic pathway described further. On the other hand, alkyne handle is necessary to attach at position 5 of nucleotide by Sonogahira or CuAAC reaction. Therefore, to the deprotected **41** was added the base K_2CO_3 and propargyl bromide. The reaction was stirred at room temperature for 4 h. The bromine of propargyl bromide was substituted by OH- group of **41** to obtain the desired reactive moiety, **AZ1** linker (**42**) in very good yield of 89%.



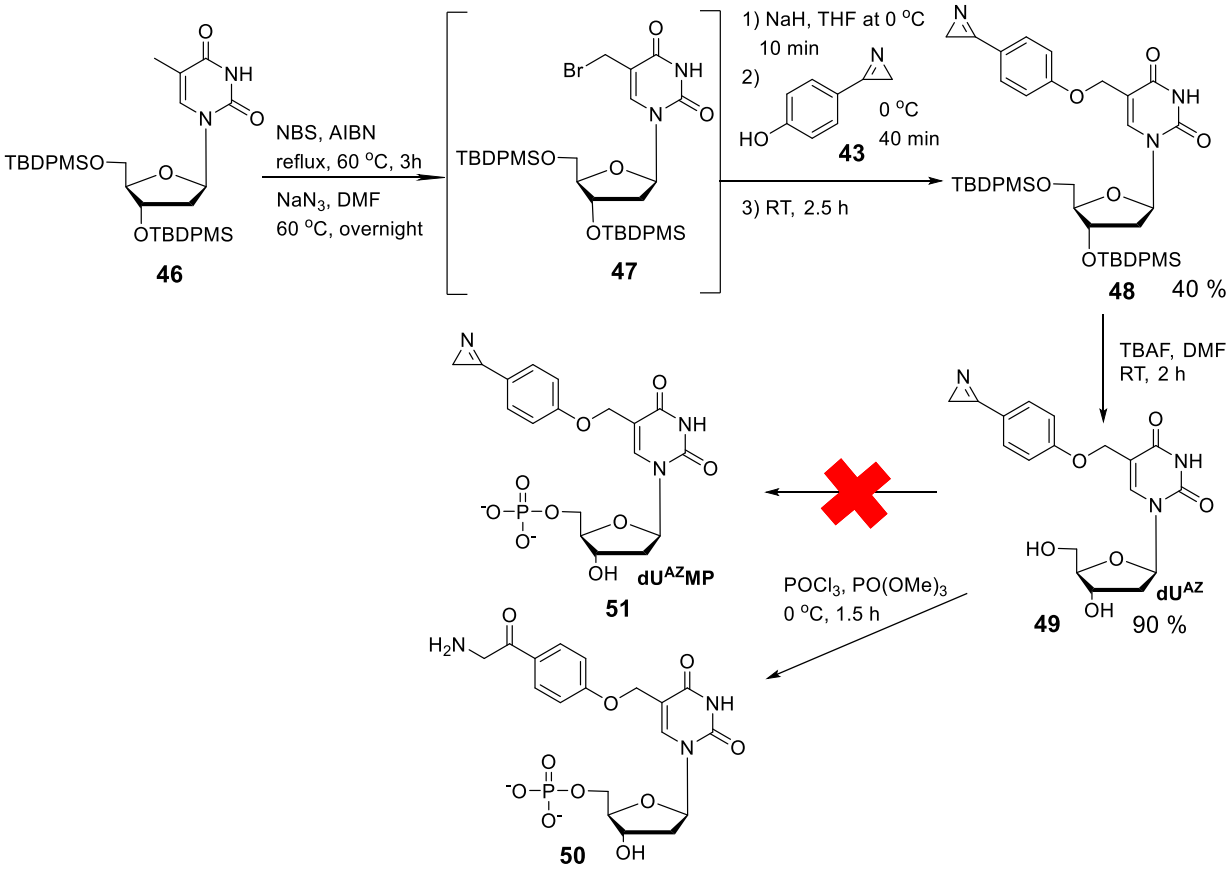
Scheme 16. A) Proposed reaction mechanism of the azirine containing moiety with *N*-Boc-protected glutamic acid, B) Synthetic scheme of **AZ1** and **AZ2** linkers.[88]

Sonogashira cross-coupling on 5-iodo-2'-deoxycytidine monophosphate (**dC^{I}MP, 25**) or CuAAC click reaction on 5-azido-2'-deoxyuridine monophosphate (**dU^{N3}MP, 31**) were attempted. In case of click reaction was observed a complex mixture which was not possible to decipher. After Sonogashira cross coupling was observed only the starting compound **dC^{I}MP** unreacted and homocoupling of the reactive linker (**Scheme 17**). Desired final compounds (**dC^{AZ}MP, 44** or **dU^{AZ}MP, 45**) were not isolated.

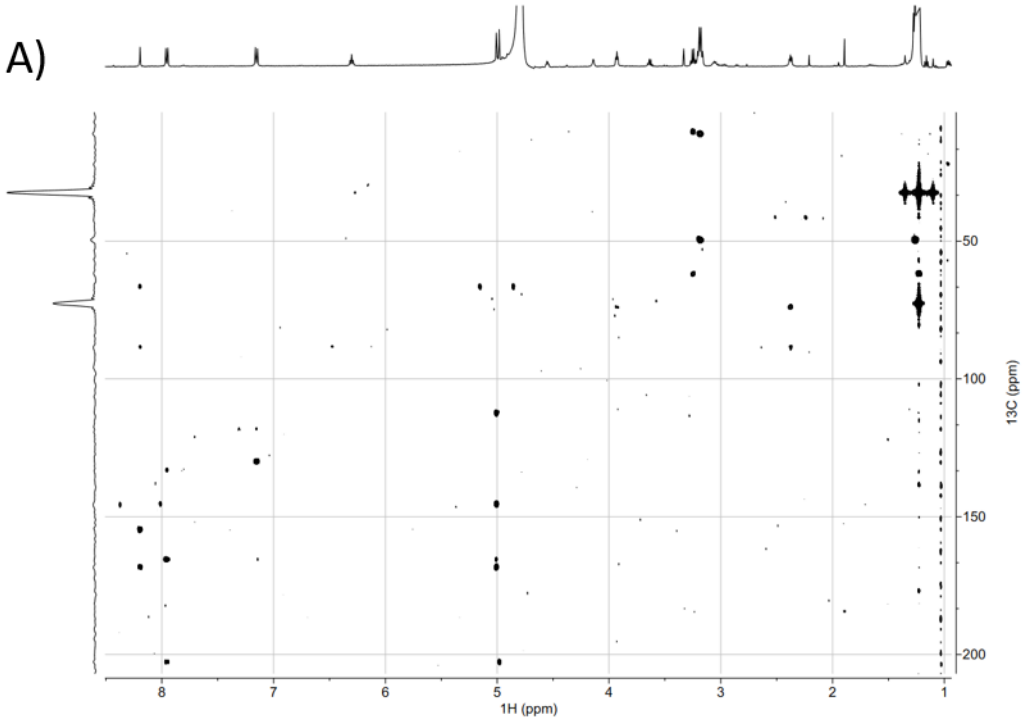


Scheme 17. Synthetic scheme of A) Sonogashira cross coupling between **dC^{I}MP** and **AZ1 linker** and of B) click reaction between **dU^{N3}MP** and **AZ1 linker**.

In order to use milder reaction conditions, an other synthetic pathway with no use of copper was attempted. The synthesis started from protected 2'-deoxythymidine (**46**) which was converted by allylic bromination into 5-bromomethyl-2'-deoxyuridine (**47**). The intermediate **47** was reacted in one pot reaction with compound **43** (described in **Scheme 17** above) and by nucleophilic substitution was introduced into nucleoside (**47**) giving the compound **48**. This latter compound was then deprotected to achieve the desired product **dU^{AZ} 49** (**Scheme 18**). Nevertheless, after monophosphorylation reaction conditions the azirine ring was hydrolyzed showing to be reactive even towards neutral nucleophile like water. HMBC and mass spectra confirmed the formation of opened form (compound **50**) instead of desired **dU^{AZ}MP 51** (**Figure 58-A,B**).



Scheme 18. Synthetic scheme of **dU^{AZ}MP** with copper free synthetic pathway.



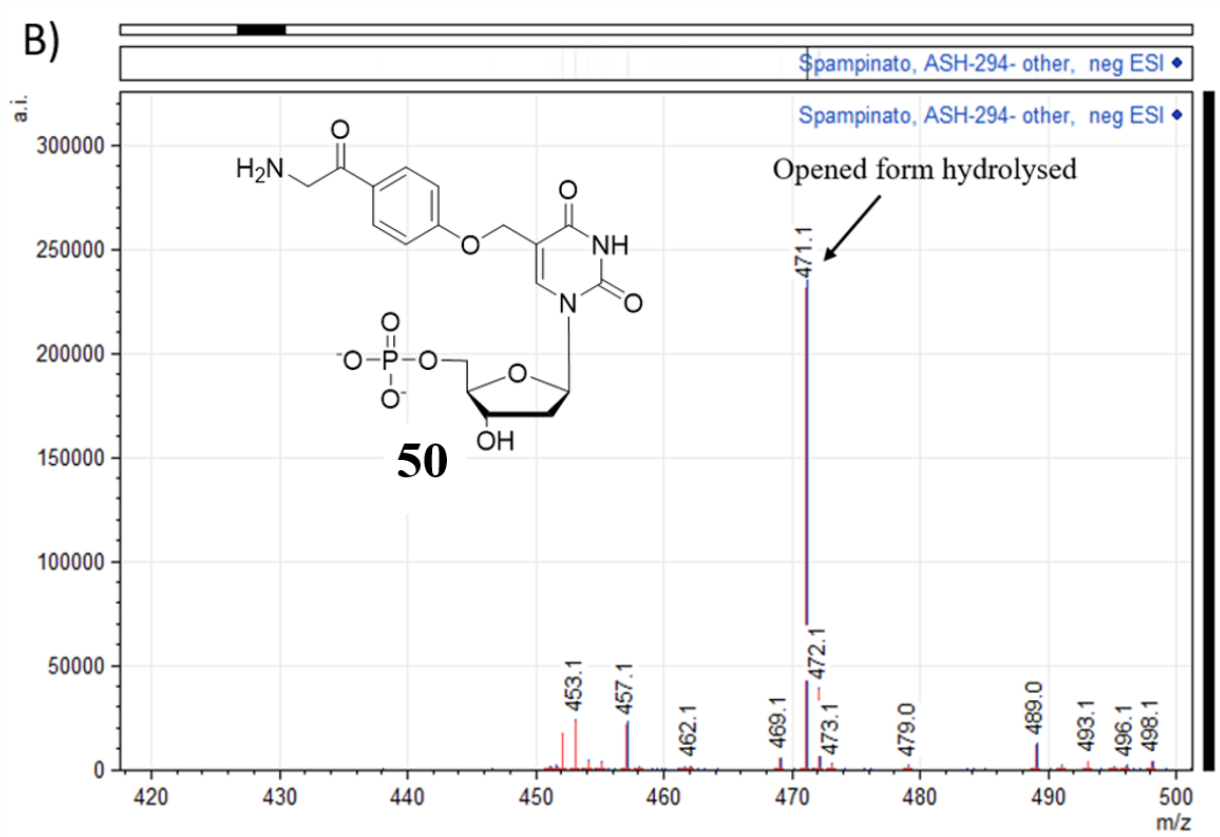


Figure 58. A) HMBC spectrum and B) mass spectrum of compound **50**.

The goal of bioconjugation projects was to synthesize new modified nucleotides and DNA bearing reactive moiety to target specific aminoacids.

Based on results from mass and NMR analyses, the azine ring was found to have been hydrolysed, meaning to be unstable and unreactive towards Glu- or Asp- aminoacids which were not used for further applications. Hence, targeting these aminoacids represents still a challenge for future research.

3.4 PART IV

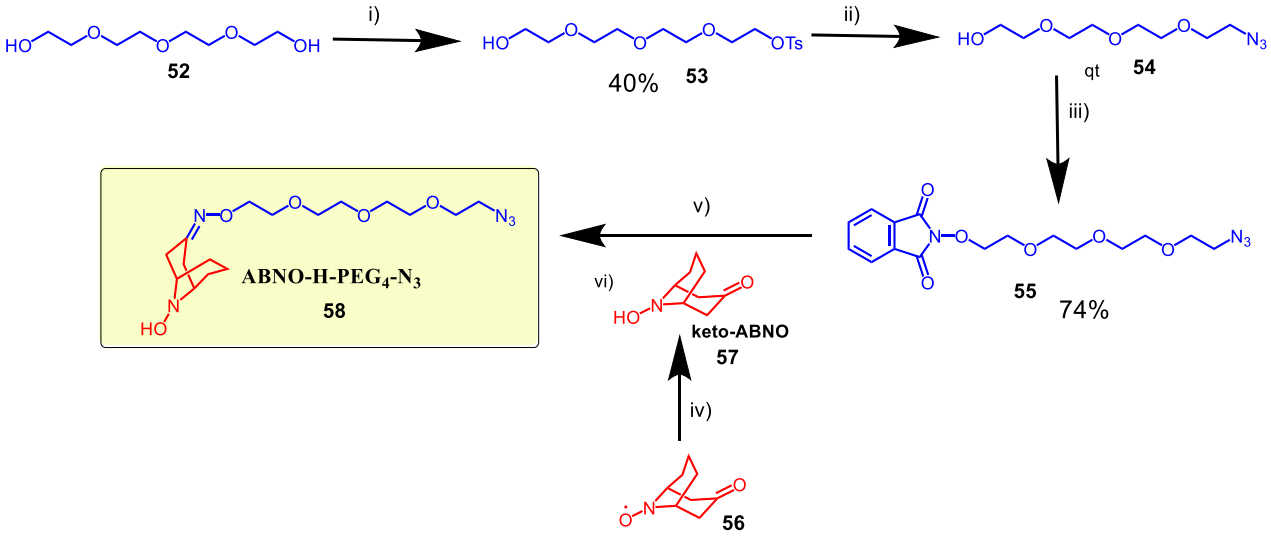
3.4.1 Design and synthesis of modified nucleotides bearing 9-hydroxy-9-azabicyclo[3.3.1]nonane (keto-ABNOH) for cross-linking with tryptophan, enzymatic incorporation and testing of reactivity

In order to achieve one of the main goal of thesis to target specific aminoacids, the next candidate chosen for our study was the keto-ABNO (OH) moiety, which was previously reported to be very reactive towards tryptophan aminoacid and containing protein.[93]

The keto-ABNO (OH) moiety was used further to functionalize nucleotides and DNA and testing their reactivity towards tryptophan (Trp) or Trp- containing peptides and proteins.

3.4.1.1 Synthesis of keto-ABNOH modified monophosphate ($dU^{ABNOH}MP$) and triphosphate ($dU^{ABNOH}TP$)

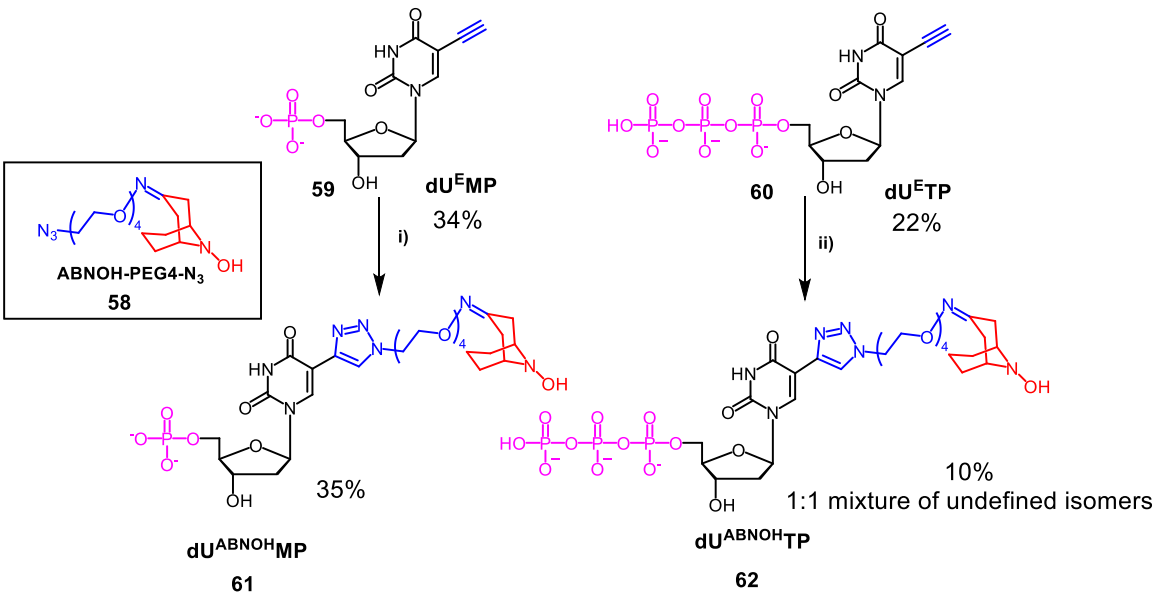
The synthesis of keto-ABNOH building block (**58**) was prepared according to published procedures (**Scheme 19**).[93]



Scheme 19. Synthesis of ABNO-H-PEG_4-N_3 moiety (**5**). Conditions: i) TosCl, THF/H_2O, NaOH at 0 °C, 5 h; ii) NaN_3, dry DMF at 70 °C, 8 h; iii) N-Hydroxyphtalimide, DIAD, PPh_3, THF at 0 °C for 20 min and then at 25 °C, 20 h; iv) ascorbic acid, MeOH at 25 °C, 1 h; v) NH_2NH_2, dry EtOH at 25 °C, 26 h; vi) keto-ABNO, AcOH at 50 °C, 16 h.

To introduce the keto-ABNO-H moiety (**58**) into 5 position of 5-ethynyl-2'-deoxyuridine monophosphate and triphosphate **dU^EMP** (compound **59**), **dU^EMP** (compound **60**), they were synthesized according to published protocol.[118] The click reaction (CuAAC) was performed on 5-ethynyl-2'-deoxyuridine monophosphate and triphosphate **dU^EMP** (compound **59**), **dU^ETP** (compound **60**) and performed in presence of $CuSO_4$*$5H_2O$, sodium ascorbate and TBTA in

DMSO. The reactions proceeded giving the desired products in acceptable yields **dU^{ABNOH}MP** (compound **61**, 35% yield) after 70 min and **dU^{ABNOH}TP** (compound **62**, 10% yield) after 3 h at 25 °C as mixture of 1:1 undesired isomers (**Scheme 20**).



Scheme 20. Synthesis of keto-ABNOH modified 2`-deoxyuridine monophosphate **dU^{ABNOH}MP** and triphosphate **dU^{ABNOH}TP**. Conditions: **dU^{E}MP** or **dU^{E}TP**, ABNOH-PEG$_4$-N$_3$, CuSO$_4$*5H$_2$O, sodium ascorbate, TBTA, DMSO, 25 °C for 70 min (i) or 3 h (ii), respectively.

3.4.1.2 Investigation on the origin of signal in NMR spectra of dU^{ABNOH}TP

By the NMR analysis of **dU^{ABNOH}TP** was observed the presence of two identical set of signals (1:1). To exclude the possibility of the second compound being an impurity or decomposition, the NMR sample was heated up to 75 °C. ^{1}H NMR and ^{31}P NMR were recorded (**Figure 59**). At that temperature (75 °C) two set of signals in ^{1}H NMR and ^{31}P NMR merge in one. Therefore, this demonstrates that at 25 °C was observed 1:1 mixture of interchangeable isomers. This behaviour was observed only in case of triphosphate, hypothesizing three phosphates groups creates interactions with the ABNO-end.

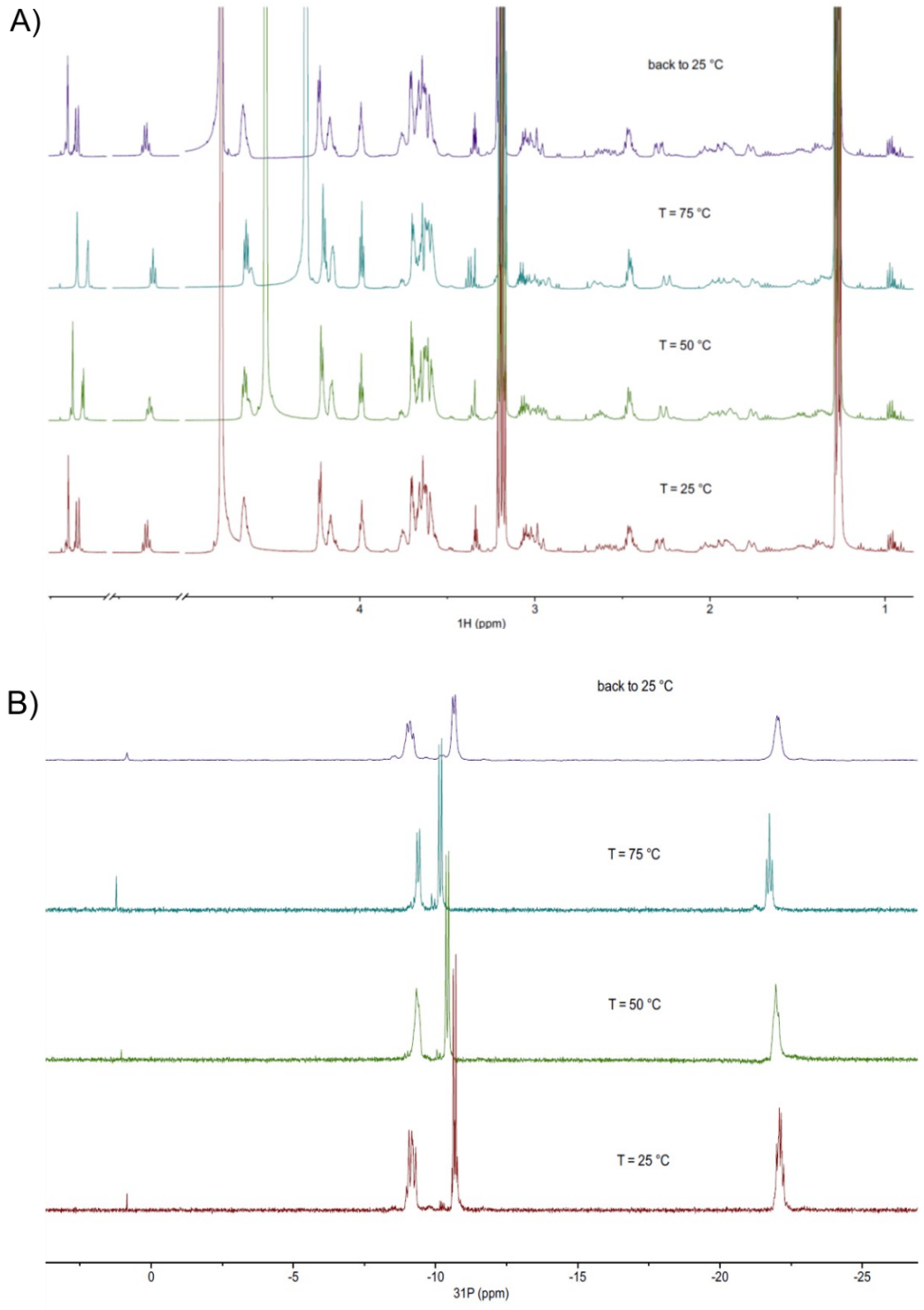
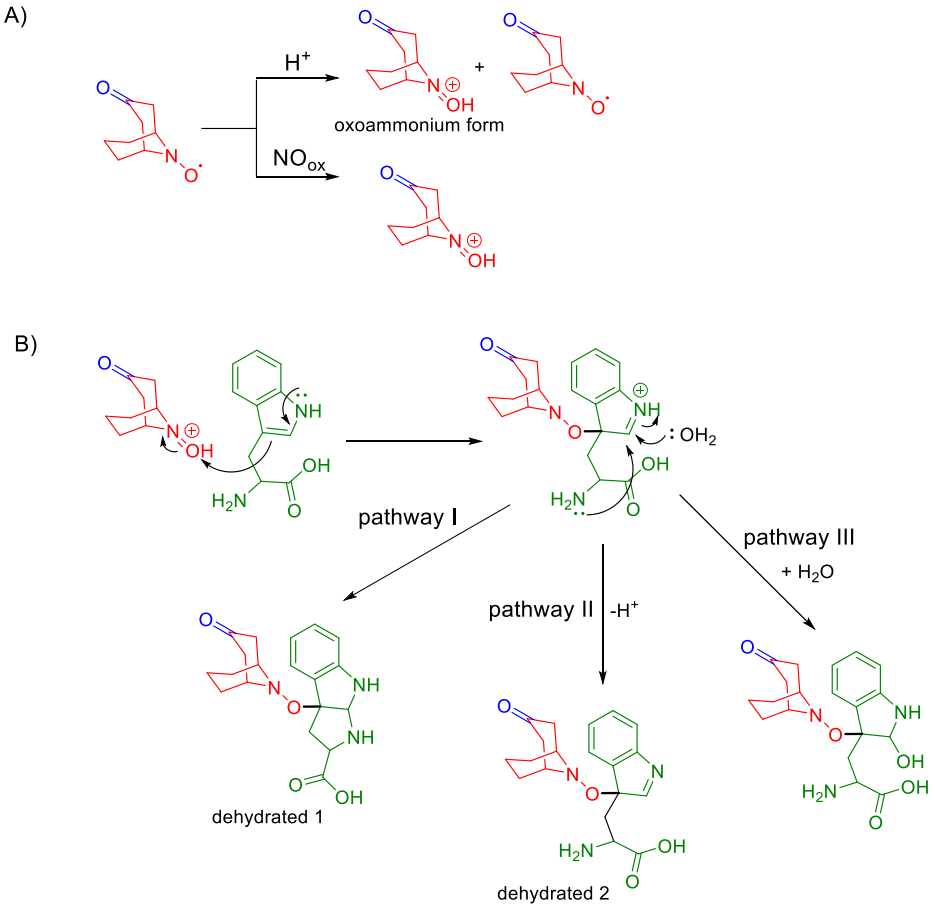


Figure 59. A) ^{1}H NMR spectra and B) ^{31}P NMR spectra at different temperature of **dU^{ABNOH}TP**.

3.4.1.3 EPR measurements of N-O radical formation on keto-ABNOH

The reactive forms of **keto-ABNOH** is the correspondent oxoammonium form, capable of reacting with tryptophan or Trp containing peptides and proteins (**Scheme 21**).

In particular, the oxoammonium form can be generated through NO-promoted in situ oxidation or acid-promoted dispropornation of N-oxyl radicals. About the mechanism of reaction, the indole ring of tryptophan nucleophilically reacts with the oxygen atom of keto-ABNO affording the iminium intermediate (**Scheme 21-A**). Either through addition of water (pathway I) or an nitrogen atom of peptide main chain (pathway II) or deprotonation (pathway III) the conjugate products were generated (**Scheme 21-B**).



Scheme 21. A) Possible mechanism of Trp-selective bioconjugation, where reactive oxonium form is generated by NO_x-promoted in situ oxidation or acid-promoted dispopornation of N-oxyl radical. B) Possible ways to formation of conjugate product either through nitrogen atom of tryptophan (pathway I) deprotonation (pathway II) or the addition of solvent water (pathway III).

As the use of $NaNO_2$ in DNA can be problematic due to possible deamination reactions, we have also studied the possibility of hydroxylamine activation to form N-O radical forms in the presence of O_2 as the only oxidant in different buffer. Thus, Keto-ABNOH (10 mM) was incubated at 60 °C for 18 h in water or AcOH (1 mM) or Tris·HCl buffer (0.02 M, pH 7.5) or $NaHCO_3$ buffer (0.05 M, pH 8.5), in a total volume of 100 uL. Afterwards, each solution was transferred in a capillary and EPR spectra were recorded in order to prove the formation of radical keto-ABNO, whereas for keto-ABNOH in water without any further treatment no radical form has been observed in EPR (**Figure 60**).

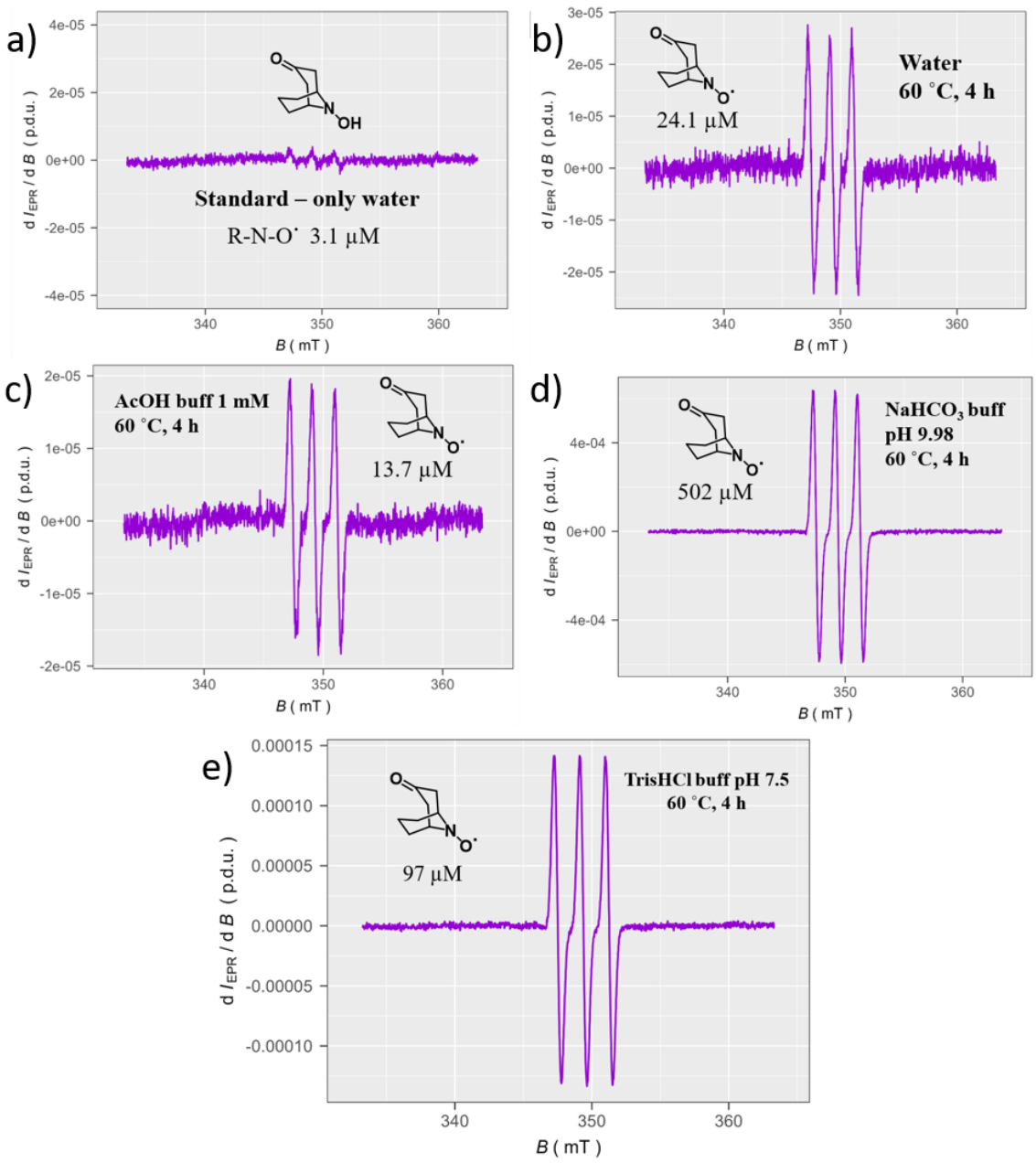
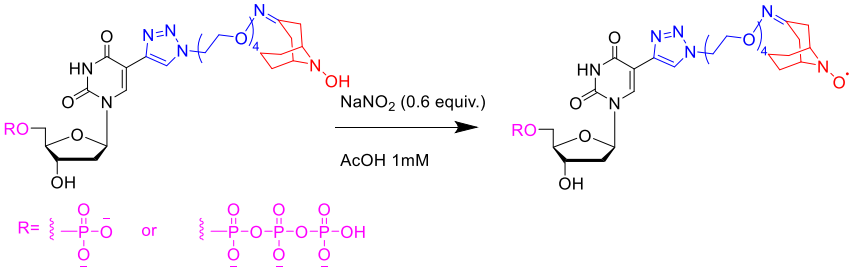


Figure 60. Formation of N-O radical on keto-ABNOH. Conditions: b) water or c) AcOH (1 mM) or d) $NaHCO_3$ pH 9.98 or e) TrisHCl buff pH 7.5 at 60 °C for 18 h. a) Control keto-ABNO in water without any thermal incubation. From the EPR measurements, we can concluded that the only presence of oxygen was enough to allow the formation of reactive radical keto-ABNO.

3.4.1.4 EPR measurements of N-O radical formation on dU^{ABNOH}MP and dU^{ABNOH}TP

The same formation of oxoammonium forms by oxidation, due to presence of O_2, and disproponation to N-oxyl radical can be ascribed for **dU^{ABNOH}MP** or **dU^{ABNOH}TP**.

In order to study the formation of radical in **dU^{ABNOH}MP** or **dU^{ABNOH}TP**, to a solution of aforementioned nucleotides (28.9 µL, 10 mM) in 1 mM AcOH was added $NaNO_2$ (2.89 µL, 6 mM), the solution was transferred in a capillary and EPR spectra were recorded at different times (2 min, 1 h, 2 h, 3.5 h) (**Figure 61**-above). From the kinetic was observed the formation of the radical in case of triphosphate was slower than monophosphate (**Figure 61**-bottom).



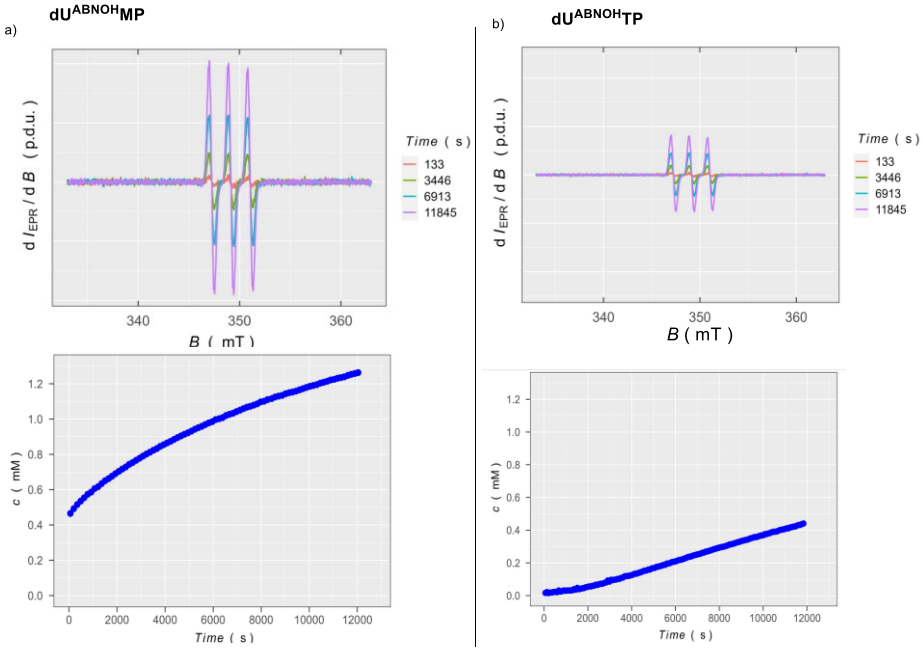
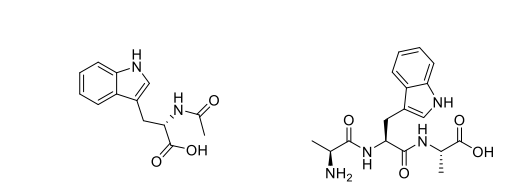


Figure 61. Kinetic formation of radical on **dU^{ABNOH}MP** (a) and on **dU^{ABNOH}TP** (b). Conditions: 1 mM AcOH, 6 mM $NaNO_2$, 37 °C, 2 min - 3.5 h.

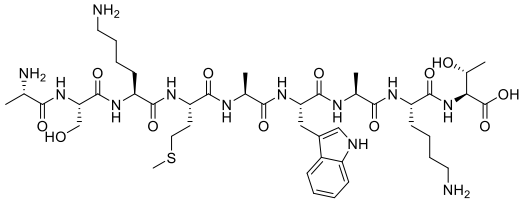
3.4.1.5 Bio-conjugation reactions

3.4.1.5.1 Structure of peptides and amino acids used in keto-ABNO study

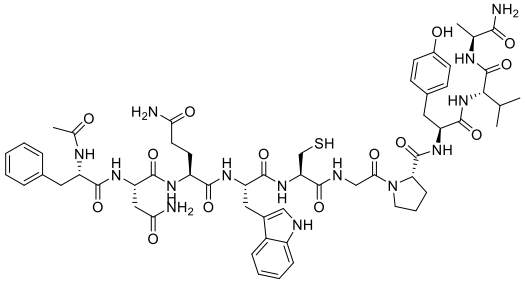


Ac-Trp
AcTrp

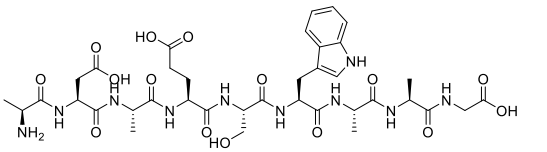
Ac-AlaTrpAla
Pept1



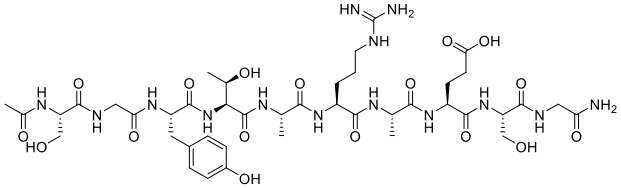
Ala-Ser-Lys-Met-Ala-**Trp**-Ala-Lys-Thr-NH_2
Pept 2



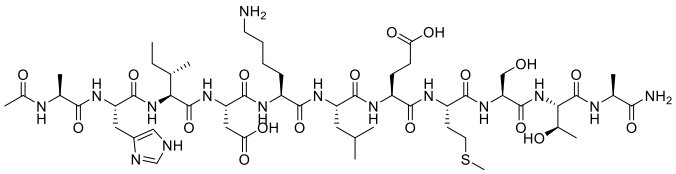
Ac-Phe-Asn-Gln-**Trp**-Cys-Gly-Pro-Tyr-Val-Ala-NH_2
Pept 3



Ala-Asp-Ala-Glu-Ser-**Trp**-Ala-Ala-Gly-NH_2
Pept 4



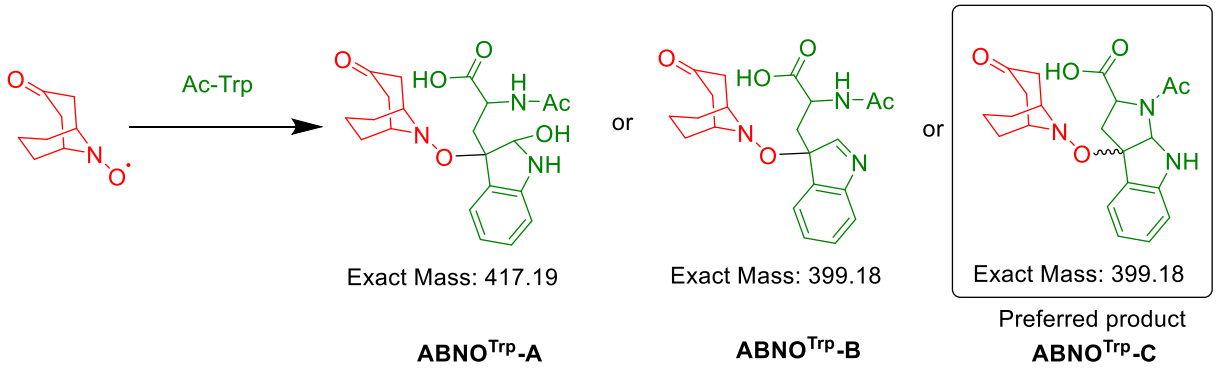
Ac-Ser-Gly-Tyr-Thr-Ala-Arg-Ala-Glu-Ser-Gly-NH_2
Pept 5



Ac-Ala-His-Ile-Asp-Lys-Leu-Glu-Met-Ser-Thr-Ala-NH_2
Pept 6

3.4.1.5.2 Cross-linking model reactions of 9-Azabicyclo[3,3,1]nonan-3-one-9-oxyl (keto-ABNO) moiety and Acetyl-Tryptophan (Ac-Trp) or Acetyl-AlaTrpAla tripeptide (Pept1), testing of reactivity with amino group

In order to understand the connectivity of bioconjugation products a model reaction of **keto-ABNO** moiety with a large excess of tryptophan (sat. solution) was performed in H_2O, at 25 °C for 2.5 h or in $NaHCO_3$ buffer, 37 °C for 3 h both the the presence or in the absence of $NaNO_2$. These conditions are usually used for bioconjugation reactions. Based on literature three different products could be formed[93] (**Scheme 22**).



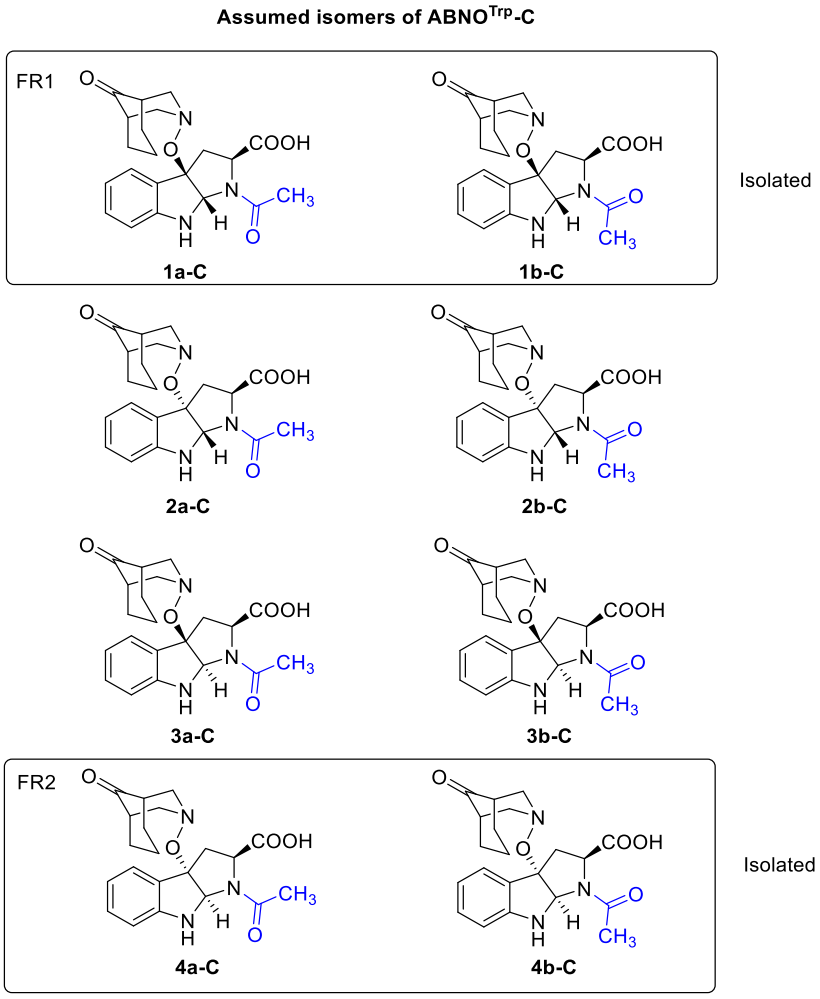
Scheme 22. Synthetic scheme of reaction between **keto-ABNO** radical and **Ac-Trp** giving three proposed products **ABNO^{Trp}-A**, **ABNO^{Trp}-B** and **ABNO^{Trp}-C**.

Despite repeated HPLC purifications, the products were isolated and characterized by NMR and mass spectroscopy as inseparable mixtures of diastereomers of tricyclic conjugates which are more stable than other structures proposed (**Scheme 22**). The reactions in presence of $NaNO_2$ gave higher yield 30% than reaction in the absence of the activator with 17% yields.

3.4.1.5.3 NMR characterization of adducts (keto-ABNO_Trp, keto-ABNO_Pept1)

By NMR and mass spectroscopy was confirmed the formation of **keto-ABNO^{Trp}-C**, whereas **keto-ABNO^{Trp}-A** and **keto-ABNO^{Trp}-B** proposed products were not isolated at all. In order to get a detail insight into products structure, we acquired and analyzed ^{1}H, ^{13}C, H,H-COSY, H,C-HSQC, H,C-HMBC and H,H-NOESY spectra.

Structure **keto-ABNO^{Trp}-A** could be ruled out based on ^{13}C chemical shift of indole C-2 (~ 80 ppm), which suggest sp^3 hybridization of indole C-2. Additionally, H,C-HMBC spectrum revealed splitting of Trp C-2 by indoline H-2 suggesting that pyrroloindoline dehydration product **keto-ABNO^{Trp}-C** is being formed during the reaction and that this compound is stable in aqueous environment. However, structure **keto-ABNO^{Trp}-C** contains two extra chiral centers and therefore formation of diastereomers is expected. Theoretically, due to the chiral centers and hindered rotation of N-Ac bond, eight isomers can be formed. By HPLC were isolated only **1a-C**, **1b-C** isomers in the first fraction and **4a-C** and **4a-C** isomers in the second fraction but no traces of **2a-C**, **2b-C**, **3a-C** and **3b-C** (**Scheme 23**).



Scheme 23. Assumed and isolated isomers of **ABNO^{Trp}-C** in the frame.

Proposed mechanism of the reaction[93] and H,H-NOESY spectra of both fractions (**Figure 62**) confirmed *syn* pyrroloindoline ring closure and presence of two rotamers due to the hindered N-Ac bond rotation.

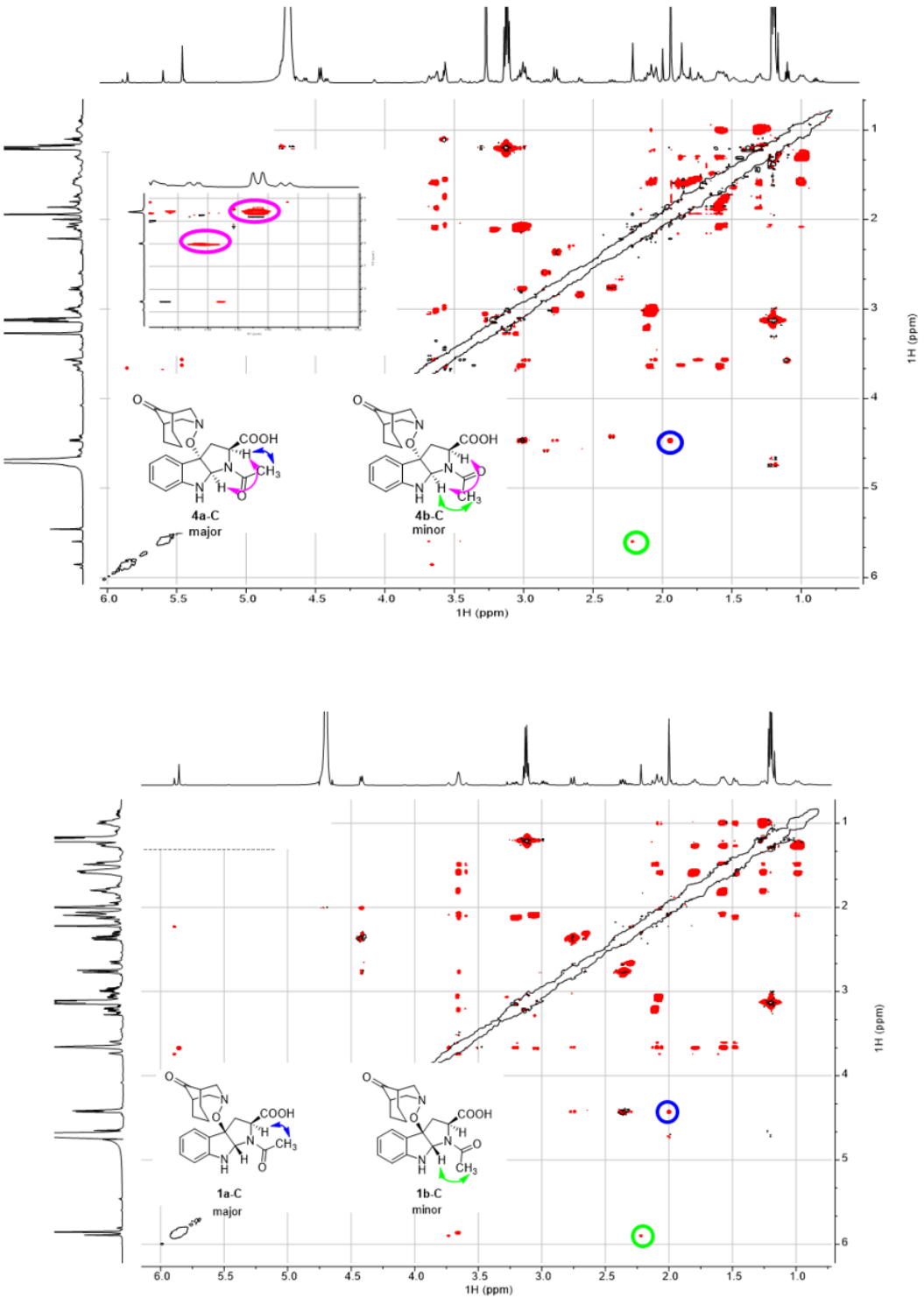


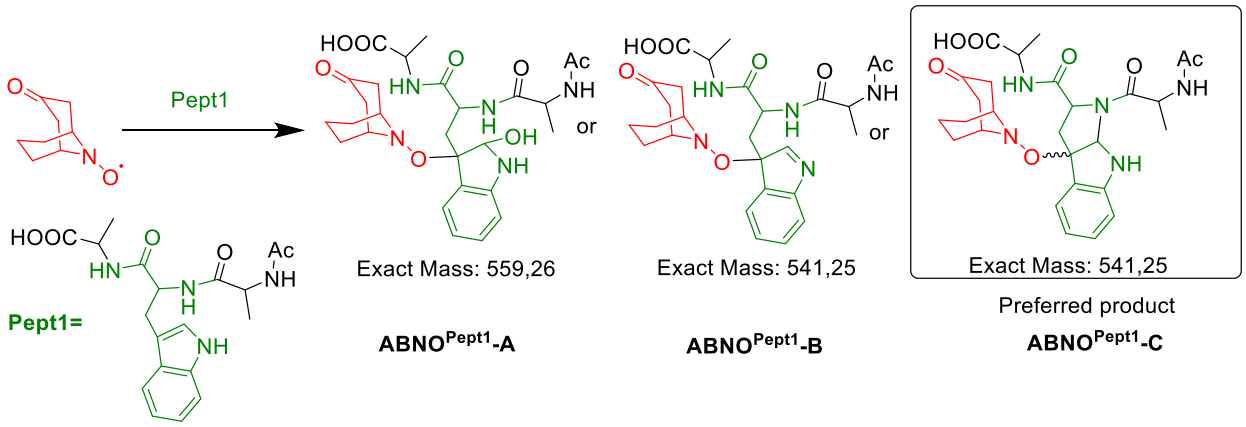
Figure 62. H,H-NOESY spectra of first **FR1** (up) and second **FR2** (down) HPLC fractions isolated after model reaction of keto-ABNO with Ac-Trp.

The isolated fractions **FR1** and **FR2** were analysed by mass and the formation of conjugate **keto-ABNO^{Trp}** was confirmed in both fractions by mass analysis.

HR/MS **FR1** (ESI$^-$) for $C_{20}H_{25}N_3O_5$ 398.18 $[M-H]^-$ calculated, found 398.17184 $[M-H]^-$

HR/MS **FR2** (ESI$^-$) for $C_{20}H_{25}N_3O_5$ 398.18 $[M-H]^-$ calculated, found 398.17193 $[M-H]^-$

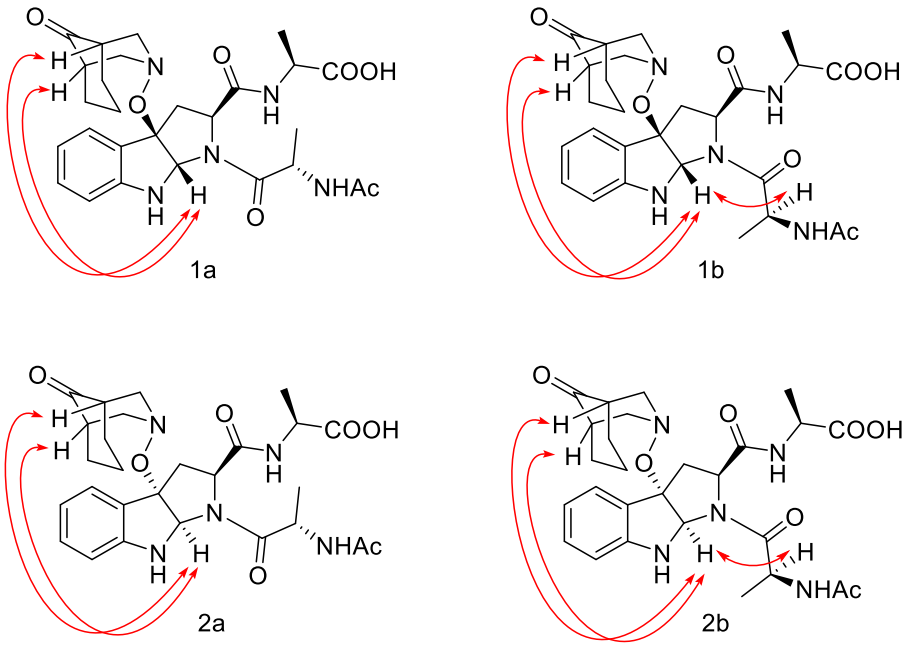
Similarly, the connectivity between **keto-ABNO-H** and **Ac-AlaTrpAla** was studied and was supposed the formation of three products (**Scheme 24**).



Scheme 24. Synthetic scheme of reaction between **keto-ABNO** radical and **Pept1** with proposed products.

The model reaction of keto-ABNO (10 mM) with tripeptide **Ac-AlaTrpAla** (250 mM) in H_2O was performed at 25 °C for 2.5 h or in $NaHCO_3$ buffer, 37 °C for 3 h both in the presence or in the absence of $NaNO_2$. The reaction in presence of $NaNO_2$ gave higher yield (14%) than reaction in the absence of the activator (8%).

The crude mixture was purified by HPLC and the formation of proposed **ABNO^{Pept1}-C** was observed as mixture of four main isomers of pyrroloindoline proposed (**Scheme 25**).



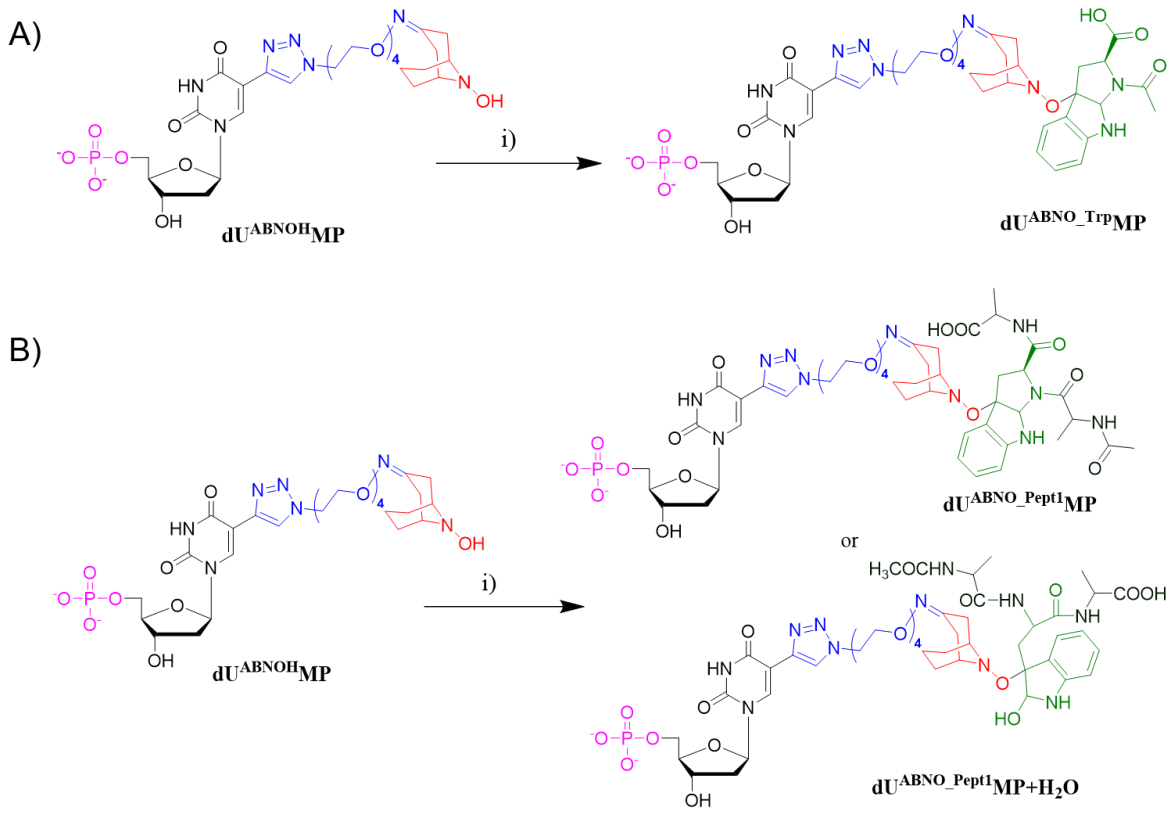
Scheme 25. Assumed products of **keto-ABNO^{Pept1}-C**. Red arrows indicate observed NOE contacts.

The structure of isomers formed were elucidated based on ^{1}H, ^{13}C, H,H-COSY, H,C-HSQC, H,C-HMBC and H,H-ROESY spectra (**Figure 97** in the Appendix, Section 7.2). The pyrroloindoline skeleton was confirmed by H,C-HMBC spectrum, in which splitting of triptophan C-2 by indoline H-2 was observed. NOE crosspeaks between indoline H-2 and protons of keto-ABNO moiety were observed in H,H-ROESY spectrum for all four isomers of **keto-ABNO^{Pept1}-C - 1a, 1b, 2a, 2b** suggesting their *syn* orientation. In addition, the assignment of rotamers **a** (minor) and **b** (major) was supported by H,H-ROESY spectrum. Set of signals of keto-ABNO moiety suggested extra isomerism of this part of molecule e.g. formation of diastereoisomers due to the presence of asymmetric nitrogen.

3.4.1.5.4 Cross-linking model reactions of dU^{ABNOH}MP and Acetyl-Tryptophan (Ac-Trp) or Acetyl-AlaTrpAla tripeptide (Pept1)

Subsequently, we tested the reactivity of model nucleoside monophosphate **dU^{ABNOH}MP** (10 µL, 10 mM aqueous solution) in reactions with either large excess of **Ac-Trp** (10 µL, saturated aqueous solution) or **Pept1** (10 µL, 250 mM aqueous solution). The reaction mixture was completed to a final volume of 50 µL with 50 mM $NaHCO_3$ pH 10 (30 µL) buffer at 37 °C for 18 h in absence of $NaNO_2$ radical activator.

The expected cyclized conjugates **dU^{ABNO_Trp}MP** and **dU^{ABNO_Pept1}MP** were analysed and characterized by UPLC-MS. In case of **Ac-Trp** the product **dU^{ABNO_Trp}MP** was formed with almost full conversion and 17% conversion in case of **Pept1** for **dU^{ABNO_Pept1}MP**. For this latter reaction also the formation of hydrated product **dU^{ABNO_Pept1}MP+H$_2$O** was observed which was proposed above as possible product (**Scheme 26**).



Scheme 26. Schematic synthetic reactions of A) **dU^{ABNO_Trp}MP** conjugate formation and of B) **dU^{ABNO_Pept1}MP** conjugate formation. Condition: **Ac-Trp** or **Pept 1**, $NaHCO_3$ pH 10, 37 °C for 18 h.

3.4.1.5.5 LC-MS characterization of conjugate products $dU^{ABNO_Trp}MP$, $dU^{ABNO_Pept1}MP$

Next, the cyclised **$dU^{ABNO_Trp}MP$** and **$dU^{ABNO_Pept1}MP$** conjugates were analysed and characterized by UPLC-MS (**Figure 63**).

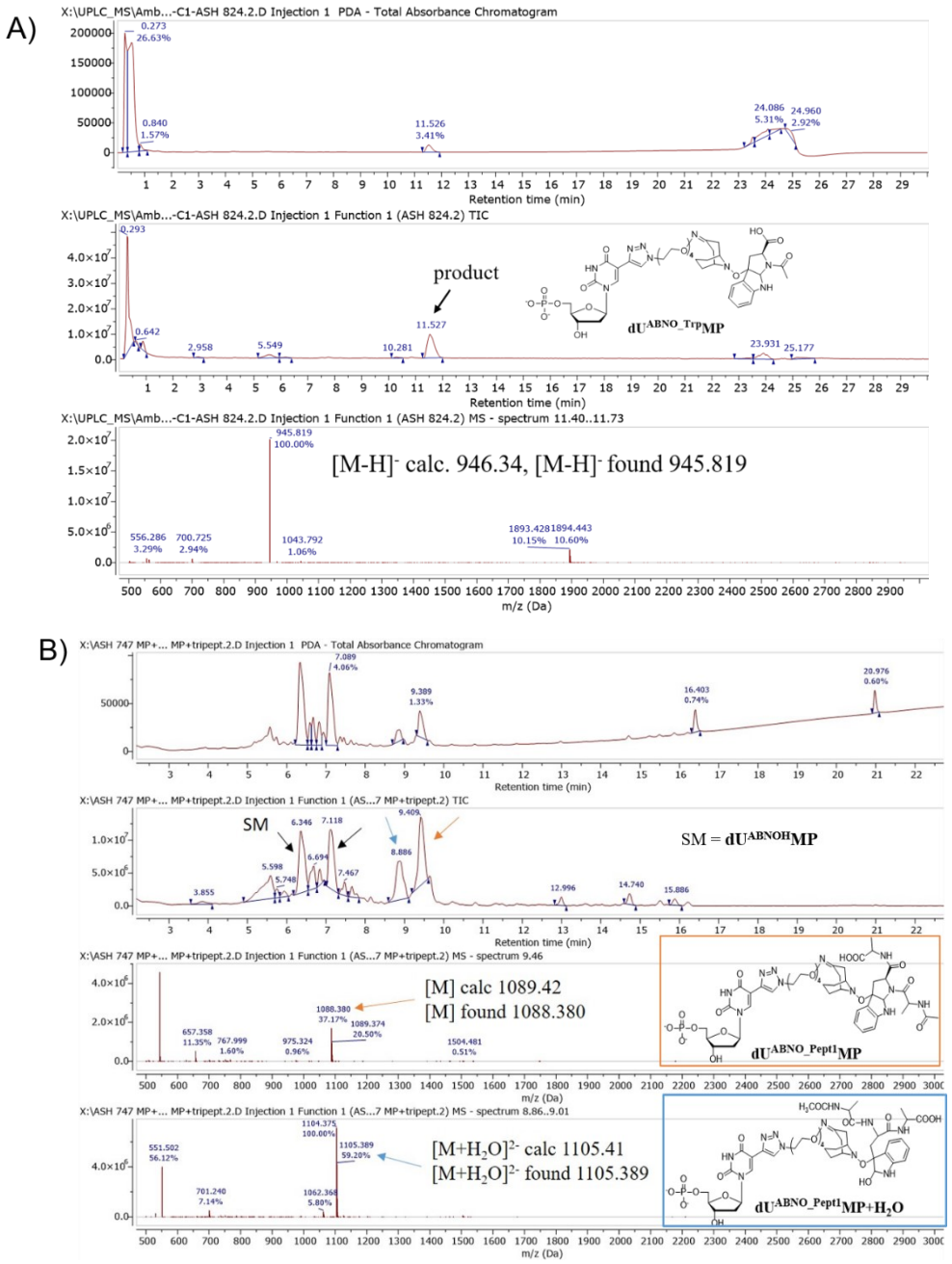
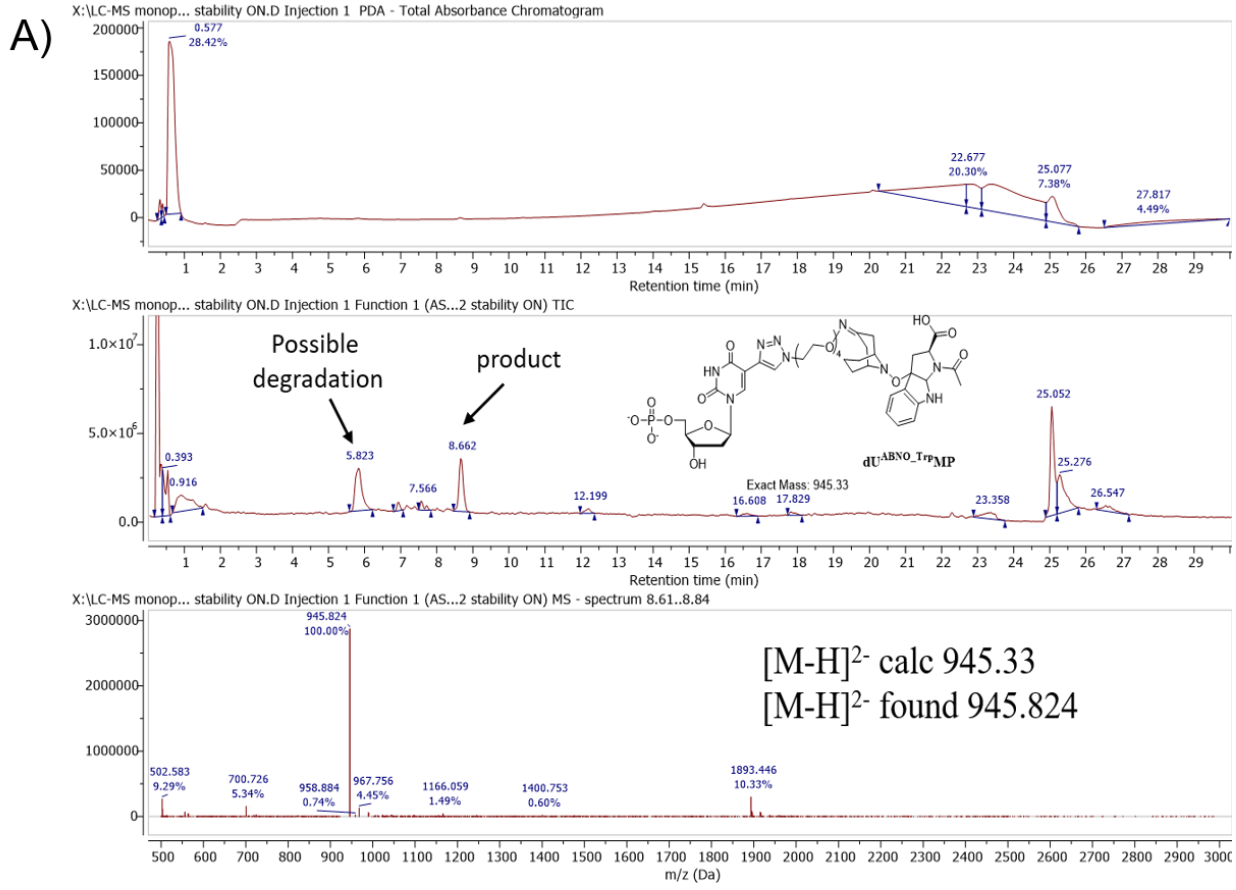


Figure 63. LC-MS chromatograms and spectra of A) cross-linking product between **$dU^{ABNOH}MP$** and **Ac-Trp** (crude mixture) or B) cross-linking product between **$dU^{ABNOH}MP$** and **Pept1** (crude mixture).

3.4.1.5.6 Stability studies of conjugated products $dU^{ABNO_Trp}MP$, $dU^{ABNO_Pept1}MP$

To confirm the stable covalent cross-linking with the indole ring of tryptophan, **$dU^{ABNO_Trp}MP$**, **$dU^{ABNO_Pept1}MP$** conjugates were isolated by HPLC and then their stability was tested in H_2O at 37 °C for 18 h and then analyzed by UPLC-MS. These conditions were used to check either under the same reaction condition the conjugates were stable after incubation overnight. The expected mass of conjugated products were found after incubation, which means they were enough to determine that linkage formed is covalent (**Figure 64**).



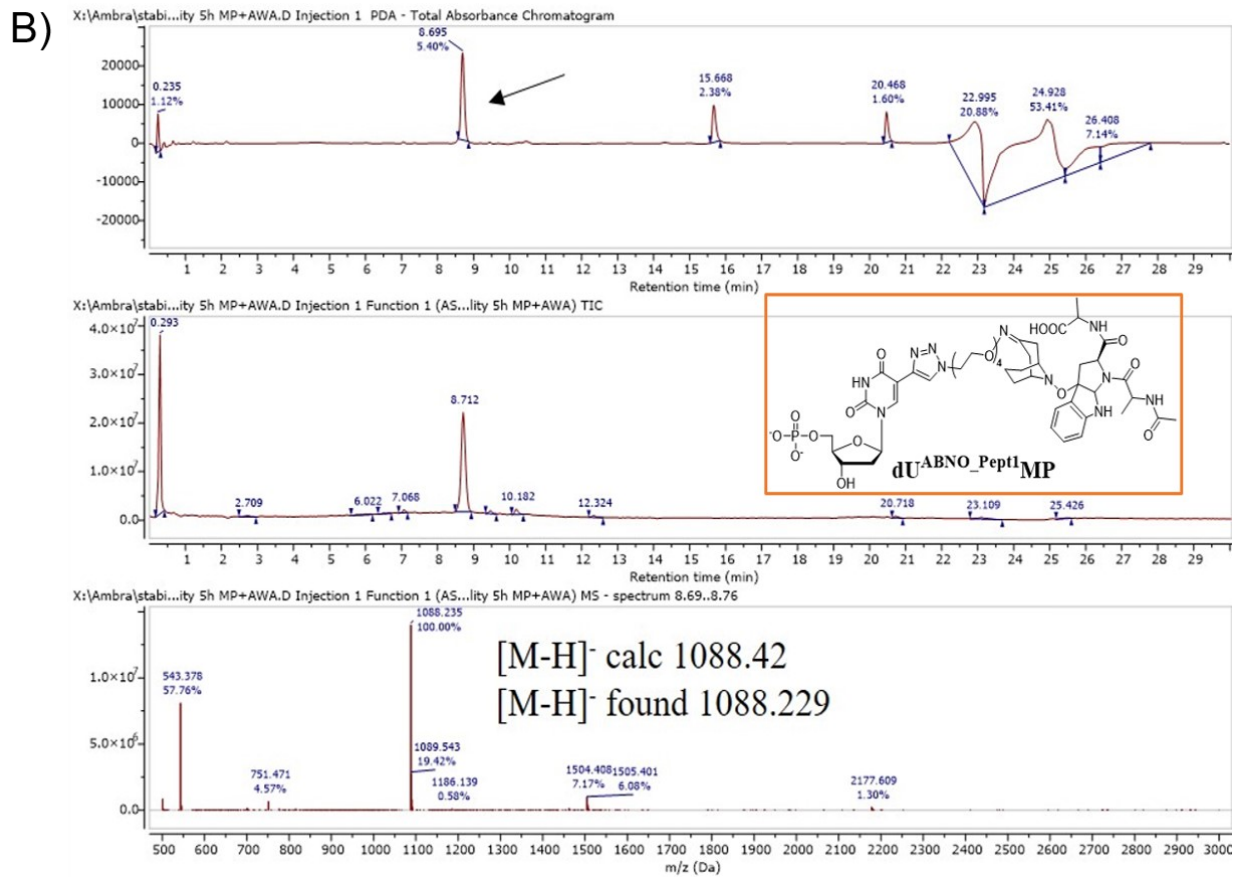


Figure 64. Stability study of A) cross-linking product between **dU^{ABNOH}MP** and **Ac-Trp** or B) cross-linking product between **dU^{ABNOH}MP** and **Pept1** in water (after HPLC isolation of the conjugated products).

The peptide conjugate **dU^{ABNO_Pept1}MP** was also characterized by Fourier Transform Ion Cyclotron Resonance (FTICR) MS/MS analysis which was performed by Marek Polák (**Figure 65**).

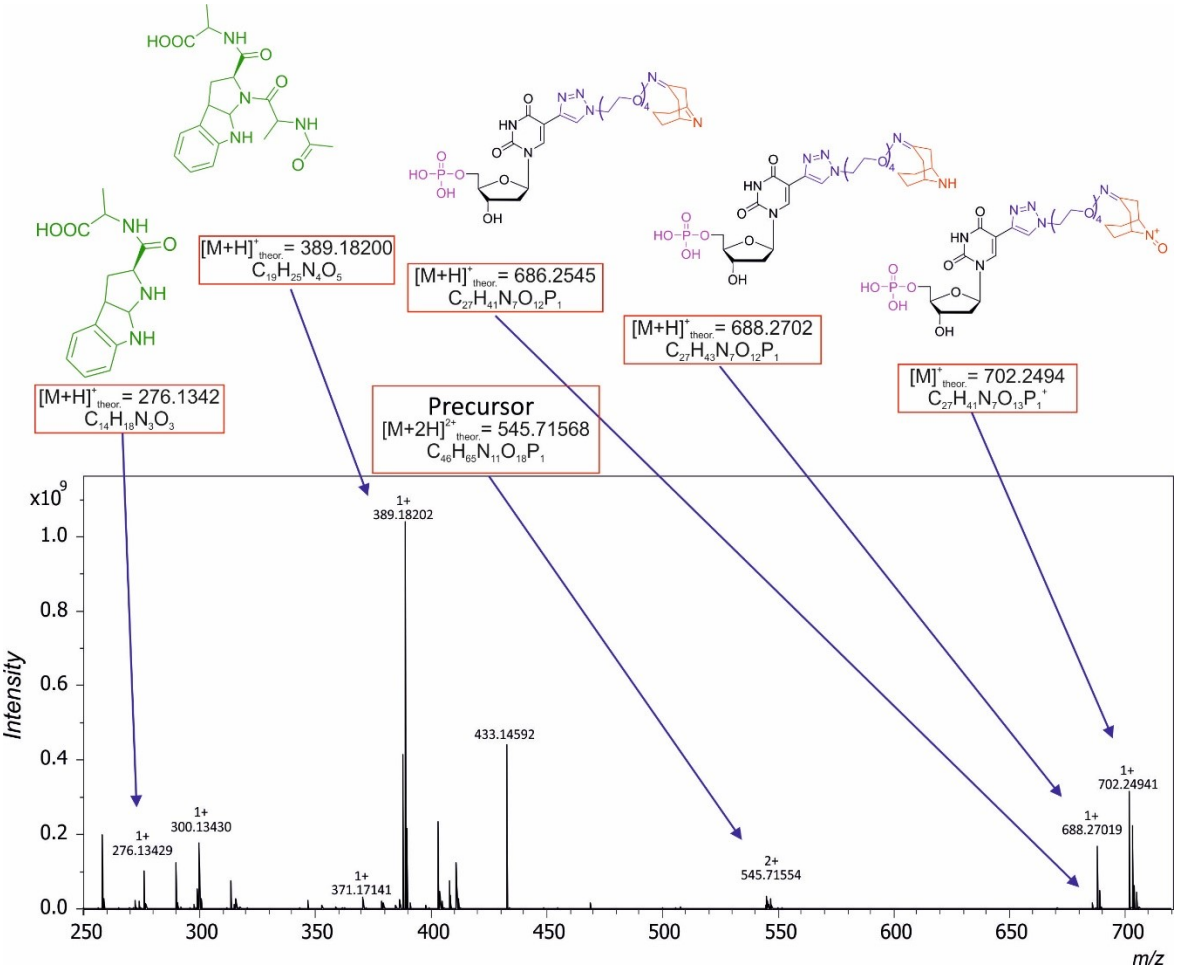
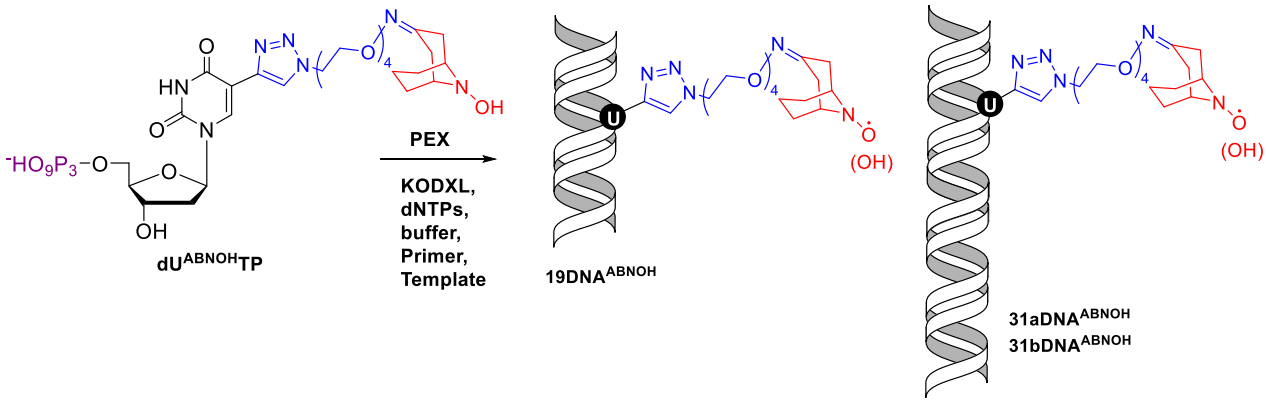


Figure 65. A CID MS/MS spectrum of conjugate **dU^{ABNO_Pept1}MP** fragmentation zoomed in m/z 250-720. Annotated compounds are denoted above the spectrum, displaying molecular formula mass in form of $[M+H]^+$.

3.4.1.6 Incorporation of dU^{ABNOH}TP into DNA by PEX and MALDI analysis

Afterwards, we tested the modified triphosphate **dU^{ABNOH}TP** as substrate for KOD XL DNA polymerase in PEX primer extension (**Scheme 27**). To synthetize **31a/31bDNA^{ABNOH}**, by PEX, 19-mer and 31- mer templates and different primers were used (**Table 12**).



Scheme 27. Enzymatic incorporation of **dU^{ABNOH}TP** via PEX.

Table 12. List of sequences of primers and templates used in keto-ABNOH study.

Name	Sequence (5'→3')	Length
Prim^{A}	5`-CATGGGCGGCATGGG-3`	15 nt
Prim^{A}_FAM [a]	5`-CATGGGCGGCATGGG-3`	15 nt
Prim^{31a}_FAM [a]	5`-ACGTTATCATACTTGTATACAAGTATAC-3`	28 nt
Prim^{31b}_FAM [a]	5`-AGGAGTATACTTGTATACAAGTATGA-3`	26 nt
Temp^{19_1T}	5`-CCCACCCATGCCGCCCATG-3`	19 nt
Temp^{19_1T}_bio [b]	5`-CCCACCCATGCCGCCCATG-3`	19 nt
Templ^{31_4T}	5`-CTAGCATGAGCTCAGTCCCATGCCGCCCATG-3`	31 nt
Templ^{31_4T}_bio	5`-CTAGCATGAGCTCAGTCCCATGCCGCCCATG-3`	31 nt
Temp^{31a}	5`-GG*A*GTATACTTGTATACAAGTATGATAACGT-3`	31 nt
Temp^{31a}_bio [b]	5`-GG*A*GTATACTTGTATACAAGTATGATAACGT-3`	31 nt
Temp^{31b}	5`-CGTT*A*TCATACTTGTATACAAGTATACTCCT-3`	31 nt
Temp^{31b}_bio [b]	5`-CGTT*A*TCATACTTGTATACAAGTATACTCCT-3`	31 nt

[a] 5`-(6-FAM) labelled, [b] 5`-biotinylated.

Respectively, 19-mer long sequence bearing one modification (**19DNA^{ABNOH}**) was chosen for model reactions, 31-mer long sequences bearing *bs*GntR recognition sequence with one modification (**31a/31bDNA^{ABNOH}**) were designed to use them further in DNA-protein cross-links (**Table 13**). Therefore the goal of the project is to test the reactivity of 19/31 mer ABNO- modified DNA towards tryptophan and Trp- containing peptides and proteins.

Table 13. List of oligonucleotides synthesized in keto-ABNOH study

Name	Sequence (3`→5`)	Length
19ON^{Nat}	3`-GGG**T**GGGTACGGCGGGTAC-5`	19 nt
19ON$^{Nat\ a}$	3`-GGG**T**GGGTACGGCGGGTAC-5`	19 nt
19ON^{ABNOH}	3`-GGG**dU***GGGTACGGCGGGTAC-5`	19 nt
19ON$^{ABNOH\ a}$	3`-GGG**dU***GGGTACGGCGGGTAC-5`	19 nt
19DNA$^{ABNOH\ a,b}$	3`-GGG**dU***GGGTACGGCGGGTAC-5`	19 nt
31DNA$^{4T_ABNOH\ a,b}$	3`GA**dU***CG**dU***AC**dU**CGAG**dU***CAGGGTACGGCGGGTAC-5`	31 nt
31ON^{4T_ABNOH}	3`GA**dU***CG**dU***AC**dU***CGAG**dU***CAGGGTACGGCGGGTAC-5`	31 nt
31aDNA$^{Nat\ a}$	3`-CCTCATATGAACATATGTTCATACTATTGCA-5`	31 nt
31bDNA$^{Nat\ a}$	3`-GCAATGTATGAACATATGTTCATATGAGGA-5`	31 nt
31aON^{ABNOH}	3`-CC**dU***CATATGAACATATGTTCATACTATTGCA-5`	31 nt
31bON^{ABNOH}	3`- GCAA**dU***AGTATGAACATATGTTCATATGAGGA-5`	31 nt
31bDNA$^{ABNOH\ a}$	3`-GCAA**dU***AGTATGAACATATGTTCATATGAGGA-5`	31 nt
31bDNA^{ABNOH}	3`-GCAA**dU***AGTATGAACATATGTTCATATGAGGA-5`	31 nt
31aDNA$^{ABNOH\ a}$	3`-CC**dU***CATATGAACATATGTTCATACTATTGCA-5`	31 nt

[a] 5`-(6-FAM)-labelled, [b] 5`-biotinylated, ON^{X}-single stranded DNA, DNA^{X}- double stranded DNA, * = **ABNOH**.

The triphosphate was shown to be a good substrate for the single nucleotide incorporation as well as 4-nucleotides incorporation giving full-length 19- or 31-mer PEX products as demonstrated by PAGE analyses (**Figure 66**).

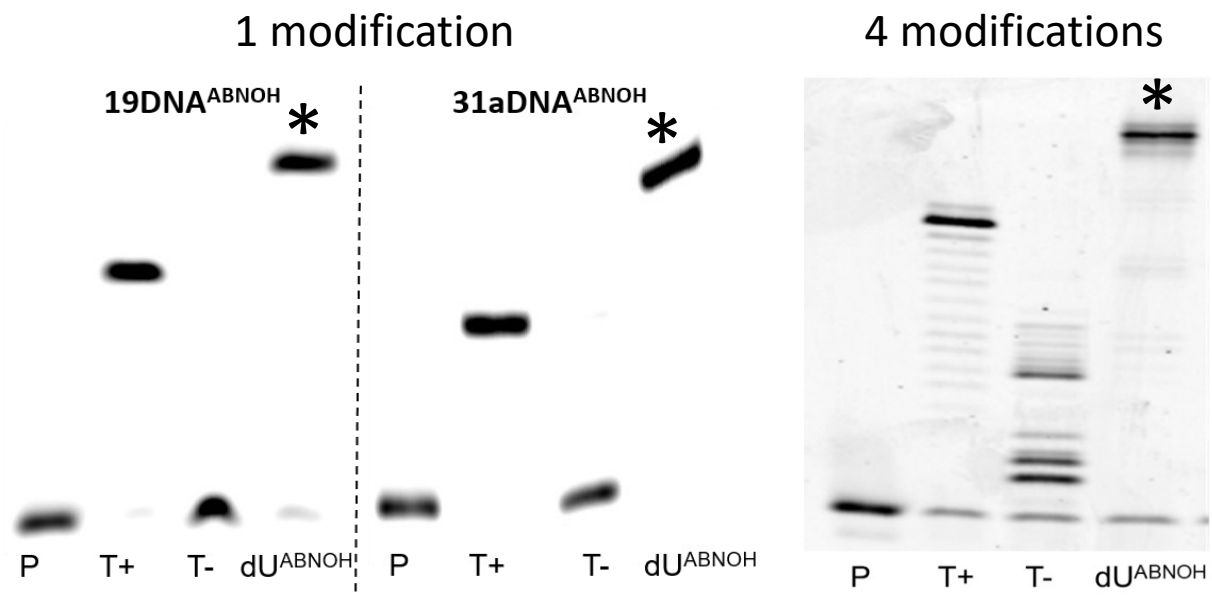


Figure 66. Denaturing SDS-PAGE analysis of PEX using KOD XL DNA polymerase and 19 or 31 mer templates for one or 4 modifications incorporated. *=full length products.

ABNOH- modified DNA were analysed and characterized by UPLC-MS, whereas ABNOH-modified single strand oligonucleotides **19/31ON^{ABNOH}** were characterized by MALDI-TOF. To obtain **19/31ON^{ABNOH}**, the PEX reaction was performed using a 5'-biothinylated template to get the modified DNA **19/31DNA^{ABNOH}** firstly and then the magnetic beads were used for magnetoseparation to separate the biotinylathed template from **19/31ON^{ABNOH}** desired products. The mass spectra confirmed the formation of full extended products and the incorporation of the modification (**Table 14**).

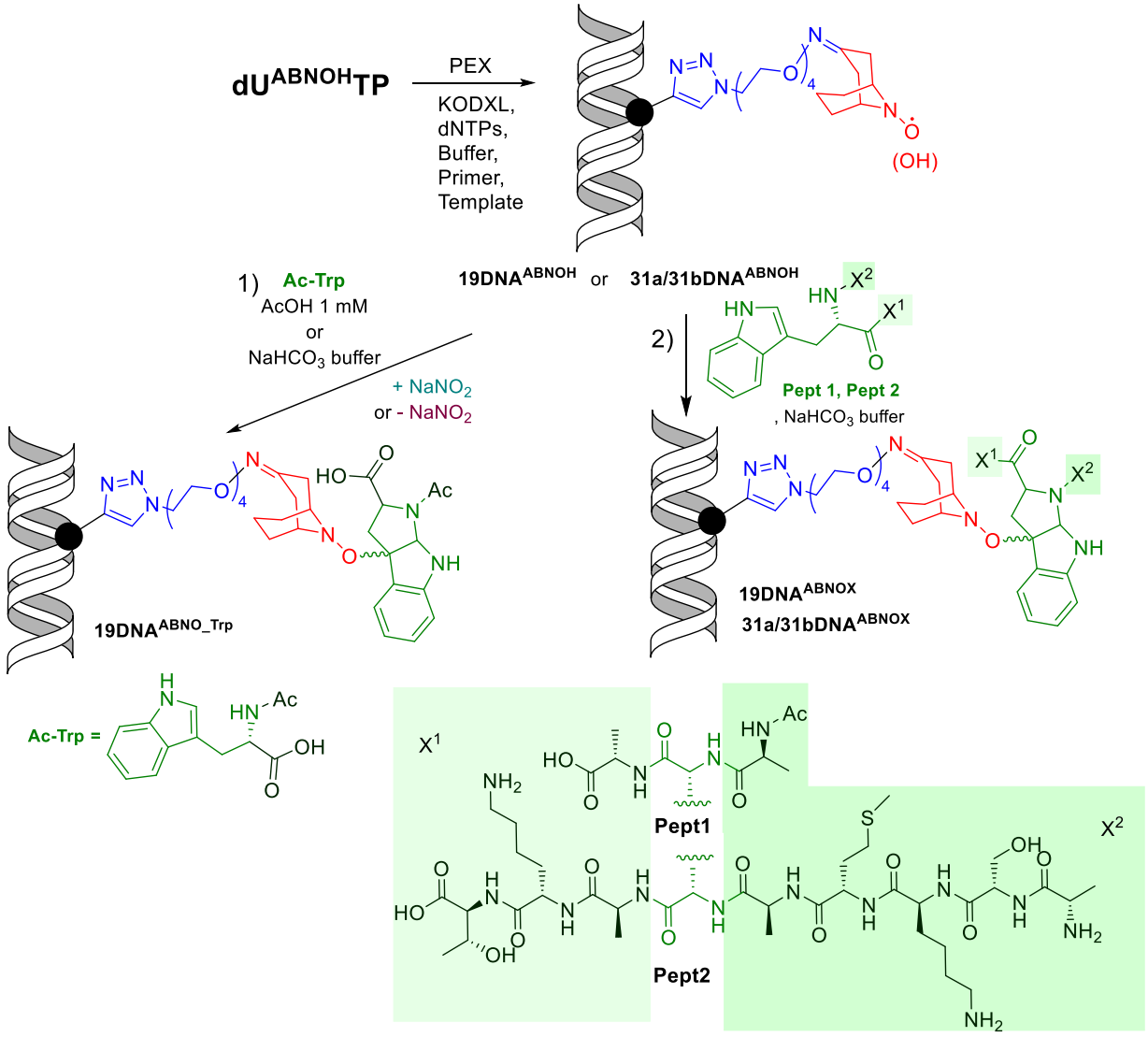
Table 14. MALDI-TOF analysis of modified nucleic acids or oligonucleotides synthetized in this study

Nucleic acids	Mw calculated [Da]	Mw found [Da]
19ON^{ABNOH}	6347.2	6349.4
19ON$^{ABNOH\ a}$	6887.9	6898.1
31ON^{4T_ABNOH}	11145.0	11146.1
31aDNA$^{ABNOH\ a}$	10360.4	10352.0
31aDNA$^{Natural\ a}$	9965.2	9968.2
31bDNA^{ABNOH}	10039.6	10039.0
31bDNA$^{ABNOH\ a}$	10538.6	10526.3
31bDNA$^{Natural\ a}$	10143.4	10145.0
31aON^{ABNOH}	9861.4	9860.0
31aON^{ABNOH}	10047.6	10048.5

[a] 5′-6-FAM-labelled products, ON^{X}-single stranded; DNA^{X}-double-stranded DNA.

3.4.1.7 Model reaction of 19DNA^{ABNOH} with Ac-Trp in presence or absence of $NaNO_2$ at different incubation temperatures

The next step of the project was to test the reactivity of **19/31DNA^{ABNOH}** towards peptides and **Pept 1** and **Pept 2** were chosen for this purpose (**Scheme 28**). For the structures of peptides used see Section 3.4.1.5.2.



Scheme 28. Enzymatic synthesis of **19DNA^{ABNOH}** or **31a/31bDNA^{ABNOH}** and their cross-linking with tryptophan and Trp-containing tripeptide (**Pept1**) and nonapeptide (**Pept 2**). Conditions: AcOH 1mM with or without 13 µM $NaNO_2$, 1 h, 25 °C, or $NaHCO_3$ buffer (50 mM, pH 10), 37 °C, 18 h.

At first, **19DNA^{ABNOH}**, after enzymatic synthesis by PEX, was reacted with a large excess of Ac-Trp (sat. sol.) in AcOH 1 mM in presence of oxidant $NaNO_2$ (13 µM) for 1 h at 25 °C. Additionally, the reaction was conducted in the same conditions in absence of $NaNO_2$ activator to compare the outcome of the cross-linking product. The cross-linking product **19DNA^{ABNO_Trp}** was observed with 51% conversion when $NaNO_2$ was present in the reaction and with 36% without activator as shown by denaturing SDS-PAGE analysis (**Figure 67**).

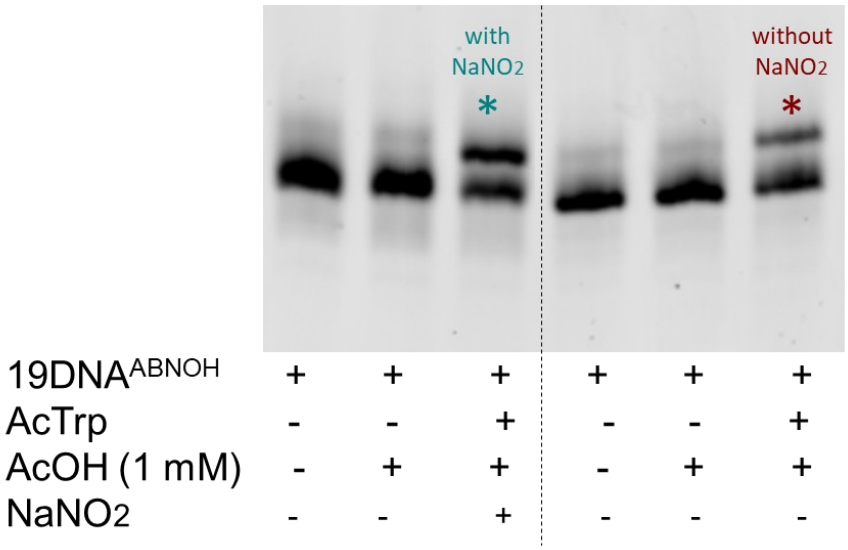


Figure 67. Cross-linking reaction between **19DNA^{ABNOH}** and **Ac-Trp** with and without the use of the activator $NaNO_2$. Conditions: AcOH (1 mM), $NaNO_2$ (13 µM), 1 h, 25 °C. * = cross-linking product.

The cross-linking of **19DNA^{ABNOH}** with **Ac-Trp** was tested at different temperatures (25 °C, 37 °C, 55 °C) in absence of activator ($NaNO_2$) in AcOH pH 5.5 1 mM or in H_2O to have a comparison of conversions. The conversions were increased progressively with higher temperature (31-36% at 25-37 °C and 51-53% at 55 °C) **Figure 68**.

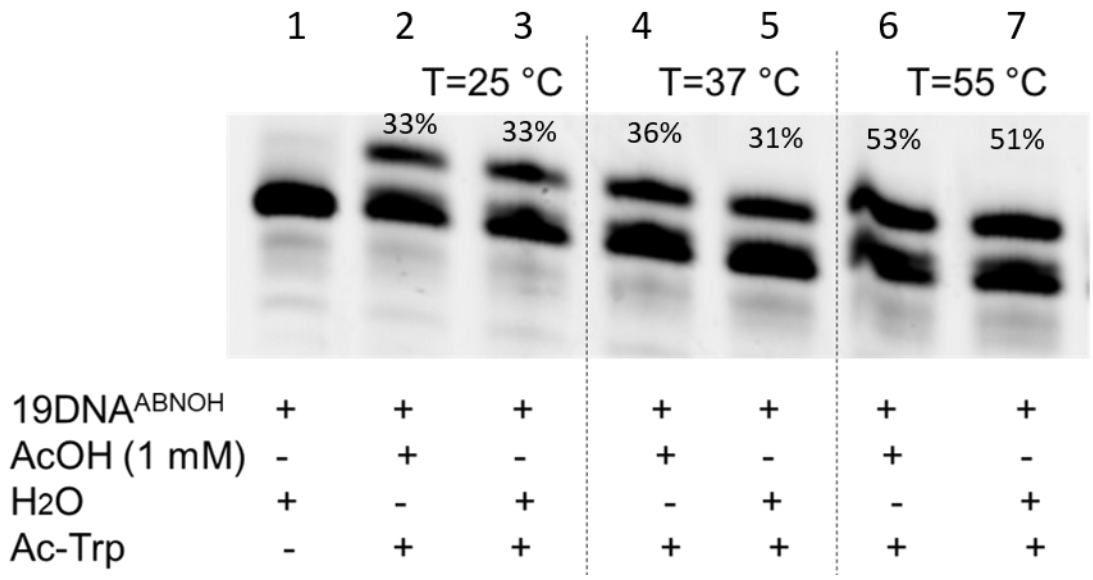


Figure 68. Denaturing PAGE analysis of cross-linking reactions between **19DNA^{ABNOH}** (lane 1-7) with **Ac-Trp** (lane 2-7) in AcOH 1mM (lane 2,4,6) or H_2O (lane 3,5,7) for 1 h at different temperature (25 °C, 37 °C, 55 °C).

Additionally, the cross-linking reaction between **19DNA^{ABNOH}** and **Ac-Trp** without $NaNO_2$ was compared with crescent concentration of $NaNO_2$ (1µM-100 mM). The reaction mixture (10 µL) containing **19DNA^{ABNOH}_FAM** (1 µM) and **Ac-Trp** (250mM) with and without $NaNO_2$ (1 µM-100 mM) in AcOH (1 mM) was incubated for 1 h at 25 °C. The conversion without $NaNO_2$ is comparable when low concentration of the same activator is used with conversion from 32 to 36 % (**Figure 69**, lanes 7-9). In contrast, larger excess of $NaNO_2$ (0.5 – 10 mM) gave the conjugate **19DNA$^{ABNO\text{-}Trp}$** in higher conversions up to 85% (**Figure 69**, lanes 11-13). However, traces of degradation by-products were observed in case of hig excess of $NaNO_2$ used (100 mM) (**Figure 69**, lane 14). Most probably, deamination of DNA was occurred as degradation by-product.

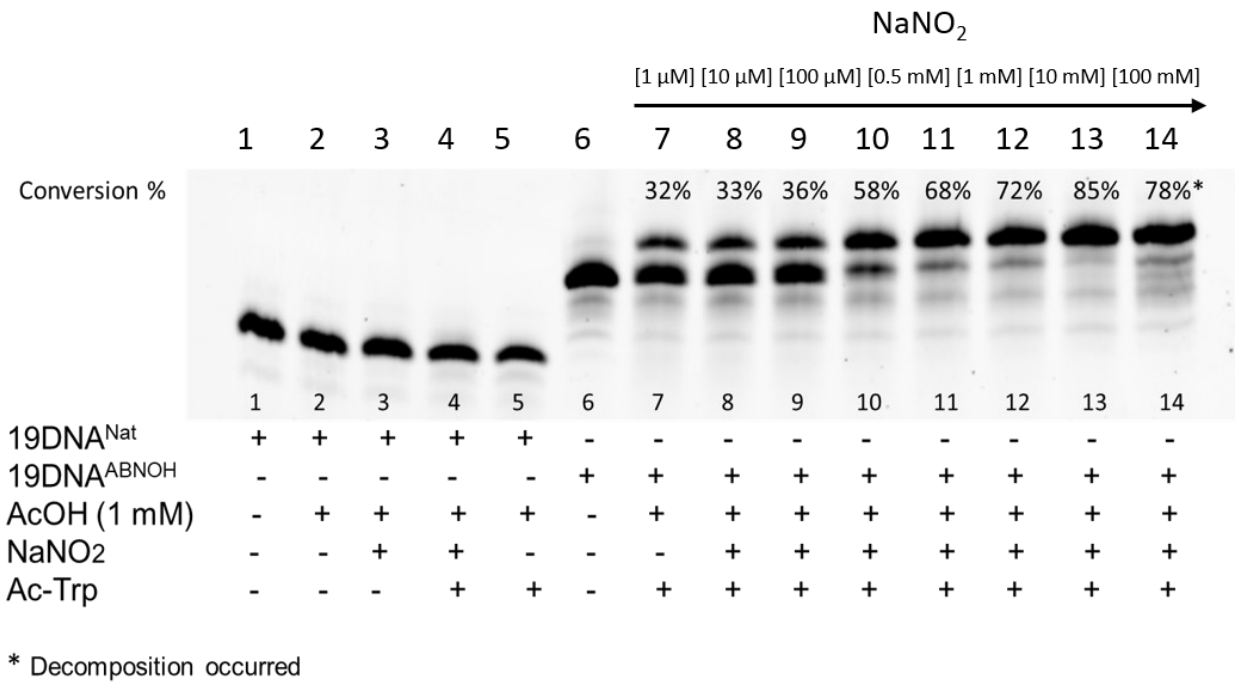


Figure 69. Denaturing SDS-PAGE analysis of cross-linking reactions between **19DNA^{Nat}_FAM** (lane 1-5), with only AcOH 1 mM (lane 2), with $NaNO_2$ (10 mM) in AcOH 1 mM (lane 3), with **Ac-Trp** (250 mM) and $NaNO_2$ (10 mM) in AcOH 1 mM (lane 4) and with **Ac-Trp** (250 mM) in AcOH 1 mM (lane 5) as negative controls. Denaturing SDS-PAGE analysis of cross-linking reactions between **19DNA^{ABNOH}_FAM** (lane 6-14) with **Ac-Trp** (lane 7-14) in AcOH 1 mM (lane 7-14) without $NaNO_2$ (lane 7) and with $NaNO_2$ [1 µM-100 mM] in AcOH 1 mM, all reactions were incubated at 25 °C for 1 h. The conversion in the cross-linked product in the reaction with $NaNO_2$ 1 µM is comparable to the one without the use of $NaNO_2$ (lane 7, conversion 32 % and lane 8, conversion 33%).

3.4.1.7.1 Cross-linking reactions of 19DNA^{ABNOH} with Trp-containing peptides and UPLC-MS characterization of adducts

After synthesis of 19- mer long ABNOH-modified double stranded DNA, **19DNA^{ABNOH}**, by PEX, the modified DNA was used for cross-linking with Trp-containing nonapeptides (**Pept2-Pept4**). Deca- and undecapeptides (**Pept5-Pept6**) were used as negative controls due to lack of Trp. The conjugation reactions were performed in different solutions: in AcOH (1 mM), H_2O for **Pept3**, **Pept5** and **Pept6** (**Figure 70-a**) and in case of **Pept2** and **Pept4** in $NaHCO_3$ (50 mM) pH 10 buffer

for 1 h or 18 h at 37 °C or NH_4Ac pH 5.5 buffer for 1 h at 37 °C (**Figure 70-b**) In this way the selectivity towards tryptophan was proved as shown from SDS-PAGE analysis according to the formation of cross-linked product only in case of reaction with tryptophan containing peptides (**Pept2**, **Pept3**, **Pept4**), as shown in **Figure 70-a**, lanes 2,6 **and Figure 70-b**, lanes 3,4,7,8,11,12.

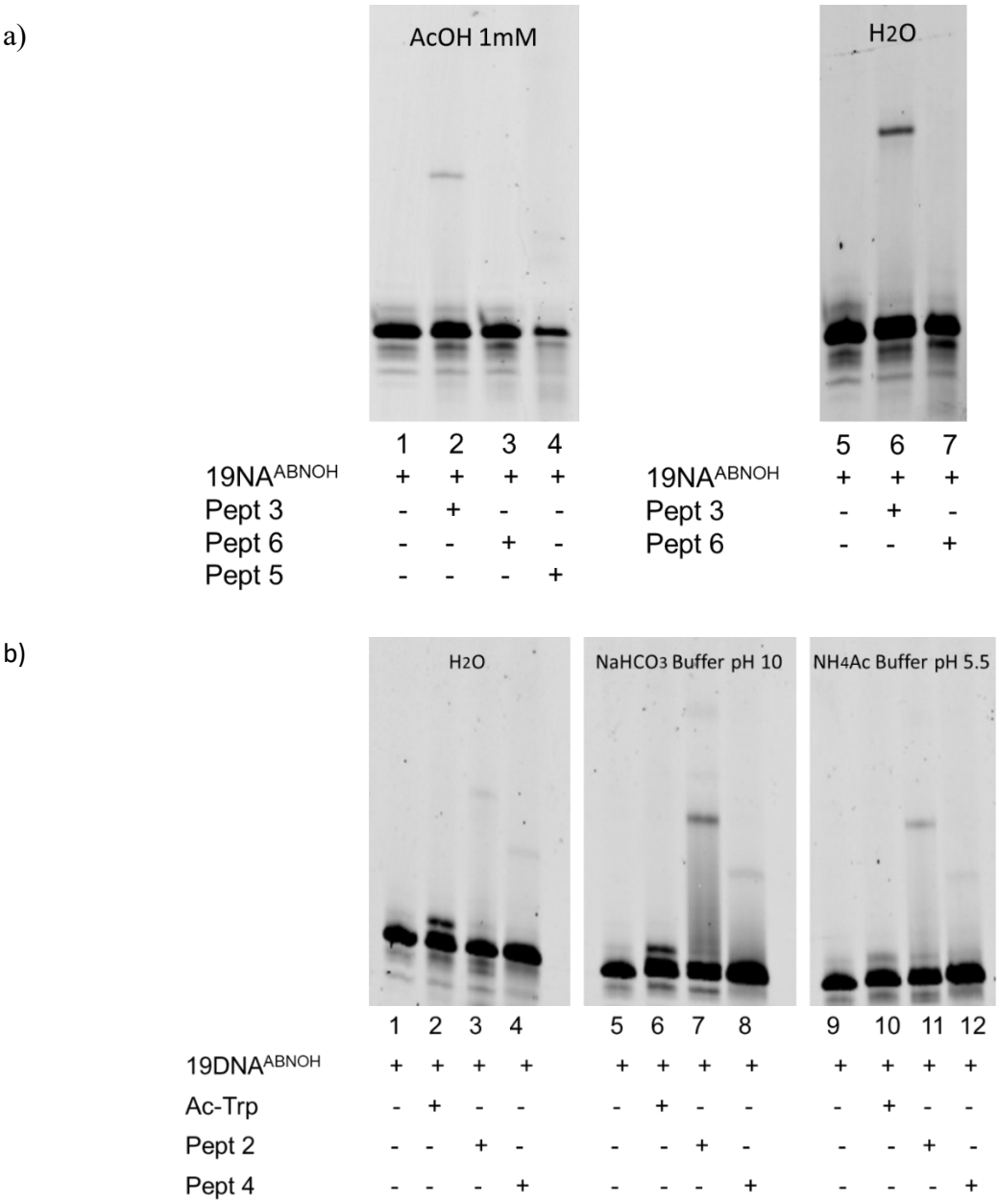


Figure 70. a) Denaturing PAGE analysis of cross-linking reactions between **19DNA^{ABNOH}** (lane 1- 7), with respectively different peptides (**Pept 3**,**5**,**6**) in AcOH 1mM (lanes 1-4) in H_2O (lanes 5-7); b) Denaturing PAGE analysis of cross-linking reactions between **19DNA^{ABNOH}** (lane 1-12),

with respectively different peptides (**Pept 2**,**4**) or **Ac-Trp** in H_2O (lanes 1-4), $NaHCO_3$ buffer pH 10 (lanes 5-8) or NH_4Ac buffer pH 5.5 (lanes 9-12) for 1 h at 37 °C.

Reaction of 19- mer long ABNOH-modified double stranded DNA, **19DNA^{ABNOH}** with **Ac-Trp** or **Pept1** or **Pept2** was tried in presence or absence of $NaNO_2$ showing better conversion in the first case (**Figure 71**, lanes 3,5,7) with 68%, 30% and 15% respectively (**Table 17**).

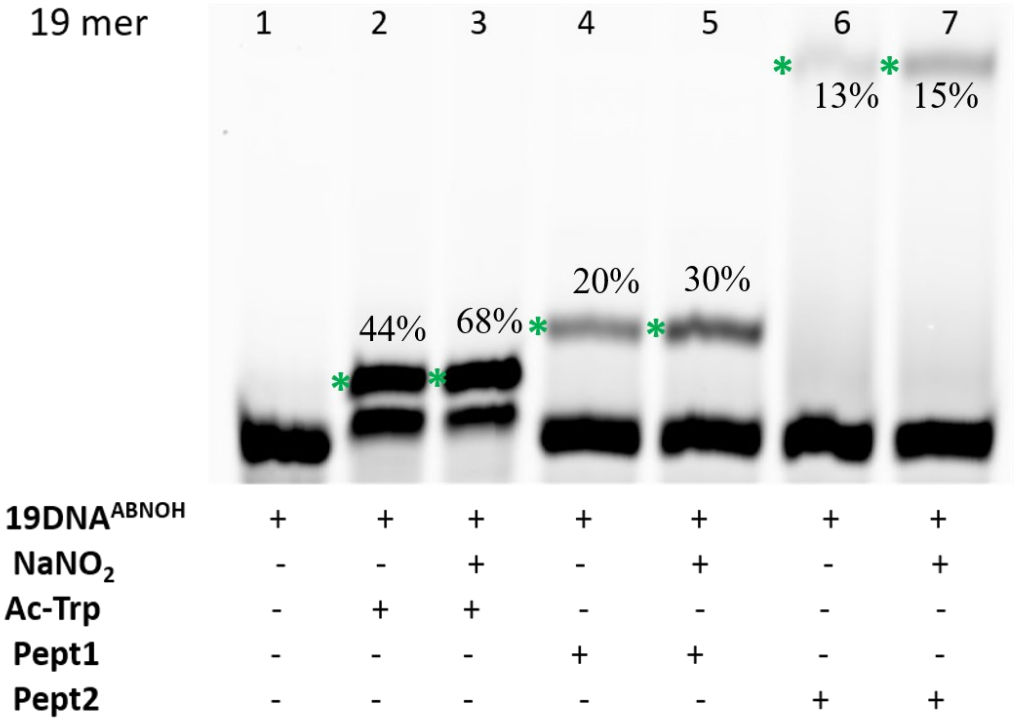


Figure 71. Denaturing 20% PAGE analysis. Conjugation of **19DNA^{ABNOH}** or (5 μM, lanes 1-7) with **Ac-Trp** (saturated aqueous solution, lane 2 without $NaNO_2$ or lane 3 with $NaNO_2$) or **Pept1** (250 mM, lane 4 without $NaNO_2$ or lane 5 with $NaNO_2$) or **Pept2** (250 mM, lane 6 without $NaNO_2$ or lane 7 with $NaNO_2$) in $NaHCO_3$ buffer (0.05 M, pH 10) at 37 °C for 18 h. Controls **19DNA^{ABNOH}** (lanes 1). *=cross-linked product

The outcome of this `trial buffers` reveals the conversions were higher when used H_2O or $NaHCO_3$ pH 10 buffer. The corresponding conjugates **19DNA$^{ABNO\text{-}Trp}$**, **19DNA$^{ABNO\text{-}Pept1}$**, **19DNA$^{ABNO\text{-}Pept2}$** or **19DNA$^{ABNO\text{-}Pept3}$** were formed in moderate conversions 10-85% for **19DNA$^{ABNO\text{-}Trp}$** and 2-44% for **19DNA$^{ABNO\text{-}Pept2}$** and **19DNA$^{ABNO\text{-}Pept3}$** (for all conversion see **Table 17**). The conjugates **19DNA$^{ABNO\text{-}Trp}$**, **19DNA$^{ABNO\text{-}Pept1}$**, **19DNA$^{ABNO\text{-}Pept2}$** were characterized by LC-MS or MALDI-TOF (**Table 15**).

Table 15. MALDI-TOF or ESI data of Trp-modified oligonucleotides or DNA adducts

ON	Mw(calc.)/[Da]	Mw(found) /[Da]
19ON$^{ABNO_Trp\ b}$	6591.9	6593.4
19DNA$^{ABNO_Pept1\ a}$	7283.9	7270.0
19DNA$^{ABNO\ _Pept2\ a}$	7888.1	7890.0
19DNA$^{ABNO\ _Pept2\ b}$	7340.4 7358.0 (+H_2O)	7353.2
19DNA$^{ABNO\ _Pept4\ a}$	7771.8	7770.0

[a] 5`-(6-FAM)-labelled, ON^X-single stranded; DNAx-double-stranded DNA. [b] Nucleic acids were measured in positive mode.

3.4.1.8 Cross-linking reaction of and 31a/31bDNA^{ABNOH} with Ac-Trp and Trp-containing peptides comparison and UPLC-MS characterization of adducts

Also, longer 31- mer ABNOH-modified double stranded DNAs, 5`- FAM labelled **31aDNA^{ABNOH}** and **31bDNA^{ABNOH}**, synthetized by PEX, were used for cross-linking with tryptophan and Trp-containing peptides and **Pept1** or **Pept2** were used as first (**Scheme 28**, above). In order to have a comparison of conjugate conversion when 19- or 31- mer ABNOH-modified DNA double stranded is present, the reactions were conducted in the same conditions as for nucleotides, in $NaHCO_3$ pH 10 buffer at 37 °C for 18 h. This latter buffer and time of 18 h were used for 31 mer-**31aDNA^{ABNOH}** and **31bDNA^{ABNOH}** with **Pept1** or **Pept 2** and further with GntR protein to have all reaction conditions unified.

Moreover, also in case of DNA the formation of conjugates was tested in presence or absence of $NaNO_2$ and analysed by SDS-PAGE gel analysis. The formation of cross-linked product is indeed increased in presence of $NaNO_2$ (**Figure 72**).

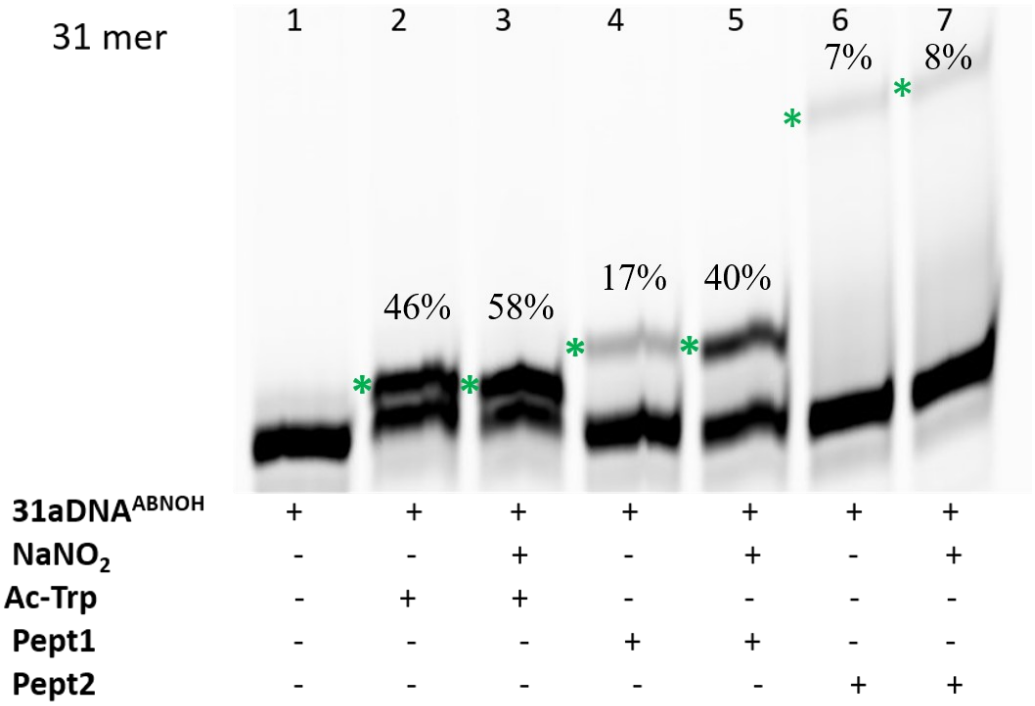


Figure 72. Denaturing 20% PAGE analysis. Conjugation of **31aDNA^{ABNOH}** (5 μM, lanes 1-7) with **Ac-Trp** (saturated aqueous solution, lane 2 without $NaNO_2$ or lane 3 with $NaNO_2$) or **Pept1** (250 mM, lane 4 without $NaNO_2$ or lane 5 with $NaNO_2$) or **Pept2** (250 mM, lane 6 without $NaNO_2$ or lane 7 with $NaNO_2$) in $NaHCO_3$ buffer (0.05 M, pH 10) at 37 °C for 18 h. Controls **31aDNA^{ABNOH}** (lanes 1). *=cross/linked product

As previously mentioned, 5`-FAM labelled **31aDNA^{ABNOH}** and **31bDNA^{ABNOH}** , synthetized by PEX, were used, in the same way, to test the reactivity towards Ac-Trp and Trp- containing peptides (**Pept2** and **Pept4**).

The conjugation reactions were proceeded in $NaHCO_3$ pH 10 buffer at 37 °C for 18 h and all adducts were visualized by SDS-PAGE gel means (**Figure 73**).

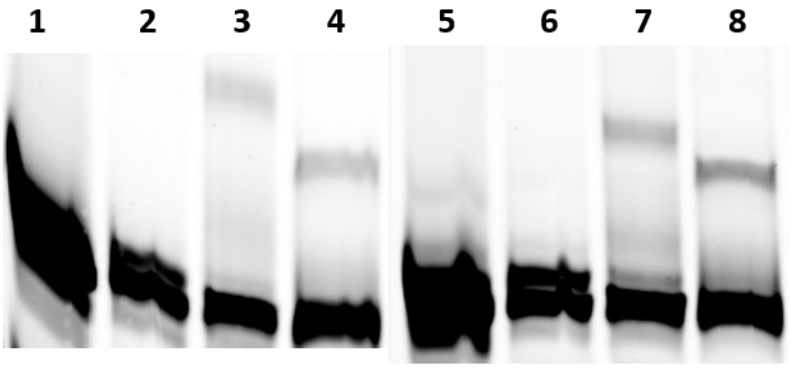


Figure 73. Denaturing 20% PAGE analysis. Conjugation of **31aDNA^{ABNOH}** (7 μM, lanes 1-4) with **Ac-Trp** (saturated aqueous solution, lane 2) or **Pept2** (250 mM, lane 3) or **Pept4** (250 mM, lane 4) in $NaHCO_3$ buffer (0.05 M, pH 10) at 37 °C for 18 h. Control **31aDNA^{ABNOH}** (lane 1); Conjugation of **31bDNA^{ABNOH}** (7 μM, lanes 5-8) with with **Ac-Trp** (saturated aqueous solution, lane 6) or **Pept2** (250 mM, lane 7) or **Pept4** (250 mM, lane 8) in $NaHCO_3$ buffer (0.05 M, pH 10) at 37 °C for 18 h. Control **31bDNA^{ABNOH}** (lane 5).

The corresponding conjugates **31a/31bDNA$^{ABNO\text{-}Trp}$**, **31a/31bDNA$^{ABNO\text{-}Pept1}$**, **31a/31bDNA$^{ABNO\text{-}Pept2}$** or **31a/31bDNA$^{ABNO\text{-}Pept4}$** were formed in moderate conversions 7-44% for **31aDNA^{ABNOH}** and 6-43% for **31bDNA^{ABNOH}** using $NaHCO_3$ pH 10 buffer. All conjugate conversions are shown in **Table 17** and characterized by LC-MS or MALDI-TOF analyses (**Table 16**).

Table 16. MALDI or LC-MS data of products of bioconjugation reactions

ON	Mw(calc.)/[Da]	Mw(found) /[Da]
31aDNA^{ABNO_Pept1}	10209.5	10200.0
31bDNA^{ABNO_Pept1}	10387.4	10370.0
31aDNA^{ABNO_Pept1} [a]	10746.5	10749.0
31bDNA^{ABNO_Pept1} [a]	10924.7	10911.5
31aDNA^{ABNO_Pept2} [a]	11350.2	11350.0
31bDNA^{ABNO_Pept2} [a]	11528.9	11530.0
31aDNA^{ABNO_Pept4} [a]	11235.9	11235.0
31bDNA^{ABNO_Pept4} [a]	11414.9	11415.0

[a] 5`-(6-FAM)-labelled; DNAx-double-stranded DNA.

Table 17. Conversion calculation table of cross-linking reactions of 19/31DNA^{ABNOH} with Ac-Trp or Trp- containing peptides

Peptide	Oligonucleotide	Buffer	$NaNO_2$	Temperature	Time	Conversion %
Ac-Trp	**19DNA^{ABNOH}** *a*	AcOH 1mM	NO	25 °C, 37 °C, 55 °C	1 h	33%, 36%, 53 %
	19DNA^{ABNOH} *a*	H_2O	NO	25 °C, 37 °C, 55 °C	1 h	33%, 31%, 51 %
	19DNA^{ABNOH} *a*	AcOH 1mM	YES	25 °C	1h	32-85%
	19DNA^{ABNOH} *a*	$NaHCO_3$ pH 10	NO	37 °C	1 h	10%
	19DNA^{ABNOH} *a*	$NaHCO_3$ pH 10	NO	37 °C	18 h	44%
	19DNA^{ABNOH} *a*	$NaHCO_3$ pH 10	YES	37 °C	18 h	68%
	31aDNA^{ABNOH} *a*	$NaHCO_3$ pH 10	NO	37 °C	3 h	37%
	31aDNA^{ABNOH} *a*	$NaHCO_3$ pH 10	NO	37 °C	18 h	46%
	31aDNA^{ABNOH} *a*	$NaHCO_3$ pH 10	YES	37 °C	18 h	58%
	31bDNA^{ABNOH} *a*	$NaHCO_3$ pH 10	NO	37 °C	3 h	43%
Pept1	**19DNA^{ABNOH}** *a*	$NaHCO_3$ pH 10	NO	37 °C	18 h	20%
	19DNA^{ABNOH} *a*	$NaHCO_3$ pH 10	YES	37 °C	18 h	30%
	31aDNA^{ABNOH} *a*	$NaHCO_3$ pH 10	NO	37 °C	3h	16%

	31aDNA^{ABNOH} [a]	$NaHCO_3$ pH 10	NO	37 °C	18 h	44%
	31aDNA^{ABNOH} [a]	$NaHCO_3$ pH 10	YES	37 °C	18 h	40%
	31bDNA^{ABNOH} [a]	$NaHCO_3$ pH 10	NO	37 °C	3 h	16%
					18 h	32%
Pept 2	**19DNA^{ABNOH}** [a]	$NaHCO_3$ pH 10	NO	37 °C	1 h	8-10%
	19DNA^{ABNOH} [a]	Phosphate pH 7.4	NO	37 °C	1 h	3-6%
	19DNA^{ABNOH} [a]	Tris pH 7.4	NO	37 °C	1 h	2-3%
	19DNA^{ABNOH} [a]	$NaHCO_3$ pH 10	NO	37 °C	18 h	13%
	19DNA^{ABNOH} [a]	$NaHCO_3$ pH 10	YES	37 °C	18 h	15%
	31aDNA^{ABNOH} [a]	$NaHCO_3$ pH 10	NO	37 °C	3 h	7%
	31aDNA^{ABNOH} [a]	$NaHCO_3$ pH 10	NO	37 °C	18 h	7%
	31aDNA^{ABNOH} [a]	$NaHCO_3$ pH 10	YES	37 °C	18 h	8%
	31bDNA^{ABNOH} [a]	$NaHCO_3$ pH 10	NO	37 °C	3 h	6%
Pept 3	**19DNA^{ABNOH}** [a]	AcOH 1mM	NO	37 °C	1 h	2%
Pept 4	**31aDNA^{ABNOH}** [a]	$NaHCO_3$ pH 10	NO	37 °C	3 h	11%
	31bDNA^{ABNOH} [a]	$NaHCO_3$ pH 10	NO	37 °C	3 h	12%

[a] 5`-(6-FAM)-labelled

By comparison of using 19- mer **19DNA^{ABNOH}** or 31 mer- **31a/31bDNA^{ABNOH}** the conversion of reactions with **Pept1** for 18 h is slightly better when **31a/31bDNA^{ABNOH}** are used, whereas with **Pept2** when used **19DNA^{ABNOH}**. As shown by PAGE gel analyses, the conversion (%) decreases as long as the aminoacidic sequence of peptide is longer which means **Pept3$_{\%}$<Pept2$_{\%}$=Pept4$_{\%}$<Pept1$_{\%}$**.

3.4.1.8.1 Stability studies of conjugates 31a/31bDNA^{ABNO_Pept1}

The stability studies of conjugates **31a/31bDNA^{ABNO_Pept1}** were performed to confirm the formation of covalent linkage between **31a/31bDNA^{ABNOH}** and **Pept1**. Firstly, the conjugates **31a/31bDNA^{ABNO_Pept1}** were obtained under the same reaction conditions, as previously (Section 3.4.1.8), in $NaHCO_3$ pH 10 buffer at 37 °C for 3 h or 18 h and observed by denaturing SDS-PAGE gel analysis (upper gel bands - **Figure 74**). Instead, the bottom gel bands corresponds to the starting

materials **31a/31bDNA^{ABNO}** unreacted. In order to isolate the conjugates **31a/31bDNA^{ABNO_Pept1}**, the upper gel bands corresponding to them were cut from gel and purified by double gel extraction using Amikon-3K column and analysed by SDS-PAGE (**Figure 75**). Then, the pieces of gels were transferred into the Amikon-3K column, having a filter on both sides, and TEAA (TriEthylAmmoniumAcetate) buffer solution was added. Afterwards, an electric field was applied on the column allowing the conjugates to move from gel to the solution buffer. As last step, **31a/31bDNA^{ABNO_Pept1}** in solution were lyophilized and then ridissolved in 10 µL volume to have a concentration of 5 µM. Therefore, after recovering of pure **31a/31bDNA^{ABNO_Pept1}** products, they were incubated in water at 37 °C overnight (18 h) which were analysed by SDS- PAGE gel and UPLC-MS analyses. By that, the stability of the conjugates was confirmed from the same electrophoretic mobility shift (**Figure 76**) and the same mass value observed before the incubation at 37 °C overnight.

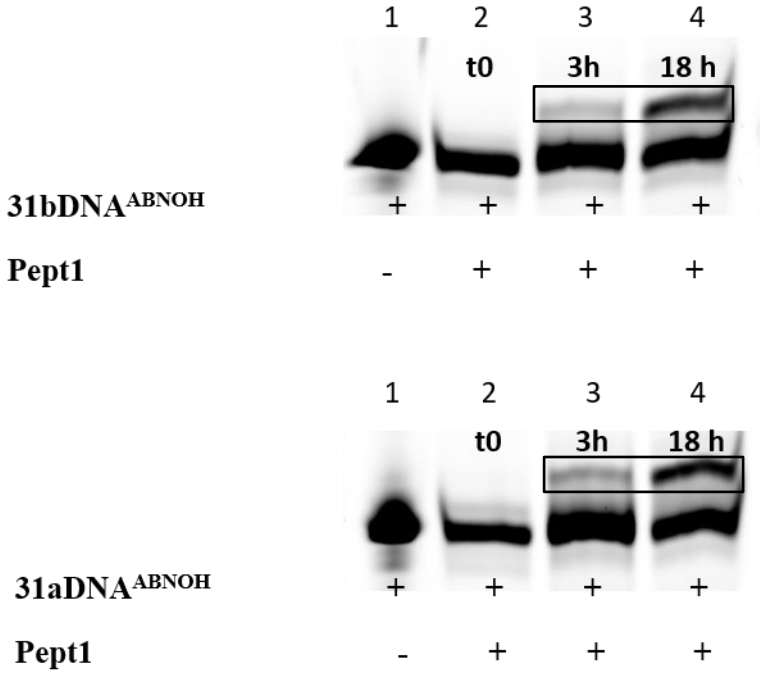


Figure 74. A) Conjugation of **31bDNA^{ABNOH}** (7 µM, lane 2,3,4) with **Pept1** (**Ac-AlaTrpAla**) (250 mM) in $NaHCO_3$ buffer (0.05 M, pH 10) at 37 °C for 0 min (lane 2), for 3 h (lane 3) for 18 h (lane 4). Control **31bDNA^{ABNOH}** (lane 1); B) Conjugation of **31aDNA^{ABNOH}** (7 µM, lane 2,3,4) with **Pept1** (250 mM) in $NaHCO_3$ buffer (0.05 M, pH 10) at 37 °C for 0 min (lane 2), for 3h (lane 3) for 18 h (lane 4). Control **31aDNA^{ABNOH}** (lane 1).

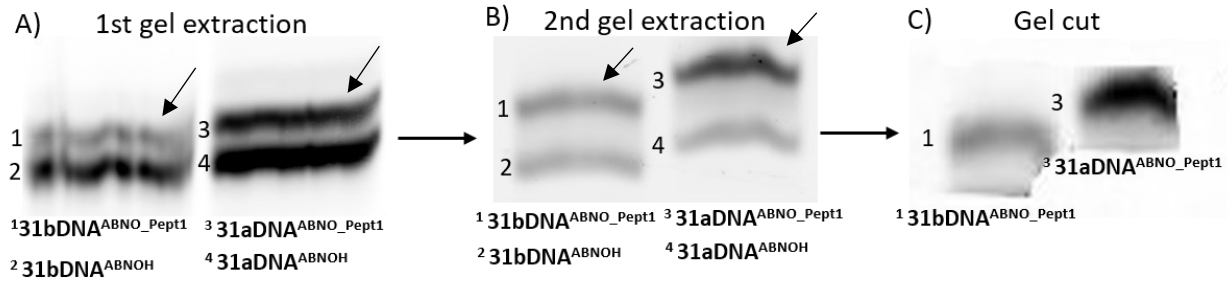


Figure 75. Denaturing 20% PAGE analysis for A) 1st gel extraction; B) analysis 2nd gel extraction; C) gel cut fragments of **31bDNA^{ABNO_Pept1}** and **31aDNA^{ABNO_Pept1}**.

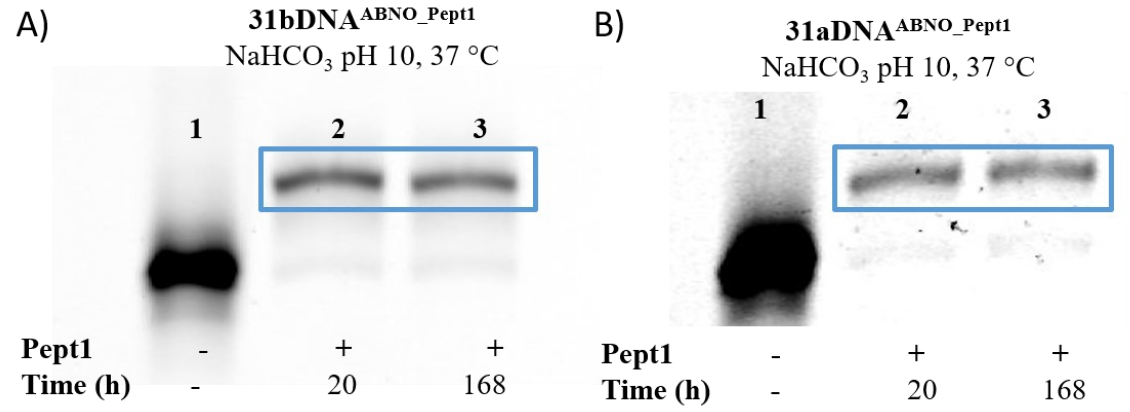


Figure 76. Denaturing 20% PAGE gel analysis of stability studies, after gel extraction, of A) **31bDNA^{ABNO_Pept1}** and of B) **31aDNA^{ABNO_Pept1}** incubated in H_2O for 20 h (lanes 2) and for 168 h (lanes 3).

Therefore, from PAGE gel analysis and UPLC-MS was confirmed even in case of DNA the formation of stable conjugates **31a/31bDNA^{ABNO_Pept1}** (undecomposed after 168 h) by covalent linkage between **31a/31bDNA^{ABNOH}** and **Pept1**.

3.4.1.9 Preparation of GntR from Bacillus subtilis (bsGntR)

In collaboration with Markéta Šoltysová from Maloy Pavlína Řezáčová's group was tested the cross-linking of **31a/31bDNA^{ABNOH}** with a Trp-containing DNA-binding protein. Tryptophan is rarely present in nature so was chosen ***bs*GntR** transcription factor as protein containing two Trp in its sequence. In particular, the ***bs*GntR** protein is a gluconate operon[119] transcriptional repressor from *Bacillus subtilis* which was prepared and provided by Marketa.

*bs*GntR (amino acid residues 2 – 243) was cloned into pET151/D-TOPO expression plasmid encoding an N-terminal hexahistidine (His6) tag and the tobacco etch virus protease (TEV PR) recognition site. In the case the His6 tag was cleaved, the six-amino-acid sequence GIDPFT remained at the N-terminus as a cloning artefact.

The recombinant protein was overexpressed in *E. coli* BL21 (DE3), then the cells were lysed by sonication and the supernatant purified by nickel-chelation chromatography on a His-select Nickel The His6-tagged *bs*GntR was eluted by step elution with the loading buffers and the less contaminated fraction was further incubated with TEV protease (TEV PR) in order to cleave off the His6 tag. The TEV PR and the His6 tag were subsequently separated by a second purification by His-select Nickel Affinity Gel (Ni-NTA) .

Final purification was performed by size-exclusion chromatography (SEC) before using the protein samples for the electrophoretic assays.

The sequence of the *bs*GntR protein no His tagged (the targeted position of W13 is highlighted in pink):

```
        10         20         30         40         50         60
MLDSKDLLYP AKWLSKASTG VRVAYELRMR IVSGLIESGT ILSENTIAAE FSVSRSPVRE

        70         80         90        100        110        120
ALKILASEKI IRLERMGAVV IGLTEKKIAE IYDVRLLLET FVFERLVKID IEPLVKDLSK

       130        140        150        160        170        180
ILEMMKVSIK YEDADEFSFQ DVLFHETIIR AIDHSYIQMI WNNLKPVMES FILLSMRVRL

       190        200        210        220        230        240
KEKYEDFTRI LDNHELYIQA IKTKDRALMI QSLHQNFDDV QDKVEDLWLS QQMLAKGAEY

NND
```

W = Trp

Moreover, according to the formation of *bs*GntR homodimers in solution which were proved for proteins from the GntR/FadR family[120], it is expected that two or three dimers bind the whole operator region. In particular the dimer represents the essential biological unit which recognizes the DNA binding site.

3.4.1.10 Design of DNA sequences 31aDNA^{Nat}/31aDNA^{ABNOH} or 31bDNA^{Nat}/31bDNA^{ABNOH} to target selectively W13 of bsGntR protein

The *bs*GntR protein forms homodimers in solution when highly diluted[119] and the dimer also represents the essential biological unit containing the effector-binding domain (EBD) and DNA-binding domain (DBD) (**Figure 77**, **A**), as it is common for proteins from the GntR/FadR family.[120,121] The *bs*GntR recognizes an operator sequence within the *gnt* promoter between positions -10 and +15 with respect to the transcription start point containing an 18 bp-long palindrome and two direct repeats 5'-GTATA-3' present in both strand directions (**Figure 77**, **B)**. In order to identify the best DNA positions for modifications, a rough model of the *bs*GntR-DNA complex by the employment of the AlphaFold 2 server[122] was constructed by Markéta Šoltysová, and using based also on existing cryo-EM model of a homologous transcriptional regulator, NanR, in a complex with its DNA operator[121] (**Figure 77, C**). The thymines at positions 31a and 31b (**Figure 77**, **D**) were identified as suitable to be modified based on two required assumptions. First, modifications of these nucleobases were not suspected to hinder the natural protein-DNA interactions. Second, they are supposed to be at a convenient distance from the targeted W13 residue, about the length of the covalent linker (21 Å).

Based on model and having considered the location of all thymines with respect to the protein, 31a and 31b positions were selected as the most suitable (**Figure 77, D**). The resulting designed and modified **31aDNA^{ABNOH}** and **31bDNA^{ABNOH}** as well as their natural analogues **31aDNA^{Nat}** and **31bDNA^{Nat}** were synthesized (Section 3.4.1.6) and used for cross-linking experiments with *bs*GntR.

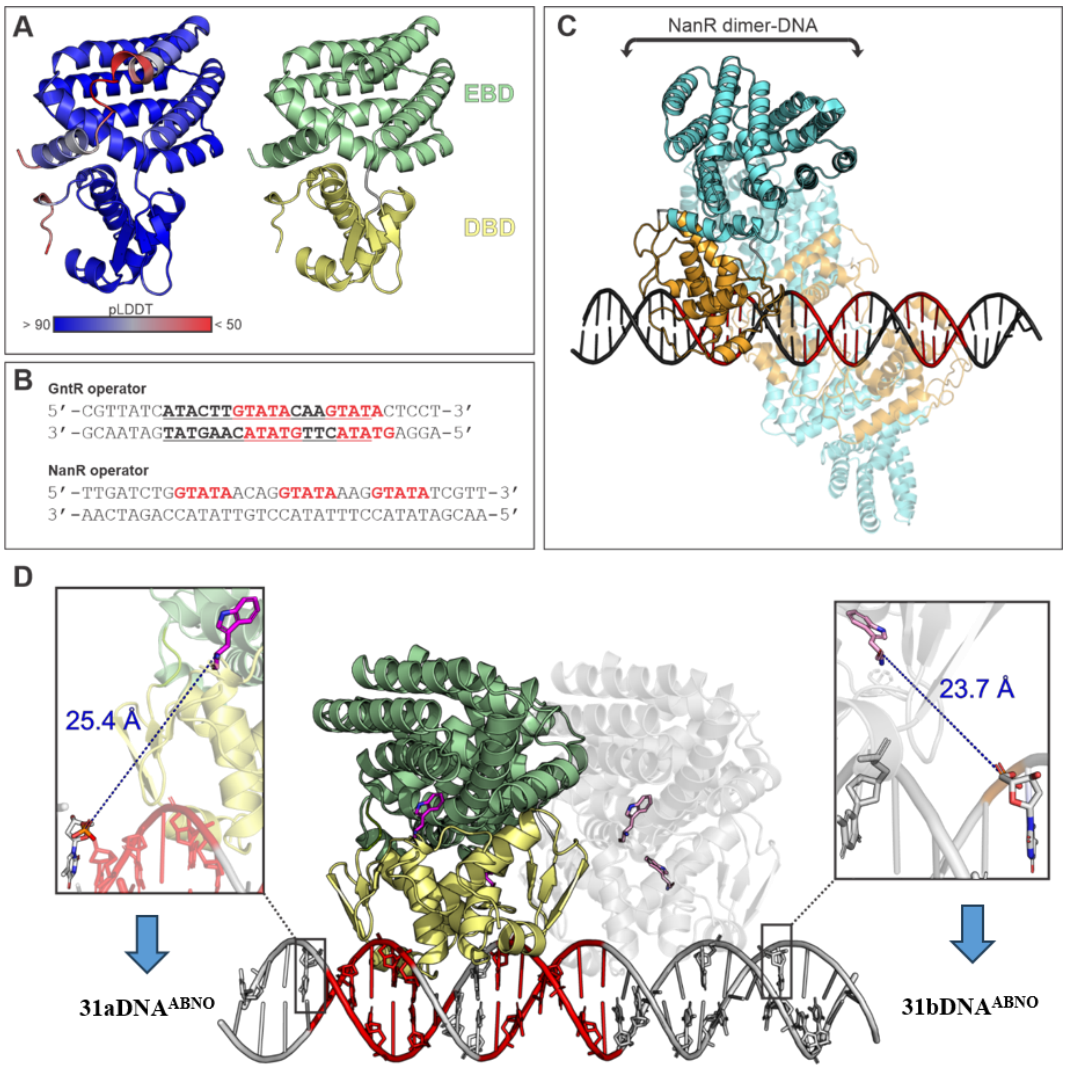
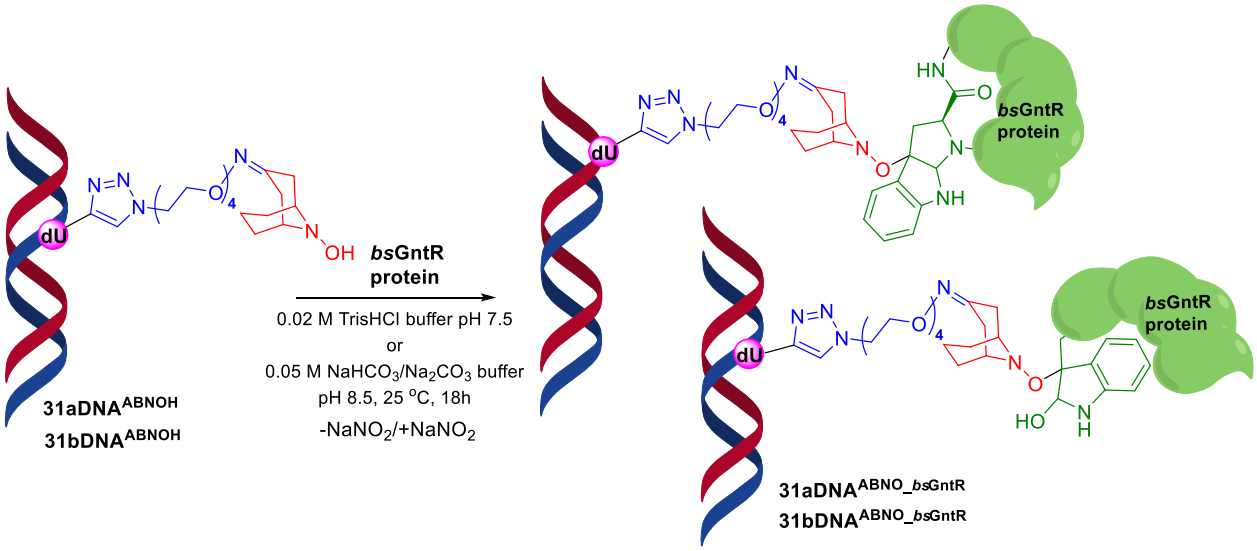


Figure 77. Reconstruction of the *bs*GntR-DNA complex for selecting suitable thymines to be modified. A) A cartoon representations of the *bs*GntR monomer predicted by AlphaFold 2 Algorithm.[122] On the left, the effector-binding domain (EBD) and DNA-binding domain (DBD) are distinguished by color and labelled. On the right, the model colored by gradient based on the prediction score (pLDDT) value shown as a gradient bar below. B) Sequences of DNA operators of GntR and NanR. In the GntR operator, the palindromic sequence is highlighted in bold and underlined. The shared patterns with NanR of "GTATA" are shown in red. C) A cartoon representation of a cryo-EM model of the NanR protein in the complex with its DNA operator.120 The dimeric unit is highlighted and the "GTATA" DNA signature patterns are shown in red. D) A resulting model of two possible positions of the *bs*GntR dimer onto an ideal B-DNA of the *bs*GntR operator sequence. Rough P - Cα distances between the selected thymines (31a and 31b) and W13 residues are shown in the zoomed-in boxes

3.4.1.10.1 Reactions of 31a/31bDNA^{ABNOH} with *bs*GntR

The proper designed DNA probes **31a/31bDNA^{ABNOH}** (previously prepared by PEX) were tested to cross-link ***bs*GntR** and the reaction was performed in TrisHCl buffer (pH 7.5) or $NaHCO_3/Na_2CO_3$ buffer (pH 10) at 37 °C for 18 h in presence or absence of $NaNO_2$ (**Scheme 29**). TrisHCl pH 7.5 and $NaHCO_3$ pH 10 buffers were tried in cross-linking reaction with the protein to compare the influence of different composition and pH buffer in the formation (%) of conjugates products **31a/31bDNA^{ABNO_bsGntR}**.



Scheme 29. Schematic reaction of **31aDNA^{ABNOH}** or **31bDNA^{ABNOH}** with ***bs*GntR** protein. Conditions: 0.02 M TrisHCl buffer (pH 7.5) or 0.05 M $NaHCO_3/Na_2CO_3$ buffer (pH 10) at 37 °C for 18 h in presence or absence of $NaNO_2$.

Also, different ratios of DNA to protein 1:1,1:4 and 1:32 were used and analysed by PAGE gel analysis (**Figure 78**, 1:1 – **Figures 79-80**, 1:4 – **Figures 81-82**, 1:32) to study the different reactivity of ABNOH modified- DNA (**31aDNA^{ABNOH}** or **31bDNA^{ABNOH}**) towards ***bs*GntR** protein.

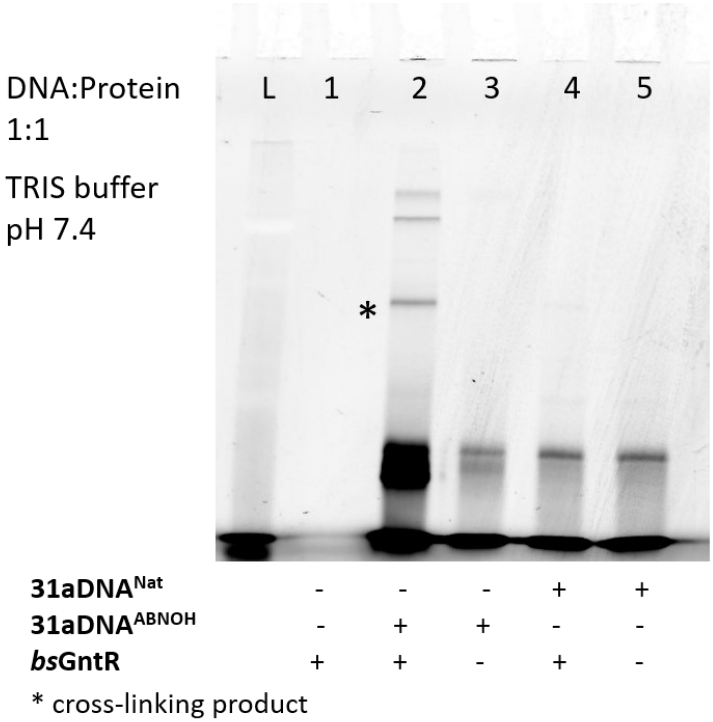


Figure 78. 10% Denaturing SDS-PAGE analysis, ratio DNA to protein 1:1, of **31aDNA^{ABNOH}** (lanes 2,3) or **31aDNA^{Natural}** (lanes 4,5) with ***bs*GntR** protein (lanes 1,2,4) incubated in Tris·HCl buffer (0.02 M, pH 7.5) at 37 °C for 18 h. L=ladder. Asterisk (*) indicates the conjugate of **31aDNA^{ABNOH}** with ***bs*GntR** protein. L=ladder.

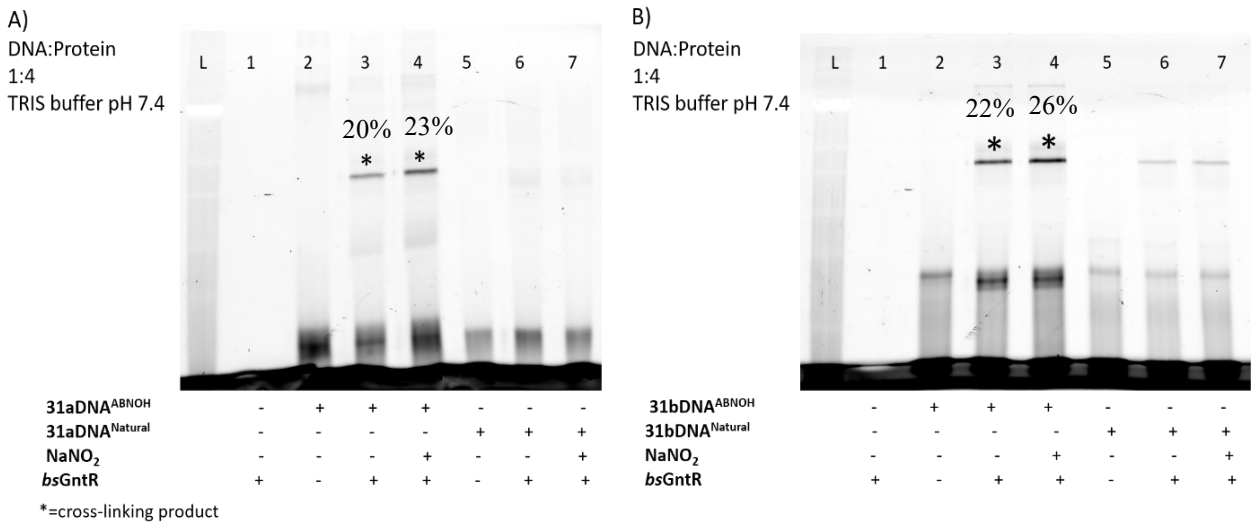


Figure 79. 10% Denaturing SDS-PAGE analysis of cross-linking of DNA to protein 1:4, comparison in presence or absence of $NaNO_2$. A) cross-linking of **31aDNA^{ABNOH}** (lanes 2-4) or **31aDNA^{Natural}** (lanes 5-7) with *bs*GntR protein (lanes 1,3,4,6,7) incubated in Tris·HCl buffer (0.02 M, pH 7.5) at 37 °C for 18 h; B) cross-linking of **31bDNA^{ABNOH}** (lanes 2-4) or **31aDNA^{Natural}** (lanes 5-7) with *bs*GntR protein (lanes 1,3,4,6,7) incubated in Tris·HCl buffer (0.02 M, pH 7.5) at 37 °C for 18 h.

Asterisk (*) indicates the conjugate of **31bDNA^{ABNOH}** with ***bs*GntR** protein. L=ladder.

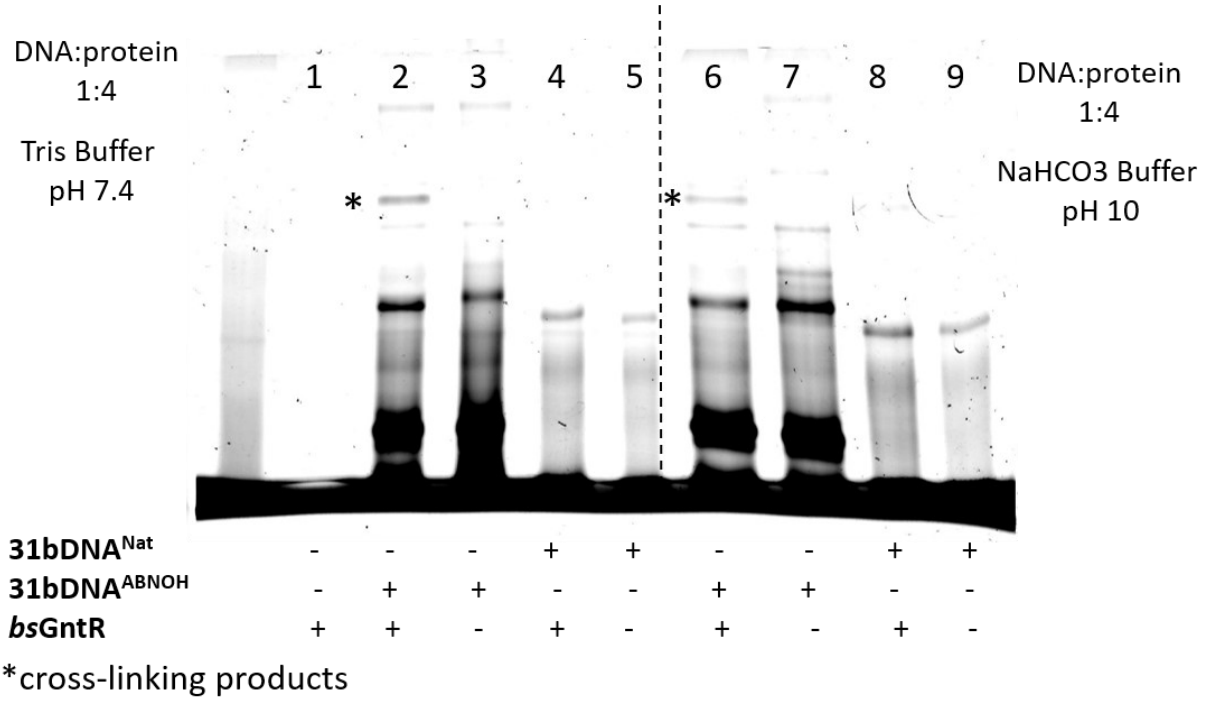


Figure 80. 10% Denaturing SDS-PAGE analysis of DNA-protein cross-linking of **31bDNA^{ABNOH}** (lanes 2,3,6,7) or **31bDNA^{Nat}** (lanes 4,5,8,9) with *bs*GntR protein (lanes 1,2,4,6,8) incubated in Tris·HCl buffer (0.02 M, pH 7.5) at 37 °C for 18 h. Right TrisHCl pH 7.4 buffer was used, left $NaHCO_3$ pH 10 buffer was used. Asterisk (*) indicates the conjugate of **31bDNA^{ABNOH}** with *bs*GntR protein. L=ladder.

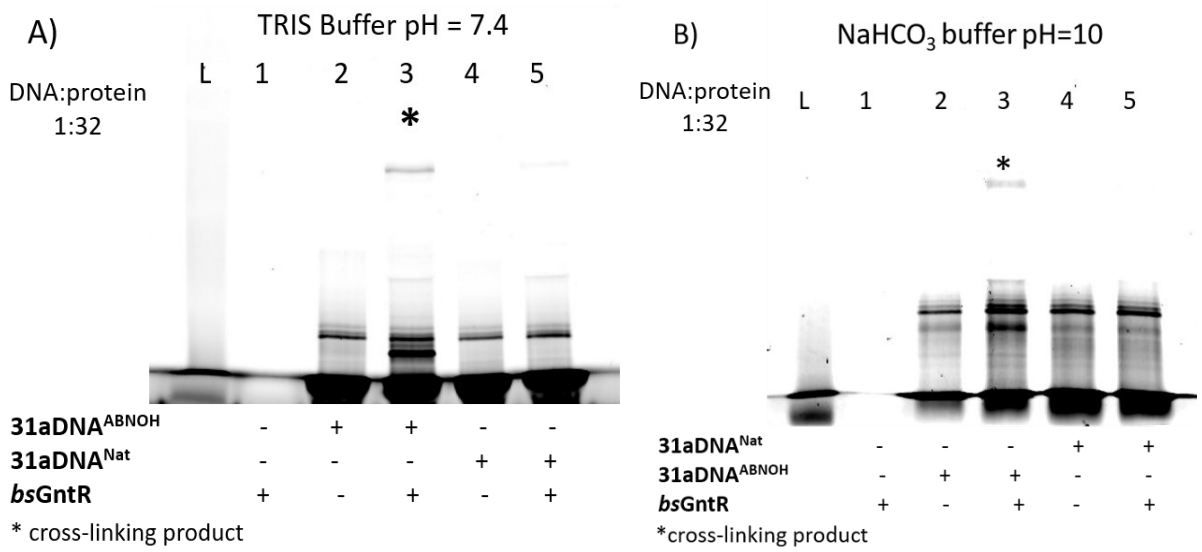


Figure 81. 10% Denaturing SDS-PAGE analysis of A) DNA-protein cross-linking of **31aDNA^{ABNOH}** (lanes 2,3) or **31aDNA^{Natural}** (lanes 4,5) with *bs*GntR protein (lanes 1,3,5) incubated in Tris·HCl buffer (0.02 M, pH 7.5) at 37 °C for 18 h and of B) DNA-protein cross-linking of **31aDNA^{ABNOH}** (lanes 2,3) or **31aDNA^{Natural}** (lanes 4,5) with *bs*GntR protein (lanes 1,3,5) incubated in $NaHCO_3$ buffer (0.05 M, pH 10) at 37 °C for 18 h. Star indicates the conjugate of **31aDNA^{ABNOH}_FAM** with *bs*GntR protein. Asterisk (*) indicates the conjugate of **31aDNA^{ABNOH}_FAM** with ***bs*GntR** protein. L=ladder.

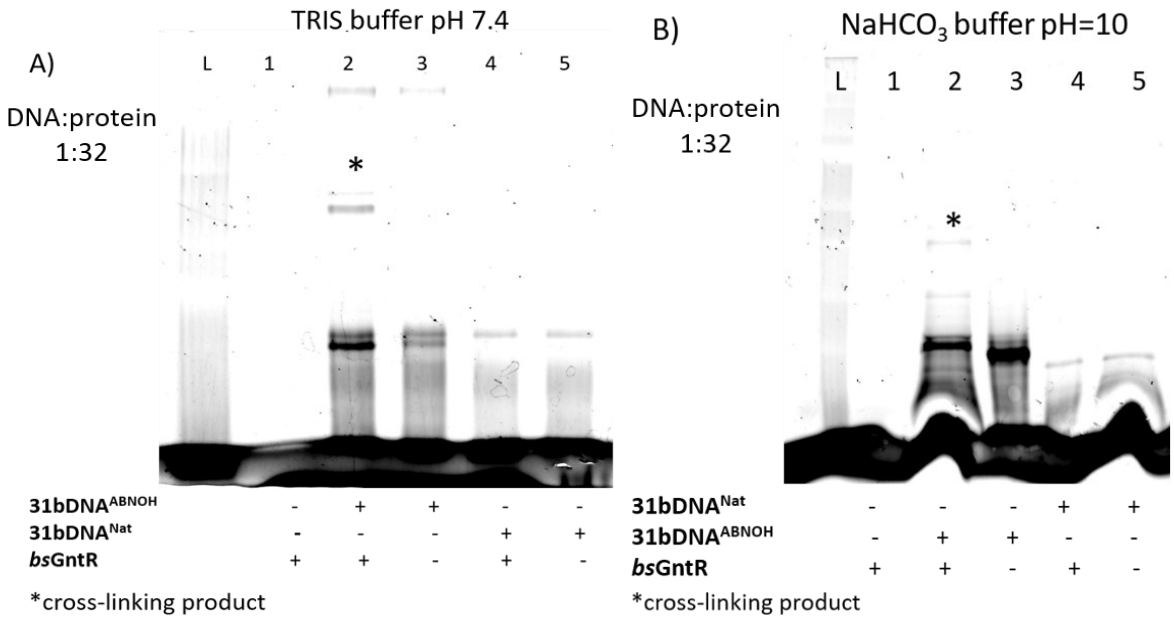


Figure 82. 10% Denaturing SDS-PAGE analysis, ratio 1:32 DNA:protein, of A) DNA-protein cross-linking of **31bDNA^{ABNOH}** (lanes 2,3) or **31bDNA^{Nat}** (lanes 4,5) with *bs*GntR protein (lanes 1,4,5) incubated in Tris·HCl buffer (0.02 M, pH 7.5) at 37 °C for 18 h and of B) DNA-protein cross-linking of **31bDNA^{ABNOH}** (lanes 2,3) or **31bDNA^{Natural}** (lanes 4,5) with *bs*GntR protein (lanes 1,2,4) incubated in $NaHCO_3$ buffer (0.05 M, pH 10) at 37 °C for 18 h. Asterisk (*) indicates the conjugate of **31bDNA^{ABNOH}** with ***bs*GntR** protein. L=ladder.

From SDS-PAGE gel analyses and by comparison of different ratios used of DNA:protein the best conversions were observed when used 1:4 DNA to protein ratio with 23-26% or 20-22% in absence or presence of $NaNO_2$ for **31aDNA^{ABNOH}_FAM** and **31bDNA^{ABNOH}** respectively, using TrisHCl pH 7.5 buffer (**Figure 79**).

The conjugates conversions were calculated for all reaction conditions used by ImageJ from intensity bands of SDS-PAGE gels (**Table 18**).

Table 18. Conversion calculation table of cross-linking reactions with *bs*GntR protein in TrisHCl pH 7.4 and $NaHCO_3$ pH 10 buffers

Protein	NaNO 2	Ratio DNA:protein	Oligonucleotide	Buffer	Temperature	Time	Conversion %
***bs*GntR**	NO	1:32	**31aDNA^{ABNOH}**	TrisHCl pH 7.4	37 °C	18 h	7%
	NO	1:32	**31aDNA^{ABNOH}**	$NaHCO_3$ pH 10	37 °C	18 h	6%
	NO	1:32	**31bDNA^{ABNOH}**	TrisHCl pH 7.4	37 °C	18 h	10%
	NO	1:32	**31bDNA^{ABNOH}**	$NaHCO_3$ pH 10	37 °C	18 h	6%
	NO	1:1	**31aDNA^{ABNOH}**	TrisHCl pH 7.4	37 °C	18 h	8%
	NO	1:4	**31aDNA^{ABNOH}**	TrisHCl pH 7.4	37 °C	18 h	22%
	YES	1:4	**31aDNA^{ABNOH}**	TrisHCl pH 7.4	37 °C	18 h	26%
	NO	1:4	**31bDNA^{ABNOH}**	TrisHCl pH 7.4	37 °C	18 h	20%
	YES	1:4	**31bDNA^{ABNOH}**	TrisHCl pH 7.4	37 °C	18 h	23%

Note: all DNA products are 5`-FAM labelled.

All conversions (%) of conjugates products were low and this can be due to not abundancy of Trp in ***bs*GntR** sequence, which makes it difficult to target it specifically and selectively. Additionally, the worse conversions resulted when $NaHCO_3$ pH 10 buffer is used, the best in case of TrisHCl pH 7.4 buffer.

3.4.1.11 Reaction of 19DNA^{Nat} or 19DNA^{ABNOH} with DNA-binding proteins not containing Trp (Histones H2A and H4) and with non-DNA-binding protein containing Trp (Lysozyme) – SDS PAGE analysis

In order to have formation of covalent bond in a cross linking reaction between DNA and protein, two requirements should be fulfilled:

The proximity effect which takes place when protein contains a DNA-binding sequence;

The presence of a specific aminoacid to target.

In order to prove these assumptions, the cross-linking reaction of **19DNA^{Nat}** or **19DNA^{ABNOH}** with Histones (H2A and H4) and lysozyme were performed.

In both cases histones (H2A, H4A) and lysozyme proteins were used as negative controls, suitable for studies because a) histones are lacking of Trp, b) lysozyme contains Trp but does not have any DNA-binding sequence. The reaction mixture (10 μL) containing **19DNA^{Nat}** or **19DNA^{ABNOH}** (1 μM), previously synthetized by PEX, was incubated with Histones H2A, H4 or Lysozyme (5 μM) in AcOH 1 mM at 25 °C for 1 h in a thermal cycler and subsequently stopped by addition of SDS-PAGE stop solution (10 μL) and heated 2 min at 99 °C. Samples were analyzed by 17.5 % denaturing SDS-PAGE and visualized using florescence imaging (**Figures 83-84**).

The use of histones (H2A, H4A) and lysozyme confirmed the importance of both, DNA-binding sequence in the protein and of a specific aminoacid to target in order to form a covalent adduct. The non formation of any cross-linking product, as shown by SDS-PAGE analyses, was the confirmation of the importance of the two aforementioned requirements (**Figures 83-84**).

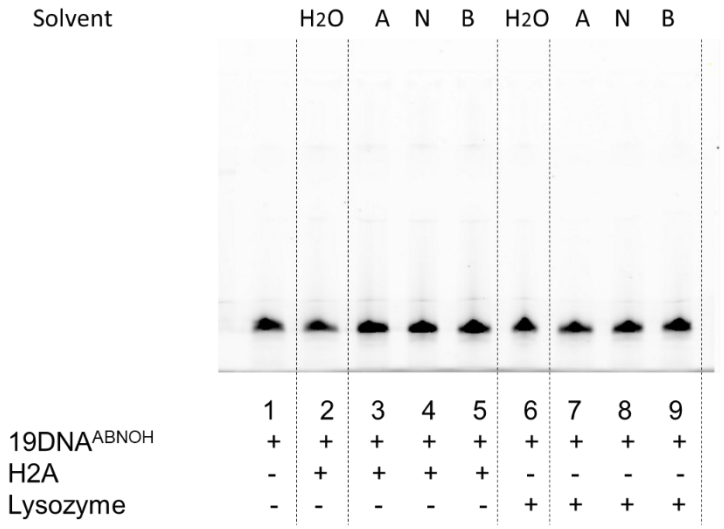


Figure 83. Denaturing SDS-PAGE analysis of cross-linking reactions between **19DNA^{ABNOH}** (lanes 1-9) with H2A (lanes 2-5), or Lysozyme (lanes 6-9) in H_2O (lanes 2,6) or 50 mM phosphate buffer pH 5 (lanes 3,7), pH 7.4 (lanes 4,8) and pH 8.5 (lanes 5,9) at 37 °C for 1 h.

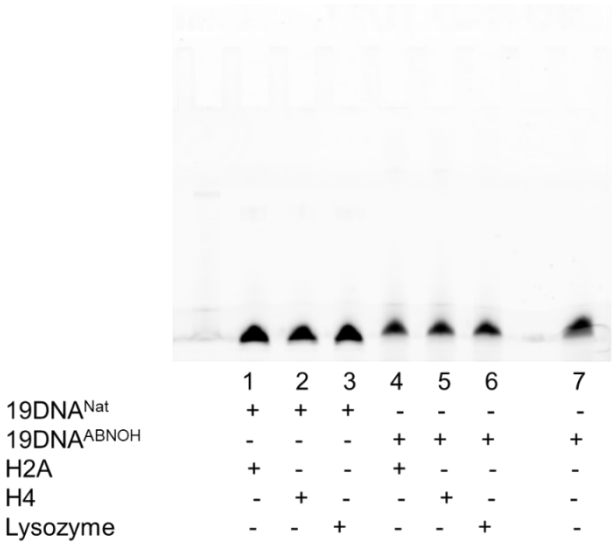


Figure 84. Denaturing SDS-PAGE analysis of cross-linking reactions between **19DNA^{Nat}** (lanes 1-3) and **19DNA^{ABNOH}** (lanes 4-7) with H2A (lanes 1,4), H4 (lanes 2,5) and Lysozyme (lanes 3,6) in AcOH 1 mM at 25 °C for 1h.

3.4.1.12 Native mass analysis of 31DNA^{ABNO_bsGntR} adduct

The native mass analysis was performed by Alan Kádek and Marek Polák from Biocev institute to characterize the **31aDNA^{ABNO_bsGntR}** conjugate. The crude reaction mixture, following the conjugation of FAM-labelled **31aDNA^{ABNOH}** with the *bs*GntR protein (ratio 1:1 of **31DNA^{ABNO}:*bs*GntR**) as described above in paragraph 3.10.14, was analysed. It was buffer exchanged into 150 mM ammonium acetate pH 7.4 (99.99% trace metals basis, Merck) by six cycles of 10-fold dilution and re-concentration on Vivaspin 500 centrifugal concentrators (30 kDa MWCO, Sartorius). Without further dilution, the samples were introduced into a Waters Synapt G2Si mass spectrometer by in-house manufactured golden-coated borosilicate glass capillaries (1B120F-4, World Precision Instruments). The spectrometer was operated in positive sensitivity ion mode over the 500 – 20,000 m/z range and was tuned to maximize ion signal, while maintaining low ion activation and non-covalent protein interactions. Namely, the ESI capillary and sampling cone were kept at 1.2 kV and 80 V, respectively. Argon was used as the collision gas at 6 ml/min flow with 40 V trap collision energy to aid in gentle ion desolvation (**Figure 85**). Data were smoothed and externally mass recalibrated on cesium iodide clusters in MassLynx 4.2 (Waters) with mass deconvolution through a combination of UniDec 6.0.4 algorithm13 and manual deconvolution[123] using the MaCSED utility from Benesch research group at the University of Oxford (http://benesch.chem.ox.ac.uk/resources.html).

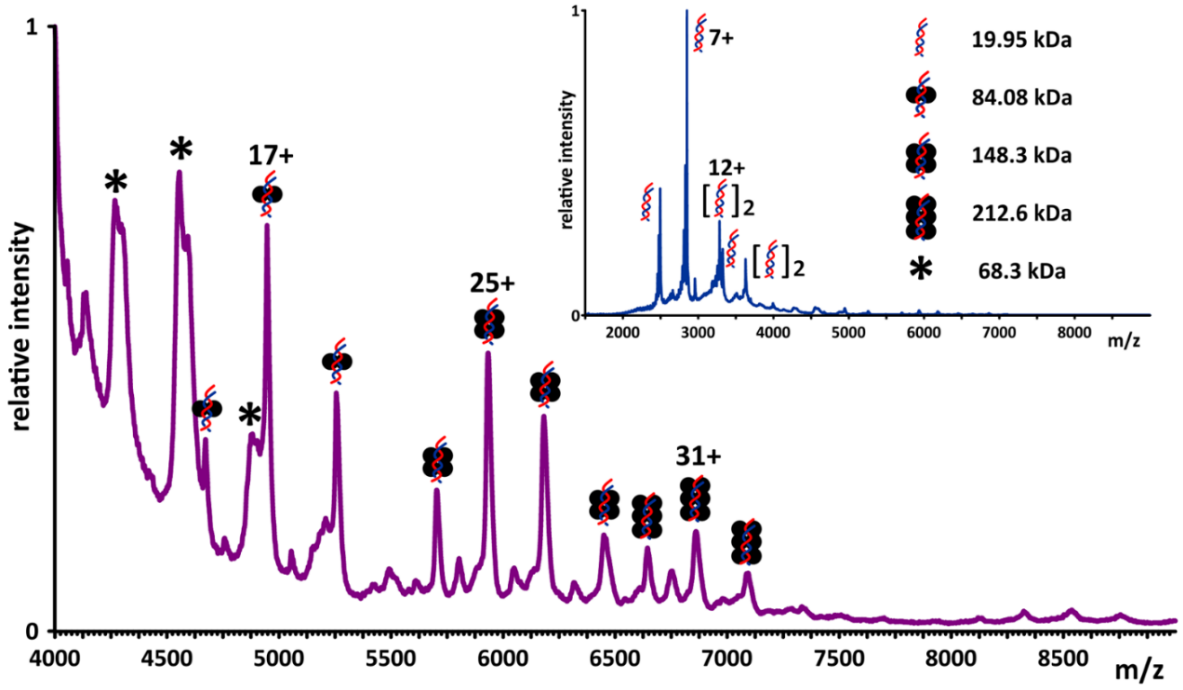


Figure 85. Native MS of **31aDNA**$^{\text{ABNO_}bs\text{GntR}}$ complex - electrosprayed from 150 mM ammonium acetate pH 7.4. Under native ESI conditions after the coupling reaction, the *bs*GntR protein (black balls) remains as a mixture of 2:1, 4:1 and to an extent 6:1 (*bs*GntR:dsDNA) stoichiometry species, always in complex with a single strain of dsDNA (red/blue helix). An excess of free dsDNA is visible in the inset. There is another contaminant protein species unrelated to GntR of about 68 kDa also present in the reaction mixture (asterisk).

As outcome of native mass analysis, the *bs*GntR was detected as non-covalent complex, still present throughout the conditions of the cross-linking reaction, in 2:1, 4:1 and 6:1 ratios (*bs*GntR:dsDNA) bound to a single dsDNA molecule **31aDNA**$^{\text{ABNO_}bs\text{GntR}}$ under the conditions of excess DNA according to the assumption of multiple dimers acting together on a single operator recognition motif (**Figure 85**). In this fourth and last part of the thesis, modified ABNOH-linked nucleotides and DNA were synthetized and reacted with Trp-containing peptides to form a stable conjugate in the presence but surprisingly even in the absence of $NaNO_2$ (most probably through activation by O_2). Therefore, the reactive ABNOH-modified DNA probe was successfully used for bioconjugations and crosslinking with Trp-containing peptides or proteins.

4 CONCLUSIONS

The first goal of the thesis has been the design and synthesis of DNA-reactive probes bearing strained dienophiles, ready to react with tetrazines by IEDDA reaction for bio-imaging application. The approach consisted in using trans-cyclooctene (2TCO, 4TCO) and byclononyne (BCN) compounds for the synthesis of modified nucleotides as building blocks for enzymatically synthetized DNA-probes. In particular, the reactive *trans*-cyclooctene needed to not isomerize in the unreactive *cis*- isomer, resisting under triphosphorylation reaction conditions and afterwards the enzymatical incorporation. The main idea was to use these strained dienophile, first *in vitro* and then *in cellulo*, to directly clicked coumarin-tetrazines to generate fluorogenic conjugates.
Two series of modified 2'-deoxyribonucleotides bearing 2TCO-, 4-TCO or BCN- modification thethering short or PEG linker were successfully synthetized starting from 5-propargylamino-2'-deoxycitidine as building block and then 2TCO-, 4TCO- or BCN-NHS carbonate were substituted by the amino group. As next step, model reaction on 4TCO- modified monophosphate with simmetric T5 tetrazine was performed to understand the connectivity of the conjugate and characterize it. Unfortunately the complete assignment of ^{1}H and ^{13}C NMR spectra was not possible because the product was a mixture of several region- and stereoisomers. The reactivity of all modified triphosphates was tested with different tetrazines (T1-T3) which were all characterized by LC-MS. The conversions were lower for dNTPs containing the shorter propargyl-carbamate linker (27–78%) and better for the compounds bearing the longer PEG3-based linker (55–90%), and the reactivity of tetrazine T2 was lower than that of tetrazines T1 and T3. In parallel the "turn-on" values of the IEDDA click reactions were studied by fluorescence spectroscopy. The largest turn-on effects, 210- and 220-fold, were observed with **dC^{4TCO}TP** and **dC^{p4TCO}TP**, respectively. Instead, the analogous 2-TCO isomers, **dC^{2TCO}TP** and **dC^{p2TCO}TP** showed 130- and 125-fold, respectively. For reaction with T2 a 5-fold light-up effect was obtained for the reaction with **dC^{4TCO}TP**, whereas a 3-fold increase of fluorescence was observed in combination with **dC^{2TCO}TP** and **dC^{pBCN}TP**. Additionally, in case of acridine orange-tetrazine conjugate **T3** was shown light-up of 15 with **dC^{p4TCO}TP** and light-up factor of 50 with **dC^{pBCN}TP**. In order to have a comparison of the fluorescence intensity increase of the products at the same concentration, fluorescence intensities at half-life of reactions were calculated. By comparison was revealed that both products of reactions of **T1** with **dC^{p2TCO}TP** and **dC^{p4TCO}TP** have comparable fluorescence

intensities. Lower reaction rate (0.0021 s^{-1}) was exhibited by product of conjugation of **dC^{pBCN}TP** with **T1** and significantly lower brightness than the TCO products at half-life. The rate constant (k_{obs} = 0.00402 s^{-1}) for the reaction of **dC^{p4TCO}TP** and **T2** was about half of that observed with **T1**. Furthermore, the modified triphosphates were enzymatically incorporated by KOD XL polymerase into DNA and the desired products were isolated and characterized by MALDI-TOF. Even for reactive functionalized dsDNA was tested the reactivity towards **T1-T4** tetrazines and the formation of conjugates were confirmed by SDS-PAGE and LC-MS. The first tetrazine tested was the **T1** giving higher conversion (81–90%) with PEG3-linked TCO-derivatives **19DNA_C^{p4TCO}** and **19DNA_C^{p2TCO}** than the **19DNA_C^{4TCO}** and **19DNA_C^{2TCO}** with short linkers (42–64%), indicating the positive role of a longer linker. Also, the BCN-linked **19DNA_C^{pBCN}** was efficiently labeled with **T1**. On the other hand, tetrazine **T2** reacted readily and completely only with BCN-linked **19DNA_C^{pBCN}**, whereas the TCO-linked DNA reacted with **T2** poorly with only incomplete conversions. As last, **T3** and **T4** tetrazines were tested revealing poor reactivity of **T3** conjugates in which higher incubation time and higher concentration of **T3** was required and for **T4** was observed complete conversion only in case of **19DNA_C^{p4TCO}**, whereas partial conversion with **19DNA_C^{p2TCO}** and **19DNA_C^{pBCN}**. The last step of this project was the post-labeling *in cellulo* and various combinations of reactive group–modified nucleotides and tetrazines **T1**–**T4** for the labeling of DNA in live U-2 OS cells. The modified **dC^{X}TPs** were delivered to cells by nucleoside triphosphate transporter **SNTT1** and incorporated in genomic DNA after incubation time of 60 min. Subsequently, treatment of cells with one of the tetrazines (**T1**–**T4**) revealed that the DNA with incorporated **dC^{p2TCO}**, **dC^{p2TCO}**, **dC^{p4TCO}**, and **dC^{pBCN}** was labeled efficiently with **T1** and **T2**. Unfortunately, cells treated with **dC^{2TCO}TP** or **dC^{4TCO}TP** did not show labeling with **T1** or with **T2**. Tetrazine **T3** labeled only the DNA with incorporated **dC^{pBCN}** requiring longer reaction time (4 h). Commercially available **T4** was able to label the DNA with incorporated **dC^{p4TCO}** or **dC^{pBCN}** in live cells but less efficiently due to low cell membrane permeability. When the cells were fixed after the incorporation of **dC^{p4TCO}TP** or **dC^{pBCN}TP**, labeling with **T4** allowed clear images of nuclei with DNA foci. Moreover, the **dC^{pBCN}TP**/**T4** system revealed the best contrast in imaging experiments and was also successfully used in a cell cycle cytometric assay allowing the quantification of cells at different stages of the cell cycle. This two-step procedure developed to label the DNA in live cells can find application particularly in cases where the direct incorporation of fluorophore bearing- dNTPs is not possible due to the

incompatibility of the fluorophore with cellular polymerases. To conclude, this newly developed method is useful to distinguish DNA-incorporated modified nucleotide from the non-incorporated one. The limitation of the method is the impossibility to monitor cells actively synthetizing DNA in real time. This goal has been accomplished in case of thiazol orange modified DNA which was synthetized afterwards in Hocek group.[124]

The second aim of the thesis has been the design and synthesis of DNA-reactive probes specific to cross-link certain aminoacids. The strategy consisted in using β-sultams, azirine or keto-ABNOH compounds for the synthesis of modified nucleotides and DNA-probes selective for serine or glutammic/aspartic acids or tryptophan, respectively. Different aminoacids were taken in consideration to extend the portfolio of this study. Modified 2'-deoxyribonucleotides bearing the β-sultams or azirine ring were synthetized by Cu-catalysed cycloadditions (CuAAC) or Sonogashira cross coupling. Unfortunately, both reactive probes were found, and proved by IR and mass analysis for β-sultams-/oxo-β-sultams- modified nucletides and by NMR and mass analysis for azirine- modified nucleotides, to be hydrolysed and unable to react with serine or glutammic/aspartic acids due to presence of copper or palladium. Furthermore, in case of β-sultams-/oxo-β-sultams- click reaction on DNA was tried in order to insert postsynthetically the reactive moiety but was unsuccessful.

An other strategy was to use keto-ABNOH as possible reactive probe for tryptophan.

ABNOH-linked 2'-deoxyuridine nucleotides **dU^{ABNOH}MP** and **dU^{ABNOH}TP** were synthetized via CuAAC reaction of 5-ethynyluracil nucleotides with ABNOH-PEG$_4$-N$_3$. Model reaction on modified monophosphate **dU^{ABNOH}MP** was performed with excess of **Ac-Trp** and **Pept1** tripeptide. The formation of cross-linking products was observed not only in presence of $NaNO_2$ but also in its absence which were proved by NMR (even though with partial assignment) and mass analyses. Notably, the oxygen itself is a sufficient oxidant for activation of the ABNOH moiety to the reactive N-oxo cation. The ABNOH- modified triphosphate **dU^{ABNOH}TP** was a good substrate for KOD XL DNA polymerase in PEX reactions to get reactive DNA probes bearing the ABNOH modification at specific position. Similarly, as model reaction on monophosphate, the ABNOH-DNA probes was reacted with saturated solution of Trp or with large excess of Trp-containing peptides (ca. 105 equiv.) with formation of cross-linked products even without the use of $NaNO_2$ radical-activator which were proved by LC-MS. On the other hand, peptides lacking

Trp did not form any cross-linked product which were used as negative controls. Proper designed ABNOH-modified DNA probes (31 mer) targeting tryptophan (W13) of bsGntR protein were formed a covalent adduct using two different buffers at pH 10 or 7.5. The covalent cross-link was proved by denaturing SDS-PAGE gel electrophoresis, whereas the stoichiometry of the protein binding was determined by native mass spectrometry which was detected exclusively in stoichiometries of 2:1, 4:1 and 6:1 (protein:DNA). In general, higher conversions of the cross-linking reactions were given in presence of $NaNO_2$ but even the reactions in the absence of this activator (which may cause toxicity and some unwanted side reactions on DNA) proceeded with acceptable yields. In conclusion, three different reactive moieties were investigated in order to synthetize modified-DNA for bioconjugation and cross-linking with peptides and proteins. The two moieties β-sultam/3-oxo-β-sultam and azirine were attempted to be attached in nucleotides and additionally for β-sultam/3-oxo-β-sultam to be inserted in DNA but they were hydrolyzed. Therefore they were not found applicability in further reactions with Ser- or Glu/Asp- containing peptides and proteins, but targeting these aminoacids remain a challenge for future research. Instead, ketoABNO- modified nucleotides and DNA were stable, and they have successfully reacted with Trp- containing peptides and proteins forming covalent conjugates. In the future this reactive ABNOH-linked nucleotide building block could be a potential tool to search for new DNA-binding proteins or bioconjugations with Trp-containing peptides and proteins. Possible applications in structural biology can be limited by the fact that a mixture of stereoisomers of the conjugate are formed.

5 EXPERIMENTAL PART

General remarks for the synthetic part – for TCO-/BCN project

^{1}H, ^{13}C and ^{31}P NMR spectra were acquired on a Bruker AVANCE IIIHD 600 (1H at 600.1 MHz, ^{13}C at 150.9 MHz), Bruker AVANCE IIIHD 500 (^{1}H at 500.0 MHz, ^{13}C at 125.7 MHz, ^{31}P at 202.4 MHz) and JEOL ECZR 500 (^{1}H at 500.2 MHz, ^{13}C at 125.8 MHz, ^{31}P at 202.5) spectrometers, as indicated. ^{1}H and ^{13}C resonances were fully assigned using H,H-COSY, H,H-ROESY, H,C-HSQC and H,C-HMBC techniques. All chemical shifts are quoted on the δ scale in ppm and referenced using residual ^{1}H solvent signal in ^{1}H NMR spectra ($\delta(CHCl_3)$ = 7.26 ppm; $\delta(CHD_2OD)$ = 3.31 ppm; and ^{13}C solvent signal in ^{13}C NMR spectra $\delta(CDCl_3)$ = 77.0 ppm; $\delta(CD_3OD)$ = 49.0 ppm). NMR spectra measured in D_2O were referenced to the signal of *t*-BuOH (10% v/v solution in D_2O, 1 drop) as the internal standard (1.24 ppm in ^{1}H, 32.43 ppm in ^{13}C). ^{31}P NMR spectra were referenced to H_3PO_4 signal (0 ppm) as the external standard. Coupling constants (*J*) are reported in Hz with the following splitting abbreviations: s = singlet, d= doublet, t = triplet, q = quartet, m = multiplet. Fluorescence measurements were performed on a FluoroMax 4 spectrofluorometer (Jobin Yvon, Horiba) from *Perkin Elmer* equipped with a 450 W xenon lamp and a single cuvette reader using data algebra formula S1c/R1c. High resolution mass spectra were measured on LTQ Orbitrap XL (Thermo Fisher Scientific) using Electrospray Ionization technique (ESI) or on on UPLC Agilent 1290 Infinity II bio system equipped with DAD and LC/MSD XT MS (operating in negative mode) detectors, with bioZen™ Oligo LC column (1.7 µm, 50×2.1 mm; Phenomenex®) using a linear gradient of 12 mM Et_3N and 300 mM HFIP in H_2O to 12 mM Et_3N and 300 mM HFIP in MeOH in 30 min.

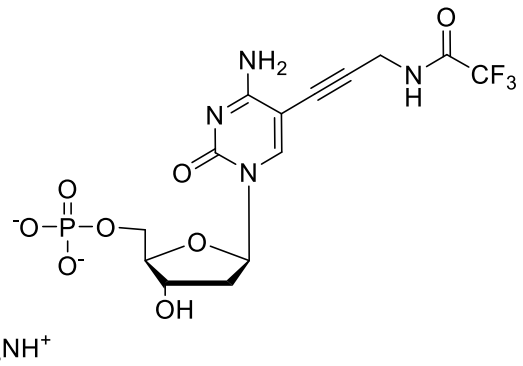
Reactions were monitored by thin layer chromatography (TLC) on TLC Silica gel 60 F254 (Merck) and detected by UV (254 nm) or by solution of 4-anisaldehyde 3.6% v/v in ethanol and 10% v/v of sulphuric acid. Reactions were monitored by Advion Expression Compact Mass Spectrometer connected with Plate Express® TLC Plate Reader using ESI. High performance flash chromatography (HPFC) was performed with Biotage SP1 apparatus on DEAE SEPHADEX A-25 sodium form columns or with ISCO Combiflash Rf system on RediSep Rf Gold Silica Gel Disposable columns. Purification of nucleoside triphosphates was performed using HPLC (Waters modular HPLC system) on a column packed with 10 µm C18 reversed phase (Phenomenex, Luna C18 100 Å). Reactions were analysed by TLC using IPAV (isopropyl alcohol /NH_4OH /H_2O, ratio 11/7/2) as mobile phase. All materials were purchased from commercial suppliers and used

without further purification unless otherwise stated. $POCl_3$ and $PO(OMe)_3$ were distilled prior to use. The water used in synthetic part was of HPLC quality. Chemicals were of analytical grade. (*E*)-cyclooct-4-en-1-yl-2,5-dioxo-1-pyrrolidinyl ester carbonic acid (**4TCO-NHS-carbonate**), (*E*)-cyclooct-2-en-1-yl-2,5-dioxopyrrolidin-1-yl ester carbonic acid (**2TCO-NHS-carbonate**), (E)-1-(cyclooct-4-en-1-yloxy)-2,5-dioxopyrrolidin-1-yl-1-(oxo-5,8,11-trioxa-2-azatetradecan)-14-oate (**4TCO-PEG3-NHS-ester**), (E)-1-(cyclooct-2-en-1-yloxy)-2,5-dioxopyrrolidin-1-yl-1-(oxo-5,8,11-trioxa-2-aza)-tetradecan-14-oate (**2TCO-PEG3-NHS-ester**) were purchased from Sirius Fine Chemicals SiChem GmbH. 2,5-dioxopyrrolidin-1-yl 1-(bicyclo[6.1.0]non-4-yn-9-yl)-3-oxo-2,7,10,13-tetraoxa-4-azahexadecan-16-oate (**endo-BCN-PEG3-NHS-ester**) was purchased from BroadPharm. Other chemicals were of analytical grade. All the EPR spectra were acquired on *Bruker* EMX^{plus} 10/12 CW (continuous wave) spectrometer equipped with a Premium-X-band microwave bridge and the standard rectangular probehead/cavity (ER160FCQ, *Bruker*). The g_{iso} value of radicals was determined using a built-in spectrometer frequency counter and a ER036TM NMR-Teslameter (*Bruker*). Typical uncertainty for the g-value determination reads ± 0.0002. Compounds were dissolved either in water or $CHCl_3$ and sampled into glass $50\,\mu L$ capillaries ($i.d. = 0.78\,\mathrm{mm}$, *Hirschmann Ringcaps*)(1). Typical EPR recording included following parameter setup: microwave frequency $9.807\,\mathrm{GHz}$, microwave power $2.5\,\mathrm{mW}$, central filed $349\,\mathrm{mT}$, sweep width $(16-30)\,\mathrm{mT}$, resolution (number of points) $1500-3000$, modulation amplitude $(0.02-0.12)\,\mathrm{mT}$, number of scans $10-150$. Kinetic measurements were performed by the preset 2D_Field_Delay experiment. The time corresponding to each EPR spectrum in the kinetic series was corrected in order coincide with the midpoint (B) of the spectrum because the preset takes the starting time of each sweep (at which the CW EPR spectrum actually does not appear). Calculations: Concentration of aminoxyl-based radicals was determined by double integration of the corresponding EPR spectra directly within the *Xenon* acquisition-processing EPR software (*Bruker)*. Prior to integration baseline correction was applied to each spectrum/sweep. The calculation procedure takes into account the geometry and calibration of the applied EPR probehead/cavity as well as all additional (instrumental + total spin quantum number S of investigated system, in our case $S = 1/2$) parameters required for absolute determination of paramagnetic centers/aminoxyl radicals with the sample.

5.1 Design and synthesis of modified nucleotides bearing 2- or 4-linked *trans*-cyclooctene (TCO) or bicyclononyne (BCN) tethered through a shorter propargylcarbamate or longer triethyleneglycol-based spacer, their enzymatic incorporation and testing of reactivity with tetrazines

5.1.1 Synthesis of modified monophosphates (dC^{PAF3}MP, dC^{NH2}MP, dC^{4TCO}MP) and modified triphosphates (dC^{4TCO}TP, dC^{2TCO}TP, dC^{p4TCO}TP, dC^{p2TCO}TP, dC^{pBCN}TP)

Synthesis of 5-(4-amino-2-oxo-5-(3-(2,2,2-trifluoroacetamido)prop-1-yn-1-yl)-2'-deoxycytidine-5-*O*-monophosphate triethylammonium salt (dC^{PAF3}MP) (4a)



Compound **dC^{PACF3}** [109] (250 mg, 687.5 µmol) was dried at 25 °C overnight in vacuo. After cooling on ice, $PO(OMe)_3$ (2.5 mL) and $POCl_3$ (200 µL) were added under argon atmosphere. The reaction mixture was stirred for 2 h at 0 °C. The product was purified by SEPHADEX Na+ using H_2O/TEAB 2 M buffer as eluent and subsequentely by RP-HPLC using a Phenomenex HPLC column (Luna 10 C8 100Å, 10×250 mm). Several co-distillations with water followed by freeze-drying gave 80% pure product **dC^{PAF3}MP** as yellowish powder (yield 30%) and 20% of hydrolyzed product which is the desired product for the next step.

Contains ~ 20% of hydrolyzed product (without trifluoroacetyl)

^{1}H NMR (500.0 MHz, D_2O, ref(*t*BuOH) = 1.24 ppm): 1.27 (t, 18H, J_{vic} = 7.3, C**H$_3$**CH_2N); 2.30 (dt, 1H, J_{gem} = 13.9, $J_{2'b,1'} = J_{2'b,3'}$ = 6.6, H-2′b); 2.43 (ddd, 1H, J_{gem} = 13.9, $J_{2'a,1'}$ = 6.2, $J_{2'a,3'}$ = 3.9, H-2′a); 3.18 (q, 12H, J_{vic} = 7.3, CH_3C**H$_2$**N); 3.90 – 3.96 (m, 2H, H-5′); 4.16 (q, 1H, $J_{4',5'}$ = 4.9,

$J_{4',5'}$ = 3.9, H-4′); 4.36(s, 2H, CH_2N); 4.52 (dt, 1H, $J_{3',2'}$ = 6.6, 3.9, $J_{3',4'}$ = 3.9, H-3′); 6.25 (dd, 1H, $J_{1',2'}$ = 6.6, 6.2, H-1′); 8.20 (s, 1H, H-6).

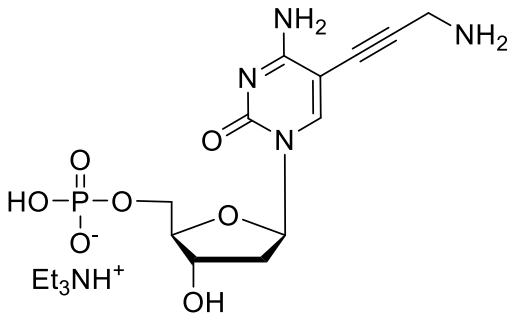
^{13}C NMR (125.7 MHz, D_2O, ref(*t*BuOH) = 32.43 ppm): 8.93 (**C**H_3CH_2N); 30.81 (CH_2N); 39.90 (CH_2-2′); 47.32 (CH_3**C**H_2N); 64.18 (d, $J_{C,P}$ = 4.5, CH_2-5′); 71.56 (CH-3′); 74.88 (**C**≡CCH_2); 86.87 (d, $J_{C,P}$ = 8.3, CH-4′); 86.96 (CH-1′); 90.67 (C≡**C**CH_2); 92.42 (C-5); 116.44 (q, $J_{C,F}$ = 286.1, **C**F_3CO); 146.02 (CH-6); 156.75 (C-2); 159.46 (q, $J_{C,F}$ = 38.6, CF_3**C**O); 165.78 (C-4).

^{31}P{^{1}H} NMR (202.4 MHz, D_2O): 4.67.

^{19}F NMR (470.4 MHz, D_2O): -72.06.

HR/MS (ESI-) for $C_{14}H_{15}O_8N_4F_3P$: calculated 455.05851 $[M-H]^-$, found 455.05884 $[M-H]^-$

Synthesis of 5-[3-(Aminoprop-1-yn-1-yl)]-2'-deoxycytidine 5-*O*-monophosphate triethylammonium salt (dC^{NH2}MP) (5)



dC^{PAF3}MP (116 mg, 255 µmol) was dissolved NH_4OH aq. (10 mL) and stirred at room temperature for 24 h. The product was purified by RP-HPLC using a Phenomenex HPLC column (Luna 10 C8 100Å, 10×250 mm). Several co-distillations with water followed by freeze-drying gave the pure product **dC^{NH2}MP** as white powder (74 mg, 81%).

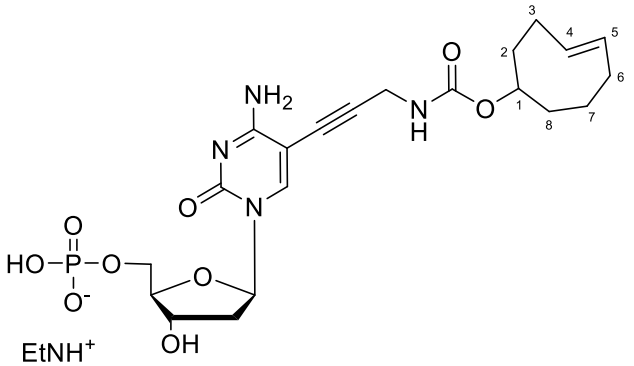
^{1}H NMR (500.0 MHz, D_2O, ref(*t*BuOH) = 1.24 ppm): 1.28 (t, 9H, J_{vic} = 7.3, C**H$_3$**CH_2N); 2.32 (dt, 1H, J_{gem} = 14.0, $J_{2'b,1'}$ = $J_{2'b,3'}$ = 6.2, H-2′b); 2.48 (ddd, 1H, J_{gem} = 14.0, $J_{2'a,1'}$ = 6.5, $J_{2'a,3'}$ = 4.7, H-2′a); 3.20 (q, 6H, J_{vic} = 7.3, CH_3C**H$_2$**N); 4.03 (d, 1H, J_{gem} = 17.0, CH_a**H$_b$**NH_2); 4.04 (ddd, 1H, J_{gem} = 11.7, $J_{H,P}$ = 4.8, $J_{5'b,4'}$ = 3.1, H-5′b); 4.06 (d, 1H, J_{gem} = 17.0, C**H$_a$**H_bNH_2); 4.14 (ddd, 1H, J_{gem} = 11.7, $J_{H,P}$ = 4.1, $J_{5'a,4'}$ = 2.6, H-5′a); 4.19 (dddd, 1H, $J_{4',3'}$ = 4.3, $J_{4',5'}$ = 3.1, 2.6, $J_{H,P}$ = 2.1, H-4′); 4.55 (ddd, 1H, $J_{3',2'}$ = 6.2, 4.7, $J_{3',4'}$ = 4.3, H-4′); 6.25 (dd, 1H, $J_{1',2'}$ = 6.5, 6.2, H-1′); 8.44 (s, 1H, H-6).

^{13}C NMR (125.7 MHz, D_2O, ref(*t*BuOH) = 32.43 ppm): 11.06 (**C**H_3CH_2N); 32.66 (CH_2NH_2); 42.89 (CH_2-2′); 49.51 (CH_3**C**H_2N); 66,62 (d, $J_{C,P}$ = 4.9, CH_2-5′); 72.71 (CH-3′); 80.71 (**C**≡CCH_2); 88.71 (d, $J_{C,P}$ = 9.0, CH-4′); 89.34 (CH-1′); 89.43 (C≡**C**CH_2); 93.84 (C-5); 149.30 (CH-6); 158.79 (C-2); 167.53 (C-4).

^{31}P NMR (202.4 MHz, D_2O): 1.57.

HR/MS (ESI-) for $C_{12}H_{16}N_4O_7P^-$: calculated 359.07621 $[M-H]^-$; found 359.07565 $[M-H]^-$.

Synthesis of 5-{[(*E*)-Cyclooct-4-en-1-yl)oxy)carbonylamino]prop-1-yl}-2'-deoxycytidine-5-*O*-triphosphate triethylammonium salt (dC^{4TCO}MP) (7)



The freeze-dried **dC^{NH2}MP** (29.2 mg, 80.7 μmol) was dissolved in 1 M TEAB buffer (pH 8, 3 mL), 4TCO-NHS-carbonate (30 mg, 112.3 μmol) in DMF (3 mL) was added to the reaction mixture. The mixture was stirred at 55 ºC for 4 h and the solvent was removed *in vacuo.* The product was purified by RP-HPLC using a Phenomenex HPLC column (Luna 10 C8 100Å, 10×250 mm). Eluent A: 0.1 M TEAB buffer (pH 7.5), B: 60% acetonitrile in 0.1 M TEAB (pH 7.5). Several co-distillations with water followed by freeze-drying gave the pure product **dC^{4TCO}MP** as yellowish powder (20 mg, 48% yield) Rf = 0.60 (IPAV).

^{1}H NMR (500.0 MHz, D_2O, ref(*t*BuOH ext) = 1.24 ppm): 1.28 (t, 9H, J_{vic} = 7.3, C**H$_3$**CH_2N); 1.51 (m, 1H, H-7b-cyclooct); 1.59 – 1.70 (bm, 2H, H-2b,7a-cyclooct); 1.73 (bm, 1H, H-8b-cyclooct); 1.84 – 1.93 (bm, 2H, H-2a,8a-cyclooct); 2.09 (m, 1H, H-3b-cyclooct); 2.13 – 2.22 (m, 2H, H-6-cyclooct); 2.29 (ddd, 1H, J_{gem} = 14.2, $J_{2'b,1'}$ = 6.8, $J_{2'b,3'}$ = 6.3, H-2′b); 2.37 (bm, 1H, H-3a-cyclooct); 2.46 (ddd, 1H, J_{gem} = 14.2, $J_{2'a,1'}$ = 6.4, $J_{2'a,3'}$ = 3.9, H-2′a); 3.20 (q, 6H, J_{vic} = 7.3, CH_3C**H$_2$**N); 4.05 (ddd, 1H, J_{gem} = 11.6, $J_{H,P}$ = 5.7, $J_{5'b,4'}$ = 4.1, H-5′b); 4.09 (ddd, 1H, J_{gem} = 11.6, $J_{H,P}$ = 4.9, $J_{5'a,4'}$ = 3.5, H-5′a); 4.12 (bs, 2H, CH_2N); 4.20 (dtd, 1H, $J_{4',5'}$ = 4.1, 3.5, $J_{4',3'}$ = 3.5, $J_{H,P}$ = 1.7, H-4′); 4.52

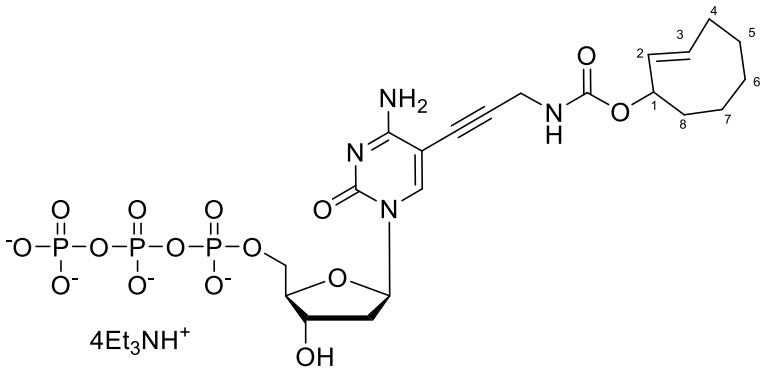
(ddd, 1H, $J_{3',2'}$ = 6.3, 3.9, $J_{3',4'}$ = 3.5, H-3′); 4.71 (m, 1H, H-1-cyclooct, overlapped with water signal); 5.65 – 5.78 (bm, 2H, H-4,5-cyclooct); 6.25 (dd, 1H, $J_{1',2'}$ = 6.8, 6.4, H-1′); 8.15 (s, 1H, H-6).

^{13}C NMR (125.7 MHz, D_2O, ref(*t*BuOH ext) = 32.43 ppm): 11.06 (**C**H_3CH_2N); 24.41 (CH_2-3-cyclooct); 27.21 (CH_2-7-cyclooct); 27.89 (CH_2-6-cyclooct); 33.64 (b, CH_2N); 36.05 (b, CH_2-8-cyclooct); 36.19 (CH_2-2-cyclooct); 42.46 (CH_2-2′); 49.51 (CH_3**C**H_2N); 67.29 (d, $J_{C,P}$ = 4.9, CH_2-5′); 73.50 (CH-3′); 75.75 (**C**≡CCH_2); 80.68 (CH-1-cyclooct); 88.53 (d, $J_{C,P}$ = 8.6, CH-4′); 89.39 (CH-1′); 94.96 (C-5); 95.41 (b, C≡**C**CH_2); 132.86 (CH-4-cyclooct); 133.12 (CH-5-cyclooct); 147.66 (CH-6); 158.62 (C-2); 161.03 (b, NHCOO); 167.78 (C-4).

^{31}P NMR (202.4 MHz, D_2O): 0.32.

HR/MS (ESI-) for $C_{21}H_{28}N_4O_9P^-$: calculated 511.17 $[M-H]^-$; found 511.2 $[M-H]^-$.

Synthesis of 5-{[(*E*)-cyclooct-2-en-1-yl)oxy)carbonylamino]prop-1-yl}-2'-deoxycytidine-5-*O*-triphosphate triethylammonium salt (dC^{2TCO}TP) (11)



The freeze-dried **dC^{NH2}TP** [110] (5 mg, 7.44 µmol) was dissolved in 1 M TEAB buffer (pH 8, 500 µL), 2TCO-NHS-carbonate (6.25 g, 24.9 µmol) in DMF (416 µL) was added to the reaction mixture. The mixture was stirred at 55 ºC for 4 h and the solvent was removed *in vacuo.* The product was purified by RP-HPLC using a Phenomenex HPLC column (Luna 10 C8 100Å, 10×250 mm). Eluent A: 0.1 M TEAB buffer (pH 7.5), B: 60% acetonitrile in 0.1 M TEAB (pH 7.5). Gradient: 30% A - 70% buffer B in 19 min. Several co-distillations with water followed by freeze-drying gave the pure product **dC^{2TCO}TP** as yellowish powder (1.9 mg, 29% yield), Rf = 0.74 (IPAV).

Mixture of diastereoisomers ~3:2 (observable on signal of H-1′, C-1′,2′,3′)

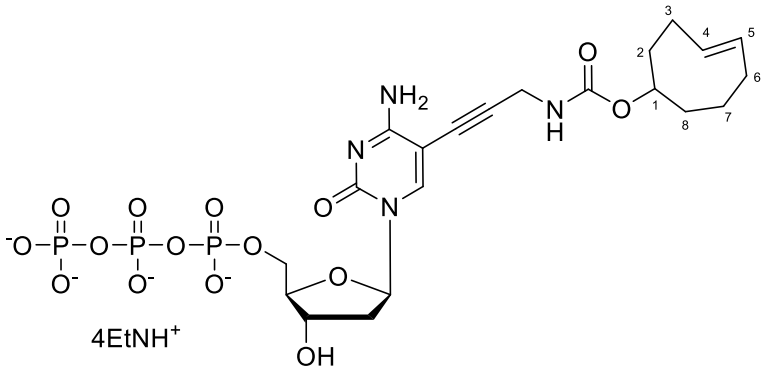
^{1}H NMR (500.0 MHz, D_2O, ref(*t*BuOH) = 1.24 ppm): 0.81 (m, 1H, H-6b-cyclooct); 1.09 (bm, 1H, H-7b-cyclooct); 1.28 (t, 36H, J_{vic} = 7.3, C**H**$_3$CH_2N); 1.49 (bm, 1H, H-5b-cyclooct); 1.65 (bm, 1H, H-7a-cyclooct); 1.73 – 1.87 (m, 2H, H-6a,8b-cyclooct); 1.92 (bm, 1H, H-5a-cyclooct); 1.96 – 2.10 (m, 2H, H-4b,8a-cyclooct); 2.31 (dt, 1H, J_{gem} = 14.3, $J_{2'b,1'} = J_{2'b,3'}$ = 6.6, H-2′b); 2.41 – 2.49 (m, 2H, H-2′a, H-4a-cyclooct); 3.21 (q, 24H, J_{vic} = 7.3, CH_3C**H**$_2$N); 4.15 (bs, 2H, CH_2N); 4.18 – 4.26 (m, 3H, H-4′,5′); 4.61 (m, 1H, H-3′); 5.21 (bs, 1H, H-1-cyclooct); 5.64 (dd, 1H, $J_{2,3}$ = 16.6, $J_{2,1}$ = 1,7, H-2-cyclooct); 5.89 (bm, 1H, H-3-cyclooct); 6.27, 6.28 (t, 1H, $J_{1',2'}$ = 6.6, H-1′); 8.17 (s, 1H, H-6).

^{13}C NMR (125.7 MHz, D_2O, ref(*t*BuOH) = 32.43 ppm): 11.06 (**C**H_3CH_2N); 26.38 (CH_2-7-cyclooct); 31.17 (CH_2-6-cyclooct); 33.68 (b, CH_2N); 38.21 (CH_2-4,5-cyclooct); 42.17, 42.22 (CH_2-2′); 42.64 (CH_2-8-cyclooct); 49.51 (CH_3**C**H_2N); 68.01 (b, CH_2-5′); 73.17, 73.22 (CH-3′); 75.81 (b, **C**≡CCH_2); 78.50 (b, CH-1-cyclooct); 88.36 (d, $J_{C,P}$ = 8.8, CH-4′); 89.09, 89.12 (CH-1′); 95.02 (b, C-5); 95.42 (b, C≡**C**CH_2); 133.98 (CH-2-cyclooct); 135.36 (b, CH-3-cyclooct); 147.61 (CH-6); 158.90 (C-2); 160.14 (b, NHCOO); 167.94 (C-4).

^{31}P NMR (202.4 MHz, D_2O): -22.55 (bdd, *J* = 21.2, 19.9, P$_\beta$); -11.31 (d, *J* = 19.9, P$_\alpha$); -6.37 (d, *J* = 21.2, P$_\gamma$).

HR/MS (ESI-) for $C_{21}H_{30}O_{15}N_4P_3$: calculated 671.09260 $[M-H]^-$, found 671.09230 $[M-H]^-$.

Synthesis of 5-{[(E)-cyclooct-4-en-1-yl)oxy)carbonylamino]prop-1-yl}-2'-deoxycytidine-5-O-triphosphate triethylammonium salt (dC^{4TCO}TP) (12)



The freeze-dried **dC^{NH2}TP** [110] (3.7 mg, 7.44 µmol) was dissolved in 1 M TEAB buffer (pH 8, 370 µL), 4TCO-NHS-carbonate (4.62 mg, 18.40 µmol) in DMF (307 µL) was added to the reaction mixture. The mixture was stirred at 55 ºC for 4 h and the solvent was removed *in vacuo.* The product was purified by RP-HPLC using a Phenomenex HPLC column (Luna 10 C8 100Å, 10×250 mm). Eluent A: 0.1 M TEAB buffer (pH 7.5), B: 60% acetonitrile in 0.1 M TEAB (pH 7.5). Gradient: 30% A - 70% B in 20 min. Several co-distillations with water followed by freeze-drying gave the pure product **dC^{4TCO}TP** as yellowish powder (1.7 mg, 35% yield) Rf = 0.72 (IPAV).

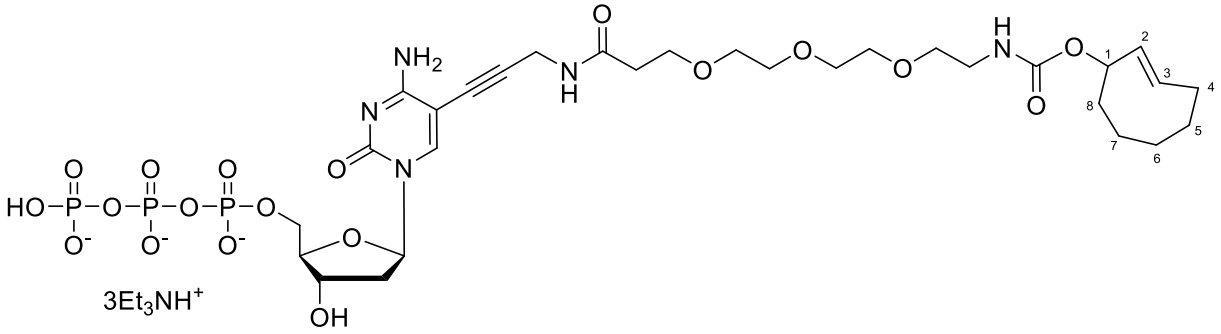
^{1}H NMR (500.0 MHz, D_2O, ref(*t*BuOH) = 1.24 ppm): 1.28 (t, 18H, J_{vic} = 7.3, C**H$_3$**CH$_2$N); 1.51 (bm, 1H, H-7b-cyclooct); 1.58 – 1.70 (bm, 2H, H-2b,7a-cyclooct); 1.74 (bm, 1H, H-8b-cyclooct); 1.84 – 1.95 (bm, 2H, H-2a,8a-cyclooct); 2.09 (bm, 1H, H-3b-cyclooct); 2.13 – 2.23 (bm, 2H, H-6-cyclooct); 2.31 (dt, 1H, J_{gem} = 14.1, $J_{2'b,1'} = J_{2'b,3'}$ = 6.5, H-2′b); 2.37 (bm, 1H, H-3a-cyclooct); 2.45 (ddd, 1H, J_{gem} = 14.1, $J_{2'a,1'}$ = 6.5, $J_{2'a,3'}$ = 4.0, H-2′a); 3.20 (q, 12H, J_{vic} = 7.3, CH$_3$C**H$_2$**N); 4.13 (bs, 2H, CH$_2$N); 4.19 – 4.23 (m, 3H, H-4′,5′); 4.60 (m, 1H, H-3′); 4.71 (m, 1H, H-1-cyclooct, partially overlapped with water signal); 5.66 – 5.78 (bm, 2H, H-4,5-cyclooct); 6.25 (t, 1H, $J_{1',2'}$ = 6.5, H-1′); 8.16 (s, 1H, H-6).

^{13}C NMR (125.7 MHz, D_2O, ref(*t*BuOH) = 32.43 ppm): 11.05 (**C**H$_3$CH$_2$N); 24.41 (CH$_2$-3-cyclooct); 27.22 (CH$_2$-7-cyclooct); 27.90 (CH$_2$-6-cyclooct); 33.69 (b, CH$_2$N); 36.17 (CH$_2$-2,8-cyclooct); 42.22 (CH$_2$-2′); 49.49 (CH$_3$**C**H$_2$N); 68.04 (d, $J_{C,P}$ = 5.5, CH$_2$-5′); 73.25 (CH-3′); 75.55 (**C**≡CCH$_2$); 80.67 (CH-1-cyclooct); 88.41 (d, $J_{C,P}$ = 8.9, CH-4′); 89.26 (CH-1′); 95.01 (C-5); 95.63 (C≡**C**CH$_2$); 132.86, 133.12 (CH-4,5-cyclooct); 147.83 (CH-6); 158.16 (C-2); 161.04 (NHCOO); 167.46(C-4).

^{31}P NMR (202.4 MHz, D_2O): -23.15 (bdd, J = 18.9, 16.0, P□); -11.47 (bd, J = 18.9, P□); -10.84 (d, J = 16.0 P□).

HR/MS (ESI-) for C21H30O15N4P3: calculated 671.09260 [M-H]$^-$, found 671.09163 [M-H]$^-$.

Synthesis of 5-{[1-((*E*)-cyclooct-2-en-1-yl)oxy]-1,14-dioxo-5,8,11-trioxa-2,15-diazaoctadec-17-yn-18-yl}-2'-deoxycytidine-5-*O*-triphosphate triethylammonium salt (dC^{p2TCO}TP) (13)



The freeze-dried **dC^{NH2}TP** [110] (7 mg, 13.46 µmol) was dissolved in 1 M TEAB buffer (pH 8, 560 µL), 2TCO-PEG3-NHS ester (9.07 mg, 21.59 µmol) in DMF (520 µL) was added to the reaction mixture. The mixture was stirred at 55 ºC for 4 h and the solvent was removed *in vacuo.* The products were purified by RP-HPLC using a Phenomenex HPLC column (Luna 10 C8 100Å, 10×250 mm). Eluent A: 0.1 M TEAB buffer (pH 7.5), B: 60% acetonitrile in 0.1 M TEAB (pH 7.5). 30% A-70% B in 32 min. Several co-distillations with water followed by freeze-drying gave the pure product **dC^{p2TCO}TP** as yellowish powder (4.15 mg, 35%), Rf = 0.68 (IPAV).

^{1}H NMR (500.0 MHz, D_2O, ref(*t*BuOH) = 1.24 ppm): 0.78 (bm, 1H, H-6b-cyclooct); 1.04 (bm, 1H, H-7b-cyclooct); 1.28 (t, 27H, J_{vic} = 7.3, C**H$_3$**CH$_2$N); 1.46 (bm, 1H, H-5b-cyclooct); 1.63 (bm, 1H, H-7a-cyclooct); 1.75 (bm, 1H, H-8b-cyclooct); 1.82 (bm, 1H, H-6a-cyclooct); 1.93 (bm, 1H, H-5a-cyclooct); 1.96 – 2.06 (m, 2H, H-4b,8a-cyclooct); 2.30 (dt, 1H, J_{gem} = 14.1, $J_{2'b,1'} = J_{2'b,3'}$ = 6.5, H-2′b); 2.41 (m, 1H, H-4a-cyclooct); 2.45 (ddd, 1H, J_{gem} = 14.1, $J_{2'a,1'}$ = 6.5, $J_{2'a,3'}$ = 4.1, H-2′a); 2.53 – 2.63 (m, 2H, COC**H$_2$**CH$_2$O); 3.20 (q, 18H, J_{vic} = 7.3, CH$_3$C**H$_2$**N); 3.29 – 3.34 (bm, 2H, NHCOC**H$_2$**CH$_2$O); 3.59 (bt, 2H, J_{vic} = 5.3, NHCOCH$_2$C**H$_2$**O); 3.64 – 3.70 (m, 8H, OC**H$_2$**C**H$_2$**O); 3.78 – 3.85 (m, 2H, COCH$_2$C**H$_2$**O); 4.16 – 4.27 (m, 5H, H-4′,5′, CH$_2$N); 4.63 (dd, 1H, $J_{3',2'}$ = 6.5, 4.1, $J_{3',4'}$ = 3.1, H-3′); 5.15 (bm, 1H, H-1-cyclooct); 5.62 (dd, 1H, $J_{2,3}$ = 16.5, $J_{2,1}$ = 2.0, H-2-cyclooct); 5.87 (m, 1H, H-3-cyclooct); 6.26 (t, 1H, $J_{1',2'}$ = 6.5, H-1′); 8.19 (s, 1H, H-6).

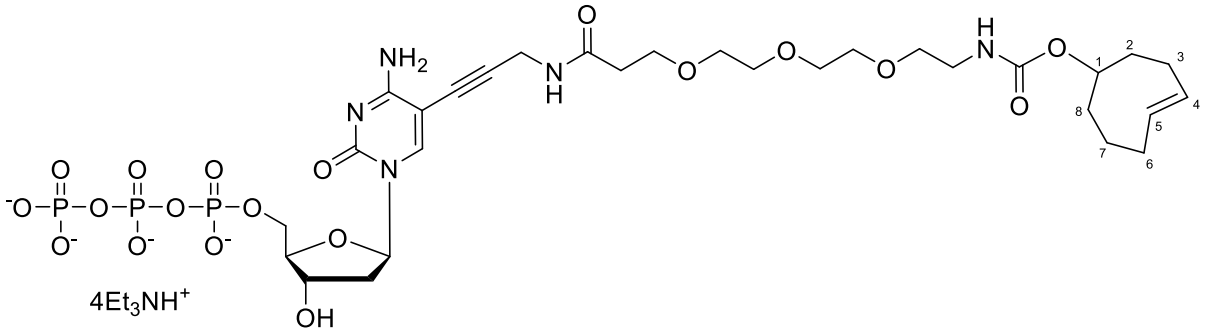
^{13}C NMR (125.7 MHz, D_2O, ref(*t*BuOH) = 30.29 ppm): 8.92 (**C**H$_3$CH$_2$N); 24.27 (CH$_2$-7-cyclooct); 29.04 (CH$_2$-6-cyclooct); 30.41 (b, CH$_2$N); 36.09, 36.21 (CH$_2$-4,5-cyclooct); 36.64 (CO**C**H$_2$CH$_2$O); 40.15 (CH$_2$-2′); 40.54 (CH$_2$-8-cyclooct); 40.67 (N**C**H$_2$CH$_2$O); 47.35 (CH$_3$**C**H$_2$N); 65.84 (d, $J_{C,P}$ = 5.6, CH$_2$-5′); 67.39 (COCH$_2$**C**H$_2$O); 69.99, 70.13, 70.26, 70.32, 70.35

(O**C**H_2**C**H_2O, NCH_2**C**H_2O); 70.99 (CH-3′); 73.77 (**C**≡CCH_2); 75.83 (CH-1-cyclooct); 86.24 (d, $J_{C,P}$ = 9.0, CH-4′); 86.98 (CH-1′); 92.57 (C≡**C**CH_2); 92.88 (C-5); 132.06 (CH-2-cyclooct); 132.98 (CH-3-cyclooct); 145.63 (CH-6); 156.62 (C-2); 158.92 (NHCOO); 165.72 (C-4); 174.83 (CONH).

^{31}P NMR (202.4 MHz, D_2O): -23.25 (bdd, J = 20.1, 19.5, P$_\beta$); -11.56 (d, J = 20.1, P$_\alpha$); -10.75 (d, J = 19.5, P$_\gamma$).

HR/MS (ESI-) for $C_{30}H_{48}N_5O_{19}P_3$: calculated 874.20836 [M-H]$^-$, found 874.20715 [M-H]$^-$.

Synthesis of 5-{[1-((*E*)-Cyclooct-4-en-1-yl)oxy]-1,14-dioxo-5,8,11-trioxa-2,15-diazaoctadec-17-yn-18-yl}-2'-deoxycytidine-5-*O*-triphosphate triethylammonium salt (dC^{p4TCO}TP) (14)



The freeze-dried **dC^{NH2}TP** [110] (4.4 mg, 8.46 µmol) was dissolved in 1 M TEAB buffer (pH 8, 352 µL), 4TCO-PEG3-NHS ester (5.71 mg, 12.13 µmol) in DMF (285 µL) was added to the reaction mixture.

The mixture was stirred at 55 ºC for 4 h and the solvent was removed *in vacuo.* The product was purified by RP-HPLC using a Phenomenex HPLC column (Luna 10 C8 100Å, 10×250 mm). Eluent A: 0.1 M TEAB buffer (pH 7.5), B: 60% acetonitrile in 0.1 M TEAB (pH 7.5). Gradient: 30% A-70% B in 32 min. Several co-distillations with water followed by freeze-drying gave the pure product **dC^{p4TCO}TP** as yellowish powder (2.70 mg 37% yield), Rf = 0.70 (IPAV).

^{1}H NMR (600.1 MHz, D_2O, ref(*t*BuOH) = 1.24 ppm): 1.24 (t, 36H, J_{vic} = 7.3, C**H$_3$**CH_2N); 1.29 (bm, 1H, H-8b-cyclooct); 1.43 (bm, 1H, H-7b-cyclooct); 1.73 (bm, 1H, H-2b-cyclooct); 1.78 (bm, 1H, H-7a-cyclooct); 1.86 (bm, 1H, H-6b-cyclooct); 2.10 (bm, 1H, H-3b-cyclooct); 2.17 – 2.26 (m, 3H, H-2a,8a,6a-cyclooct); 2.30 (dt, 1H, J_{gem} = 14.0, $J_{2'b,1'} = J_{2'b,3'}$ = 6.5, H-2′b); 2.31 (m, 1H, H-3a-

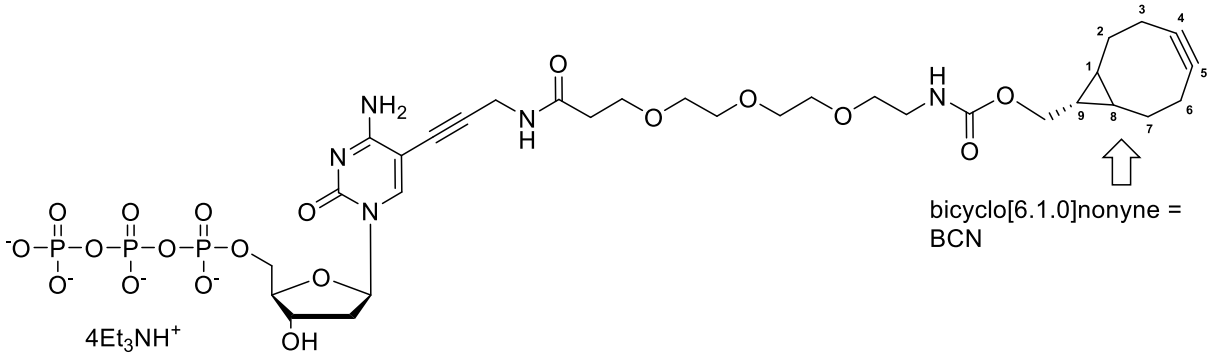
cyclooct); 2.44 (ddd, 1H, J_{gem} = 14.0, $J_{2'a,1'}$ = 6.5, $J_{2'a,3'}$ = 4.7, H-2′a); 2.54 – 2.63 (m, 2H, COC**H₂**CH₂O); 3.09 (q, 24H, J_{vic} = 7.3, CH₃C**H₂**N); 3.27 – 3.36 (bm, 2H, NHCOC**H₂**CH₂O); 3.57 – 3.62 (bm, 2H, NHCOCH₂C**H₂**O); 3.63 – 3.70 (m, 8H, OC**H₂**C**H₂**O); 3.77 – 3.87 (m, 2H, COCH₂C**H₂**O); 4.17 – 4.30 (m, 5H, H-4′,5′, CH₂N); 4.64 (dt, 1H, $J_{3',2'}$ = 6.5, 4.7, $J_{3',4'}$ = 4.7, H-3′); 4.76 (m, 1H, H-1-cyclooct); 5.60 – 5.69 (bm, 2H, H-4,5-cyclooct); 6.25 (t, 1H, $J_{1',2'}$ = 6.5, H-1′); 8.21 (s, 1H, H-6).

^{13}C NMR (150.9 MHz, D₂O, ref(*t*BuOH) = 30.29 ppm): 8.92 (**C**H₃CH₂N); 27.95 (CH₂-7-cyclooct); 29.93 (CH₂-3-cyclooct); 30.11 (CH₂N); 32.46 (CH₂-8-cyclooct); 34.18 (CH₂-6-cyclooct); 36.35 (CO**C**H₂CH₂O); 39.71 (CH₂-2′); 40.38 (N**C**H₂CH₂O); 40.67 (CH₂-2-cyclooct); 46.85 (CH₃**C**H₂N); 65.18 (d, $J_{C,P}$ = 5.6, CH₂-5′); 67.10 (COCH₂**C**H₂O); 69.80 (NCH₂**C**H₂O); 69.87, 69.98 (O**C**H₂**C**H₂O); 70.02 (CH-1-cyclooct); 70.04, 70.09 (O**C**H₂**C**H₂O); 71.74 (CH-3′); 73.54 (**C**≡CCH₂); 86.01 (d, $J_{C,P}$ = 9.0, CH-4′); 86.48 (CH-1′); 92.29 (C≡**C**CH₂); 92.55 (C-5); 132.85, 136.23 (CH-4,5-cyclooct); 145.42 (CH-6); 156.39 (C-2); 158.924 (NHCOO); 165.49 (C-4); 174.53 (CONH).

^{31}P NMR (202.4 MHz, D₂O): -22.59 (bdd, J = 21.2, 20.0, P$_\beta$); -11.40 (d, J = 20.0, P$_\alpha$); -6.37 (d, J = 21.2, P$_\gamma$).

HR/MS (ESI-) for $C_{30}H_{48}N_5O19P_3$: calculated 874.20836 [M-H]$^-$, found 874.20807 [M-H]$^-$.

Synthesis of 5-{[(bicyclo[6.1.0]non-4-yn-9-yl)-3,16-dioxo-2,7,10,13-tetraoxa-4,17-diazaicos-19-yn-20-yl)-2'-deoxycytidine-5-*O*-triphosphate triethylammonium salt (dC^{pBCN}TP) (15)



The freeze-dried **dC^{NH2}TP** [110] (4.6 mg, 2.29 µmol) was dissolved in 0.5 M TEAB buffer (pH 7.5, 319 µL), endo-BCN-PEG3-NHS ester (6 mg, 3.81 µmol) in DMF (258 µL) was added to the reaction mixture. The mixture was stirred at 55 ºC for 4 h and the solvent was removed *in vacuo.*

The products were purified by RP-HPLC using a Phenomenex HPLC column (Luna 10 C8 100Å, 10×250 mm). Eluent A: 0.1 M TEAB buffer (pH 7.5), B: 60% acetonitrile in 0.1 M TEAB (pH 7.5). 30% A-70% B in 32 min. Several co-distillations with water followed by freeze-drying gave the pure product **dC^{pBCN}TP** as yellowish powder (1.36 mg, 20%), Rf = 0.68 (IPAV).

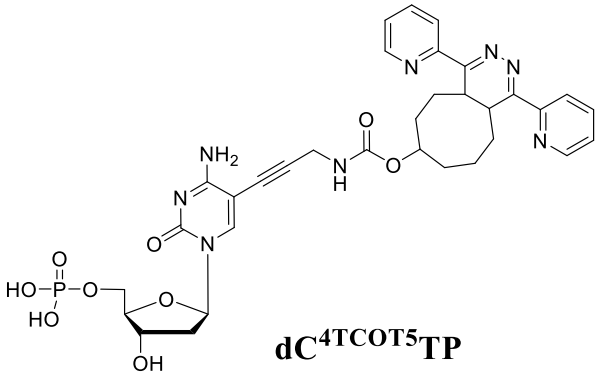
^{1}H NMR (600.1 MHz, D_2O, ref(*t*BuOH) = 1.24 ppm): 0.92 – 0.98 (m, 2H, H-1,8-BCN); 1.25 (t, 36H, J_{vic} = 7.3, C**H$_3$**CH$_2$N); 1.37 (m, 1H, H-9-BCN); 1.49 – 1.60 (bm, 2H, H-2b,7b-BCN); 2.15 – 2.29 (bm, 6H, H2a,3,6,7a-BCN); 2.31 (dt, 1H, J_{gem} = 14.0, $J_{2'b,1'} = J_{2'b,3'}$ = 6.5, H-2′b); 2.44 (ddd, 1H, J_{gem} = 14.0, $J_{2'a,1'}$ = 6.5, $J_{2'a,3'}$ = 4.6, H-2′a); 2.54 – 2.63 (m, 2H, COC**H$_2$**CH$_2$O); 3.12 (q, 24H, J_{vic} = 7.3, CH$_3$C**H$_2$**N); 3.29 – 3.33 (bm, 2H, NHCOC**H$_2$**CH$_2$O); 3.57 – 3.60 (bm, 2H, NHCOCH$_2$C**H$_2$**O); 3.64 – 3.69 (m, 8H, OC**H$_2$**C**H$_2$**O); 3.78 – 3.85 (m, 2H, COCH$_2$C**H$_2$**O); 4.14 – 4.24 (m, 6H, H-4′,5′b, CH$_2$N, CH$_2$O); 4.27 (ddd, 1H, J_{gem} = 11.5, $J_{H,P}$ = 6.3, $J_{5'a,4'}$ = 3.8, H-5′a); 4.63 (dt, 1H, $J_{3',2'}$ = 6.5, 4.6, $J_{3',4'}$ = 4.6, H-3′); 6.24 (t, 1H, $J_{1',2'}$ = 6.5, H-1′); 8.21 (s, 1H, H-6).

^{13}C NMR (150.9 MHz, D_2O, ref(*t*BuOH) = 30.29 ppm): 9.11 (**C**H$_3$CH$_2$N); 17.88 (CH-9-BCN); 20.41 (CH-1,8-BCN); 21.43 (CH$_2$-3,6-BCN); 29.24 (CH$_2$-2,7-BCN); 30.38 (CH$_2$N); 36.65 (CO**C**H$_2$CH$_2$O); 40.03 (CH$_2$-2′); 40.80 (N**C**H$_2$CH$_2$O); 47.22 (CH$_3$**C**H$_2$N); 64.46 (CH$_2$O); 66.47 (d, $J_{C,P}$ = 5.4, CH$_2$-5′); 67.39 (COCH$_2$**C**H$_2$O); 69.99 (NCH$_2$**C**H$_2$O); 70.12, 70.24 (O**C**H$_2$**C**H$_2$O); 70.31 (CH-3′); 70.33 (O**C**H$_2$**C**H$_2$O); 73.85 (**C**≡CCH$_2$); 86.28 (d, $J_{C,P}$ = 9.0, CH-4′); 86.81 (CH-1′); 92.59 (C≡**C**CH$_2$); 92.81 (C-5); 100.92 (C-4,5-BCN); 145.73 (CH-6); 156.68 (C-2); 159.51 (NHCOO); 165.78 (C-4); 174.77 (CONH).

^{31}P NMR (202.4 MHz, D_2O): -22.42 (dd, *J* = 21.0, 19,9, P$_\beta$); -11.38 (d, *J* = 19,9, P$_\alpha$); -6.33 (d, *J* = 21.0, P$_\gamma$).

HR/MS (ESI-) for $C_{32}H_{48}N5O_{19}P_3$: calculated 898.20836 $[M-H]^-$, found 898.20729 $[M-H]^-$.

5.1.2 IEDDA model reaction of dC^{4TCO}MP with T5 (3,6-di-2-pyridyl)-tetrazine (16)



dC^{4TCOT5}TP

The freeze-dried **dC^{4TCO}MP** (20 mg, 39 µmol) was dissolved in water (3 mL).
3,6-di-2-pyridyl-1,2,4,5-tetrazine or **T5** tetrazine (30 mg, 127 µmol) was dissolved in 2 mL ACN/H_2O (3:1). The mixture was stirred at 37 ºC for 2 h and the solvent was removed in vacuo. The product was purified by RP-HPLC using a Phenomenex HPLC column (Luna 10 C8 100Å, 10×250 mm). Eluent A: 0.1 M TEAB buffer (pH 7.5), B: 60% acetonitrile in 0.1 M TEAB (pH 7.5). The complicated chromatogram gave mixture of diasteroisomers of the desired product (3:7, DCM/MeOH).
MS and NMR confirmed formation of the desired triphosphate-tetrazine conjugate **dC^{4TCOT5}MP**, however, complete assignment of ^{1}H and ^{13}C NMR spectra was complicated by several factors. First, concentration of the sample was not sufficient for direct observation of ^{13}C NMR spectrum. Instead, 2D correlation experiments such as H,C-HSQC and H,C-HMBC were used to identify structural fragments of the product in 1H NMR spectrum, such as 2-deoxyribose, nucleobase, linker, cyclooctane ring and pyridyl moiety (see **Figure 25**, Section 3.1.1.4.2). Second, based on reaction mechanism, click reaction of substituted trans-cyclooctene with tetrazine provides a mixture of several isomers that contributes to complexity of NMR spectra as well.

5.1.3 IEDDA click reactions of dC^{4TCO}TP or dC^{2TCO}TP with T1 or T2 tetrazines and of dC^{p4TCO}TP or dC^{p2TCO}TP or dC^{pBCN}TP with T1, T2 or T3 tetrazines (dC^{XTY}TP)

Procedure: Freshly prepared stock solution of **T1** tetrazine (1 mM in DMSO) was used for the measurements. The turn-on fluorescence measurements were performed as follows: 1 µL of the tetrazine stock solution was diluted in 1 mL PBS buffer (1 X, pH 7.4, giving 1 µM final

concentration) and the fluorescence was measured using 370 nm excitation (slit 3) and 400-650 nm for emission.

To this solution was added solution of the triphosphates (**dC^{4TCO}TP** or **dC^{2TCO}TP** or **dC^{p4TCO}TP** or **dC^{p2TCO}TP** or **dC^{pBCN}TP)**: 1 µL from 1 mM fresh stock in water, giving 1 µM final concentration, 1 equiv. to tetrazine). The cuvette was inserted into the fluorescence spectrophotometer and the measurement was started. The spectra were recorded at several time points and all measurements were typically repeated two-three times. The final measurement was performed after ca. 30 minutes. The data were processed using Origin software. All spectra were subtracted from the baseline (PBS buffer as the blank). The turn-on values were calculated from the observed fluorescence intensities of the click products at the emission maximum (485 nm) after ca. 15 min of reaction divided by the highest residual fluorescence of the quenched coumarin tetrazine dye. These values are presented for **dC^{4TCO}TP**, **dC^{2TCO}TP**, **dC^{p4TCO}TP**, **dC^{p2TCO}TP** in **Figure 86-A,B,C,D** (Section 7.1, Appendix) and for **dC^{pBCN}TP** in **Figure 87** (Section 7.1, Appendix) and **Table 2** (Section 3.1.1.5). The highest fluorescence signal intensity was recorded for reaction with the axial **dC^{p4TCO}TP** as listed in **Table 2** (Section 3.1.1.5). Fluorescence time lapse measurements showing changes in the fluorescence signal of the click products over 30 min for **T1**, shown in **Figure 88** (Section 7.1, Appendix).

Due to the lower reactivity of the **T2** tetrazine, the fluorescence measurements were performed as follows:

Procedure: 1 µL of 1 mM solution of **dC^{4TCO}TP**, **dC^{2TCO}TP**, **dC^{pBCN}TP**, **dC^{p4TCO}TP** or **dC^{p2TCO}TP** were combined with 1 µL of 1 mM solution of **T2** tetrazine in 8 µL PBS and the reaction mixtures were incubated in the dark at room temperature for 1 hour. Then, the reaction mixtures were diluted to 1 mL with PBS to give final 1 µM solution and the fluorescence was measured using 370 nm excitation wavelength and 400-650 nm emission window (slit 3). The values are presented in **Figures 89-91** (Section 7.1, Appendix) and **Table 2** (Section 3.1.1.5).

Due to the lower reactivity of the **T3** tetrazine, the fluorescence measurements were performed as follows:

Procedure: 1 uL of 1 mM solution of **dC^{pBCN}TP** or **dC^{p4TCO}TP** were combined with 1 uL of 1 mM solution of **T3** tetrazine and the reaction mixtures were incubated in the dark at room

temperature for 2 hours. Then, the reaction mixtures were diluted to 1 mL with PBS and the fluorescence was measured using 500 nm excitation wavelength and 525-750 nm emission window (slit 3). The values are presented in **Table 2** (Section 3.1.1.5) and shown in **Figure 92**.

5.1.4 Preparation of nucleotide-tetrazine conjugates $dC^{XTY}TP$ for LC/MS characterization

$dC^{X}TP$ (**X= -4TCO** or **-2TCO** or **-p4TCO** or **-p2TCO** or **-pBCN**) dissolved in water (10 μM, 5 μL) was mixed with excess of **T1** or **T2** or **T3** tetrazines dissolved in DMSO (100 μM, 5 μL). The mixture containing **T1** or **T2** was left stirring at 37 °C for 30 min. The mixture containing **T3** was left stirring at 37 °C for 18h. The identity of the **$dC^{X}TP$** controls and **$dC^{XTY}TP$** click products was verified by analytical LC/MS and by high resolution MS. From MS spectra addition of water to the click product (**$dC^{XTY}TP$**) was observed. Estimated formation (%) of **$dC^{XTY}TP$** click products are shown in **Table 1**, Section 3.1.1.4.4

5.1.5 Enzymatic synthesis of TCO-, BCN-modified DNA by PEX

General remarks for the biochemical part

All gels were analysed by fluorescence imaging using Typhoon FLA 9500 (GE Healthcare). Mass spectra of short DNAs were measured by UltrafleXtreme MALDI-TOF/TOF (Bruker) mass spectrometer with 1 kHz smartbeam II laser technology or on UPLC Agilent 1290 Infinity II bio system equipped with DAD and LC/MSD XT MS (operating in negative mode) detectors, with bioZen™ Oligo LC column (1.7 μm, 50×2.1 mm; Phenomenex®) using a linear gradient of 12 mM Et_3N and 300 mM HFIP in H_2O to 12 mM Et_3N and 300 mM HFIP in MeOH in 30 min. Mass spectra were deconvoluted using OpenLAB CDS CS LC/MS Deconvolution Bioanalysis software.[119] The matrix consisted of 3-hydroxypicolinic acid (HPA)/ picolinic acid (PA)/ ammonium tartrate in ratio 9/ 1/ 1. UV-Vis spectra were measured at room temperature on NanoDrop1000 (ThermoScientific). Fluorescence was measured on a Fluoromax 4 spectrofluorimeter (HORIBA Scientific). Samples were concentrated on CentriVap Vacuum Concentrator system (Labconco). Synthetic oligonucleotides (primers, templates and biotinylated templates; for sequences see Table S1) were purchased from Generi Biotech (Czech Republic).

Natural nucleoside triphosphates (dATP, dGTP, dTTP, dCTP) were purchased from Thermo Scientific. KOD XL DNA polymerase and corresponding polymerase reaction buffer from Merck Life Sciences, streptavidin magnetic particles were obtained from Sigma Aldrich (Merck), QIAquick® Nucleotide Removal Kit QIAGEN (Biotech, Czech Republic). Milli-Q water was used for all experiments. PAGE stop solution used after PEX reactions contains: 95% [v/v] formamide, 0.5 mM EDTA, 0.025% [w/v] bromophenol blue, 0.025% [w/v] xylene cyanol, 0.025% SDS in MilliQ water. Samples after PEX reaction were analyzed by 12.5% PAGE (acrylamide/bisacrylamide 19:1, 25% urea) under denaturing conditions (50 min, 50 °C, 1 X TBE buffer) or by 12.5% or 20 % native PAGE (acrylamide/bisacrylamide 19:1, no urea) under following conditions of 3 h, 10 mA, RT, 1 X TBE buffer. 6X Loading buffer for native PAGE contains: 40% [w/v] saccharose, 0.02% [w/v] bromophenol blue, 0.02% [w/v] xylene cyanol in Milli-Q water. Conversions were calculated using Image J. analyzer. Tetrazine-coumarine conjugates **T1** and **T2** and PINK tetrazine **T3** were synthetized by Dr. Milan Vrabel`s group according to published methods.[46] Pyrimidyl-Tetrazine-5-TAMRA **T4** was purchased from (Jena Bioscience (CLK-097). Other chemicals were of analytical grade. Samples after conjugation with Histones were analyzed by 17.5% SDS denaturing PAGE (acrylamide/methylenebisacrylamide 29:1; 1.92 M glycine, 0.25 M Tris, 0.1% SDS- 230 V/1 h). Other chemicals were of analytical grade.

Single incorporation of $dC^{4TCO}TP$ or $dC^{2TCO}TP$ or $dC^{p4TCO}TP$ or $dC^{p2TCO}TP$ or $dC^{pBCN}TP$ using 19-mer template-Analytical scale

The reaction mixture (20 μL) contained primer **Prim A_FAM** (3 μM, 1.5 μL), template **Temp19_1C** (3 μM, 1 μL), KOD XL DNA polymerase (0.25 U/μL, 0.11 μL), natural dGTP (4 mM, 0.5 μL), either natural or modified dCTP (4 mM, 0.3 μL) in enzyme reaction buffer (10 X, 2.05 μL) and water 14.65 μL. The reaction mixture was incubated for 20 min at 60 °C in a thermal cycler. The PEX reaction was stopped by addition of PAGE stop solution (20 μL) and heated for 3 min at 95 °C. Samples were separated with 12.5% denaturing PAGE and visualized using fluorescence imaging (**Figures 28-29**, Section 3.1.1.7).

Single incorporation of dC^{4TCO}TP or dC^{2TCO}TP or dC^{p4TCO}TP or dC^{p2TCO}TP or dC^{pBCN}TP using 19-mer template-Semi-preparative scale

The reaction mixture (20 µL) contained KOD XL DNA polymerase (0.25 U/µL, 0.2 µL), primer **Prim A** (50 µM, 1 µL), template **Temp19_1C_2xbio** (50 µM, 1.5 µL), dGTP (4 mM, 0.5 uL), dCTP (modified; 4 mM, 0.3 µL each), in KOD XL reaction buffer (10 X, 2.05 µL). The reaction mixture was incubated in a thermal cycler for 20 min at 60 °C and stopped by cooling to 4 °C. Biotinylated PEX products were purified using DB Streptavidin magneto-separation procedure and then analysed by MALDI-TOF mass spectrometry and used for fluorescence turn-on measurements in case of **19ON_C^{p4TCO}** and **19ON_C^{pBCN}** (See **Table 7**, Section 3.1.1.10).

5.1.6 DB Streptavidin magneto-separation procedure:

Streptavidin particles (Roche; 50 µL) were washed with binding buffer (3 × 200 µL; 10 mM Tris, 1 mM EDTA, 100 mM NaCl, pH 7.5). The PEX solution (50 µL) was mixed with binding buffer (200 µL) and incubated for 30 min at 15 °C and 1400 rpm. The magnetic beads were collected on a magnet (DynaMagTM-2, Invitrogen), and washed with wash buffer (3 × 300 µL; 10 mM Tris, 1 mM EDTA, 500 mM NaCl, pH 7.5) and water (4 × 300 µL). Then water (50 µL) was added and the sample was denatured for 2 min at 900 rpm and 75 °C. The beads were collected on a magnet and the solution was transferred into a clean vial. The product was evaporated to dryness, then dissolved in water and analysed by MALDI-TOF mass spectrometry.

5.1.7 Reaction of 19DNA_C^{X} (X = 4TCO or 2TCO or p4TCO or p2TCO or pBCN) with tetrazines

Reaction of 19DNA_C^{X} (X = 4TCO or 2TCO or p4TCO or p2TCO or pBCN) with T1 or T2 tetrazines and of 19DNA_C^{X} (X = p4TCO or pBCN) with T4 tetrazine

19DNA_C^{natural} and **19DNA_C^{X}** were prepared in semi-preparative scale using 6-FAM-labelled primer as described above (Section 5.1.5) and firstly used as negative controls (without treatment with **T1**, **T2**, **T4**) and then modified **19DNA_C^{X}** were used for post-synthetic labelling with **T1**, **T2** and **T4** tetrazines. The crude mixture for each **19DNA_C^{X}**, after PEX, was purified by QIAGEN column and subsequently **T1** or **T2** or **T4** tetrazine (each either 10 µM or 100 µM in 5

µL DMSO) was added to the purified **19DNA_C^{X}** (0.4 µM in 5 µL H_2O). The reaction (10 µL) with final concentration of **19DNA_C^{X}** 0.2 µM and of **T1** or **T2** or **T4** tetrazine 5 µM or 50 µM was incubated for 30 min at 37 °C in a thermal cycler. After incubation with tetrazine, the reaction was purified again by QIAGEN column to remove the excess of unreacted tetrazine. PAGE stop solution (10 µL) was added to the purified reaction (10 µL) and samples were analysed by 12.5% denaturing PAGE and visualized using fluorescence imaging **(Figure 30** – click reaction with **T1**, **Figure 33** – click reaction with **T2**, **Figure 34** – click reaction with **T4**).

Reaction of 19DNA_C^{X} (X = p4TCO or pBCN) with T3 tetrazine

19DNA_C^{natural} and **19DNA_C^{X}** (**X** = **p4TCO** or **pBCN**) were prepared in semi-preparative scale using 6-FAM-labelled primer as described above (Section 5.1.5) and firstly used as negative controls (without treatment with **T3**) and then modified **19DNA_C^{X}** were used for post-synthetic labelling with T3 tetrazine. The crude mixture, after PEX, was purified by QIAGEN column and subsequently **T3** tetrazine (100 µM or 1 mM in 5 µL DMSO) was added to the purified **19DNA_C^{X}** (0.4 µM in 5 µL H_2O). The reaction (10 µL) with final concentration of **19DNA_C^{X}** 0.2 µM and of **T3** tetrazine 50 µM or 500 µM was incubated for 18 h (due to low reactivity) at 37 °C in a thermal cycler. After incubation with tetrazine, the reaction was purified by QIAGEN column to remove the excess of unreacted tetrazine. Then, magneto-separation was applied to have single-stranded clicked **19ON_C^{X}** products. PAGE stop solution (10 µL) was added to the purified reaction (10 µL) and samples were analysed by 20% NATIVE PAGE and visualized using fluorescence imaging (**Figure 34**). All click products were confirmed by MALDI (See **Table 6**, Section 3.1.1.9).

5.1.8 Formation of tetrazine-ONs click products (19ON_C^{XTY}) for MALDI-TOF analysis

Procedure: Using **Temp19_1C_2xbio** and **primer PrimA**, **19DNA_C^{4TCO}**, **19DNA_C^{2TCO}**, **19DNA_C^{p4TCO}** and **19DNA_C^{p2TCO}** or **19ON_C^{pBCN}** were prepared by PEX in semi-preparative scale as described above (Section 5.1.5). The reaction mixture (20 µL) containing **19DNA_C^{4TCO}** or **19DNA_C^{2TCO}** or **19DNA_C^{p4TCO}** or **19DNA_C^{p2TCO}** or **19ON_C^{pBCN}** (0.5 µM in water) and tetrazine (**T1** or **T2**, **T3**[6] or **T4**, each 100 µM in DMSO) were stirred at 37 °C for 30 min, except for **T3** which was kept overnight due to lower reactivity. Subsequently, the products were purified by magneto-separation, as described above (Section 5.1.6) to get **19ON_C^{4TCOT1}** or

19ON_C^{4TCOT2}, **19ON_C^{2TCOT1}** or **19ON_C^{2TCOT2}**, **19ON_C^{p4TCOT1}** or **19ON_C^{p4TCOT2}** or **19ON_C^{p4TCOT3}** or **19ON_C^{p4TCOT4}**, **19ON_C^{p2TCOT1}** or **19ON_C^{p2TCOT2}**, **19ON_C^{pBCNT1}** or **19ON_C^{pBCNT2}** or **19ON_C^{pBCNT3}** or **19ON_C^{pBCNT4}** tetrazine-linked products. The resulted ONs were analysed by MALDI-TOF mass spectrometry (See **Table 6**, Section 3.1.1.9).

5.1.9 Oligonucleotide fluorescence turn-on measurement of 19ON_C^{p4TCO} or 19ON_C^{pBCN} with T1 or T3 tetrazine (19ON_C^{XTY})

19ON_C^{p4TCO} and **19ON_C^{pBCN}** were prepared in semi-preparative scale as described in Section 5.1.5 using template **19ON_C_2xbio** and primer **PrimA**, then magneto-separation was applied. Freshly prepared stock solution of the tetrazine **T1** (0.1 mM in DMSO) was used for the measurement. The turn-on fluorescence measurements were performed as follows:
1 µL of the tetrazine stock solution (0.1 mM in DMSO) was diluted in 100 µL PBS buffer (1 X, pH 7.4, giving 1 µM final concentration) and the fluorescence was measured using 370 nm excitation (slit 3) and 400-650 nm for emission. To this solution was added solution of the **19ON_C^{p4TCO}** or **19ON_C^{pBCN}** (5.5 µL from 18 µM fresh stock in water, giving 1 µM final concentration, 1 equiv. to **T1** tetrazine). The spectra were recorded at several time points (final measurements after 30 minutes) and all measurements were typically repeated two-three times.
Due to the lower reactivity of the **T3** tetrazine, the fluorescence measurements were performed as follows: 5.5 µL of 18 µM solution of **19ON_C^{p4TCO}** or 7 µL of 14 µM solution of **19ON_C^{pBCN}** was combined with 1 µL of 0.1 mM solution of **T3** tetrazine and the reaction mixtures were incubated in the dark at room temperature for 3 hours. Then, the reaction mixtures were diluted to 100 µL with PBS and the fluorescence was measured using 500 nm excitation wavelength and 525-750 nm emission window (slit 3).
The cuvette was inserted into the fluorescence spectrophotometer and the measurement was started. The spectra were recorded at several time points (final measurements after 30 minutes) and all measurements were typically repeated two-three times. The data were processed using OriginPro software. All spectra were subtracted from the baseline (PBS buffer as the blank). The turn-on values were calculated from the observed fluorescence intensities of the click products at the emission maximum (475 nm) after ca. 30 min of reaction divided by the highest residual fluorescence of the quenched coumarin tetrazine dye at the same wevelength (475 nm). The

fluorescence turn-on values of the click products formed in the reaction of **19ON_C^{p4TCO}** and **19ON_C^{pBCN}** with **T1** tetrazine are shown in **Figures 93-94** or of **19ON_C^{pBCN}** with **T3** tetrazine is shown in **Figure 95** and **Table 7**. In case of **19ON_C^{p4TCO}** with **T3** was not observed any turn on in the fluorescence.

5.1.10 Labelling of DNA using dC^{p4TCO} or dC^{p2TCO}TP or dC^{pBCN}TP derivatives delivered to U-2 OS cells by SNTT1

General remarks for cellulo experiments

Materials

Cell cultures: U-2 OS (ATCC HTB-96)

SNTT1; Synthetic Nucleoside Triphosphate Transporter, identical with commercially available BioTracker NTP-*Transporter* molecule (Merck *SCT064*), was used as a 10 mM stock solution in tricine buffer.

L15 medium (Gibco, 21083-027)

FBS (Capricorn, cat.no. FBS-HI-12A)

GlutaMAX-I (Gibco, cat.no. 35050-038)

Tricine buffer

Preparation procedure – dissolve following compounds (cell culture quality) in 980 mL of pure DI (MilliQ) water:

Tricine (0.896 g; 5 mmol),

Glucose (2 g; 11 mmol),

NaCl (7.325 g; 125 mmol),

$CaCl_2$ anhydrous (0.200 g; 1.8 mmol),

$MgSO_4$ anhydrous (0.100 g; 0.8 mmol),

KCl (0.400 g; 5.4 mmol),

Adjust pH of the buffer to 7.40 (at 31 °C) with NaOH (1 mol/L; approx. 850 µL).

Add pure (MilliQ) water to 1000 mL.

Filter the solution through 0.22 µm sterile filter. Keep under sterile conditions at 0-4 °C.

Tetrazines **T1**, **T2** and **T3**: working solutions of **T1**, **T2** and **T3** were prepared from the corresponding stock solutions in DMSO (**T1**, **T2** - 1 mM in DMSO; **T3** - 10 mM in DMSO) by dilution with L15 complete medium to the desired concentrations prior to use.

Tetrazine **T4**: Pyrimidyl-Tetrazine-5-TAMRA (Jena Bioscience; CLK-097) was diluted with DMSO to 10 mM stock solution which was than diluted with L15 complete medium to the desired concentrations prior to use.

DNA-staining solution: Prepare stock solution of DAPI in water (1 mg/ml). Dilute this stock solution with PBS in ratio 1:100 (1µL of DAPI stock solution per 100 µL of PBS) to make the final working reagent, further referred to as "DNA-staining solution".

Instrumentation

Confocal microscope ZEISS LSM 980 fitted with an incubator and objective 40x/1.2 W C-Apochromat (Carl Zeiss Jena GmbH, Germany)

LUNA Automated Cell Counter

Labelling of DNA using dC^{p4TCO}TP or dC^{pBCN}TP and tetrazines T1 or T2 in live U2 OS cells

1) Preparation of the cell culture

U-2 OS cells (ATCC HTB-96) were cultured in a 96 Well Glass Bottom Plate (Cellvis, P96-1.5H-N) in complete medium [L-15 (Gibco, 2C3-027) supplemented with 10% FBS (Capricorn, cat.no. FBS-HI-12A) and 1% GlutaMAX-I (Gibco, cat.no. 35050-038)] in an incubator (37°C, no additional CO_2) to the confluency of about 70%.

General procedure for the incorporation of dC^{X}TP (X= 2TCO, 4TCO, p2TCO, p4TCO, pBCN) into DNA of U2OS cells and subsequent labelling with a tetrazine T1 or T2

The medium was aspirated and the cells were washed with pre-warmed (37°C) tricine buffer. Then the cells were treated with 40 µL of an equimolar mixture of **dC^{X}TP** and **SNTT 1** (10 µM) in tricine buffer (37°C, 5 min). The mixture was then removed and complete L-15 medium (80 µL) was added to the cells. The cell culture was incubated in complete medium (37°C) for 60 min and then stock solution of a tetrazine (**T1** or **T2**) was added to achieve the final desired concentration (1 µM) in media. The plate with live cells was mounted onto microscope (ZEISS LSM 980) and incubated (37°C) for a further 15-minute period. The samples were irradiated sequentially with a laser beam at 405 nm, the detector range was set to 410 – 499 nm for the coumarine detection.

Procedure for the labeling with dC^{p4TCO}TP and tetrazine T1 and subsequent staining of nuclei with DRAQ5

The medium was aspirated and the cells were washed with pre-warmed (37°C) tricine buffer. Then the cells were treated with 40 µL of an equimolar mixture of **dC^{p4TCO}TP** and **SNTT 1** (10 µM) in tricine buffer (37°C, 5 min). The mixture was then removed and complete L-15 medium (80 µL) was added to the cells. The cell culture was incubated in complete medium (37°C) for 60 min and then stock solution of tetrazine **T1** was added so that the final concentration in media was 1 µM. The plate with live cells was mounted onto microscope (ZEISS LSM 980) and incubated (37°C) for a further 10-minute period. Then DRAQ5 (2.5 µM) was added to the medium, the cells were further incubated for 5 minutes before they were imaged. The samples were irradiated sequentially with a laser beam at 405 nm, the detector range was set to 410 – 499 nm for the coumarine detection. The DRAQ5 dye was excited at 639 nm, detected at 642 – 756 nm.

Procedure for the labeling with dC^{pBCN}TP and tetrazine T4 (pyrimidyl-tetrazine-5-TAMRA) in live cells

The medium was aspirated and the cells were washed with pre-warmed (37°C) tricine buffer. Then the cells were treated with 40 µL of an equimolar mixture of **dC^{pBCN}TP** and **SNTT 1** (10 µM) in tricine buffer (37°C, 5 min). The mixture was then removed and complete L-15 medium (80 µL) was added to the cells. The cell culture was incubated in complete medium (37°C) for 60 min and then stock solution of tetrazine **T4** was added to the well so that the final concentration in media was 20 µM. The plate was mounted onto microscope (ZEISS LSM 980) and further incubated

(37°C; 45 min in total); in the course of incubation the cells were imaged by confocal microscopy in 10 min periods to check for optimal incubation time. After 45 min of incubation, the samples were irradiated with a laser beam at 561 nm, the detector range was set to 534 – 690 nm.

Procedure for the labeling with dC^{pBCN}TP and tetrazine T4 (pyrimidyl-tetrazine-5-TAMRA) in fixed cells

The medium was aspirated and the cells were washed with pre-warmed (37°C) tricine buffer. Then the cells were treated with 40 µL of an equimolar mixture **dC^{pBCN}TP** and **SNTT 1** (10 µM) in tricine buffer (37°C, 5 min). The mixture was then removed and complete L-15 medium (80 µL) was added to the cells. The cell culture was incubated in complete medium (37°C) for 60 – 90 min. After this period the cells were washed with cold PBS and cold methanol (-20°C) was added. The plate was kept at 0° C for 60 min, then methanol was aspirated, the fixed cells were washed with PBS and allowed to warm up to room temperature and then stock solution of tetrazine pyrimidyl-tetrazine-5-TAMRA (2 µM) in PBS was added. The plate was mounted onto microscope (ZEISS LSM 980) and further incubated (37°C); in the course of incubation the cells were imaged by confocal microscopy in 10 min periods to check for optimal incubation time (30 – 40 min). After 30 min of incubation, HOECHST 33342 (3 µM) was added to the medium, the cells were further incubated for 5 minutes before they were imaged. The samples were irradiated sequentially with a laser beam at 405 nm, the detector range was set to 410 – 499 nm for HOECHST 33342 detection. TAMRA dye was excited at 561 nm, detected at 534 – 690 nm.

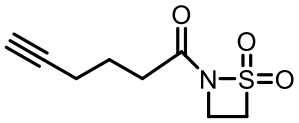
5.1.10.1 Protocol for the cell cycle progression experiment

The cells were cultured in Leibovitz's (1X) L-15 Medium, no phenol red (cat.no. 21083-027, Gibco^{TM}, Fisher Scientific, Sweeden) supplemented with 10 % Fetal bovine serum (cat. no. F9665, Sigma-Aldrich, USA). Aproximately 0.5 million of U 2-OS cells were seeded in a well of a 6-well plate (cat. no. 3046 FALCON, Becton Dickinson, USA) and cultured for 24 hours (37°C, no additional CO_2). Then the cells were washed with pre-warmed tricine buffer and treated with 500 µL an equimolar mixture **dC^{p4TCO}TP** and **SNTT1** (10 µM) in tricine buffer (37°C, 5 min). The mixture was then removed and complete L-15 medium (2 mL) was added to the cells and the cells were incubated for 60 min (37°C, no additional CO_2). After this period, the cells were washed with PBS (phosphate buffered saline), dissociated with TripLE^{TM} (cat. no. 12604-013, Gibco^{TM}, Fisher Scientific, Sweeden) and fixed with methanol for 60 min at 4 °C as described above. Subsequently

the cells were washed with PBS and incubated with 2 μM pyrimidyl-tetrazine-5-TAMRA **(T4)** in PBS for 30 minutes at 37 °C in thermoblock. The cells were then centrifuged (300 x g ; 3 min) to remove **T4** solution and PBS (0.25 mL) was added. DNA-staining solution (100 μL; see above for the preparation of working solution of DAPI) was added and the cell suspension and the cells were analyzed by flow cytometry (BD LSRFortessaTM). For TAMRA, yellow-green laser (561 nm) was used with detection filters 582/15-nm band pass. For DAPI, UV laser (355 nm) was used with detection filter 450/50-nm band pass. Flow rate was set lower than 400 events/second and 20 000 events were collected. For analysis FlowJo™ v10.6.1 (Becton, Dickinson & Company) was used.

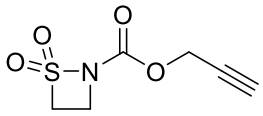
5.2 Design and synthesis of modified nucleotides bearing thiazetidine (β-sultam) for cross-linking with serine, their enzymatic incorporation and testing of reactivity

Synthesis of 1-(1,1-dioxido-1,2-thiazetidin-2-yl)hex-5-yn-1-one (22)



1,2-thiazetidin-1,1-dioxide (1.86 gr, 14.24 mmol) was added to **21** (1.02 gr, 9.49 mmol) and DMAP (2.32 gr, 18.99 mmol) in DCM (15 mL) at 0 °C and stirred for 2 h. The reaction mixture was diluted with DCM (50 mL) and washed with saturated acqueous NH_4Cl (1x50 mL) and brine (1x50 mL). The organic layer was dried over anhydrous Na_2SO_4 and concentrated in vacuo. Purification by flash colum chromatography gave the product **22** as a white solid (78% yield).
All spectra are consisted with published literature.[80]

Prop-2-yn-1-yl 1,2-thiazetidine-2-carboxylate 1,1-dioxide (24)



nBuLi (1.6 M in hexane, 1.18 mL, 4.67 mmol, 1.0 eq) was added to a solution of 1,2-thiazetidin-1,1-dioxide (200 mg, 4.67 mmol, 1.0 eq) in THF dry (10 mL) at -78 ºC for 10 min. Propargyl chloroformate (0.18 mL, 4.67 mmol, 1 eq) in THF dry (0.8 mL) was added and the reaction mixture was allowed to warm up to RT. Then brine (10 mL) was added, the phases were separated and the acqueous layer was extracted with EtOAc (2x10 mL). The combined organic fractions were dried

over anhydrous Na_2SO_4 and concentrated in vacuo. Purification by flash column chromatography (DCM) gave the desired compound as white solid (80 mg, 35 % yield)
All spectral data are consistent with the published literature.[80]

5.2.1 Synthesis of modified monophosphates ($dC^{SU1}MP$, $dC^{SU2}MP$, $dU^{SU2}MP$) and triphosphates ($dC^{SU1}TP$, $dC^{SU2}TP$)

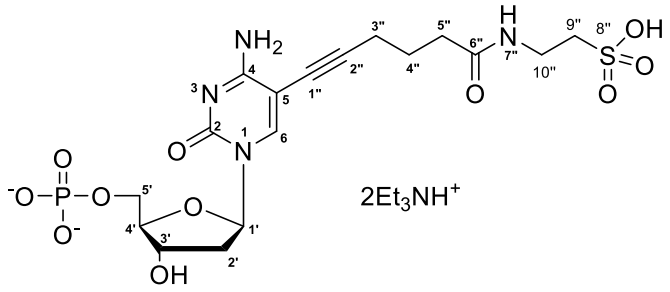
Synthesis of nucleoside monophosphate

General procedure A: Preparation of 2'-deoxynucloside 5'-O-phosphates (dN^XMP). Compound dN^X was dried at 80 °C for 2 h in vacuo. After cooling, dry $PO(OMe)_3$ and $POCl_3$ (1.2 equiv) were added and the reaction mixture was stirred under an argon atmosphere for 1 h on ice.
Phosphorylation was stopped by the addition of TEAB (2M, 1mL) and water (2 mL). After purification, several codistillations with water followed by freeze-drying from water, gave the desired **dC^XMP** as a white powder.

Synthesis of modified nucleoside monophosphate or triphosphate

General procedure B: Nucleoside monophosphate analogue (dNMP or dNTP), modified alkyne CuI, $Pd(OAc)_2$, TPPTS were dissolved in mixture water/acetonitrile (2:1, 2.5 ml) under argon atmosphere followed by addition of DIPEA. The reaction mixture was stirred at RT for 1 hour and then evaporated in vacuo. The products were purified by C18 reverse phase HPLC using water-methanol (5 to 100 %) containing 0.1M TEAB buffer as eluent. Several co-distillations with water and freeze-drying from water gave the desired products as pale yellow powders.

Synthesis of 5-{6-oxo-6-[(2-sulfoethyl)amino)hex-1-yn-1-yl]}-2'-deoxycitidine-5'-*O*-monophosphate (dC^{SU1}MP) (27a)



dC^{SU1}MP: According to *General procedure C*, **dC^{SU2}MP** was prepared from **dC^{N3}MP** (50.0 mg, 200 µmol), **22** (301 µmol), sodium ascorbate (15 mg, 75 µmol) and copper(II) sulfate pentahydrate (4 mg, 16 µmol) as white solid (30% yield).

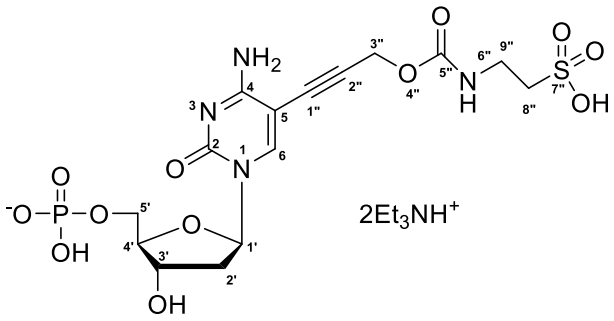
^{1}H NMR (500.0 MHz, D_2O, ref(*t*BuOH) = 1.24 ppm): 1.27 (t, 18H, J_{vic} = 7.3, C**H$_3$**CH$_2$N); 1.89 (tt, 2H, $J_{4'',5''}$ = 7.3, $J_{4'',3''}$ = 6.9, H-4″); 2.29 (ddd, 1H, J_{gem} = 13.9, $J_{2'b,1'}$ = 7.2, $J_{2'b,3'}$ = 6.5, H-2′b); 2.41 (t, 2H, $J_{5'',4''}$ = 7.3, H-5″); 2.44 (ddd, 1H, J_{gem} = 13.9, $J_{2'a,1'}$ = 6.3, $J_{2'a,3'}$ = 3.7, H-2′a); 2.50 (t, 2H, $J_{3'',4''}$ = 6.9, H-3″); 3.07 (t, 2H, $J_{9'',10''}$ = 6.9, H-9″); 3.20 (q, 12H, J_{vic} = 7.3, CH$_3$C**H$_2$**N); 3.56 (t, 2H, $J_{10'',9''}$ = 6.9, H-10″); 3.99 – 4.06 (m, 2H, H-5′); 4.18 (tdd, 1H, $J_{4',5'}$ = 4.3, $J_{4',3'}$ = 3.7, $J_{H,P}$ = 1.1, H-4′); 4.52 (dt, 1H, $J_{3',2'}$ = 6.5, 3.7, $J_{3',4'}$ = 3.7, H-3′); 6.26 (dd, 1H, $J_{1',2'}$ = 7.2, 6.3, H-1′); 8.05 (s, 1H, H-6).

^{13}C NMR (125.7 MHz, D_2O, ref(*t*BuOH) = 32.43 ppm): 11.06 (**C**H$_3$CH$_2$N); 21.17 (CH$_2$-3″); 26.72 (CH$_2$-4″); 37.67 (CH$_2$-5″); 37.93 (CH$_2$-10″); 42.18 (CH$_2$-2′); 49.50 (CH$_3$**C**H$_2$N); 52.50 (CH$_2$-9″); 67.15 (d, $J_{C,P}$ = 4.8, CH$_2$-5′); 73.74 (CH-3′); 73.96 (**C**≡C-CH$_2$); 88.57 (d, $J_{C,P}$ = 8.5, CH-4′); 89.18 (CH-1′); 96.09 (C-5); 99.59 (C≡**C**-CH$_2$); 146.61 (CH-6); 158.93 (C-2); 168.10 (C-4); 179.00 (C-6″).

^{31}P{^{1}H}NMR (202.4 MHz, D_2O): 1.41.

MS (ESI-) calc. $C_{17}H_{23}N_4O_{11}PS_2^-$ calc. 522.08 , found [M-H+Na] 545.1.

Synthesis of 5-{3-[(2-sulfoethyl)carbamoyloxy]prop-1-yn-1-yl}-2'-deoxycitidine-5'-O-monophosphate (dC^{SU2}MP) (28a)



dC^{I}MP was prepared according to *General procedure A,* and then the final compound **dC^{SU2}MP** was obtained applying the *General procedure B* from **dC^{I}MP** (50 mg, 115 µmol), **24** (160 µmol), CuI (1.62 mg, 5 µmol), $Pd(OAc)_2$ (0.95 mg, 4.2 µmol), TPPTS (5.0 mg, 50 µmol) and DIPEA (150 µl, 0.112 mmol). After purification by C18 reverse phase HPLC using water/methanol (5 to 100 %) containing 0.1M TEAB buffer as eluent the final compound was isolated in 35% yield but only in opened ring form which was confirmed by mass.

^{1}H NMR (500.0 MHz, D_2O, ref(*t*BuOH) = 1.24 ppm): 1.27 (t, 18H, J_{vic} = 7.3, C**H$_3$**CH_2N); 30 (ddd, 1H, J_{gem} = 14.0, $J_{2'b,1'}$ = 7.0, $J_{2'b,3'}$ = 6.7, H-2′b); 2.45 (ddd, 1H, J_{gem} = 14.0, $J_{2'a,1'}$ = 6.3, $J_{2'a,3'}$ = 3.8, H-2′a); 3.10 (t, 2H, $J_{8'',9''}$ = 6.9, H-8″); 3.20 (q, 12H, J_{vic} = 7.3, CH_3C**H$_2$**N); 3.55 (t, 2H, $J_{9'',8''}$ = 6.9, H-9″); 4.01 (dd, 2H, $J_{H,P}$ = 5.4, $J_{5',4'}$ = 4.2, H-5′); 4.18 (td, 1H, $J_{4',5'}$ = 4.2, $J_{4',3'}$ = 3.8, H-4′); 4.51 (dt, 1H, $J_{3',2'}$ = 6.7, 3.8, $J_{3',4'}$ = 3.8, H-3′); 4.91 (s, 2H, H-3″); 6.24 (dd, 1H, $J_{1',2'}$ = 7.0, 6.3, H-1′); 8.16 (s, 1H, H-6).

^{13}C NMR (125.7 MHz, D_2O, ref(*t*BuOH) = 32.43 ppm): 11.06 (**C**H_3CH_2N); 39.21 (CH_2-9″); 42.20 (CH_2-2′); 49.50 (CH_3**C**H_2N); 52.94 (CH_2-8″); 56.37 (CH_2-3″); 67.03 (d, $J_{C,P}$ = 4.0, CH_2-5′); 73.69 (CH-3′); 79.82 (**C**≡C-CH_2); 88.67 (d, $J_{C,P}$ = 8.3, CH-4′); 89.37 (CH-1′); 93.41 (C≡**C**-CH_2); 94.44 (C-5); 148.30 (CH-6); 158.86 (C-2); 160.23 (C-6″); 167.95 (C-4).

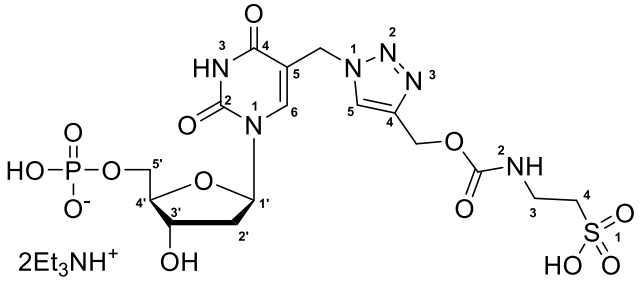
^{31}P{^{1}H}NMR (202.4 MHz, D_2O): 1.88.

MS (ESI-) calc. For $C_{15}H_{20}N_4O_{12}PS^-$ 511.04, found [M-H] 511.06.

General procedure C: Copper-catalyzed click reactions to dU^{N3}MP

dU^{N3}MP [24] alkyne cointaining the reactive mojety, sodium ascorbate and copper(II) sulfate pentahydrate were dissolved in a mixture of water/tert-butanol. The reaction mixture was stirred at room temperature for 2 h (complete consumption of the starting material was proved by TLC with propan-2 ol/water/aqueous ammonia (11:2:7) as the eluent). The products were purified by C18 reverse-phase HPLC with a mixture of water/methanol (5 to 100%) containing 0.1 M TEAB buffer as eluent.

5-{4-[(2-sulfoethyl)carbamoyloxy]methyl}-1H-1,2,3-triazol-1-yl)methyl)-2'deoxyuridine-5'-O-monophosphate (dU^{SU2}MP) (33a)



dU^{SU2}MP: According to *General procedure C*, **dU^{SU2}MP** was prepared from **dU^{N3}MP** (50.0 mg, 200 µmol), **24** (301 µmol), sodium ascorbate (15 mg, 75 µmol) and copper(II) sulfate pentahydrate (4 mg, 16 µmol). After HPLC purification the product was isolated as white solid (20%, yield).

^{1}H NMR (500.0 MHz, D_2O, ref(*t*BuOH) = 1.24 ppm): 1.27 (t, 18H, J_{vic} = 7.3, C**H$_3$**CH$_2$N); 2.41 (ddd, 1H, J_{gem} = 14.1, $J_{2'b,1'}$ = 6.4, $J_{2'b,3'}$ = 3.8, H-2′b); 2.45 (ddd, 1H, J_{gem} = 14.1, $J_{2'a,1'}$ = 7.3, $J_{2'a,3'}$ = 5.9, H-2′a); 2.96 (t, 2H, $J_{4,3}$ = 6.9, H-4, CH_2); 3.20 (q, 12H, J_{vic} = 7.3, CH_3C**H$_2$**N); 3.51 (t, 2H, $J_{3,4}$ = 6.9, H-3, CH_2); 4.01 (ddd, 1H, J_{gem} = 11.5, $J_{H,P}$ = 5.1, $J_{5'b,4'}$ = 3.6, H-5′b); 4.05 (ddd, 1H, J_{gem} = 11.5, $J_{H,P}$ = 4.5, $J_{5'a,4'}$ = 3.3, H-5′a); 4.19 (tdd, 1H, $J_{4',3'}$ = 3.6, $J_{4',5'}$ = 3.6, 3.3, $J_{H,P}$ = 1.7, H-4′); 4.60 (ddd, 1H, $J_{3',2'}$ = 5.9, 3.8, $J_{3',4'}$ = 3.6, H-3′); 5.17 (s, 2H, CH_2O); 5.40 (s, 2H, CH_2N); 6.33 (dd, 1H, $J_{1',2'}$ = 7.3, 6.4, H-1′); 8.09 (s, 1H, H-5-triazole); 8.37 (s, 1H, H-6).

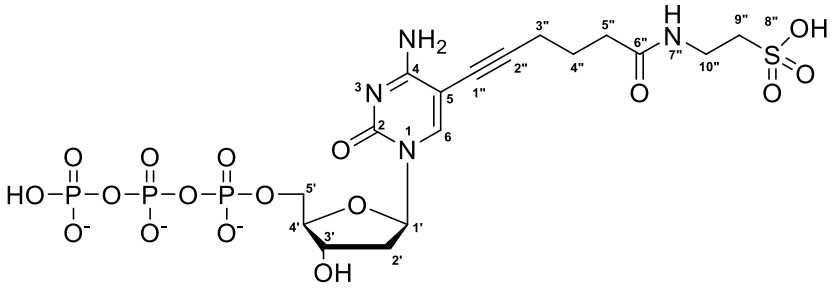
^{13}C NMR (125.7 MHz, D_2O, ref(*t*BuOH) = 32.43 ppm): 11.06 (**C**H_3CH_2N); 39.13 (CH_2-3); 41.90 (CH_2-2′); 49.49 (CH_3**C**H_2N); 49.89 (CH_2N); 52.99 (CH_2-4); 60.64 (CH_2O); 66.78 (d, $J_{C,P}$ = 4.7,

CH_2-5′); 73.92 (CH-3′); 88.49 (CH-1′); 89.05 (d, $J_{C,P}$ = 8.6, CH-4′); 111.27 (C-5); 128.11 (CH-5-triazole); 145.26 (CH-6); 145.76 (C-4-triazole); 154.33 (C-2); 160.41 (NCOO); 167.38 (C-4).
$^{31}P\{^1H\}$ NMR (202.4 MHz, D_2O): 2.34.
MS (ESI+) calcd. For $C_{16}H_{21}N_6O_{13}PS^{2-}$ 568.06 (exact mass), found [M+2Na] 614.20.

Synthesis of nucleoside triphosphate

General procedure D: Preparation of 2'-deoxynucloside 5'-*O*-phosphates (dN^XTP). Compound dN^X was dried at 80 °C for 2 h in vacuo. After cooling, dry $PO(OMe)_3$ and $POCl_3$ (1.2 equiv) were added and the reaction mixture was stirred under an argon atmosphere for 1 h on ice. Therefore pyrophosphate and Et_3N were added and the reaction is kept stirring further for 2 h. Phosphorylation was stopped by the addition of TEAB (2M, 1mL) and water (2 mL). After purification, several codistillations with water followed by freeze-drying from water, gave the desired **dN^XTP** as a white powder.

Synthesis of 5-{[6-oxo-6-(2-sulfoethyl)amino]hex-1-yn-1-yl}-2'-deoxycitidine-5'-*O*-triphosphate ($dC^{SU1}TP$) (29a)



dC^ITP was prepared according to *General procedure D* and subsequently final compound **$dC^{SU1}TP$** prepared following the *General procedure B* from **dC^ITP** (50 mg, 115 µmol), **22** (160 µmol), CuI (1.62 mg, 5 µmol), $Pd(OAc)_2$ (0.95 mg, 4.2 µmol), TPPTS (5.0 mg, 50 µmol) and DIPEA (150 µl, 0.112 mmol). After purification by C18 reverse phase HPLC using water/methanol (5 to 100 %) containing 0.1M TEAB buffer as eluent the final compound was isolated in 10% yield but only in opened ring form which was confirmed by mass.

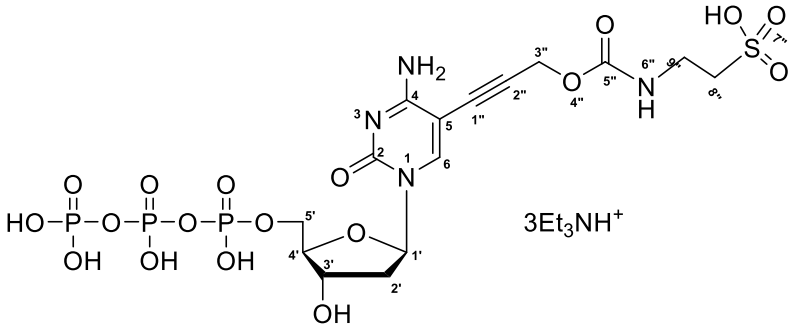
^{1}H NMR (600.1 MHz, D_2O, ref(*t*BuOH) = 1.24 ppm): 1.84 – 1.94 (m, 2H, H-4″); 2.30 (dt, 1H, J_{gem} = 14.0, $J_{2'b,1'} = J_{2'b,3'}$ = 6.9, H-2′b); 2.40 (t, 2H, $J_{5'',4''}$ = 7.2, H-5″); 2.42 (ddd, 1H, J_{gem} = 14.0, $J_{2'a,1'}$ = 6.9, $J_{2'a,3'}$ = 4.2, H-2′a); 2.51 (t, 2H, $J_{3'',4''}$ = 7.0, H-3″); 3.08 (t, 2H, $J_{9'',10''}$ = 6.9, H-9″); 3.56 (t, 2H, $J_{10'',9''}$ = 6.9, H-10″); 4.17 – 4.25 (m, 3H, H-4′,5′); 4.60 (ddd, 1H, $J_{3',2'}$ = 6.9, 4.2, $J_{3',4'}$ = 3.3, H-3′); 6.25 (dd, 1H, $J_{1',2'}$ = 6.9, 6.4, H-1′); 8.03 (s, 1H, H-6).

^{13}C NMR (150.9 MHz, D_2O, ref(*t*BuOH) = 32.43 ppm): 21.15 (CH_2-3″); 26.69 (CH_2-4″); 37.67 (CH_2-5″); 37.89 (CH_2-10″); 41.82 (CH_2-2′); 52.50 (CH_2-9″); 67.98 (d, $J_{C,P}$ = 5.6, CH_2-5′); 72.98 (CH-3′); 73.89 (**C**≡C-CH_2); 88.72 (d, $J_{C,P}$ = 8.5, CH-4′); 88.96 (CH-1′); 96.04 (C-5); 99.64 (C≡**C**-CH_2); 146.63 (CH-6); 158.88 (C-2); 168.09 (C-4); 178.94 (C-6″).

^{31}P{^{1}H}NMR (202.5 MHz, D_2O): -20.77 (bt, *J* = 17.5, P_β); -10.20 (d, *J* = 17.5, P_α); -5.23 (bd, *J* = 17.5, P_γ).

MS (ESI-) calc. For $C_{17}H_{24}N_4O_{17}P_3S^{3-}$ 699.01 found [M-H] 699.0.

Synthesis of 5-{3-[(2-sulfoethyl)carbamoyloxy]prop-1-yn-1-yl}-2'-deoxycitidine-5'-*O*-triphosphate (dC^{SU2}TP) (30a)



dC^{I}TP was prepared according to *General procedure D* and subsequently final compound **dC^{SU2}TP** prepared following the *General procedure B* from **dC^{I}TP** (50 mg, 115 µmol), **24** (160 µmol), CuI (1.62 mg, 5 µmol), $Pd(OAc)_2$ (0.95 mg, 4.2 µmol), TPPTS (5.0 mg, 50 µmol) and DIPEA (150 µl, 0.112 mmol). After purification by C18 reverse phase HPLC using water/methanol (5 to 100 %) containing 0.1M TEAB buffer as eluent the final compound was isolated in 20% yield but only in opened ring form which was confirmed by mass.

^{1}H NMR (500.0 MHz, D_2O, ref(*t*BuOH) = 1.24 ppm): 1.28 (t, 27H, J_{vic} = 7.3, C**H**$_3$CH$_2$N); 2.31 (ddd, 1H, J_{gem} = 14.1, $J_{2'b,1'}$ = 6.9, $J_{2'b,3'}$ = 6.6, H-2′b); 2.44 (ddd, 1H, J_{gem} = 14.1, $J_{2'a,1'}$ = 6.3, $J_{2'a,3'}$ = 4.1, H-2′a); 3.10 (t, 2H, $J_{8'',9''}$ = 6.9, H-8″); 3.20 (q, 18H, J_{vic} = 7.3, CH$_3$C**H**$_2$N); 3.54 (t, 2H, $J_{9'',8''}$ = 6.9, H-9″); 4.17 – 4.25 (m, 3H, H-4′,5′); 4.60 (bm, 1H, H-3′); 4.92 (s, 2H, H-3″); 6.24 (dd, 1H, $J_{1',2'}$ = 6.9, 6.3, H-1′); 8.17 (s, 1H, H-6).

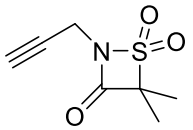
^{13}C NMR (125.7 MHz, D_2O, ref(*t*BuOH) = 32.43 ppm): 11.06 (**C**H$_3$CH$_2$N); 39.20 (CH$_2$-9″); 42.02 (CH$_2$-2′); 49.50 (CH$_3$**C**H$_2$N); 52.97 (CH$_2$-8″); 56.43 (CH$_2$-3″); 68.01 (d, $J_{C,P}$ = 5.0, CH$_2$-5′); 73.11 (CH-3′); 79.78 (**C**≡C-CH$_2$); 88.34 (d, $J_{C,P}$ = 8.6, CH-4′); 89.22 (CH-1′); 93.49 (C≡**C**-CH$_2$); 94.47 (C-5); 148.28 (CH-6); 158.87 (C-2); 160.27 (C-6″); 167.96 (C-4).

^{31}P{^{1}H}NMR (202.4 MHz, D_2O): -22.73 (bs, P$_\beta$); -11.29 (d, J = 19.4, P$_\alpha$); –8.88 (vbs, P$_\gamma$).

MS (ESI-) calc. For $C_{15}H_{23}N_4O_{18}P_3S$ 671.99, found [M-H] 671.98854

5.3 Design and synthesis of modified nucleotides bearing 1,1-dioxido-3-oxo-4,4-dimethyl-thiazetidine (4,4-dimethyl-3-oxo-β-sultam) for cross-linking with serine, their enzymatic incorporation and testing of reactivity

4,4-dimethyl-2-(prop-2-yn-1-yl)-1,2-thiazetidin-3-one 1,1-dioxide (37)



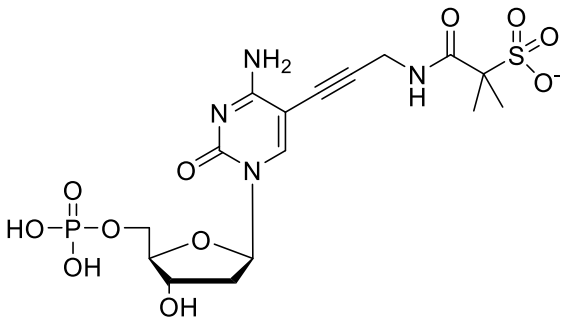
2-(Chlorosulfonyl)-2-methylpropanoyl chloride (**36**) (300 mg, 4.9 mmol) was dissolved in dry ether (50 ml) and the mixture was cooled to −78 ℃. A solution of propargylamine (0.16 ml, 4.9 mmol) in ether (3 ml) was added dropwise over 20 minutes at −78 ◦C with stirring. The mixture was stirred for 10 minutes before triethylamine (0.4 ml, 14.7 mmol) in ether (3 ml) was added dropwise over 10 minutes to the mixture at −78 ℃. The reaction mixture was stirred at −78 ◦C for 1 hour and then for 3 hours at room temperature and the resultant $Et_3N{\cdot}HCl$ salt filtered off.

The solvent was removed by reduced pressure rotary evaporation at 30 ℃ to yield a creamy white residue which was purified using column chromatography (30 g silica) (3 : 2 hexane–EtOAc, *R*f = 0.44) (20%). All spectral data are consistent with the published literature.[85]

IR: *v* max (cm−1) ($CHCl_3$) 3287, 1777, 1326, 1176, 1118, 582. ^{1}H NMR: *d* (DMSO): 4.40 (2H, d, CH_2), 3.48 (1H, t, CH), 1.65 (6H, s 2 x CH_3). ^{13}C NMR: *d* (CDCl3): 163.83 (C=O), 83.64 (quaternary carbon), 76.33 (≡CH), 75.63 (quaternary carbon), 29.32 (CH_2), 11.12 (2 x CH_3).
HRMS (ESI-) calcd. For $C_7H_9NO_3S$ 187.03 (exact mass), found 186.02316 [M - H].

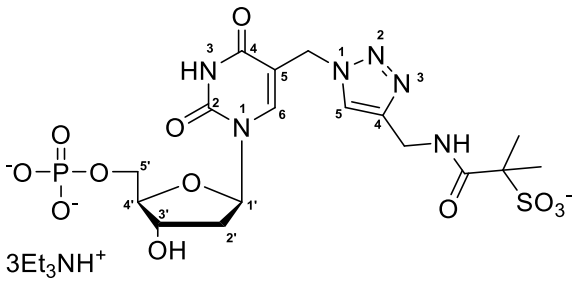
5.3.1 Synthesis of modified monophosphate (dC^{OXOSU}MP, dU^{OXOSU}MP)

5-(3-{[(2,2-dimethyl-1-oxo-2-sulfoxy)prop-1-yn-1-yl]amino)-2'-deoxycytidine-5'-O-monophosphate (dC^{OXOSU}MP) (38a)



Prepared according to *General procedure A* **dC^{OXOSU}MP** was prepared from **dC^{I}MP** (30 mg, 69 µmol), *N*-(propargyl)-3-oxo-b-sultam **37** (7.9 mg, 42 µmol), CuI (0.95 mg, 5 µmol), $Pd(OAc)_2$ (0.56 mg, 2.5 µmol), TPPTS (2.84 mg, 28 µmol) and DIPEA (87 µl, 0.064 mmol). After purification by C18 reverse phase HPLC using water/methanol (5 to 100 %) containing 0.1M TEAB buffer as eluent **dC^{OXOSU}MP** was isolated but only in opened ring form as shown from IR and mass.
IR: *v* max (cm−1) (net): 3398, 2677, 1653, 1506, 1398, 1193, 1036
MS (ESI-) calcd. For $C_{16}H_{22}N_4O_{11}PS^-$ 509.07 (exact mass), found [M-H] 509.1

Synthesis of 5-{[1H-1,2,3-triazol-4-yl)methyl)amino)-2-methyl-1-oxopropyl-2-sulfonate]-2'deoxyuridine-5'-O-monophosphate (dU^{OXOSU}MP) (39a)



According to *general procedure C*, **dU^{OXOSU}MP** was prepared from **dU^{N3}MP** (50.0 mg, 138 µmol), **37** (20.0 mg, 106 µmol), sodium ascorbate (4.2 mg, 21 µmol) and copper(II) sulfate pentahydrate (5 mg, 17 µmol). After HPLC purification the poduct was isolated as white solid (48% yield).

^{1}H NMR (500.0 MHz, D_2O, ref(*t*BuOH) = 1.24 ppm): 1.26 (t, 27H, J_{vic} = 7.3, C**H$_3$**CH$_2$N); 1.53 (s, 6H, $(CH_3)_2C$); 2.37 (ddd, 1H, J_{gem} = 14.0, $J_{2'b,1'}$ = 6.3, $J_{2'b,3'}$ = 3.7, H-2′b); 2.44 (ddd, 1H, J_{gem} = 14.0, $J_{2'a,1'}$ = 7.5, $J_{2'a,3'}$ = 6.2, H-2′a); 3.18 (q, 18H, J_{vic} = 7.3, CH_3C**H$_2$**N); 3.93 – 3.98 (m, 2H, H-5′); 4.19 (tdd, 1H, $J_{4',5'}$ = 4.2, $J_{4',3'}$ = 3.6, 3.7, $J_{H,P}$ = 1.0, H-4′); 4.48 (s, 2H, C**H$_2$**NH); 4.58 (dt, 1H, $J_{3',2'}$ = 6.2, 3.7, $J_{3',4'}$ = 3.7, H-3′); 5.32, 5.35 (2 × d, 2 × 1H, J_{gem} = 14.6, CH_2N); 6.34 (dd, 1H, $J_{1',2'}$ = 7.5, 6.3, H-1′); 7.91 (s, 1H, H-5-triazole); 8.25 (s, 1H, H-6).

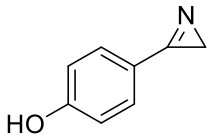
^{13}C NMR (125.7 MHz, D_2O, ref(*t*BuOH) = 32.43 ppm): 11.10 (**C**H_3CH_2N); 24.18 ((**C**H_3)$_2$C); 37.94 (CH_2NH); 41.63 (CH_2-2′); 49.46 (CH_3**C**H_2N); 50.37 (CH_2N); 66.43 (d, $J_{C,P}$ = 4.5, CH_2-5′); 66.64 ((CH$_3$)$_2$**C**); 74.01 (CH-3′); 88.32 (CH-1′); 88.95 (d, $J_{C,P}$ = 8.4, CH-4′); 111.43 (C-5); 126.65 (CH-5-triazole); 144.87 (CH-6); 147.39 (C-4-triazole); 157.23 (C-2); 171.05 (C-4); 175.64 (CONH).

^{31}P{^{1}H} NMR (202.4 MHz, D_2O): 4.56.

MS (ESI-) calc. For $C_{17}H_{22}N_6O_{12}PS_3$- 565.08, found [M-H] found 567.1

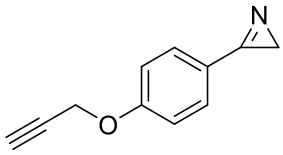
5.4 Design and synthesis of modified nucleotides bearing azirine for cross-linking with glutamic and aspartic acids, enzymatic incorporation and testing of reactivity

4-(2H-azirin-3-yl)phenol (AZ2 linker, 43)



To a solution of the compound **41** (234.0mg, 0.64 mmol) in THF (2.0mL) was added a TBAF solution (1.0 M in THF, 1.3 mL, 1.30mmol). The resulting yellow solution was stirred at room temperature for 30 min and quenched by saturated NH_4Cl solution. After removal of THF in vacuo, the aqueous layer was extracted with ethyl acetate. The combined organic layers were washed with brine and dried over anhydrous Na_2SO_4. After evaporation of solvents, the resulting crude product was used for the next step directly without further purification.

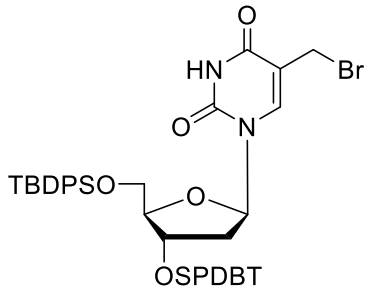
3-(4-(prop-2-yn-1-yloxy)phenyl)-2H-azirine (AZ1 linker, 42)



To a stirred solution of product obtained in the last step in DMF (5.0) mL was added K_2CO_3(115.0mg, 0.83 mmol) and 3-bromopropyne (99.0mg, 0.83 mmol). The mixture was stirred at RT for 4 h followed by addition of water (10.0 mL). The mixture was extracted with ethyl acetate, the combined organic layers were washed with brine and dried over anhydrous Na_2SO_4. After evaporation of the solvent, the residue was purified by flash chromatography to give **AZ1 linker (42)** as a white solid (89% yield). All spectral data are consistent with the published literature.[88]

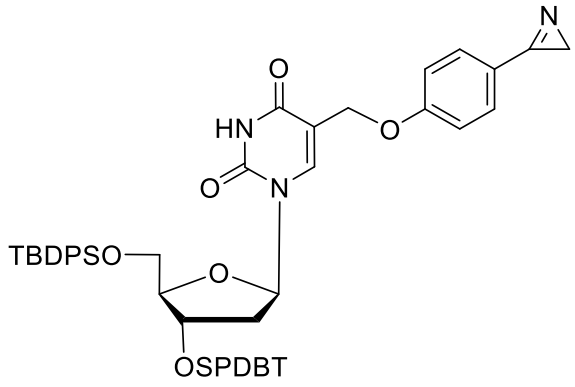
5.4.1 Synthesis of modified dU^{AZ} nucleoside

Synthesis of 5-(bromomethyl)-5'-O-di(terbutyldiphenylsilyl)-2'-deoxyuridine (47)



To a solution of 5'-tertbutyldiphenylsilyl-2'-deoxyuridine (4.0 gr, 9.0 mmol) in benzene was added AIBN (0.224 g, 1.0 mmol) and NBS (2.0 g, 11.0 mmol) and the reaction mixture was stirred for 3 h at reflux at 70 ºC. The reaction was diluted with EtOAc and then extracted with saturated ammonium chloride solution (3x20 mL). All collected organic fractions were washed with brine (3x20 mL). After drying with sodium sulfate and concentration in vacuo, the crude product as yellow/orange foam was used for the next step without any further purification.

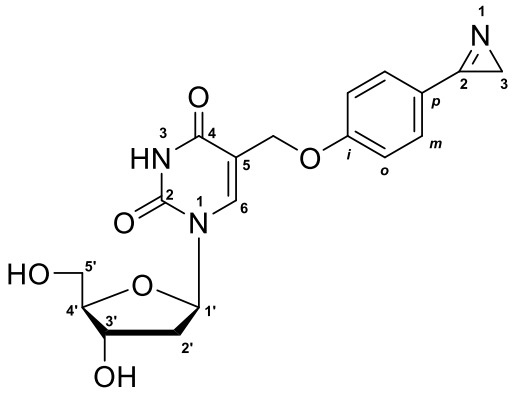
Synthesis of 5-{[4-(2H-azirin-3-yl)phenoxy]methyl}-3',5'-O-di(tertbutyldiphenylsilyl)-2'-deoxyuridine (48)



NaH (75 mg, 60% suspension in oil, 1.25 mmol, 2.50 eq) was dissolved in THF at 0 ºC for 10 min, after evolution of gas was added a solution of **7** (153 mg, 0.50 mmol) in absolute THF and leave stirring at RT for 40 min. Then a solution of compound from the previous step (525 mg, 0.60

mmol, 1.20 eq) in THF was added to reaction mixture and stirred under an inert atmosphere at room temperature for 2 h. The reaction was quenched by the addition of saturated ammonium chloride solution, diluted with EtOAc and washed with brine. After drying with sodium sulfate and concentration in vacuo, the crude product was used for the next step without any further purification.

Synthesis of 5-{[4-(2H-azirin-3-yl)phenoxy]methyl}-2'-deoxyuridine (dU^{AZ}) (49)



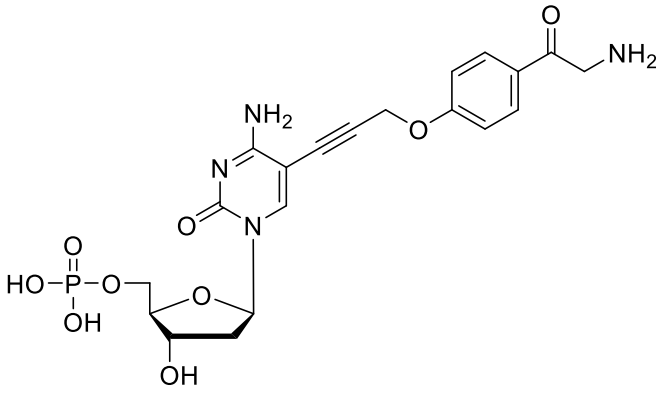
To a solution of **48** (317 mg, 1 eq) in THF (4 mL) was added TBAF (0.54 mL, 2 eq) and the reaction mixture was stirred for 2 h at RT. After concentration in vacuo the crude product was purified by flash column chromatography (DCM/MeOH 8:2) to give the desired compound **55** with 90% yield as a solid yellow.

^{1}H NMR (600.1 MHz, CD_3OD): 1.74 (s, 2H, H-3-azirine); 2.22 (ddd, 1H, J_{gem} = 13.6, $J_{2'b,1'}$ = 7.1, $J_{2'b,3'}$ = 6.2, H-2′b); 2.31 (ddd, 1H, J_{gem} = 13.6, $J_{2'a,1'}$ = 6.2, $J_{2'a,3'}$ = 3.5, H-2′a); 3.73 (dd, 1H, J_{gem} = 12.0, $J_{5'b,4'}$ = 3.5, H-5′b); 3.80 (dd, 1H, J_{gem} = 12.0, $J_{5'a,4'}$ = 3.1, H-5′a); 3.94 (td, 1H, $J_{4',3'}$ = 3.5, $J_{4',5'}$ = 3.5, 3.1, H-4′); 4.40 (dt, 1H, $J_{3',2'}$ = 6.2, 3.5, $J_{3',4'}$ = 3.5, H-3′); 4.88, 4.90 (2 × dd, 2 × 1H, J_{gem} = 11.4, ^{4}J = 0.9, CH_2O); 6.29 (dd, 1H, $J_{1',2'}$ = 7.1, 6.2, H-1′); 7.21 – 7.24 (m, 2H, H-*o*-phenylene); 7.85 – 7.88 (m, 2H, H-*m*-phenylene); 8.29 (t, 1H, ^{4}J = 0.9, H-6).

^{13}C NMR (150.9 MHz, CD_3OD): 19.77 (CH_2-3-azirine); 41.62 (CH_2-2′); 62.70 (CH_2-5′); 64.24 (CH_2O); 72.09 (CH-3′); 86.72 (CH-1′); 89.06 (CH-4′); 110.54 (C-5); 116.77 (CH-*o*-phenylene); 119.10 (C-*p*-phenylene); 132.65 (CH-*m*-phenylene); 142.15 (CH-6); 152.03 (C-2); 164.09 (C-*i*-phenylene); 164.96 (C-4); 165.72 (C-2-azirine).

HRMS (ESI+) calcd. for $C_{18}H_{19}N_3O_6$ 373.13, found [M + Na] 396.11684, [M + H] 374.13466

Synthesis of 5-[3-(4-acetylphenoxy)prop-1-yn-1-yl]-2'-deoxycytidine-5'-*O*-monophosphate (50)



Prepared according to *General procedure A* from **dC^{I}MP** (50 mg, 0.1 mmol), **AZ1 linker 42** (20 mg, 0.1 mmol), CuI (1.62 mg, 10 μmol), $Pd(OAc)_2$ (1.12 mg, 5.0 μmol), TPPTS (5 mg, 10 μmol) and DIPEA (0.13 mL, 0.84 mmol). After purification by C18 reverse phase HPLC using water/methanol (5 to 100 %) containing 0.1M TEAB buffer as eluent and the compound **50** was isolated as solid pale yellow instead of desired compound **dU^{AZ}MP (51)**.
MS (ESI-) calcd. for $C_{20}H_{22}N_3O_9P$ 479.11 (exact mass), found 478.11 [M - H].

5.5 Design and synthesis of modified nucleotides bearing 9-hydroxy-9-azabicyclo[3.3.1]nonane (keto-ABNOH) for cross-linking with tryptophan, enzymatic incorporation and testing of reactivity

General remarks for the chemistry part - keto-ABNOH project

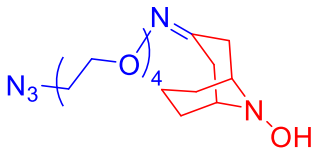
^{1}H, ^{13}C and ^{31}P NMR spectra were acquired on a Bruker AVANCE IIIHD 600 (^{1}H at 600.1 MHz, ^{13}C at 150.9 MHz) and Bruker AVANCE IIIHD 500 (^{1}H at 500.0 MHz, ^{13}C at 125.7 MHz, ^{31}P at 202.4 MHz), spectrometers, as indicated. ^{1}H and ^{13}C resonances were fully assigned using H,H-COSY, H,H-ROESY, H,C-HSQC and H,C-HMBC techniques. All chemical shifts are quoted on the δ scale. NMR spectra measured in D_2O were referenced to the signal of *t*-BuOH (10 % v/v solution in D_2O, 1 drop) as the internal standard. ^{31}P NMR spectra were referenced to H_3PO_4 signal (0 ppm) as the external standard. Coupling constants (*J*) are reported in Hz with the following splitting abbreviations: s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet. High

resolution mass spectra were measured on LTQ Orbitrap XL (Thermo Fisher Scientific) using Electrospray Ionization Technique (ESI). Reactions were monitored by thin layer chromatography (TLC) on TLC Silica gel 60 F254 (Merck) and detected by UV (254 nm) or by solution of 4-anisaldehyde 3.6 % v/v in ethanol and 10% v/v of sulphuric acid. Reactions were monitored by Advion Expression Compact Mass Spectrometer connected with Plate Express® TLC Plate Reader using ESI. High performance flash chromatography (HPFC) was performed with Biotage SP1 apparatus on DEAE SEPHADEX A-25 sodium form columns or with ISCO Combiflash Rf system on RediSep Rf Gold Silica Gel Disposable columns. Purification of nucleoside triphosphates and monophosphates was performed using HPLC (Waters modular HPLC system) on a Kinetex column (RP 18, 5µm EVO C18 100 LC-Column 250x21.2 mm) or on X-Bridge-prep (RP 18, 5µm OBD 19x150 mm). Reactions were analysed by TLC using IPAV (isopropyl alcohol /NH_4OH /H_2O, ratio 11/7/2) as mobile phase. All materials were purchased from commercial suppliers and used without further purification unless otherwise stated. $POCl_3$ and $PO(OMe)_3$ were distilled prior to use. The water used in synthetic part was of HPLC quality. Chemicals were of analytical grade.

All the EPR spectra were acquired on *Bruker* EMX^{plus} 10/12 CW (continuous wave) spectrometer equipped with a Premium-X-band microwave bridge and the standard rectangular probehead/cavity (ER160FCQ, *Bruker*). The g_{iso} value of radicals was determined using a built-in spectrometer frequency counter and a ER036TM NMR-Teslameter (*Bruker*). Typical uncertainty for the g-value determination reads ± 0.0002. Compounds were dissolved either in water or $CHCl_3$ and sampled into glass $50\,\mu L$ capillaries ($i.d. = 0.78\,\mathrm{mm}$, *Hirschmann Ringcaps*)(1). Typical EPR recording included following parameter setup: microwave frequency $9.807\,\mathrm{GHz}$, microwave power $2.5\,\mathrm{mW}$, central filed $349\,\mathrm{mT}$, sweep width $(16-30)\,\mathrm{mT}$, resolution (number of points) $1500-3000$, modulation amplitude $(0.02-0.12)\,\mathrm{mT}$, number of scans $10-150$. Kinetic measurements were performed by the preset 2D_Field_Delay experiment. The time corresponding to each EPR spectrum in the kinetic series was corrected in order coincide with the midpoint (B) of the spectrum because the preset takes the starting time of each sweep (at which the CW EPR spectrum actually does not appear). Calculations: Concentration of aminoxyl-based radicals was determined by double integration of the corresponding EPR spectra directly within the *Xenon* acquisition-processing EPR software (*Bruker)*. Prior to integration baseline correction was applied to each spectrum/sweep. The

calculation procedure takes into account the geometry and calibration of the applied EPR probehead/cavity as well as all additional (instrumental + total spin quantum number S of investigated system, in our case $S = 1/2$) parameters required for absolute determination of paramagnetic centers/aminoxyl radicals with the sample.

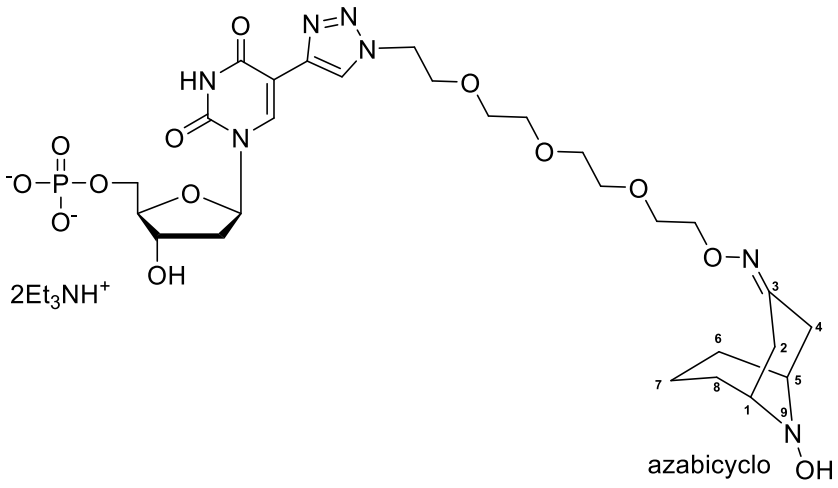
Synthesis of ABNOH-PEG$_4$-N$_3$ (58)



To a solution of 2-(2-(2-(2-(2-azidoethoxy)ethoxy)ethoxy)ethoxy)isoindoline-1,3-dione (82 mg, 0.216 mmol) in dry ethanol (11.2 mL) was added hydrazine monohydrate (12.5 µL, 0.257 mmol) and the resulting mixture was stirred at room temperature for 26 h. The resulting white precipitate was filtrated off and the filtrate was evaporated. To the residue was added hexane/$CHCl_3$ (1:1) and the generated white precipitate was filtrated off and the filtrate was evaporated (repeated twice) to the formation of a colorless oil. To a solution of obtained oil in degassed MeOH (11.3 mL) was added ABNOH (38.43 mg, 248 mmol) and AcOH (26 µL, 0.455 mmol) and the resulting mixture was stirred for 16 h at 50 °C. The reaction solution was concentrated under reduced pressure and the residue was purified by silica gel column chromatography CH_2Cl_2/EtOAc (3:2) to EtOAc/MeOH (5:1) to yield the desired product **ABNOH-PEG$_4$-N$_3$ (58)** as a colourless oil (20.5 mg, 26 % yield over 2 steps). The data were in accordance with the published literature.[93]

5.5.1 Synthesis of modified monophosphate (dU^{ABNOH}MP) and triphosphate (dU^{ABNOH}TP)

Synthesis of dU^{ABNOH}MP (61)



dU^{E}MP (15 mg, 0.037 mmol) was dissolved in a pressure tube in 0.3 mL DMSO (previously degassed). ABNOH-PEG$_4$-N$_3$ (22.3 mg, 0.06 mmol) was dissolved in a second pressure tube in 0.3 mL of DMSO (previously degassed). In a third pressure tube $CuSO_4$*5H_2O (3.74 mg, 0.02 mmol) and tris(benzyltriazolylmethyl)amine (TBTA, 7.95 mg, 0.015 mmol and sodium ascorbate (7.3 mg, 0.036mmol) were dissolved in 0.5 mL of degassed DMSO. The first and the second solution were added to the third dropwise almost at the same time. The reaction was stirred at 25 °C for 1.10 h and subsequently stopped by the addition of 0.1 mL of 5 M EDTA (ethylenediaminetetraacetic acid tetrasodium salt dihydrate) solution. The product was purified by RP-HPLC on X-Bridge-prep column, with use of linear gradient of 0.1 M TEAB (triethylammonium bicarbonate in H_2O) to 0.1 M TEAB in H_2O/MeOH (1:1) to MeOH as eluent. Several co-distillations with water followed by freeze-drying from water gave the product **dU^{ABNOH}MP** as white powder (10.5 mg, 35% yield). Rf (IPAV) = 0.8.

^{1}H NMR (500.0 MHz, D_2O, ref (ext. *t*BuOH) = 1.24 ppm): 1.28 (t, 18H, J_{vic} = 7.3, C**H$_3$**CH$_2$N); 1.34, 1.47 (2 × m, 2 × 1H, H-7-azabicycloo); 1.74 (m, 1H, H-6b-azabicyclo); 1.82 – 1.98 (m, 3H, H-6a,8-azabicyclo); 2.21 (d, 1H, J_{gem} = 16.1, H-2b-azabicycloo); 2.41 – 2.51 (m, 2H, H-2′); 2.61

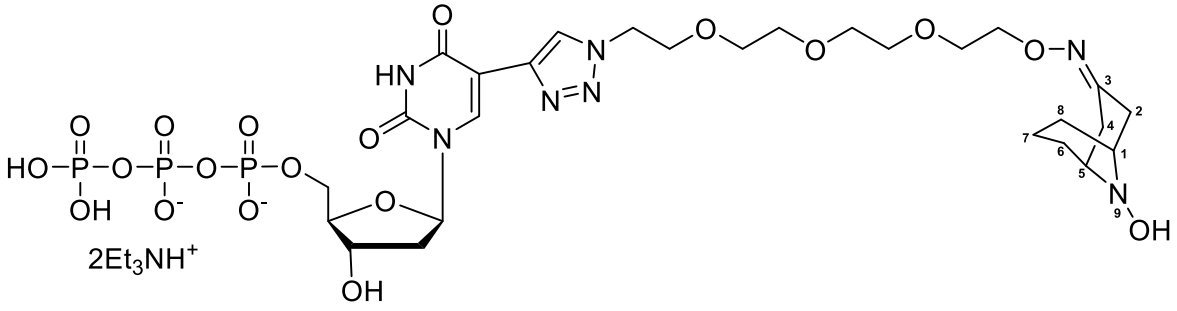
(dd, 1H, J_{gem} = 16.8, $J_{4b,5}$ = 5.9, H-4b-azabicyclo)o; 2.92 (dd, 1H, J_{gem} = 16.8, $J_{4a,5}$ = 4.0, H-4a-azabicyclo); 3.01 (m, 1H, H-2a-azabicycloo)); 3.20 (q, 12H, J_{vic} = 7.3, $CH_3C\mathbf{H_2}N$); 3.51 (m, 1H, H-5-azabicyclo); 3.54 – 3.71 (m, 11H, H-1-azabicyclo, $OC\mathbf{H_2}CH_2ON$, $OC\mathbf{H_2}C\mathbf{H_2}O$); 3.97 – 4.03 (m, 2H, $OC\mathbf{H_2}CH_2N$); 4.09 (ddd, 1H, J_{gem} = 11.4, $J_{H,P}$ = 5.1, $J_{5'b,4'}$ = 3.9, H-5′b); 4.10 – 4.18 (m, 3H, H-5′a, $OCH_2C\mathbf{H_2}ON$); 4.25 (m, 1H, H-4′); 4.62 (dt, 1H, $J_{3',2'}$ = 6.2, 3.3, $J_{3',4'}$ = 3.3, H-3′); 4.63 – 4.72 (m, 2H, $OCH_2C\mathbf{H_2}N$); 6.38 (dd, 1H, $J_{1',2'}$ = 7.6, 6.4, H-1′); 8.43 (s, 1H, H-5-triazole); 8.47 (s, 1H, H-6).

^{13}C NMR (125.7 MHz, D_2O, ref(ext. *t*BuOH) = 32.43 ppm): 11.06 ($C\mathbf{H_3}CH_2N$); 17.41 (CH_2-7-azabicyclo); 26.53 (CH_2-4-azabicyclo); 30.87 (CH_2-2-azabicyclo); 33.10 (CH_2-8-azabicyclo); 33.66 (CH_2-6-azabicyclo); 42.10 (CH_2-2′); 49.51 ($CH_3C\mathbf{H_2}N$); 53.11 ($N\mathbf{C}H_2CH_2O$); 59.82 (CH-5-azabicyclo); 60.44 (CH-1-azabicyclo); 67.60 (d, $J_{C,P}$ = 4.7, CH_2-5′); 71.53 ($NCH_2\mathbf{C}H_2O$); 71.81, 72.31, 72.43, 72.67, 72.70 ($O\mathbf{C}H_2\mathbf{C}H_2O$, $O\mathbf{C}H_2CH_2ON$); 74.24 (CH-3′); 74.92 ($OCH_2\mathbf{C}H_2ON$); 88.89 (d, $J_{C,P}$ = 8.8, CH-4′); 88.91 (CH-1′); 108.56 (C-5); 127.29 (CH-5-triazole); 140.47 (CH-6); 141.54 (C-4-triazole); 153.68 (C-2); 163.35 (C-3-azabicyclo); 165.63 (C-4).

^{31}P{1H} NMR (202.4 MHz, D_2O): 1.73.

HR/MS (ESI^-) for $C_{27}H_{41}N_7O_{13}P$ 702.25054 $[M-H]^-$ calculated, found 702.25040 $[M-H]^-$.

Synthesis of dU^{ABNOH}TP (62)



~1:1 mixture of undefined isomers

dU^{E}TP (18.2 mg, 0.037 mmol) was dissolved in a pressure tube in 0.3 mL DMSO (previously degassed). ABNO-PEG_4-N_3 (26.5 mg, 0.07 mmol) was dissolved in a second pressure tube in 0.3 mL of DMSO degassed. In a third pressure tube $CuSO_4$*$5H_2O$ (4.45 mg, 0.018 mmol) and tris(benzyltriazolylmethyl)amine (TBTA 9.46 mg, 0.018 mmol) and sodium ascorbate (8.7 mg,

0.044 mmol) were dissolved in 0.6 mL of degassed DMSO. The first and the second solutions were added to the third dropwise almost at the same time. The reaction was stirred at 25 °C for 3 h and subsequently stopped by the addition of 0.1 mL of 5 M EDTA (ethylenediaminetetraacetic acid tetrasodium salt dihydrate) solution. The product was purified by RP-HPLC on X-Bridge-prep column, with use of linear gradient of 0.1 M TEAB (triethylammonium bicarbonate in H_2O) to 0.1 M TEAB in H_2O/MeOH (1:1) to MeOH as eluent. Several co-distillations with water followed by freeze-drying from water gave the product as white powder **dU^{ABNOH}TP** (3 mg, 10% yield). Rf (IPAV) = 0.3. According to NMR analysis the product was isolated as ~1:1 mixture of undefined isomers which was further investigated (see section 3.4.1.2).

^{1}H NMR (600.1 MHz, D_2O, ref(*t*BuOH) = 1.24 ppm): 1.27 (t, 36H, J_{vic} = 7.3, C**H$_3$**CH$_2$N); 1.33 – 1.40 (m, 2H, H-7b- azabicycloo); 1.44 – 1.52 (m, 2H, H-7a-ABN); 1.73 – 1.79, 1.85 – 2.06 (2 × m, 8H, H-6,8-ABN); 2.24 – 2.32 (m, 2H, H-2b/4b-ABN); 2.41 – 2.53 (m, 4H, H-2′b); 2.54 – 2.66 (m, 2H, H-2b/4b-ABN); 2.94 – 3.09 (m, 4H, H-2a,4a-ABN); 3.20 (q, 24H, J_{vic} = 7.3, CH$_3$C**H$_2$**N); 3.55 – 3.78 (m, 28H, H-1,5-ABN, OC**H$_2$**C**H$_2$**O); 3.95 – 4.03 (m, 4H, OC**H$_2$**CH$_2$N); 4.13 – 4.19 (m, 4H, OC**H$_2$**CH$_2$O); 4.20 – 4.26 (m, 6H, H-4′,5′); 4.62 – 4.69 (m, 3H, H-3′, OCH$_2$C**H$_2$**N); 6.31 – 6.37 (m, 2H, H-1′); 8.34, 8.35 (2 × s, 2 × 1H, H-6); 8.40 (s, 2H, H-5-triazole).

^{13}C NMR (150.9 MHz, D_2O, ref(*t*BuOH) = 30.29 ppm): 8.92 (**C**H$_3$CH$_2$N); 14.90 (CH$_2$-7-ABN); 24.28, 24.30, 28.67, 28.70 (CH$_2$-2,4-ABN); 30.38, 30.43, 30.93, 30.95 (CH$_2$-6,8-ABN); 39.19, 39.29 (CH$_2$-2′); 47.36 (CH$_3$**C**H$_2$N); 50.84, 50.85 (N**C**H$_2$CH$_2$O); 57.82, 57.47 (CH-1,5-ABN); 66.21, 66.26 (2 × d, $J_{C,P}$ = 5.8, CH$_2$-5′); 69.48, 69.51 (O**C**H$_2$CH$_2$N); 69.67, 70.20, 70.23, 70.33, 70.58, 70.60 (OCH$_2$CH$_2$O); 71.37, 71.42 (CH-3′); 72.94 (OCH$_2$CH$_2$O); 86.17, 86.26 (2 × d, $J_{C,P}$ = 8.8, CH-4′); 86.43, 86.64 (CH-1′); 106.57, 106.66 (C-5); 125.11, 125.16 (CH-5-triazole); 138.75, 138.83 (CH-6); 139.57, 139.60 (C-4-triazole); 151.57, 151.58 (C-2); 159.50, 159.62 (C-3-ABN); 163.54, 163.60 (C-4).

^{31}P{^{1}H} NMR (202.4 MHz, D_2O): -22.15, -22.08 (2 × t, J = 19.3, P$_\beta$); -10.67 (d, J = 19.3, P$_\alpha$); -9.25, -9.11 (2 × d, J = 19.3, P$_\gamma$).

HR/MS (ESI$^-$) for $C_{27}H_{43}N_7O_{19}P_3$ 862.18320 $[M-H]^-$ calculated, found 862.18165 $[M-H]^-$.

5.5.2 Model reactions

Model reaction with keto-ABNO and Tryptophan

Keto-ABNO (20 mg, 0.129 mmol) and N-Ac-Trp (32 mg, 0.129 mmol) were stirred in 2.5 mL H_2O for 1.5 h at 25 °C. The reaction was purified by RP-HPLC on X-Bridge-prep column, with use of linear gradient of 0.1 M TEAB (triethylammonium bicarbonate in H_2O) to 0.1 M TEAB in H_2O/MeOH (1:1) to MeOH as eluent.

Model reaction of cross linking between keto-ABNO with tripeptide Ac-AlaTrpAla (Pept1)

9-oxyl-9-azabicycloo[3.3.1]nonan-3-one, keto-ABNO (20 mg, 0.17 mmol) was dissolved in 1.5 mL H_2O/$NaHCO_3$ buffer (2:1). Subsequently, **Pept1** (37 mg, 0.095 mmol) was dissolved in 0.5 mL $NaHCO_3$ buffer pH 10, and added to keto-ABNO solution. The whole reaction mixture was kept stirring at 37 °C for 3 h and then purified on RP-HPLC (XBridge C18, 5 µm
19 x 250 mm) with use of linear gradient of 0.1 M TEAB (triethylammonium bicarbonate in H_2O) to 0.1 M TEAB in H_2O/MeOH (1:1) to MeOH as eluent. Several co-distillations with water followed by freeze-drying from water gave a mixture of several products, Rf (IPAV) = 0.2, stained in $KMnO_4$.

Model reaction between dU^{ABNOH}MP and Ac-Trp or Pept1, isolation and stability study of the products via LC-MS

The reaction mixture (∑V = 150 µL) was prepared by mixing of **dU^{ABNOH}MP** (10 µL, 10 mM aqueous solution) with either **Ac-Trp** (50 µL, saturated aqueous solution) or **Pept1** (10 µL, 250 mM aqueous solution). The reaction mixture was completed to a final volume of 50 µL with 50 mM $NaHCO_3$ buffer pH 10 (50 µL) and further incubated at 37 °C for 18 h. The crude mixture was directly injected in UPLC and analyzed via LC-mass (**Figure 63, A-B**, section 3.4.1.5.5). For the conjugated product **dU^{ABNO_Trp}MP** was observed almost full conversion, only traces of **dU^{ABNO}MP** from LC-MS analysis of crude mixture. **dU^{ABNO_Pept1}MP** was observed in 17% conversion and additionally the hydrated form **dU^{ABNO_Pept1}MP+H_2O** in 9% conversion.
The corresponding product of either **dU^{ABNO_Trp}MP** or **dU^{ABNO_Pept1}MP** was purified by HPLC (Luna Omega 5 µM Polar C18 100 A 150 x 4.6 mm, with use of linear gradient of 0.1 M TEAB (triethylammonium bicarbonate in H_2O) to 0.1 M TEAB in H_2O/MeCN (95:5) to MeCN as eluent.

The MS spectra were measured on UPLC Agilent 1290 Infinity II bio system equipped with DAD and LC/MSD XT MS (operating in negative mode) detectors, with bioZen™ Oligo LC column (1.7 μm, 50×2.1 mm; Phenomenex®) using a linear gradient of 12 mM Et_3N and 300 mM HFIP in H_2O to 12 mM Et_3N and 300 mM HFIP in MeOH in 30 min.

Stability studies of dU^{ABNO_Trp}MP or dU^{ABNO_Pept1}MP in H_2O at 37 °C for 18 h

dU^{ABNO_Trp}MP or **dU^{ABNO_Pept1}MP** were prepared according to procedure shown above in the paragraph 5.6.2.
The HPLC-purified product **dU^{ABNO_Trp}MP** was lyophilized and then re-dissolved in 50 μL of H_2O and incubated at 37 °C for 18 h, followed by characterization by LC-MS (**Figure 64**, section 3.4.1.5.5).

MS/MS analysis of dU^{ABNOPept1}MP fragmentation

dU^{ABNO_Pept1}MP was prepared according to the procedure shown above in the paragraph 5.4.12.
The solution of reaction mixture was mixed with 100 μL of 0.1% formic acid and desalted on a peptide reverse-phase microtrap column (Optimize Technologies, USA) with elution by 80% acetonitrile, 0.1% formic acid. SolariX XR mass spectrometer equipped with 15 T magnet (Bruker Daltonics), which was externally calibrated using sodium trifluoroacetate clusters to achieve sub-ppm mass accuracy. The analyte ions were introduced into the spectrometer at 2.5 μl/min through an ESI source operated in positive ion polarity. The time-of-flight was set to 0.6 ms, transfer RF frequency of collision cell and transfer optics were set to 2 MHz and 6 MHz, respectively, with the collision voltage −0.5 V for ion thermalization, drying gas temperature was 200 °C, and data were acquired in broadband magnitude mode (*m/z* 92−2000) with 2 M data point transient (transient length 0.42 s). For MS, ions were accumulated for 0.1 s and 32 scans were summed. For MS/MS, doubly charged ions corresponding to the theoretical mass of **dU^{ABNO_Pept1}MP** were isolated in the quadrupole with the isolation window of 3 Da (**Figure 65**, section 3.4.1.5.6), fragmented by collision induced dissociation with argon at -3.5 V collision energy. and the products accumulated for 5 s prior to MS analysis, during which 64 scans were averaged.

Data analysis

Data were analyzed using Data Analysis 5.3 (Bruker Daltonics) with the use of ChemDraw 22.0.0.22 suite (PerkinElmer) for the annotation of theoretical structures and CID fragments. Elemental formulas for individual precursor and fragment species were directly derived from fitting the isotopic fine structure peaks using Bruker Daltonics SmartFormula algorithm.

General remarks for the biochemical part – ketoABNO project

All gels were analyzed by fluorescence imaging using Typhoon FLA 9500 (GE Healthcare). Mass spectra of oligonucleotides were measured on UltrafleXtreme MALDI-TOF/TOF (Bruker) mass spectrometer with 1 kHz smartbeam II laser or on UPLC Agilent 1290 Infinity II bio system equipped with DAD and LC/MSD XT MS (operating in negative mode) detectors, with bioZen™ Oligo LC column (1.7 µm, 50×2.1 mm; Phenomenex®) using a linear gradient of 12 mM Et3N and 300 mM HFIP in H2O to 12 mM Et3N and 300 mM HFIP in MeOH in 30 min. Mass spectra were deconvoluted using OpenLAB CDS CS LC/MS Deconvolution Bioanalysis software.[87] UV-Vis spectra were measured at room temperature on NanoDrop1000 (ThermoScientific). Samples were concentrated on CentriVap Vacuum Concentrator system (Labconco). Synthetic oligonucleotides (primers, templates and biotinylated templates; for sequences see Table S1) were purchased from Generi Biotech (Czech Republic). Natural nucleoside triphosphates (dATP, dGTP, dTTP, dCTP) were purchased from Thermo Scientific. Histone human recombinant proteins (H_2A, H_4) were purchased from New England Biolabs. KOD XL DNA polymerase and corresponding polymerase reaction buffer from Merck Millipore, streptavidin magnetic particles were obtained from Sigma Aldrich (Merck), QIAquick Nucleotide Removal Kit QIAGEN (Biotech, Czech Republic) Micro-Bio Spin P-30 (BioRad Czech Republic) and Pur-A-Lyze Maxi Dialysis Kit for gel extraction (Sigma Aldrich Czech Republic). Milli-Q water was used for all experiments. PAGE stop solution used after PEX reactions contains: 95% [v/v] formamide, 0.5 mM EDTA, 0.025% [w/v] bromophenol blue, 0.025% [w/v] xylene cyanol, 0.025% SDS. VPS loading buffer used for protein SDS gels contains: 0.05 M TRIS (pH 6.8), 17% glycerol, 16 mM mercaptoethanol, 3.5% SDS, and bromophenol blue. PageBlue™ protein staining solution was obtained from Thermofisher Scientific. Samples after PEX reaction were analyzed by either 12.5% or 20% PAGE (acrylamide/bisacrylamide 19:1) under denaturing conditions (1 h, 50 °C, 1xTBE buffer).

Conversions were calculated using Image J. analyzer. Samples after conjugation with Histones were analyzed by 17.5% SDS denaturing PAGE (acrylamide/methylenebisacrylamide 29:1; 1.92 M glycine, 0.25 M Tris, 0.1% SDS- 230 V/1 h). Other chemicals were of analytical grade.

5.5.3 Enzymatic synthesis of ABNO- modified DNA

Single incorporation of dU^{ABNOH}TP using 19-mer template - Analytical scale

The reaction mixture (20 µL) contained primer **Prim^{A}_FAM** (4 µM, 1 µL), template **Temp^{19_1T}** (4 µM, 1.5 µL), KOD XL DNA polymerase (0.25 U/µL, 0.5 µL), natural dGTP (4 mM, 0.5 µL), either natural or modified dTTP (4 mM, 1 µL) and enzyme reaction buffer (10X, 2 µL). The reaction mixture was incubated for 30 min at 60 °C in a thermal cycler. The PEX reaction was stopped by addition of PAGE stop solution (20 µL) and heated for 5 min at 95 °C. Samples were analyzed by 20% denaturing PAGE and visualized using florescence imaging (**Figure 66**-left, section 3.4.1.6).

Multiple incorporation of four modified dU^{ABNOH}TP using 31-mer template - Analytical scale

The reaction mixture (20 µL) contained primer **Prim^{A}_FAM** (3 µM, 1 µL), template **Temp^{31_4T}** (3 µM, 1.5 µL), KOD XL DNA polymerase (0.25 U/µL, 0.5 µL), natural dNTPs (dGTP, dATP, dCTP; 4 mM, 0.5 µL), either natural or modified dTTP (4 mM, 1 µL) in enzyme reaction buffer (10 X, 2 µL). The reaction mixture was incubated for 30 min at 60°C in a thermal cycler. The PEX reaction was stopped by addition of PAGE stop solution (20 µL) and heated for 5 min at 95 °C. . Samples were analyzed by 20% denaturing PAGE and visualized using florescence imaging (**Figure 66-**right).

Single incorporation of dU^{ABNOH}TP using 31-mer template-Analytical scale

The reaction mixture (20 µL) contained KOD XL DNA polymerase (0.25 U/µL, 0.2 µL), primer **Prim^{31a}_FAM** or **Prim^{31b}_FAM** (3 µM, 1 µL), template **Temp^{31a}** or template **Temp^{31b}** (3 µM, 1.5 µL), dNTPs (either natural or modified; 4 mM, 0.5 µL each) in KOD XL reaction buffer (10X, 2.1 µL). The reaction mixture was incubated in a thermal cycler for 30 min at 60 °C and stopped by cooling to 4 °C.

5`-(6-FAM)-labelled PEX products were purified using QIAquick Nucleotide Removal Kit using

the standard procedure. The PEX reaction was stopped by addition of PAGE stop solution (20 µL) and heated for 3 min at 95°C. Samples were analyzed by 20% denaturing PAGE and visualized using florescence imaging (**Figure 66-**centre).

Single incorporation of dU^{ABNOH}TP using 19-mer template-Semipreparative scale

The reaction mixture (50 µL) contained KOD XL DNA polymerase (0.25 U/µL, 0.85 µL), primer **Prim^{A}** or primer **Prim^{A}_FAM** (100 µM, 1.7 µL), template **Temp^{19_1T}**, template **Temp^{19_1T}_bio**, natural dGTP (4 mM, 2.38 µL), either natural or modified dTTP (4 mM, 2.38 µL) in KOD XL reaction buffer (10X, 5 µL). The reaction mixture was incubated in a thermal cycler for 40 min at 60 °C and stopped by cooling to 4° C.

Single incorporation of dU^{ABNOH}TP using 31-mer template-Semipreparative scale

The reaction mixture (20 µL) contained KOD XL DNA polymerase (0.25 U/µL, 0.5 µL), primer **Prim^{31a}_FAM** or **Prim^{31b}_FAM** (50 µM, 1 µL), template **Temp^{31a}** or template **Temp^{31b}** (50 µM, 1.5 µL), dNTPs (either natural or modified; 4 mM, 0.5 µL each) in KOD XL reaction buffer (10X, 2.1 µL). The reaction mixture was incubated in a thermal cycler for 30 min at 60 °C and stopped by cooling to 4 °C.

Multiple incorporation of dU^{ABNOH}TP using 31-mer template with 4 modification - Semipreparative scale

The reaction mixture (50 µL) contained KOD XL DNA polymerase (0.25 U/µL, 0.3 µL), primer **Prim^{A}** or primer **Prim^{A}_FAM** (100 µM, 2 µL), template **Temp^{31_4T}** or template **Temp^{31_4T}_bio** (100 µM, 2 µL), dNTPs (either natural or modified; 4 mM, 0.375 µL each) in KOD XL reaction buffer (10X, 5 µL). The reaction mixture was incubated in a thermal cycler for 40 min at 60 °C and stopped by cooling to 4 °C.

A) Biotinylated PEX products **19DNA^{ABNOH}_bio or 31DNA^{4T_ABNOH}_bio**, **31aDNA^{ABNOH}_bio, 31bDNA^{ABNOH}_bio** were purified using DB Streptavidin magneto-separation procedure (see paragraph 5.2.3) to generate single-stranded products **19ON^{ABNOH} or 31ON^{4T_ABNOH}** and then analyzed by MALDI-TOF mass spectrometry (**Table 14**).

B) 5`-(6-FAM)-labelled PEX products **19DNA^{ABNOH}_FAM, 31aDNA^{ABNOH}_FAM, 31bDNA^{ABNOH}_FAM, 31aDNA^{Natural}_FAM**, **31bDNA^{Natural}_FAM** were purified using QIAquick Nucleotide Removal Kit using the standard procedure @quiagen (https://www.qiagen.com/us) and then analyzed by MALDI-TOF mass spectrometry or by LC-MS (**Table 14**).

5.5.4 Bio-conjugation reactions

Model reaction of 19DNA^{ABNOH}_FAM with Ac-Trp with and without use of $NaNO_2$ PAGE analysis

19DNA^{ABNOH}_FAM was prepared by PEX in semi-preparative scale as described above (Section 5.5.3). The reaction mixture (5 µL) containing **19DNA^{ABNOH}_FAM** (1 µM) and **Ac-Trp** (250 mM) in AcOH (1 mM) and $NaNO_2$ (13 µM) was incubated for 1 h at 25 °C in a thermal cycler and subsequently stopped by addition of PAGE stop solution (5 µL) and heated 5 min at 95 °C. Samples were analyzed by 20% denaturing PAGE and visualized using florescence imaging (**Figure 67**).

Reaction of 19DNA^{ABNOH}_FAM with Ac-Trp without use of $NaNO_2$ in H_2O and AcOH 1mM at different incubation temperature– PAGE analysis

19DNA^{ABNOH}_FAM was prepared by PEX in semi-preparative scale as described above (Section 5.5.3). The reaction mixture (5 µL) containing **19DNA^{ABNOH}_FAM** (1 µM) and **Ac-Trp** (250 mM) in AcOH (1 mM) or in H_2O was incubated for 1 h at 25 °C, 37°C and 55°C in a thermal cycler and subsequently stopped by addition of PAGE stop solution (5 µL) and heated 5 min at 95 °C. Samples were analyzed by 20% denaturing PAGE and visualized using florescence imaging (**Figure 68**).

Reaction of 19DNA^{Nat}_FAM and 19DNA^{ABNOH}_FAM with Ac-Trp with and without $NaNO_2$ at different concentrations [1 µM-100 mM] in AcOH 1 mM– PAGE analysis

19DNA^{Nat}_FAM and **19DNA^{ABNOH}_FAM** were prepared by PEX in semi-preparative scale as described above (Section 5.5.3). The reaction mixture (10 µL) containing **19DNA^{Nat}_FAM** (1 µM) was incubated with and without $NaNO_2$ (10 mM) and **Ac-Trp** (250 mM) in AcOH (1 mM) for 1 h at 25 °C in a thermal cycler. Instead the reaction mixture (10 µL) containing **19DNA^{ABNOH}_FAM**

(1 μM) and **Ac-Trp** (250mM) with and without $NaNO_2$ (1 μM- 100 mM) in AcOH (1 mM) was incubated for 1 h at 25 °C in a thermal cycler and subsequently stopped by addition of PAGE stop solution (10 μL) and heated 5 min at 95 °C. Samples were analyzed by 20% denaturing PAGE and visualized using florescence imaging (**Figure 69**).

Reaction of 19DNA^{Nat}_FAM and 19DNA^{ABNOH}_FAM with Ac-Trp and other Trp-containing peptides at different conditions – PAGE analysis

19DNA^{ABNOH}_FAM were prepared by PEX in semi-preparative scale as described above (Section 5.5.3). The reaction mixture (10 μL) containing **19DNA^{ABNOH}_FAM** (1 μM) and **Ac-Trp** (250 mM) or 250 mM of peptides (**Pept 2,3,4,5,6** see structures in Section 3.4.1.5.1) in AcOH, H_2O, $NaHCO_3$ buffer or NH_4Ac buffer was incubated for 1 h at 37 °C in a thermal cycler and subsequently stopped by addition of PAGE stop solution (10 μL) and heated 5 min at 95 °C. Samples were analyzed by 20% denaturing PAGE and visualized using florescence imaging (**Figure 70**).

Reaction of 19DNA^{ABNOH} with Trp or Trp-containing peptides

19DNA^{ABNOH} was prepared by PEX in semi-preparative scale as described above (Section 5.5.3). The reaction mixture (100 μL) containing **19DNA^{ABNOH}** or **19DNA^{ABNOH}_bio** (1 μM) and **Ac-Trp** (sat. sol.) or **Pept1** or **Pept2** 250 mM in 50 mM $NaHCO_3$ buffer pH 10, was incubated in presence or absence of $NaNO_2$ for 18 h at 37 °C in a thermal cycler. Samples were analyzed by 20 % denaturing PAGE and visualized using florescence imaging (**Figure 71**).

Reaction of 31aDNA^{ABNOH} with N-acetyl-Trp and N-acetyl-Trp-containing peptides in presence or absence of $NaNO_2$– PAGE analysis and LC-MS analysis

31aDNA^{ABNOH}, **31aDNA^{ABNOH}_FAM**, **31bDNA^{ABNOH}**, **31bDNA^{ABNOH}_FAM** were prepared by PEX in semi-preparative scale as described above (Section 5.5.3). Products were purified by QIAquick Nucleotide Removal Kit according to standard protocol and then analyzed by LC-MS. The reaction mixture (50 μL), containing **31aDNA^{ABNOH}** or **31aDNA^{ABNOH}_FAM** or **31bDNA^{ABNOH}** or **31bDNA^{ABNOH}_FAM** (7 μM, 10 μL) and **Ac-Trp** (saturated solution) or **Pept1** or **Pept2** or **Pept4** (250 mM, 10 uL for sequences see **Table 13**) in 30 μL $NaHCO_3$ buffer (0.05

M, pH 10) was incubated in presence or absence of $NaNO_2$ in a thermal cycler at 37 °C for 18 h and subsequently stopped by addition of PAGE stop solution (10 µL) and heated 5 min at 95 °C. Samples were analyzed by 20 % denaturing PAGE and visualized using florescence imaging **Figures 72-73**. Subsequently, reaction mixtures were purified by magnetoseparation as described above and then analyzed by LC-MS mass spectrometry (See **Table 16**).

Stability studies and reactivity studies of 31aDNA^{ABNO_Pept1} and 31bDNA^{ABNO_Pept1}

6-FAM-labelled **31aDNA^{ABNOH}_FAM** and **31bDNA^{ABNOH}_FAM** were prepared by PEX in semi-preparative scale as described above in section 5.5.3 and purified by QUIAGEN columns according standard protocol. The reaction mixture (50 µL) containing **31aDNA^{ABNOH}_FAM** or **31bDNA^{ABNOH}_FAM** (7 µM) and **Pept1**, in $NaHCO_3$ buffer (0.05 M, pH 10) was incubated for 18 h at 37 °C in a thermal cycler. The crude mixture was mixed with PAGE stop solution (see General remarks) and loaded on 20% preparative denaturing PAGE under denaturing conditions (1 h, 50 °C, 1X TBE buffer) and visualized using florescence imaging (**Figure 74**). After the run, a modified strand and a template were visualized by UV lamp, and the area containing FAM-modified product was cut out and gel extraction was performed twice (**Figure 75**). In order to extract the product, Pur-A-Lyze Maxi Dialysis Kit was used. The cut-out piece of gel was placed into a prehydrated Pur-A-Lyze Maxi Dialysis Kit containing 3 mL of 1X TAE buffer. Further, the dialysis kit was placed into a horizontal electrophoretic system in a way that its membrane was in a perpendicular position to the current direction. Then, current was applied (45 min, 100 V) and the product was eluted from the piece of gel and trapped on the dialysis column membrane. To detach DNA molecules from the dialysis column membrane, the current was applied in the reverse direction (1 min, 100 V). The purified conjugate product was desalted from the buffer by washing 3 times with 380 µL of H_2O, using Amicon® Ultra 0.5 mL-3 kDa MWCO centrifugal filters and concentrated to the volume of H_2O 30 µL. The product was incubated for 20 h and 168 h at 37 °C to test the stability. After incubation the PAGE stop solution (10 µL) was added to samples and heated for 5 min at

95 °C. Samples were separated with denaturing 20% PAGE (TBE 1X, 42 mA, 2 h) and visualized using florescence imaging (**Figure 76**).

Cross-linking of 31aDNA^{ABNOH}_FAM or 31bDNA^{ABNOH}_FAM and *bs*GnRT protein-(SDS analysis)

Preparation of GntR from *Bacillus subtilis* (*bs*GntR)

*bs*GntR (amino acid residues 2 – 243) was cloned into pET151/D-TOPO expression plasmid (Invitrogen, USA) encoding an N-terminal hexahistidine (His6) tag and the tobacco etch virus protease (TEV PR) recognition site. In the case the His6 tag was cleaved, the six-amino-acid sequence GIDPFT remained at the N-terminus as a cloning artefact.

The recombinant protein was overexpressed in *E. coli* BL21 (DE3) grown on Luria-Bertani broth (Sigma Aldrich, USA) supplemented with 0.4% *(v/v)* glycerol and 100 μg/ml ampicillin. The bacteria were cultured at 37°C until the optical density at 550 nm reached about 1. Then, the gene expression was induced with 0.1 mM ethyl-β-D-thiogalactopyranoside at 18°C, and the cultivation continued for approximately 14 h. Bacterial cells were harvested by centrifugation (4000 rpm /30 min/4 °C), resuspended in the lysis buffer (50 mM Tris·HCl, pH 7.5, 50 mM NaCl, 5% *(v/v)* glycerol, 0.02% *(v/v)* β-mercaptoethanol, 1 mM phenylmethylsulfonyl fluoride, 2 mM $MgCl_2$, and 5 μg/ml deoxyribonuclease I (Sigma Aldrich, USA) in a volume of 5 ml per 1 g of harvested cells. Afterwards, the cells were lysed by sonication. The purification process followed previously published protocols[125,126] lysate was centrifuged (14,000 rpm/40 min/4 °C), and the supernatant was subjected to nickel-chelation chromatography on a His-select Nickel Affinity Gel (Ni-NTA) column (Sigma Aldrich, USA) equilibrated in the loading buffer (50 mM Tris·HCl, pH 7.5, 500 mM NaCl, 5% *(v/v)* glycerol and 0.02% *(v/v)* β-mercaptoethanol). The His6-tagged *bs*GntR was eluted by step elution with the loading buffers supplemented with 30 mM, 200 mM and 350 mM imidazole. The less contaminated fraction eluted with 200 mM imidazole was further incubated with TEV PR in order to cleave off the His6 tag (24-hour dialysis at 4 °C against the loading buffer). The TEV PR and the His6 tag were subsequently separated from the sample by the second run of Ni-NTA chromatography. The purified *bs*GntR solution was dialyzed against the protein storage buffer (20 mM Tris·HCl, pH 7.5, 250 mM NaCl, 0.02% *(v/v)* β-mercaptoethanol), whose suitability was proved by differential scanning fluorimetry (Thermofluor).[127] Finally, the protein sample was concentrated on Amicon Ultra concentrators (Millipore, USA) to a concentration of approximately 9 mg/ml and 15 mg/ml in the case of His6-tag cleaved and uncleaved *bs*GntR, respectively. Before using the protein samples for the

electrophoretic assays, sample aliquots were subjected to preparative size-exclusion chromatography (SEC) and re-concentrated. The final concentration was determined by amino acid analysis.[125,126,127]

Reaction of 31aDNA^{ABNOH} or 31bDNA^{ABNOH} with *bs*GntR protein

6-FAM-labelled **31aDNA^{ABNOH}_FAM** or **31bDNA^{ABNOH}_FAM** and **31aDNA^{Nat}_FAM** or **31bDNA^{Nat}_FAM** were prepared by PEX in semi-preparative scale as described above (Section 5.5.3). **31aDNA^{Natural}_FAM/31aDNA^{ABNOH}_FAM** or **31bDNA^{Nat}_FAM /31bDNA^{ABNOH}_FAM** (50 µM) was incubated with *bs*GntR (50 µM) either in 10 µL of Tris·HCl buffer (0.02 M, pH 7.5), or $NaHCO_3$ buffer (0.05 M, pH 10) at 37 °C for 18 h in presence or absence of $NaNO_2$. The reaction was diluted in 2X VPS loading buffer, denatured for 10 min at 75 °C and analysed by 10 % SDS denaturing PAGE at room temperature (200 V, 50 min). Different ratios of **31aDNA^{Nat}_FAM/31aDNA^{ABNOH}_FAM** to protein (**1:1 Figure 78, 1:4 Figures 79-A, and 1:32 Figure 81-A,B**) or **31bDNA^{Nat}_FAM/31bDNA^{ABNOH}_FAM** to protein (1:4 **Figures 79-B**, **Figure 80**, 1:32 **Figure 81-A**,**B, Figure 82-A,B**) were tested and visible bands of conjugated products were observed by fluorescence imaging according to ladder. Gels comparison in presence or absence of $NaNO_2$ in case of cross-linking between **31aDNA^{Nat}_FAM/31aDNA^{ABNOH}_FAM** and *bs*GntR (**Figure 79-A**) or in case of cross-linking between **31bDNA^{Nat}_FAM/31bDNA^{ABNOH}_FAM** and *bs*GntR (**Figure 79-B**).

Reaction of 19DNA^{ABNOH}_FAM with H2A or Lysozyme in different buffers– SDS PAGE analysis

19DNA^{ABNOH}_FAM was prepared by PEX in semi-preparative scale as described above (Section 5.5.3). The reaction mixture (10 µL) containing **19DNA^{ABNOH}_FAM** (1 µM) was incubated with H2A or Lysozyme (5 µM) in H_2O or 50 mM phosphate buffer at different pH of 5, 7.4 or 8.5 at 37 °C for 1 h in a thermal cycler and subsequently stopped by addition of SDS-PAGE stop solution (10 µL) and heated 2 min at 99 °C. Samples were analyzed by 17.5 % denaturing SDS-PAGE and visualized using florescence imaging (**Figure 83**).

Reaction of 19DNA^{Nat}_FAM or 19DNA^{ABNOH}_FAM with DNA-binding proteins not containing Trp (Histones H2A and H4) and with non-DNA-binding protein containing Trp (Lysozyme) – SDS PAGE analysis

6-FAM-labelled **19DNA^{Nat}_FAM** and **19DNA^{ABNO}_FAM** were prepared by PEX in semi-preparative scale as described above (Section 5.5.3). The reaction mixture (10 µL) containing **19DNA^{Nat}_FAM** (1 µM) was incubated with Histones H2A, H4 or Lysozyme (5 µM) in AcOH 1 mM at 25 °C for 1 h in a thermal cycler and subsequently stopped by addition of SDS-PAGE stop solution (10 µL) and heated 2 min at 99 °C. Samples were analyzed by 17.5 % denaturing SDS-PAGE and visualized using florescence imaging (**Figure 84**).

Sequences of the proteins used in the experiment:

H2A

SGRGK QGGKA RAKAK SRSSR AGLQF PVGRV HRLLR KGNYS ERVGA GAPVY LAAVL EYLTA EILEL AGNAA RDNKK TRIIP RHLQL AIRND EELNK LLGRV TIAQG GVLPN IQAVL LPKKT ESHHK AKGK

Genbank accession number: AAN59960

H4

SGRGK GGKGL GKGGA KRHRK VLRDN IQGIT KPAIR RLARR GGVKR ISGLI YEETR GVLKV FLENV IRDAV TYTEH AKRKT VTAMD VVYAL KRQGR TLYGF GG

Genbank accession number: AAM83108

Lysozyme

MRSLLILVLC FLPLAALGKV FGRCELAAAM KRHGLDNYRG YSLGN**W**VCAA KFESNFNTQA TNRNTDGSTD YGILQINSR**W** **W**CNDGRTPGS RNLCNIPCSA LLSSDITASV NCAKKIVSDG NGMNA**W**VA**W**R NRCKGTDVQA **W**IRGCRL

CAS Number: 12650-88-3

Native mass spectrometry of 31aDNA^{ABNOH} with *bs*GntR protein

Crude reaction mixture following the conjugation of FAM-labelled **31aDNA^{ABNOH}** with the *bs*GntR protein (as described above) were buffer exchanged into 150 mM ammonium acetate pH

7.4 (99.99% trace metals basis, Merck) by six cycles of 10-fold dilution and re-concentration on a Vivaspin 500 centrifugal concentrator (30 kDa MWCO, Sartorius). Without further dilution, the samples were introduced into a Waters Synapt G2Si mass spectrometer by ESI from in-house manufactured golden-coated borosilicate glass capillaries (1B120F-4, World Precision Instruments) as described previously[128]. The spectrometer was operated in positive sensitivity ion mode over the 500 – 20,000 m/z range and was tuned to maximize ion signal, while maintaining low ion activation and non-covalent protein interactions. Namely, the ESI capillary and sampling cone were kept at 1.2 kV and 80 V, respectively. Argon was used as the collision gas at 6 ml/min flow with 40 V trap collision energy to aid in gentle ion desolvation (**Figure 85**). Data were smoothed and externally mass recalibrated on cesium iodide clusters in MassLynx 4.2 (Waters) with mass deconvolution through a combination of UniDec 6.0.4 algorithm and manual deconvolution using the MaCSED utility from Benesch research group at the University of Oxford (http://benesch.chem.ox.ac.uk/resources.html).

6 LIST OF ABBREVIATIONS AND SYMBOLS

BCN bicycle[6.1.0]non-4-yne

CuAAC copper catalyzed cycloaddition

Cy3 Cyanine 3

Cy5 Cyanine 5

DAPI 4’,6-diamidino-2-phenylindole

DBU 1,8-diazabicyclo[5.4.0]undec-7-ene

DMF N,N-dimethylformamide

DMSO dimethyl sulfoxide

DNA deoxyribonucleic acid

dNMPs 2’-deoxyribonucleoside monophosphates

dNTPs 2’-deoxyribonucleoside triphosphates

DRAQ-5 deep red antraquinone 5

DPCs DNA-Protein cross-linking

dsDNA double stranded DNA

EDTA ethylendiamine-tetraacetic acid

EDG electron donating group

EPR electron spin resonance

Equiv. equivalents

EWG electron-withdrawing group

FAM fluorescein

FG functional group

FRET fluorescence resonance energy transfer

FTICR fourier transform ion cyclotron resonance

HOMO highest occupied molecular orbital

HMBC heteronuclear multiple bond correlation

HRMS high resolution mass spectroscopy

HSQC heteronuclear spin quantum coherence spectroscopy

HPLC high performance liquid chromatography

IEDDA inverse electron-demand Diels Alder

LC liquid chromatography

LUMO lowest occupied molecular orbital

MALDI matrix-assisted laser desorption/ionization

MeCN acetonitrile

NMR nuclear magnetic resonance

ON oligonucleotide

PAGE polyacrylamide gel electrophoresis

PBS phosphate buffer saline

PCR polymerase chain reaction

PEX primer exenstion reaction

RNA ribonucleic acid

SDS-PAGE sodium dodecyl sulfate polyacrylamide gel electrophoresis

SNTT synthetic nucleoside triphosphate transporter

SPAAC strain promoted azido alkyne cycloaddition

SPANC strain promoted alkyne nitrone cyloaddition

ssDNA single stranded DNA

T tetrazine

TAMRA tetramethylrhodamine

TBTA tris(benzyltriazolylmethyl)amine

TBET through band energy transfer

TCO *trans*-cyclooctene

TEAB triethylammonium bicarbonate

TIC total ionization chromatogram

TOP1 Topoisomerase I

TPPTS Tris(3-sulfophenyl)phosphine trisodium salt

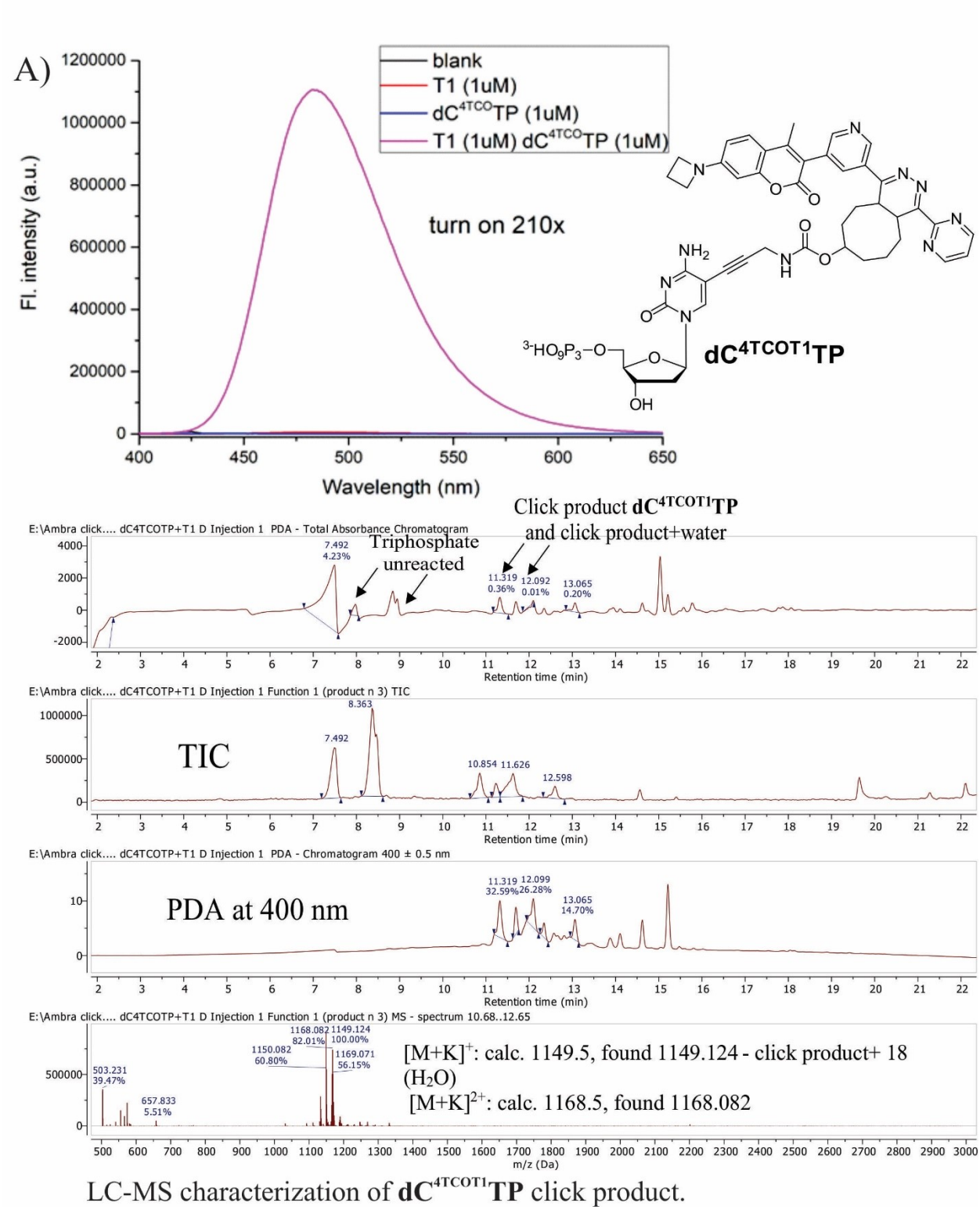
TRIS tris(hydromethyl)aminoethane

UV ultraviolet

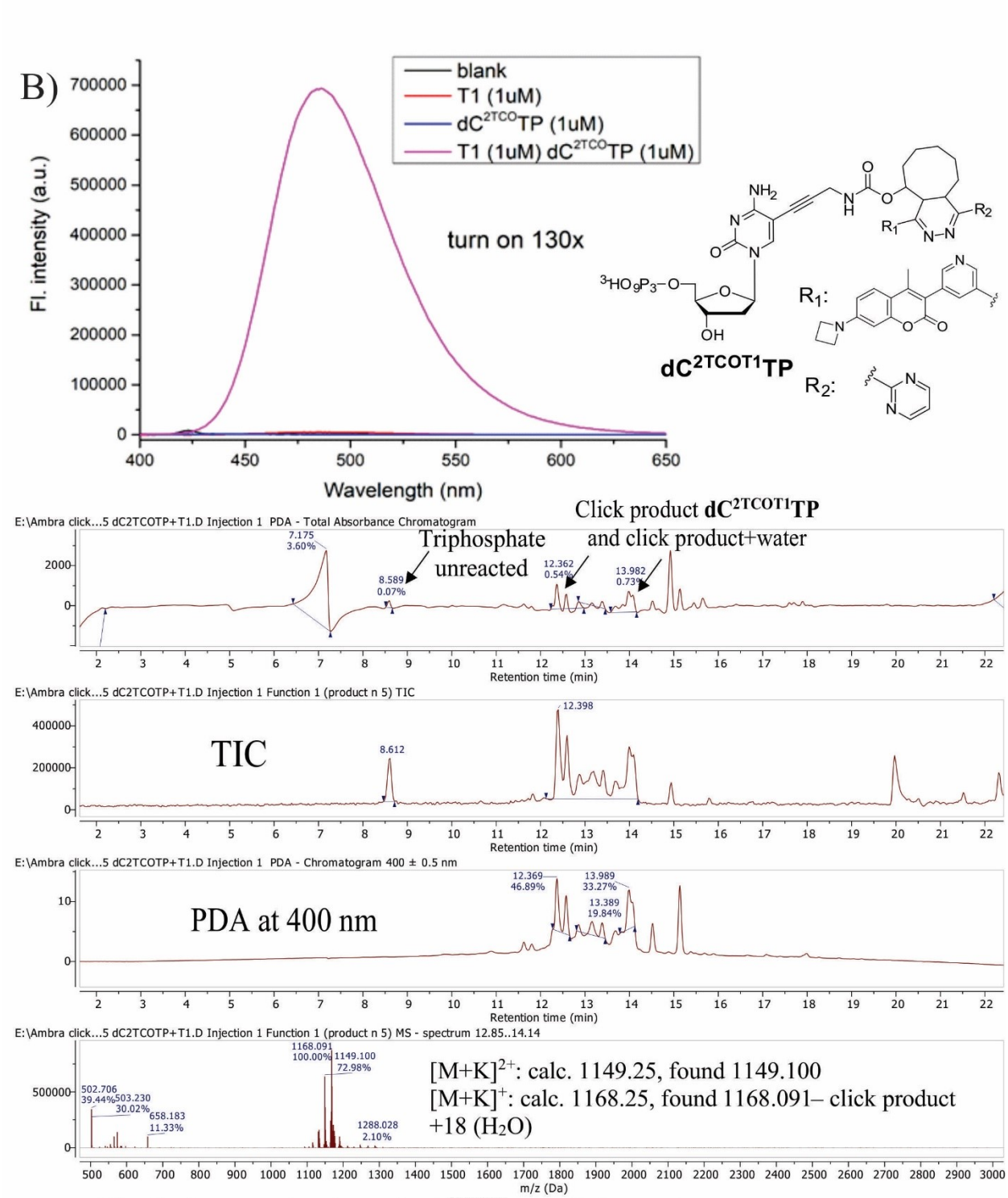
U-2 OS human bone osteosarcoma epithelial cell

7 APPENDIX

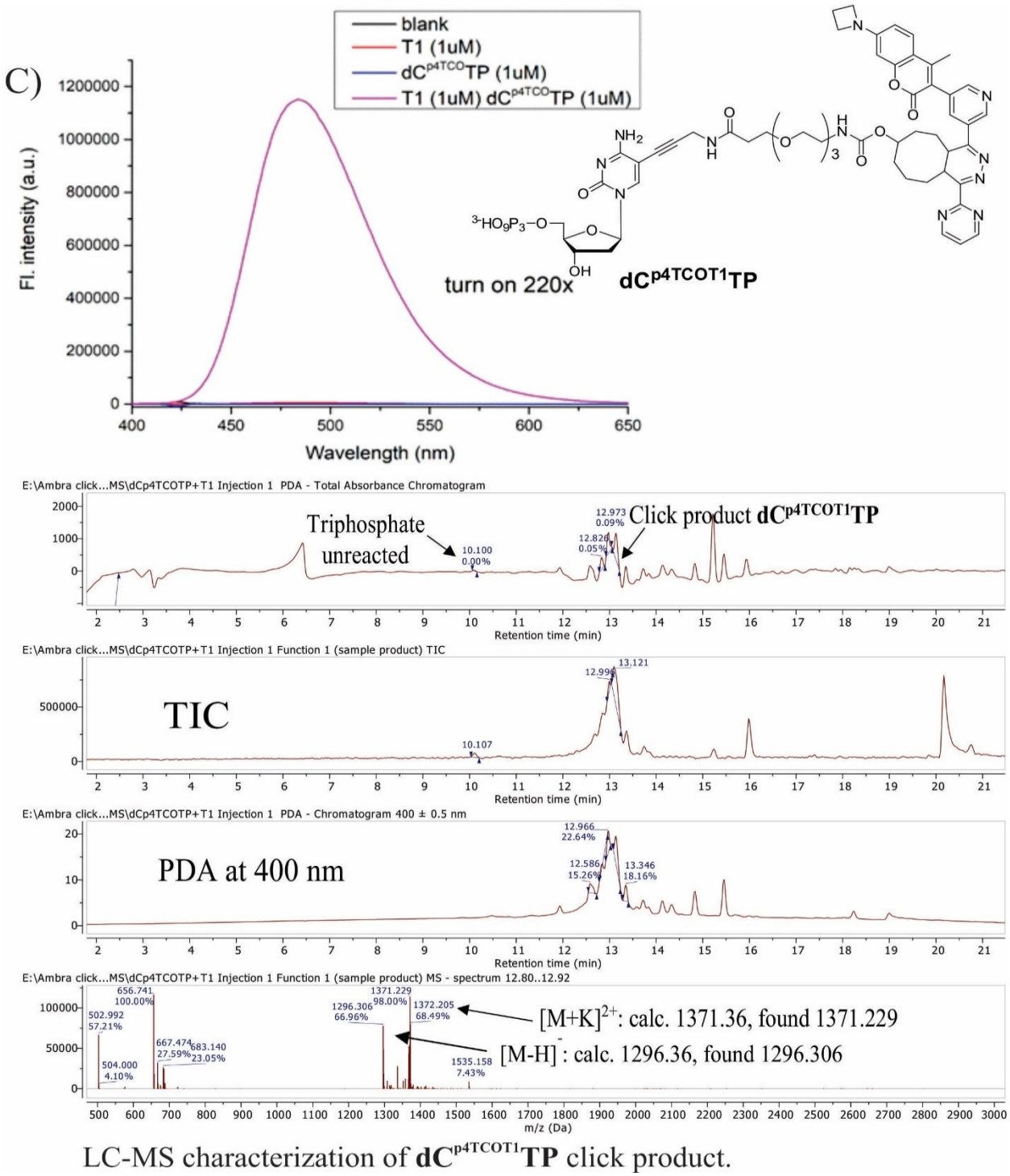
7.1 Fluorescence and LC-MS spectra



LC-MS characterization of **dC^{4TCOT1}TP** click product.



LC/MS characterization of **dC^{2TCOT1}TP** click product.



LC-MS characterization of **dC^{p4TCOT1}TP** click product.

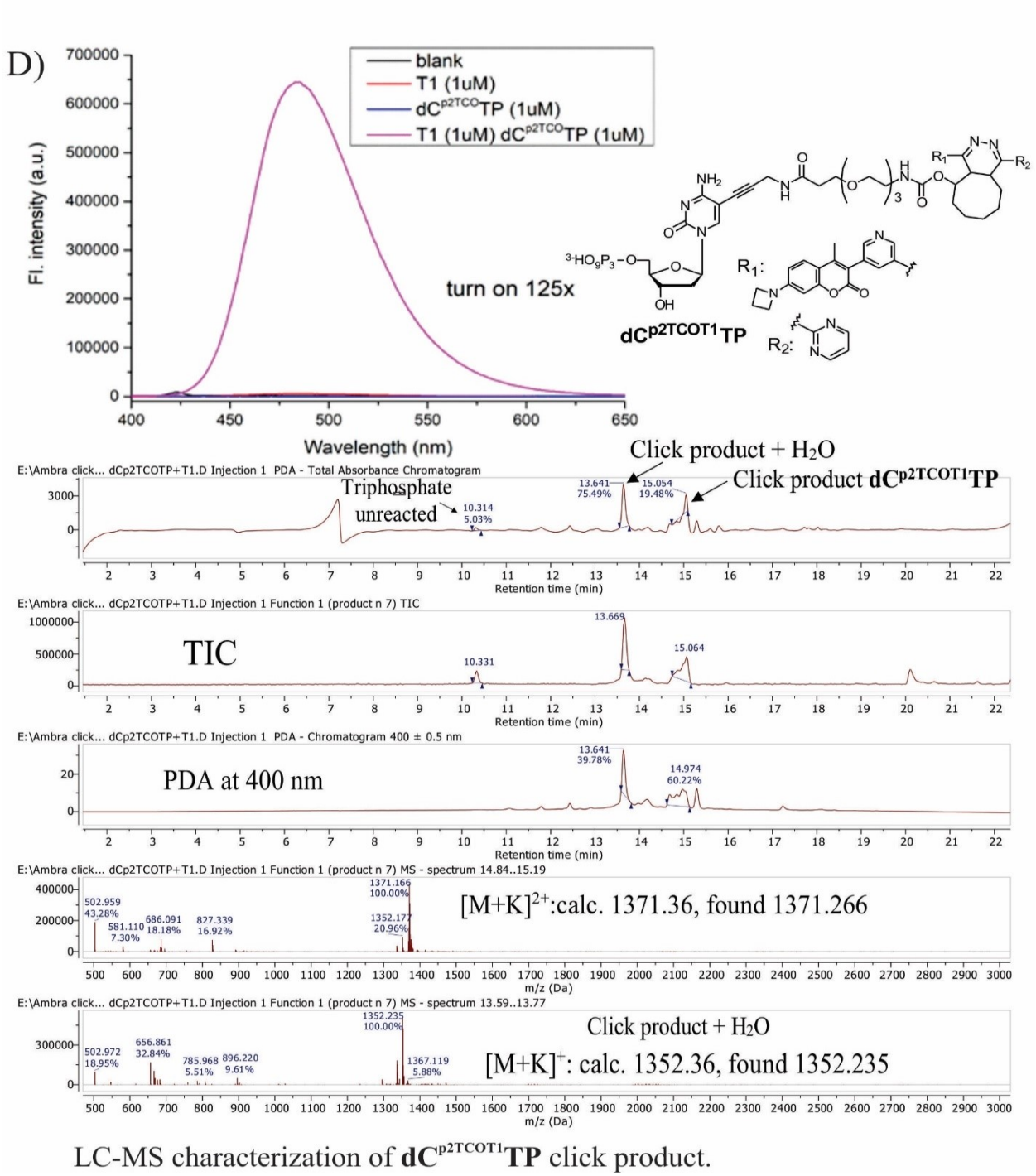
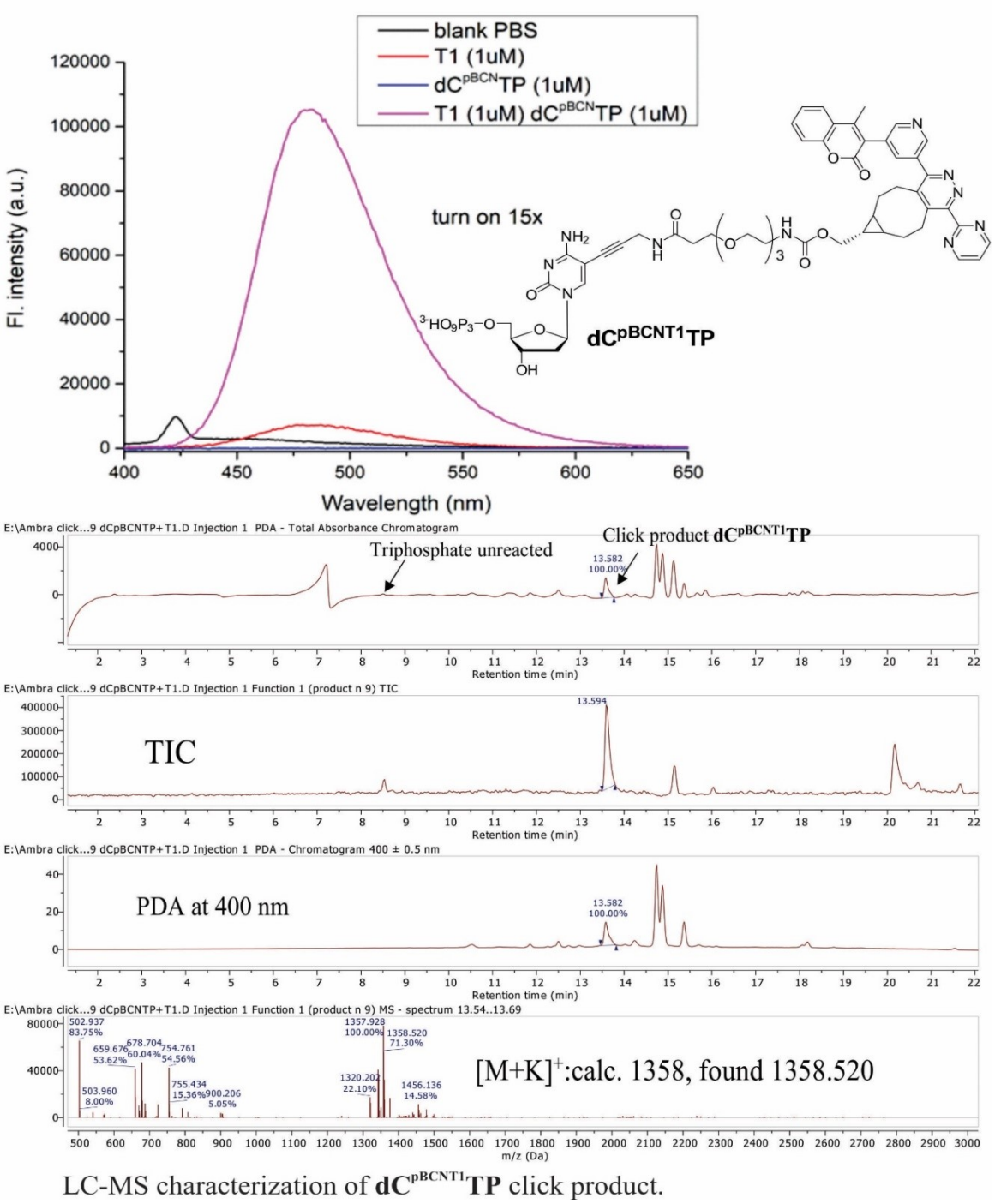


Figure 86. Fluorescence turn on measurements and LC-MS spectra of the click products formed in the reaction of **T1** tetrazine with A) **dC^{4TCO}TP**, B) **dC^{2TCO}TP**, C) **dC^{p4TCO}TP**, D) **dC^{p2TCO}TP** (indicated as number-x = x-fold). Fluorescence spectra of the click products (magenta line) are shown in relation to the respective quenched tetrazine (red line), the triphosphate alone (blue line) and PBS blank (black line) to indicate the fluorescence turn-on ratio.



LC-MS characterization of **dC^{pBCNT1}TP** click product.

Figure 87. Fluorescence turn on meaurement and LC-MS spectrum of the click product formed in the reaction of **T1** tetrazine with **dC^{pBCN}TP** (indicated as number-x = x-fold). Fluorescence spectra of the click product (magenta line) are shown in relation to the respective quenched tetrazine (red line), the triphosphate alone (blue line) and to PBS blank (black line) to indicate the fluorescence turn-on ratio.

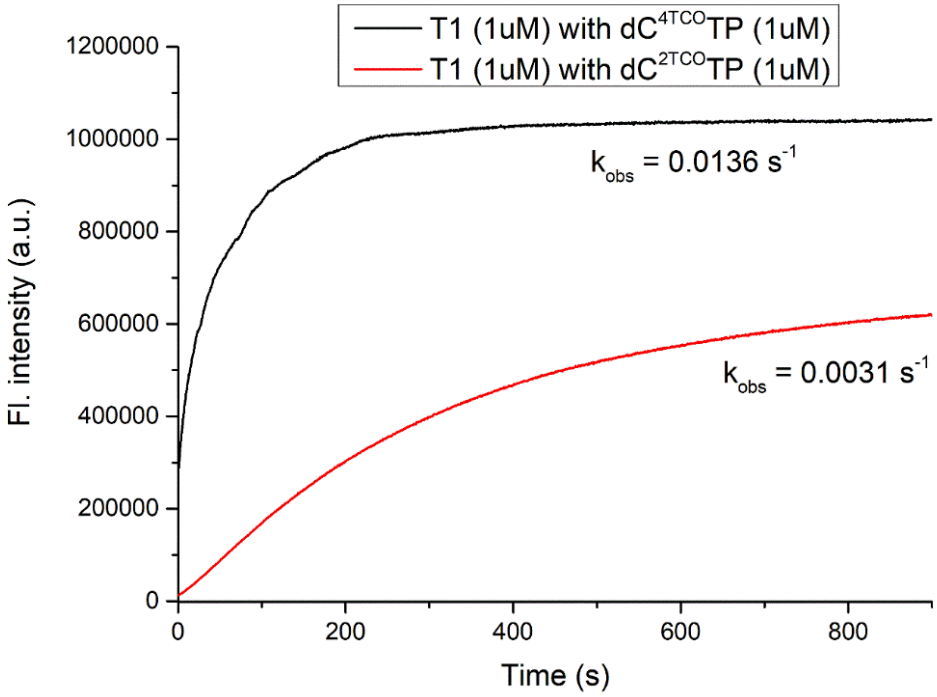
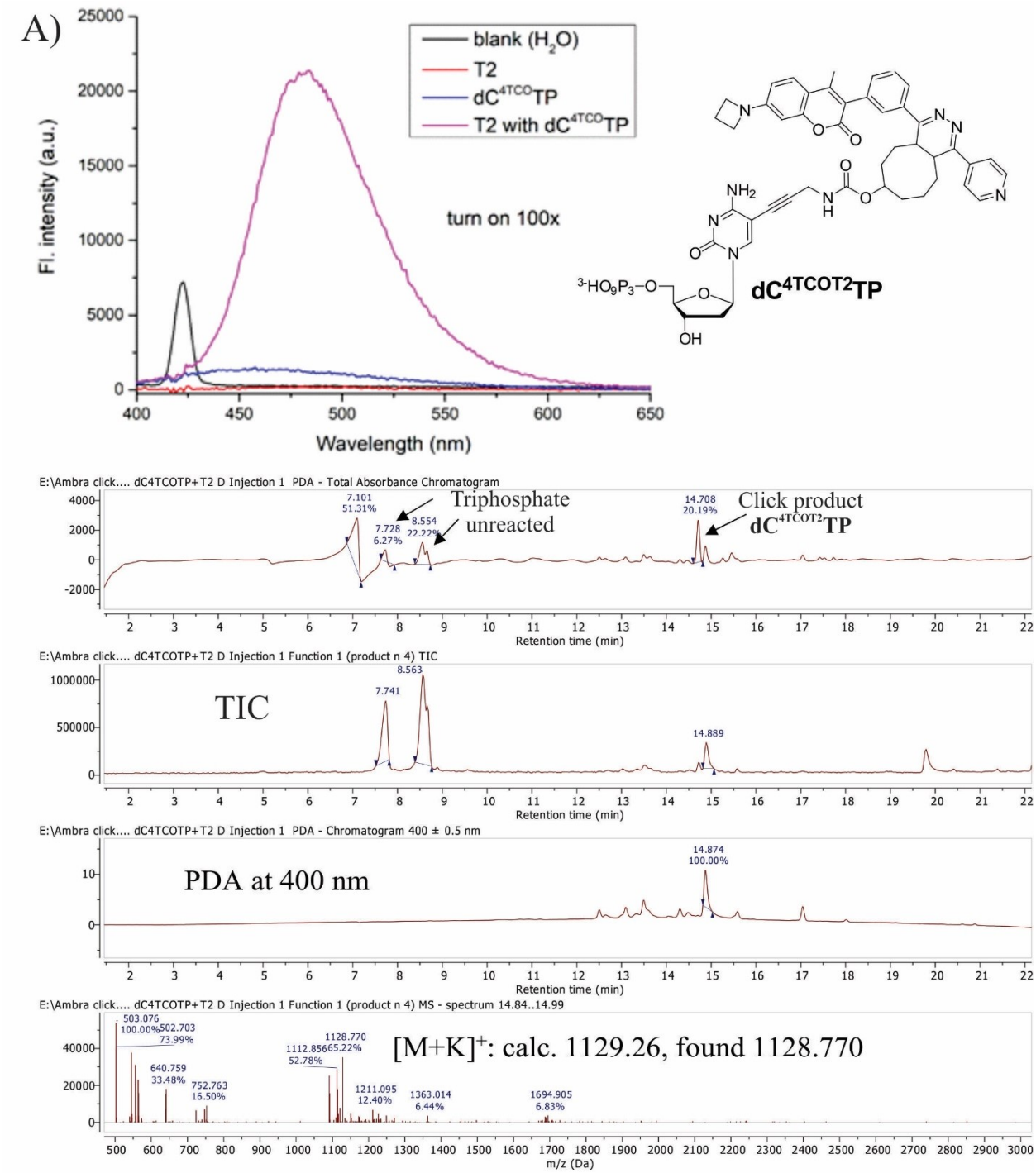


Figure 88. Fluorescence time lapse measurements showing changes in the fluorescence signal of A) the click products over 30 min for **T1** or B) 1h for **T2**. A) The fluorescence intensities were finally plotted as the function of time showing that **dC^{4TCO}TP** reacts with **T1** 4.5x faster than **dC^{2TCO}TP**. Excitation was set to 370 nm (slit 3) and the fluorescence was collected at 485 nm (slit 3) over 30 minutes (interval time 0.5 s).



LC-MS characterization of **dC^{4TCOT2}TP** click product.

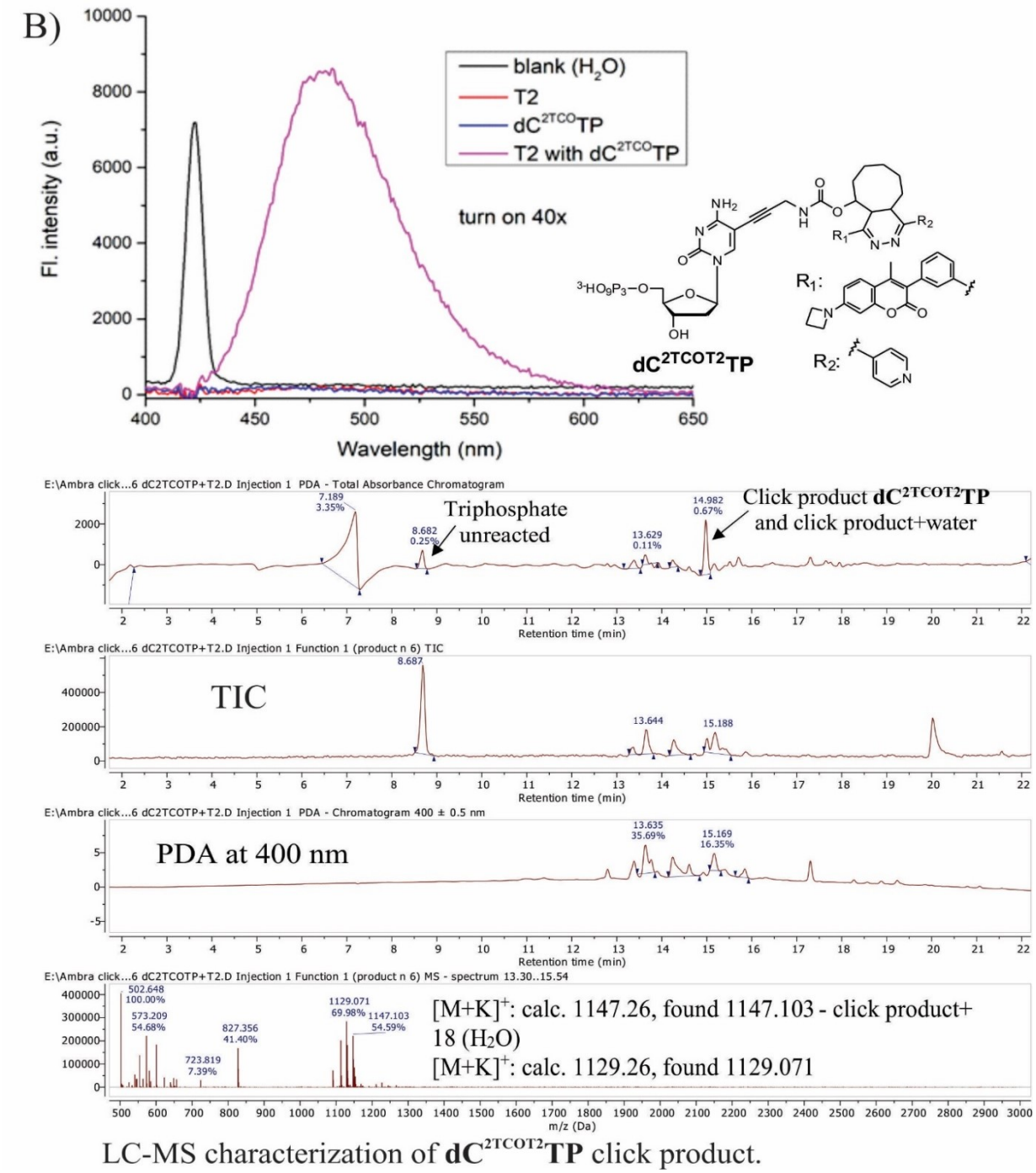
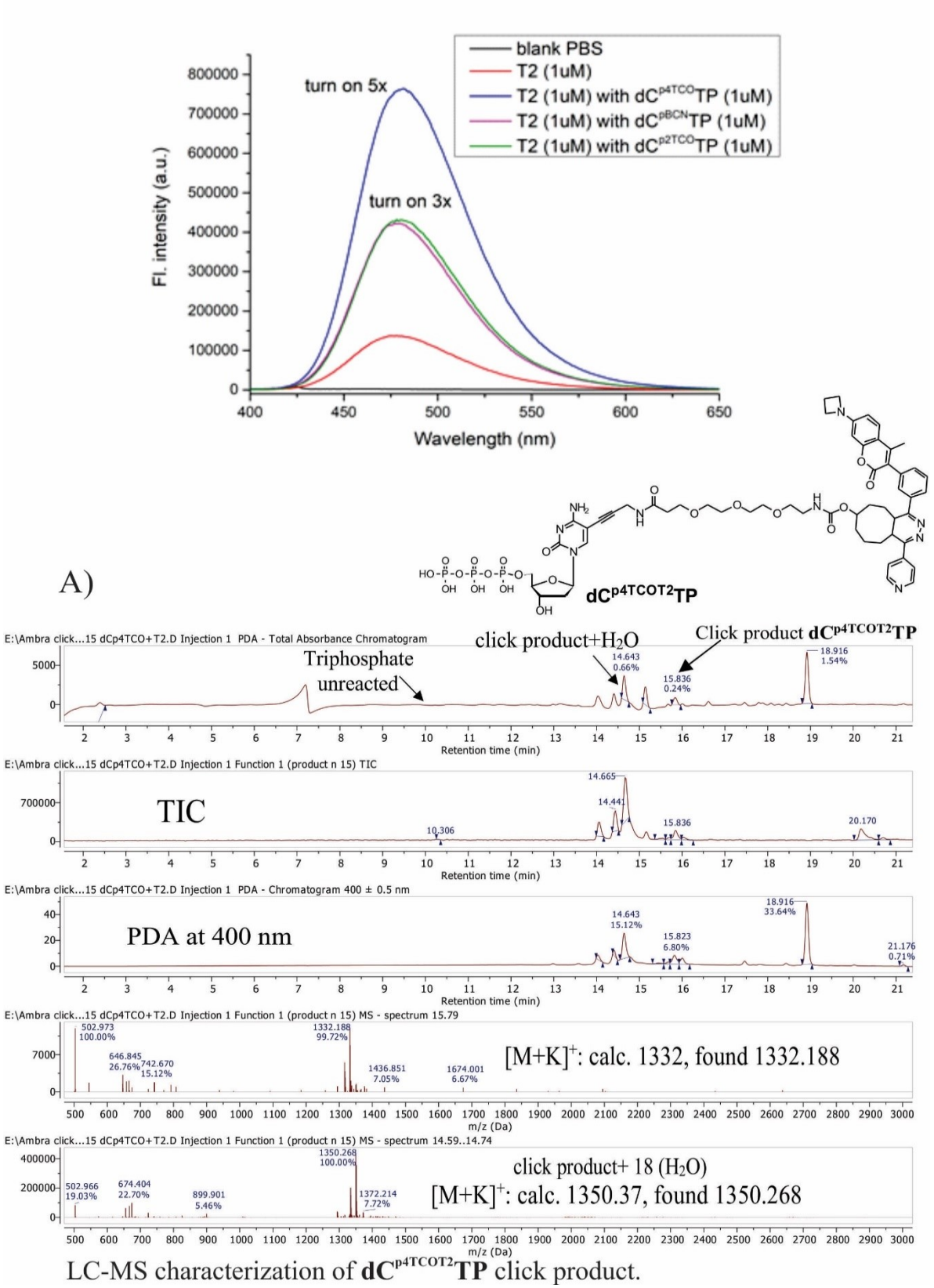
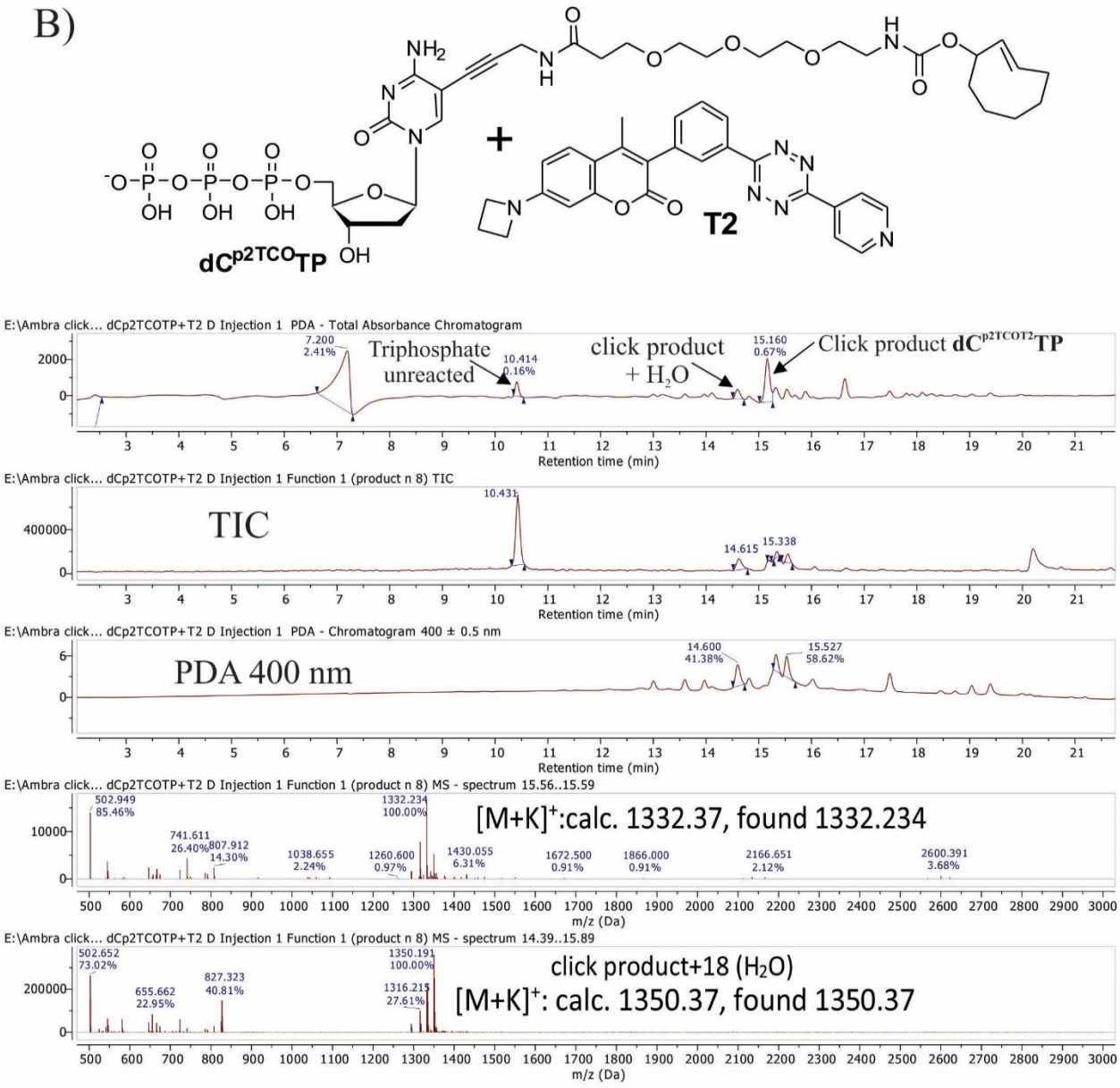


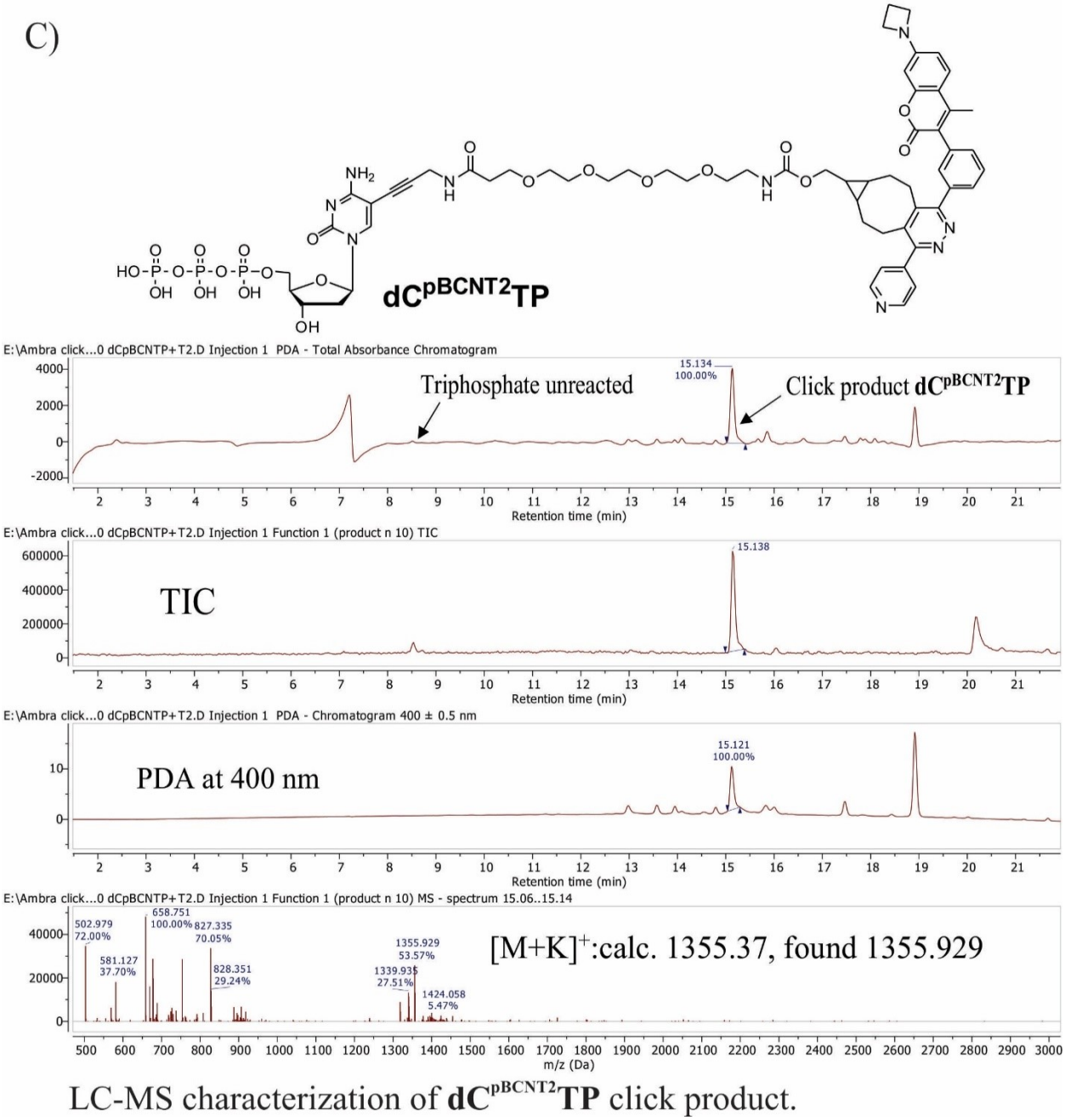
Figure 89. Fluorescence turn on measurement and LC-MS spectra of the click product formed in the reaction of **T2** tetrazine with A) **dC^{4TCO}TP** (indicated as number-x = x-fold), B) **dC^{2TCO}TP**. Fluorescence spectra of the click product (magenta line) are shown in relation to the respective quenched tetrazine (red line), the trisphosphate alone (blue line) and to PBS blank (black line) to indicate the fluorescence turn-on ratio.



LC-MS characterization of **dC^{p4TCOT2}TP** click product.



LC-MS characterization of **dC^{p2TCOT2}TP** click product.



LC-MS characterization of **dC^{pBCNT2}TP** click product.

Figure 90. Fluorescence turn on meaurements and LC-MS spectra of the click products formed in the reaction of **T2** tetrazine with A) **dC^{p4TCO}TP** (indicated as number-x = x-fold), B) **dC^{p2TCO}TP** and C) **dC^{pBCN}TP**. Fluorescence spectra of the click products (blue, magenta and green line) are shown in relation to the respective quenched tetrazine (red line) and to PBS blank (black line) to indicate the fluorescence turn-on ratio.

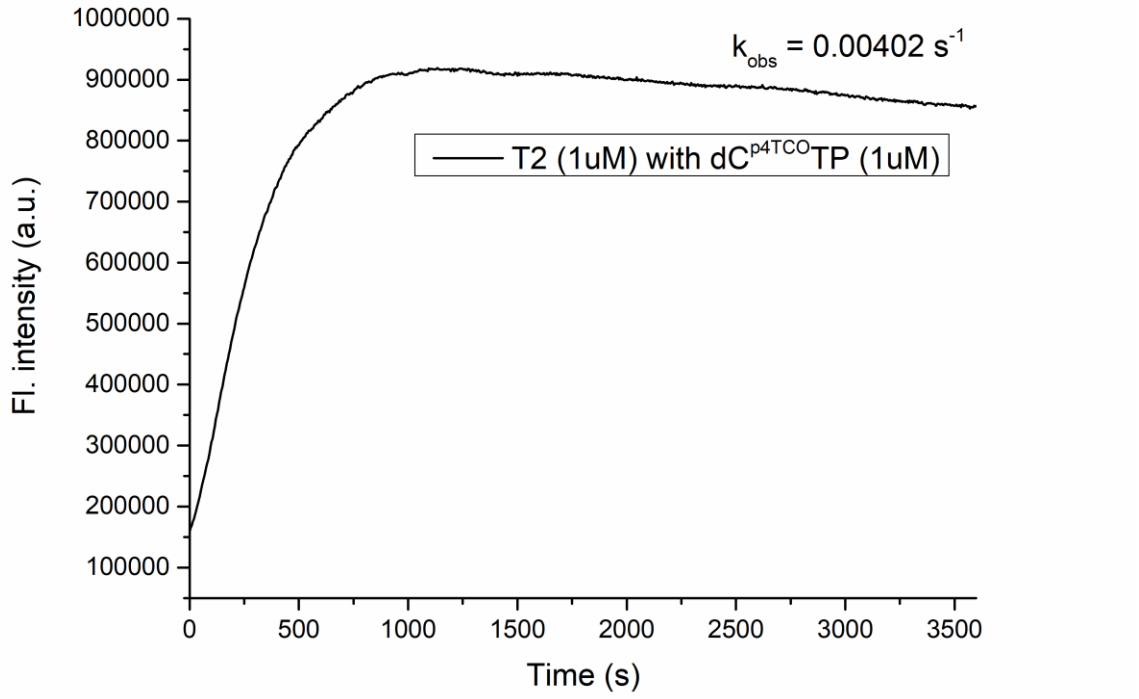
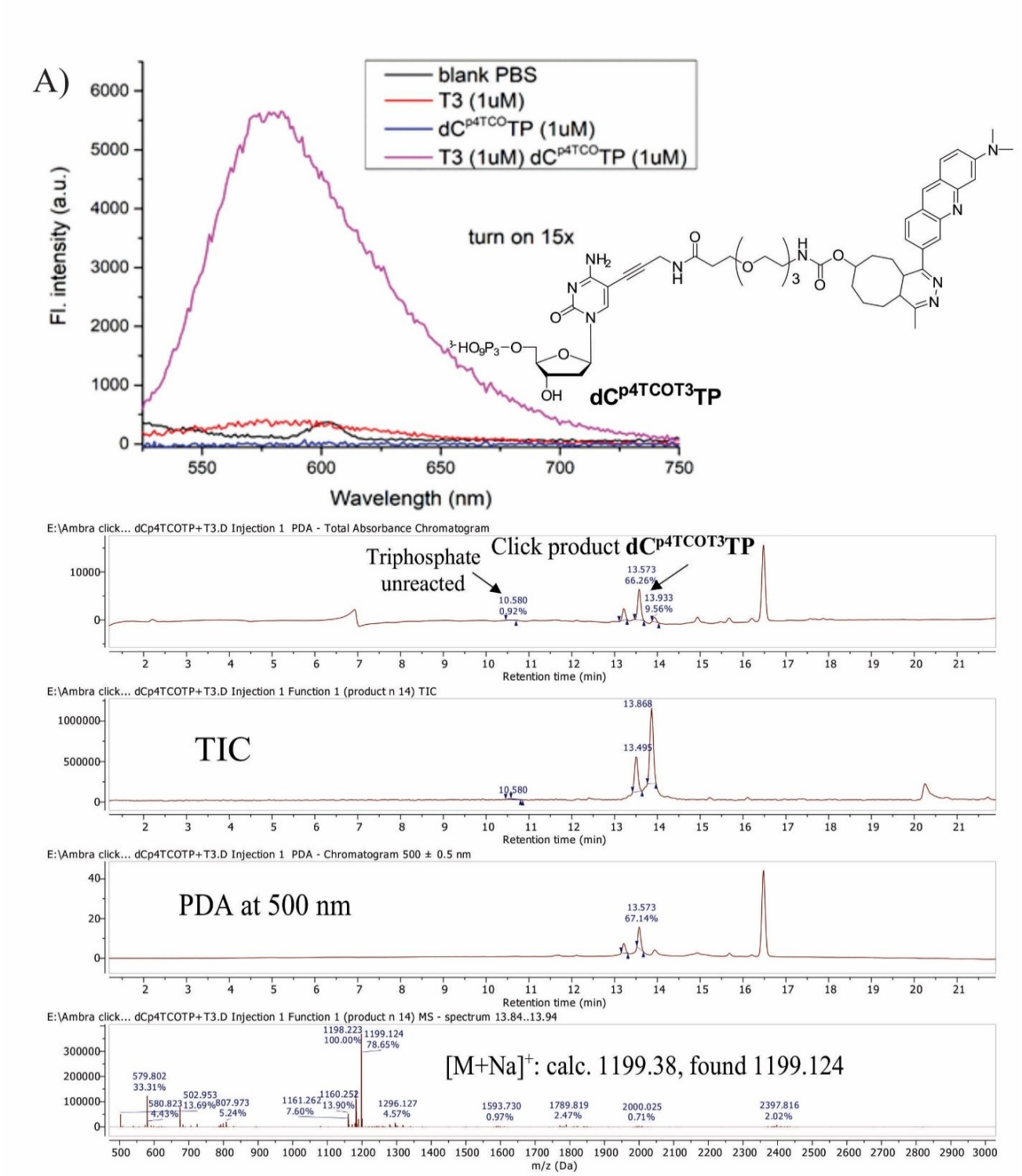
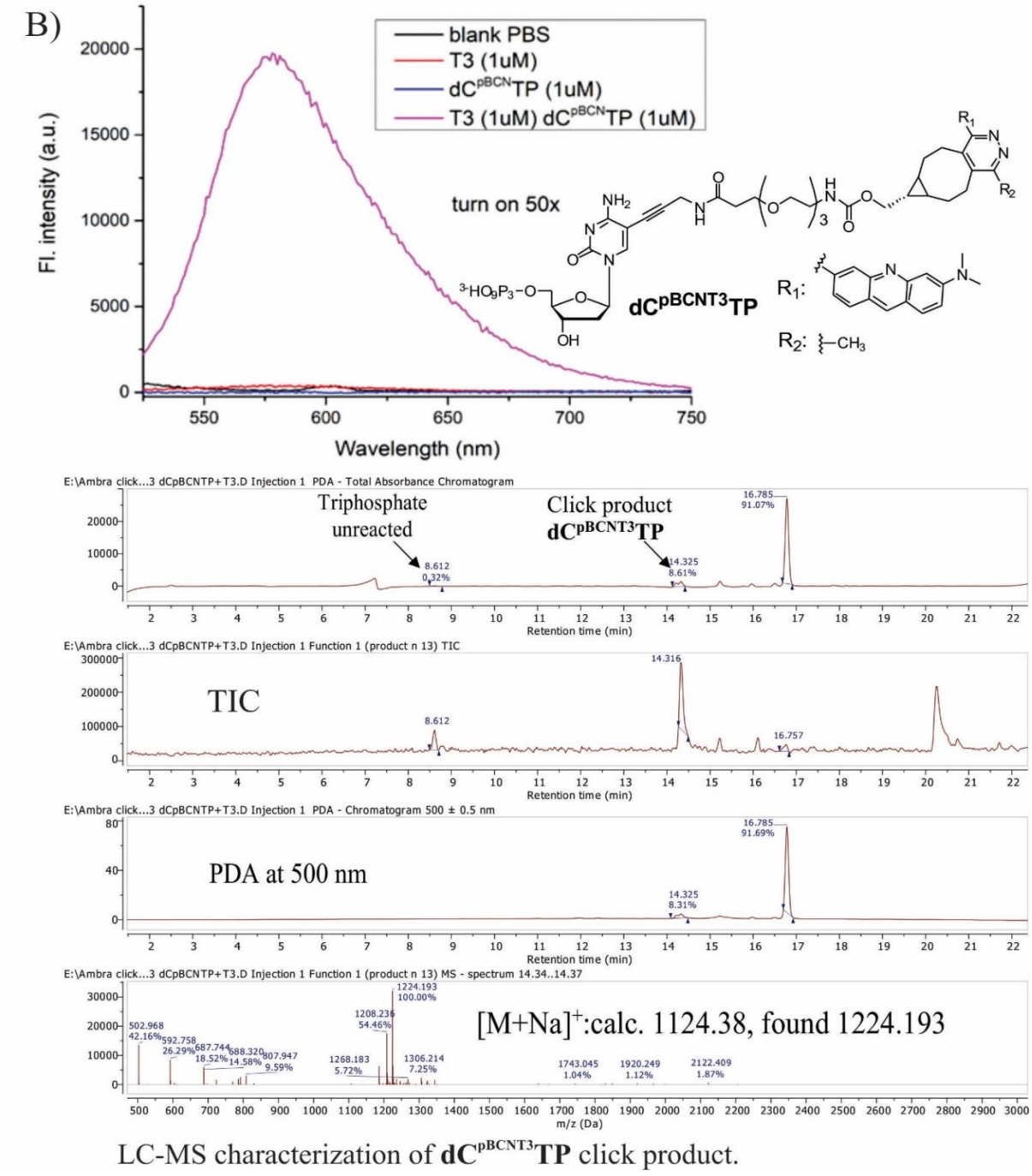


Figure 91. Fluorescence time lapse measurements showing changes in the fluorescence signal of the click products over 1h for **T2**. Excitation was set to 370 nm (slit 3) and the fluorescence was collected at 485 nm (slit 3) over 1h for **T2** (interval time 2s). The data were subtracted from the background fluorescence signal of PBS. The observed rate constants were calculated by fitting the curves with single exponential equation using the Origin Pro software.



LC-MS characterization of **dC^{p4TCOT3}TP** click product.



LC-MS characterization of **dC^{pBCNT3}TP** click product.

Figure 92. Fluorescence turn on measurements and LC-MS spectra of the click products formed in the reaction of **T3** tetrazine with A) **dC^{p4TCO}TP**, B) **dC^{pBCN}TP** (indicated as number-x = x-fold). Fluorescence spectra of the click products (magenta line) are shown in relation to the respective quenched tetrazine (red line), the triphosphate alone (blue line) and PBS blank (black line) to indicate the fluorescence turn-on ratio.

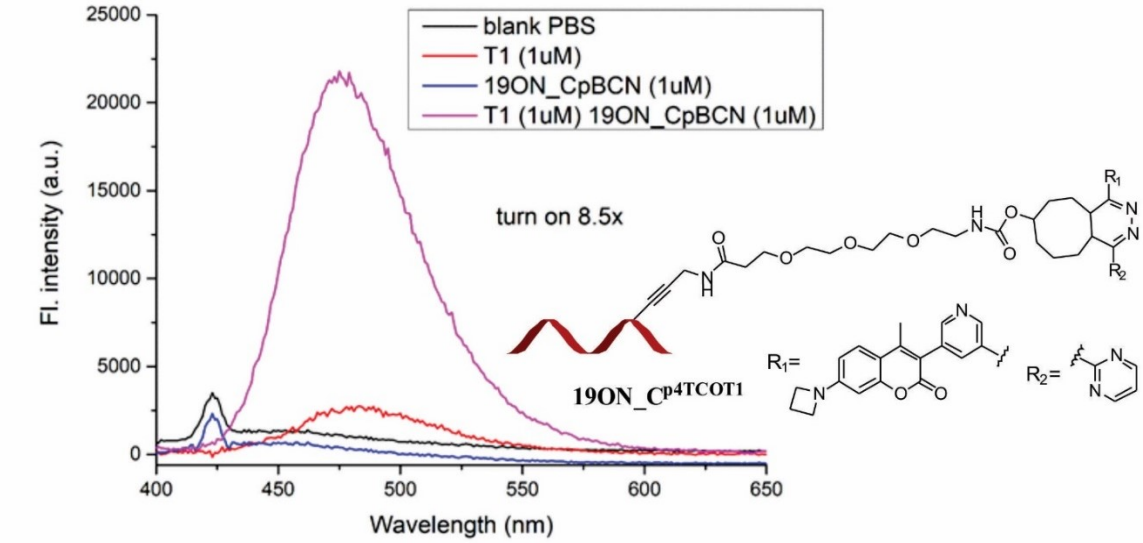


Figure 93. Fluorescence spectra of the click product formed in the reaction of tetrazine **T1** with **19ON_C^{p4TCO}** (indicated as number-x = x-fold). Fluorescence spectra of the tetrazine-linked DNA after reaction with **T1** (magenta line) are shown in relation to the respective quenched tetrazine (red line), to the **19ON_C^{p4TCO}** alone (blue line) and to PBS blank (black line) to indicate the fluorescence turn-on ratio. Only one isomer of the click product is shown for clarity.

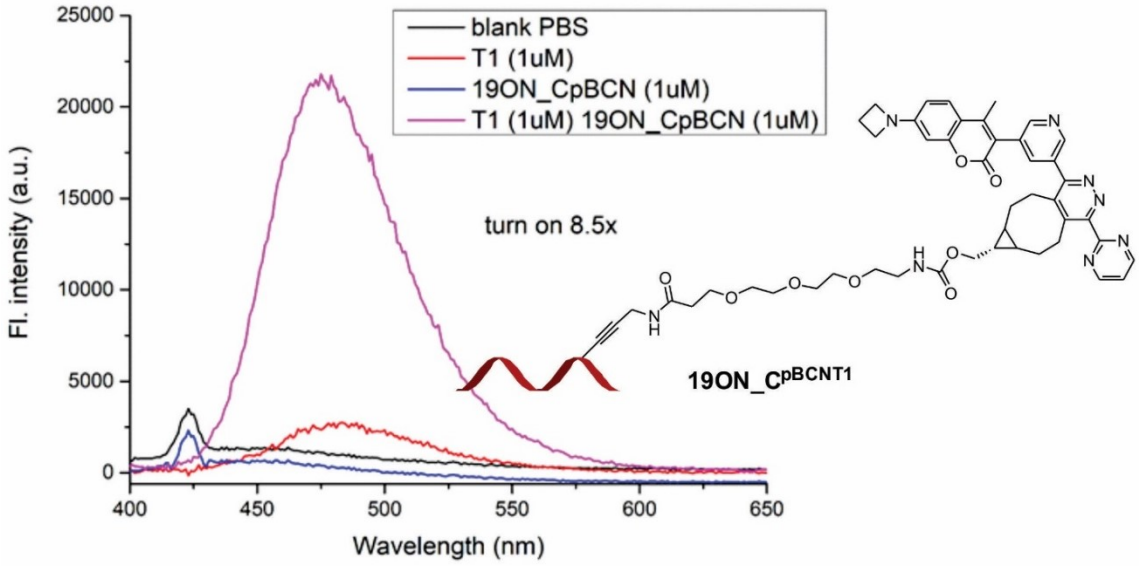


Figure 94. Fluorescence spectra of the click product formed in the reaction of **T1** tetrazine with **19ON_C^{pBCN}** (indicated as number-x = x-fold). Fluorescence spectra of the tetrazine-linked DNA after reaction with **T1** (magenta line) are shown in relation to the respective quenched tetrazine

(red line), to the **19ON_C^{pBCN}** alone (blue line) and to PBS blank (black line) to indicate the fluorescence turn-on ratio.

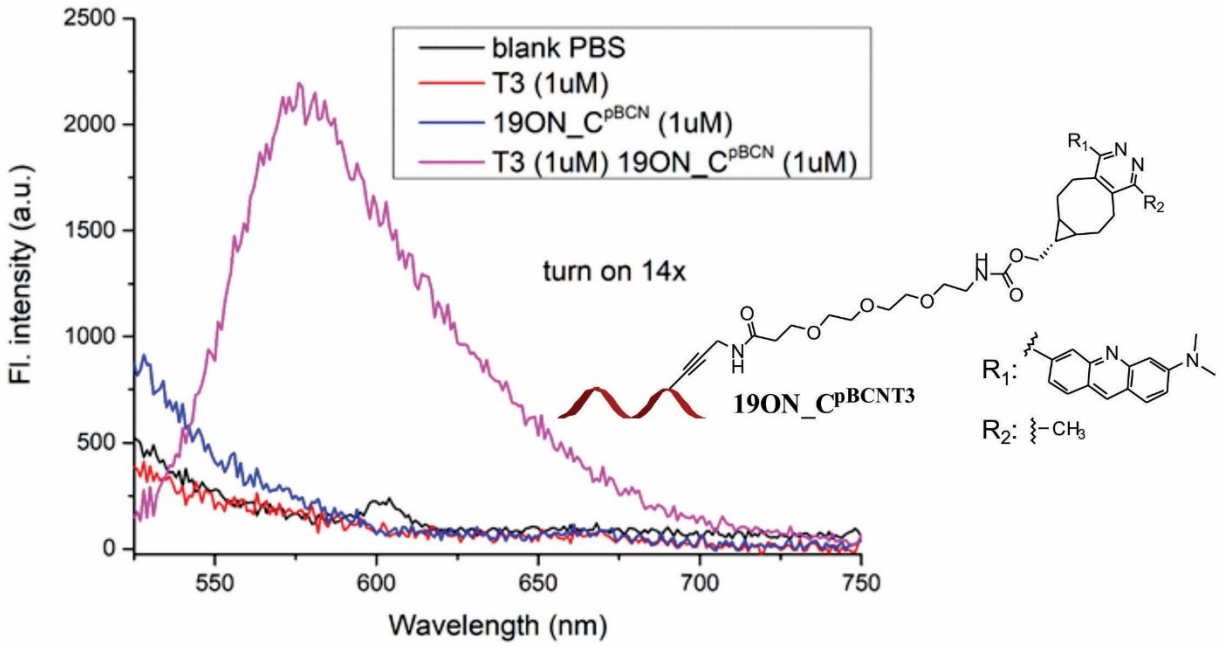


Figure 95. Fluorescence spectra of the click product formed in the reaction of **T3** tetrazine with **19ON_C^{pBCN}** (indicated as number-x = x-fold). Fluorescence spectra of the tetrazine-linked DNA after reaction with **T3** (magenta line) are shown in relation to the respective quenched tetrazine (red line), to the **19ON_C^{pBCN}** alone (blue line) and to PBS blank (black line) to indicate the fluorescence turn-on ratio.

7.2 Cell imaging

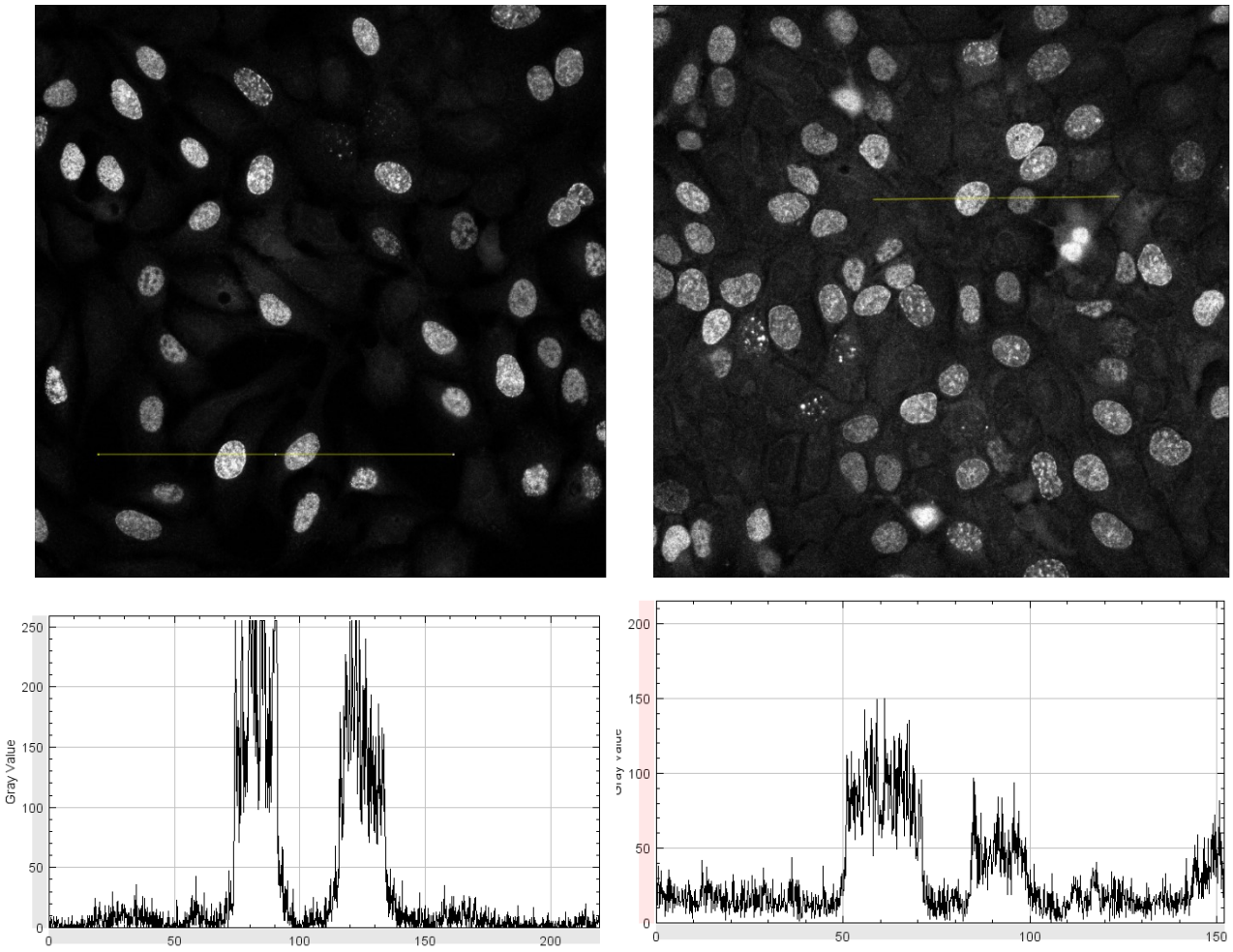
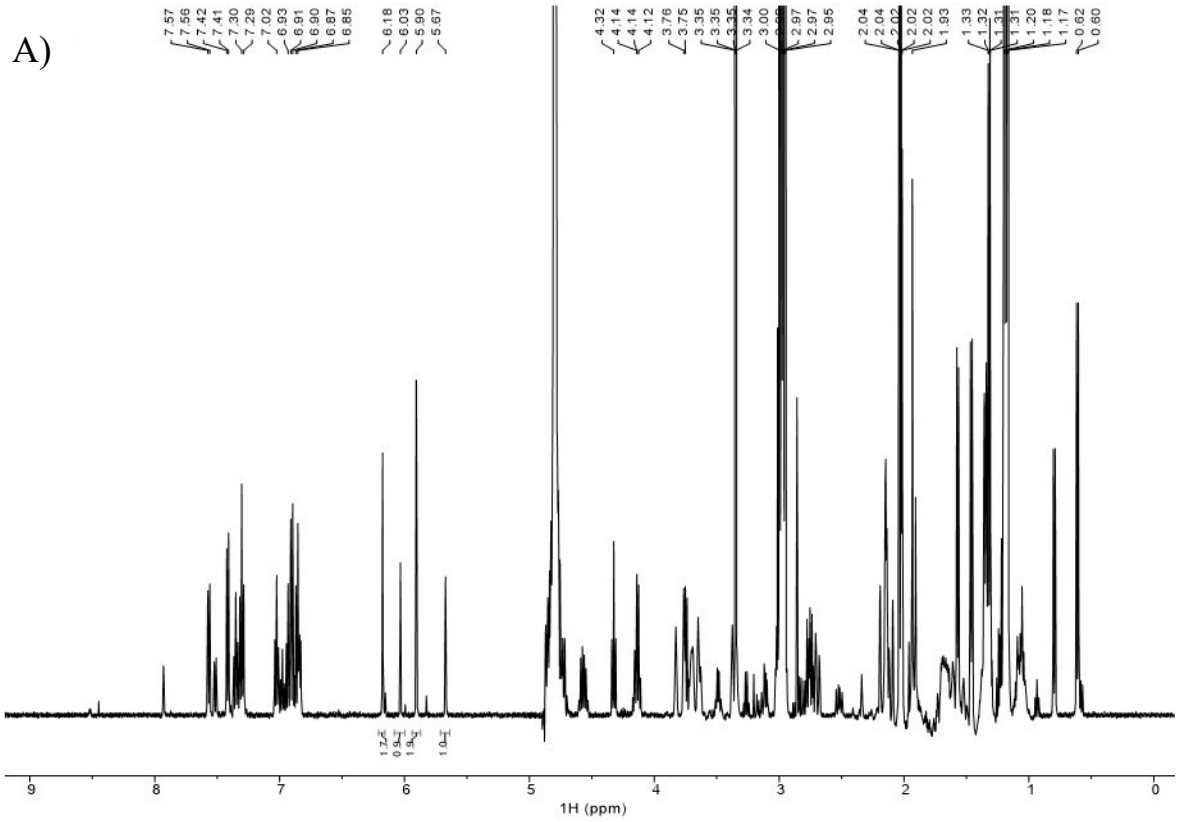


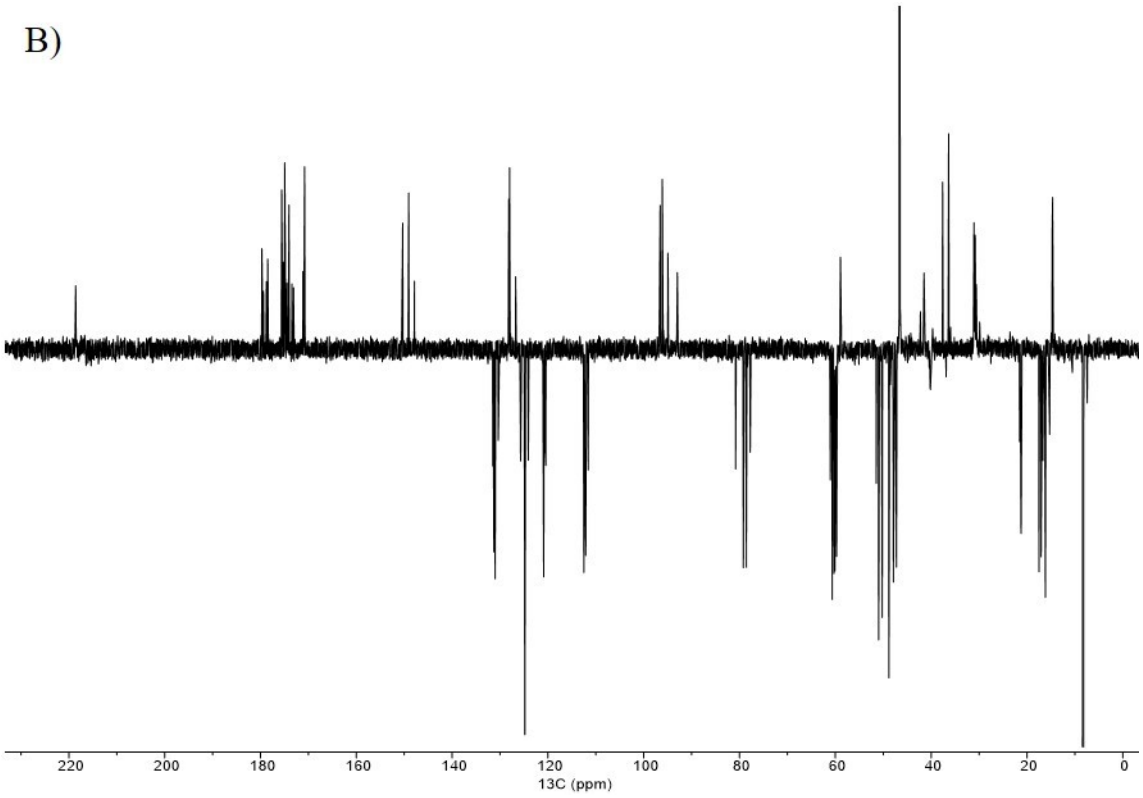
Figure 96. Comparison of contrast in images of U2OS cells labeled with A) **dC^{pBCN}TP** + **T4** and B) **dC^{p4TCO}TP** + **T4.** Images were acquired under identical conditions, acquisition parameters were adjusted according to the brightest nuclei in Image A. Image B) exhibits the lower intensity of labeled DNA foci and higher background fluorescence than image A). Image recorded by Dr. Tomáš Kraus.

7.3 NMR spectra

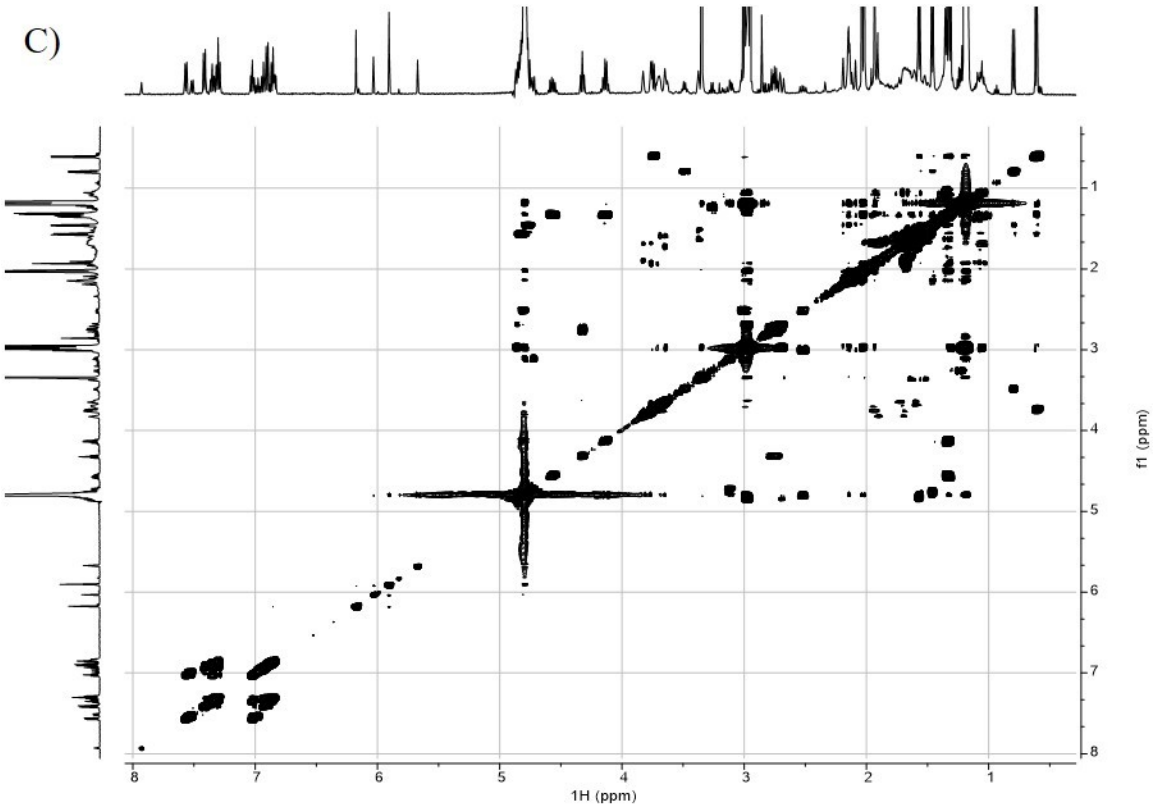
A)



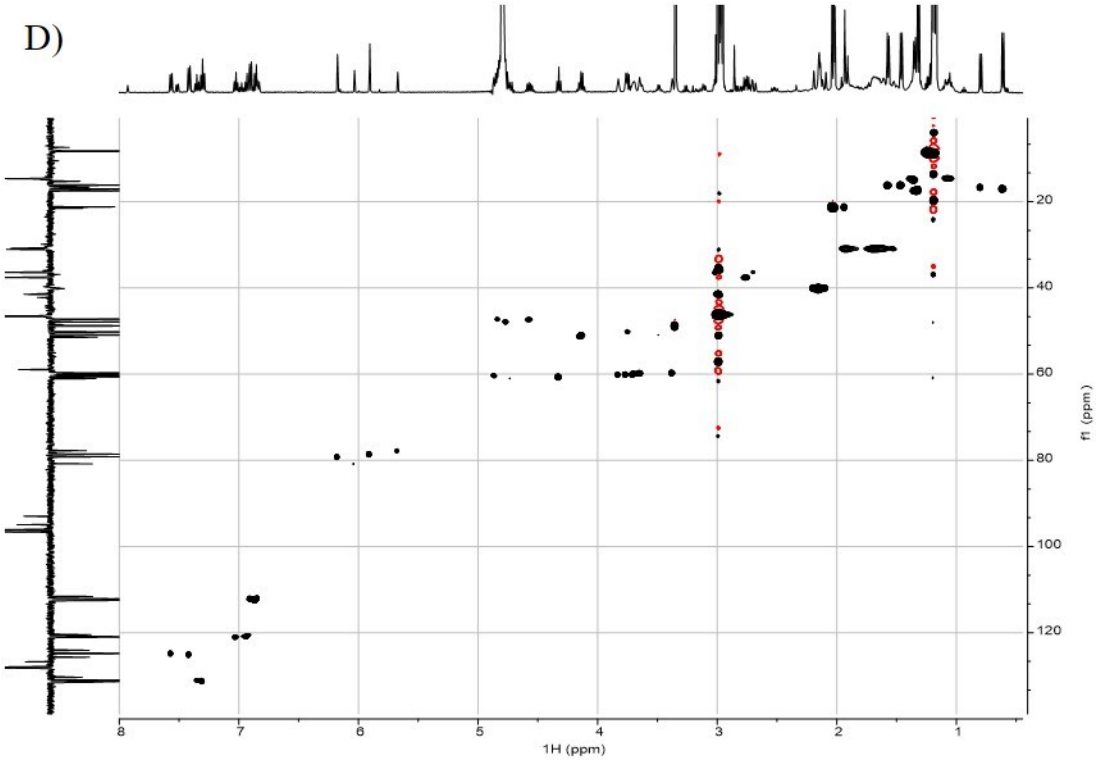
B)


220
200
180
160
140
120
100
80
60
40
20
0
13C (ppm)

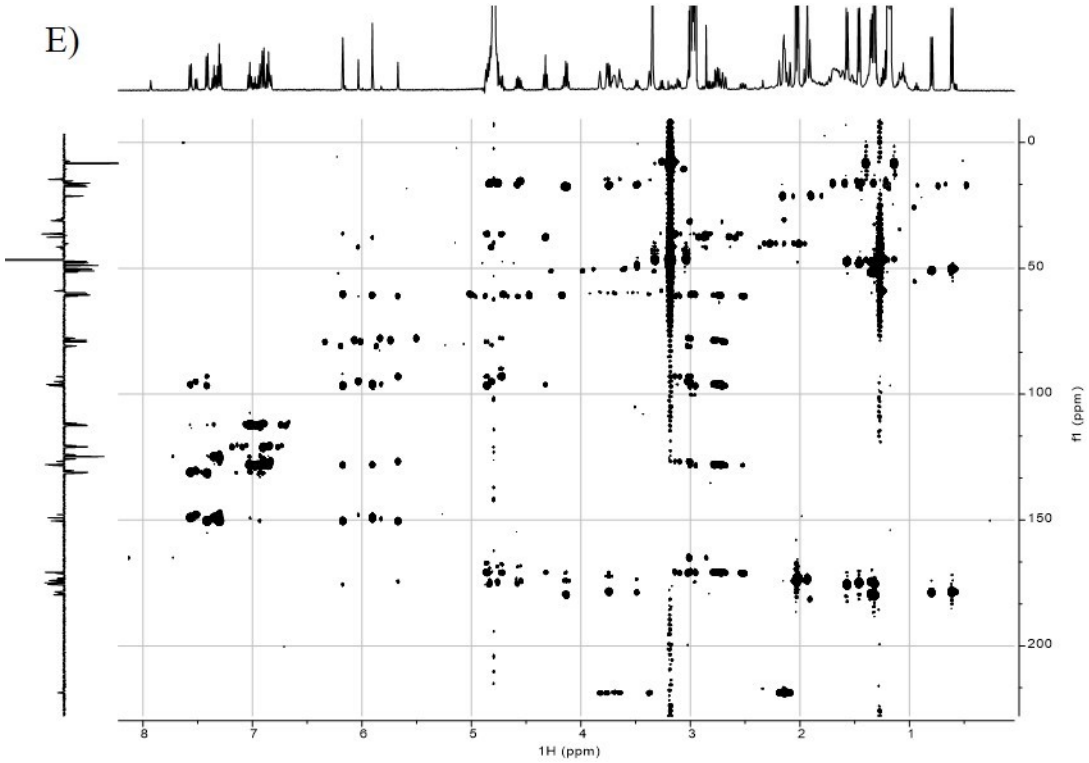
C)


8
7
6
5
4
3
2
1
1H (ppm)
f1 (ppm)

D)


20
40
60
80
100
120
f1 (ppm)
8
7
6
5
4
3
2
1
1H (ppm)

E)


0
50
100
150
200
f1 (ppm)
8
7
6
5
4
3
2
1
1H (ppm)

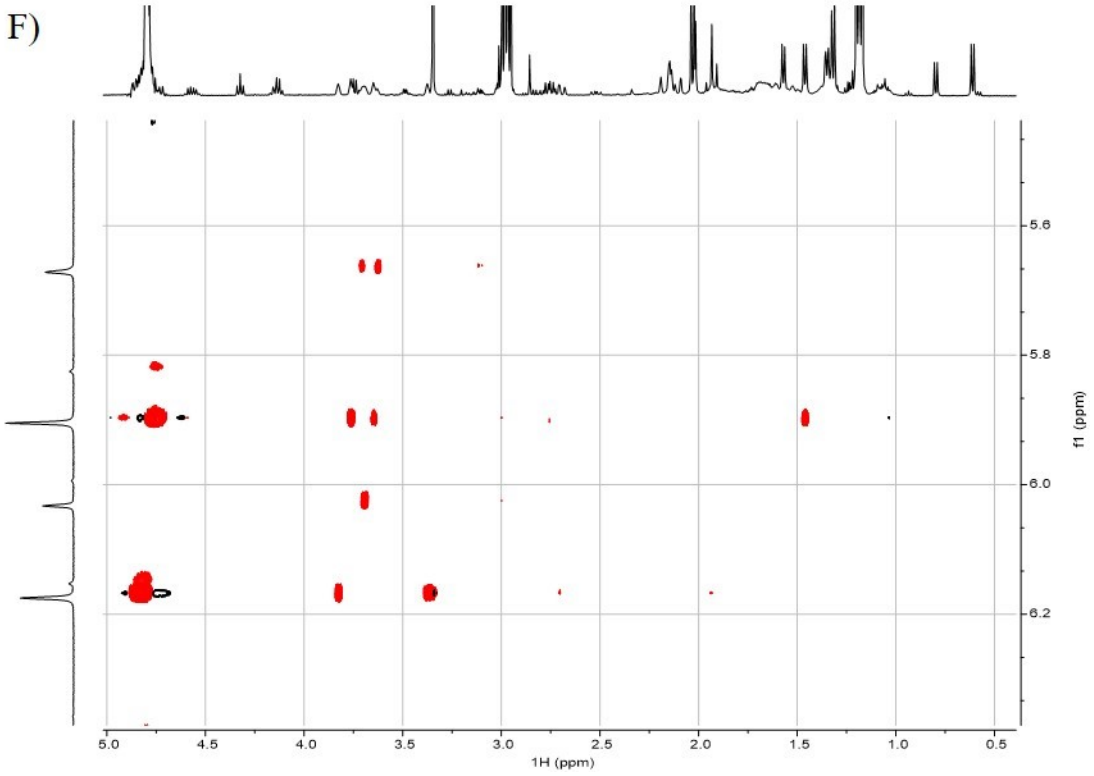


Figure 97. A) ^{1}H NMR of **ABNO^{Pept1}-C** (500.0 MHz, in D_2O), B) ^{13}C NMR of **ABNO^{Pept1}-C** (125.7 MHz, in D_2O), **C)** ^{1}H,^{1}H-COSY of **ABNO^{Pept1}-C** (500.0 MHz, in D_2O), D) ^{1}H,^{13}C-HSQC of **ABNO^{Pept1}-C** (125.7 MHz, in D_2O): E) ^{1}H,^{13}C-HMBC of **ABNO^{Pept1}-C** (125.7 MHz, in D_2O): F) Expansion of ^{1}H,^{1}H-ROESY of **ABNO^{Pept1}-C** (500.0 MHz, in D_2O) showing important NOE cross peaks in red.

HR/MS (ESI$^-$) for $C_{27}H_{34}O_7N_5$ 540.25 $[M-H]^-$ calculated, found 540.24637 $[M-H]^-$

8 REFERENCES:

1 Wang G, Vasquez KM. Z-DNA, *Front Biosci*., **2007**, 1;12:4424-38.

2 Jain A, Wang G, Vasquez KM, *Biochimie*, **2008**, 90(8):1117-30.

3 Pavlova AV, Kubareva EA, Monakhova MV, Zvereva MI, Dolinnaya NG, *Biomolecules,* **2021**, 11(9):1284.

4 https://microbenotes.com/different-forms-of-dna-b-form-a-form-z-form/png.

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

6 Tasara T., Angerer B., Damond M., Winter H., Dörhöfer S., Hübscher U. and Amacker M., *Nucleic Acids Research*, **2023**, 31(10):2636-46.

7 K. Mullis, F. Faloona, S. Scharf, R. Saiki, G. Horn, and H. Erlich, *Cold Spring Harb Symp Quant Biol*, **1986**, 51, 1, 263-73.

8 Dadová, J., Cahová, H., Hocek, M., *Nucleic Acids and Molecular Biology Series,* 31, Springer **2016**, 123 – 144.

9 Ondruš M, Sýkorová V, Bednárová L, Pohl R, Hocek M., *Nucleic Acids Res.*, **2020**, 48(21):11982-11993.

10 Wu Y, Fa M, Tae EL, Schultz PG, Romesberg FE., *J Am Chem Soc.*, **2002**, 124(49):14626-30.

11 Kore, A. R., Shanmugasundaram, M., Senthilvelan, A., & Srinivasan, B., *Nucleosides, Nucleotides & Nucleic Acids*, **2012**, 31(5), 423–431.

12 Yoshikawa M., Kato T., Takenishi T. *Tetrahedron Letters*, **1967**, 8, 50, 5065-5068.

13 Hocek M., Fojta M. *Org. Biomol. Chem.*, **2008**, 6, 2233-224.

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

15 Hasník Z., Pohl, R., Klepetářová B., Hocek M., *Synthesis*, **2009**, 1309-1317.

16 Dadová, J.; Vidláková, P.; Pohl, R.; Havran, L.; Fojta, M.; Hocek, M. *J. Org. Chem.* **2013**, *78*, 9627 – 9637.

17 a) Xue L., Karpenko I. A., Hiblot J. and Johnsson K., *Nat. Chem. Biol.*, **2015**, 11 , 917 —923; b) Li J. and Chen P. R., *Nat. Chem. Biol.*, **2016**, 12, 129 —137; c) Krall N., da Cruz F. P, Boutureira O. and Bernardes G. J. L., *Nat. Chem.*, **2016**, 8, 103 —113; d) Spicer C. D. and Davis B. G., *Nat. Commun.*, **2014**, 5, 4740.

18 Lang K., and Chin V, *Chem. Rev.*, **2014**, 114, 4764 —4806.

19 Rostovtsev V. V., Green L. G, Fokin V. V. and Sharpless K. B, *Angew. Chem., Int. Ed.*, **2002**, 41, 2596 —2599.

20 Tornoe C. W., Christensen C. and Meldal M., *J. Org. Chem.*, **2002**, 67, 3057 —3064.

21 Agard N. J., Prescher J. A. and Bertozzi C. R., *J. Am. Chem. Soc.*, **2004**, 126, 15046 —15047.

22 Agard J. N., Prescher J. A., Bertozzi C.R. *J. Am. Chem. Soc.* **2004**, 126, 46, 15046–15047.

23 Johansson J. R., Beke-Somfai T., Stalsmeden S. A., Kann N. *Chem. Rev.* **2016**, 116, 23, 14726–14768.

24 Neef A. B., Luedtke N. W. *ChemBioChem*, **2014**, 15, 789-793.

25 Obeid S., Buβkamp H., Welte W., Diederichs K., Marx A. *Chem.Comm.,* **2012**, 48, 8320-8322.

26 Panattoni, A.; Pohl, R.; Hocek, M. *Org. Lett.* **2018**, *20*, 3962 - 3965.

27 Qu D., Wang G., Wang Z., Zhou L., Chi W., Cong S., Ren X., Liang P., Zhang B. Anal. Biochem. **2011**, 417, 112-121.

28 Slavíčková M., Janoušková M., Šimonová A., Cahová H., Kambová M., Šanderová H., Krásný L., Hocek M. *Chem. Eur. J.* **2018**, *24*, 8311 - 8314.

29 Saxon E. and Bertozzi C. R., *Science*, **2000**, 287 , 2007 —2010.

30 Nilsson B. L., Kiessling L. L. and Raines R. T., *Org. Lett.*, **2000**, 2, 1939 —1941.

31 Saxon E., Armstrong J. I. and Bertozzi C. R., *Org. Lett.*, **2000**, 2, 2141 —2143.

32 Van Berkel S. S. , M. B. van Eldijk M. B. and van Hest J. C., *Angew. Chem., Int. Ed.*, **2011**, 50, 8806 —8827.

33 Axup J. Y., Bajjuri K. M., Ritland M., Hutchins B. M., Kim C. H., Kazane S. A., Halder R., Forsyth ., J. S., Santidrian A. F., Stafin K., Lu Y., Tran H. , Seller A. J., Biroc S. L., Szydlik A., Pinkstaff J. K., Tian F., Sinha S. C., Felding-Habermann B., Smider V. V. and Schultz P. G., *Proc. Natl. Acad. Sci. U. S. A.*, **2012**, 109 , 16101 —16106.

34 Song W., Wang Y., Qu J., Madden M. M. and Lin Q. , *Angew. Chem., Int. Ed.*, **2008**, 47, 2832 —2835.

35 Herner A. and Lin Q., *Top. Curr. Chem.*, **2016**, 374, 1.

36 MacKenzie D. A., Sherratt A. R., Chigrinova M., Cheung L. L., and Pezacki J. P. , *Curr. Opin. Chem. Biol.*, **2014**, 21 , 81 —88.

37 M. Yang , J. Li and P. R. Chen , *Chem. Soc. Rev.*, **2014**, 43, 6511 —6526.

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

39 Carboni R. A. and Lindsey R. V. *J. Am. Chem. Soc.*, **1959**, 81, 4342 —4346.

40 B.L. Oliveira, Z. Guo, Bernandes G. J. L. *Chem. Soc. Rev.*, **2017**, 46, 4895-4950.

41 a) D. L. Boger , R. P. Schaum and R. M. Garbaccio , *J. Org. Chem.*, **1998**, 63, 6329 —6337;
b) A. Hamasaki, R. Ducray and D. L. Boger, *J. Org. Chem.*, **2006**, 71 , 185 —193.

42 D. Wang , W. Chen , Y. Zheng , C. Dai , K. Wang , B. Ke and B. Wang , *Org. Biomol. Chem.*, **2014**, 12 , 3950 —3955.

43 J. Balcar , G. Chrisam , F. X. Huber and J. Sauer , *Tetrahedron Lett.*, **1983**, 24, 1481 —1484.

44 J. Balcar , G. Chrisam , F. X. Huber and J. Sauer , *Tetrahedron Lett.*, **1983**, 24, 1481 —1484.

45 L. R. Domingo, M. T. Picher and J. A. Saez, *J. Org. Chem.*, **2009**, **7**4, 2726 —2735.

46 Galeta J. R., Dzijak J., Obořil M., Dračínský M., Vrábel M. *Chemistry - A European Journal*, **2020**, 26 (44): 9945-9953.

47 Sauer J. and Lang D., *Angew. Chem., Int. Ed. Engl.*, **1964**, **7**6 , 603.

48 Sauer J., Heldmann D. K., Hetzenegger J., Krauthan J., Sichert H. and Schuster J., *Eur. J. Org. Chem.*, **1998**, 2885 —2896.

49 Soenen D. R., Zimpleman J. M. and Boger D. L., *J. Org. Chem.*, **2003**, 68, 3593 —3598.

50 a) F. Liu , Y. Liang and K. N. Houk, *J. Am. Chem. Soc.*, **2014**, 136 , 11483 —11493;

b) F. Liu, R. S. Paton , S. Kim , Y. Liang and K. N. Houk, *J. Am. Chem. Soc.* **2013**, 135 , 15642 —15649.

51 a) Wagner J. A., Mercadante D., Nikic I., Lemke E. A. and Grater F., *Chem. – Eur. J.*, **2015**, 21 , 12431 —12435; b) Hoffmann J. E., Plass T., Nikic I., Aramburu I. V., Koehler C., Gillandt H., Lemke E. A. and Schultz C., *Chem. – Eur. J.*, **2015**, 21, 12266 —12270; c) Li J., Jia S. and Chen P. R., *Nat. Chem. Biol.*, **2014**, 10, 1003 —1005; d) Versteegen R. M., Rossin R., ten Hoeve W., Janssen H. M. and Robillard M. S., *Angew. Chem., Int. Ed.*, **2013**, 52 , 14112 —14116.

52 Reiner T, Zeglis B. M. The inverse electron demand Diels-Alder click reaction in radiochemistry. *J Labelled Comp Radiopharm*. **2014**, 57(4):285-90.

53 Qiu L., Mao W., Yin H., Tan H., Cheng D., Shi H. *Contrast Media Mol Imaging.* **2019**, 25, 9182476.

54 Kang D., Lee S., Kim J. *Chem*, **2022**, 8, 8, 2260-2277.

55 Zhu L, Mahato RI. *Expert Opin Drug Deliv*. **2010**, 7(10):1209-26.

56 Cai X, Dou R, Guo C, Tang J, Li X, Chen J, Zhang J. **2023**, *Pharmaceutics*,15(5):1502.

57 a) Zawada Z., Tatar A., Mocilac P., Budesinsky M., Kraus T. *Angew. Chem., Int. Ed.* **2018**, *130*, 10039– 10043. b) Güixens-Gallardo P., Zawada Z., Matyašovský J., Dziuba D., Pohl R., Kraus T., Hocek M. *Bioconjugate Chem.* **2018**, *29*, 3906– 3912. c) Kužmová E., Zawada Z., Navrátil M., Günterová J., Kraus T. *Anal. Biochem.* **2021**, *614*, 114002.

58 Wang G., Wang F., Huang Q., Li Y., Liu Y., Wang Y. *Biomed Res Int.* **2015**; 2015:757530.

59 Aloisi C.M.N., Escher N.A., Kim H.S., Geisen S.M., Fontana G.A., Yeo J.E., Schärer O.D., Sturla S.J. *DNA Repair (Amst).* **2022**, 110:103262.

60 Weickert P., Stingele J. *Annu Rev Biochem.* **2022**; 91:157-181.

61 Brown A, Kumar S, Tchounwou PB. *J Cancer Sci* Ther, **2019**, 11(4):97.

62 Ward J.F. *Int. J. Radiat Biol.* **1994**, 66:427–32.

63 Pommier Y, Sun Y, Huang SN, Nitiss JL. *Nat. Rev. Mol. Cell Biol,* **2016**.

64 Jacobs AL, Schär P. *Chromosoma*; **2012**, 121(1):1-20.

65 Smith S. S., Kaplan B. E., Sowers L. C., Newman E. M. *Proc. Natl. Acad. Sci. U. S. A.* **1992**, 89, 4744– 4748.

66 Jung Y., Lippard S. J. *Chem. Rev.* **2007**, 107, 1387– 1407.

67 Huang H., Woo J., Alley S. C., Hopkins P. B. *Bioorg. Med. Chem.* **1995**, 3, 659– 669.

68 Yeo J. E., Wickramaratne S., Khatwani S., Wang Y. C., Vervacke J. S., Distefano M. D., Tretyakova N. Y. *ACS Chem. Biol*., **2014**, 9, 1860– 1868.

69 Ivancová I., Leone D-L., Hocek M., *Current Opinion in Chemical Biology,* **2019**, 52, 136-144.

70 Daďová J., Orság P., Pohl R., Brázdová M., Fojta M., Hocek M. *Angew. Chem. Int. Ed*. **2013**, *52*, 10515-10518.

71 Olszewska A., Pohl R., Brázdová M., Fojta M., Hocek M. *Bioconjugate Chem.* **2016**, *27*, 2089 – 2094.

72 Brunderová M., Krömer M., Vlková M., and Hocek, M. *Angew. Chem. Int. Ed.* **2023**, *62*, e202213764.

73 Raindlová V., Pohl R., Hocek M. *Chem. Eur. J.* **2012**, *18*, 4080-4087.

74 a) Matyašovský J., Pohl R., Hocek M., *Chem. Eur. J.* **2018**, *24*, 14938-14941; b) Matyašovský, J.; Hocek, M., *Org. Biomol. Chem.* **2020**, *18*, 255-262.

75 Ivancová, I.; Pohl, R.; Hubálek, M.; Hocek, M. *Angew. Chem. Int. Ed.* **2019**, *58*, 13345-13348.

76 Leone D.-L., Hubálek M., Pohl R., Sýkorová V., Hocek, M. *Angew. Chem. Int. Ed.* **2021**, *60*, 17383-17387.

77 Leone D.-L., Pohl R., Hubálek M., Kaderábková M., Krömer M., Sýkorová V., Hocek M. *Chem. Eur. J.* **2022**, *28*, e202104208.

78 Rajagopalan S, Wang C, Yu K, Kuzin AP, Richter F, Lew S, Miklos AE, Matthews ML, Seetharaman J, Su M, Hunt JF, Cravatt BF, Baker D. *Nat Chem Biol.,* **2014**, 10(5):386-91.

79 Bachovchin DA, Cravatt BF. *Nat Rev Drug Discov.,* **2012**, 11(1):52-68.

80 a) Beardsell M., Hinchliffe P.S., Wood J. M., Wilmouth R., Shofield C. J., Page M.I., *Chem. Commun.*, **2001**, 497-498. b) Kolb R., Bach C. N., Sieber A. S. *Chem. Commun.*, **2014**, 50, 427-429.

81 Barwick M., Tarek A-Z.and Novak I. *J. Phys. Chem. A*, **2008**, 112, 10993–10997**.**

82 Llinas A., Ahmed N., Cordaro M., Laws A. P., Frere J. M., Delmarcelle M., Silvaggi N. R., Kelly J. A. and. Page M. I, *Biochemistry*, **2005**, 44, 7738–7746.

83 Tsang W.Y., Ahmed N., Hemming K., Page M.I. *Org Biomol Chem*, **2007**, 5(24):3993-4000

84 Page M. I., Hinchliffe P. S., Wood J. M., Harding L. P. and Laws A. P. *Bioorg. Med. Chem. Lett.*, **2003,** 13, 4489–4492.

85 Carvalho A.R.L, Almeida T. V., Brito A.J., Lurn M.K., Oliveira T.F., Guedes R.C., Goncalves L.M., Lucas D.S., Cravatt B.F., Archer M. and Moreira R. *ACS Chem. Biol.* **2020**, 15, 4, 878–883.

86 Vantourout J. C., Adusumalli S. R.,. Knouse K. W, Flood D. T., Ramirez A., Padial N. M., Istrate A., Maziarz K., Justine N., Merchan R.R., Qiao J. X., Schmidt M. A., Deery M. J., Eastgate M. D., Dawson P. E., G. Bernardes J. L., Baran P. S., *J. Am. Chem. Soc.* **2020**, 142, 17236–17242.

87 a) Cheng K., Lee J., Hao P., Yao S.Q., Ding K., Li Z. *Angew. Chem. Int. Ed.* **2017**, 56, 15044–15048; b) Marjanovic J., Lewandowski T.M., Marin V., Patterson M., Miesbauer L., Ready D., Williams J., Vasudevan A., Lin Q. *J. Am. Chem. Soc.* **2016**, 138, 14609–14615.

88 Ma N., Hu J., Zhang Z.M., Liu W., Huang M., Fan Y., Yin X., Wang J., Ding K., Ye W., Li Z. J Am Chem Soc., **2020**, 142(13):6051-6059.

89 Sornay C., Hessmann S., Erb S., Dovgan I., Ehkirch A., Botzanowski T., Cianférani S., Wagner A. and Chaubet G. *Chem. Eur. J.,* **2020**, 26, 13797–13805.

90 King J. L. and Jukes T. H. *Science*, **1969**, 164 , 788 —798.

91 Antos J. M., Francis M. B. *Curr Opin Chem Biol.* **2006** 10, 3, 253-62.

92 Perez-Rizquez C., Abian O., Palomo J. M. *Chem. Commun.*, **2019**, 55, 12928-12931.

93 a) Seki Y., Ishiyama T., Sasaki D., Abe J., Sohma Y., Oisaki K., Kanai M. *J. Am. Chem. Soc.* **2016**, *138*, 10798–10801. b) Maruyama K., Malawska K. J., Konoue N., Oisaki K., Kanai M. *Synlett* **2020**, *31*, 784–787.

94 Portas CM, Bjorvatn B, Ursin R. *Prog Neurobiol.*, **2000**, 60(1):13-35.

95 Gratzner H. G. *Science* **1982**, *218*, 474– 475.

96 Mead T. J. and Lefebvre V. *Methods Mol Biol.*, **2014**, 1130:233-243.

97 Ligasová A., and Koberna K. *Molecules*, **2018**, *23*, 3007.

98 Fantoni N. Z., El-Sagheer A. H., Brown T. *Chem. Rev.* **2021**, *121*, 7122– 7154.

99 Krell K., Harijan D., Ganz D., Doll L., Wagenknecht H.A. *Bioconjugate Chem.* **2020**, 31, 990–1011.

100 Geng, P.; List, E.; Rönicke, F.; Wagenknecht, H.A. *Chem.─Eur. J.* **2022**, 29, e202203156.

101 Bujalska, A.; Basran, K.; Luedtke, N. W. *RSC Chem. Biol.* **2022**, 3, 698– 701.

102 Tera, M.; Glasauer, S. M. K.; Luedtke, N. W. *ChembioChem* **2018**, 19, 1939– 1943.

103 Verdine G. L., Norman D.P., *Annu. Rev. Biochem.,* **2003**, 72, 337-66.

104 Winnacker M., Breeger S., Strasser R., Carell T. *ChemBioChem*, **2009**, 10, 109-18.

105 Lee H.S., Dimla D., Schultz P. G., *Bioorg. Med. Chem. Lett*., **2009**, 19, 5222-5224.

106 Lagrange T., Kim T. K., Orphanides G., Ebright Y.W., Ebright R.H., Reinberg D., *PNAS* **1996**, 93, 10620-10625.

107 Niemeyer C. M. *Angew. Chem. Int. Ed.*, **2010**, 49, 1200-1216.

108 Krömer M., Bártová K., Raindlová V., Hocek M. *Chem. Eur. J.* **2018**, 24, 11890-11894.

109 Garg, N. K., Woodroofe, C. C., Lacenere, C. J., Quake, S. R., and Stoltz, B. M. *Chem. Commun.,* **2005***,* 4551-4553.

110 Jäger S., Rasched G., Kornreich-Leshem H., Engeser M., Thum O., Famulok M. *J. Am. Chem. Soc.* **2005**, *127*, 15071– 15082.

111 Loehr M. O., and Luedtke N. W. *Angew. Chem. Int.* **2022**, 61, 22.

112 Bakkum T., van Leeuwen T., Sarris A. J. C., Van Elsland D. M., Poulcharidis D., Overkleeft, H. S., van Kasteren S. I. *ACS Chem. Biol.* **2018**, *13*, 1173– 1179.

113 Brázdilová P., Vrábel M., Pohl R., Pivonková H., Havran L., Hocek M, Fojta M. *Chemistry*, **2007**, 13(34):9527-33.

114 Elliott AD. *Curr Protoc Cytom.* **2020**, 92(1):e68.

115 Kuba, M.; Pohl, R.; Kraus, T. and Hocek, M. *Bioconjugate Chem.* **2023**, *34 (1)*, 133-139.

116 Ligasová A, Frydrych I, Koberna K. Basic Methods of Cell Cycle Analysis. *Int J Mol Sci.* **2023**, 24(4):3674.

117 Neef A. B., Luedtke N. W. *ChemBioChem* **2014**, 15, 789-793.

118 Borsenbenger V., Kukwikila M. and Howorka S. *Org. Biomol. Chem.*, **2009**, **7**, 3826-3835.

119 a) Miwa Y., Fujita Y. *J Biol Chem*., **1988**, 15, 263(26):13252-7. b) Fujita Y, Fujita T, Miwa Y, Nihashi J, Aratani Y. *J Biol Chem.*, **1986**, 261(29):13744-53.

120 Rigali S., Derouaux A., Giannotta F., Dusart J. *J Biol Chem.*, **2002**, 12;277(15):12507-15.

121 Horne C.R., Venugopal H., Panjikar S., Wood D.M., Henrickson A., Brookes E., North R.A., Murphy J.M., Friemann R., Griffin M.D.W., Ramm G., Demeler B., Dobson R.C.J. *Nat Commun*., **2021**, 31;12(1):1988.

122 Jumper J., Evans R., Pritzel A., Green T., Figurnov M., Ronneberger O., Tunyasuvunakool K., Bates R., Žídek A., Potapenko A., Bridgland A., Meyer C., Kohl S., Ballard J. A., Cowie A., Romera-Paredes B., Nikolov S., Jain R., Adler J., Back T., Petersen S., Reiman D., Clancy E., Zielinski M., Steinegger M., Pacholska M., Berghammer T., Bodenstein S., Silver D., Vinyals O., Senior A.W., Kavukcuoglu K., Kohli P., Hassabis D. *Nature*, **2021**, 596, 583–589.

123 Marty M. T., Baldwin A. J., Marklund E. G., Hochberg G. K. A., Benesch J. L. P., Robinson C. V. *Anal. Chem.* **2015**, 87, 8, 4370–4376.

124 Kuba M., Khoroshyy P., Lepšík M., Kužmová E., Kodr D., Kraus T. and Hocek M. *Angew. Chem. Int.*, **2023**, 62, 38. e202307548.

125 Písačková J., Procházková K., Fábry M., Řezáčová, P. *Cryst. Growth Des.* **2013**, 13, 844-848.

126 Procházková K., Čermáková K., Pachl P., Sieglová I., Fábry M., Otwinowski Z. Řezáčová P. *Acta Crystallogr D Biol Crystallogr* **2012**, 68, 176-185.

127 Ericsson, U. B., Hallberg, B. M., Detitta, G. T., Dekker, N. & Nordlund, P. *Anal Biochem* **2006**, 357, 289-298.

128 Hernández H, Robinson CV. *Nat Protoc.*, **2007**, 2(3):715-26.