Exploring copper(I) as a catalyst for nucleic acid alkylation

Inauguraldissertation

zur

Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch-Naturwissenschaftlichen Fakultät

der Universität Basel

von

Stefanie N. Geigle

aus Deutschland

Basel, 2018

Originaldokument gespeichert auf dem Dokumentserver der Universität Basel

edoc.unibas.ch

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät auf Antrag von

Prof. Dr. Dennis G. Gillingham

Prof. Dr. Florian P. Seebeck

Basel, den 23. Mai 2017

Prof. Dr. Martin Spiess
Dekan der Philosophisch-Naturwissenschaftlichen Fakultät

Nur wenige wissen, wie viel man wissen muss,
um zu wissen, wie wenig man weiß.

Werner Heisenberg (1901 – 1976)

Acknowledgement

I would like to thank *Prof. Dr. Dennis G. Gillingham* for the opportunity to do my PhD in his group. I am deeply grateful for the interesting project he allowed me to independently work on, but also to evolve my skills and knowledge in biochemistry.

I would like to thank *Prof. Dr. Florian P. Seebeck* for accepting the co-examination of my PhD thesis and for the great atmosphere in the third floor, as well as the knowledge transfer for biological related problems.

I would like to thank the whole Gillingham, Seebeck and Mayor groups for the great atmosphere in and outside the university. I am thankful for the time we spent together, you all made my time at the university as an unbelievable good memory (especially the legendary mayor parties).

I am grateful for all the students *Daniel Moser, Florian Leisinger* and *Cedric Stress* who worked with me and contributed to this work.

I want to thank *Anja Stampfli*, *Basilius Sauter* and *Kiril Tishinov* for taking their time to proof-read this thesis.

Furthermore, I would like to thank *Dr.Daniel Häusinger* and his group who always supported me with their excellent NMR service, *Dr. Heinz Nadig* for measuring HRMS and also allowing me to use his device, the secretary and Werkstatt team for taking care of management and infrastructures.

I want to thank all my colleagues and friends who joined me on this roller coaster of joy, frustration and happiness. Special thanks and kisses go to the four Ts of the petrol station: *Ina Bodoky*, *Samantha Brianza*, *Livia Glanzamm* and *Annika Büttner*. From the beginning of our studies we always supported each other and I am really thankful to meet such awesome personalities. Especially, I will miss our extensive learning sessions with the never ending fights about the right answer. Thank you for feeling always emotionally-supported also besides the studies; our friendships are really special and will continue until we are old and ugly, love ya my dick popos!

I would like to say thank you to *Ronan Rocaboy* who came in my live at the right time and supported me always with food, jokes and happiness, I love you little fox!

My biggest thanks go to my family, my parents *Rike* and *Detlef* and my siblings *Franzi*, *Tobi* and *Sebi*, who always believed in me. Thank you for kicking out the stones from this roller coaster for me, when it was necessary. You always respected my choices and supported me in every direction I decided to go; without you my life would not be as happy and successful.

Abstract

Chemical manipulations of nucleic acids have been crucial in the development of methodologies to study and understand epigenetics and transcriptomes. However, a complete understanding of these complex and dynamic systems is still missing and therefore we believe new synthetic tools for selective nucleic acid engineering are required. We report here the discovery that copper(I) carbenes derived from α-diazocarbonyl compounds selectively alkylate the O^6-position of guanine (O^6-G) in mono- and oligonucleotides. This new methodology allows the targeting of only purine-type lactam oxygen, whereas other types of amides or lactams are poorly reactive under the smooth guanine alkylation conditions. Mechanistic studies point to a substrate-directed alkylation reaction with the N7G as the key functionality of the purine-nucleobase that controls the high chemoselectivity. We used copper(I)-catalyzed O^6-G alkylation to readily engineer O^6-G derivatives to study two open questions in the biochemistry of O^6-G adducts: the repair by alkylguanine transferases and their incorporation efficiency during DNA replication. Furthermore, with a reactivity and stability screen of functionalized water soluble N-heterocyclic carbene ligands we could identify tight-binding variants that permanently stabilized copper(I) in water. This system allowed alkylation of oligonucleotides in a reducing agent free environment maintaining the selectivity for the O^6-G position. Most importantly, the tight-binding ligands that stabilize copper(I) provide access to further functionalization such as bio-conjugation. Given the importance of O^6-G lesion in biology and the need for simple methods to engineer nucleic acids, we believe that copper(I)-catalyzed O^6-G alkylation will find broad applicability.

Table of Contents

List of abbreviations

AA	Amino acid
Ac_2O	Acetic anhydride
A	Adenine
AdoMet	S-Adenosylmethionine
Ag_2O	silveroxide
AGT	Alkylguanine transferase
Ala	Alanine
Arg	Arginine
Asn	Asparagine
Asp	Aspartate
BCA	Bathocuproinedisulfonic acid disodium salt
BrAcBr	Bromoacetyl bromide
Bu	Butyl
CuAAC	Copper(I) alkyne- azide cycloaddition
CH_2Cl_2	Dichloromethane
CH_2O	Formaldehyde
CO	Carbon monoxide
Cs_2CO_3	Casium carbonate
Cu	Copper
CuCl	Copper chloride
CuO	Copperoxide
Cu(OAc)	Copper diacetate
$Cu(II)SO_4$	Copper(II)sulfate
Cys	Cysteine
C	Cytosine
DAAs	Diazoacetamides
DBU	1,8-Diazabicyclo(5.4.0)undec-7-ene
DFT	Density functional theory
DIPEA	N, N-diisopropylethyl amine
DMF	Dimethylformaide
DMSO	Dimethylsulfoxide
DMT	4,4`-Dimethoxytrityl
DNA	2`-Deoxyribonucleic acid

dNTP	2`-Deoxynucleoside triphosphate
(d)N*TP	Modified (2`-deoxy)nucleoside triphosphate
ds	Double-stranded
DTT	Dithiothreitol
EDA	Ethyl α-diazoacetate
EDC	1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide
EDTA	Ethylenediaminetetraacetic acid
ee	Enantiomeric excess
Et	Ethyl
EtOH	Ethanol
Ethylcm	Ethyl carboxymethyl
ESI	Electrospray ionization
G	Guanine
Gln	Glutamine
Glu	Glutamate
GlaTgs	Trimethylguanosine synthases
Gly	Glycine
hAGT	Human AGT
HBF_4*OEt_2	Tetrafluoroboric acid diethyl ether complex
HBr	Hydrogen bromide
HCl	Hydrogen chloride
$HC(OEt)_3$	Triethyl orthoformate
H_2O	Water
H_2SO_4	Sulfuric acid
His	Histidine
HMBC	Heteronuclear multiple-bond correlation
HOBt	1-Hydroxybenzotriazol
HPLC	High performance liquid chromatography
I	Inosine
KI	Kalium iodide
KOtBu	Potassium tertbutoxide
LiOH	Lithium hydroxide
Lys	Lysine
MALDI	Matrix-assisted laser desorption/ionization

Me	Methyl
MeCN	Acetonitrile
MES	2-(N-morpholino)ethanesulfonic acid
Met	Methionine
MeOH	Methanol
m^5C	5-Methylcytosine
MNNG	N-methyl-N`-nitro-N-nitrosoguanine
MNNU	N-nitroso-N-methylurea
MP	Monophosphate
mRNA	Messenger RNA
MS	Mass spectrometry
MTases	Methyltransferases
MTIC	Monomethyl triazene
NA	Nucleic acid
Na_2CO_3	Sodium carbonate
NaCl	Sodiumchloride
NaH	Sodium hydride
NaN_3	Sodiumazide
NaOAc	Sodium acetate
NEt_3	Triethylamine
NH_3	Ammoniak
NH_4Cl	Ammonium chloride
NHC	N-heterocyclic carbene
NH_2OH	Hydroxylamine
NHI	Nitrogen hydrogen insertion
NMR	Nuclear magnetic resonance
nt	Nucleotides
NTP	Nucleoside triphosphate
Nu	Nucleophilic group
OHI	Oxygen hydrogen insertion
OSPS	Oligonucleotide solid phase synthesis
OTf	Trifluoromethanesulfonate
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffer saline

PCR	Polymerase chain reactions
PEX	Primer extension experiments
PF_6	Hexafluorophosphate
Phe	Phenylalanie
PMHS	Polymethyldydrosilan
$POCl_3$	Phosphorylchloride
quant	quantitative
Rh_2OAc_4	Dirhodium tetraacetate
RNA	Ribonucleic acid
RP	Revere phase
Ser	Serine
S_N	Nucleophilic substitution
ss	Single-stranded
T	Thymine
TAMRA	Carboxytetramethylrhodamine
TBDMS	*Tert*-butyldimethysiliyl
*t*BuNHOH	*Tert*-butylhydroxyl amine
^{t}BuOH	*Tert*-butanol
TBE	Tris/Borate/EDTA
TEAA	Triethylamonium acetate
THPTA	Tris(3hydroxypropyltriazolylmethyl)amine
TMZ	Temozolomide
THF	Tetrahydrofuran
tht	Tetrahydrothiophene
TEP	Tolman electronic parameter
Tris	Tris(hydroxymethyl)aminomethane
tRNA	Transfer RNA
Trp	Tryptophan
TsNHNHTs	*N,N*-ditosylhydrazine
Tyr	Tyrosine
U	Uracil
UPLC	Ultra performance liquid chromatography
WC	Watson-Crick
XHI	X-H insertion

Chapter1. Introduction: synthetic approaches to chemically modified nucleic acids

Nucleic acids are outstanding molecules created by nature to hold all information about the growth, development, functioning and reproduction of a living organism. An additional, recently appreciated, facet of these informational molecules is that their function can be regulated by chemical modification. For example the inhibition of gene expression in cells is associated with methylation of cytosine.[1] Many natural transcriptomic and genomic modifications have been found. The complexity and dynamics of these systems makes them difficult to study and a complete understanding of their function is still missing.[2] One way to study nucleic acid functioning is by selective modification, thus creating a need for new chemistry to selectively target nucleic acids. Selective NA engineering is not only important to study their biological relevance it can be also applied for therapeutic issues such as chemical gene regulation, which can be important for diseases that are caused by overexpression of certain genes.[3]

Solid-phase synthesis of oligonucleotides was a paradigm-shifting change for building short nucleic acids. This approach also allows the introduction of a huge variety of modifications but is often limited by the size of the molecule and the ease of the non-canoncial phosphoramidite synthesis. The direct modification of nucleic acids (NAs) may be more efficient especially for large complex oligonucleotides. Therefore, we are interested in the development of novel ways to directly target both: deoxyribonucleic acid (DNA) and ribonucleic acid (RNA). Evolving selective modification strategies of NAs is difficult due to multiple similar functionalities, however, new reactivities of the NAs may be revealed and can be used to change their biological role.

1.1 Structure and key physical properties of nucleic acids

Before we have a detailed look at the reaction properties of nucleic acids, the general structure of the polymers will be described. Nucleic acids are not as diverse as protein since they only consist of four repeating building blocks: adenine (A), guanine (G), cytosine (C) are common in both DNA and RNA, while thymine (T) and uracil (U) are specific for DNA and RNA, respectively. These so called nucleobases are connected to a sugar moiety, the ribose (or deoxyribose in the case of DNA). Linkage of the sugars via phosphodiester bonds comprise the backbone of the nucleic acids. DNA duplexes, built out of two single-stranded DNAs in antiparallel direction, that are right-handed helices stabilized by Watson-Crick (WC) hydrogen bonding in different tertiary structures (Figure 1). The most common tertiary structure of dublex DNA is the B-form. To expand the structural diversity in dublex DNA, Hoogsten base-pairs can be formed by a conformation flip of the purine bases: adenine and guanine (Figure 1).[4] In addition to the four canonical blocks in DNA and RNA, a 5th nucleoside was detected: inosine (I). While inosine is a premutagenic event

in DNA (false base-pairing to C) it is an essential modification in RNA.[5] Inosine is introduced by enzymes via deamination of adenine, one function seems to be to offer higher transcriptome diversity.

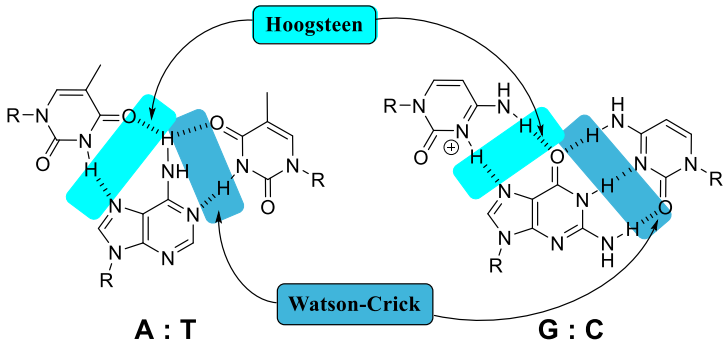


Figure 1. Watson-Crick base pairing that stabilize the double helices in DNA: adenine (A) and thymine (T), guanine (G) and cytosine (C). R corresponds to the deoxyribose. Hoogsteen base pairing that results from a conformation flip of the purine bases: A and G.

For a better understanding of how nucleic acids can react it is important to know their key physical properties. If we look at the first proton association of adenosine monophosphate (AMP) and cytidine monophosphate (CMP) they are protonated at a pH around 3.5-4.2 (Figure 2A), whereby, the other bases are protonated at pH of 2.[6] Comparing the pK_a values of both monophosphates (MPs), CMP is slightly more basic than AMP. However, in a longer DNA or RNA strand the nucleobases in proximity can have an influence on the basicity and switch the order.[7] RNA should not be treated with strong acids because at a pH below 5 the RNA starts to decompose. Summarized protonation in a nucleic acid occurs first at the N1 position of adenine and N3 position of cytosine.

The first deprotonation in DNA and RNA occurs on the amide like functionality in inosine › guanine ≈ uracil › thymine with inosine as the most acidic nucleobase (acidity order) (Figure 2B). Again, the differences in the pK_a values are small enough that the neighboring nucleobases can have an influence on the order of acidity. A second protonation or deprotonation is not described for the nucleobases because these would be in unrealistic pH ranges.

Depending on the environment a pK_a perturbation of the nucleic acid can occur, thereby change the reactivity of the oligonucleotide. For example, in DNA duplexes, the pK_a values of the nucleobases are shifted to neutrality[7]. The uncharged nucleobases in double-stranded (ds) DNA are important, for example, for the stability of the genome because structural mistakes that result from proton gain or loss might lead to replication errors.[8] RNA is different because it can exist as unstructured single-strands or more complex folded molecules with single- and double-stranded regions. The folded structures in RNA can create protein

like pockets that result in pK_a shifts. Furthermore, wobble or Hoogsten base-pairing can shift nucleobase pK_as. To achieve catalysis RNA largely relies on co-factors or acid-base chemistry.[9] For example the hepatitis delta virus ribozyme shows a catalytically essential cytidine with pK_a shifted from 4.1 to 6.4, which plays an essential role in the mechanism of ribozymes.[10-11]

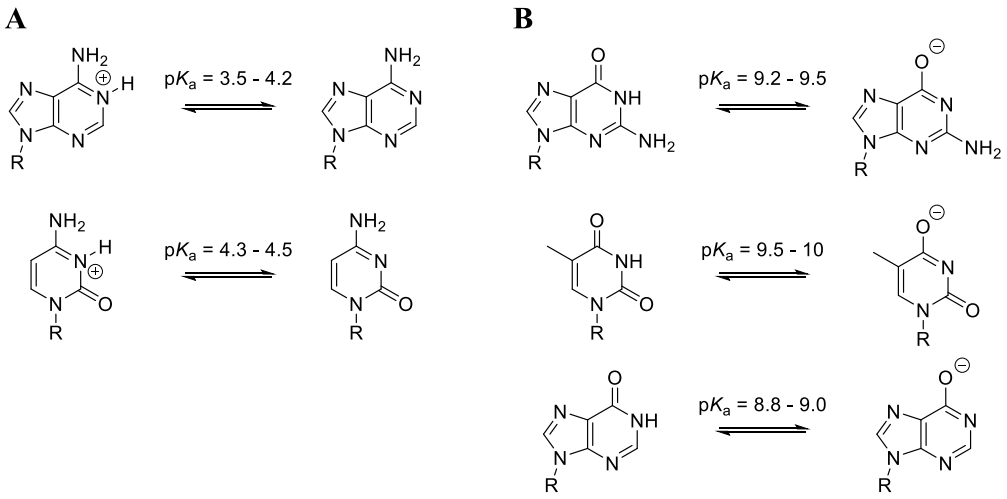


Figure 2. (A) First proton associations in nucleic acids occur at N1 in adenine and N3 in cytosine. (B) First proton dissociations in nucleic acids occur at the lactam-like unit of guanine, thymine and inosine.

Computational studies were also done on the nucleobases to gain an insight into the dipole moments and polarizibalities of the nucleobases.[8] The different calculation methods were not in in agreement for all cases. However, the take home message is that guanine has the highest dipole moment and the greatest polarizability. These findings would suggest guanine as the common target in modification reactions. The following section will support these predictions because guanine is the primary target of oxidants and electrophiles.[8]

1.2 Reaction and modification mechanism of nucleic acids

Generally, there are two main strategies to introduce modifications to nucleic acids: pre- and post-synthetically. Depending on the complexity of the modification and the target as well as the modification site, one strategy can be advantageous to the other. For example pre-synthetic approaches are rather beneficial to introduce selectively single modifications to nucleic acids. An overview about the most important methodologies to introduce modification sites to nucleic acids is shown in Figure 3. The different methods will be discussed in detail in the following chapter.

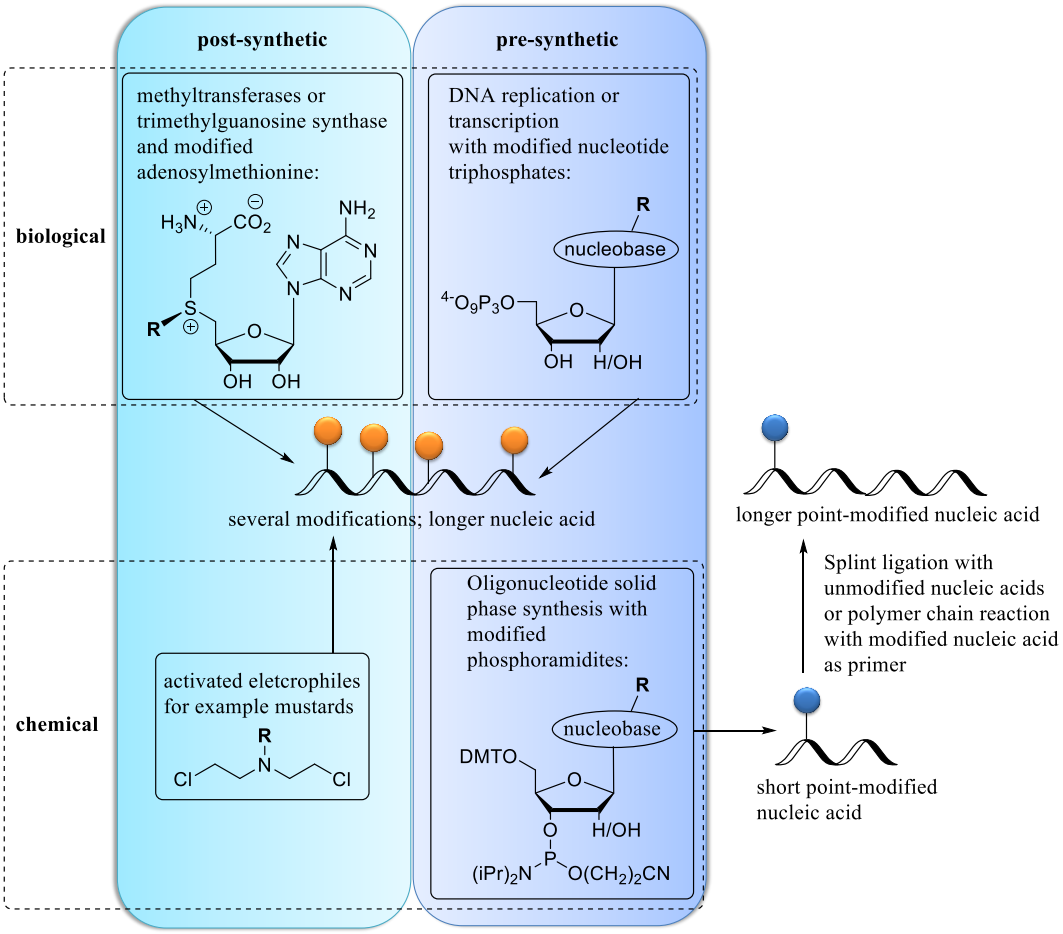


Figure 3. An overview of the most important methodologies to introduce modification sites to nucleic acids.

1.2.1 Modification mechanism of nucleobases

Nucleic acids, bearing a diversity of functional groups, can serve as nucleophiles, electrophiles and also proceed in radical reactions. However, the big challenge in post-synthetically modifying them is to obtain site-selectivity. Generally, post-synthetic modification methods are achieved with activated electrophiles that react with the phosphate backbone or the nucleophilic groups of the nucleobases. If we have a look at the nucleobases, they all bear similar nucleophiles, such as the exocyclic amine in A, G and C; the lactam-like functionalities in G, T and C as well as the N7- and N3- position of the purine bases. To find new reactions that discriminate between all these nucleophiles without the use of protecting groups is very difficult. However, some reactivity properties of NAs can be extracted out of the investigations with different electrophiles.

In DNA duplexes, the helix is used as protection shield and most of nucleophilic groups are rather unreactive because they are already occupied by Watson-Crick base pairing. Guanine is reported as the most electron rich nucleobase and in DNA duplexes it is the primary target of electrophiles.[12] The accessible guanines in the major groove are then modified at the most nucleophilic N7- position.

The nucleobase nucleophilicity of the biopolymers can be better estimated in single-strands due to missing secondary structures and Watson-Crick base pairing. The typical chemical modification sites of NAs with electrophiles are presented in Figure 4A. Again guanine is the most targeted nucleobase and no real site-selectivity can be observed. However, modification of single-stranded NAs with different electrophiles leads to some reactivity trends. For example the change from S_N2 to S_N1 type alkylating agents can change the modification sites (Figure 4C-D). The S_N2-type electrophiles (alkyl sulfonates) predominantly modify N7-G but also target to a much smaller extent the N2-, and N4-G, as well as N1-, N3-, and N7-A. The S_N1 type electrophiles (N-methyl-N`-nitro-N-nitrosoguanidine (MNNG); temozolomide) also mainly modify the N7-G position but O^6-G, N3-A and the phosphate backbone is targeted too. The use of π-electrophiles resulted in inosine, uridine as well as in pseudouridine modification (Figure 4E).

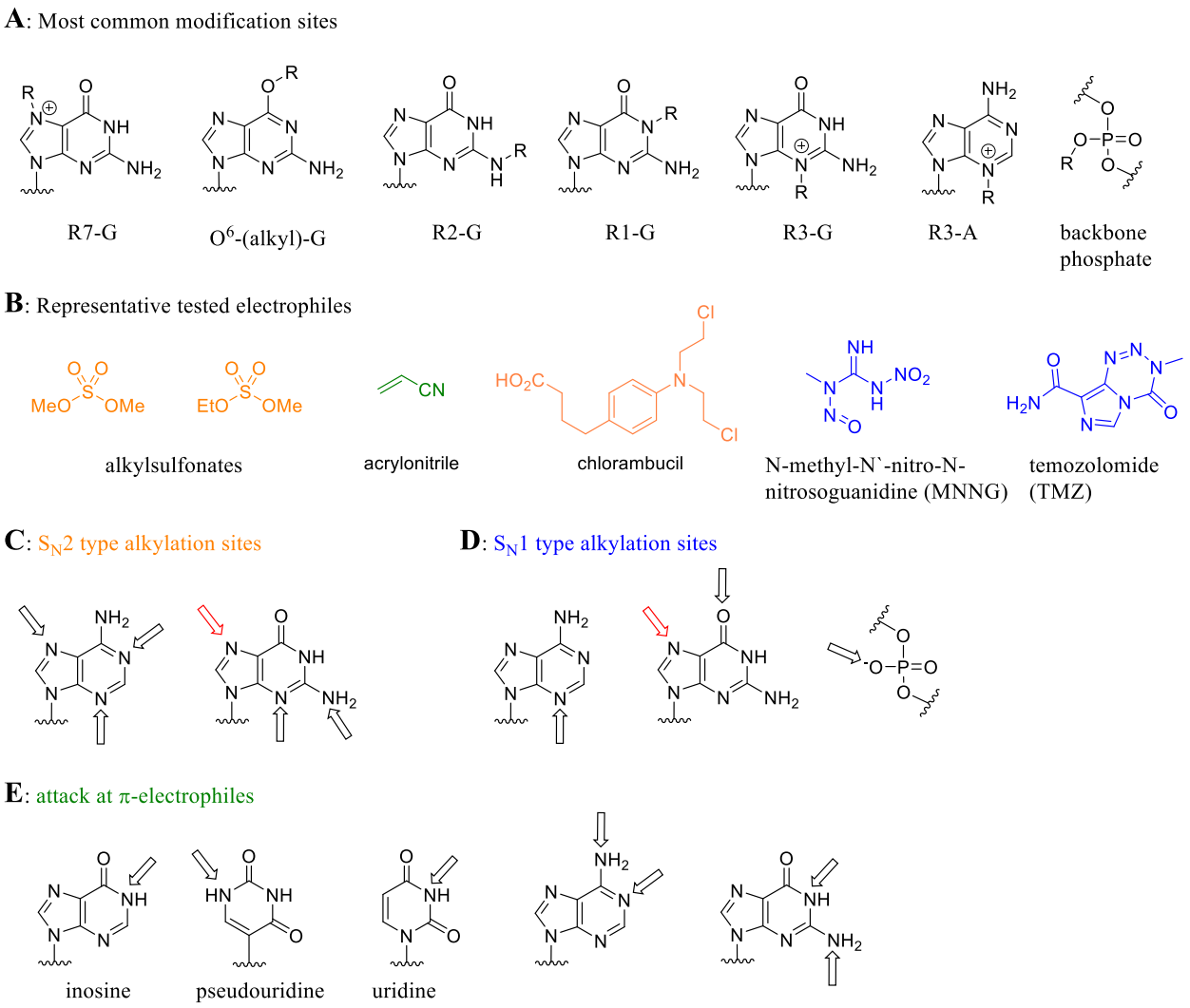


Figure 4. (A) The main nucleophilic sites that are modified in nucleic acids. (B) The representative electrophiles that react with nucleic acids. (C) Alkylation sites of nucleobases with S_N2 types electrophiles. (D) S_N1 types electrophiles. (E) Attack at π-electrophiles.

1.2.2 Alkylation of nucleic acids by natural products

Earlier I presented modification trends of nucleic acids with different common electrophiles to get an insight how these biomolecules possibly act as nucleophiles. Now we will have a look how nature found solutions to target nucleic acids and then see how researchers developed chemical approaches inspired by nature's modification mechanisms.

Considering the essential role of nucleic acids it is not surprising that nature has evolved mechanisms to target them. We will give an overview of alkylation reactions of DNA and RNA that have been evolved by organisms. Nature has found some brilliant solutions to suppress the reactivity of those molecules until it meets its target, the nucleic acid. There are six different classes of natural products that alkylate DNA in a different mechanism: activated cyclopropyl groups, quionone methides, three-membered heteroatom rings, reactive iminium ions, photochemical cycloadditions and diazo compounds. One representative example for every different alkylation class is shown in Figure 5.

(+)-Duocarmycin A bears the dienone cyclopropane motif that is quite stable until it binds to the minor groove of A-T rich strands of DNA.[13] The molecule can only bind to the DNA after amide rotation. This rotation results in increased electrophilicity of the dienone cyclopropane. Then alkylation occurs at the N3-A in close proximity (Figure 5A).[14-15]

Mitomycin C is used as potent cytotoxin against various cancers.[16] A reductive activation followed by methanol elimination and aziridine ring-opening yields the electrophilic quinone-methide. Binding to the minor groove of DNA the N2-G (good minor groove nucleophile for π-electrophiles) is alkylated. Subsequent elimination of the carbamoyl moiety generates the α,β-unsaturated iminium functionality that can be attacked by a second N2-G (Figure 5B).[17-19] The N2-G alkylation can occur both inter-and intramolecularly, with the intermolecular cross-link reported as more toxic.[20-22]

Three-membered ring heterocycles (epoxides, aziridines and episulfoniums) are natural products that also target DNA (Figure 5C-E). Aflatoxin B_1 represents probably the most toxic member of this class and is produced by fungal food contamination. This toxin is the major cause of liver cancer, especially in countries that lack appropriate food storage. After uptake in food, the epoxidation of aflatoxin B_1 occurs in the liver. DNA is then alkylated by nucleophilic attack of N7-G.[23-24]

Another example of this group is azinomycin A, which bears a reactive aziridine and epoxide functionality. These three-membered ring heterocycles are reported to alkylate N7-G and N7-A. The presence of two electrophiles in one molecule can result in DNA cross-links, whereby the primary alkylation product is formed between the aziridine unit and the N7-G.[25-26]

Leinamycin alkylates DNA with the corresponding created episulfonium ion preferentially at GG and GT regions of dsDNA. Nucleophillic attack of thiol delivers the reactive thioester and sulfenic acid. Subsequent attack of the sulfenic acid on the ester forms the oxathiolanone and the episulfonium is built by the attack of the nearby olefin.[27-29]

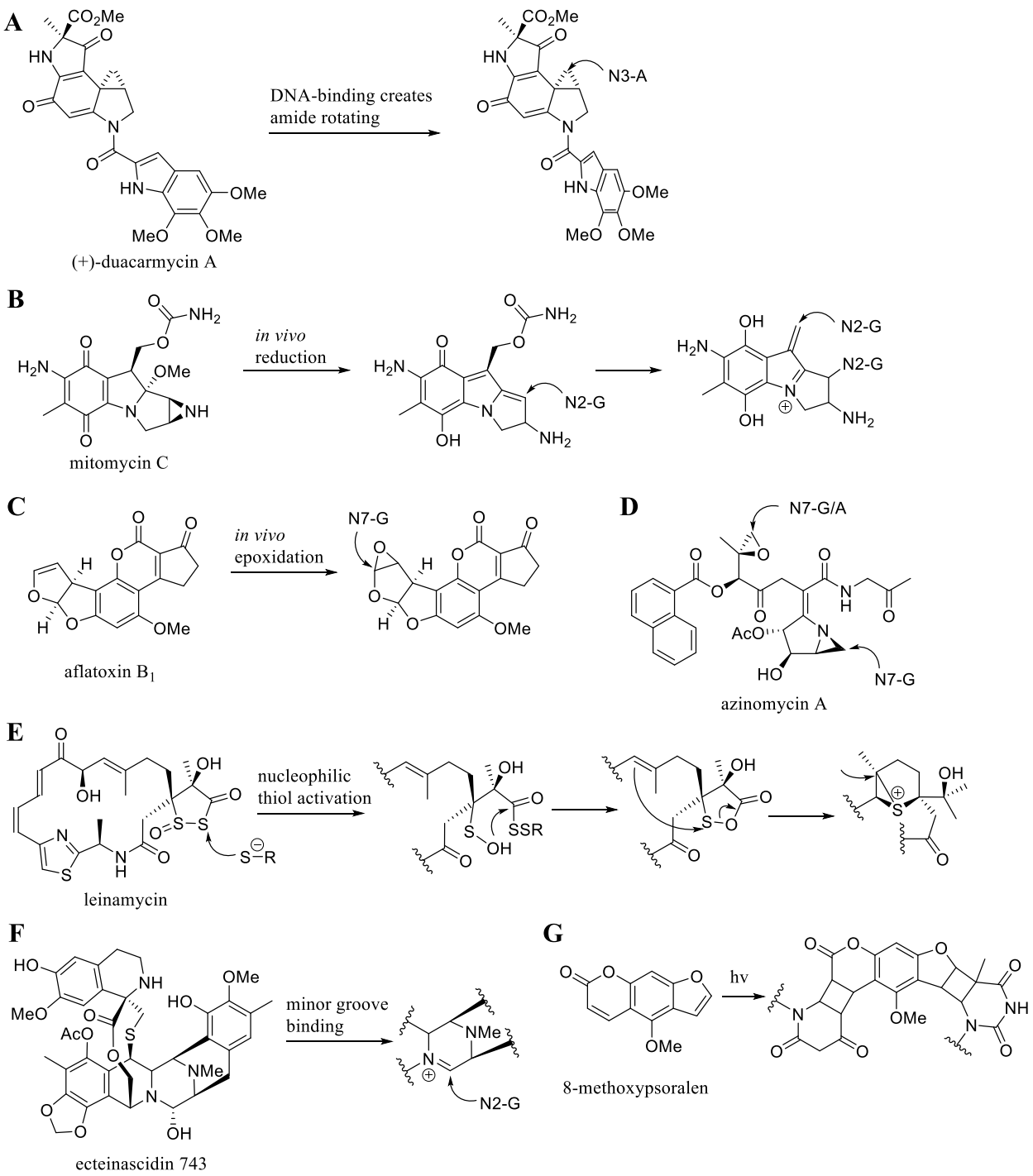


Figure 5. (A) (+)-Duocarmycin A is activated by DNA binding and alkylates it through a cyclopropyl enone electrophile. (B) Mitomycin C is a bifunctional electrophile. Biological reduction delivers the quinone-methide that alkylates N2-G in the minor groove; subsequent elimination leads to the electrophile that is attacked by another N2-G. (C-E) Representative natural product bearing three membered rings (epoxide, aziridine and episulfonium) for DNA

alkylation. (F) Ecteinascidin 743 generates a reactive iminium species that alkylates N2-G upon binding to minor groove of dsDNA. (G) The furanocoumarin moiety can lead to photocross-linking at A-T sites in dsDNA.

Ecteinascidin 743 is an example of a natural product that alkylates the N2-G position of DNA with a reactive iminium ion. The reactive iminium ion is formed upon binding of the substrate with the minor groove of dsDNA (Figure 5F).[30]

Furocoumarins are known to react with DNA via [2+2] cycloaddition. An example of this mechanism shows psoralen that is found in various plants, alkylating dsDNA by photocycloaddition to the thymidine residues (Figure 5G). These cycloadditions can lead to toxic DNA cross-links. The toxicity of furocoumarins (8-methoxypsolaren) has proven for blood cancer treatment and psoriasis.[31-34]

The diazo functionality is of particular interest for the present work. As we will see in later sections diazo compounds were chosen as alkylation source for nucleic acids. There are only a small number of natural compounds bearing stabilized diazo groups.[35-36] The most prominent examples are kinamycin and lomaiviticin. These natural products seem to induce strand-breaking by radical reactions, these however will not be further discussed.[37] More important for the present work, are the diazo natural products consisting of diazopeptides. The structurally simplest example is azaserine, a diazotisated α-amino acid (Figure 6A). These natural diazo compounds are stable enough to target intracellular structures. However, they fell out of clinical studies due to their toxicity.[38-40] The primary mode-of-action of natural α-diazo carbonyl compounds was reported to act as glutamine antimetabolite but their mutagenic activities suggest that DNA is a likely target as well.[41-43]

Diazo actetate formation is reported to happen frequently in humans. Nitrosation of the N-terminus of peptides or glycine itself results in diazo acetates that may cause O^6-(carboxymethyl)-G (O^6-(cm)-G) adducts in the presence of DNA (Figure 6B).[44-45] These DNA damages were especially found in people with a high red meat diet. Summarized, the report shows that DNA alkylation by α-diazo carbonyl compounds is likely, however precise mechanistic studies are missing to conclude whether there is a biological relevance.

The natural product streptozotocin contains a *N*-nitroso-*N*-methylurea functionality (MNNU).[46] This moiety is related to synthetic guanidine derivatives which are precursors to the highly reactive methyldiazonium electrophile (Figure 6C). Unstabilized α-diazo carbonyl compounds have a very short half-life of a few seconds in water.[47] Nevertheless, the created highly reactive alkyldiazonium ion targets many nucleophilic positions in DNA including O^6-G modification. This lesion is reported as highly mutagenic and toxic.[12] The glucose transporter protein GLuT2, highly expressed on the surface of beta

cells, is responsible for recognition of the glucosamine side chain in streptozotocin.[48-49] This specific recognition has made the natural product an effective drug in pancreatic cancer.

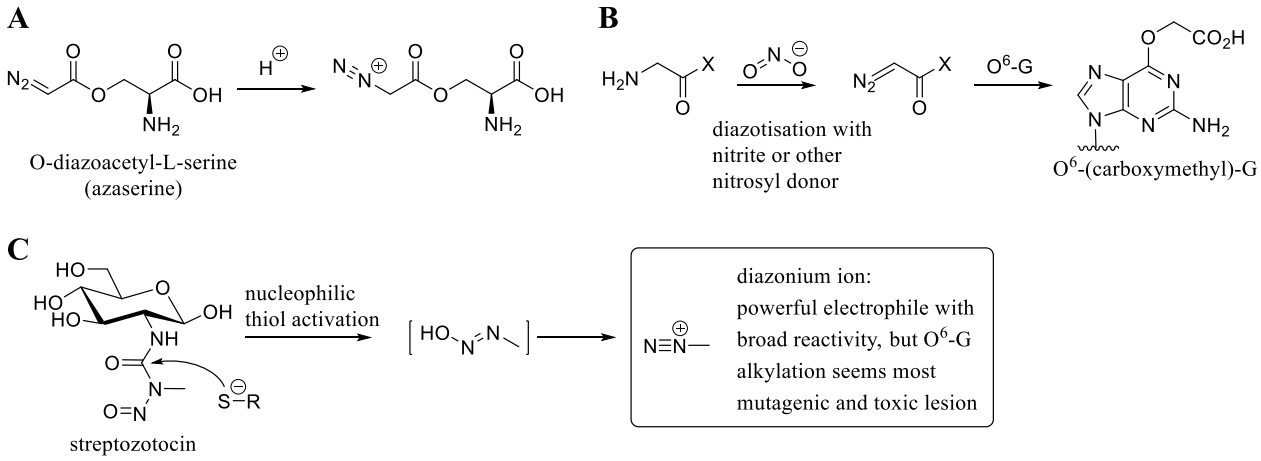


Figure 6. DNA alkylation by diazo and diazonium groups. (A) Structural simplest example for diazo "peptide" azaserine and diazonium formation under acidic condition. (B) Diazo acetates or diazopeptides formation by nitrite or other nitrosyl donor and O^6-G adduct formation. (C) Streptozotocin as example for natural product that creates a highly reactive diazonium ion after thiol activation.

1.2.3 Chemical alkylation of nucleic acids

Inspired by the mechanism of natural products to alkylate DNA, researchers developed a number of molecules to target DNA in similar mechanism.

For example the synthetic developed mustards follow a similar alkylation mechanism as the natural products that modify DNA with a reactive three membered ring heterocycles. The general structure of mustards includes two chloro-ethyl side chains and a central heteroatom (nitrogen or sulfur) (Figure 7A). Intermolecular displacement of one chloride leads to the formation of the three membered ring heterocycles (aziridine or episulfonium ion). Comparable to the natural products leinamycin and azinomycin, the N7-G position in DNA is modified (Figure 7B).[50-53] The second chloride can react in the same mechanism to form inter-or-intrastrand DNA cross-links causing cellular damage that triggers apoptosis.[54] Sulfur mustards were widely used as chemical weapons. Despite the fact that a different mode-of-action (hydrolysis and conjugation involving β-lyase and S-oxidation) was reported, many studies also indicated the related toxicity to DNA alkylation.[55] Nitrogen mustards found application in chemotherapy and chloromethine is the simplest member that still remains important as an anti-cancer drug (Figure 7A).[56-57] The electrophilicity

of the nitrogen mustard can be tuned by varying the substituent on the nitrogen from aliphatic to aromatic.[58] For example, the activation of chloromethine is much easier compared to chlorambucil.

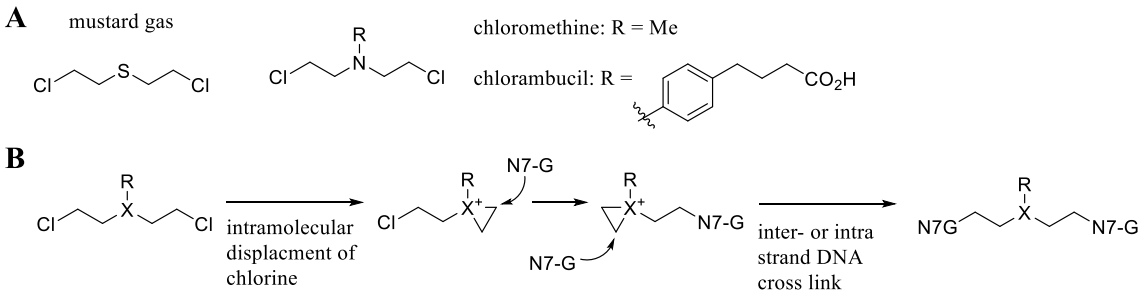


Figure 7. Three-membered ring heterocycles designed by researchers to alkylate DNA. (A) Representative examples for sulfur and nitrogen mustards: mustard gas, chloromethine and chlorambucil. (B) Intramolecular displacement of chlorine lead to DNA alkylation preferentially at N7-G. A second N7-G alkylation causes inter-or intrastrand cross-links.

Temozolomide (TMZ) represents another example of a DNA alkylation agent that reacts with a similar mechanism compared to the natural product bearing a reactive diazo group. TMZ is a synthetic anticancer prodrug, which acts as a methyl transfer agent. In the presence of water a spontaneous break down of TMZ occurs to monomethyl triazene (MTIC) and the highly reactive methyl diazonium electrophile is formed (Figure 8).[59] The TMZ induced methylation of DNA preferentially targets the N7-G position in G rich regions. However, O^6-G and N3-A can also be methylated. The O^6-G alkylation is important for the drug action of TMZ. This modification is reported as highly cytotoxic and mutagenic because polymerases can misincorporate T opposite the O^6-modified G during DNA replication resulting in mutagenesis or apoptosis.[60-61] TMZ is used for brain cancer and glioblastoma multiform (a tumor located in the central nervous system) treatments. Fortunately, the effect of TMZ can be reduced or completely suppressed by DNA repair-mechanism: methylguanine-DNA methyltransferase or low activity of DNA mis-match repair.

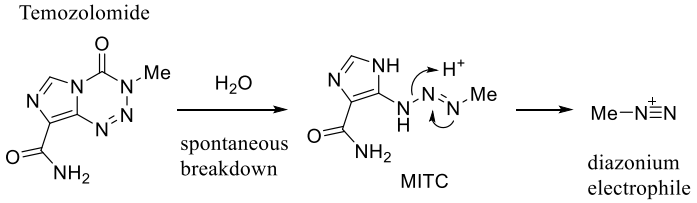


Figure 8. DNA alkylation by diazonium groups derived from synthetic precursors.

In the pharmacopeia there are numerous DNA alkylating agents for example from the nitrosourea family that follow a similar modification mechanism. Decomposition under physiological conditions creates the highly reactive DNA alkylation species.[56, 62]

1.3 The pre-synthetic approach for the synthesis of modified nucleic acids

In the previous section it was demonstrated how nature and researchers inspired by nature evolved strategies to target nucleic acids. All these modification strategies are useful for chemoselective nucleic acid alkylation but not to introduce sequence-selective functionalities. The pre-synthetic modification approaches are the most common ways to address this problem.

The synthesis of modified nucleic acids in a pre-synthetic manner can be subdivided in two main classes:

(1) Automated oligonucleotide solid phase synthesis (OSPS) with modified phosphoramidites.
(2) Enzyme-catalyzed primer extension with synthetic modified triphosphates.

An overview of both pre-synthetic approaches is presented in Figure 9. To obtain the target functionalized nucleic acid, the synthesis of either the modified phosphoramidite (Figure 9, left box) or modified nucleoside triphosphate ((d)N*TP) is necessary (Figure 9, right box). The two representative structures bear the functional group on the nucleobases. However, as later discussed other modification sites are also possible. The modified phosphoramidite is then incorporated in OSPS cycle and the modified (d)N*TP incorporated in DNA replication or transcription. Selecting between biosynthesis and chemical synthesis is determined by the complexity of modification, substrate availability as well as realizable chemical synthesis. For example, the enzyme-catalyzed approach strongly depends on the steric demand of the modification that has to be tolerated and accepted by the polymerase.

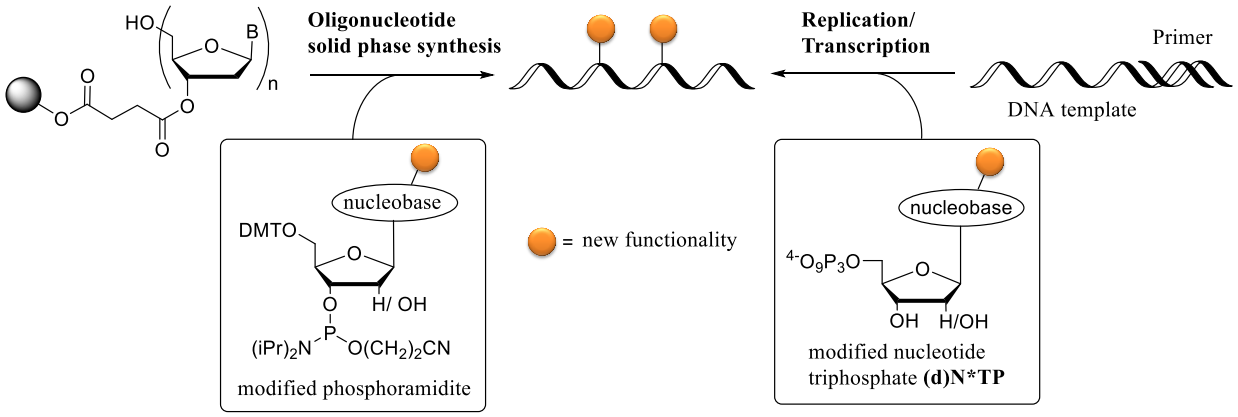


Figure 9. Synthesis of modified NAs: Incorporation of modified phosphoramidites in oligonucleotide solid phase synthesis (left); Incorporation of modified triphosphates in DNA replication or transcription (right).

Generally the modification can be placed at any site of the nucleic acid as long as the chemical synthesis of the corresponding nucleoside phosphoramidite or triphosphate is possible. There are examples in the literature that show modification at the terminal 3`- or 5`- end hydroxyl groups as well as inside the nucleic acid chain. The modifications inside the chain are commonly introduced at the nucleobases (Figure 10B+C). Modification of the phosphate backbone is also described in the literature however this modification site is rarely applied due to perturbation of the resulting helix.[63-64] Ribose modification at the 2`hydroxyl group presents another possibility to introduce the desired functionalities. Using this method, alkyne functionalities as well as dyes could be introduced to nucleic acid chains (Figure 10A).[65-66] Nevertheless, functionalization at the nucleobases represents the most common strategy. The pyrimidine bases are often modified at position 5, since it does not interfere with Watson-Crick base pairing (Figure 10B).[67-68] The purine bases are often modified at the deaza 7-G and A positions (Figure 10C).[69-72] Again, the interference of the modification with Watson-Crick base-pairing is minimized. Especially, for the polymerase mediated nucleic acid synthesis, the incorporation efficiency depends strongly on proper Watson-Crick base pairing. There are also examples of the synthesis of modified purines dNTPs at position 8. They are reported as poor substrates for polymerases and therefore rarely applied.[73]

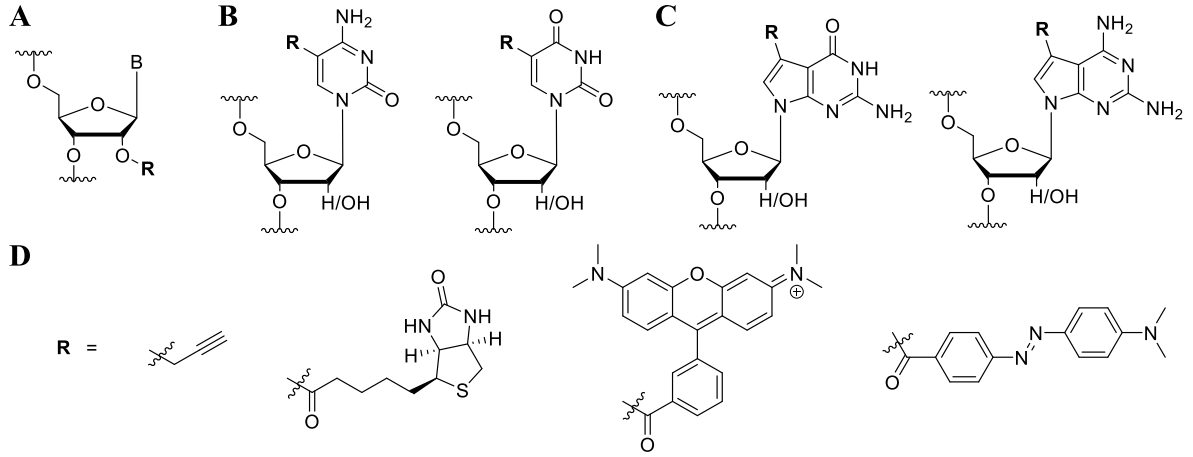


Figure 10. (A) Modification at the 2`-hydroxyl group of the ribose moiety. (B) Modification at pyrimidine bases 5-C and 5-U. (C) Modification at purine bases 7-deaza-G and 7-deaza-G A. (D) Examples for modifications that can be introduced: alkyne, biotin as pull-down tag or fluorophores.

The general process for nucleic acid synthesis with phosphoramidite chemistry is explained first and then how modified phosphoramidites are introduced into this cycle. The oligonucleotide solid-phase synthesis (OSPS) is an automated process that is most widely used for nucleic acid synthesis. In a four step cycle: (1) detritylation, (2) coupling, (3) oxidation and (4) capping the nucleic acid chain is elongated in a 3`-5`direction to synthesize small and medium sized nucleic acids (up to 200 bases) (Figure 11).[74-75] The commercial solid supports for nucleic acid synthesis bear the nucleoside of choice (A, G, T, C, U), which

are detritylated under acidic conditions in the first step of synthesis. The second monomer is then coupled as commercial phosphoramidite with a 5`-OH dimethoxytrityl protection group. Tetrazoles are used as activation agents for the coupling step. Subsequent treatment with iodine results in the desired oxidation state of the phosphate backbone. The yield of the coupling step is never quantitative and therefore a capping step is necessary. Treatment with acetic anhydride acetylates the unreacted 5`-OH groups, avoiding elongation of incomplete reaction products from previous synthesis cycles. The acetylated products (aka. truncated products) will not react further and are excluded from the synthetic cycle. In the next step another monomer can be coupled to the chain or the chain can be cleaved from the resin including nucleobase deprotection under basic conditions. In a final reverse-phase (RP)-column chromatography purification step the truncated products can be separated yielding the target nucleic acid.

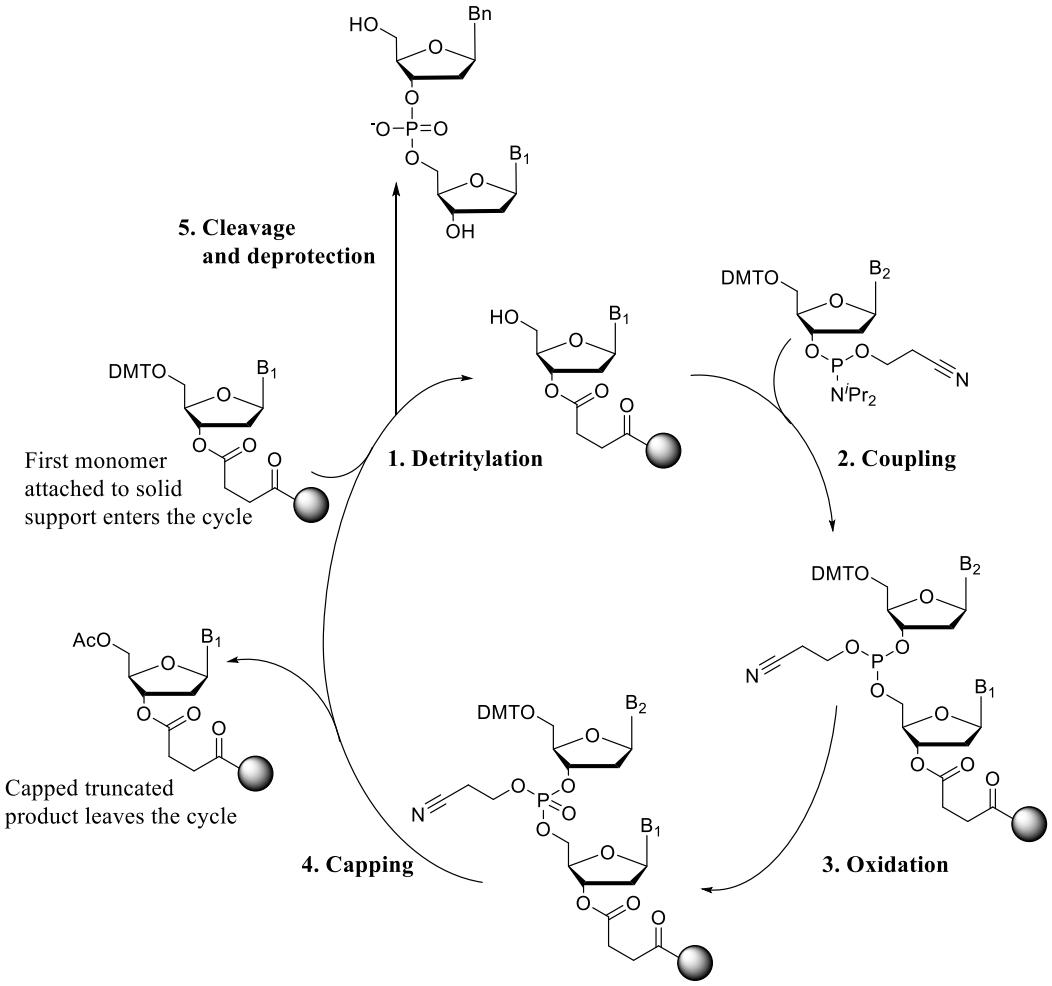


Figure 11. Synthetic cycle for automated oligonucleotide solid phase synthesis using phosphoramidite chemistry with DNA as example. (1) Detritylation: Trichloracetic acid in CH_2Cl_2. (2) Coupling: 5-(Benzylthio)-1H-tetrazole in MeCN. (3) Oxidation: Iodine in THF/pyridine/H_2O. (4) Capping: (a) Ac_2O, THF/acetic anhydride, (b) Methylimidazole, THF/pyridine. (5) Cleavage and deprotection: 32 % NH_3 in H_2O, rt 2h then 55°C for 16 h.

The pure chemical pre-synthetic modification approach uses the OSPS strategy to introduce single modification sites to nucleic acids (Figure 12A). To synthesize the modified nucleic acid of choice the corresponding modified phosphoramidite is incorporated in the OSPS cycle. The advantage of this method is that point modifications can be introduced at the desired positions in the nucleic acid strand. Using this strategy, access modified oligonucleotides of arbitrary complexity is possible, but syntheses of the precursor phosphoramidite are often difficult. Out of numerous examples of modified phosphoramidites that are successfully incorporated in the OSPS cycle, one is presented in Figure 12B. This example was chosen to demonstrate the laborious synthesis of O^6-G modified DNA without the ability for direct modification. Johnson and coworkers synthesized the O^6-G alkylated phosphoramidites in 6 steps and the final O^6-G modified nucleic acids were then tested as substrates for the repair protein O^6-alkylguanine transferase (AGT).[76] This protein transfers the alkyl group from the O^6-G position to a reactive cysteine residue and regenerates the native DNA.[77] The repair of O^6-G alkylation is important because these lesions are reported as highly mutagenic and carcinogenic.[78] The human AGT (hAGT) protein is also expressed in tumor cells and its inhibition can increase the efficiency of DNA alkylating chemotherapeutic drugs. Therefore determination of hAGT activity in cell extracts is important.

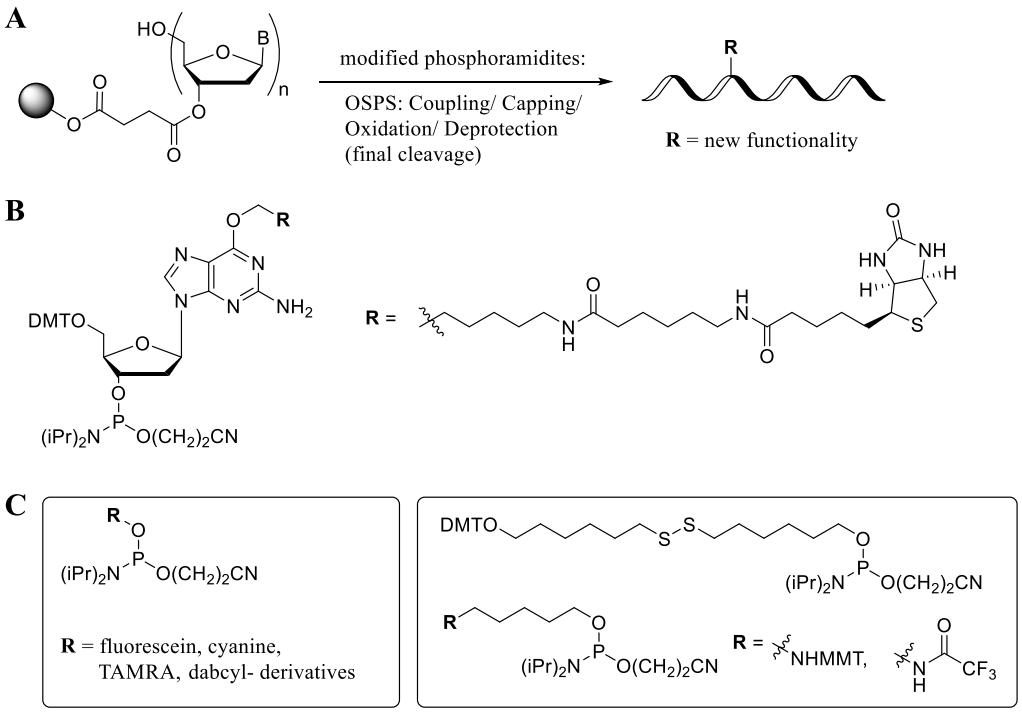


Figure 12. (A) Synthetic procedure for the synthesis of modified nucleic acid with OSPS. (B) Structure examples of synthetized O^6-G modified phosphoramidites that were successfully incorporated in OSPS cycle. (C) Structure examples of commercially available phosphoramidites for direct introduction in OSPS cycle and nucleic acid labelling (left box). Commercially available phosphoramidites that bear nucleophilic groups for further functionalization after oligonucleotide synthesis (right box).

Figure 12C illustrates several examples for commercially available monomers that bear either a fluorophore or a nucleophile moiety.[64] Using the OSPS the fluorophores can be introduced at 3`- and 5`-end and as well as in the chain. Labelling of nucleic acids is important for bioanalytic applications, genomics and diagnostics. The introduction of a nucleophile to the nucleic acid chain is used for further functionalization. Often this method is reported as post-synthetic modification strategy but in reality the modification is introduced pre-synthetically.[64]

The chemical synthesis of nucleic acids with OSPS is facile, robust and scalable. However there are also limitations: difficulties in the synthesis of longer oligonucleotides (›100nt); the chemical synthesis of the desired modified phosphoramidite can be very difficult and a laborious multistep procedure is necessary.[76, 79] Additionally the modification site has to be stable during the whole OSPS process (deprotection, coupling, oxidation, capping and cleavage).

One way to prepare more complex nucleic acid structures and to overcome the length limitations is to ligate several nucleic acid fragments together. In this approach a short modified nucleic acid fragment is synthesized with OSPS and then ligated to a larger oligonucleotide that is 5`phosphorylated (Figure 13). The ligation is catalyzed by T4 DNA polymerase and a complementary nucleic acid template (splint) is used to bridge both fragments.

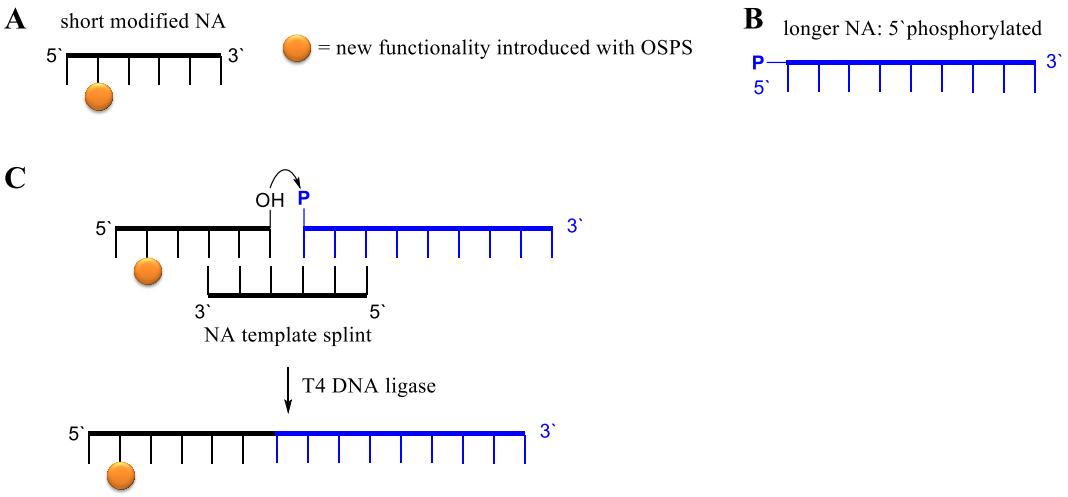


Figure 13. Approach for the synthesis of longer single-point modified nucleic acids: (A) The shorter fragment bearing the modification is synthesized with OSPS and a modified phosphoramidite. (B) Commercial longer NA that is 5`phosphorylated. (C) The shorter fragment bearing the modification is enzymatically ligated to the second 5`phosphorylated nucleic acid fragment using the splint template and T4 DNA polymerase.

An alternative to the classical phosphoramidite chemistry to pre-synthetically modify nucleic acids is enzyme catalysis (therefore called biological pre-synthethic approach see Figure 3). Scientists have taken inspiration from nature and used DNA or RNA polymerases to introduce modified (deoxy) nucleoside triphosphates (dN*TP or N*TP) to nucleic acid chains. The incorporation efficiency of the new synthesized modified triphosphate (dN*TP or N*TP) is always first tested in primer extension experiments (PEX) (Figure 14). In this approach a primer is hybridized with a complementary DNA template. The functionalization is than installed if the modified **dN*TP** is incorporated by the polymerase instead of the natural dNTP (Figure 14B, representative for modified **DNA** synthesis). For successful nucleic acid synthesis it is not only important that the functionalized dN*TP is accepted and incorporated by the polymerase, further extension is mandatory. A lot of investigations were done to develop new polymerases that tolerate a broader substrate scope.[80-81] However, there is no polymerase that can be universally applied. For every new synthetic (d)N*TP investigation of several polymerases are necessary and the reaction conditions (amount of polymerase, ratio of natural to modified (d)N*TP) have to be optimized.

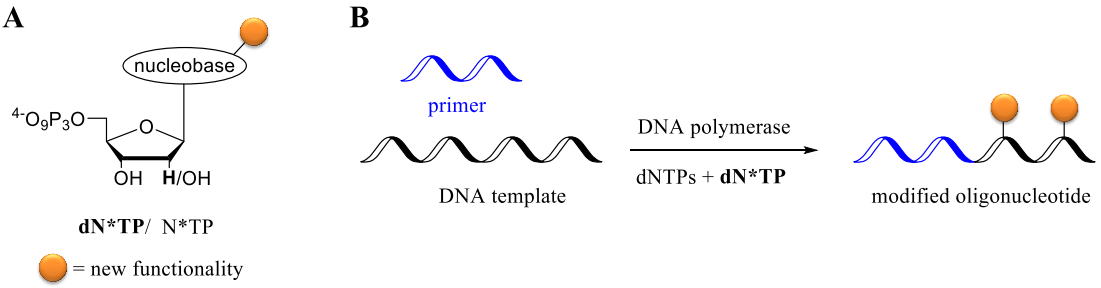


Figure 14. (A) General structure of nucleobase functionalized dN*TP and N*TP. (B) Synthesis of modified DNA via primer extension (PEX) in the presence of dN*TP and natural dNTPs catalyzed by DNA or RNA polymerase.

The first investigations in this field were done in 1981. Langer and coworkers incorporated a biotinylated dUTP in DNA synthesis.[82] Later, Famulok and colleagues were the first ones who presented a full set of all four modified dN*TPs and studied them towards PEX.[83] Different functionalities including guanidine,[84] amino acids,[85] ferrocenes[86] and saccharide[87] were successfully incorporated. Carell and coworkers demonstrated that alkyne functionalization is also possible with this approach. The alkyne moiety was then further modified by copper-catalyzed alkyne cycloaddition.[67-68, 72, 88-91]

Base-modified nucleotide triphosphates (N*TPs) were also investigated as substrates for RNA polymerases but they gained less attention due to the difficult handling of modified RNA compared to DNA. However, a groundbreaking example was demonstrated by Hirao, who expanded the genetic alphabet with the development of new base pairs such as 2-amino-6-(2-thienyl)purine (s) and 2-oxo-(1H) pyridine (y) (Figure 15A).[92-93] The bulky 6-thienyl group disfavors s-T pairing (Figure 15B). The transcriptional specificity for

s-y pairing is comparable to the natural base pairs and site selective incorporation of functionalized yTP opposite s by T7 transcription yielded the single modified RNA. In the follow-up studies transcribed RNA from dsDNA could be site selective labeled with fluorophores such as TAMRA and biotin (Figure 15C+D).[94-96] In the first step the target dsDNA is amplified by polymerase chain reaction (PCR) with a 5`primer bearing the T7-promoter sequence and a 3`primer bearing the artificial base s. Transcription of the PCR product with the modified yTPs yields the target functionalized RNA. This method represents so far the best solution to introduce single functionalization to RNA molecules.

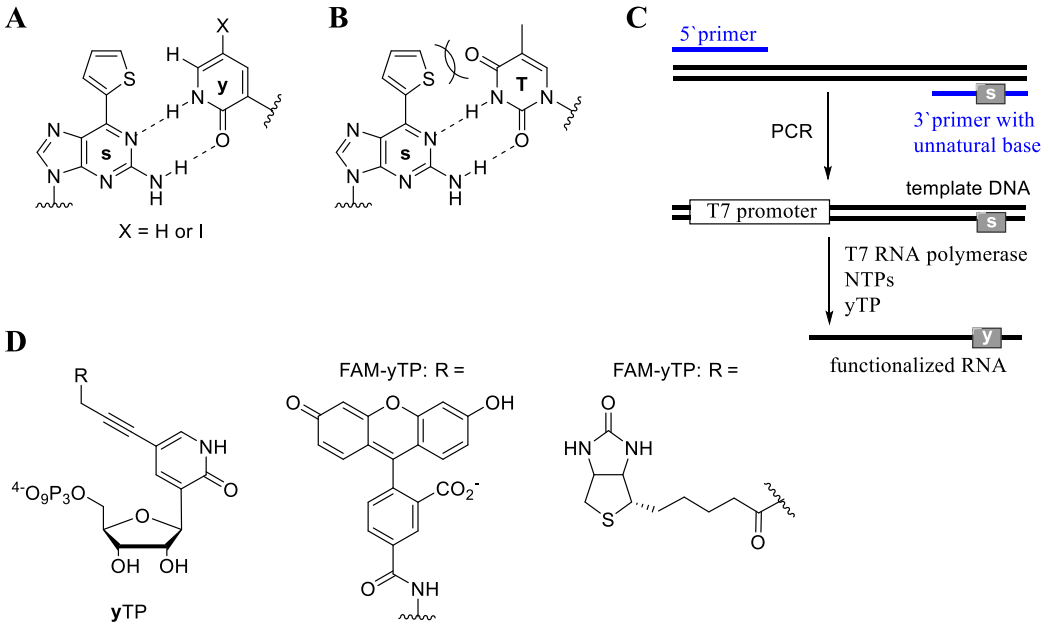


Figure 15. (A) Non-canonical base pair s-y. (B) Bulky 6-thienyl group disfavors s-T pairing (C) Synthesis of single modified RNA by s-y pairing and primer extension catalyzed by T7 transcription.

The major drawbacks of nucleic acid modification by PEX are the limited scale of the reaction and obstacles in the synthesis of the functionalized (d)N*TPs. There are many strategies for their preparation. However, depending on the modification site (especially for O^6-G), there is only the possibility to introduce the functionality to the nucleoside first and then a final chemical triphosphorylation step is necessary. Already in the `80s, Yoshokawa showed that the reaction of nucleosides with phosphorylchloride ($POCl_3$) followed by pyrophosphate and triethylammoniumcarbonate yielded the desired 5`triphosphate.[97] There are alternatives in the literature for triphosphorylation, but the method developed by Yoshokawa is most commonly used.[98] In some cases protection groups are required, for example for nucleophilic functionalities, due to the use of the strong electrophilic phosphorous reagents in the triphosphorylation step. Then the final deprotection step has to be mild and compatible with the TP. Considering all these facts, laborious multistep procedures are often necessary for the preparation of functionalized (d)N*TPs.

A better synthetic alternative was developed in 2003: the direct functionalization of (d)NTPs via cross coupling reaction. Burgess and coworkers presented the Sonogashira alkynylation of dUTP.[99] Afterwards efficient protocols for Suzuki-Miyaura reactions of halogenated (d)NTPs were published;[100-102] Challenges were faced with hydrolysis of the triphosphate, but these could be mitigated by using mild cross-coupling conditions as well as fast isolation of the product. It should be mentioned that these cross-coupling approaches are only applicable for modification of position 5 in pyrimidines and 7 in purines.

The PEX method allows the synthesis of medium-size DNA (15-100 bp). Polymerase chain reactions (PCR) are necessary to obtain longer modified strands. In this method a forward and reverse primer are used for DNA amplification. The DNA synthesis is run in several cycles (20-40) that always repeat: denaturation, annealing and extension using a thermostable DNA polymerase. PCR amplifications are generally more challenging for the enzymes than PEX because the polymerase has also to read through the modified template in the 2nd cycle. Therefore smaller modifications or functionalities bearing a flexible linker are substrates for PCR amplifications.[91, 103]

Unfortunately, neither the PEX nor the PCR modification methods are suitable for internal site-specific point functionalization as the OSPS phosphoramidite chemistry approach. One way to introduce site-selective modification to longer nucleic acids is the combination of both: the synthesis of the modified primer with the OSPS followed by PEX with (d)NTPs and replication or transcription (Figure 16).

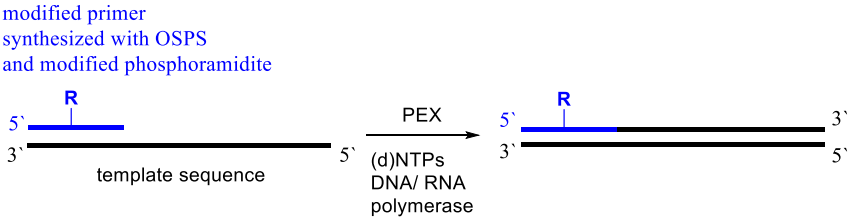


Figure 16. Synthesis of longer point modified nucleic acids: combination of chemical OSPS approach with biological PEX catalyzed by DNA/RNA polymerases with (d)NTPs.

1.4 Chemical sequence-selective alkylation of nucleic acids

The combination of pre-and post-synthetic modifications strategies allowed researchers to develop sequence-selective alkylation of nucleic acids. A short nucleic acid strand (or nucleic acid mimic) bearing a reactive moiety is designed to be complementary to the target nucleic acid and guides the reactive group to the desired location (Figure 17A). The nucleic acid strand bearing the reactive group is synthesized using OSPS with the modified phosphoramidite. Sasaki and coworkers presented a templated cross-link

formation via a Michael type addition of cytidine to the reactive allyl unit (Figure 17B).[104] Another example was reported by Rokita and coworkers. They used quinone methides to target nucleic acids and form cross-links (Figure 17B).[105-106] Both alkylation methods are site-selective and the functional group reacts after annealing of the strands. However, after the successful reaction the guiding sequence is still covalently bound to the target sequence and will change the native function of it. Therefore, the presented methods are likely useful for nucleic acid labeling and detection. Sasaki and coworkers found a solution for this problem. They developed a new strategy using a functionality transfer technique where the guiding sequence is released after the addition reaction (Figure 17C).[107] In this technique a carbon-sulfur bond connects the electrophile (α,β-unsaturated diketone) to the guiding sequence and in the transfer process this bond is cleaved, releasing the guiding sequence. This technique is unique and allows DNA and RNA alkylation without affecting their function by the covalent bound guiding sequence.

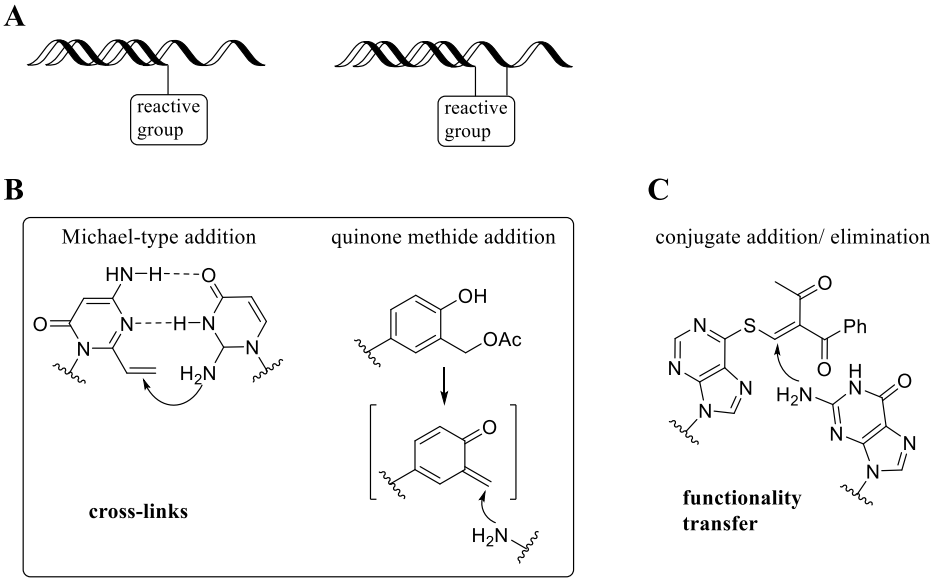


Figure 17. (A+B) Sequence selective modification of nucleic acid by a short nucleic acid complement bearing a reactive group. The methods highlighted in the box lead to cross-links and are used as detection systems, however the functionality transfer of malonyldienone-type electrophile releases the guiding sequence (C).

1.5 Catalytic approaches for nucleic acid modification

As already mentioned in the previous section, direct modification of nucleic acids would be more efficient and may provide access to more complex structures. However, the development of such strategies also has challenges due to multiple similar reactive groups in nucleic acids. Nature has solutions to introduce sequence-selective modifications to nucleic acids. For example, methyltransferases (MTases) catalyze the

methyl transfer from the cofactor *S*-adenosylmethionine (AdoMet) to various positions of DNA and RNA (C, A and G). The cofactor AdoMet is the main natural source of electrophilic methyl groups (Figure 18A). Various AdoMet analogues were synthesized and investigated in the reaction of alkyl transfer. The first alkyl transfer analogues targeted DNA with an aziridine moiety.[108] This strategy was rather useful to target DNA than functionalized RNA. However, optimization of the AdoMet design allowed sequence-specific targeting of dsDNA, transfer RNA (tRNA), dsRNA and ssRNA with alkyl, alkenyl, alkynyl and azide functionalities (Figure 18A, B). Furthermore, Rentmeister and coworkers developed a strategy to use human trimethylguanosine synthases (GlaTgs) to transfer alkyl groups to 5`-cap structures of messenger RNAs (mRNAs) (Figure 18C).[109-110] Recent developments in this field published by the same group showed the successful transfer of a vinylbenzyl motif. The vinyl group can be further functionalized e.g. using tetrazine inverse electron demand Diels-Alder reaction.[111]

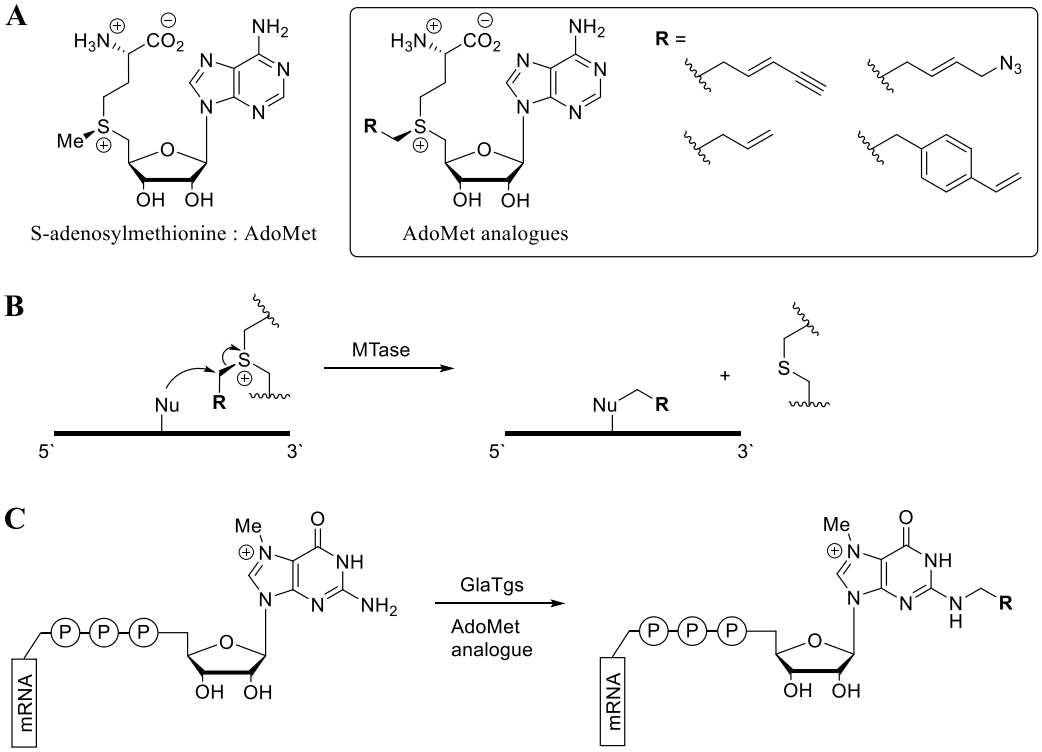


Figure 18. Natural enzymes for transcript labelling. (A) S-adenosylmethione (AdoMet) and its analogues (highlighted in the box). (B) Reaction of nucleophilic group (Nu) of nucleic acid with AdoMet analogues catalyzed by methyltransferases (MTase). (C) mRNA 5`-cap targeting with AdoMet analogues catalyzed by trimethylguanosine synthases (GlaTgs).

1.6 Chemical nucleobase modification for sequencing of natural nucleic acid modification

In this section one example is presented to show how chemical modification of nucleic acids has been used to sequence non-canonical bases in the whole genomes at single base resolution.

The chemical reactivity of nucleic acids was explored to understand which chemicals can be dangerous for genomes. Already in the `70s the different reactivity of uracil and cytosine with bisulfite was reported. The cytosine addition product is unstable and hydrolyses to uracil. In contrast the uracil addition product eliminates the bisulfite and restores uracil (Figure 19A). These different reactivities were used to develop a sequencing strategy for modified cytosine (Figure 19B). 5-methylcytosine (m^5C) is widespread in DNA[112] and RNA.[113] This modification is known to regulate gene expression in eukaryotes. The existence of this modification has been known since 1925, but only with the development of modern DNA sequencing methods were many biological roles revealed. It was shown that m^5C could react with bisulfite but the reaction is slow enough to discriminate between C and m^5C (Figure 19).[114-115] The target isolated DNA is sequenced by normal methods. This method cannot discriminate between C and m^5C. Then the DNA is treated with bisulfite and all Cs are converted to Us that are read as Ts after PCR amplification and sequencing. The first data set is then compared with the data set from bisulfite treatment and all m^5C sites can be determined. The Cs that are not converted were m^5C in the original DNA strand.

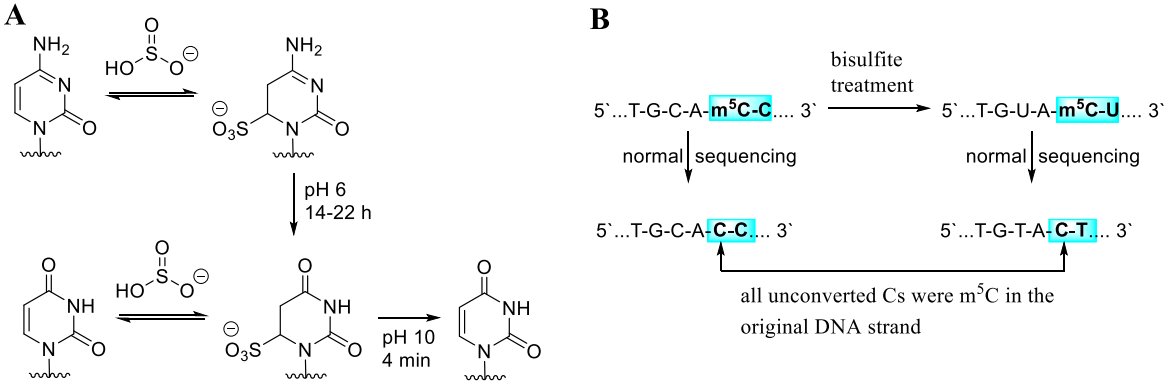


Figure 19. "Selective" reaction of DNA with bisulfite enabled the development of a sequencing method to discriminate between Cs and m^5Cs. (A) Bisulfite adds to the position 6 in pyrimidines resulting in deamination of cytosine and restorage of uracil. (B) Concept of bisulfite sequencing strategy to map m^5Cs.

The optimization of the reaction condition allowed the detection of m^5Cs in whole epigenomes at single-base resolution.[116-118] In addition the construction of oxidative and reductive bisulfite sequencing allowed the detection of all m^5C oxidation derivatives (hydroxymethyl, formyl and carboxyl). For a more detailed description about oxidative and reductive bisulfite sequencing as well as other sequencing methods derived from chemical nucleobase modification the reader is referred to our recently published review.[8] The

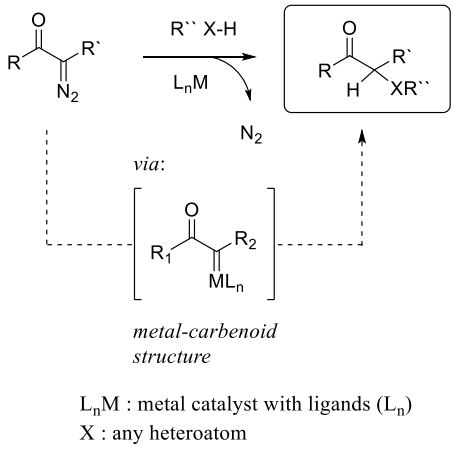
bisulfite sequencing methods was shown to demonstrate how important chemical modifications are for structure elucidation of nucleic acids. Limitations of these methods are the multistep procedure and the loss of the analyte due to harsh conditions. Therefore deep sequencing is needed in order to achieve enough reads for reliable results.

Inspired by nature, researchers have developed fantastic solutions to introduce modifications to nucleic acids in post-synthetic as well as in pre-synthetic approaches. There are plenty of modification methods presented in the literature, however, a complete understanding of the properties and reactivities of nucleic acids is missing. Therefore, new selective chemistry is needed to find answers to these questions and potentially also find new modification sites. With the present work we want to contribute to this field with the study of copper(I) as catalyst for nucleic acid modification.

Chapter 2: Electrophilic Cu(I) carbenes react chemoselectively with guanine through a substrate directed reaction

2.1 Catalytic X-H insertion reactions of diazo carbonyl compounds: early developments and mechanistic insights.

The bond formation between a heteroatom and a carbon is an important process for both natural and man-made molecules. Construction of this type of bond is crucial for chemists and has numerous applications in chemical biology. The extensive use of Buchwald-Hartwig C-N and C-O couplings catalyzed by palladium is an important example of these methodologies.[119,120] Another strategy to form carbon-heteroatom bonds is carbenoid based X-H insertion (XHI). In this transformation the nucleophile X can be a nitrogen, oxygen, sulfur, selenium, phosphorous or halogen atom. The present work is focused on metal carbenoid chemistry. This strategy to link a heteroatom to a carbon is the method of choice for selective nucleic acid modification. Formation of the crucial metal-carbenoid complex typically occurs *in situ* starting from a diazo precursor. Subsequent nucleophilic attack of the XH group in a concerted or stepwise reaction, results in the desired X-C bond (Scheme 1). The underutilization of this type of reaction in synthetic chemistry is rather surprising considering the broad reaction scope and functional group tolerance of XHI catalyzed by rhodium, copper, ruthenium or iron. Additionally, recent developments such as the enantioselective variant expand the scope of potential application in chemical synthesis.



Scheme 1. (A) General scheme for X-H insertion reaction with metals.

The first example of XHI was documented by Reichstein and Casanova in 1950.[121] They showed that treatment of an α-diazo ketone with copper(I) oxide in the presence of methanol yielded the corresponding α-methoxyketone (Figure 20A) instead of the expected Wolf rearrangement product. Two years later, Yates presented a more detailed investigation into XHI reactions with α-diazo ketones catalyzed by copper. He

illustrated the possibility to use various alcohols and amines (thiophenol, aniline, piperidine, ethanol) to yield the corresponding XHI products (Figure 20B).[122] For the first time a carbene-type mechanism was proposed. Further studies demonstrated that rhodium(II) carbenoids derived from dirhodium tetraacetate (Rh_2OAc_4) were excellent catalyst for insertion into O-H bonds (Figure 20C). Already in the `80s it was shown that XHI was applicable to industrial processes. The key step in the synthesis of the antibiotic (+)-thienamycin is a rhodium(II) catalyzed N-H insertion reaction (Figure 20D).[123] However, most investigations in the carbenoid field were carried out on C-H insertion reactions and cyclopropanation. Reasons why catalyzed XHI reactions are underexploited could be that there are already reliable classical methods available to form a carbon-heteroatom bond e.g. the substitution reactions without the need of a catalyst. However, catalytic pathways have unique advantages over stoichiometric ones, like controlling reaction selectivity (chemo-, enantio-, diasterio-, and regioselectivety) by choosing the appropriate ligand system. This is particularly advantageous in the synthesis of large molecules, like natural products, where selective reactions are very important. Besides the selective synthesis of complex molecules, applications of XHI in chemical biology can be crucial to study biological systems. In this chapter, the examples for XHIs are focused on X = N, O and S due to their relevance in biological systems.

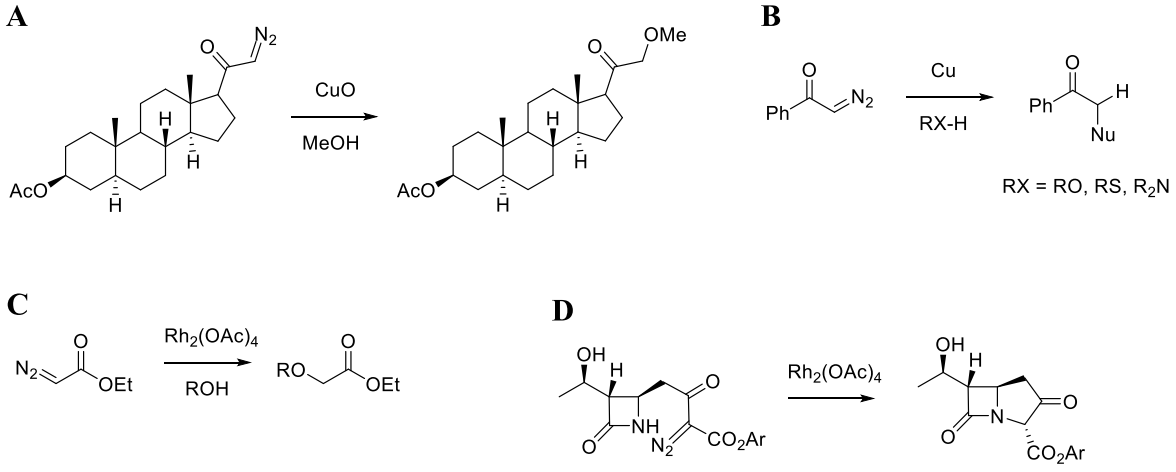


Figure 20. Initial investigations on XHI catalyzed by copper or rhodium (A-D).

2.2 Metal carbenoids: different diazo compound as carbenoid precursors and electronic features of the metal carbenoid bonds.

In this section the reactivity and stability of diazo compounds is briefly discussed to present the structural influence on carbene formation that can be relevant for XHI. As already mentioned in the previous sections, diazo compounds are reactive building blocks and used as carbene precursors for XHI reactions catalyzed by transition metals. Their reactivity and thermal stability can be tuned by the electronic character of the substituents in the α-position of the diazo group. Generally, the substituents can be used to classify the diazo substrates in four different groups (Figure 21A). The acceptor/acceptor species bearing two carbonyl functionalities show the highest stability, whereas, species with only electron donating groups are very unstable. Several resonance structures contribute to the stabilization of the diazo compound via delocalization of the negative charge. The anion of stabilized diazo compounds shows poor interaction with metals (Figure 21A). Thus, acceptor/acceptor and acceptor/donor stabilized diazo substrates need more active catalysts to form the corresponding metal-carbene.

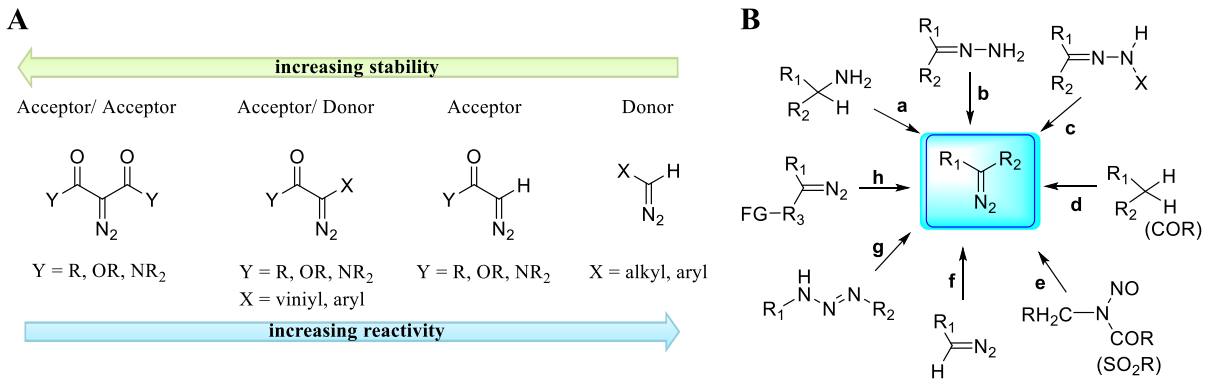


Figure 21. (A) Structure and stability of α-diazo compounds. (B) General synthetic routes for α-diazo compounds. (a) diazotization of α-acceptor-substituted primary amines with nitrous acid; (b) oxidation (dehydrogenation) of hydrazones with e.g silver oxide; (c) cleavage of sulfonylhydrazones under basic conditions; (d) diazo-group transfer onto activated methylene or methine substrates; (e) cleavage of N-alkyl-N-nitroso sulfonamides, carboxamides, ureas or urethanes in the presence of a base; (f) electrophilic substitution at diazomethyl substrates; (g) triazene fragmentation; (h) post synthetic modification of an existing diazo compound.

The simplest and most reactive example of a diazo compound is diazomethane. However, this compound is also the most hazardous diazo substrate and should be handled with care. Generally, unstabilized diazo compounds are considered as potentially explosive and their thermal stability should be taken into consideration during synthesis. The major synthetic routes for diazo compounds are summarized in Figure 21B. For a detailed description of the synthesis of diazo compounds, the reader is referred to the excellent review by Maas.[124]

2.3 Copper(I) catalyzed X-H insertion reactions with α-diazocarbonyl compounds.

The development of copper carbene chemistry is of particular interest for the present study and is therefore described in detail.

The first important application of X-H insertion reactions with copper(I) carbenoids derived from α-diazocarbonyl compounds was reported in `60s by Rajagopalan. He presented the alkylation of the active site carboxyl residue of pepsin with α-diazocarbonyl compound **1** and a large excess of copper(II) salt (Figure 22A).[125] Unfortunately, the copper(I) carbenoid chemistry was overshadowed by other transition metal carbenes until the discovery of Jørgonsen and coworkers, who used chiral copper(I) complexes catalytically for stereoselective insertion reactions into aniline with diazo compound **2**.[126] Investigations of diverse chiral ligands (diphosphines, diamines, diimines, bisoxazolines or P-N ligands) resulted in an ee up to 28 % (Figure 22B). Two representative chiral ligands **3** and **4** with the highest ee are shown in Figure 22B. Furthermore, the Jørgenson group showed that a change in either the diazo substrate or amine source did not improve the enantioselectivity. However, variation of the copper source and chiral ligand was later demonstrated by other groups to achieve an enantioselectivity up to 98% ee for NHI and OHI (Figure 22C and D).[127-128] In all stereoselective XHI reactions, the chiral catalyst was formed *in situ* by mixing a copper salt with the chiral ligand.

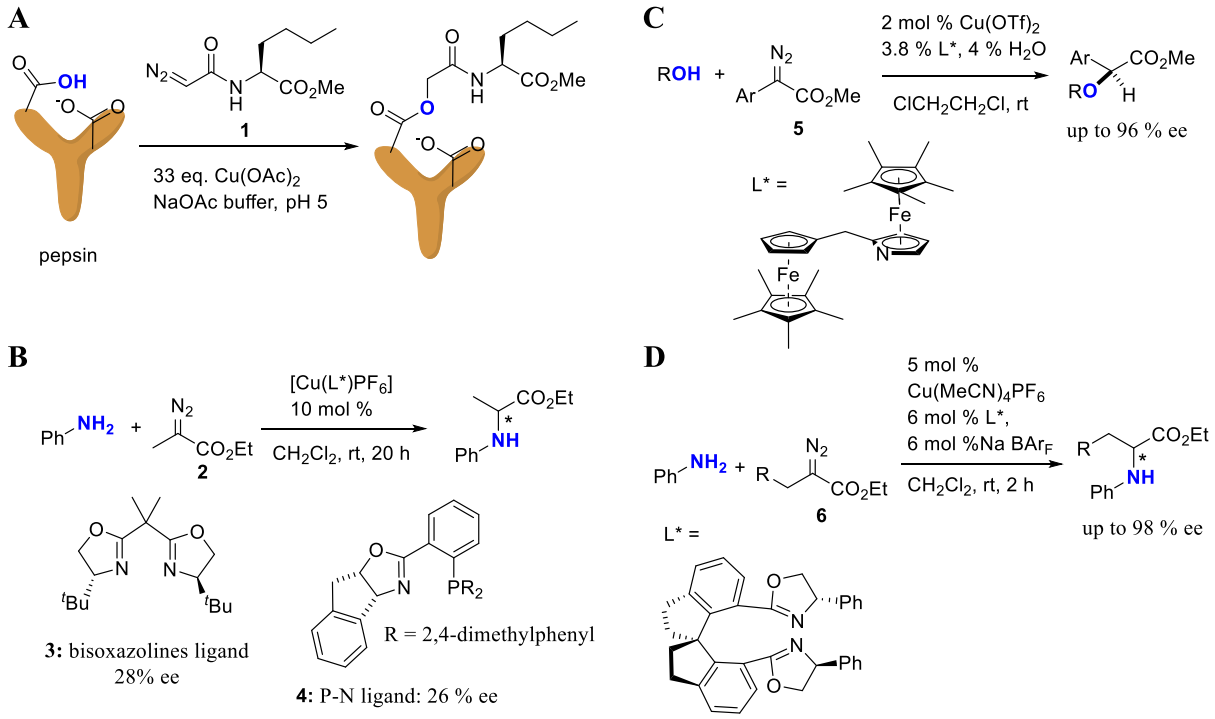


Figure 22. Copper(I) catalyzed carbenoid XHI. (A) Pepsin modification via OHI catalyzed by excess of copper. (B) First example of stereoselective NHI catalyzed by copper(I). (C) Stereoselective OHI catalyzed by copper(I). (D) Recent example of stereoslective NHI catalyzed by copper.

In the mechanism of XHI reactions different transition metals can lead to slightly different reaction pathways. However, there are two general mechanisms: a concerted insertion into non-polar bonds (as in the case of C-H bonds), and a stepwise ylide formation (typically occurs with polar bonds). The mechanism for the copper catalyzed carbenoid insertion reaction is shown in Figure 23. Kochi[129] and Tessi[130] demonstrated that the catalytically active species is copper(I). If necessary, a reductant such as ascorbate can be added to create the desired oxidation state of copper. However, it was also shown that α-diazocarbonyl compounds can act as reducing agent by undergoing a sacrificial coupling.

A number of scenarios, for example some coordination modes of the substrate or Lewis bases can potentially slow or inhibit the XHI reaction (top line in Figure 23). A crystal structure of an unproductive substrate binding shows one way that copper catalysis could be blocked (Figure 23, dashed box).[131] However, if the XHI reaction is successful, it will follow the general mechanism for XHIs *via* steps A-D. The first step is the coordination of the negatively polarized diazo substrate to the metal, (Figure 23, step A), which is then followed by loss of nitrogen to form the metal carbenoid (Figure 23, step B). A representative example of a metal carbenoid characterized by X-Ray and NMR is shown in Figure 23 (blue box).[131] In rhodium(II) catalysis, step B, the loss of nitrogen, is the rate-determining step (RDS). However, the computational studies for XHI catalyzed by copper showed similar transition state energies for all steps A-D.[132-133] Therefore, it is likely that the rate is determined by the copper source. Coordination of the nucleophilic XH group to the metal gives the corresponding ylide (Figure 23, step C) and electrophilic insertion of the carbene bond into XH followed by a proton shift (Figure 23, step D) yields the final XHI product. Computational studies by Liang and coworkers suggested that the insertion step D includes a migration of the copper to the carbonyl oxygen (Figure 23, green box).[133]

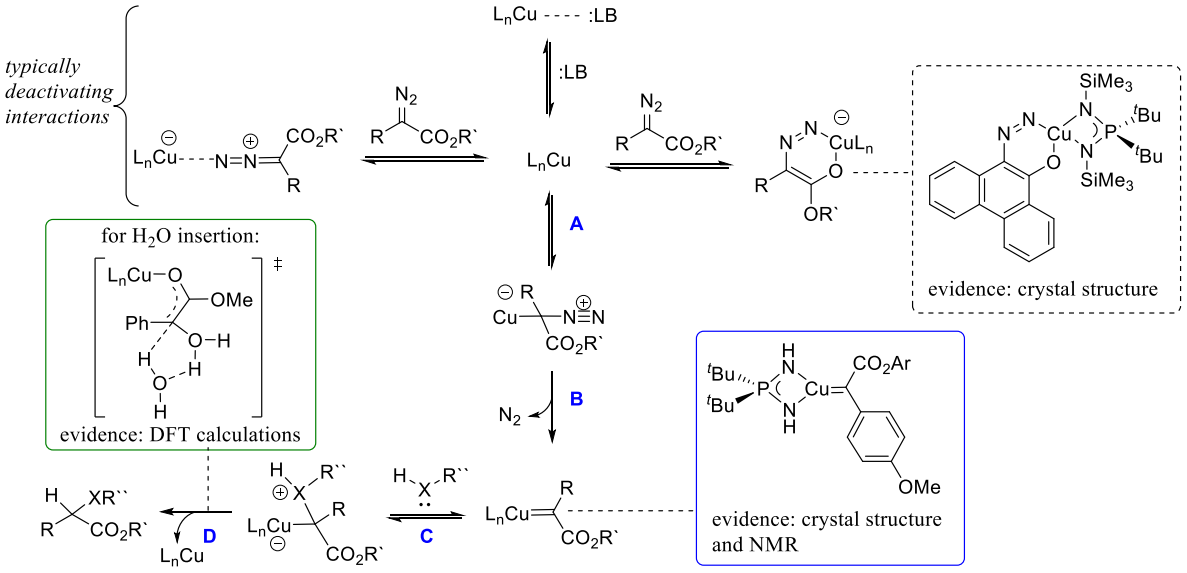


Figure 23. Mechanistic insights of copper(I) catalyzed carbenoid chemistry.

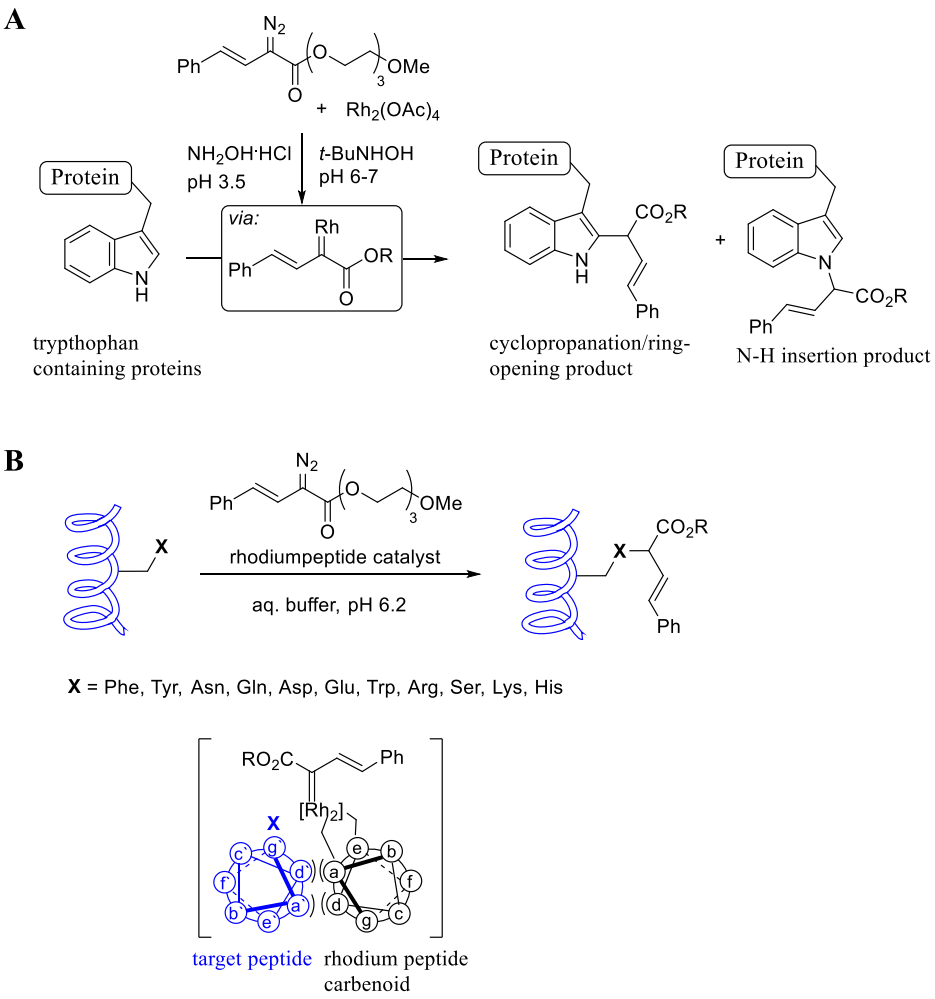
2.4 Metal-carbenoid XHI reactions in chemical biology

In chemical biology the selective chemical manipulation of biomolecules is used to label, modify, probe or pull-down a molecule of interest. Methods for nucleic acid modifications have many applications, one of which is use as tools to understand their biological role or for adaption in diagnostics and therapeutics. The major challenge in modification of biomolecules for biorthogonal chemistry is to achieve high efficiencies and the right selectivity without destroying the corresponding activity of the biomolecule. Considering that most of these reactions should take place under mild and physiological conditions: neutral pH, ambient temperature and in an aqueous environment; the possibility for competing side reactions is very high. Furthermore, these complex biomolecules show diverse functionalities and contain multiple reactive sites. Therefore the development of a selective biomolecular modification is difficult.

Nevertheless, there are examples in the literature of bioconjugation using metal carbenoid chemistry for nucleic acids and proteins. Both substrate classes have nucleophilic sites for possible XHIs. In proteins the amino acid side chains or the N-terminus can be targeted and the nucleic acids provide nucleophiles on the nucleobases as well as in the ribophosphate.

The challenge in the bioconjugation of proteins is the selective introduction of the new functionality without affecting its structure and activity. Thus, the metal-carbenoid method should remain efficient at reasonable protein concentrations, and ideally should target a specific single site or a single functional group while leaving the remainder of the protein untouched.

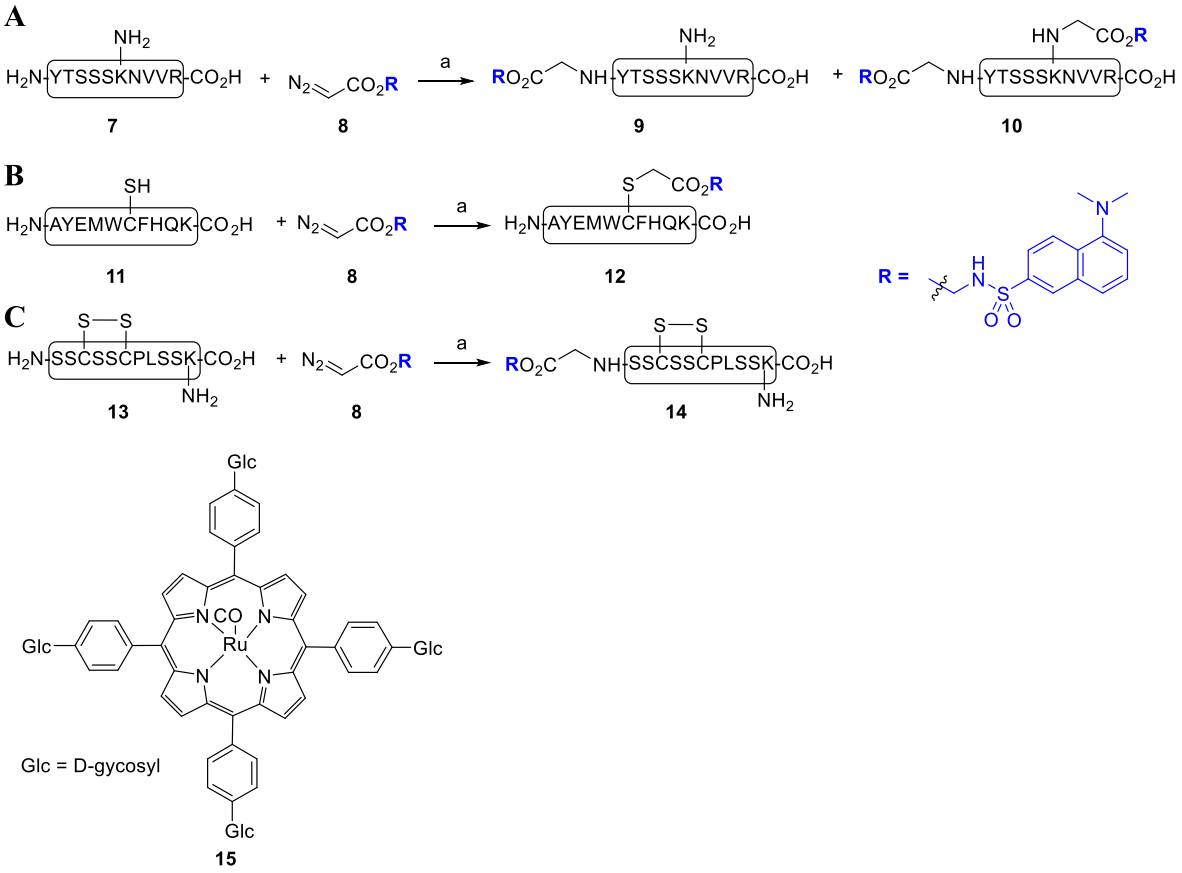
Following the discovery by Rajagopalan and coworkers in 1966 that protein can be studied by diazo modification,[125] Antos and Francis presented a rhodium carbenoid system for protein labelling in 2004. The proteins (myoglobin and subtilisin) were successfully modified at their tryptophan units using Rh_2OAc_4 and α-diazoesters at low pH.[134-135] The reaction conditions yielded not only the NHI product of tryptophan, but also some cyclopropanation of the tryptophan indole moiety (Scheme 2A). The presence of the buffer salt hydroxylamine hydrochloride was important for this transformation to proceed, however, its detailed role in the catalytic mechanism was never discovered. In the follow-up studies improved reaction conditions (*t*-butylhydroxyl amine) for smooth alkylation of mellitin, lysozyme and FKBP-mellitin were presented. Yet, protein alkylation using rhodium carbenoid chemistry is limited to proteins bearing solvent accessible tryptophans (else the protein denaturation is required before alkylation). Later, Ball and coworkers reported a catalytic system where the rhodium carbenoid is generated from a rhodium bearing peptide α-helix delivering the alkylation group in direct proximity to the target functional group via coiled-coil interactions of both helices (Scheme 2B).[136-137] This strategy allowed modification of nearly all natural amino acid side-chains bearing a nucleophile.



Scheme 2. Rhodium carbenoid-mediated modification of proteins. (A) Rh(II) carbenoids for covalent modification of tryptophan residues in proteins. (B) Site-specific modification of amino acid chains by specific binding of Rh(II) catalyst, that brings the resulting Rh(II) carbenoid in close proximity to the modification site.

In 2010, Che and coworkers presented a protein labelling method for site-selective alkylation of the N-terminus using dansyl-functionalized diazoacetate **8**, catalyzed by a water soluble ruthenium porphyrin complex **15** (Scheme 3).[138] Water solubility was introduced by attaching α-D-glucosyl groups to the porphyrin ring. The first model peptide **7** bearing multiple reaction sites (primary and secondary alcohols, phenol, guanyl and amino groups) was efficiently alkylated at the N-terminus. Only traces of lysine side chain alkylation **10** were detected (Scheme 3A). The high selectivity for terminal NHI was interesting considering that the pK_a value of the amino group in the lysine side chain is more basic (pK_a = 9.3-9.5) compared to the N-terminus (pK_a = 7.6-8.0). However, the reaction was run at pH 7.4-8.4 and therefore most of the lysine side chains were protonated. The absence of O-alkylation can be explained by the higher nucleophilicity of the nitrogen. More investigations with different peptides, for example peptide **11,** showed

that the selectivity could be switched to SHI if only one cysteine was present (Scheme 3B). SHI was shown to be faster than NHI using metal carbene chemistry. However, the selectivity could be tuned back to terminal NHI by introducing a disulfide link between two cysteine side chains (Scheme 3C). Further investigation extended the method to allow modification of the terminal amine in different proteins. Those examples include the NHI of terminal lysine in RNase A; glycine in insulin A; and phenylalanine in Insulin B. The limitations of this alkylation method were shown by the failure of modifying myoglobin and ubiquitin where only traces of the corresponding alkylation product were detected.

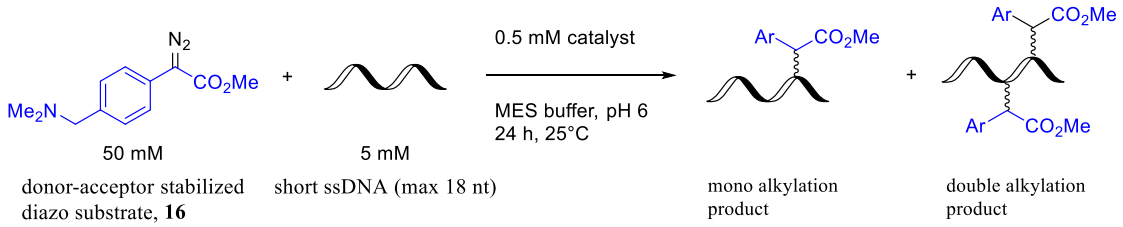


Scheme 3. Ruthenium carbenoid substrate-dependent NHI or SHI derived from a dansyl functionalized diazoacetate **8** and peptide **7**/**11** and **13**. (a) Ru catalyst **15**, phosphate buffer saline (PBS), room temperature, 1 h.

The discovery of metal carbene chemistry for biomolecules functionalization, especially nucleic acids, is of outstanding interest for the present study. In 2012, a method for post-synthetic nucleic acid alkylation was developed in our group. The method allowed modification of short and medium sized single-stranded DNAs and RNAs using rhodium(II) carbenoids derived from α-diazocarbonyl compound **16** and Rh_2OAc_4 (Scheme 4).[139] This alkylation method targets only the nucleobases resulting in a number of mono and

double alkylation products. NMR studies revealed that only the exocyclic amino groups in adenine, guanine and cytosine were modified.

One year later, Dr. Tishinov investigated other transition metals in nucleic acid modification. Particularly, the copper(I) carbenoid alkylation reactions showed promising results. Using the donor-acceptor stabilized diazo substrate **16**, Cu(II)SO_4 and sodium ascorbate (as the reductant) several model oligonucleotides were successfully alkylated.[140] The copper(I) catalytic system showed to be more efficient than rhodium(II) catalysis. The presence of the reducing agent is important to sustain the copper(I) oxidation state, which is the catalytically active species in this system.



Scheme 4. Rh(II)/ Cu(I) carbenoid nucleic acid alkylation using the donor-acceptor substituted diazo substrate **16**.

The copper(I) catalyzed alkylation reactions of nucleic acids proceed under similar conditions as the copper(I)-catalyzed alkyne-azide cycloaddition (CuAAC) for bioconjugation. The tris(3-hydroxypropyltriazolylmethyl)amine (THPTA) ligand was used to scavenge hydroxyl radicals generated by transition metals and ascorbate in the presence of molecular oxygen.[141,142] In the absence of a radical scavenger the formed radicals by an copper(I)-catalyzed electron transfer to molecular oxygen can lead to oxidation or cleavage of the nucleic acids.

The potential of copper(I) catalyzed carbenoid alkylation compared to rhodium catalysis was shown with a small series of model oligonucleotides (Table 1). In all alkylation reactions the modification was more efficient in the presence of copper(I) rather than rhodium(II). Additional differences are shown in d(TGT) and the T 4-mer alkylation (Table 1, Entry4/5 and 9/10). Modification of d(TGT) by Rh(II) carbenoids was only possible upon the addition of a second nucleotide (Table 1, Entry 3) and thymine could not be targeted at all (Table 1, Entry 9).

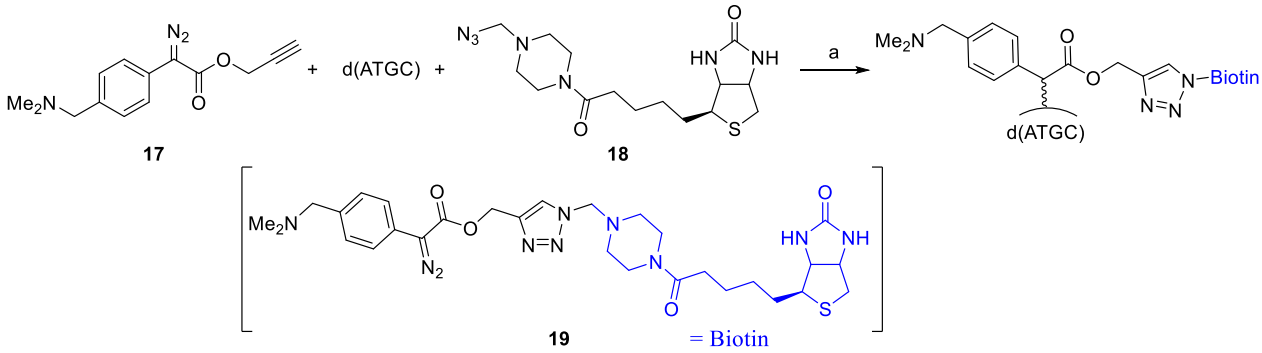
As previously discussed the alkylation products of d(TGT), d(TCT) and d(TAT) catalyzed by rhodium(II) were characterized by NMR. By comparison of the MS and HPLC traces for the rhodium(II) method with the copper(I) catalyzed reactions, the same regioselectivity and modification of the exocyclic amino group in cytidine, adenine, and guanine, was assumed.

A further advantage of the copper(I) catalyzed alkylation was the ability to run the reaction in tandem with CuAAC in a one-pot reaction mixture. Reaction of d(ATGC) with an alkyne functionalized diazo substrate **17** in the presence of an azide bearing biotin **18** yielded several NHI/click products. Unfortunately, the efficiency of the one pot reaction was rather low. It is likely that the click product **19** is formed first and the alkylation then took place. The ssDNA modification with a more bulky diazo substrate **19** could slow down the reaction rate and decrease its efficiency (Scheme 5).

Table 1. Comparison of copper(I) and rhodium(II) alkylation with model substrates and α-diazocarbonyl compound 16.

Entry	**Metal source**	**Asc [mM]**	**ssDNA**	**Number mono:**	**double:**	**ssDNA conversion, %[a]**
1	Rh_2OAc_4	-	d(TAT)	2	-	20
2[b]	$CuSO_4$	10	d(TAT)	2	-	37
3[c]	Rh_2OAc_4	-	d(TGT)	2	-	25
4[b]	$CuSO_4$	10	d(TGT)	4	1	46
5	Rh_2OAc_4	-	d(TCT)	2	-	16
6[b]	$CuSO_4$	10	d(TCT)	4	-	27
7	Rh_2OAc_4	-	d(ATGC)	4	-	56
8[b]	$CuSO_4$	10	d(ATGC)	7	4	70
9	Rh_2OAc_4	-	$d(T_4)$	-	-	-
10[b]	$CuSO_4$	10	$d(T_4)$	7	1	19

Reaction conditions: 5 mM ssDNA, 0.5 mM metal source, 50 mM diazo substrate **16**, 50 mM MES buffer pH 6.0, 24 h, room temperature. [a] Determined by HPLC analysis of the reaction mixture. [b] 2.5 mM THPTA. [c] 5 mM d(TAT), 5 mM d(TGT), 1mM Rh_2OAc_4.



Scheme 5. Tandem CuAAC/NHI labeling of d(ATGC) with alkyne conjugated diazo substrate **17** and biotin-azide derivative **18**.

2.5 Copper carbene selective alkylation of the O^6 position of guanine in mono- and oligonucleotides

The aim of the work discussed within this chapter was the development of a selective post-synthetic alkylation method for nucleic acids with copper(I) carbenoids.

2.5.1 Reaction discovery and scope of O^6 alkylation reaction

As presented in the last section, single-stranded nucleic acids were successfully alkylated by rhodium(II) and copper(I) carbenoids derived from the donor-acceptor stabilized α-diazocarbonyl compound **16**. These NHI reactions primarily targeted the exocyclic amino group of C, A and G. So far, only donor-acceptor stabilized diazo substrates were investigated in this catalytic system. Here, we present that changing the α-diazocarbonyl compound **16** to more reactive acceptor stabilized diazoacetates and amides dramatically changed the alkylation profile and chemoselectivity: the O^6 position in guanine (G) and inosine (I) was selectively targeted. As an example for the dramatic change in the alkylation profile TGT modification in the presence of ethyl α-diazoacetate (EDA) or diazo compound **16** was compared (Figure 24). The crude HPLC trace of TGT alkylation in the presence of the acceptor stabilized diazo compound shows one single peak for the O^6 G alkylation product (Figure 24A) compared to several alkylation products in the presence of the donor-acceptor stabilized diazo compound **16** (Figure 24B).

The potential of copper(I) carbenes derived from ethyl α-diazoacetate (EDA) selective alkylating the O^6 position of guanine and inosine in monophosphates (MPs), triphosphates (TPs) and short ssDNAs is presented in Table 2. All alkylation products were analyzed by tandem mass fragmentation and showed the corresponding ions for nucleobase (guanine and inosine) alkylation (representative examples: Figure 25 and Figure 26). A detailed verification of the alkylation site at the nucleobase is presented in section 2.5.2 pointing to the O^6 position. We explored the efficiency and selectivity of O^6-G alkylation by characterizing reactions with nucleotide monophosphates. Treatment of dGMP or GMP with 20 mol% copper(I) and commercial EDA in aqueous buffer gave the corresponding O^6 alkylation products in high yields within 30 min reaction time (Table 2, Entry 1 and 2). Copper(I) was created *in situ* by mixing $Cu(II)SO_4$ and sodium ascorbate. Under the same conditions, the structural analogues IMP, GTP and dGTP were also efficiently modified at their O^6 positions (Table 2, Entry 6, 11 and 12). However, the absence of the copper(I) catalyst did not led to modification (Table 2, Entry 5); and replacing the starting material with AMP, CMP, UMP or TMP resulted in little or no detectable alkylation (Table 2, Entry 7-10). In short, only substrates that contain either a guanine or inosine in the structure were converted to the desired alkylation product, while the other substrates were partially converted through unselective pathways that we could not completely characterize. In each case the mass for alkylation was usually detected, but these did not give substantial peaks in the HPLC that we could isolate. A similar observation was found for d(TAT) (Table 2, Entry 14). Some starting material was consumed but no reasonable amount of any alkylation product could be isolated.

In contrast, when guanine was present post-synthetic modifications of a DNA 3-mer d(T**G**T) and 9-mer d(TTTT**G**TTTT) with copper(I) carbenes derived from EDA yielded efficiently a single alkylation product (Table 2, Entry 13 and 17, Figure 25D+E).

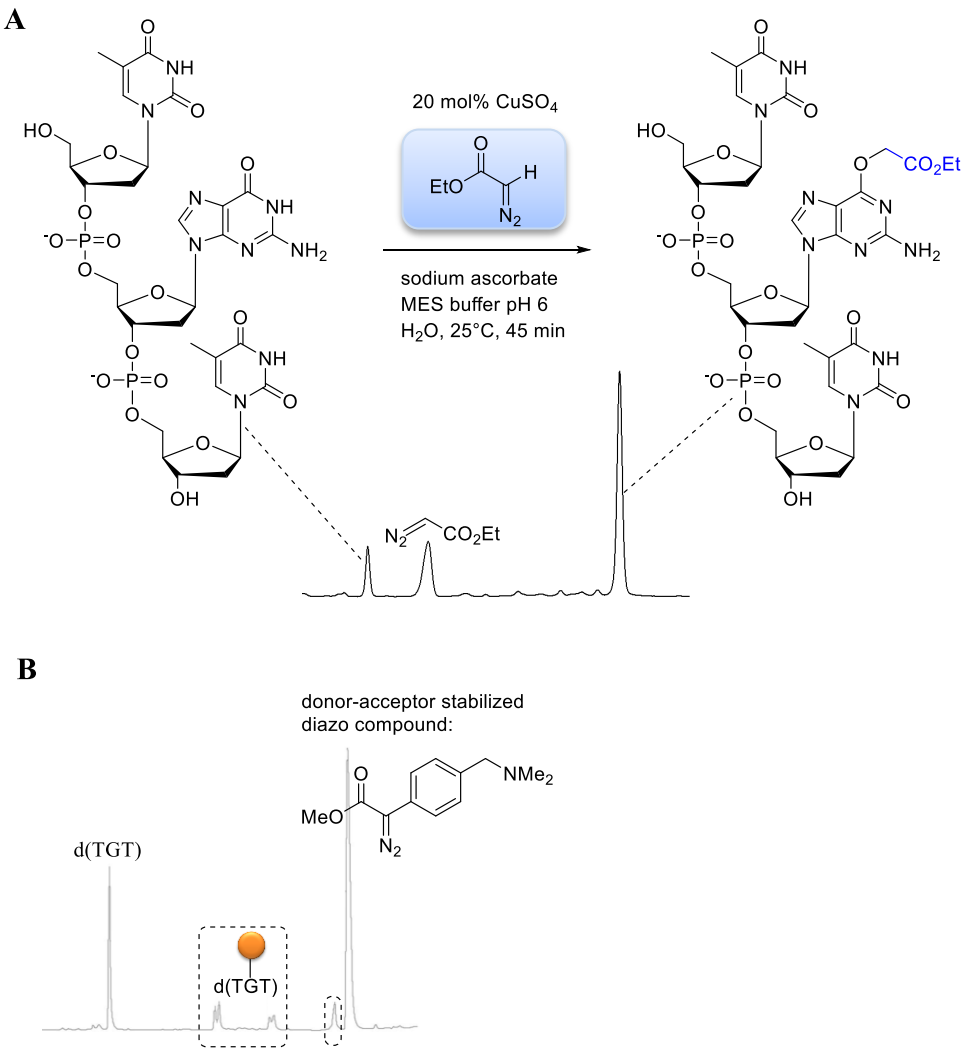
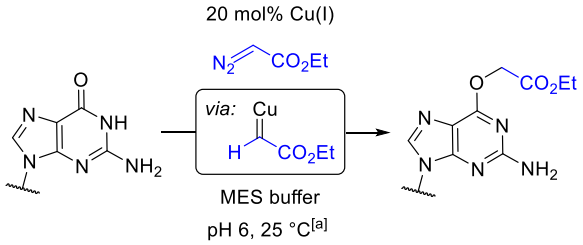


Figure 24. (A) HPLC analysis of d(TGT) alkylation in the presence of the acceptor stabilized diazo compound ethyl α-diazoacetate (EDA) and *in situ* formed copper(I) after 45 min reaction time (Table 2, Entry 13). (B) HPLC analysis of d(TGT) alkylation with the donor-acceptor stabilized diazo compound **16** and *in situ* formed copper(I) after 24 h reaction time (Table 1, Entry 4).

Table 2. Alkylation potential of copper(I) carbenoids derived from ethyl α-diazoacetate (EDA) with mono- and oligonucleotides.



Entry	substrate	Time [min]	Cosolvent	Conversion of substrate [%]	Yield [%]
1	d**G**MP	30	none	98[d]	86
2	**G**MP	30	none	96[d]	92
3	**G**MP	30	DMSO	93[d]	-
4[b]	**G**MP	30	none	87	-
5[c]	**G**MP	150	none	0	-
6	**I**MP	30	none	92[d]	81
7	AMP	30	none	40	-
8	CMP	30	none	38	-
9	UMP	30	none	21	-
10	TMP	30	none	13	-
11	**G**TP	30	^{t}BuOH	98[d]	73
12	d**G**TP	30	^{t}BuOH	97[d]	65
13	d(T**G**T)	15/45	none	62/8[d]	-
14	d(TAT)	120	none	35[e]	-
15	d(AT**G**C)	150	^{t}BuOH	38[d]	-
16	d(AT**G**C)	150	dioxane	64[d]	-
17	d(TTTT**G**TTTT)	30	DMSO	82[d]	-
18	**G**	19[f]	DMSO	70	-
19	A	19.5[f]	DMSO	11	-

[a] Complete reaction conditions: 5 mM substrate, 1 mM $CuSO_4$, 5 mM sodium ascorbate, 50 mM EDA, 100 mM MES buffer pH 6, H_2O, 20% (v/v) cosolvent, 25°C. [b] No sodium ascorbate added. [c] No $CuSO_4$ and sodium ascorbate added. [d] G-alkylation site was verified on the basis of tandem MS fragmentation ions. [e] Six mono alkylation products were detected by ESI, but the amount was too low for isolation. [f] Reaction time corresponds to hours not minutes.

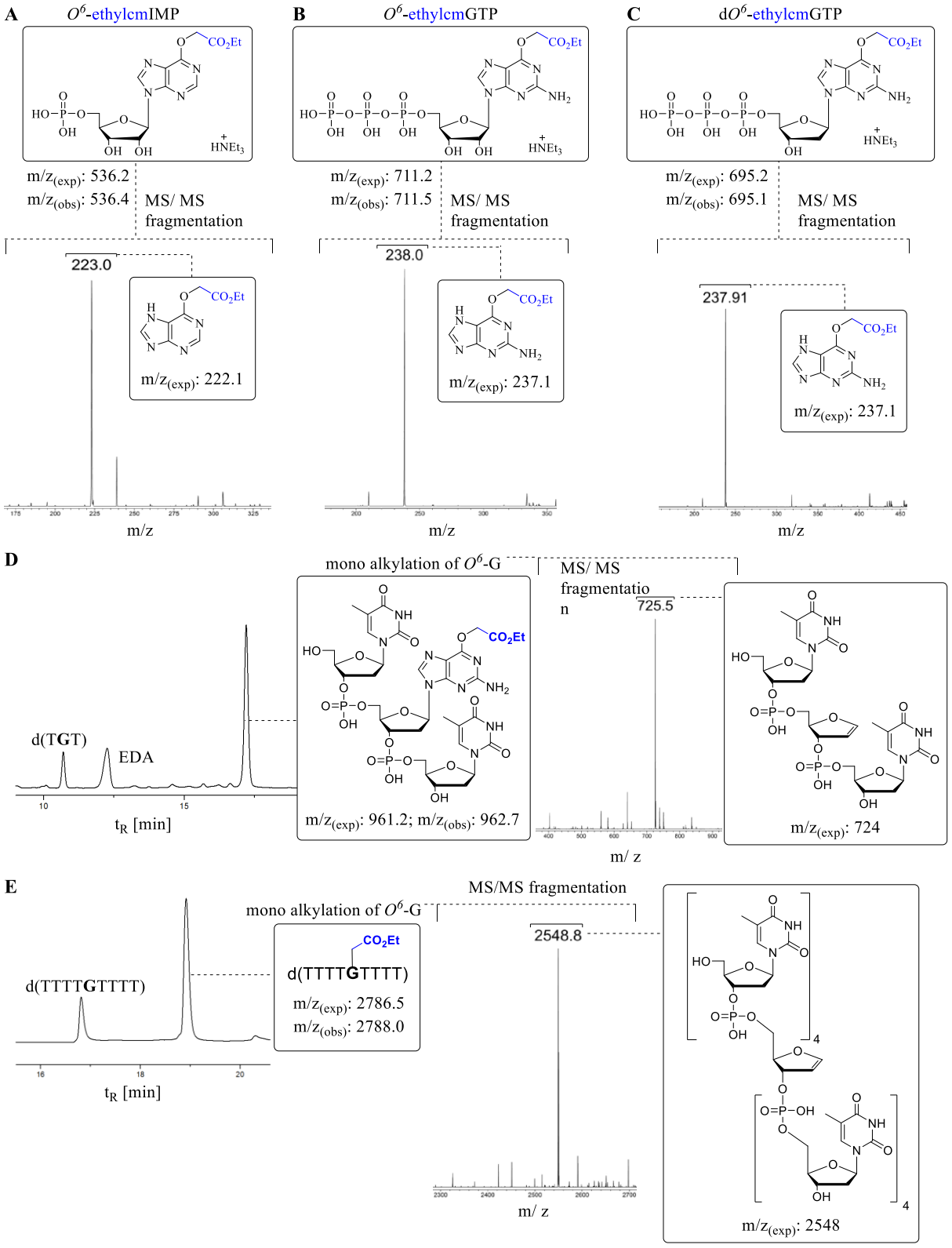


Figure 25. HPLC-MS/MS analysis of O^6-G alkylation with copper(I) carbenoids derived from EDA (Table 2) (A) IMP (Entry 6). (B) GTP (Entry 11). (C) dGTP (Entry 12). (D) d(TGT) (Entry 13). (E) d(TTTTGTTTT) (Entry 17).

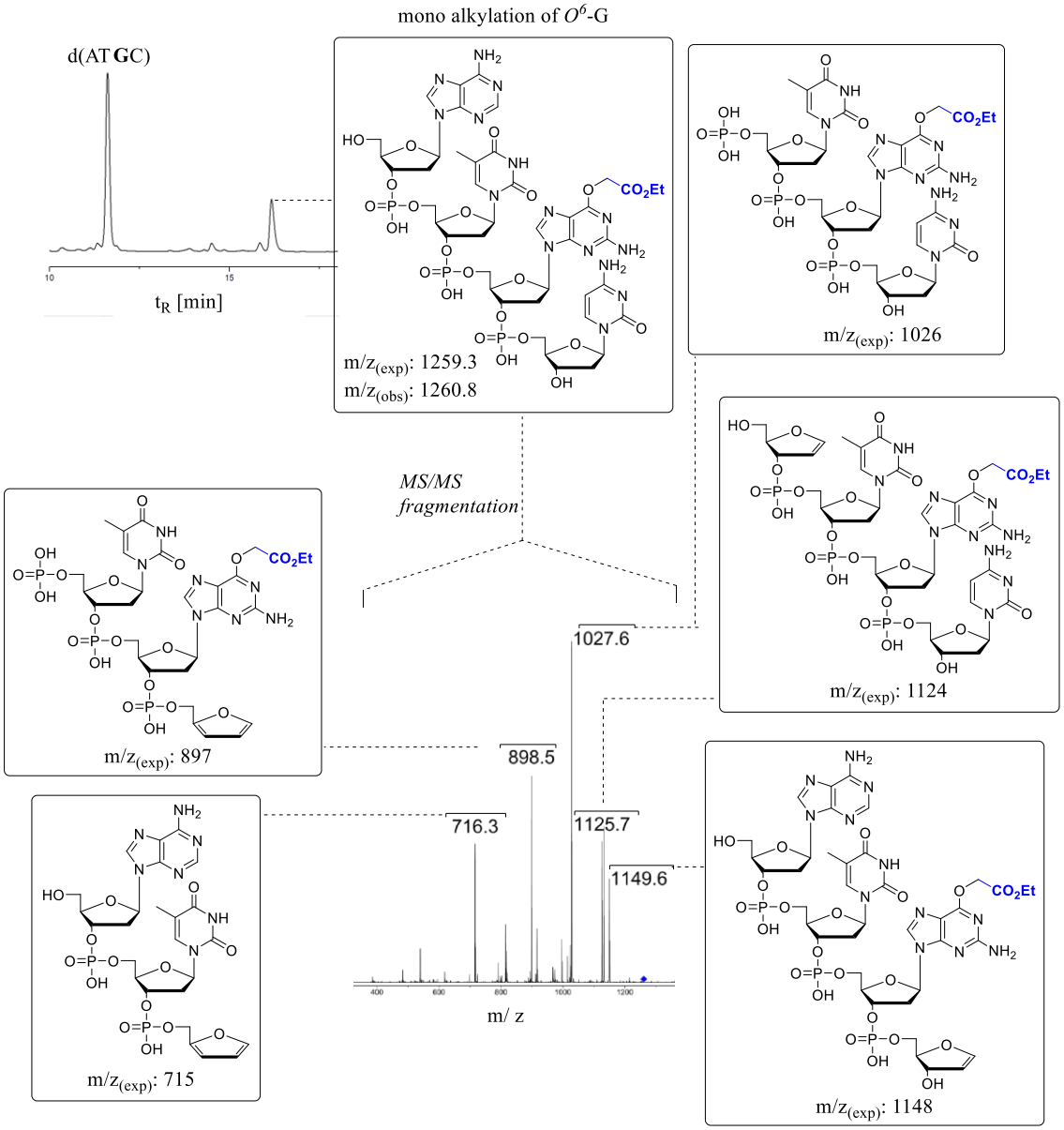


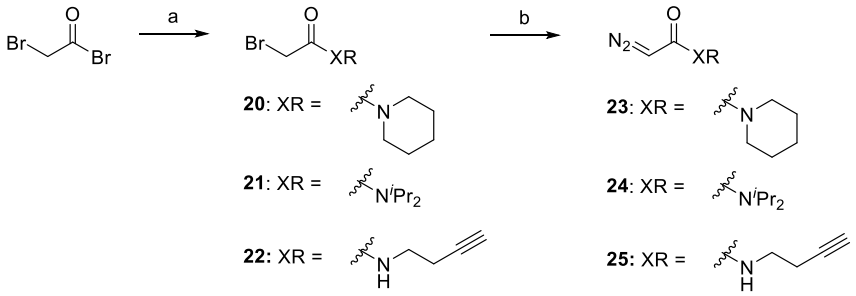
Figure 26. d(ATGC) HPLC-MS/MS analysis of O^6-G alkylation reactions with copper(I) carbenoids derived from EDA (Table 2, Entry 15).

While oligonucleotides were also selectively alkylated at the O^6-G position, the reactions were slower. For example, the O^6-G alkylation of d(ATGC) with copper(I) carbenoids (Table 2, Entry 15; Figure 26), required longer reaction time and the conversion of the ssDNA 4-mer was lower compared to mononucleotides (cf. Table 2, Entries 11 and 15). This reduced reactivity of d(ATGC) can likely be attributed to the unproductive sequestration of copper(I) to other binding sites in the oligonucleotide, for example the nucleobases A, T and C as well as the ribophosphate backbone. These results already indicated the limitation of this alkylation method by the length of the oligonucleotide. However, the lower reactivity

of oligonucleotides could be diminished by the addition of 1,4-dioxane as co-solvent and showed almost twice the amount of d(ATGC) conversion when using a stock solution of EDA in 1,4-dioxane instead of 10 % ^{t}BuOH in water (Table 2, Entry 16).

Diazo substrate variation is another way to improve the reaction efficiency with challenging nucleic acids. In the previous section it was mentioned that the electronic character of the substituents in α-position to the diazo functionality has an influence on its stability and reactivity. Here, we demonstrate that the diazo substrate change from the diazoacetate (EDA) to acetamides has an influence on the catalytic system. The alkylation of ssDNAs with diazoacetamides (DAAs) led to higher conversion numbers and the high O^6 selectivity was maintained (Table 3). Modification of the model substrate d(ATGC) for example, delivered the O^6-G adduct with more than 60% conversion of the starting material with both DAAs **23** and **24** as carbene source (Table 3, Entry 1 and 2), whereas the same substrate led to only 38% conversion with EDA (Table 2, Entry 15).

The DAAs were efficiently synthesized in two steps, according to an adapted procedure by Fukuyama and coworkers.[143] Reaction of commercially available bromoacetyl bromide (BrAcBr) with the secondary amines (piperidine, *N*,*N*-diisopropylamine and 1-amino-3-butyne) followed by treatment with *N*,*N*-ditosylhydrazine in the presence of DBU yielded the final DAAs **23-25** (Scheme 6).

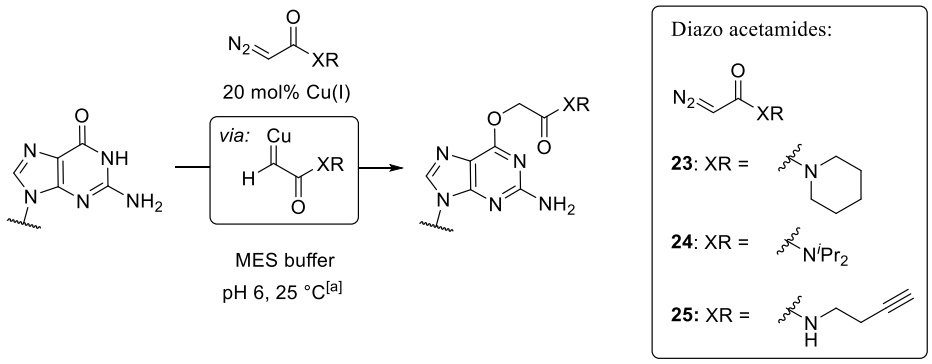


Scheme 6. Synthesis of diazo acetamides (DAAs). (a) *N*,*N*-diisopropylamine/piperidine, CH_2Cl_2, -60°C to 25°C, 30 min (**20**: 31%), (**21**: 52%), (**22**: 75%). (b) *N*,*N*-ditosylhydrazine, DBU, THF, 0°C-25°C, 20 min, (**23** and **24**: 60%), (**25**: 34%).

While the DAAs are better substrates in nucleic acid modification, they are also less stable. For example, the preparation of accurate stock solutions in 10 % ^{t}BuOH and water was impossible for the DAAs due to rapid decomposition. These findings can be explained by the change of the "acceptor" substituent in α-position to the diazo group. While the amide group in DAAs stabilizes the diazo group, it does this to a smaller extent than a carboxy group due to the higher electron density of amides. Therefore, DMSO was used as co-solvent to stabilize the highly reactive DAAs in the alkylation reactions. Unfortunately, the coordinating co-solvent lowers the reactivity of the nucleic acids. Changing the reaction medium to a less

coordinating co-solvent such as dioxane was the best compromise between these competing affects, d(ATGC) was efficiently modified in a fifth of the time (cf. Table 3 Entry 3 and 4). Summarized, the DAAs are better substrates for nucleic acid modification and even longer ssDNA 16-mers were efficiently alkylated (Table 3 Entry 8 and 9).

Table 3. Alkylation reactions of diazoacetamides (DAAs) with oligonucleotides

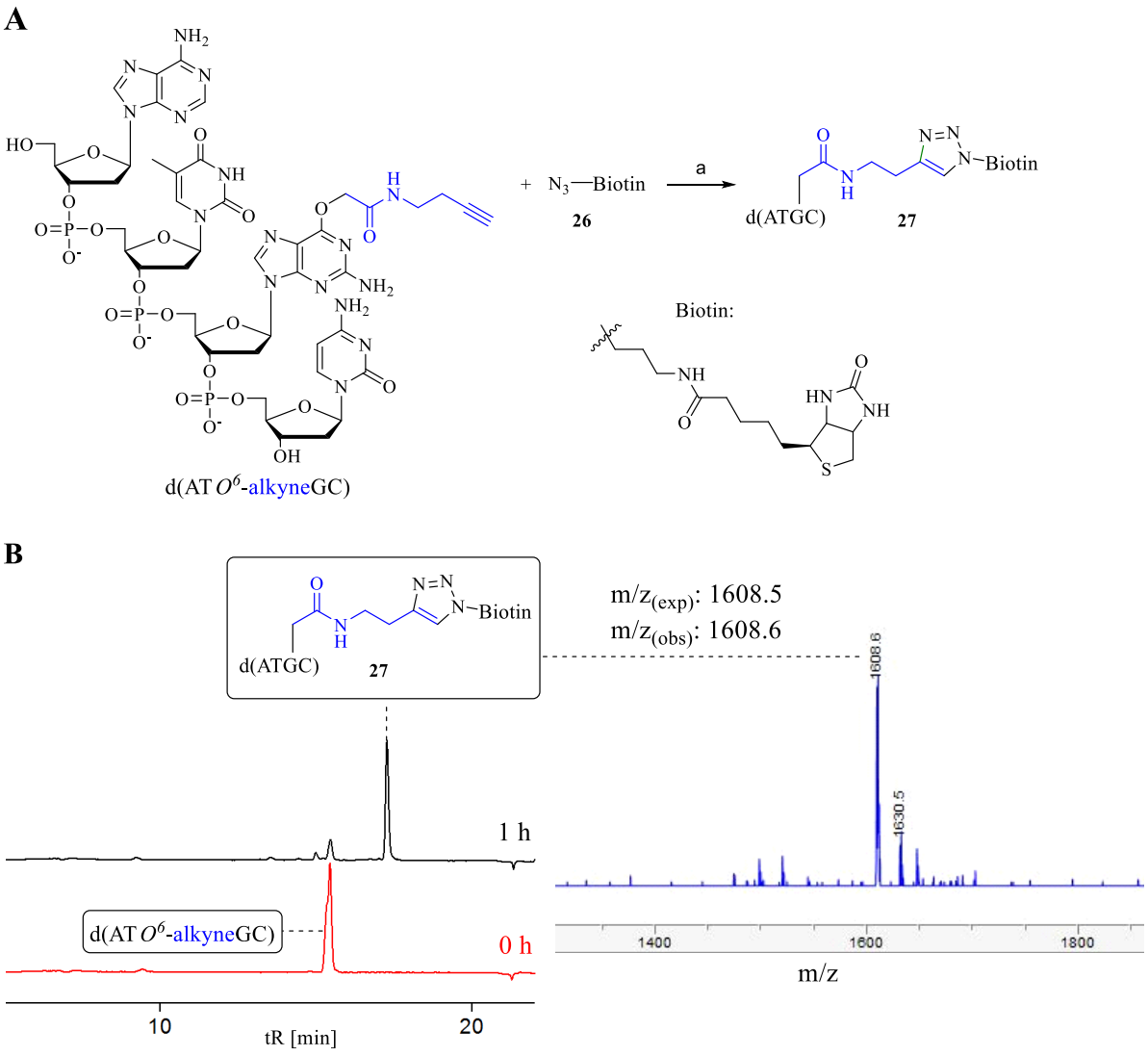


Entry	ssDNA	Diazo-acetamide	Time [min]	Conversion of ssDNA [%][d]
1	d(T**G**T)	**23**	15	78[b]
2	d(T**G**T)	**24**	15	78[b]
3	d(AT**G**C)	**23**	150	62[b]
4	d(AT**G**C)[c]	**23**	30	74[b]
5	d(AT**G**C)	**24**	150	61[b]
6	d(TTTT**G**TTTT)	**23**	30	92[b]
7	d(TTTT**G**TTTT)	**24**	30	81[b]
8	d(AACA**G**TCATATCCTTA)	**23**	180	48
9	d(AACA**G**TCATATCCTTA)	**24**	180	40
10	d(AT**G**C)[c]	**25**	150	66[b]

[a] Complete reaction conditions: 5 mM substrate, 1 mM $CuSO_4$, 5 mM sodium ascorbate, 50 mM diazoacetamide, 100 mM MES buffer pH 6, H_2O, DMSO 20% (v/v), 25°C. [b] G-modification was verified on the basis of tandem MS fragmentation ions. [c] Same conditions as *a* but 20 % (v/v) of dioxane was used instead of DMSO. [d] Determined by HPLC-analysis of the reaction mixtures.

The stability of the amide bond in the DAAs offers also the advantage to introduce different functional groups by simply changing the amine source in the diazo synthesis. The attached functionalities can be further used for bio-conjugation reaction, labelling or pull downs. As an example, the model substrate d(ATGC) was modified at the O^6-G position with a DAA bearing a terminal alkyne **25** (Table 3 Entry 10)

and further functionalized by click reaction with biotin azide derivative **26**. In summary, our developed copper(I) carbene chemistry allowed the efficient post-synthetic O^6-G modification of a ssDNA 4-mer **27** with a complex pull-down tag in only two steps (Scheme 7).



Scheme 7. Example for DAA substrate modularity - can be used for pull down enrichment. (A) Click reaction of O^6-G alkyne modified d(ATGC) with biotin azide. (B) HPLC analysis of the click reaction mixture after 0 h (red trace) and after 1h (black trace) with corresponding mass trace for click product detected by UPLC-MS analysis. (a) $CuSO_4$, sodium ascorabate, DMSO, TEAA pH 7, 25°C, 1 h.

2.5.2 Determination of the alkylation site

The surprising selectivity of copper(I) carbenoids derived from α-diazocarbonyl compounds prompted us to investigate the alkylation site in detail. The tandem mass analysis of the short ssDNA and MP already showed that the alkylation took place at the nucleobases guanine and inosine, respectively. Therefore a model substrates dGMP was chosen to determine the modification site on the nucleobase via 2D-NMR analysis of the isolated alkylation products **28** (Figure 27A). The arrows in Figure 27A show all possible alkylation sites of guanine. The obtained coupling over four bonds in the heteronuclear multiple-bond correlation (HMBC) spectra from the imidazole proton H8 of guanine led us to identify the chemical shift of the carbon C6 (Figure 27B). A second coupling over three bonds from this carbon C6 to the CH_2 group of the newly attached moiety narrowed the possible points of attachment to either the O^6 or N1 position. The two possible structures are presented in the black box in Figure 27B. Furthermore, the ^{13}C chemical shifts were calculated for both possible alkylation products **29** and **30** using Chemdraw (Figure 27B dashed box). Comparison of the calculated ^{13}C chemical shifts (highlighted with blue numbers) with the measured one from the CH_2 group of compound **28**, suggested modification at the O^6 position. Additionally, no correlation between N^1 and the attached CH_2 functionality was obtained in the ^{15}N-HMBC analysis. This correlation over two bonds would be expected if N1 is alkylated. (Figure 27B).

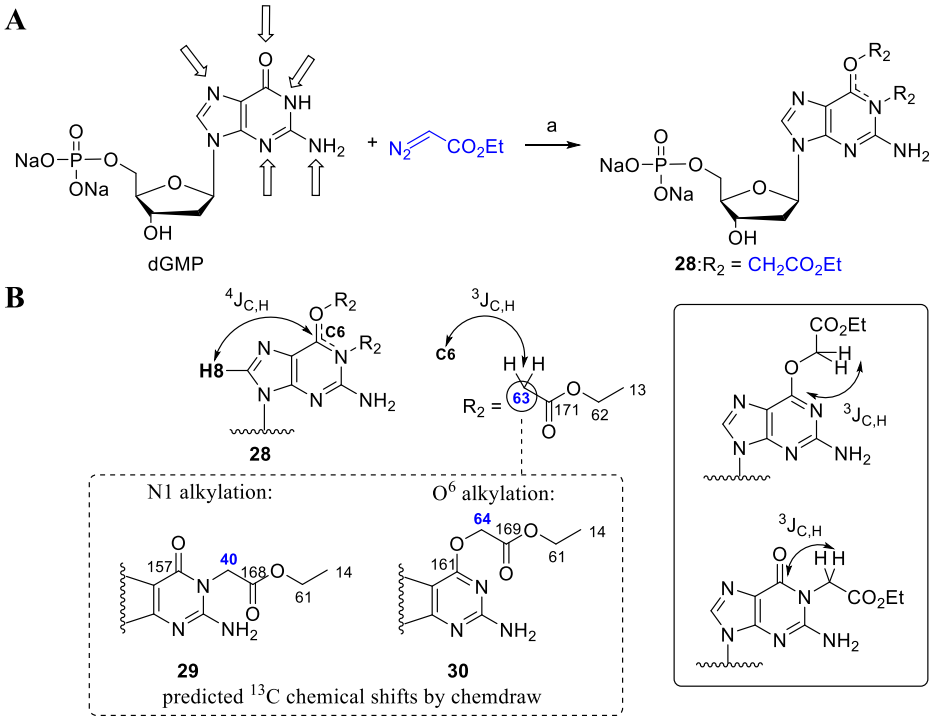


Figure 27. (A) Determination of the alkylation site with the model substrates dGMP. (B) Important signals for structure elucidation isolated from 2D-NMR analysis. Predicted ^{13}C-NMR chemical shift with Chemdraw (dashed box); two representative ^3J couplings for either O^6 or N1 alkylation, respectively (black box). (a) $CuSO_4$, sodium ascorbate, MES buffer pH 6, H_2O, 25°C, 5 h (60 %).

Nevertheless, we decided to synthesize an O^6-(ethyl carboxymethyl)-inosine (O^6-(ethylcm)-I) adduct **31** by a traditional synthetic route as additional verification of the alkylation site (Figure 28, pathway B). Commercially available 6-chloropurine riboside was protected as *tert*-butyldimethysiliyl (TBDMS) ether at the 2`, 3`and 5` hydroxyl groups. Nucleophilic aromatic substitution of the chloride bearing ribonucleoside **32** with ethyl glycolate followed by deprotection of the silyl groups yielded the O^6-(ethylcm)I **31** in 37% over three steps. Furthermore, the copper(I) carbene chemistry derived from EDA was used to yield the inosine alkylation product **31** in 60% in one step (Figure 28, pathway A). Subsequent structure elucidation of both compounds **31** isolated from the different pathways A and B showed identical NMR spectra as well as HPLC traces (Figure 29) and confirmed the proposed O^6 modification site.

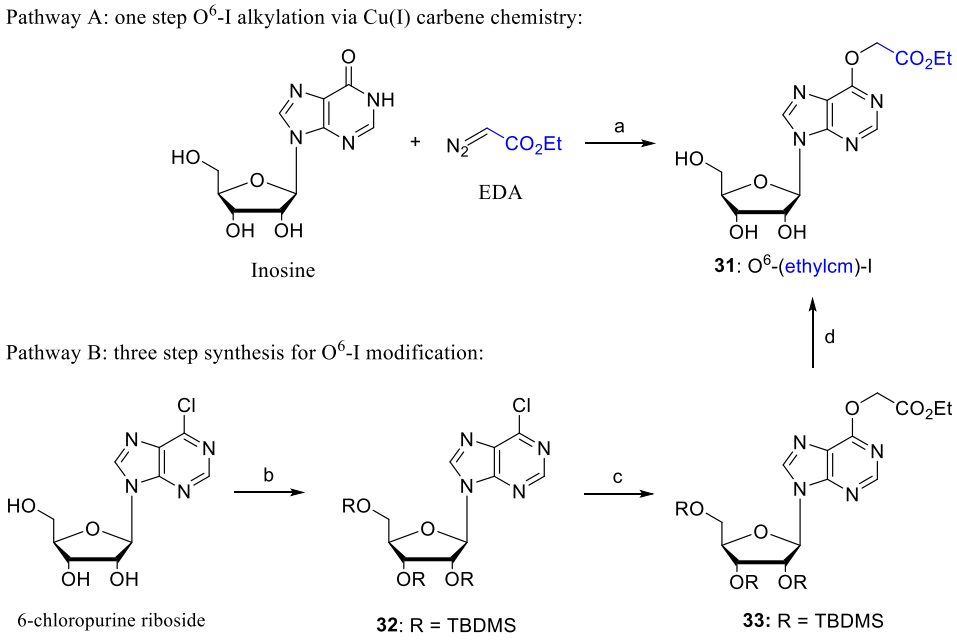


Figure 28. Verification of O^6 alkylation. Pathway A: synthesis of O^6-(ethylcm)I **31** using copper(I) carbenoid chemistry in one step. Pathway B: synthesis of O^6-(ethylcm)I **31** via 3 steps synthesis. (a) $CuSO_4$, sodium ascorbate, MES buffer pH 6, H_2O, 6% (v/v) DMSO, 25°C, 5 h (60 %). b) TBDMS-Cl, imidazole, DMF, 25°C, on. (77 %). c) ethylglycolate, NaH, THF, reflux, 4 h (48 %). d) Pyridine·HF, pyridine, 25°C, 16 h (99%).

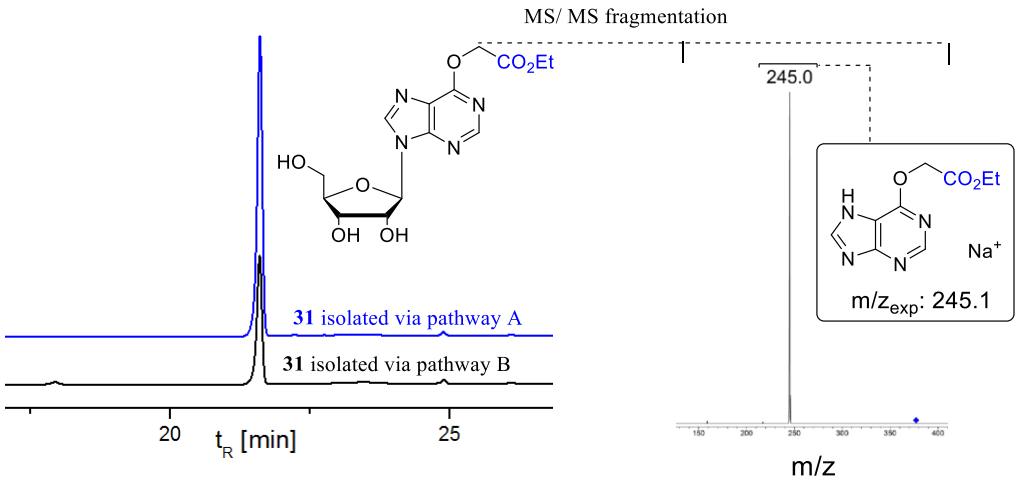


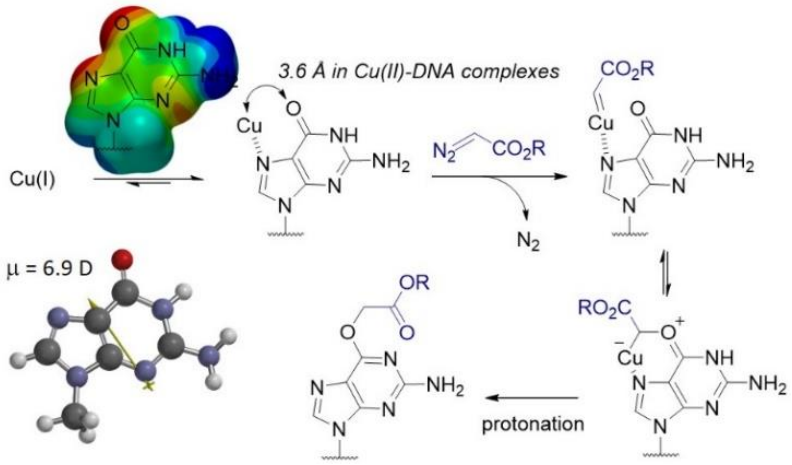
Figure 29. HPLC analysis O^6-(ethylcm)-I **31** synthesized via pathway A (blue trace) and pathway B (black trace) and corresponding tandem MS fragmentation.

2.5.3 Mechanistic hypothesis accounting for the high O^6-G chemoselectivety

The obtained selectivity for copper carbenes derived from acceptor-stabilized diazo compounds does not reflect the reported nucleophilicity for the DNA nucelobases.[8] The N7-G for example is reported as the most potent nucleophile in dsDNA and ssDNA and targeted by other alkylating agents five to tenfold higher than the O^6-G position. We assumed that a pre-association of the copper(I) catalyst with guanine is the key for the resulting selectivity, since unstabilized copper(I) carbene should readily react with water. Guanine is reported as the most electron rich nucleobase and is able to form chelates with divalent metals such as copper(II).[144-145] Additionally, DFT calculations of the molecular dipoles of nucleobases indicated that guanine has the largest dipole (6.9 D), negatively polarized between O^6 and N7 (Scheme 8). Unfortunately, not much is known about the interaction between copper(I) and DNA.[146] The sensitive oxidation state of copper(I) in aqueous buffer at neutral pH could be the reason for the lack of investigations.[147] However, copper(II) is reported to bind to the N7 position of guanine and therefore we hypothesize that the selectivity derives from the carbene formation with an N7-G bound copper(I) catalyst. The crystal structure of copper(II)[148] and other divalent metals[145] demonstrate a strong preference for N7-G coordination placing the copper center approximately 3.5 Å away from the O^6 position. The formed copper carbene would be poised in the same manner to form the metal ylide with the O^6 position (Scheme 8). Subsequent protonation would yield the O^6 alkylation product.

The obtained selectivity reflects also the preferred site for copper-catalyzed oxidative DNA damage (induced by copper with peroxide and superoxide) that shows enhanced oxidation at and around guanines.[149]

There are two possible pathways reported for copper-mediated redox damage. The first pathway goes through a hydroxyl radical that is formed by unbound copper(I) induced hydrogen peroxide cleavage or reduction of reactive oxygen compounds. If the hydroxyl radical is close to DNA it can lead to DNA fragmentation. This unselective mechanism poses a constant danger for DNA and RNA degradation but it is also inefficient. In the other pathway the copper(I) is bound to guanine and forms a copper(II) oxo complex with peroxide that guides the oxidative damage to guanine itself or the neighbouring bases. Unfortunately, the reports about oxidative damage of DNA gave little insight into the mechanism or why guanine is targeted or where exactly copper is bound. We assume that the specificity in our alkylation reaction with copper(I) carbenes and the oxidative damage induced by copper is a results from the same thing: the ability of guanine to coordinate to the copper species.



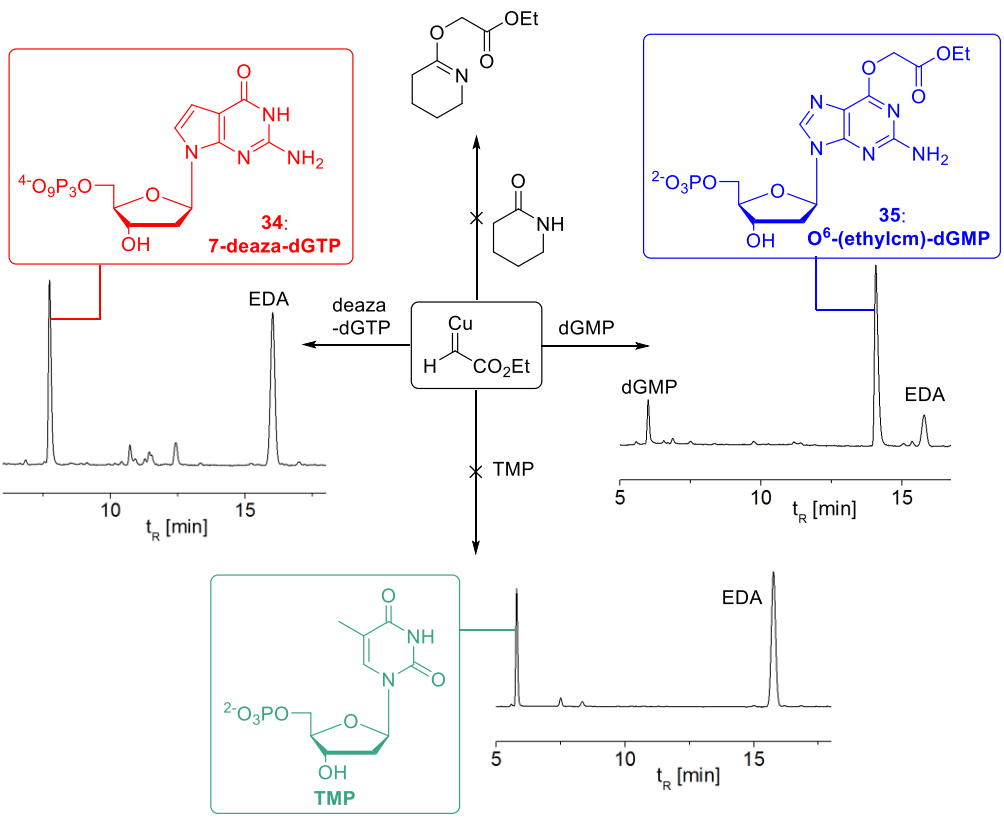
Scheme 8. Mechanistic hypothesis accounting for the selective O^6-modification in guanine.

Our hypothesis that the pre-coordination of copper to guanine guides the selectivity of the reaction would predict the following:

(1) Poor reactivity of other amides or lactones that do not bear the same coordination ability
(2) Blocking the coordination site of the copper center by strongly chelating ligands would extinguish the reactivity because coordination of both substrates the guanine and the chelating ligand would not leave coordination space to form the copper(I) carbene.

1. Poor reactivity of other amides or lactones that do not bear the same coordination ability. The first substrate of choice to investigate the influence of the coordination ability of guanine in the alkylation reaction was the 7-deaza variant **34** of dGTP. Comparing the alkylation reaction of normal guanine in dGTP (HPLC traces Scheme 9) with the reaction outcome of the 7-deaza variant **34** a substantial reduced reactivity and selectivity is observed. Additionally, the dTMP and the lactam piperidine-2-alkylation are similarly

unreactive. These results add further weight that the N7 coordination is the important factor for selective O^6 alkylation.



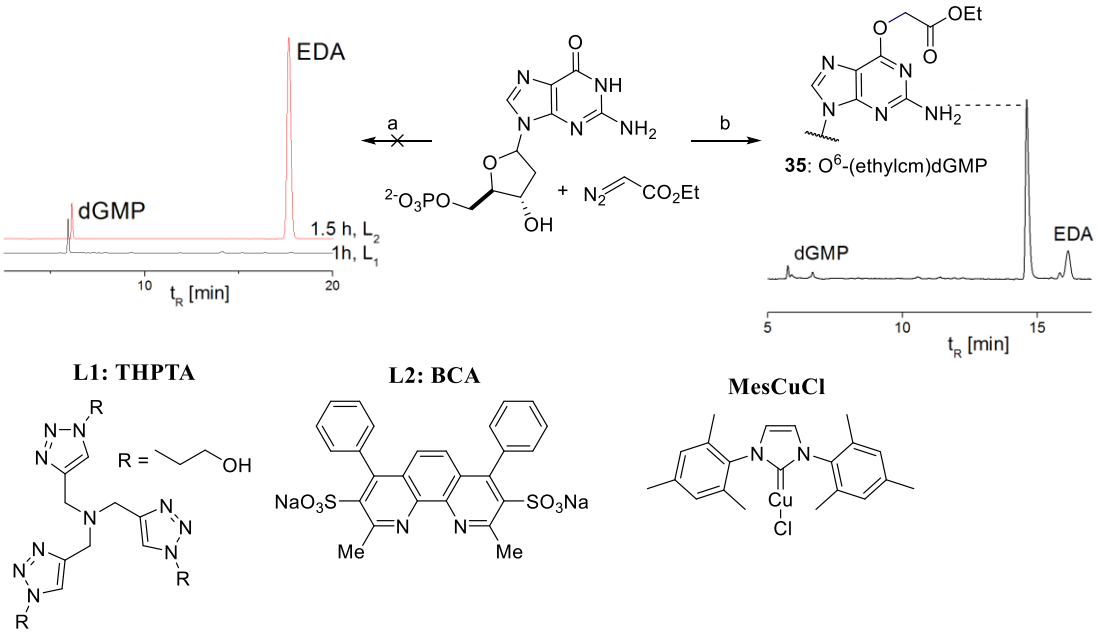
Scheme 9. Mechanistic studies that point to the important role of the N7 in guanine guiding the alkylation reaction to the O^6 position. The HPLC traces were recorded after 30 min reaction time. The reactions conditions: 5 mM substrate (dGMP, 7-deaza dGTP **34**, TMP or a-valerolactone), 1 mM $CuSO_4$, 5 mM sodium ascorbate, 50 mM EDA, 100 mM MES pH 6, dioxane 20 % (v/v), H_2O, 25 °C.

2. The guanine alkylation is inhibited in the presence of multidentate ligands, however, promoted with N-heterocyclic carbene (NHC) ligands.

In the CuAAC that is widely used for biochemical transformation, ligands such as tristriazoles and bathocuproine are employed for copper(I) stabilization and acceleration of the click reaction.[150-152] However, the presence of strong coordinating ligands such as tris(3-hydroxypropyltriayolylmethyl)amine (THPTA) or bathocuproinedisulfonic acid disodium salt (BCA) prohibited alkylation of dGMP (Scheme 10, left pathway). Even after several hours, however, little to no alkylation products were detected. In the

presence of THPTA, EDA was completely consumed after one hour but no dGMP alkylation occurred. The copper carbene derived from EDA is formed in this case but decomposed in an unproductive manner, likely through reaction with water.

In contrast, *N*-hetreocyclic carbene (NHC) stabilized copper(I) behaved different in dGMP alkylation. The reaction of dGMP with EDA in the presence of commercial mesitylimidazolium copper chloride (MesCuCl) yielded a similar alkylation profile as the standard conditions ($CuSO_4$ and ascorbate) with O^6-(ethylcm)dGMP **35** as main product (Scheme 10, right pathway). The NHC, as monodentate ligand, still allows the coordination to the N7-G and therefore the desired alkylation can proceed. These types of ligands have other favorable properties such as, the reduction of oxidative damage; These are explored in detail in chapter 3 in nucleic acid alkylation reactions.



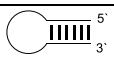
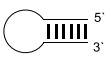
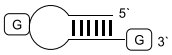
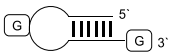
Scheme 10. Multidentate ligands such as THPTA and BCA prevent dGMP alkylation (left pathway) and efficient dGMP alkylation catalyzed by MesCuCl (right pathway). (a) 20 mol % $CuSO_4$, 20 mol% Ligand **L1** or **L2**, 10 eq. EDA, 1 eq. sodium ascorbate, Mes buffer pH 6, 20% (v/v) dioxane, H_2O, 25°C. (b) 20 mol % MesCuCl, 10 eq. EDA, Mes buffer pH 6, 30% (v/v) dimethoxyethane, H_2O, 25°C, 30 min. The HPLC traces were recorded after 30 min.

2.5.4 Copper(I) carbenoid alkylation of double-stranded DNAs (hairpin motifs, mismatch double stranded DNA)

Double-stranded (ds) DNA is primary structure present in cells if the DNA is not being actively replicated. In that sense, the copper(I) alkylation system to modify double-stranded DNA is an important feature concerning the reaction scope of nucleic acids alkylation.

In order to test the post-synthetic alkylation reaction on dsDNA, three model substrates that form hairpin structures were selected (Table 4). The advantages of using hairpin instead of linear dsDNAs are reflected in the synthesis and pre-preparation as the duplex has not to be formed by annealing equimolar amounts of two single-strands. The hairpin is synthesized as single strand that ensures the quantitative duplex formation. Two similar hairpin structures were chosen bearing a small change in the loop and at the 3`position. Both hairpins contained a small loop of five nucleotides and a duplex region of six base pairs. The third hairpin had a longer duplex region with a length of 10 base pairs and an additional 5` guanine overhang.

Table 4. Scope of copper(I) alkylation with double stranded DNA

Entry	Hairpin substrate	Hairpin Structure	$CuSO_4$ [mM]	Time [h]	Conversion[a] [%]
1	d(CGAACGTTTTTCGTTCG)		1 mM	10	8
2	d(CGAACGTTTTTCGTTCG)		5 mM	24	40
3	d(CGAACGTT**G**TTCGTTCG**G**)		1 mM	-	19[b]
4	d(CGAACGTT**G**TTCGTTCG**G**)		5mM	22	38
5	d(GCGGAATTCCGTTTTTCGGAATTCCG)		1 mM	6	48[c]
6	d(GCGGAATTCCGTTTTTCGGAATTCCG)		5 mM	2	60[c]

Reaction conditions: 5 mM hair-pin, 50 mM EDA, 5 eq. of sodium ascorbate compared to $CuSO_4$ concentration, 100 mM Mes buffer pH 6, 25°C. [a] The conversion was determined by HPLC analysis of the reaction mixtures. [b] The conversion was determined at the time point when all EDA was consumed, due to co-elution of EDA with the hair-pin. [c] 100 mM DAA **23** was used instead of EDA.

The alkylation of the hairpins was investigated with copper(I) carbenes derived from EDA under the standard conditions that were worked out for the MPs/TPs and single-stranded DNAs. The hairpin bearing no guanine in the loop or as overhang was not modified under these conditions (Table 4, Entry 1, Figure

30A), indicating that guanines engaged in double strands are not targeted. The hairpin structure bearing two accessible guanines (in the loop and as 3`overhang) was reacted under the same condition resulting in moderate conversion of the starting material (Table 4, Entry 3). Unfortunately, the concentration of the newly formed species was too low to characterize it (Figure 30C).

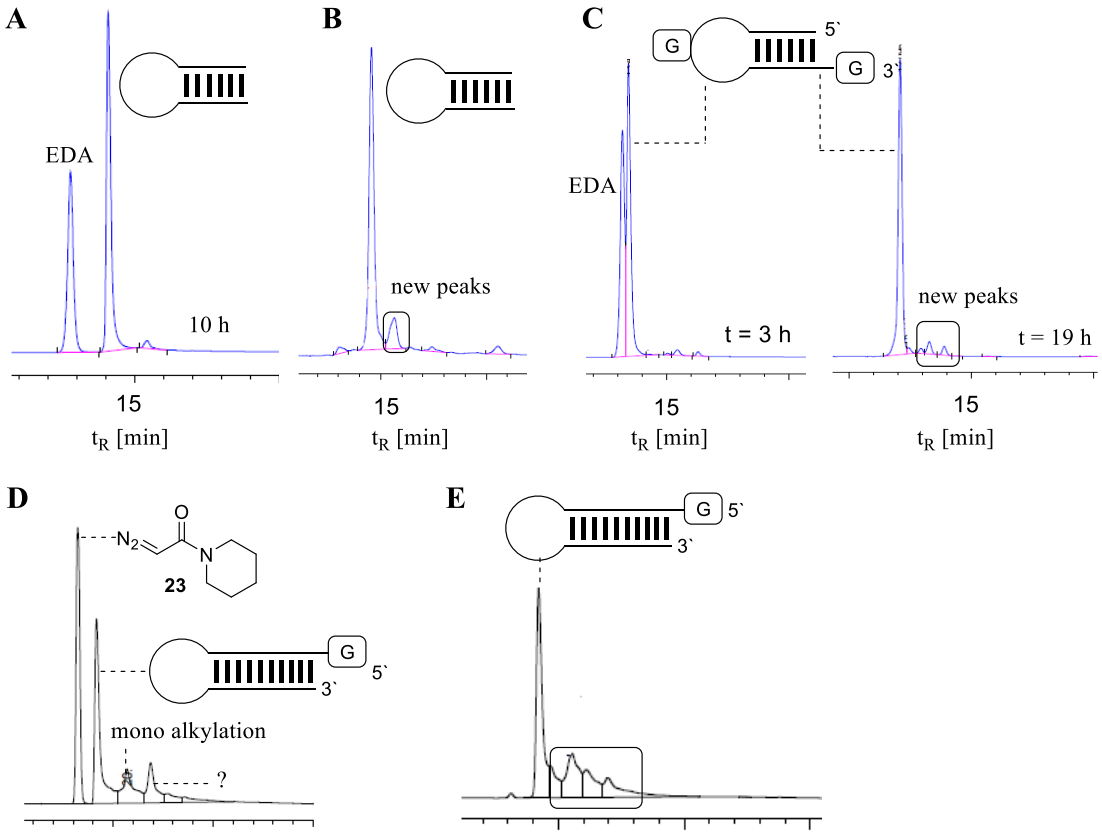


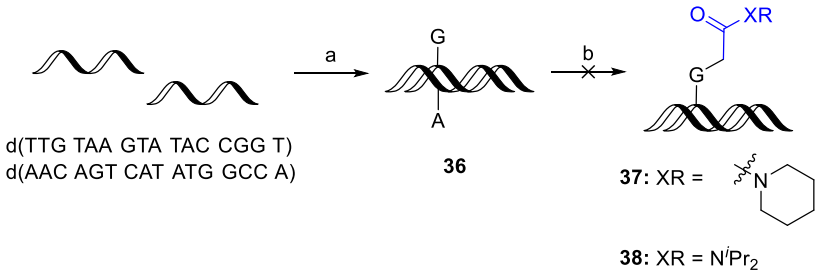
Figure 30. HPLC traces of hair-pin alkylation reaction in the presence of copper(I) carbenoids. (A) hair-pin (T-loop) alkylation reaction with 20 mol% copper(I) and EDA (Table 4, Entry 1). (B) hair-pin (T-loop) alkylation reaction with 1 eq. of copper(I) and EDA (Table 4, Entry 2). (C) hair-pin (3`G and G in the loop) alkylation reaction with 20 mol% copper(I) and EDA (Table 4, Entry 3). (D) hair-pin (5`G) alkylation reaction with 20 mol% copper(I) and DAA **23** (Table 4, Entry 4). (E) hair-pin (5`G) alkylation reaction with 1 eq. copper(I) and DAA **23** (Table 4, Entry 5).

In order to improve the conversion and isolate a sufficient amount of product, the reaction conditions were slightly changed. Attempts with stoichiometric amounts of copper(I) enhanced the conversion of the hairpin (Table 4, Entry 4). However, a similar increasing reactivity was also observed for the hairpin that was unreactive with catalytic amounts of copper(I) (Table 4, Entry 2). It is likely that the higher amount of copper induces single-strand formation in the hairpins. We assume that the coordination of copper to the double strand can interfere with Watson-Crick base pairing resulting in duplex opening and accessible guanines for alkylation.

DAA **23** was also investigated as carbene source in hairpin alkylation. Treatment of the hair-pin bearing 10 base pairs in the duplex and a 5` guanine overhang with the copper(I) carbene derived DAA **23** yielded a single monoalkylation product (Table 4, Entry 5 Figure 30D). Analysis of the second new peak by MALDI TOF did not show a reasonable mass that could be correlated to any expected alkylation product. The presence of stoichiometric instead of catalytic amounts of copper(I) showed again a different alkylation profile (Figure 30E). This supports the hypothesis that single-strands are formed based on copper concentration.

Nevertheless, selectively targeting single-stranded guanines over double stranded is an important characteristic and could be used for site selective structure depending modification of DNA.

As previously discussed, guanines in double-stranded DNAs are not targeted as long the copper concentration is catalytic and does not interfere with Watson-Crick base pairing. A single mismatch base pair such as G:A in a DNA duplex could free the guanine from the duplex and make it accessible for alkylation. A DNA 16-mer duplex **36** was synthesized bearing a mismatch base-pair G:A at position five and the reactivity in alkylation reaction was investigated (Scheme 11). Unfortunately, the mismatch DNA **36**was unreactive in the presence of copper(I) carbenoids derived from DAAs **23** and **24**. It is likely that only one mismatch pair is not enough to make the guanine accessible for modification.



Scheme 11. Alkylation of a double-stranded DNA bearing one G:A mismatch in the duplex. (a) NaCl, MES buffer pH 6, 80°C for 10 min then to 25°C over 3 h. (b) DAA **23** or **24**, $CuSO_4$, sodium ascorbate, 25°C, 24 h.

2.5.5 Biochemical studies with O^6-G alkylated substrates

The developed post-synthetic alkylation method allows the ready synthesis of O^6-G derivatives. In the next step we investigated two open questions in the biochemistry of these adducts: the repair of O^6-G modification by the alkylguanine transferase (AGT) and the incorporation of O^6-G modified triphosphates during DNA replication.

AGT proteins are found in all domains of life and play an important role in repairing genomic O^6-G damage.[153-155] This protein is able to repair O^6-(alkyl)-G adducts from single and preferentially from double-stranded DNA. Removing of the alkyl group from the O^6 position can be essential to prevent mutagenesis. DNA replication relies on proper Watson-Crick base pairing of the nucleobases. For example, during replication T can be misincorporated opposite O^6-G adducts due to altered proton donor and acceptor properties of the modified nucleobase (Figure 31A). This misincorporation can be highly mutagenic if not repaired because after a second cycle of replication a permanent G to A transition mutation takes place. AGT can prevent this point mutation by nucleophilic attack of the active site cysteine residue that removes the modification in a S_N2 manner and restores the native guanine. This process is irreversible and therefore hAGT is a suicide protein that is degraded from the cell after it performs repair.

Diazotized peptides produced by bacteria may be a possible alkylation source for nucleic acid modification.[156-158] Therefore, we investigated whether the amide functionality can also be repaired by the alkyl guanine transferase protein hAGT (Scheme 12A-C). Treatment of the O^6-(alkyl)-G adduct **39** (Table 3, Entry 6) with the repair protein at 37°C showed the desired restoration of the natural G. The slow removal of the modification by hAGT was not surprising because reduced reaction rates are expected for short ssDNAs, due to lower binding constants to the protein.[159] Nevertheless, guanine in the ssDNA 9-mer **39** was successfully repaired and the carboxymethyl amide functionality can be added to the substrate list of adducts repaired by hAGT.

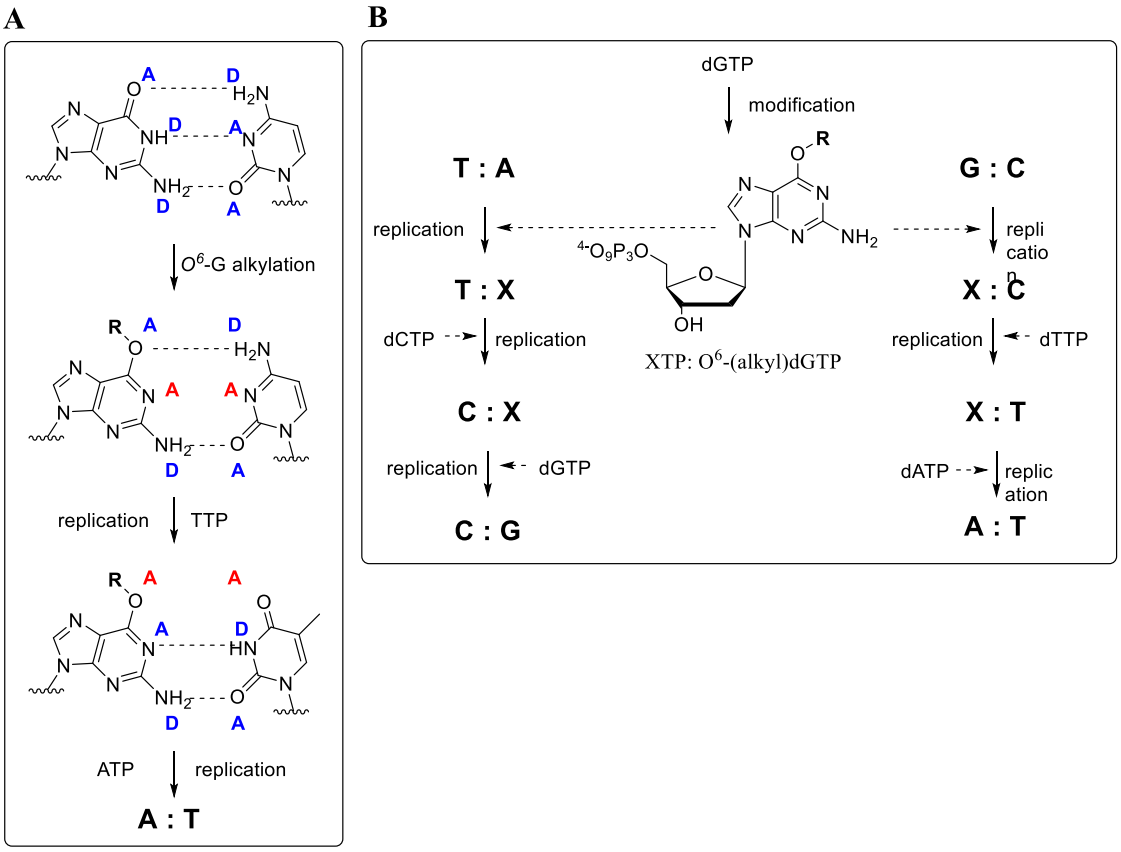
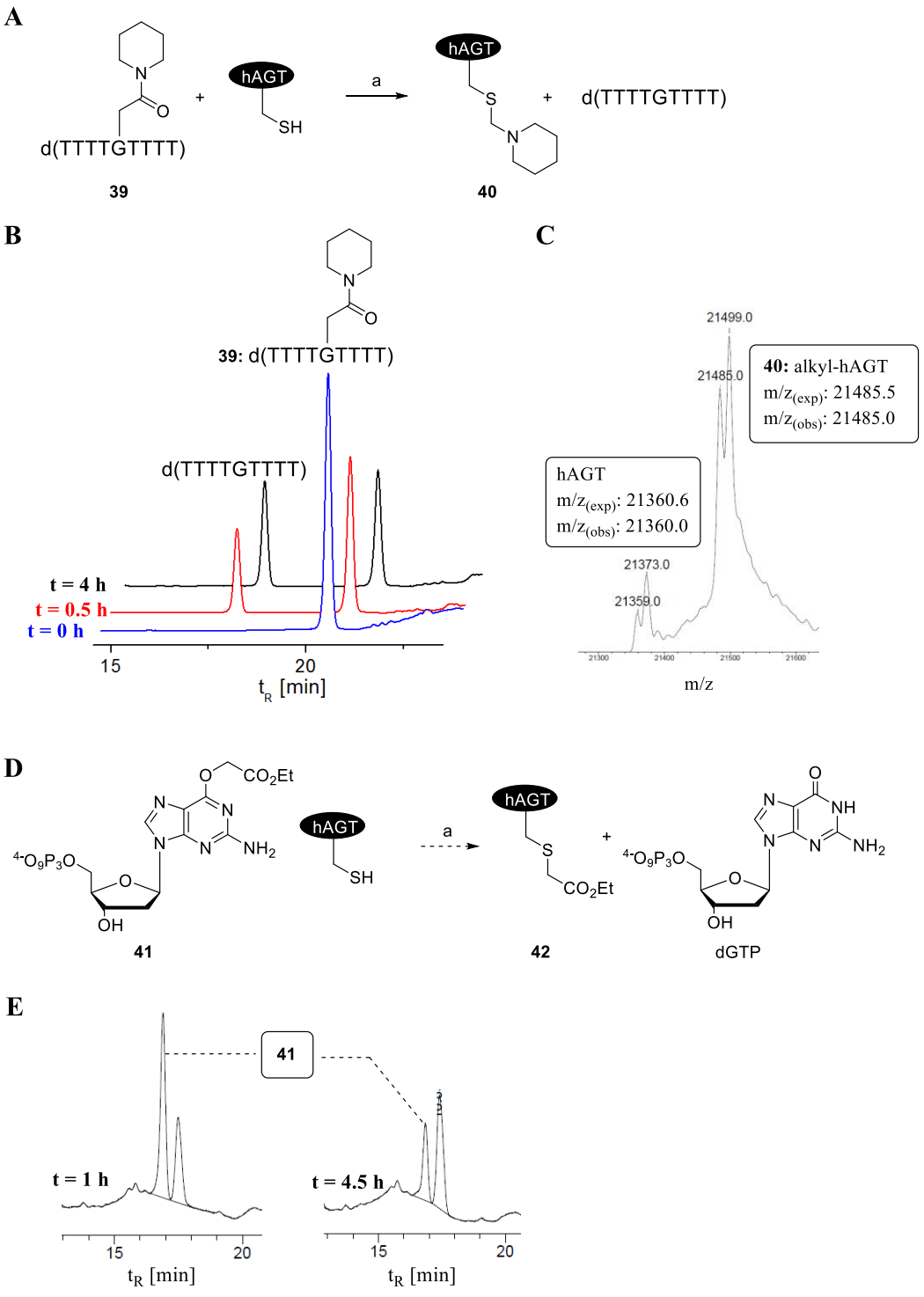


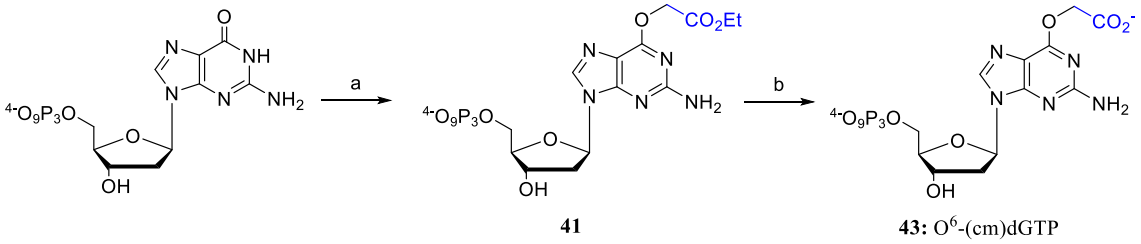
Figure 31. Two possible pathways for pyrimidine transition mutations. (A) Guanine is modified (R = methyl, ethyl carboxymethyl etc) at the O^6-position in the DNA template and after the 2nd cycle of replication a G to A transition mutation may occur. D = proton Donor; A = proton acceptor. (B) The guanine triphosphate dGTP is modified (R = methyl, ethyl carboxymethyl etc) at the O^6-position and wrong misincorporation opposite C or T led to G:A or A:G transition mutation after the third cycle of replication.

Another pathway for pyrimidine transition at modified G is presented in Figure 31B. Instead of a damaged template the natural dGTP is modified at the O^6-position and may be either incorporated opposite cytidine (right route) or opposite thymine (left route). Both routes lead to an A:G or G:A transition mutation after the third cycle of replication. To investigate if also a repair mechanism for damaged dGTP exists, we tested the O^6-dGTP substrate **41** as a potential substrate for hAGT. Treatment of the modified triphosphate **41** with hAGT in aqueous buffer showed 70% conversion of the starting material after 4.5 hours at 37°C (Scheme 12 D-E), whereas the control reaction without the enzyme showed 27 % decomposition of compound **41** after 6.5 hours. Unfortunately, HR-MS analysis of hAGT from the reaction mixture did not show the target mass for the modified protein **42**. With these results in hand the question whether hAGT can repair O^6-modified dGTPs remains unanswered.



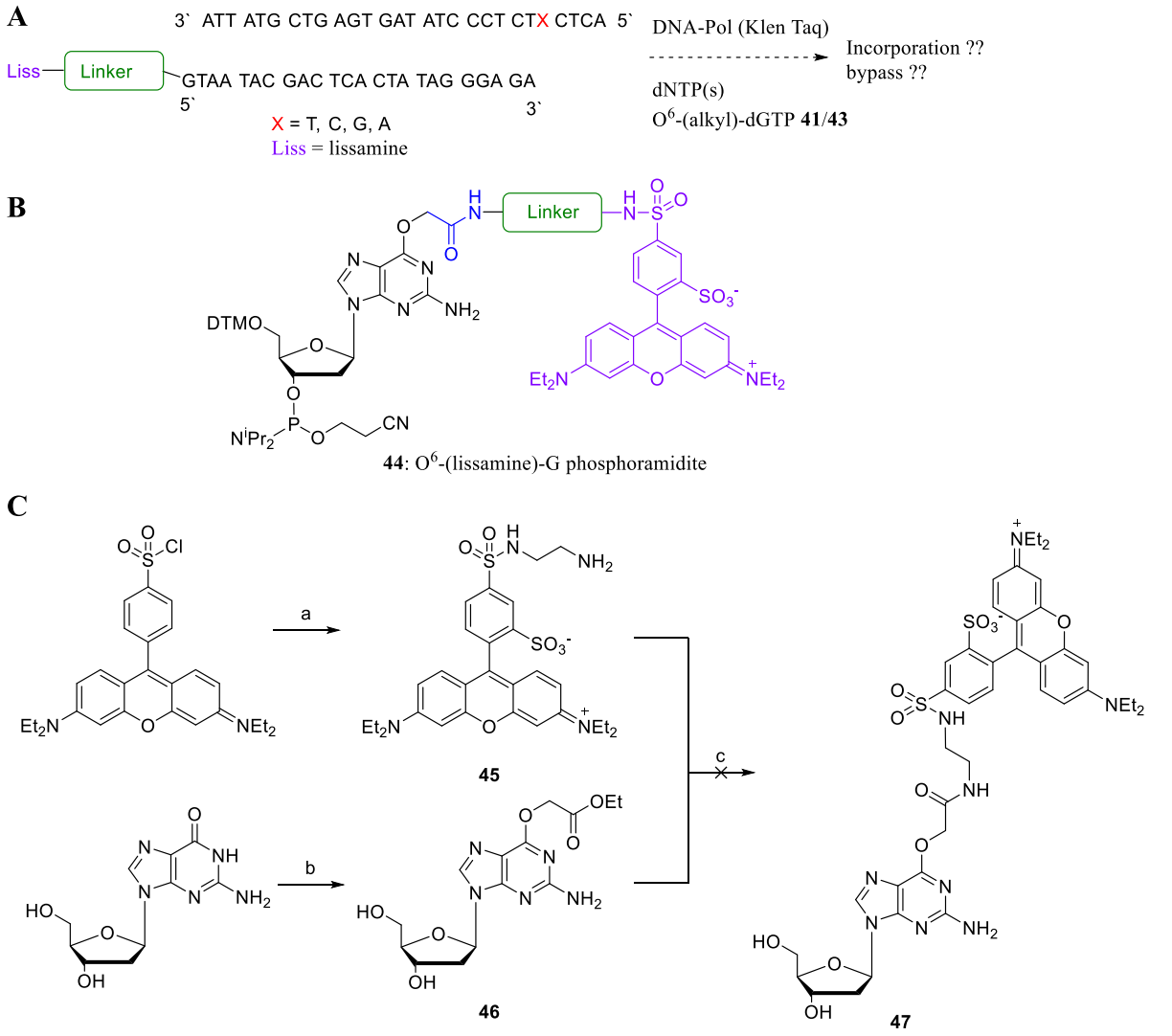
Scheme 12. (A) De-alkylation of O^6-G modified ssDNA 9-mer **39** with hAGT. (B) HPLC analysis of the reaction mixture after 0 (blue), 0.5 (red) and 4 hours (black). (C) High resolution ESI-MS from hAGT after 4 h at 37°C. (D) De-alkylation of O^6-G modified dGTP **41** with hAGT. (E) HPLC analysis of de-alkylation reaction from after 1 and 4.5 hours. (a) 50 mM TrisHCl buffer pH 7.6, 250 mM NaCl, 5 mM DTT, 37°C.

We tested than the impact of O^6-alkylation on nucleotide incorporation during DNA synthesis and characterized O^6-modified dGTPs as polymerase substrates. We synthesized the O^6-modified dGTPs **41** and **43** with our copper(I) carbenoid methodology. Commercial dGTP was selectively alkylated at the O^6-position with EDA as carbene source. Hydrolysis of the ethylester with lithium hydroxide yielded the carboxymethyl (cm) modified dGTP **43** in two steps. Our methodology provides the alkylated O^6-dGTP motifs **41** and **43** very efficiently without the use of any protecting group or the need of a phosphorylation step. The O^6-G modified nucleosides and triphosphates reported in the literature were synthesized in a laborious four to seven steps.[160-163] Especially, for the synthesis of O^6-alkylated dGTPs the triphosphate is always introduced at the end with strong electrophilic phosphorous reagents resulting in low yield due to side product formation.[98]

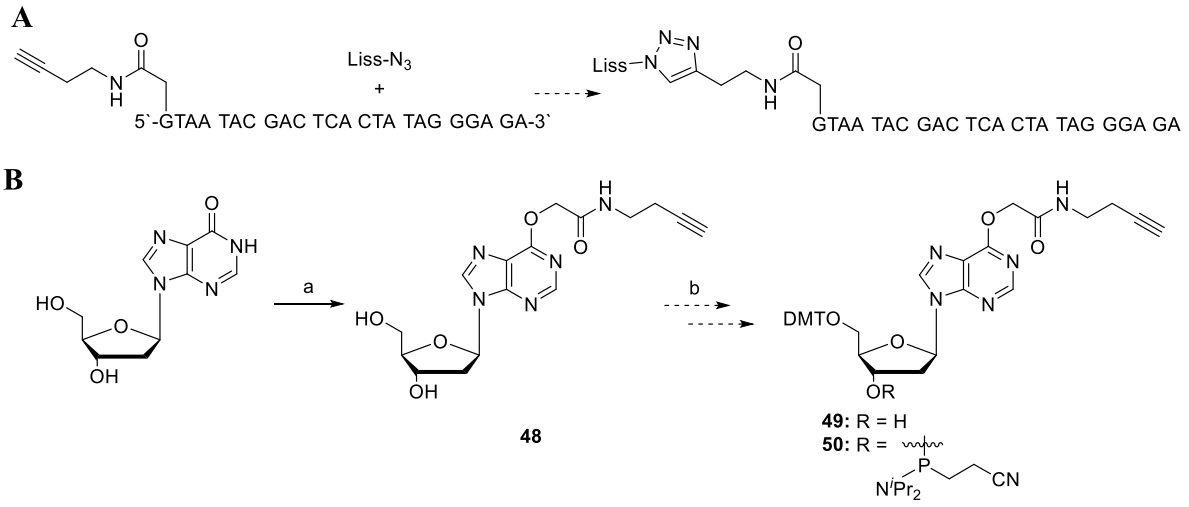


Scheme 13. Synthesis of dO^6-(alkyl)-GTPs **41** and **43** with Cu(I) carbenoid chemistry. (a) EDA, $CuSO_4$, sodium ascorbate, MES buffer pH 6, H_2O, rt, 30 min (65 %). (b) LiOH, H_2O, rt. 1h (57 %).

To study the incorporation of the O^6-G adducts **41** and **43** in DNA synthesis, we planned single nucleotide primer extension experiments (PEX) using DNA polymerases (Scheme 14A). The primer for PEX is usually labeled with a fluorophore or radiolabel. Synthesis of the primer was planned with OSPS. Coupling of an O^6-G modified phosphoramidite **44** bearing the fluorophore, lissamine would introduce the desired label to the primer (Scheme 14B). Commercial Sulforhodamine B acid chloride was treated with ethylenediamine to yield the corresponding sulfonamide **45** (Scheme 14C). Commercial guanosine was alkylated with copper(I) carbene derived from EDA to yield the O^6-G adduct **46**. Unfortunately, subsequent treatment with sulfonamide **45** did not yield the desired aminolysis product **47** (Scheme 14C). Several conditions (different solvents and bases) were investigated, however, the two derivatives **45** and **46** could not be connected together. The strategy was changed to an alkyne functionalized inosine phosphoramidite **50** that should be introduced in the OSPS cycle and post-labeled with a lissamine azide derivative *via* click reaction (Scheme 15A). Commercial inosine was efficiently alkyne functionalized at the O^6-position to yield **48** (Scheme 15B). However, dimethoxytrityl (DMT) protection of the O^6-modified inosine derivative **48** yielded only 4 % of the DMT protected product **49**.



Scheme 14. (A) Primer extension experiments of O^6-(alkyl)-dGTPs using DNA polymerases. (B) Structure of O^6-(lissamine)-G phosphoramidite **44**. (C) Synthetic strategy for O^6 lissamine functionalized guanosine **47** via aminolysis of lissamine sulfonamide **45** with O^6-G adduct **46**. (a) ethylenediamine, DMAP, NEt_3, CH_2Cl_2, rt, on (55 %). (b) EDA, $CuSO_4$, sodium ascorbate, MES buffer pH 6, H_2O, 4 % (v/v) DMSO, rt, 5h (56 %). (c) Na_2CO_3, MeOH, reflux, 1.5 d or Na_2CO_3 in H_2O/MeCN, rt, 2d or in H_2O/MeCN, rt, 2d.



Scheme 15. (A) Post-functionalization of DNA *via* click reaction with Lissamine azide derivative. (B) Synthetic strategy for O^6 alkyne functionalized inosine phosphoramidite **50**. (a) **25**, $CuSO_4$, sodium ascorbate, MES buffer pH 6, H_2O, 4 % (v/v) DMSO, rt, 4 h (72 %). (b) DMT-Cl, pyridine, rt, on (4-7 %).

In collaboration with the group of Prof. Sturla we could finally investigate the incorporation of O^6-modified dGTPs **41** and **43** in primer (5`-end [y-$^{32}$P] radiolabeled) extension catalyzed by the KlenTaq mutant: *KTqM747K*. This enzyme is reported to bypass several modifications.[164-166] Several unmodified 28-mer ssDNA templates (5`-ATT ATG CTG AGT GAT ATC CCT CTX CTC A) with X = G, A, T or C were annealed with the complementary 5`-end [y-$^{32}$P] radiolabeled 23-mer primer (5`-AGA GGG ATA TCA CTC AGC ATA AT) and treated with the O^6-(alkyl)-dGTPs **41** and **43** (Figure 32A). The primer extension reaction were quenched after 10 min at 55°C and separated by denaturing polyacrylamide gel electrophoreses and visualized with a phosphor imager (Figure 32B). Both O^6-(alkyl)-dGTPs **41** and **43** were efficiently incorporated opposite C but also to a smaller extent opposite T. After successful incorporation of the O^6-(alkyl)-dGTPs opposite T the polymerase stalled, however, the O^6-(alkyl)-G:C mismatch was tolerated and a second extension occurred. These experiments demonstrated as a proof-of-concept that O^6-(cm)-G damaged nucleotides are substrates for DNA replication and can be incorporated in DNA synthesis.

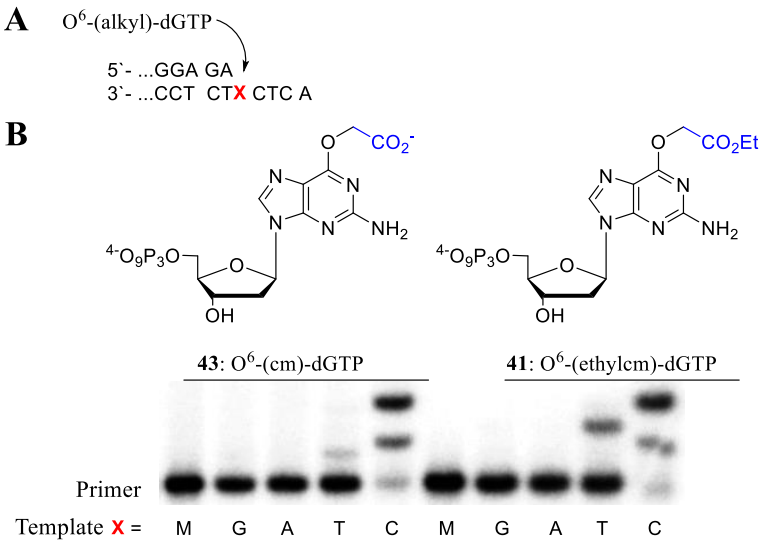
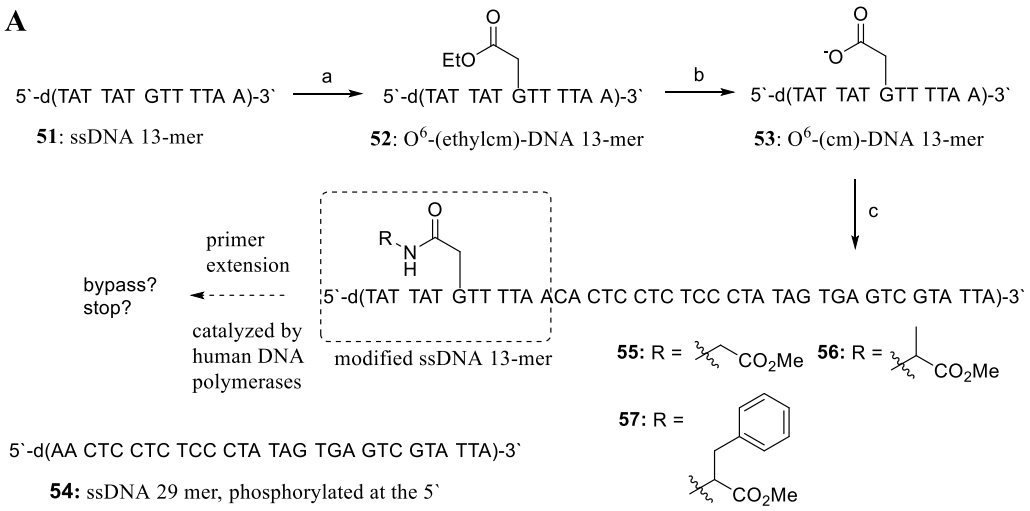
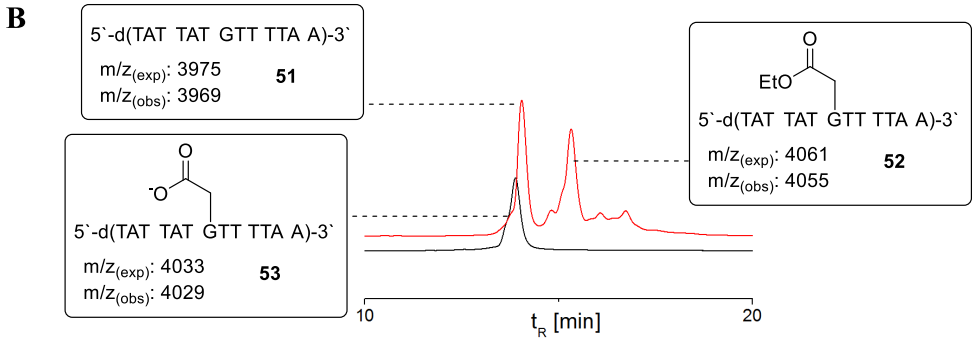


Figure 32. (A) single nucleotide primer extension experiment with O^6-(alkyl)-dGTPs **41** and **43** opposite DNA templates with X = G, A, T, C catalyzed by *KTqM747K*. Reaction were run at 55°C for 10 min and contained 50 µM O^6-(alkyl)-dGTPs **41** and **43**. (B) PAGE analysis of the primer extension experiments with M corresponding to blank primer.

As a next step, a full study about how different human polymerases deal with a variety of carboxymethyl damaged DNA during replication was planned in collaboration with the group of Prof. Sturla. Post-synthetic modification of ssDNA 13-mer **51** using our copper(I) carbene chemistry should yield the O^6-G adduct and subsequent ligation elongates the DNA strand to a 42-mer. Finally the different O^6-G modifications are introduced by amide coupling (Scheme 16).

Treatment of commercially available ssDNA 13-mer **51** with copper(I) carbene derived from EDA and subsequent hydrolysis of the ethyl ester with lithium hydroxide yielded the O^6-cm-DNA 13 mer **53**. The ligation with the commercial 5`-phosphorylated DNA 29-mer **54** was done by Dr. Zhanybekova. Unfortunately, the ligation was pretty inefficient. The reason for this inefficiency could be the low quality of the commercial 5`-phosphorylated DNA 29-mer **54**: around 50 % corresponded to the non-phosphorylated species. The subsequent amide couplings were investigated although purification of the ligation mixture was unsuccessful. Treatment of the ligation mixture with three different amino acids glycine (Gly), alanine (Ala) and phenylalanine (Phe) in the presence of a large excess of coupling reagents, however, yielded only the modified ssDNA 13-mer (highlighted with the dashed box) instead of the modified 42-mer. These results showed that the amide coupling was successful; however the ligation step was too inefficient and has to be improved. Therefore, the question whether these amide functionalities are substrates for human DNA polymerase is still not answered.





Scheme 16. (A) Synthesis of O^6-G modified ssDNA 42-mer via ligation of O^6-(cm) DNA 13-mer 53 with commercial 5`phosphorylated 29-mer £130 and subsequent amide coupling with different AA such as Gly, Ala and Phe. (B) HPLC traces of the alkylation reaction mixture after 1 h (red) and hydrolysis of the ethyl ester after 1h (black). (a) EDA, $CuSO_4$, sodium ascorbate, Mes buffer pH 6, H_2O, 25°C. (b) LiOH, H_2O, 25°C, 1 h. (c) 1000 eq. EDC, 1000 eq, HOBt, 1000 eq amino acid (AA), 1000 eq. DIPEA, 25°C, on.

2.6 Conclusion and future directions

A new way to create O^6-G adducts post-synthetically from nucleic acids of different size starting from mono- or triphosphates to longer oligonucleotides was discovered. The method allows an efficient and clean modification of the O^6-G position using $CuSO_4$ (reduced with sodium ascorbate) and α-diazoacetates or acetamides. The ability to synthesize precisely defined O^6-G adducts in high yields allowed studies in their biochemistry. This we presented with the first synthesis of O^6-(cm)-dGTP and the first analysis of its effect on DNA replication.

The developed method, the copper(I) catalyzed O^6-G alkylation, represent an easy synthesis set-up including readily available material. The stability of α-diazoacetates is high enough that some are commercially available, otherwise an easy two step protocol facilitates the synthesis of nearly any α-diazoester. To study O^6-G adducts in biology or biotechnology a ready synthesis of these compounds is needed. The method we described here is the simplest reported so far.

Future work on the copper(I) carbenoid system could be continued in the following directions:

(a) Modification of longer oligonucleotides bearing several reactive guanines, such as tRNAs.
   - Reverse transcription of the O^6-G adducts with DNA polymerase enzymes that do not tolerate the modification site and result in a replication stop. Length directed separation of the DNA fragments using electrophoresis would provide an insight about the modification site.
   - Digestion to shorter oligonucleotides or nucleotides using enzymes without affecting the alkylation site. Comparing the digestion profile with the unmodified oligonucleotide

(b) Synthesis of other biologically relevant O^6-G adducts
   - Test the substrates as either templates or O^6-modified dNTPs in primer extension experiments using human DNA polymerases
   - Test the O^6-G adducts as substrates for de-alkylation reactions with hAGT

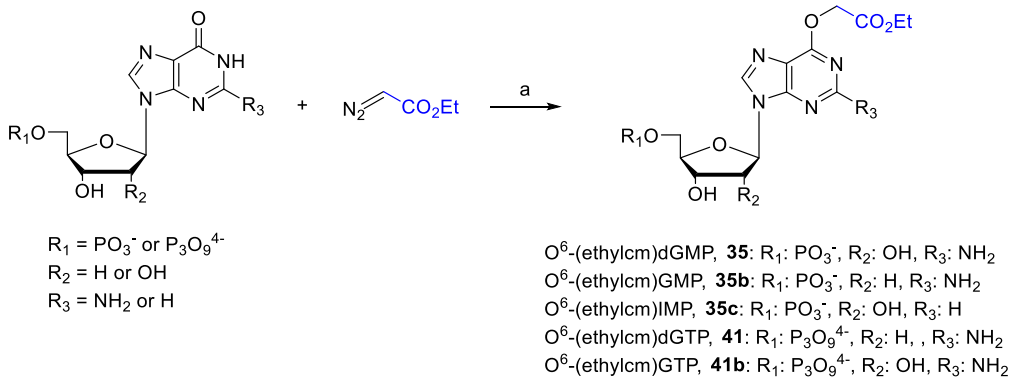
(c) Studying the catalyst scope with the use of enzymes that contain copper(I).

2.7 Experimental

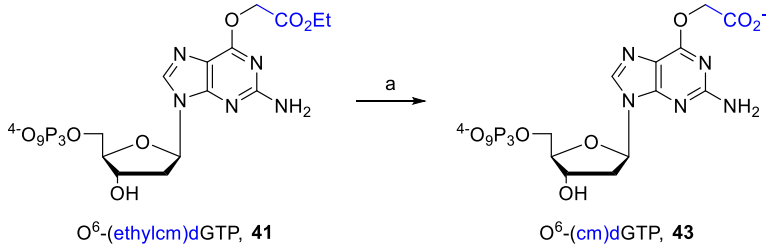
2.7.1 General Experimental Information and Synthetic Schemes

All reagents and solvents used were of analytical grade. Buffers were prepared with ultrapure water. All chemicals were purchased from Sigma-Aldrich, Alfa Aesar or Acros and used as received. Unlabeled dNTPs were obtained from Invitrogen and [γ-^{32}P]ATP was purchased from PerkinElmer Life Sciences. KTqM747K DNA polymerase was purchased from myPOLS Biotec. Oligonucleotides used in primer extension experiments were obtained from VbC Biotech. Solid-phase oligonucleotide synthesis was carried out on an Expedite 8909 nucleic acid synthesis system (PerCeptiveBiosystems) on 1µmol CPG columns using standard phosphoramidite chemistry with 0.3 M 5-benzylthio-1-H-tetrazole as activator. Purification was carried out on a preparative Shimadzu UFLC system with a Gemini-NX 5 µm C18 21.2 x 250 mm (Phenomenex) column using 100 mM triethylammonium acetate (TEAA pH 7.25)/acetonitrile (MeCN) gradients as mobile phase. Elution was carried out at a flow rate of 20 mL/min monitored at 254 nm using **Method A:** 0-15 % MeCN in 30 min, 15-95 % MeCN in 10 min, 95 % MeCN for 5 min for short ssDNA and **Method B**: 0-12% MeCN in 25 min, 12-16 % MeCN in 10 min, 16-95 % MeCN in 5 min, 95 % MeCN for 5 min for longer ssDNA. Alkylation and de-alkylation reactions were analyzed on an analytical Shimadzu UFLC system or Agilent 1100 LC system with an Eclipse XDB-C8 4.6 x 150 mm (Agilent) column using 100 mM TEAA/MeCN gradient. Elution was carried out at a flow rate of 1 mL/min using **Method C**: 0-16 % MeCN in 18 min, 16-80 % MeCN in 5 min, 80 % MeCN for 3 min with peak detection at 254 nm. Purification was carried out on the preperative Shimadzu UFLC system as described above using **Method D**: 0-16 % MeCN in 20 min, 16-20 % MeCN in 5 min, 20-95 % in 1 min, 95 % MeCN for 5 min. Aqueous product fractions were freeze dried on a Christ Alpha 2-4 LDplus flask lyophilizer at 0.1 mbar. ^{1}H, ^{13}C and ^{31}P-NMR, HMQC and HMBC spectra were acquired on a BrukerAvance (400, 500 or 600 MHz proton frequency) spectrometer at 298.15 K. Chemical shifts relative to TMS were referenced to the solvent's residual peak and are reported in ppm. ESI MS spectra were measured on a Bruker Esquire3000plus mass spectrometer by direct injection in positive polarity of the ion trap detector. Polyacrylamide gel electrophoresis (PAGE): PAGE was done with 12.5 % gel (29:1 w/w acrylamide/bisacrylamide) in 1×TBE (89 mM Tris-borate, 2 mM EDTA) with Orange G, Bromophenol blue and XylenCynol FF as tracer dyes. The samples were prepared by mixing 5 µL of sample with 2.5 µL 4×TBE in 40 % (v/v) glycerol. Five microliters from the resulting mixtures were then applied on the gel without further pre-treatment. Trace dye mixtures were applied on separate wells. The gels were run at 30 V/cm until the Orange G migrates to approximately 1 cm from the end of the gel. The gels were then washed briefly with deionized water and soaked with Coomassie Brilliant Blue R-250 for 1 h, destained and visualized using Bio-RadChemiDoc MP system equipped with Image Lab 5.0 software.

Synthetic Schemes:



Scheme 17. Synthesis of *O*^{6} modified GMP 1, dGMP 2, IMP 3, dGTP 4 and GTP 5 using Cu(I) carbene chemistry. a) 20 mol% Cu(I), MES buffer pH 6, H_2O, 25°C, 1 h, 1: 92 %, 2: 86 %, 3: 81 %, 4: 73 %, 5: 65 %.



Scheme 18. Synthesis of O^6-(cm)dGTP **43**. a) LiOH, 1 h, 25°C, 57 %.

2.7.2 Chemical synthesis

General Procedure: Synthesis of Diazoacetamides (DAAs)

A: To a solution of bromacetyl bromide (BrAcBr) in dry CH_2Cl_2, amine (freshly filtered over neutral aluminium oxide) was added at -60°C. The cooling bath was removed and the reaction mixture was stirred for an additional hour. Water was added and the reaction mixture was extracted with EtOAc. The solvent was removed under reduced pressure and the residue purified by flash chromatography on Si_{60} in cyclohexane and EtOAc.

B: To a solution of bromoacetamide derivative in THF, *N,N*-ditosylhydrazine (TsNHNHTs) was added. The reaction mixture was cooled to 0°C, DBU was added and the mixture was stirred for 20 min at room temperature. Saturated $NaHCO_3$ was added while precipitation was observed. The mixture was centrifuged

at 4400 g for 2 min. The supernatant was extracted with Et_2O (3 x 10 mL). The combined organic layers were dried over Na_2SO_4, filtered and the residue purified by flash chromatography on Si_{60} in CH_2Cl_2 and MeOH (prior to purification, the column was neutralized with two column volumes CH_2Cl_2 containing 5 % NEt_3).

2-bromo-*N,N*-diisopropylacetamide (20). According to the general procedure A BrAcBr (100 µL, 1.15 mmol) and *N,N*-diisopropylamine (324 µL, 2.30 mmol) were reacted to yield **20** as colorless liquid (133 mg, 52 %). The analytical data were in agreement with literature.[2] ^{1}H-NMR (500 MHz, $CDCl_3$) δ/ppm: 3.96 (dt, *J* = 13.2, 6.7 Hz, 1H), 3.82 (s, 2H), 3.43 (p, *J* = 7.2, 6.6 Hz, 1H), 1.38 (d, *J* = 6.9 Hz, 6H), 1.25 (d, *J* = 6.7 Hz, 6H). ^{13}C-NMR (126 MHz, $CDCl_3$) δ/ppm: 165.57, 50.62, 46.47, 28.71, 20.78, 20.20.

2-bromo-1-(piperidin-1-yl)ethan-1-one (21). According to the general procedure A, BrAcBr (100 µL, 1.15 mmol) and piperidine (228 µL, 2.30 mmol) were reacted to yield **21** as colorless liquid (74.5 mg, 31 %). The analytical data were in agreement with the literature.[3] ^{1}H-NMR (500 MHz, $CDCl_3$) δ/ppm: 3.86 (s, 2H), 3.58 – 3.52 (m, 2H), 3.50 – 3.40 (m, 2H), 1.65 (p, *J* = 2.8 Hz, 4H), 1.56 (t, *J* = 5.6 Hz, 2H). ^{13}C-NMR (126 MHz, $CDCl_3$) δ/ppm: 166.15, 48.05, 43.39, 26.31, 26.28, 25.49, 24.39.

2-bromo-N-(but-3-yn-1-yl)acetamide (22). According to the general procedure A, BrAcBr (252 µL, 2.90 mmol) and 1-amino-3-butyne (250 µL, 2.90 mmol) were reacted to yield **22** as colorless liquid (415 mg, 75 %). The analytical data were in agreement with the literature.[3] ^{1}H-NMR (250 MHz, $CDCl_3$) δ 6.84 (s, 1H), 3.91 (s, 1H), 3.46 (q, J = 6.3 Hz, 1H), 2.45 (td, J = 6.4, 2.6 Hz, 1H), 2.05 (t, J = 2.6 Hz, 1H).

2-diazo-*N,N*-diisopropylacetamide (23). According to the general procedure B, **20** (123.3 mg, 555 µmol), TsNHNHTs (378 mg, 1.11 mmol) and DBU (414 µL, 2.78 mmol) were reacted to yield **23** as yellow liquid (56.0 mg, 60 %). ^{1}H-NMR (500 MHz, $CDCl_3$) δ/ppm: 4.94 (s, 1H), 3.65 (p, *J* = 7.2 Hz, 2H), 1.28 (s, 12H). ^{13}C-NMR (101 MHz, $CDCl_3$) δ/ppm: 164.68, 47.83, 47.10, 21.25. HRMS (ESI): $C_8H_{15}N_3NaO^+$ *calcd.*: 192.1107, found: 192.1110.

2-diazo-1-(piperidin-1-yl)ethan-1-one (24). According to the general procedure B, **21** (66.3 mg, 322 µmol), TsNHNHTs (219 mg, 643 µmol) and DBU (240 µL, 1.61 mmol) were reacted to yield **24** as yellow waxy solid (29.5 mg, 60%). ^{1}H-NMR (500 MHz, $CDCl_3$) δ/ppm 1H NMR (400 MHz, CDCl3) δ/ppm: 4.99 (s_{br}, 2H), 3.35 (s_{br}, 4H), 1.67 – 1.62 (m, 2H), 1.58 – 1.51 (m, 4H). ^{13}C-NMR (126 MHz, $CDCl_3$) δ/ppm: 164.34, 59.87, 48.39, 46.31,43.40,25.88,24.47. HRMS (ESI): $C_7H_{12}N_3O^+$ *calcd.*: 154.0975, *found*: 154.0976.

2-diazo-1-(piperidin-1-yl)ethan-1-one (25). According to the general procedure B, **22** (414 mg, 2.18 mmol), TsNHNHTs (1.48 g, 4.36 mmol) and DBU (1.63 mL, 10.9 mmol) were reacted to yield **25** as yellow waxy solid (102 mg, 34%). ^{1}H-NMR (500 MHz, $CDCl_3$) δ/ppm 1H NMR (400 MHz, CDCl3) δ/ppm: 5.42 (s_{br}, 1H), 4.75 (s, 1H), 3.45 (q, J = 6.2 Hz, 2H), 2.42 (td, J = 6.3, 2.6 Hz, 2H), 2.01 (s, 1H). ^{13}C-NMR (126 MHz, $CDCl_3$) δ/ppm: 165.84, 81.60, 70.19, 47.34, 38.62, 19.95. LRMS (ESI): $C_6H_7N_3NaO^+$ *calcd.*: 160.0, *found*: 159.8.

O^6-(ethylcm)I (31) (*via* pathway A, Figure 28). To a solution of $CuSO_4$ (27.9 mg, 112 µmol) in H_2O (94 mL) and MES buffer (10.8 mL), inosine (150 mg, 559 µmol, dissolved in 7 mL DMSO), EDA (676 µL, 5.59 mmol) and ascorbate (111 mg, 559 µmol) were added and the reaction mixture stirred at room temperature. Analysis of the reaction mixture after 5 h via RP-HPLC showed 66% conversion of inosine. The reaction was stopped, lyophilized and taken up in CH_2Cl_2 while a precipitate was observed. The resulting suspension was centrifuged at 4400 g for 5 min followed by lyophilizing of the supernatant. The residue was purified by flash chromatography on Si 60 in CH_2Cl_2 and MeOH. Residual MeOH sticking to the product was removed by adding H_2O and lyophilization to obtain **31** as white solid (119 mg, 60 %). ^{1}H-NMR (400 MHz, MeOD) δ/ppm: 8.59 (s, 1H), 8.48 (s, 1H), 6.10 (d, *J* = 5.8 Hz, 1H), 5.19 (s, 2H), 4.74 (t, *J* = 5.4 Hz, 1H), 4.36 (dd, *J* = 5.1, 3.3 Hz, 1H), 4.22 (q, *J* = 7.1 Hz, 2H), 4.17 (q, *J* = 3.1 Hz, 1H), 3.91 – 3.3.88 (m, 1H), 3.78 – 3.753 (m, 1H), 1.26 (t, *J* = 7.1 Hz, 3H). ^{13}C-NMR (101 MHz, MeOD) δ/ppm: 169.76, 160.99, 153.06, 152.70, 144.35, 122.65, 91.01, 87.79, 75.73, 72.29, 64.05, 63.12, 62.44, 14.40. HRMS (ESI): $C_{14}H_{18}N_4NaO_7^+$ *calcd.*: 377.1068, *found*: 377.1073.

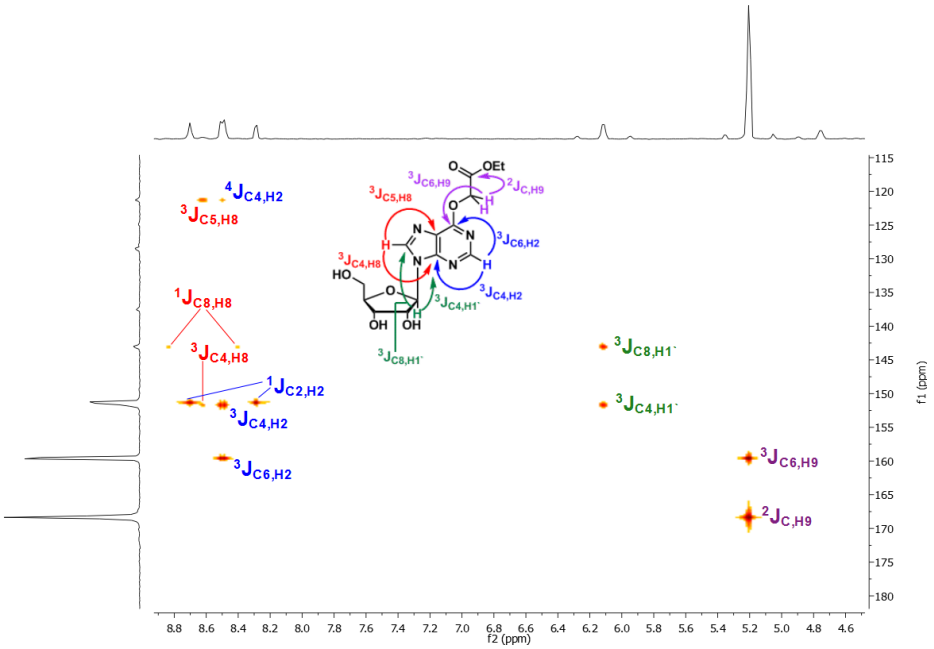
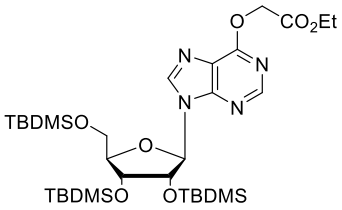


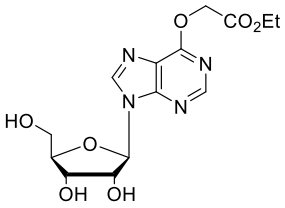
Figure 33. Important HMBC correlations of O^6-ethylcmI **31** synthesized *via* pathway A (Figure 28).

2`, 3`, 5` (TBDMS) 6-chloropurine riboside (32). To a solution of 6-chloropurine riboside (108 mg, 376 µmol) in dry DMF (1.5 mL), imidazole (256 mg, 3.76 mmol) and TBDMS-Cl (283 mg, 1.88 mmol) were added and the reaction stirred at room temperature for 22 h. The reaction was diluted with CH_2Cl_2 (5 mL) and H_2O (5 mL). The water layer was extensively extracted CH_2Cl_2 (5 x 5 mL) and the combined organics were dried over Na_2SO_4, filtered and concentrated under reduced pressure. The residue was purified by flash chromatography on Si 60 in cyclohexane/EtOAc to obtain **32** as white solid (182 mg, 289 µmol, 77 %). ^{1}H-NMR (250 MHz, $CDCl_3$) δ/ppm: 8.8 (s, 1H), 8.51 (s, 1H), 6.08 (d, *J* = 4.6 Hz, 1H), 4.54 (t, *J* = 4.5 Hz, 1H), 4.27 (t, *J* = 4.0 Hz, 2H), 4.12 (dd, *J* = 6.1, 3.2 Hz, 1H), 3.98 (dd, *J* = 11.5, 3.5 Hz, 1H), 3.76 (dt, *J* = 11.6, 2.2 Hz, 1H), 0.91 (s, 9H), 0.88 (s, 9H), 0.74 (s, 9H), 0.08 (m, 12H), -0.07 (s, 3H), -0.28 (s, 3H). ^{13}C-NMR (101 MHz, $CDCl_3$) δ/ppm: 152.08, 151.65, 151.14, 144.21, 132.23, 88.73, 85.87, 76.54, 71.96, 62.54, 26.26, 25.97, 25,76, 18.71, 18.22, 17.97, -4.24, -4.54,-4.91, -5.19. HRMS (ESI): $C_{28}H_{53}ClN_4NaO_4Si_3^+$ *calcd.*: 651.2955, *found*: 651.2956.

2`, 3`, 5` (TBDMS) O^6-(ethylcm)I (33). NaH (15.9 mg, 397 µmol, 2.50 eq.) was placed in a pressure tube and THF (1.5 mL) was added. Then ethyl glycolate (82.7 mg, 794 µmol) was added and the reaction mixture stirred for 30 min at room temperature. Then chloropurine riboside derivative **32** (100 mg, 159 µmol) was added and the reaction heated to reflux for 4 h. The solvent was removed under reduced pressure and the residue was purified by flash chromatography on Si 60 cyclohexane/EtOAc to obtain **33** as white solid (53.0 mg, 76.0 µmol, 48 %). ^{1}H-NMR (500 MHz, $CDCl_3$) δ/ppm: 8.47 (s, 1H), 8.39 (s, 1H), 6.08 (d, *J* = 4.7 Hz, 1H), 5.12 (s, 1H), 5.10 (s, 1H), 4.57 (t, *J* = 4.4 Hz, 1H), 4.32 (t, *J* = 4.2 Hz, 1H), 4.23 (q, *J* = 7.1 Hz, 2H), 4.14 (td, *J* = 3.9, 2.5 Hz, 1H), 4.04 (dd, *J* = 11.4, 3.7 Hz, 1H), 3.79 (dd, *J* = 11.5, 2.6 Hz, 1H), 1.25 (m, 2H), 0.95 (s, 9H), 0.92 (s, 9H), 0.80 (s, 9H), 0.14 (s, 3H), 0.13 (s, 3H), 0.10 (s, 3H), 0.09 (s, 3H), -0.03 (s, 3H), -0.19 (s, 3H). ^{13}C-NMR (101 MHz, $CDCl_3$) δ/ppm: 168.19, 159.52, 152.25, 151.8, 141.64, 121.47, 88.80, 85.41, 76.39, 71.69, 63.14, 62.42, 61.45, 26.27, 25.99, 25.82, 18.72, 18.22, 18.00, 14.25, -4.21, -4.57, -4.60, -4.84, -5.18, -5.25. HRMS (ESI): $C_{32}H_{60}N_4NaO_7Si_3^+$ *calcd.*: 719.3662, *found*: 719.3672.



O^6-(ethylcm)I (31) (pathway B, Figure 28). The TBDMS protected inosine derivative **33** (26.0 mg, 37.0 µmol) was co-evaporated with pyridine (2 mL), dissolved in pyridine (0.5 mL) followed by dropwise addition of hydrogen fluoride pyridine (73.4 µL, 530 µmol) and the reaction mixture was stirred at room temperature for 15 h. The solvent was removed under reduced pressure and the residue purified by flash chromatography Si_{60} CH_2Cl_2/ MeOH. Residual pyridine was co-evaporated with toluene then H_2O was added and the residue was lyophilized to obtain **31** as yellowish solid (13.2 mg, 37.0 µmol, 99%.). ^{1}H-NMR (500 MHz, MeOD) δ/ppm: 8.59 (s, 1H), 8.48 (s, 1H), 6.10 (d, *J* = 5.8 Hz, 1H), 5.19 (s, 2H), 4.74 (t, *J* = 5.5 Hz, 1H), 4.36 (dd, J = 5.1, 3.3 Hz, 1H), 4.23 (q, *J* = 7.1 Hz, 2H), 4.17 (q, *J* = 3.1 Hz, 1H), 3.90 (dd, *J* = 12.4, 2.8 Hz, 1H), 3.77 (m, 1H), 1.26 (t, *J* = 7.1 Hz, 3H). ^{13}C-NMR (101 MHz, MeOD) δ/ppm: 169.76, 160.99, 153.06, 152.7, 144.35, 122.65, 91.01, 87.79, 75.73, 72.29, 64.05, 63.12, 62.44, 14.40. HRMS (ESI): $C_{14}H_{18}N_4NaO_7^+$ *calcd.*: 377.1068, *found*: 377.1071.



General Procedure for Cu(I)-catalyzed monophosphates (MPs) and triphosphates (TPs) alkylation using EDA (small scale)

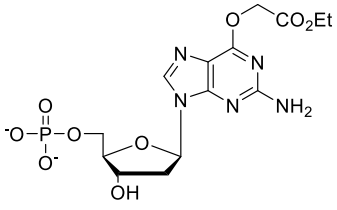
Typically 20 µL of total reaction volume contained 1 mM of $CuSO_4$, 5 mM of MP or TP, 100 mM MES buffer, 50 mM EDA. Stock solutions of each compound were prepared and combined in an Eppendorf tube in the given order. The reactions were monitored by injection of 1 µL of the reaction mixture into RP-HPLC using Method C. The conversion of MP or TP`s were directly calculated from the corresponding peak areas in the HPLC traces (Table 2). The peaks were separated using Method C (6 µL injection of reaction mixture), collected and analyzed by ESI-MS/MS.

Stock solutions: $CuSO_4$: 5 mM in H_2O; MP/TP: 50 mM in H_2O; MES buffer: 500 mM in H_2O, pH 6
EDA: 250 mM in DMSO; ascorbate: 100 mM in H_2O

General Procedure for Cu(I)-catalyzed MP and TP alkylation using EDA (big scale)

To a solution of $CuSO_4$ (1 mM) in H_2O (degassed), MES buffer stock solution (100 mM), MP/TP (5 mM), EDA (50 mM) and ascorbate (5 mM) were added and the reaction mixture was stirred at room temperature for 1 h. The reaction mixture was lyophilized and the residue was purified by preparative RP-HPLC using Method D. The corresponding product fractions were combined and lyophilized. The residue was dissolved in H_2O and $NaClO_4$ (15 eq.) was added. The mixture was stirred for 16 h at room temperature followed by addition of acetone (400 % v/v). The mixture was centrifuged at 4400 x g for 3 min. The supernatant was carefully discarded and the pellet was dissolved in ultra pure H_2O and lyophilized to give the final modified MP/TP.

O^6-(ethylcm)dGMP (35). According to the general procedure $CuSO_4$ (5.62 mg, 22.5 µmol), dGMP (50.0 mg, 112 µmol), EDA (136 µL, 1.12 mmol) and ascorbate (22.3 mg, 112 µmol) in H_2O (25.5 mL) and MES buffer stock solution (51.0 µL) were reacted to yield **35** as white solid (41.8 mg, 86 %.). ^{1}H-NMR (500 MHz, MeOD) δ/ppm: 8.27 (s, 1H), 6.34 (t, *J* = 6.7 Hz, 1H), 5.04 (s, 2H), 4.71 (s, br, 1H), 4.22 (q, *J* = 7.1 Hz, 2H), 4.10 (s, br, 1H), 4.01 (s, br, 2H), 2.84 (dt, *J* = 13.2, 6.5 Hz, 1H), 2.38 (ddd, *J* = 13.2, 6.0, 3.1 Hz, 1H), 1.25 (t, *J* = 7.1 Hz, 2H). ^{13}C-chemical shifts were extracted from 2D HMQC and HMBC spectra (500 MHz, MeOD) δ/ppm: 168.81, 159.59, 153.94, 139.15, 86.98, 83.79, 71.84, 63.88, 61.91, 60.79, 38.98, 13.02. ^{31}P-NMR (202 MHz, MeOD) δ/ppm: 5.05 (s, 1P). HRMS (ESI): $C_{14}H_{19}N_5Na_2O_9P^+$ *calcd.*: 478.0710, *found*: 478.0716.



O^6-(ethylcm)GMP (35b). According to the general procedure $CuSO_4$ (4.78 mg, 19.2 µmol), GMP (50.0 mg, 95.8 µmol), EDA (116 µL, 958 µmol), ascorbate (18.98 mg, 95.8 µmol) in H_2O (24.7 mL) and MES buffer stock solution (50.0 µL) were reacted and purified to yield **35b** as white solid (45.7 mg, 92 %.). ^{1}H-NMR (500 MHz, MeOD) δ/ppm: 8.32 (s, 1H), 5.97 (d, *J* = 5.8 Hz, 1H), 5.06 (s, 2H), 4.77 (t, *J* = 5.4 Hz, 1H), 4.49 (dd, *J* = 5.1, 3.1 Hz, 1H), 4.24 (q, *J* = 7.1 Hz, 2H), 4.21 (s, 1H), 4.12 – 4.04 (m, 1H), 1.27 (s, 3H). ^{13}C-chemical shifts were extracted from 2D HMQC and HMBC spectra (500 MHz, MeOD) δ/ppm: 168.82, 159.63, 154.34, 139.34, 113.68, 87.81, 84.92, 74.21, 71.14, 63.67, 61.96, 60.91, 13.08. ^{31}P-NMR (202 MHz, MeOD) δ/ppm: 4.89 (s, 1P). HRMS (ESI): $C_{14}H_{19}N_5Na_2O_{10}P^+$ *calcd.*: 494.0659, *found*: 494.0668.

O^6-(ethylcm)IMP (35c). According to the general procedure $CuSO_4$ (4.73 mg, 18.9 µmol), IMP (50.2 mg, 94.7 µmol), EDA (115 µL, 947 µmol), ascorbate (18.7 mg, 94.7 µmol) in H_2O (25.4 mL) and MES buffer stock solution (51.0 µL) were reacted to yield **35c** as white solid (33.1 mg, 81 %.). ^{1}H-NMR (500 MHz, MeOD) δ/ppm: 8.88 (s, 1H), 8.47 (s, 1H), 6.21 (d, *J* = 5.6 Hz, 1H), 5.17 (s, 2H), 4.78 (t, *J* = 5.1 Hz, 1H), 4.49 (s_{br}, 1H), 4.29 – 4.17 (m, 3H), 4.07 (s_{br}, 2H), 1.25 (t, *J* = 7.1 Hz, 3H). ^{13}C-chemical shifts were extracted from 2D HMQC and HMBC spectra (500 MHz, MeOD) δ/ppm: 168.51, 159.10, 152.44, 151.45, 143.06, 87.92, 85.39, 75.16, 71.23, 63.64, 62.61, 61.0, 13.03. ^{31}P-NMR (202 MHz, MeOD) δ/ppm: 4.96 (s, 1P). HRMS (ESI): $C_{14}H_{19}N_4NaO_{10}P^+$ *calcd.*: 457.0731, *found*: 457.0730.

O^6-(ethylcm)dGTP (41). According to the general procedure $CuSO_4$ (5.24 mg, 20.9 µmol), dGTP (70.0 mg, 101 µmol), EDA (122 µL, 1.01 mmol), ascorbate (20.0 mg, 101 µmol) in H_2O (22.9 mL) and MES buffer stock solution (46.0 µL) were reacted to yield **41** as white solid (39 mg, 65 %). ^{1}H-NMR (600 MHz, D_2O) δ/ppm: 8.28 (s, 1H), 6.41 (t, *J* = 6.8 Hz, 1H), 5.12 (s, 1H), 5.11 (s, 1H), 4.83 (dt, *J* = 7.0, 3.6 Hz, 1H), 4.32 – 4.27 (m, 3H), 4.24 (ddd, *J* = 10.4, 6.2, 4.1 Hz, 1H), 4.21 – 4.15 (m, 1H), 2.83 (dt, *J* = 13.8, 6.8 Hz, 1H), 2.54 (ddd, *J* = 14.1, 6.3, 3.6 Hz, 1H), 1.26 (t, *J* = 7.1 Hz, 3H). ^{13}C-chemical shifts were extracted from 2D HMQC and HMBC spectra (600 MHz, D_2O) δ/ppm: 171.02, 159.71, 153.64, 139.38, 114.1, 85.72, 83.43,

70.84, 65.27, 63.26, 62.74, 38.50, 13.22. ^{31}P-NMR (243 MHz, D_2O) δ/ppm: -6.74 (s, 1P), -10.9 (d, *J* = 19.4 Hz, 1P), -22.04 (s, 1P). HRMS (ESI): $C_{14}H_{19}N_5Na_4O_{15}P_3^+$ *calcd.*: 681.9676, *found*: 681.9685.

O^6-(ethylcm)GTP (41b). According to the general procedure $CuSO_4$ (6.15 mg, 24.6 µmol), GTP (70.0 mg, 118 µmol), EDA (143 µL, 1.18 mmol), ascorbate (23.5 mg, 118 µmol) in H_2O (26.7 mL) and MES buffer stock solution (54.0 µL) were reacted to yield **41b** as white solid (53.2 mg, 73 %). ^{1}H-NMR, peak broadening caused by stacking of compound **41b** in D_2O (400 MHz, D_2O) δ/ppm: 8.40 (s_{br}, 1H), 6.11 (s_{br}, 1H), 5.14 (s_{br}, 2H), 4.56 (s_{br}, 1H), 4.37 (s_{br}, 1H), 4.36 – 4.08 (m, 5H), 1.24 (t, *J* = 6.6 Hz, 3H). ^{13}C-chemical shifts extracted from 2D HMQC and HMBC spectra, (600 MHz, D_2O) δ/ppm: 173.78, 84.66, 74.42, 71.15, 66.06, 64.19, 63.29, 30.85, 13.78 (four carbon signals could not be detected due to the peak broadening). ^{31}P-NMR (243 MHz, D_2O) δ/ppm: -5.39 (s_{br}, 1P), -10.49 (s_{br}, 1P)), -19.39 (s_{br}, 1P). HRMS (ESI): $C_{20}H_{38}N_6O_{16}P_3^+$ *calcd.*: 711.1552, *found*: 711.1551.

O^6-(cm)dGTP (43). O^6-(ethylcm)dGTP **41** (15.3 mg, 22.5 µmol) was placed in an Eppendorf tube (0.5 mL) dissolved in 1M LiOH (33.7 µL, 33.7 µmol) and stirred at 25°C. The reaction mixture was monitored by RP-HPLC using method C. After 1 h full conversion of the starting material was observed. The mixture was freeze dried and purified by preperative RP-HPLC using method D. The corresponding fraction were combined and lyophilized. The residue was dissolved in H_2O (600 µL) and $NaClO_4$ (41.1 mg, 338 µmol,) was added. The mixture was stirred for 5 h at room temperature and then acetone (4 mL) was added. The obtained suspension was centrifuged at 4400 g for 2 min. The supernatant was carefully discarded. The pellet was re-dissolved in ultra pure H_2O and lyophilized to yield **6** as a white solid (12.8 mg, 57%). ^{1}H-NMR (400 MHz, D_2O) δ/ppm: 8.35 (s_{br}, 1H), 6.40 (d, *J* = 7.7 Hz, 1H), 4.83 (s, 2H), 4.33 – 4.06 (m, 3H), 2.81 (dd, *J* = 14.3, 7.3 Hz, 1H), 2.56 (d, *J* = 10.2 Hz, 1H). ^{13}C-chemical shifts were extracted from 2D HMQC and HMBC spectra (600 MHz, D_2O) δ/ppm: 176.20, 160.56, 138.94, 85.42, 83.21, 70.83, 65.14, 63.34, 38.52. ^{31}P-NMR (243 MHz, D_2O) δ/ppm: -5.29 (s, 1P), -10.42 (s, 1P), -19.09 (s, 1P). HRMS (ESI): $C_{18}H_{33}N_6\,NaO_{15}P_3^+$ *calcd.*: 689.1109, *found*:. 689.1107.

2.7.3 Alkylation of ssDNA catalyzed by Cu(I) carbenes

Synthesis of ssDNA

Solid-phase DNA synthesis was carried out on 1-µmol CPG columns using standard phosphoramidite chemistry with 0.3 M 5.benzylthio-1-H-tetrazole as activator. The DNA was cleaved from the resin with 32 % (v/v) aqueous ammonia for 2 h at room temperature. Deprotection was carried out at 55°C for 18 h. The residue was freeze dried and purified as described in Methods A and B. The identity of all ssDNA was confirmed by ESI-MS or MALDI TOF MS.

Cu(I)-catalyzed ssDNA O^6-G modification using EDA or DAAs

General procedure for ssDNA alkylation

Typically 20 µL of total reaction volume contained 1 mM of $CuSO_4$, 5 mM of ssDNA, 100 mM MES buffer, 50 mM EDA or DAA and 5 mM ascorbate. Stock solutions of each compound were prepared and combined in an Eppendorf tube in that order. The reactions were monitored by injecting 1 µL of the reaction mixture into RP-HPLC using Method C. The conversions of the ssDNA`s were directly calculated from the corresponding peak areas in the HPLC traces. The peaks were separated using Method C (6 µL injection of reaction mixture), collected and analyzed by ESI-MS or UPLC-MS.

Stock solutions: $CuSO_4$: 5 mM in H_2O; MES buffer: 500 mM in H_2O, pH 6; EDA: 250 mM in H_2O containing 10 % (v/v) ^{t}BuOH or DMSO; DAA: 250 mM in DMSO or dioxane; ascorbate: 100 mM in H2O.

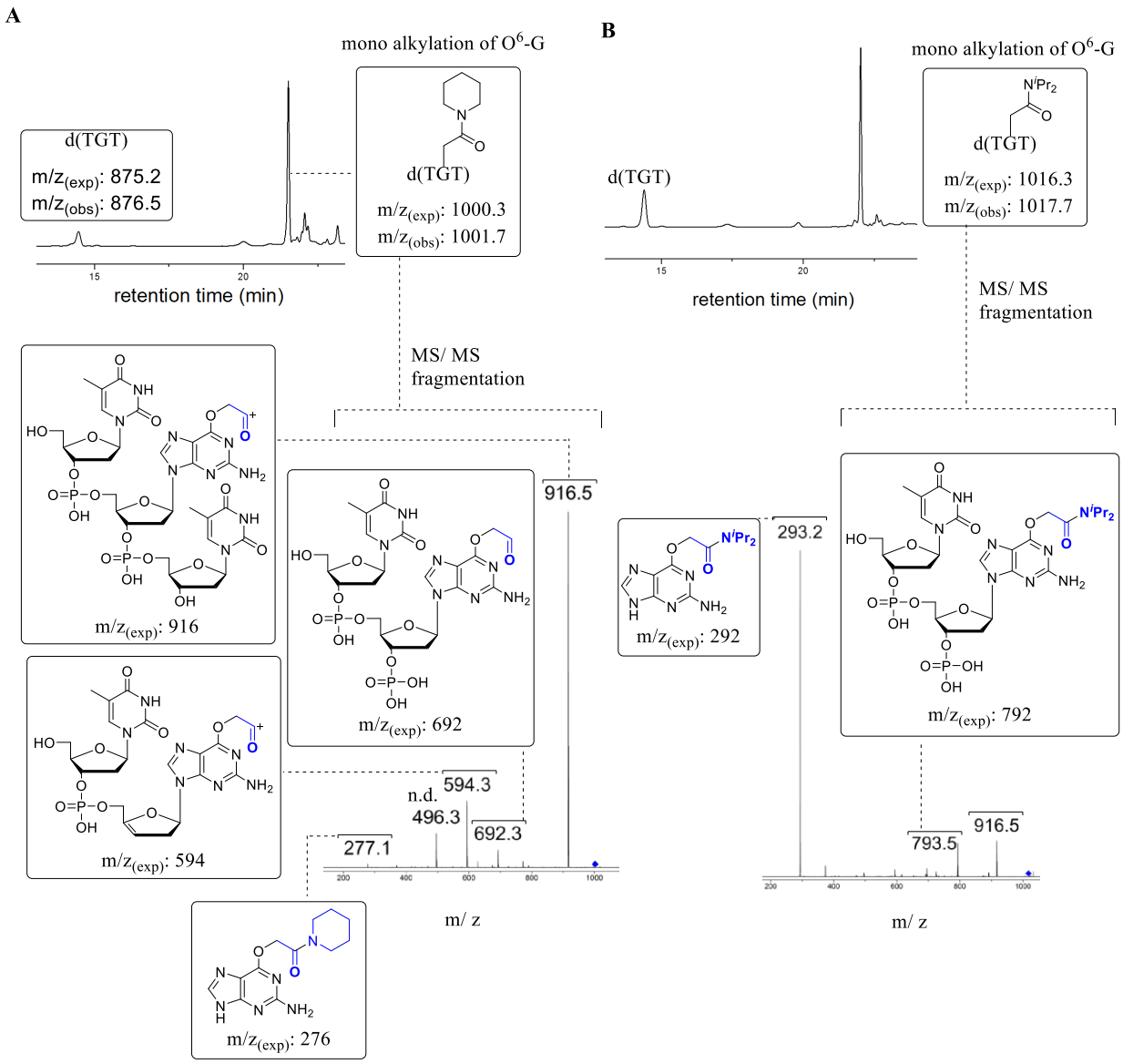


Figure 34. (A) HPLC-MS/MS analysis of O^6-G alkylation of d(TGT) with DAA **23** following the general procedure (conversion see Table 3, Entry 1). (B) HPLC-MS/MS analysis of O^6-G alkylation of d(TGT) with DAA **24** following the general procedure (conversion see Table 3, Entry 2).

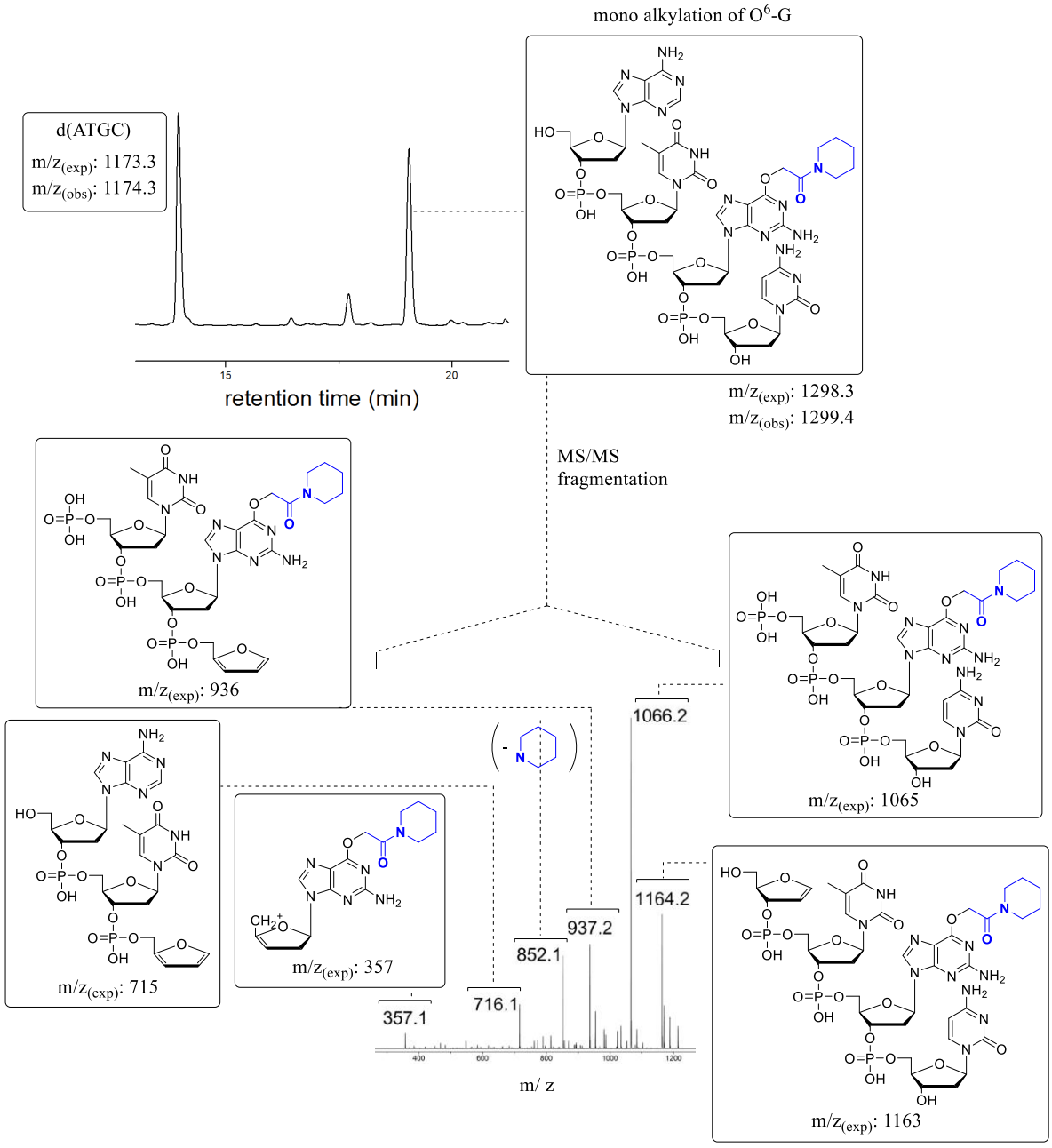


Figure 35. HPLC-MS/MS analysis of O^6-G alkylation reaction of d(ATGC) with DAA **23** following the general procedure (conversion see Table 3, Entry 3).

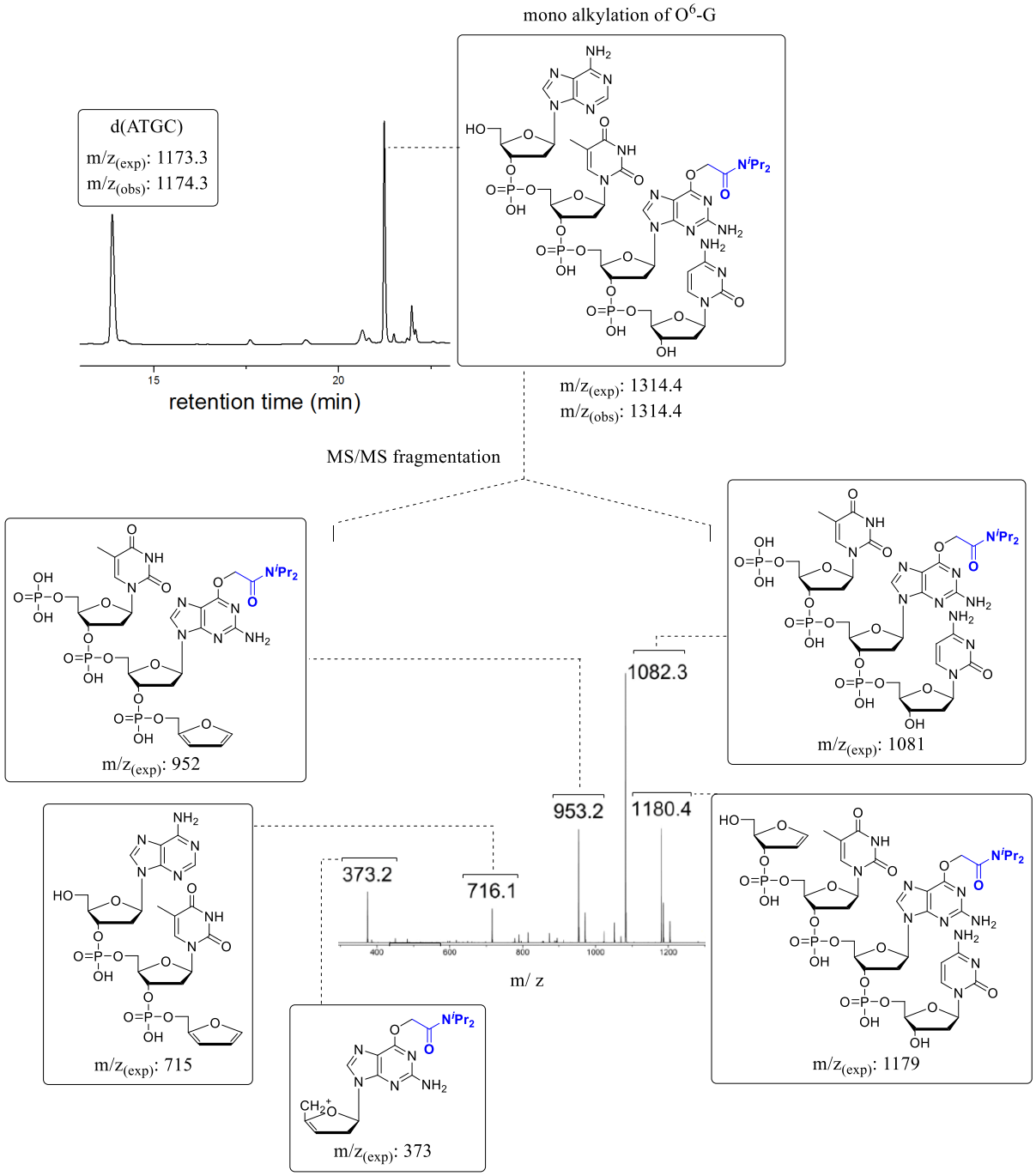


Figure 36. HPLC-MS/MS analysis of O^6-G alkylation reaction of d(ATGC) with DAA **24** following the general procedure (conversion see Table 3, Entry 5).

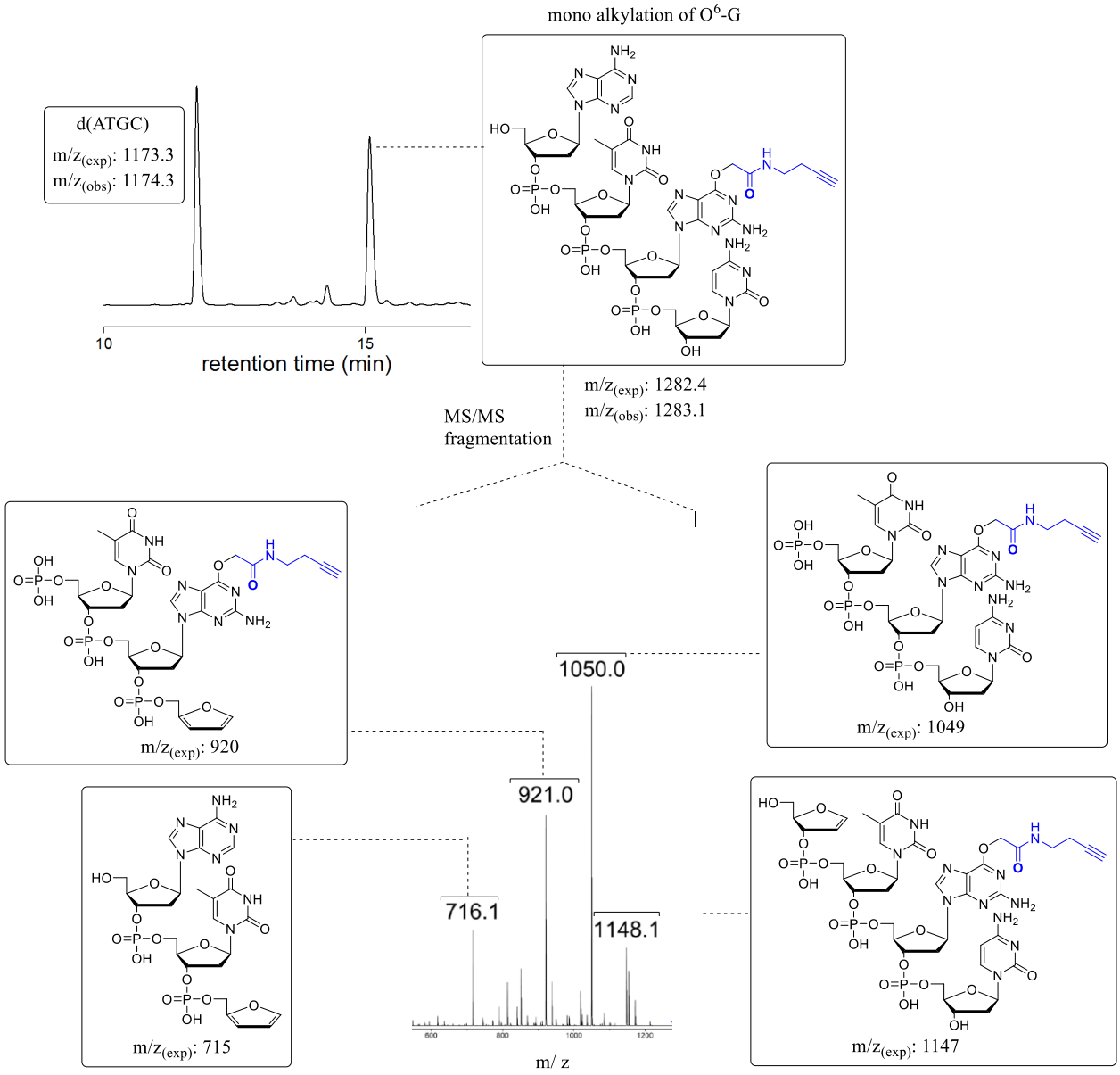


Figure 37. HPLC-MS/MS analysis of O^6-G alkylation reaction of d(ATGC) with DAA **25** following the general procedure (conversion see Table 3, Entry 10).

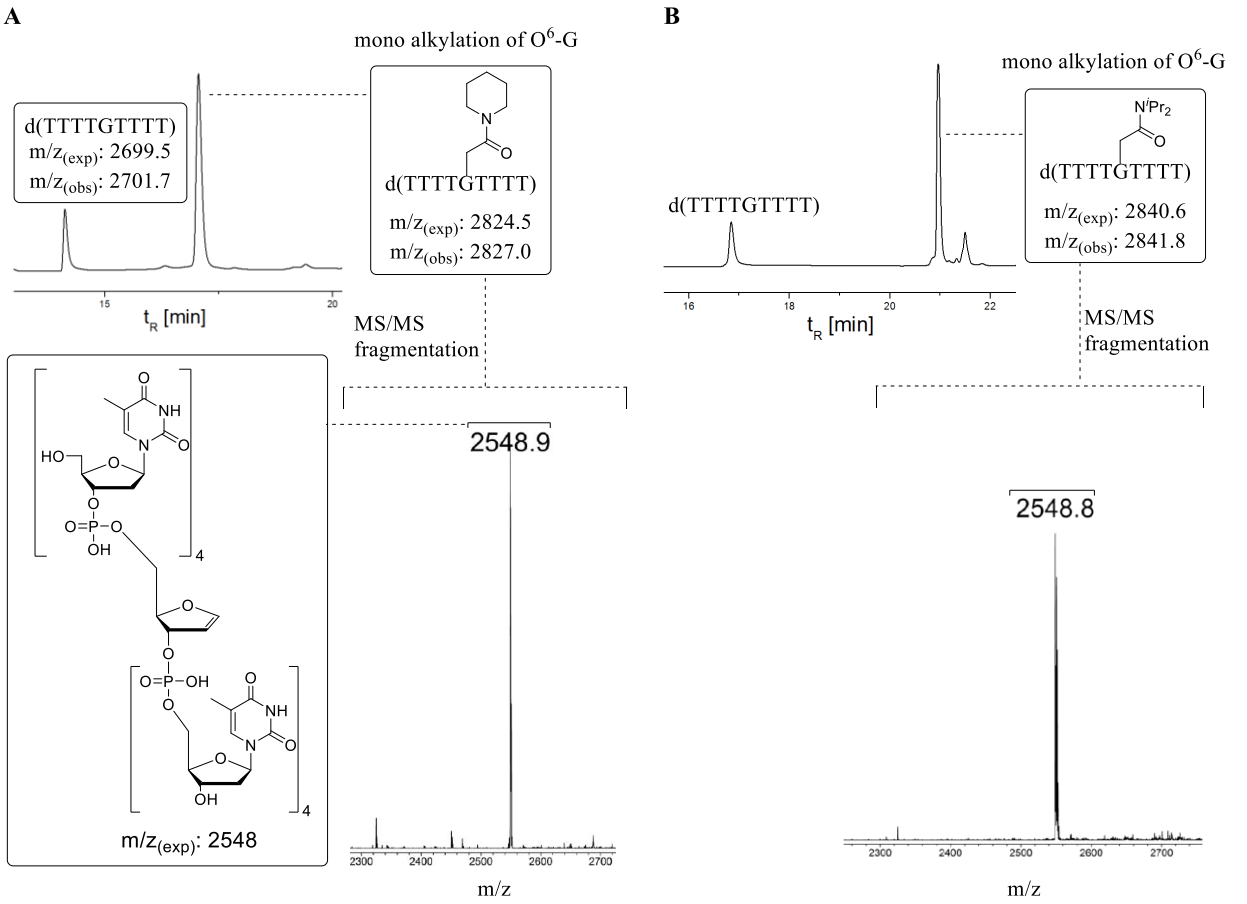


Figure 38. (A) HPLC-MS/MS analysis of O^6-G alkylation reaction of d(TTTTGTTTT) with DAA **23** following the general procedure (conversion see Table 3, Entry 6). (B) HPLC-MS/MS analysis of O^6-G alkylation reaction of d(TTTTGTTTT) with DAA **24** following the general procedure (conversion see Table 3, Entry 7).

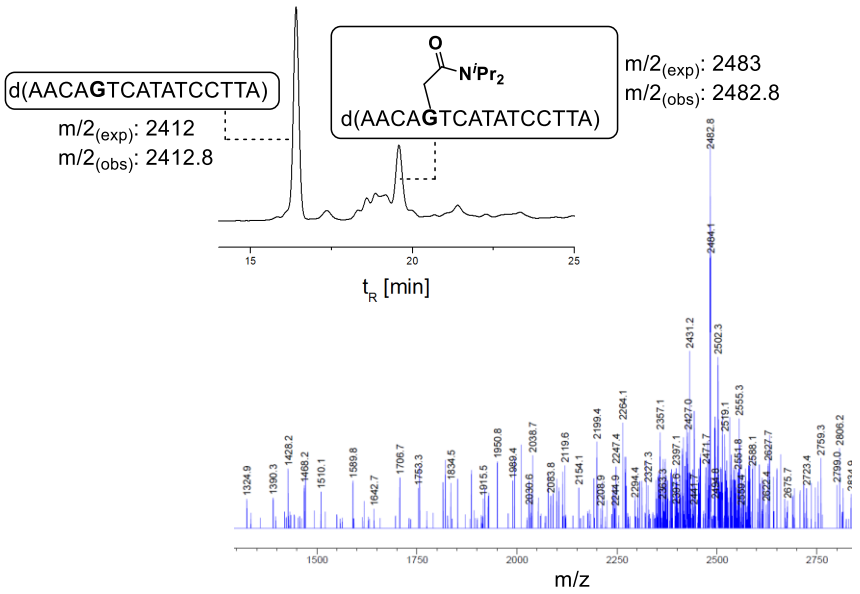


Figure 39. HPLC-UPLC-MS analysis of O^6-G alkylation reaction of d(AACAGTCATATCCTTA) with DAA **24** following the general procedure (conversion see Table 3, Entry 8).

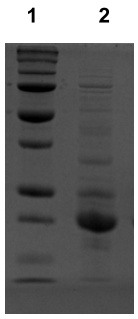
2.7.4 Click reaction with alkyne O^6-G modified d(ATGC) and azide biotin conjugate

According to the genereal procedure (see section 2.2), d(ATGC) was modified using Cu(I) carbene chemistry and DAA **12.** The alkylation product was purified using Method C and directly used for click reaction with azide biotin conjugate **16**, synthesized by a literature known procedure.[4] d(AT*O*^{6}-akyneGC) stock solution (100 µL, 34.8 µM stock solution in 1mM TEAA buffer pH 7.25) was mixed with azide biotin conjugate **16** (53 µL, 100 µM stock solution in DMSO), sodium ascorbate (17.4 µL, 1 mM stock solution in H_2O), DMSO (28.2 µL) and degassed for 30 seconds. The $CuSO_4$ stock solution (1.40 µL, 12.5 mM stock solution in H_2O) was added (see Figure S15, panel A) and the mixture was stirred at room temperature for 1 h. Analysis of the reaction mixture by RP-HPLC using method C (40 µL injection) showed 81 % conversion of the starting material to the target click product (Scheme 7B).

2.7.5 Reparation of O^6-G modified ssDNA 9-mer by hAGT

Expression and purification of hAGT

hAGT WT carrying an *N*-terminal hexa-histidine-tag was produced in *Escherichia coli* (*E.coli*) BL21 cells were grown in LB medium containing 50 mg/L ampicillin and 34 mg/L chloramphenicol. Protein expression was induced by adding 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) at 18°C. After 16 h at 18°C the cells were harvested, re-suspended in lysis buffer (20 mM TrisHCl pH 8 and 300 mM NaCl), lysed by sonication and centrifuged for 20 min at 4°C. The protein was purified by nickelnitrilotriaceticacid affinity (Ni-NTA) by mixing the clear lysate with Ni(II)-NTA agarose (Qiagen) for 20 min at 4°C. The agarose beads were washed with lysis buffer containing 10 mM and 20 mM imidazole. After elution of the protein using 200 mM imidazole the fractions found to contain the hAGT using SDS PAGE (Figure 40) were pooled and precipitated with 72% ammonium sulfate. The pellet was re-dissolved in urea and dialyzed in 50 mM Tris-HCl pH 7.6, 250 mM NaCl and 5 mM DTT and stored at -20°C.



1: Protein ladder 10-250 kDA
2: Elution with 200 mM imidazole

Figure 40. SDS PAGE analysis of the purified hAGT protein.

De-alkylation of O^6-G modified ssDNA 9-mer

The O^6-G modified ssDNA 9-mer **39** (79 µM) was incubated with hAGT (86 µM) for 4 h at 37°C (Scheme 12A+B). After 0.5 and 4 h aliquots of 30 µL were taken, centrifuged, the supernatant analyzed by RP-HPLC using Method C and the pellet re-dissolved in 0.1 % formic acid containing water and analyzed by HR-MS (ESI).

2.7.6 Single primer extension experiment with synthetic O^6-(cm)dGTP

Oligonucleotides and DNA sequences

Table 5. Oligonucleotides and DNA sequences for primer extension experiment.

Primer	**Sequence (5' - 3')**
23-mer	TAA TAC GAC TCA CTA TAG GGA GA
Templates	
28-mer G	ACT C**G**T CTC CCT ATA GTG AGT CGT ATT A
28-mer A	ACT C**A**T CTC CCT ATA GTG AGT CGT ATT A
28-mer T	ACT C**T**T CTC CCT ATA GTG AGT CGT ATT A
28-mer C	ACT C**C**T CTC CCT ATA GTG AGT CGT ATT A

Primer extension assays

Radioactive labeling of primer strands at their 5'end was performed using T4 polynucleotide kinase (Promega) and [γ-^{32}P]ATP following the manufacturer protocol. 1 μM of primer was annealed with 1.5 μM of corresponding templates at 95°C for 5 minutes followed by a slow cooling over 2 hours. Standard primer extension reactions (10 μL) contained 1x reaction buffer (50 mM Tris HCl pH 9.2, 16 mM $(NH_4)_2SO_4$, 2.5 mM $MgCl_2$, 0.1% Tween 20), 5 nM DNA polymerase, 15 nM DNA (15 nM primer and 22.5 nM template), and 10 μM dNTPs. Reactions were initiated by adding dNTP(s) to the enzyme/DNA mixtures and allowed to react at 55°C for 10 min. Reactions were then quenched by the addition of 10 μL PAGE gel loading buffer (80% formamide, 20 mM EDTA, 0.05% bromophenol blue, 0.05% xylene cyanole FF) and the product mixtures were analyzed by 15% polyacrylamide/7M urea denaturing gels and subjected to autoradiography (Bio-Rad).

Chapter 3: *N*-heterocyclic carbene complexes

The aim of this work was the development and synthesis of a water soluble copper complex for oligonucleotide modifications. In contrast to the Cu(I) catalyst *in situ* formed by mixing $Cu(II)SO_4$ with ascorbate we sought to identify a tight-binding ligand that permanently stabilize Cu(I). Such system might be able to overcome challenges with oxidative damage occurring by reducing the ability of Cu(I) to activate oxygen. Furthermore, a tight-binding ligand makes the catalyst modular, a feature that becomes important for further bio-conjugation such as linking the catalyst to a short oligonucleotide and use it as guiding catalyst in a templated directed modification of a second oligonucleotide.

3.1 *N*-heterocyclic carbene properties and their complexes

N-heterocyclic carbene (NHC) ligands derived from imidazole (Figure 41A) are known to form strong bonds with transition metals and to tolerate a broad scope of functional groups,[167] thereby providing an excellent choice of ligand to stabilize Cu(I). These carbenes form excellent catalysts for a large number of homogenous processes such as metathesis[168], hydrosilation[169], hydrogenation[170], cross couplings[171] and enatioselective transformations.[172] The majority of theses transformations were traditionally carried out with phosphine ligands, but the use of NHC complexes has in many cases improved thermal stability as well as sensitivity towards oxygen. In 1962 the reactivity and stability of NHCs were investigated for the first time by Wanzlick.[173] Six years later he reported the first synthesis of a NHC metal complex (Figure 41B).[174] Surprisingly, in following two decades NHCs were not further investigated in transition metal chemistry until Arduengo and coworkers isolated the first crystalline adamantyl substituted NHC (Figure 41C).[175] This bench-stable carbene showed extremely high stability in the absence of moisture and oxygen.

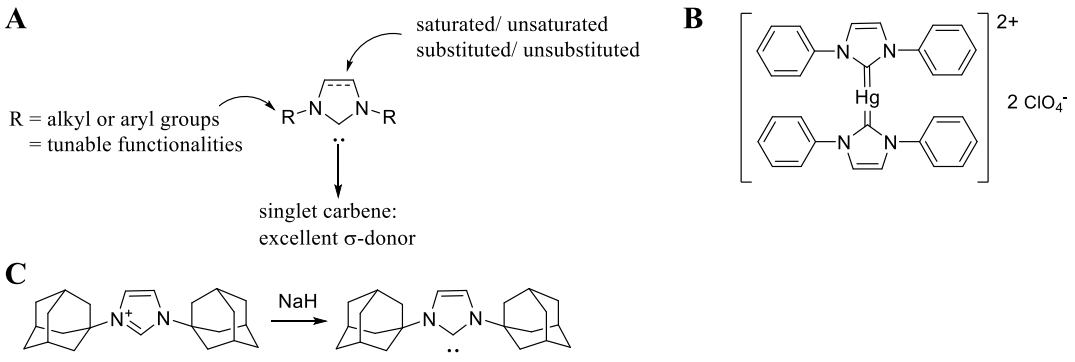


Figure 41. (A) Structure of imidazole NHC ligand. (B) First synthesized NHC-metal complex. (C) First crystalline and storable NHC.

Following Arduengo`s discovery the interest in NHCs increased dramatically and the electronic and steric contributions to the NHC-metal bond were investigated. At the beginning, σ-donation was considered as

the sole contributor to the NHC-metal bond. Later, more precise investigations showed that also π-interactions contributed. The most important electronic contributions to the NHC-metal bond, namely σ -donation, π -donation and π * back-donation are summarized in Figure 42. The electronic contributions to the metal carbene bond have also been extensively investigated by calculations.[176,177] Jacobsen and coworkers reported that in a system with a low d electron count both π interactions contribute to stability of the metal carbon bond.[178] However, in a system with a high d electron count the π -back donation alone represents the major contribution. Measuring IR spectroscopy and X-ray diffraction are possible methods to compare the electronic properties of NHC ligands.[179,180]

There are examples in the literature that correlate the bulkiness on the nitrogens with the stability of the resulting complexes.[181] Bulky substituents have the additional benefit that they can protect the free carbene from dimerization.[12] The majority of NHCs are stronger donors than the strongest phosphine ligand, however the behavior differences between NHCs themselves are more likely related to their steric properties.[182]

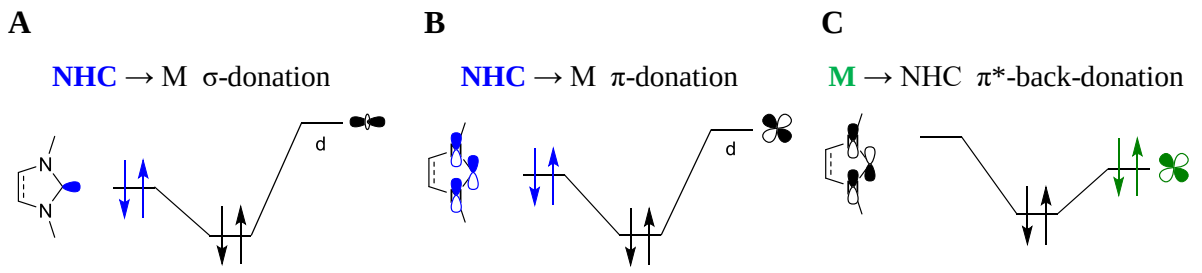


Figure 42. MO diagram representing the electronic contributions for NHC-metal complex stability. (A) NHC to metal σ-donation. (B) NHC to metal π-donation. (C) Metal to NHC π*-backdonation.

3.2 *N*-heterocyclic copper carbene complexes as catalysts in chemical transformations

The synthesis of the first *N*-heterocyclic copper carbene complex as a cationic di-carbene $[(NHC)_2Cu]^+$ was reported in 1993 by Arduengo (Figure 43).[183] One year later, Raubenheimer presented a neutral mono-carbene copper complex [(NHC)CuCl].[184] Since then, many modifications have been made to the NHC ligands in order to tune their steric and electronic properties. Two structural alterations became very important in this field: the introduction of chiral centers to create steroeselective catalysts and the addition of functional groups to form, for example, chelating ligands.[185]

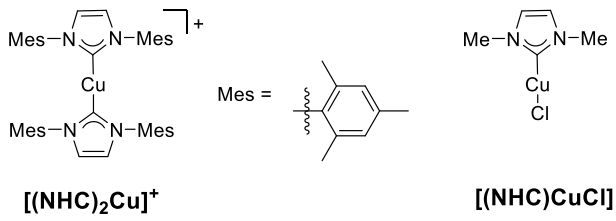


Figure 43. Structures of the first cationic **[(NHC)$_2$Cu]$^+$** and neutral **[(NHC)CuCl]** complexes.

Since 2001 [(NHC)-CuX] complexes have been employed in homogenous catalysis. In addition to their use as catalysts these Cu(I) species also displayed toxicity against human tumor cells.[186] The pioneering work of Woodward showed that these complexes can catalyze the conjugate 1,4-addition of diethylzinc to enones (Figure 44A).[187] This conjugation proceeded with *in situ* catalyst formation, which still represents a common approach nowadays. Two years later, Buchwald and Sadighi presented a reductive addition catalyzed by a NHC-Cu complex (Figure 44B).[188] Besides conjugate addition many other reactions are reported for NHC-Cu(I) catalytic systems (Figure 44C) namely hydrosilation reactions,[189,190] carboxylation and carbonylation,[191] allylic substitutions[192] and the [3+2] cycloaddition of a terminal alkyne with organic azides.[193]

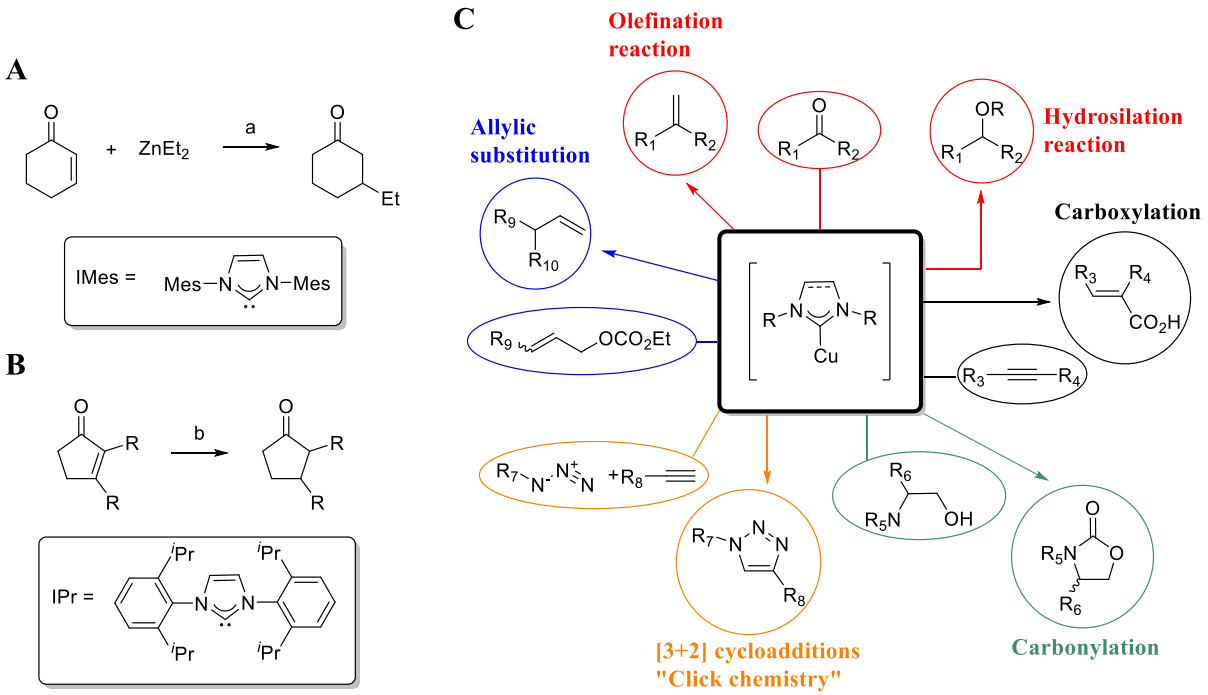


Figure 44. (A). The first example of an *in situ* formed NHC-Cu catalyst. (B) The first example of a [(NHC)Cu] complex used in catalysis. (C) Various transformation catalyzed by [(NHC)Cu] complexes. a) IMes·HCl/KOtBu, Cu(OTf)$_2$. (b) [(IPr)CuCl], KO^{t}Bu, polymethyldydrosilan (PMHS).

Related to this work [(NHC)CuX] complexes are reported to catalyse carbene transformations with diazo compounds. However, only a few NHC-metal systems are suitable to form the carbene from the diazo substrate. Pérez and Nolan developed the first carbene formation with the [(NHC)CuX] catalysts IPrCuCl and EDA in the presence of styrene (Figure 45).[194] Suprisingly, no consumption of EDA was detected without the metal coordinating substrate (Figure 45). In addition, aniline, pyrolidine, ethanol and sec-butanol were successfully alkylated under these conditions. It is likely that the [(NHC)CuCl] catalyst does not represent the catalytic active species, instead the nucleophilic substrate helps to form it.

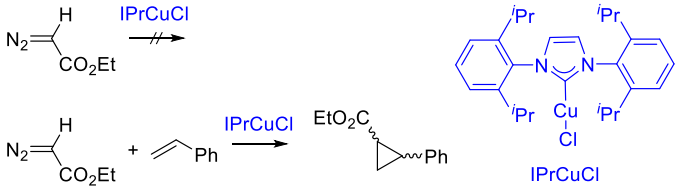


Figure 45. Reactions of EDA in the presence of IPrCuCl.

Two years later Pérez and Nolan expanded the scope of carbene transfer to include C-H insertion, copper carbenoids derived from EDA and [(NHC)CuCl] or [(NHC)AuCl] alkylate alkanes in the presence of a halide scavenger.[195] This methodology showed that the metal source, the counter ion and bulkiness of the NHC ligand have a big influence on the overall yield and regioselectivity of C-H functionalization. In 2008 nitrene transfer to aliphatic alkenes was reported by Apella and coworkers.[196] The nitrene is created *in situ* by iodoxybenzene and sulfamide (trichloroethylsulfamate ester). Further investigations into different [(NHC)CuX] sources showed that Cu(II) was a better catalyst for this transformation (Figure 46).

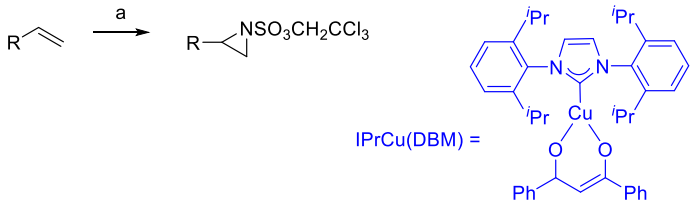


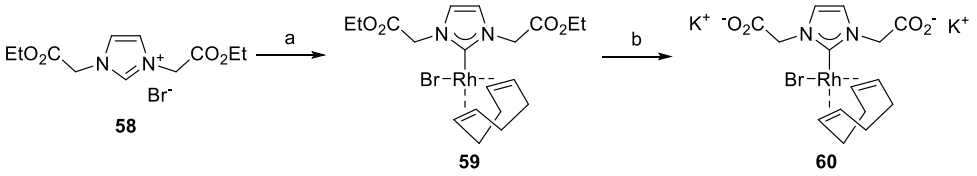
Figure 46. Aziridination of aliphatic alkenes using *in situ* formed nitrenes catalyzed by NHC-Cu(II) complex. (a) IPrCu(DBM), PhI=O, $H_2NSO_3CH_2CCl_3$.

3.3 Water soluble *N*-heterocyclic carbene metal complexes

As already mentioned in chapter 2 the nucleic acid modification using carbene chemistry proceed in water. Therefore, we were concentrated on the synthesis of NHC bearing water solubilizing functionalities.

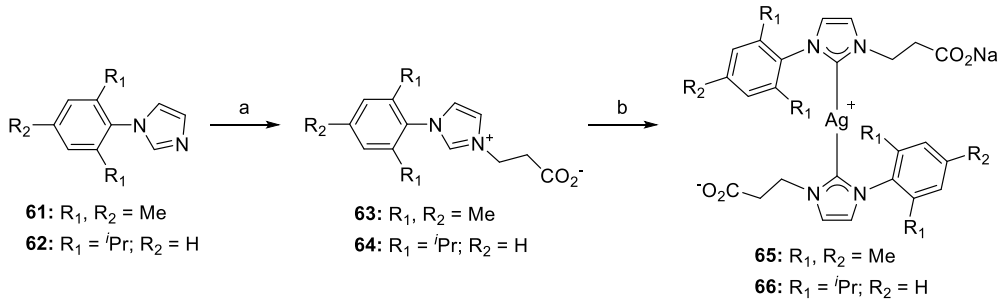
3.3.1 Carboxylate or sulfonate functionalized NHC and corresponding metal complexes

The first water soluble NHC metal complex **60** was reported in 1997 by Gossen and coworkers.[197] Water solubility was achieved by the introduction of two carboxylate side chains (Scheme 19). The carboxylates were formed, following rhodium complexation, via ethyl ester hydrolysis. The NHC functionalized ligand **58** was synthesized via standard alkylation of commercially available imidazole with the corresponding bromo alkyl source.



Scheme 19. Structure of the first water soluble rhodium complex. (a)[((cod)Rh(OEt))$_2$]. (b) KOH.

Later similar strategies were developed to synthesize the carboxylate functionalized NHC ligands **63** and **64** prior to metal complexation (Scheme 20).[198] Using water as the reaction medium the corresponding silver complexes were **65** and **66** synthesized by Shaugnessey and coworkers. NaCl was added to the reaction medium to avoid cationic coordination of silver to the carboxylate and polymerization. Transmetalation of these complexes led to the corresponding gold species.

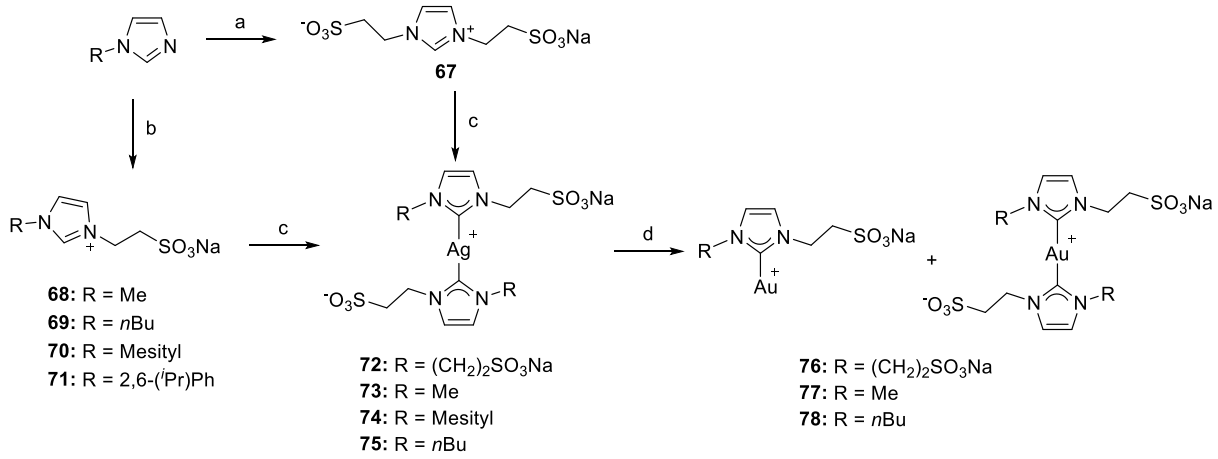


Scheme 20. Synthesis of carboxylate functionalized NHC **63** and **64** and the corresponding silver complex **65** and **66**. (a) $Br(CH_2)_2CO_2$ then Na_2CO_3. (b) Ag_2O, NaCl.

As demonstrated in both Scheme 19 and Scheme 20, the metal source of carboxylate-functionalized NHCs is introduced as metal precursor (silver complex) with an internal base. To date no strategy has been

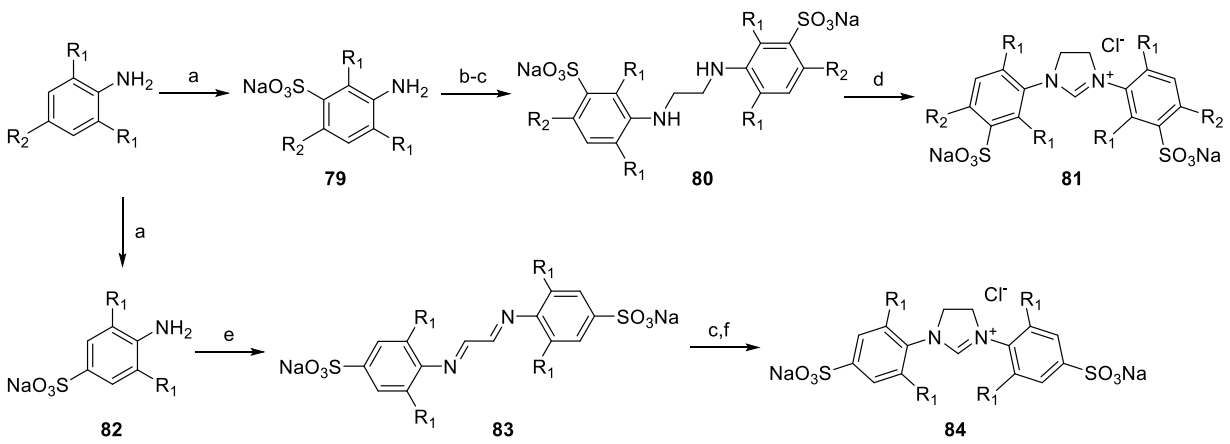
described to synthesize the corresponding complexes via free carbene; deprotonation under basic condition followed by metal complexation. A possible explanation might be that nucleophilic attack on the carbonyl is prevented using the transmetallation strategy.

Another possibility to introduce polarity to the NHC ligand is sulfonate functionalization. The sulfonate group is known to induce excellent water solubility due to its changed nature at most pHs. Two types of sulfonate functionalized NHCs are reported: the alkyl and aryl substituted NHC ligands (Scheme 21 and Scheme 22). Synthesis of the alkyl variant was reported in 1995 via two routes: the alkylation of commercially available imidazole with 2-bromoalkylsulfonates to obtain the symmetric NHC ligand **67**; Cleavage of sultones to yield NHC ligands **68**-**71**.[199,200] Comparable to carboxylate functionalized NHCs, the silver substrates **72**-**75** are available via complexation with silver oxide in the presence of NaCl. Transmetalation of the NHC-silver complexes **72**,**73** and **75** with chloro(tetrahydrothiophene)gold(I) resulted in a mixture of the dimeric and monomeric carbene gold complexes **76**-**78**.[201] In contrast to NHCs bearing carboxylates the synthesis of sulfonate functionalized NHC-metal complexes via a free carbene is reported.[202,203]



Scheme 21. Synthesis of alkyl sulfonate functionalized NHCs and corresponding gold complexes. (a) $Br(CH_2)_2SO_3Na$. (b) propane/butane sultone or C_2H_4CX then Na_2SO_3. (c) Ag_2O, NaCl. (d) [Au(tht)Cl], NaCl.

In 2007 Hoveyda and coworkers presented the synthesis of a chiral arylsulfonate substituted NHC ligand with good water solubility.[204] Plenio`s group reported the synthesis of aryl-sulfonate functionalized NHCs **81** and **84**.[205] These ligands were synthesized by the conventional condensation route starting from an aniline source (Scheme 22). Subsequent sulfonation, condensation with glyoxal to the Schiff base, reduction to the diamine and final cyclization with triethyl orthoformate yielded the target arylsulfonates. In the follow-up studies direct sulfonation of the NHC ligands was reported.[206]

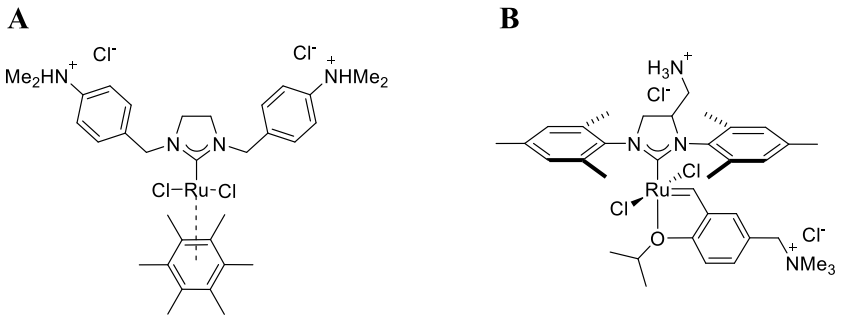


Scheme 22. Synthesis of arylsulfonate functionalized NHCs. a) oleum. b) glyoxal. c) H_2, Pd/C. d) $HC(OEt)_3$. e) 1,4-dioxane-2,3-dio. f) chloromethyl pivalate.

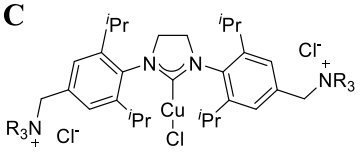
Comparable to alkylsufonate substituted NHC ligands, the majority of the corresponding complexes were synthesized by Ag_2O complexation followed by transmetalation.[207] Roy and coworkers showed that synthesis of Pd complexes starting from these ligands is possible via the free carbene intermediate.[206]

3.3.2 Amine/ Ammonium functionalized NHC and corresponding copper(I) complexes

Amine or ammonium functionalization represents a different strategy to introduce water solubility to NHC-metal complexes. Quaternization of amines to the corresponding ammonium species is necessary to achieve the desired polarity of the ligands. The first amine functionalized NHC was reported in 2001 by Oezdemir and coworkers (Figure 47A).[208] Treatment with HCl yielded the desired water solubility of the Ru(II) complex as an ammonium chloride salt. However, the complex stability in water is described to be low. Later, a backbone modified NHC-Ru(II) complex was synthesized by Grubbs et al (Figure 47B**)**.[209] The backbone functionalization allows introduction of the favored polarity without affecting the steric and electronic properties around the metal center.

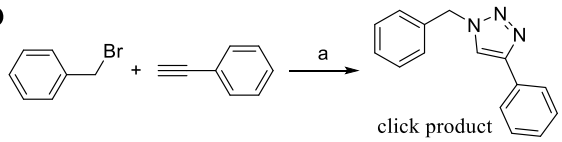


C



85: R = Et, X = Cl, **86:** R = Et, X = Br
87: R = Et, X = I, **88:** R = nBu, X = Cl
89: R = nBu, X = Br, **90:** R = nBut, X = I

D



entry	catalyst	Time [h]	Yield click product, [%]
1	**85**	3	98
2	**86**	6	Trace
3	**87**	6	92

Figure 47. (A) The first amine functionalized NHC-metal complex. (B) Backbone functionalized water soluble NHC-Ru complex. (C) Ammonium tagged [(NHC)CuX] complexes **85**-**90** used in three component click reaction. (D) CuAAC catalyzed by NHC-Cu(I) catalysts **85**-**87**. (a) 5 mol % catalyst, NaN_3, H_2O, rt.

Sharpless introduced the concept of "Click" chemistry in 2001[210] and since then Cu(I) catalyzed azide-alkyne cycloadditions (CuAAC) has become the most popular click reaction due to its high reliability.[211] Wang and co-workers reported the synthesis of different ammonium functionalized [NHC-CuX] complexes **85**-**90** (Figure 47C) that catalyze the three component cycloaddition of benzyl bromide, phenylacetylene and NaN_3 in water (Figure 47D).[212] A clear effect of the electronic properties of the NHC ligand as well as the counter ion is reflected in the reaction yields and times. Of these compounds **85** proceeded as the best catalyst. Reuse of catalyst **85** following liquid-liquid extractions with ether was possible for up to four cycles. Nevertheless, an argon atmosphere was necessary to maintain a similar catalytic activity. One year later, Gaulier and coworkers presented the first ammonium functionalized [(NHC)CuI] complex **91**, catalyzing the CuAAC of peptides bearing sensitive side chains (Figure 48A).[213] This catalyst showed high stability towards the oxidation of Cu(I). The cycloaddition was able to proceed under open air conditions. Many solvents including MeOH, DMSO, NMP and HFIP were compatible with catalyst **91**, only α-amino acids prevented the click reaction. The test substrates, three pentapeptides containing sensitive amino acids towards copper-mediated oxidation (His, Met, Cys), were investigated (Figure 48B). Five-fold excess of the corresponding azides and 40 to 100 mol % catalyst loading yielded the target click products **92-96**. No oxidation of sulfide and thiol containing peptides occurred. Additionally, the *N*-terminal octapeptides were successfully modified under these conditions (Figure 48B). Surprisingly, the [(NHC)CuI] catalyst **91** was synthesized via free carbene intermediate using NaOH and CuI in the presence of two alcohol functional groups, which are known to form alcohol adducts with these types of free carbenes.[214] Unfortunately neither the proton and carbon NMR spectra (only chemical shifts) nor a crystal structure of complex **91** has been published. Complete structural elucidation, monomeric or dimeric Cu carbene complex, is impossible without a crystal structure. Additionally, the mass for compound **91** reported in the paper is incorrect.

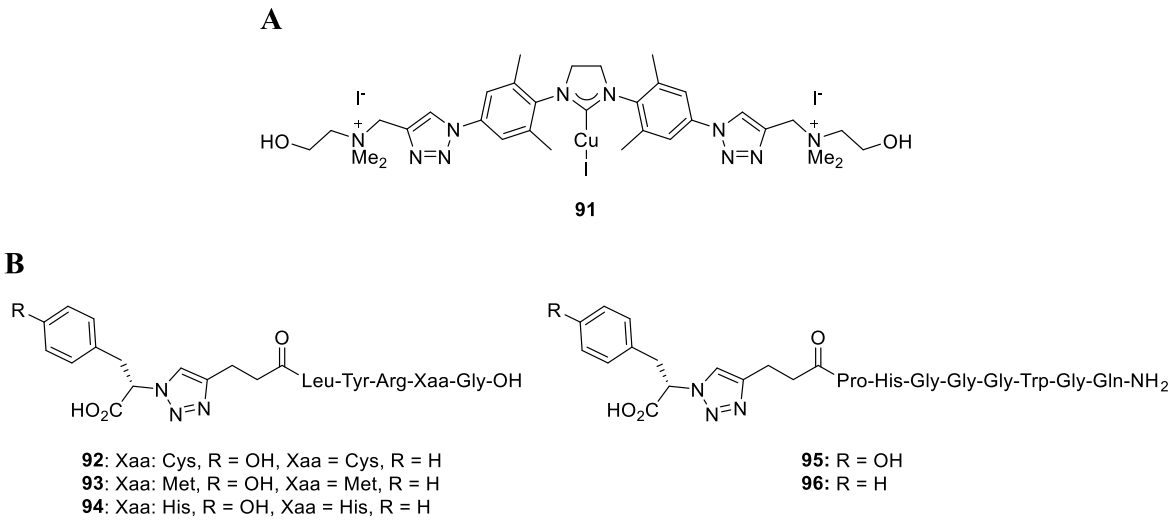
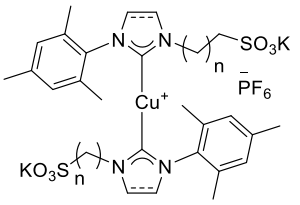


Figure 48. (A) Ammonium functionalized [(NHC)CuI] 91catalyzing the CuAAC of unprotected peptides. (B) Synthesized peptides using catalyst 91in the presence of 0.2 M HEPES pH 7.6.

Considering all of this information I believe the results from this paper require further examination. It is impossible to make any conclusive statements about the origin of the high complex stability without full structural elucidation.

Nonionic water-solubilizing elements can be introduced by alcohol or ether functionalization of the NHCs. To achieve the desired water solubility a variety of NHCs were substituted with either glycol or ether groups. For a detailed report the reader is referred to reviews by Schaper and coworkers.[215-216]

The most recent data on the synthesis of water soluble [(NHC)CuX] catalysts was reported in 2014 by Velazquez and coworkers. The [$(NHC)_2$Cu] compounds **97** and **98** bearing an alkyl sulfonate on one side (Figure 49)[216] successfully catalyze the standard CuAAC reaction of benzyl azide and phenylacetylene. However, once again mass spectrometry was used for structural determination instead of crystal structures. Ionization of [(NHC)CuX] complexes can cause ligand dissociation or dimerization products which not represent the actual structure. Additionally, the quality of the corresponding proton and carbon NMR spectra are low. It is therefore impossible to identify the corresponding structures with the given data.



97: n = 3
98: n = 4

Figure 49. Bis-carbene structure of water soluble [$(NHC)_2Cu$] **97**-**98** bearing alkyl sulfonate on one side.

In conclusion, only a few water soluble [(NHC)CuX] complexes **85-91, 97** and **98** have been described to date. All these reported water soluble [(NHC)CuX] complexes were developed for CuAAC chemistry. Unfortunately, their structure validation is poor, especially for complex **91**, preventing any detailed insight into their properties. In the remainder of this chapter I will discuss my efforts towards creating water soluble NHC complexes and testing them in catalysis.

3.4 Synthesis of water soluble NHC ligands and their Cu(I) complexes

Our first design of a [(NHC)CuX] catalyst to proceed in reducing agent free oligonucleotide alkylation is shown in Figure 50. Two hydrophillic functionalities were introduced in a para position to the nitrogens, to minimize their influence on the electronic properties of the metal center. The diisopropyl variant was choosen to maintain the correct steric properties, that would prevent carbene dimerization during complexation (Figure 50A). Our focus was the synthesis of a monomeric water soluble [(NHC)CuX] complex to use a similar synthetic strategy for the long term goal the development of a [(NHC)Cu] catalyst bound to single stranded oligonucleotides. This catalyst should further be used for template directed oligonucleotide modification (Figure 50B).

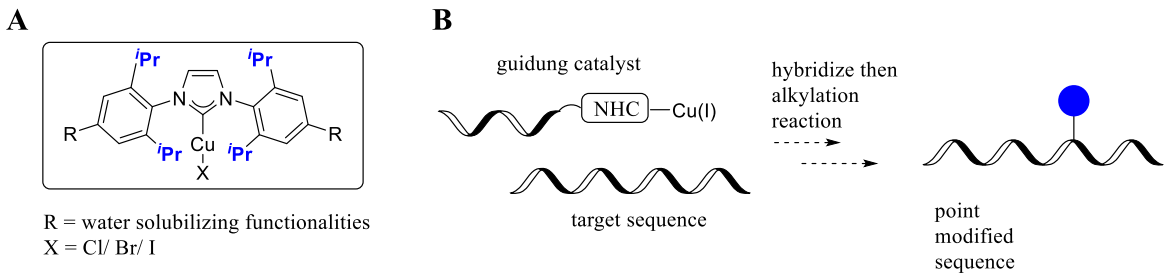
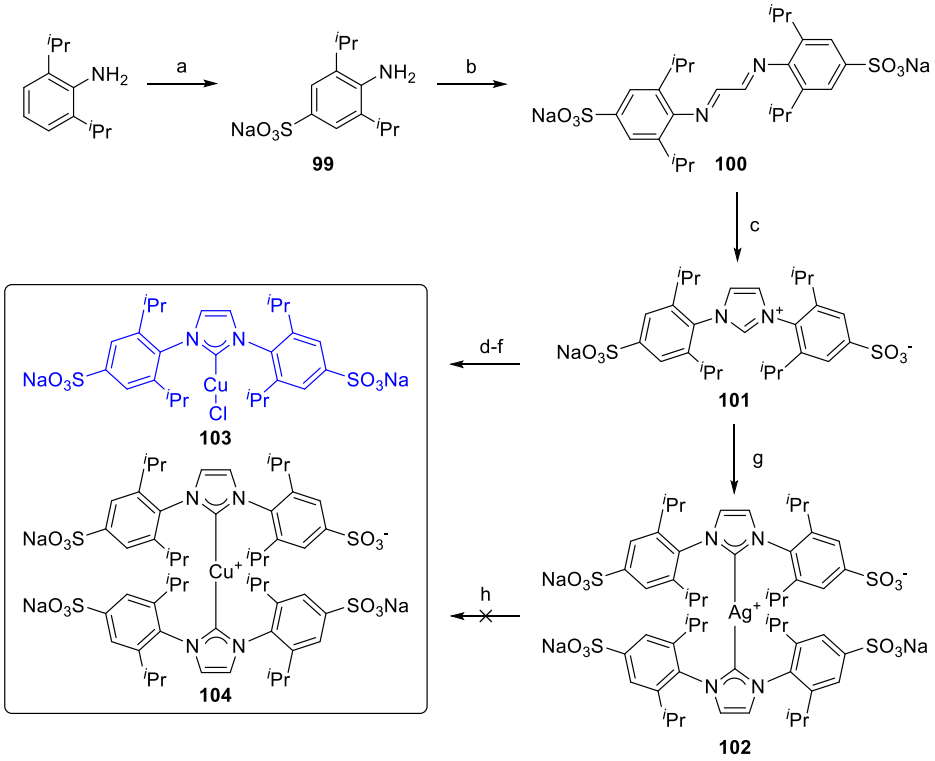


Figure 50. (A) General design of water soluble [(NHC)CuX] catalyst for oligonucleotide alkylation. (B) Template-directed oligonucleotide modification with guiding NHC-Cu(I) catalyst.

3.4.1 Synthesis of NHC-Cu(I) complex bearing sulfonate functionalization

Our first choice of functionality to promote water solubility of the [(NHC)CuX] catalyst were sulfonates. As reported by Plenio and coworkers in 2007 the NHC ligand **101** bearing two sulfonates was synthesized starting from commercial 2,6-diisopropylaniline.[205] Selective sulfonation at the para position, subsequent transformation to the Schiff base with dihydroxydioxane and the final cyclization using chloromethyl pivalate yielded the corresponding imidazolium ring **101** (Scheme 23). The synthesis of the sulfonate functionalized NHC ligand **101** was straight forward and the yield for cyclization could be improved, compared to literature yield, by extending the reaction time. The first challenges were faced during complexation with Cu(I). Singlet carbene formation using different bases including NH_3, NaOH and sodium *tert*-pentoxide, in the presence of CuCl did not yield full conversion to the target [(NHC)CuCl] complex **103**. Additionally, the formation of dicarbene Cu(I) complex **104** was detected. Interestingly, deprotonation with NaOH yielded solely the biscarbene species **104**. Use of an excess of sodium *tert*-pentoxide improved the conversion to compound **103** and prevented di-carbene **104** formation, however purification by column chromatography was still necessary. Unfortunately, partially oxidation of Cu(I) to Cu(II) occurred during

this process resulting in a low yield. The alternative route to install Cu(I) via transmetalation was unsuccessful. The silver replacement by copper only resulted in decomposition to the free NHC ligand **101**. Nevertheless, the water soluble [(NHC)CuCl] complex **103** was successfully isolated after 4 steps in an overall yield of 10% and tested towards oligonucleotide alkylation (see section 3.5).



Scheme 23. Synthesis of water soluble NHC-Cu complex 103. a) H_2SO_4, 180°C, 5 h then 1.7 M NaOH, reflux, 1.5 h (43 %). b) dihydroxydioxane, formic acid, 25°C, EtOH (38 %). c) chloromethypyvalate, 45°C, 5 d, (95 %). d) aqu. NH_3, CuCl, H_2O, rt - 60°C, on. e) NaOH, CuCl, MeOH (63 % of **104**). f) sodium *tert*-pentoxide, THF, rt, 1h then CuCl, rt, on. g) Ag_2O, H_2O, reflux, 24 h. h) CuCl or CuSPh.

3.4.2 Ammonium functionalized NHC-Cu(I) catalyst

In order to broaden the scope and test the reactivity and stability of [(NHC)CuX] compounds as catalysts towards oligonucleotide alkylation two different ammonium functionalized [(NHC)CuCl], **105** and **106**, were synthesized (Figure 51). We thought the introduction of a positive charge through ammonium groups would be beneficial since the positively charged catalyst would be coulombically attracted to negatively charged oligonucleotides. The steric demand of the NHCs around the metal center are similar to the sulfonate functionalized [(NHC)CuCl] complex **103**, however, with the ammonium series, we chose as well to investigate different electronic parameters in the NHC skeleton by looking at both saturated and unsaturated backbones. The saturated [(NHC)CuCl] version **105** was synthesized according to a published procedure by Wang and coworkers.[212] The target Cu(I) catalyst was isolated in 9 % yield from a laborious synthesis of seven steps (see steps a-g in Scheme 24). Additionally, the complex stability was very poor in its solid state. Rapid decomposition was detected after isolation at room temperature.

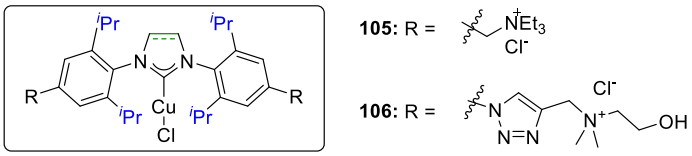
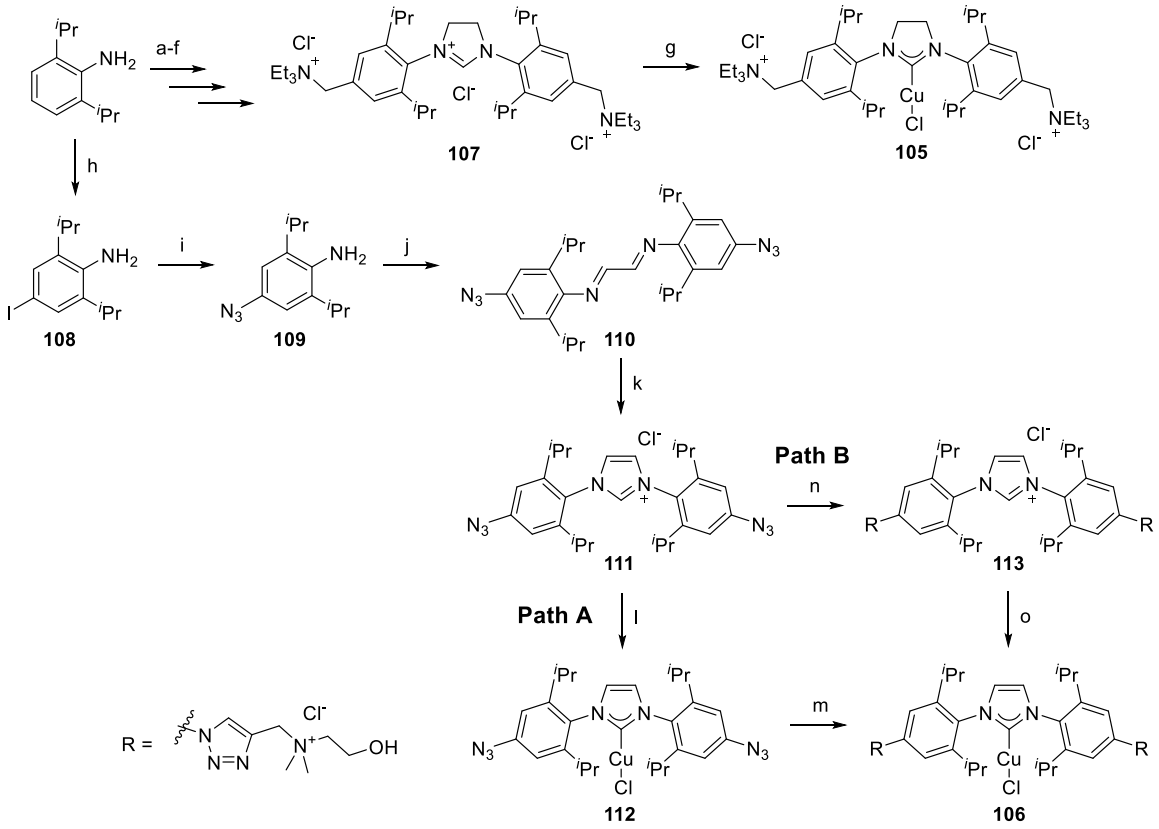


Figure 51. Structure of ammonium functionalized [(NHC)CuCl] complexes 105 and 106.

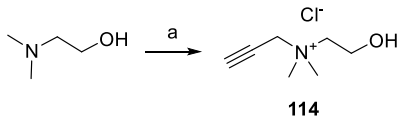
In contrast to complex **105**, the unsaturated ammonium modified [(NHC)CuCl] catalyst **106** showed improved stability towards oxygen and moisture. Purification by column chromatography was possible after post-synthetic functionalization via azide-alkyne cycloaddition. The azide functionalized NHC ligand **111** was synthesized starting from commercial 2,6-diisopropylaniline according to a published procedure by Gautier et al.[217] Iodation at the para position followed by aromatic substitution with sodium azide yielded aniline **109** (Scheme 25). The corresponding Schiff base **110** was formed in the presence of glyoxal. After the final cyclization with paraformaldehyde, the NHC ligand **111** was isolated in 11% over six steps.

There are two possibilities to introduce the water solubilizing ammonium functionality **106**: before (Scheme 25 path A) or after complexation with CuCl (Scheme 25 path B). The ammonium functionalized alkyne **114** was synthesized in one step via alkylation of commercial dimethylaminoethanol (Scheme 25). Both strategies were tested, but the post-complexation functionalization was found to be more efficient. First, the corresponding [$(NHC)_2Ag$] complex was synthesized and then Cu(I) was introduced by

transmetallation. The final click reaction between the [(NHC)CuCl] species **112** with the ammonium functionalized terminal alkyne **114** yielded the target water soluble [(NHC)CuCl] complex **106**.



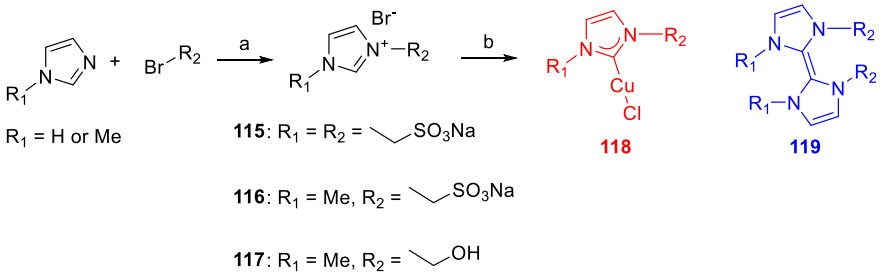
Scheme 24. Synthesis of ammonium functionalized NHC-Cu(I) catalyst 105 and £14 (a) morpholine, CH_2O, EtOH, H_2O, reflux, 14 d. (b) glyoxal, MeOH, HCO_2H cat., rt, 2h (55%). (c) $LiAlH_4$, THF, rt, 24 h (85 %) (d) $HC(OEt)_3$, NH_4Cl, mw, 5 min, 145°C (76 %). (e) ethylchloroformate, MeCN, rt, 1 h (59%). (f) NEt_3, MeOH, rt, 24 h (65 %). (g) KO^tBu, CuCl, MeOH, reflux, 13 h (69 %). (h) Iodine, $NaHCO_3$, Et_2O, 2h, rt (98%). (i) NaN_3, CuI, sodiumascorbate, N,N-dimethylethylenediamine, DMSO, H_2O, 50°C, on (96 %). (j) glyoxal, acetic acid, MeOH, rt, 4 h (41 %). (k) paraformaldehyde, chloro(trimethyl)silane, EtOAc, 70°C, 16 h (74 %). (l) Ag_2O, LiCl, THF, 70°C, 2 h then CuCl, CH_2Cl_2, 37°C, on (69 %). (m) alkyne **114**, $CuSO_4$, sodium ascorbate, DMSO, H_2O, 55°C, 2h (58 %).



Scheme 25. Synthesis of ammonium functionalized alkynes **114**. (a) propargyl chloride, toluene, rt (64%).

3.4.3 Alkyl substituted NHC-Cu(I) catalysts

We also made alkyl *N*-substitution in the NHC ligand since this promised a simple way to access a diverse range of NHC ligands. In one step commercial imidazole or methyl imidazole was alkylated with different bromo alkyl species to yield the target water soluble NHC ligands **115-117** (Scheme 26). Unfortunately complexation reactions of all NHC ligands **115-117** in the presence of CuCl and KO^{t}Bu only yielded the NHC dimer **118** instead of the target Cu(I) complex **119**.



Scheme 26. Synthesis of alkyl substituted NHC ligands **115-117**. (a) DMF, 3 h, reflux (**115**: 40 %), (**116**: 44%), (**117**: 73%). (b) KO^{t}Bu, CuCl, MeOH, reflux, on.

Complete structural elucidation of the isolated compounds after treatment with KO^{t}Bu and CuCl was difficult (Figure 52A). In all cases there was no signal for the imidazole proton in the ^{1}H-NMR spectra, indicating carbene formation occurred, however, the other signals did not show the expected chemical shift, that would result from copper(I) complexation. The reaction mixtures were also analyzed by UPLC-MS, but no reasonable mass could be detected. Finally, comparison of the ^{13}C chemical shifts isolated from HMBC correlation with reported chemical shift of different carbenes (Figure 52B, C and D) was suggestive of dimeric NHC carbenes **119** (Figure 52). Therefore, alkyl substituted NHC ligands showed to be inappropriate to stabilize Cu(I) in the corresponding carbene complexes.

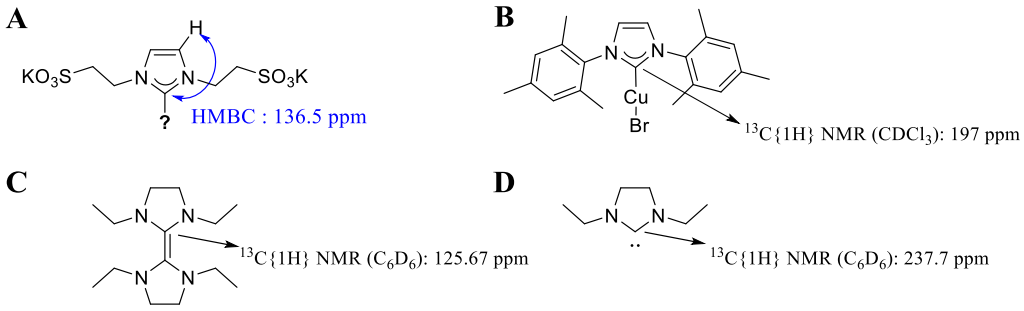


Figure 52. (A) ^{13}C chemical shift from HMBC correlation of the substrate £18 after complexation reaction. Examplex for ^{13}C chemical shifts from literature. (B) Aryl substituted [(NHC)Cu] complex. (C) Alkyl substituted NHC dimer. (D) Alkyl substituted free carbene.

3.4.4 Synthesis of chelating NHC ligands to improve Cu(I) stability

Given the failure of alkyl substitution, we returned to more bulky aryl substitution on the nitrogens for the next series. Complex formation of stability was a constant problem for our copper complexes so here we thought to introduce chelating groups such as hydroxyl, sulfonate and pyridyl on one site of the NHC ligand to improve the stability of the final [(NHC)Cu] complex (Figure 53). The sulfonate functionality in the para position was maintained to ensure water solubility of the final catalyst. Two isopropyl groups were kept in the design to prevent ligand dimerization.

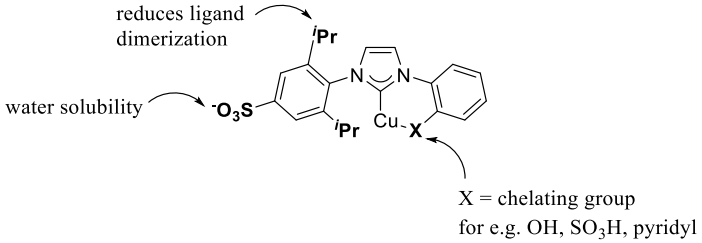
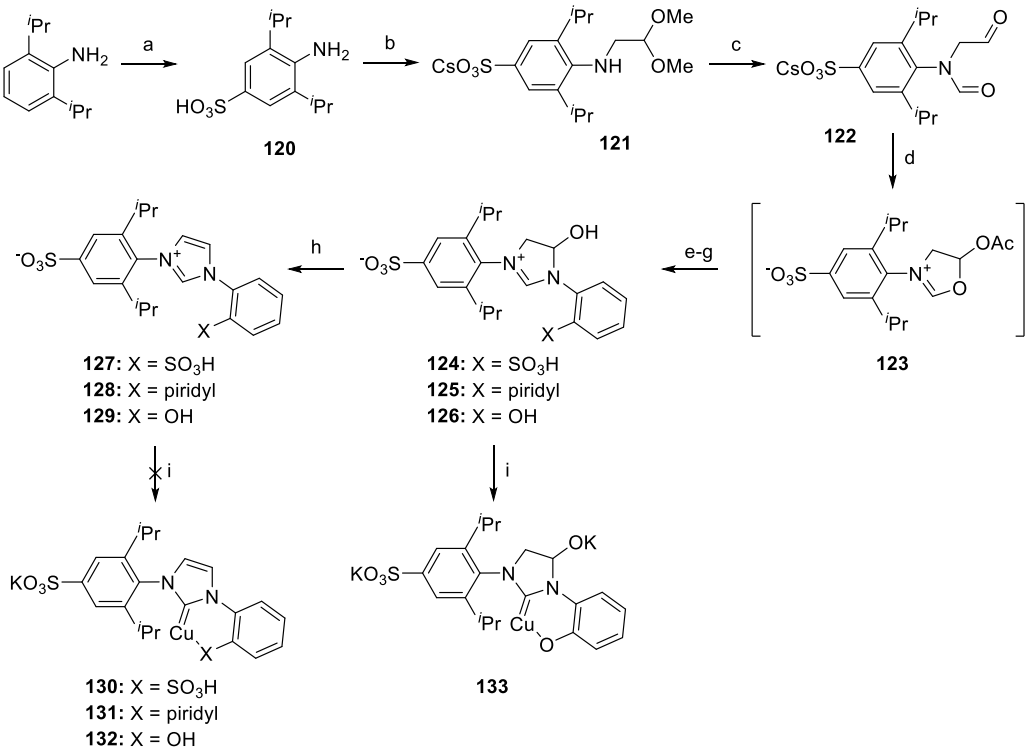


Figure 53. General structure of chelating [(NHC)Cu] complex.

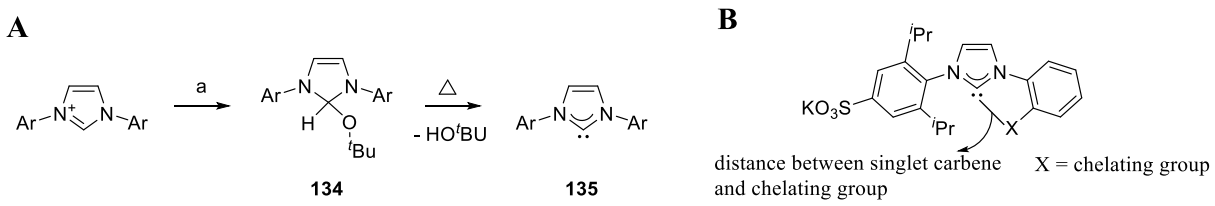
Three different NHC ligands **127-129** were successfully synthesized (overall yields **127**: 4 %, **128**: 2 %, **129**: 6 %) according to an adapted procedure by Fürstner et al.[218] Sulfonation at the para position, starting from commercial 2,6-diisopropylaniline was followed by *N*-alkylation with commercial bromoacetaldehyde diethylacetal under basic conditions yielding the aminoacetal **121** (Scheme 27). Starting from the polar compound **120**, it was necessary to change the solvent system and base for successful *N*-alkylation. Different solvent and bases were tested, however only Cs_2CO_3 in DMF resulted in a decent amount of product **121**. All investigations (different bases and solvents, excess of reagents) to improve the yield were unsucessfull. Either decomposition of the starting material or side product formation like the double alkylated species was observed. The formamide **122** was synthesized with acetic anhydride and further cyclized to the oxazolinium salt **123** using the strong mineral acid HBF_4. Direct reaction with the second amine species, without isolation of **123**, yielded the hemiaminals **124-126**. Final treatment with an excess of HBF_4 at 80°C yielded the target unsymmetric imidazolium salts **127-129**. However, complexation with CuCl was impossible. Neither direct complexation with CuCl in the presence of KO^{t}Bu nor carbene transfer reaction from the silver complexes yielded the target [(NHC)Cu] complexes **130**-**132**. NMR studies revealed that deprotonation had occurred in all cases in the presence of KO^{t}Bu without formation of the Cu carbene bond. One possibility is that only the alkoxy protected compound **134** was present because 70°C was not the required temperature to form the singlet carbene **135** (Scheme 28A). However, for the reported NHCs, 60-80°C was enough to form carbene **135** from the corresponding alkoxy species **134**.[219] It is more

likely that the distance between the singlet carbene and the chelating groups are not appropriate for both interactions, the Cu(I) carbene and Cu(I) coordination to the chelating group, to occur (Scheme 28B).



Scheme 27. Synthesis of unsymmetric water soluble NHC-Cu(I) complexes. (a) conc. H_2SO_4, 5 h, 180°C, 47 %. (b) bromoacetaldehyde diethylacetal, Cs_2CO_3, DMF, 120°C, on (39 %). (c) Ac_2O, formic acid, CH_2Cl_2, reflux, 6 h then formic acid, room temperature, 24 h (51 %). (d) HBF_4*OEt_2, Ac_2O, room temperature, 15 min (quant). (e) 2-aminobenzenesulfonic acid, DMF, rt, 5 h (19 %). (f) 2-aminophenol, DMF, rt, 5 h (33 %). g) 2-(pyridin-2-yl)aniline, DMF, rt, 5 h. (h) HBF_4*OEt_2, DMF, 80°C, 3 d, (**127**: 43 %), (**128**: 21 % over 3 steps), (**129**: 71 %). (i) 3.00 eq. KO^{t}Bu, CuCl, MeOH, reflux, 1 d.

The hemiaminals **124-126** were also treated with KO^{t}Bu and CuCl. Surprisingly, the NHC-Cu(I) catalyst **133** could be isolated and was shown to be bench stable for longer than two weeks. Unfortunately, complex **133** was unreactive in alkylation reactions (see section 3.5).



Scheme 28. (A) Singlet carbene formation **135** by thermal elimination of alkoxy species **134** in the presence of KO^{t}Bu. (a) KO^{t}Bu, THF, 60-80°C. (B) Demonstration of distance between NHC model and the chelating group X.

3.4.5 Synthesis of skeleton functionalized NHC-Cu(I) complex

Increasing the [(NHC)CuX] catalyst stability by chelating groups was unsuccessful due to failure in the complex formation step. It is reported that the NHCs donor properties are influenced by the nature of the heterocyclic skeleton.[220] In 2014 Zhang and coworkers reported a new set of skeleton modified NHC ligands.[221] Inspired from earlier work modification with one or two amino groups in the heterocycle should enhance the electronic donor properties of the ligand. IR-stretching frequencies of the corresponding [(NHC)RhCl(CO)$_2$] complexes **136** and **137** (Figure 54) was consistent with this idea giving a 2 cm^{-1} change of the Tolman electronic parameter (TEP) compared to the non-substituted analogues. The TEP is described as the change of IR stretching frequency of the carbonyl ligands depending on the electron density at the metal center. The crystal structures of the carbene ligands and their corresponding calculated buried volume revealed the steric contribution of the skeleton modification resulting in two air and moisture stabile palladium complexes **138** and **139** (Figure 54).

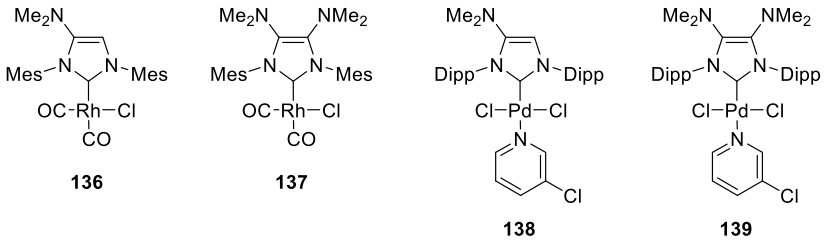
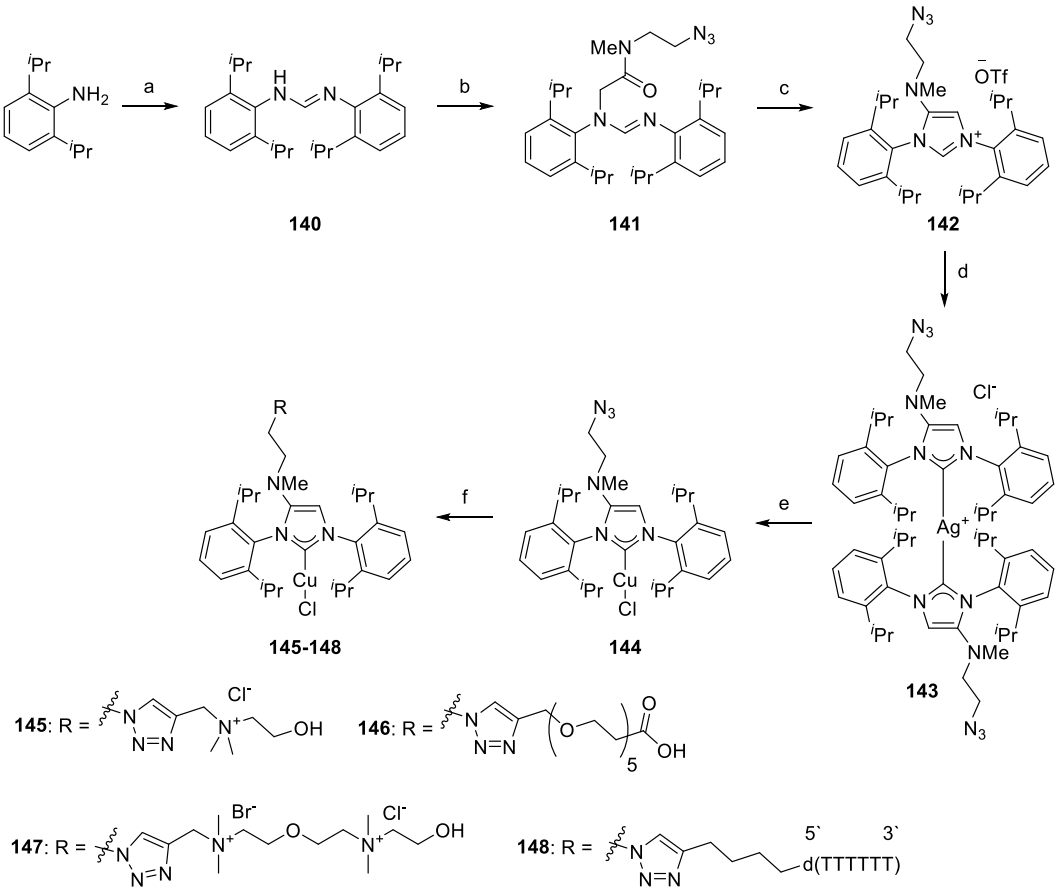


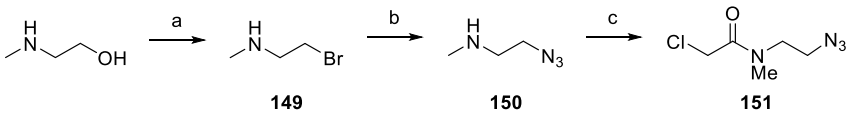
Figure 54. Structures of skeleton modified NHC-metal complexes **136-139**.

To reach similar stability of a water soluble [(NHC)CuX] complex we designed an amino skeleton modified NHC, bearing an azide group **142** (Scheme 29). Post-synthetic functionalization of the [(NHC)CuX] complex **144** with polar groups should increase the water solubility of the final catalyst. The NHC ligand **142** was synthesized by Dr. Rasale starting from commercial 2,6-diisopropylaniline which was transformed to the formamidine **140** using triethylorthoformate. Further alkylation with the azide **151** followed by cyclization in the presence of triflic anhydride yielded the target functionalized NHC **142**. The azide tag **151** was synthesized in three steps. Bromination of commercial 2-(methylamino)ethanol followed by substitution with sodium azide and final reaction with chloro acetylchloride yielded **151** (Scheme 29). Firstly, the [(NHC)$_2$Ag] complex **143** was synthesized using Ag_2O and final transmetalation with CuCl yielded the skeleton functionalized copper(I) complex **144** (Scheme 29). Already the hydrophobic [(NHC)CuCl] **144** showed high stability towards oxygen. A stock solution of **144** in $CDCl_3$ (not degassed) was monitored over three days by recording several proton NMR spectra. No decomposition could be detected. Final click reactions in the presence of functionalized terminal alkynes **114** and **153-155** yielded the skeleton modified [(NHC)CuCl] compounds **145-148**. Unfortunately, introduction of one ammonia group was not enough to increase the polarity to provide water solubility to the final complex **145**.

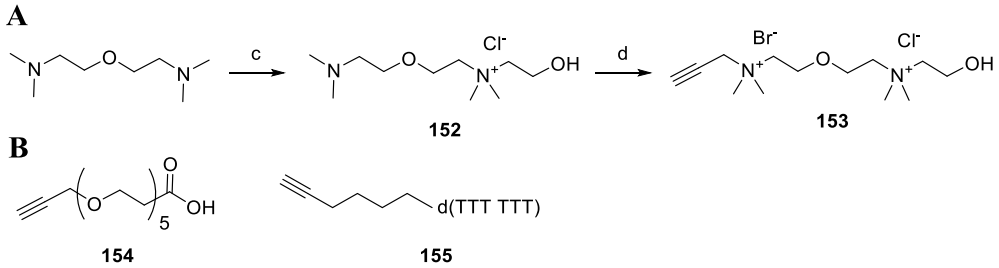
Completion of the CuAAC with the PEG functionalized alkyne bearing a carboxylic acid group **154** was detected by UPLC-MS after 2 h. Unfortunately, the target complex **146** could not be isolated in an amount sufficient enough to test its catalytic activity towards oligonucleotide modification. Click reaction with a di-ammonium functionalized tag **153**, synthesized in two steps starting from bis[2-(N,N-dimethylamino)ethyl] ether via alkylation using chloroethanol and propargylbromide (Scheme 31), yielded the first **water soluble** skeleton modified [(NHC)CuCl] complex **147**. Furthermore, it was also possible to post functionalize the [(NHC)CuCl] **144** complex with a short single-stranded six-mer d(TTTTTT) to introduce the desired water solubility. The CuAAC reaction of a commercial modified T-mer **155** with a hexynyl at the 5`end yielded the water soluble [(NHC)CuCl] catalyst **148**.



Scheme 29. Synthesis of skeleton modified NHC-Cu(I) complexes **145**-**148**. (a) triethylorthoformate, acetic acid, reflux, on (31%). (b) 1042, KI, NEt_3, DMF, 100°C, on. (62%). (c) triflic anhydride, 2,6-lutidine, -78°C (37%). (d) Ag_2O, THF, 70°C, 2 h. (e) CuCl, CH_2Cl_2, 35°C, on, (78%). (f) **114**/ **153**/ **154**/ **155**/, $CuSO_4$, sodium ascorbate, DMSO, H_2O, 25-50°C.



Scheme 30. Synthesis of azide functionalized tag 1042. (a) HBr, 150°C (51 %), 11 h. (b) NaN_3, H_2O, 75°C, 15 h (60 %). (c) chloro acetylchloride, NEt_3, CH_2Cl_2, room temperature, 15 h (63%).



Scheme 31. (A) Synthesis of ammonium functionalized alkyne **153**. (B) Structure of commercial alkynes **154** and **155** (B). (a) propargyl chloride, toluene, rt (64%). (b) chloro ethanol, MeCN, 80°C, on (48%). (c) propargylbromide, MeCN, 80°C, on (62%).

Post functionalization of the [(NHC)CuCl] complex **144**, via a click reaction, proofed to be the most efficient method to introduce the required hydrophilicity to the resulting catalyst. Additionally, skeleton modification showed to be an appropriate strategy to increase the stability of the final [(NHC)CuCl] complex. Using this approach two ammonium functionalized complexes **145** and **147** and a [(NHC)CuCl] catalyst covalent bound to ssDNA **148** were successfully isolated and tested towards oligonucleotide modification.

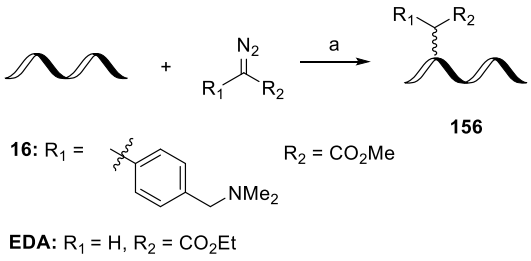
3.5 Alkylation of short single stranded oligonucleotides catalyzed by NHC-Cu(I) complexes

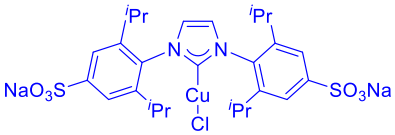
Despite all of the synthetic challenges encountered, we managed to create a series of water soluble NHC-Cu(I) complexes. Our next goal was to look at their catalytic performance in nucleic acid alkylation.

3.5.1 Symmetrical NHCCu(I) catalyst bearing water functionalities in para position to nitrogens towards nucleic acid modification

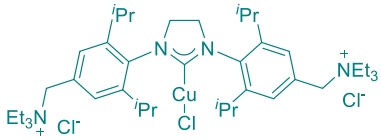
The alkylation potential of water soluble [(NHC)CuCl] complexes **103** and **105** towards oligonucleotide modification was investigated (Table 6). These alkylation reactions represent the first studies of [(NHC)CuCl] catalysts in carbene transfer reactions in the presence of ssDNA. Trinucleotides were used as test substrates to probe the regioselectivity of the alkylation reaction. All three short single stranded DNAs (d(TAT), d(TCT) and d(TGT)) were successful modified using the donor-acceptor stabilized α-diazocarbonyl compound **16** catalyzed by **103** (Table 6, entries 1-6). The diazo species **16** was synthesized in three steps according to literature.[139] Comparable to the reports by Dr. Tishinov using *in situ* reduced copper(I) (by mixing $CuSO_4$ and sodium ascorbate) the nucleobases adenine, guanine and cytosine were targeted. In contrast to that catalytic system also thymine alkylation was detected (Figure 55-57). For example, d(TAT) modification resulted in six mono and one double alkylation (Table 6, Entry 1) instead of two mono alkylation products using *in situ* reduced Cu(I).[140] Tandem mass analysis of these products point to five modification located on thymine and one on adenine (Figure 55). Assuming that the exocyclic amine in adenine is alkylated as reported by Dr. Tishinov, thymine modification most likely occurred at the carbonyl oxygens. Furthermore, d(TGT) and d(TCT) were more efficiently targeted and different amount of alkylated species were detected (Table 6, Entry 3 and 5; Figure 56 and Figure 57). The higher conversion of the oligonucleotides d(TGT) and d(TCT) can`t be directly correlated with higher catalytic efficiency of **103** because a 20 mol% catalyst loading was used instead of 10 mol%. Nevertheless, the conversion of d(TAT) stayed similar in the presence of higher Cu(I) concentration (Table 6, Entry 1). Modification of d(TGT) showed only one main signal for the mono alkylation product and MS/MS analysis revealed the ion for guanine modification (Figure 56A). At this stage no further experiments were done to identify the exact location of guanine alkylation. It was assumed that the alkylation took place at the exocyclic amine as reported for adenine but there is also the possibility of O^6 or N-1 modification (Figure 56B). The addition of a reducing agent (sodium ascorbate) to the [(NHC)CuCl] catalytic system resulted in oxidative damage of the ssDNAs about 25-27 % (Table 6, entry 2, 4 and 6). Without reductant in the reaction mixture more alkylation product was formed.

Table 6. Scope of short nucleic acid alkylation catalyzed by [(NHC)CuCl]





103



105

Entry	catalyst	Asc [mM]	ssDNA	Diazo	Time [h]	Conversion ssDNA [%]	Number of products 156	Oxidative damage [%]
1	**103**	-	d(TAT)	**16**	22	40/ 37[a]	7/ 2[b]	-
2	**103**	20	d(TAT)	**16**	23.5	50	7	25
3	**103**	-	d(TGT)	**16**	22.5	60/ 46[a]	3/ 4[b]	-
4	**103**	20	d(TGT)	**16**	24	72	3	28
5	**103**	-	d(TCT)	**16**	23	62/ 27[a]	6/ 4[b]	-
6	**103**	20	d(TCT)	**16**	24.5	82	6	27
7	**105**	-	d(TCT)	**16**	16	27	6	-
8	**103**	-	d(TGT)	**EDA**	2	74	1	-
9	$CuSO_4$	20	d(TGT)	**EDA**	2	84	1	-
10	Rh_2OAc_4	-	d(TGT)	**EDA**	2	0	0	-
11	no	-	d(TGT)	**EDA**	2	0	0	-
12	**103**	-	d(TCT)	**EDA**	1	52	8	-
					3.5	58		
13	**105**	-	d(TCT)	**EDA**	1	8	8	-
					16	44		

(a) 1 mM [(NHC)CuCl] catalyst **103** and **105**, 5 mM ssDNA, 50 mM diazo, 100 mM MES buffer pH 6, H_2O, room temperature. [a] Conversion of ssDNA after 24h, reaction condition: 0.5 mM $CuSO_4$, 2.5 mM THPTA, 5mM ssDNA, 50 mM diazo, 10 mM sodium ascorbate, 100 mM MES pH 6, 24 h, room temperature. [b] Reported number of alkylated species using *in situ* formed Cu(I) catalytic system.

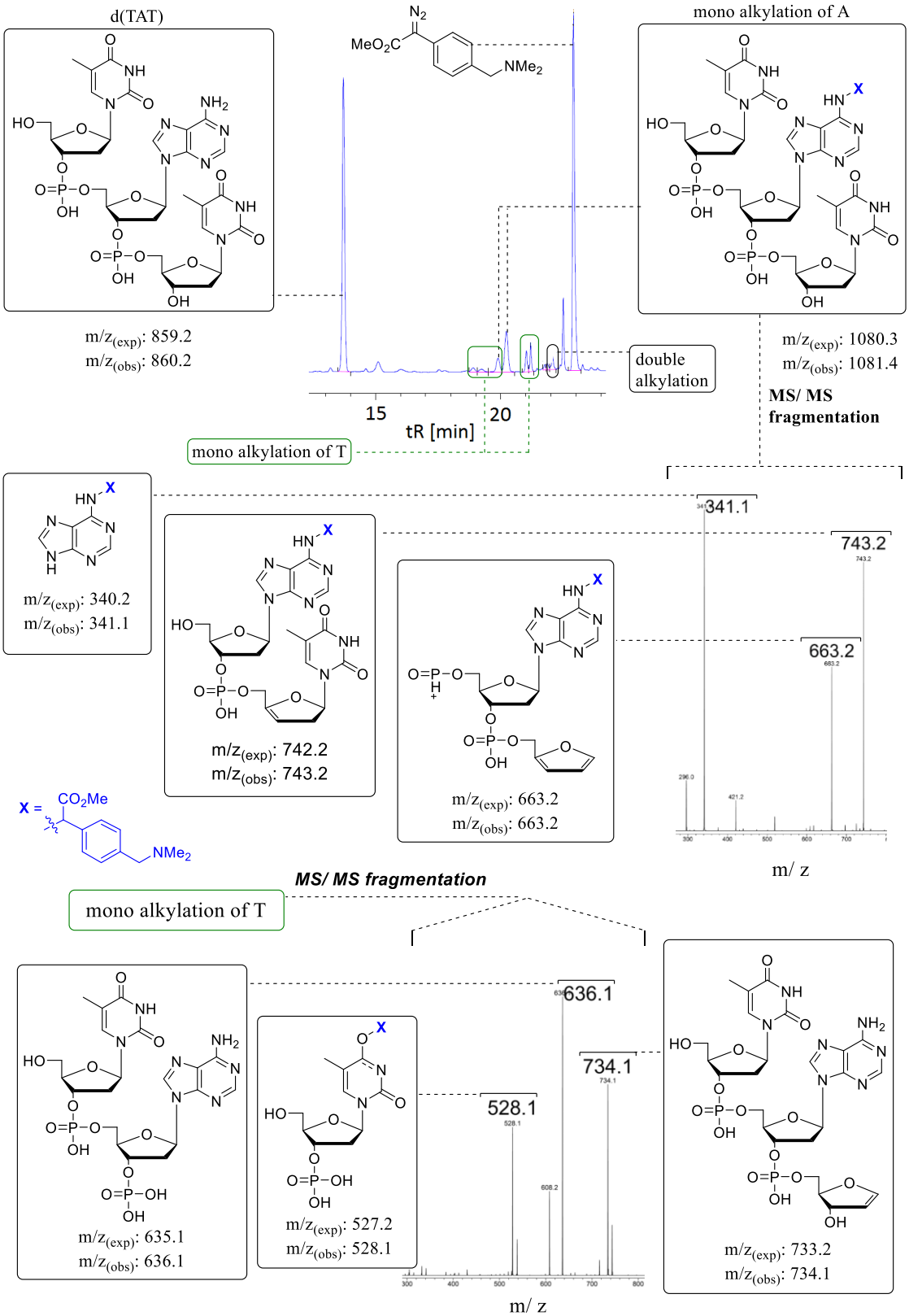


Figure 55. HPLC-MS/MS analysis of d(TAT) alkylation reaction with **16** catalyzed by **103** following the general procedure (Table 6, Entry 1).

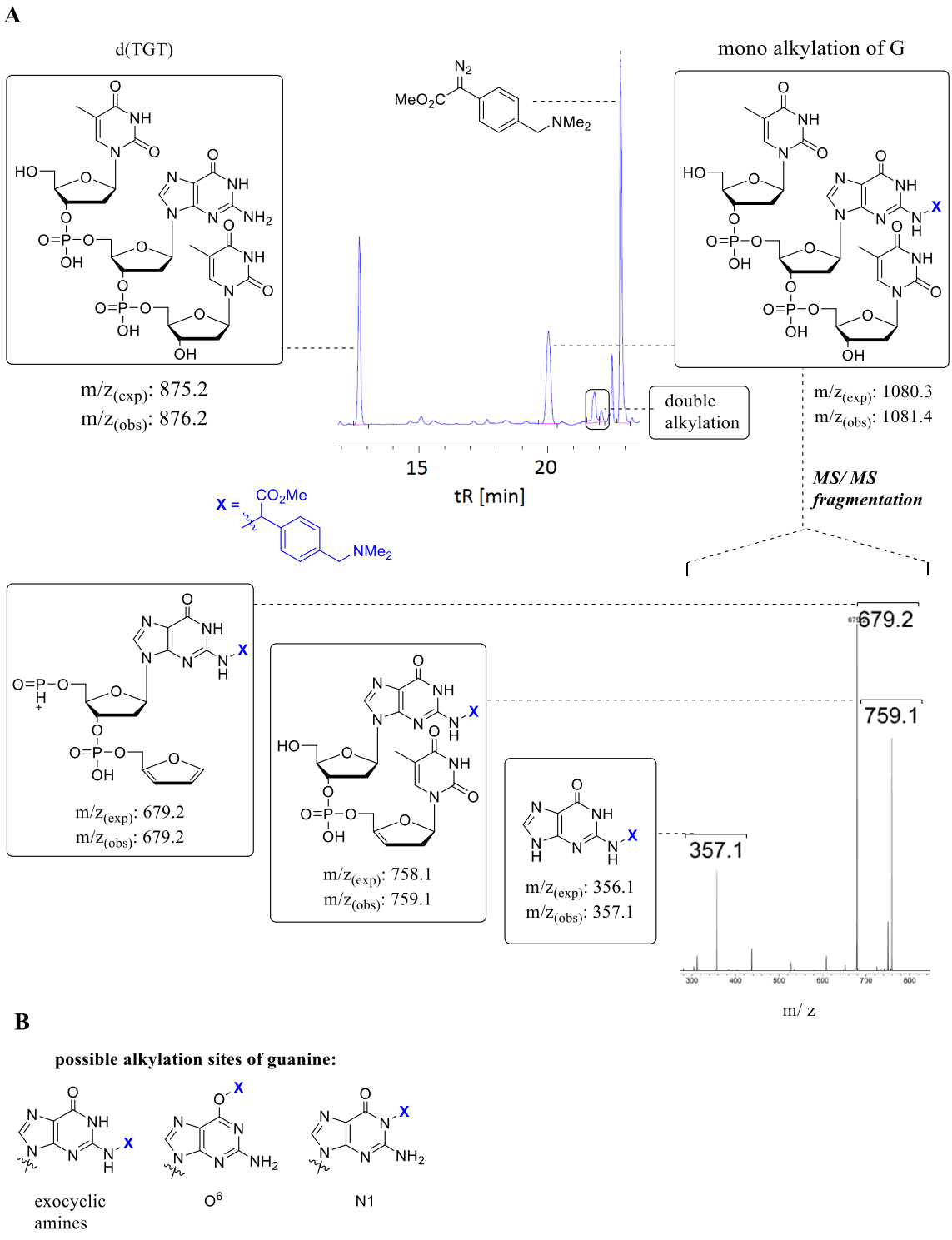


Figure 56. (A) HPLC-MS/MS analysis of d(TGT) alkylation reaction with **16** catalyzed by **103** following the general procedure (Table 6, Entry 3). (B) Possible alkylation sites of guanine.

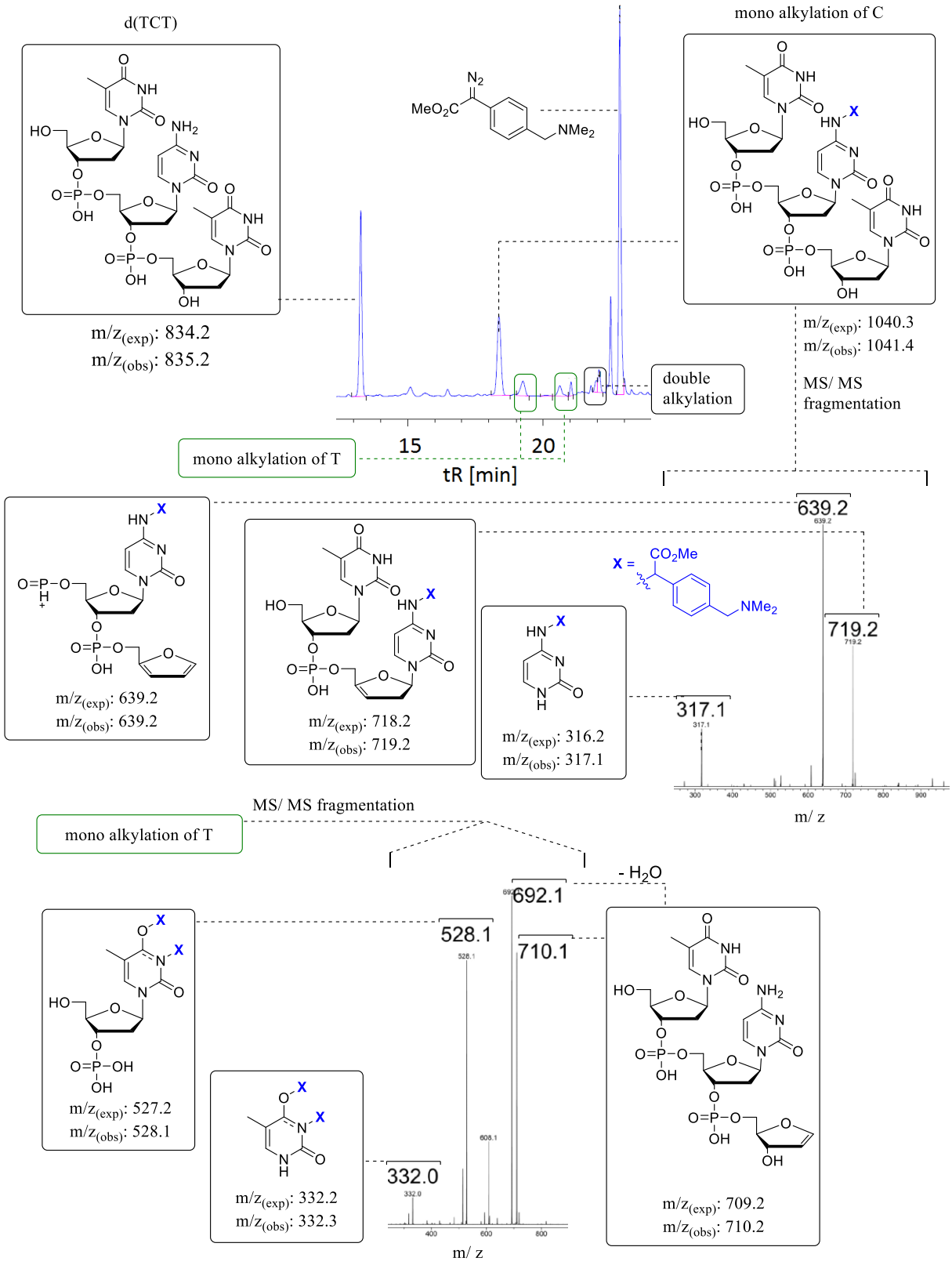


Figure 57. HPLC-MS/MS analysis of d(TCT) alkylation reaction with **16** catalyzed by **103** following the general procedure (Table 6, Entry 5).

In the next step, the effect of the diazo substrate structure on the alkylation process was investigated. For the first time commercial α-ethyl diazoacetate (EDA) was investigated as carbene source for oligonucleotide modification (Table 6, Entry 8-13). The alkylation proceeds much faster using diazoacetate EDA and d(TGT) modification yielded one single peak with high efficiency (Table 6, Entry 8; Figure 59A). ESI-MS analysis of the corresponding isolated peak showed surprisingly the mass for the double alkylation product but unfortunately tandem MS analysis was unsuccessful to reveal the alkylation site. Diazo compounds, especially the unstabilized diazoacetates, can dimerize (Figure 58A). d(TGT) was therefore treated with diethyl fumurate (dimerization product of EDA) in the presence of **103** to exclude diethyl fumurate as the reactive species responsible for double alkylation in d(TGT) (Figure 58B). Under these conditions no conversion of the short ssDNA was detected in presence or absence of **103**. The same alkylation profile for d(TGT) modification with EDA was observed using *in situ* reduced Cu(I), only the conversion of the nucleic acid was higher (Table 6, Entry 9). Surprisingly, testing d(TGT) alkylation in the presence of Rh_2OAc_4 only yielded decomposition of EDA and no alkylation product was formed (Table 6, Entry 10). The carbene transfer reaction was also more efficient towards d(TCT) modification using unstabilized EDA (Table 6, Entry 12). About six mono alkylation products could be isolated and tandem mass fragmentation pointed to about four different thymine and two different cytidine modifications (Figure 59B). The number of alkylation products per nucleobase can`t be exactly determined due to co-elution of the alkylation products.

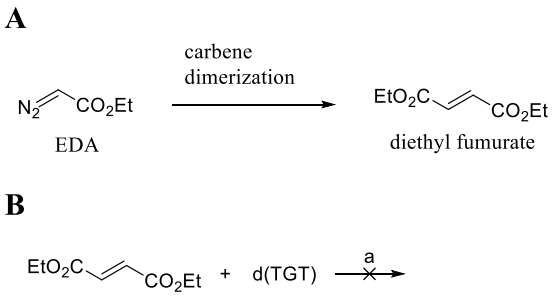


Figure 58. (A) Possible carbene dimerization of EDA. (B) Test reaction of d(TGT) in the presence of diethylfumurate and catalyst **103**. (a) 20 mol % **103**, MES buffer pH 6, 20 v/V % ^{t}BuOH, rt, 4h.

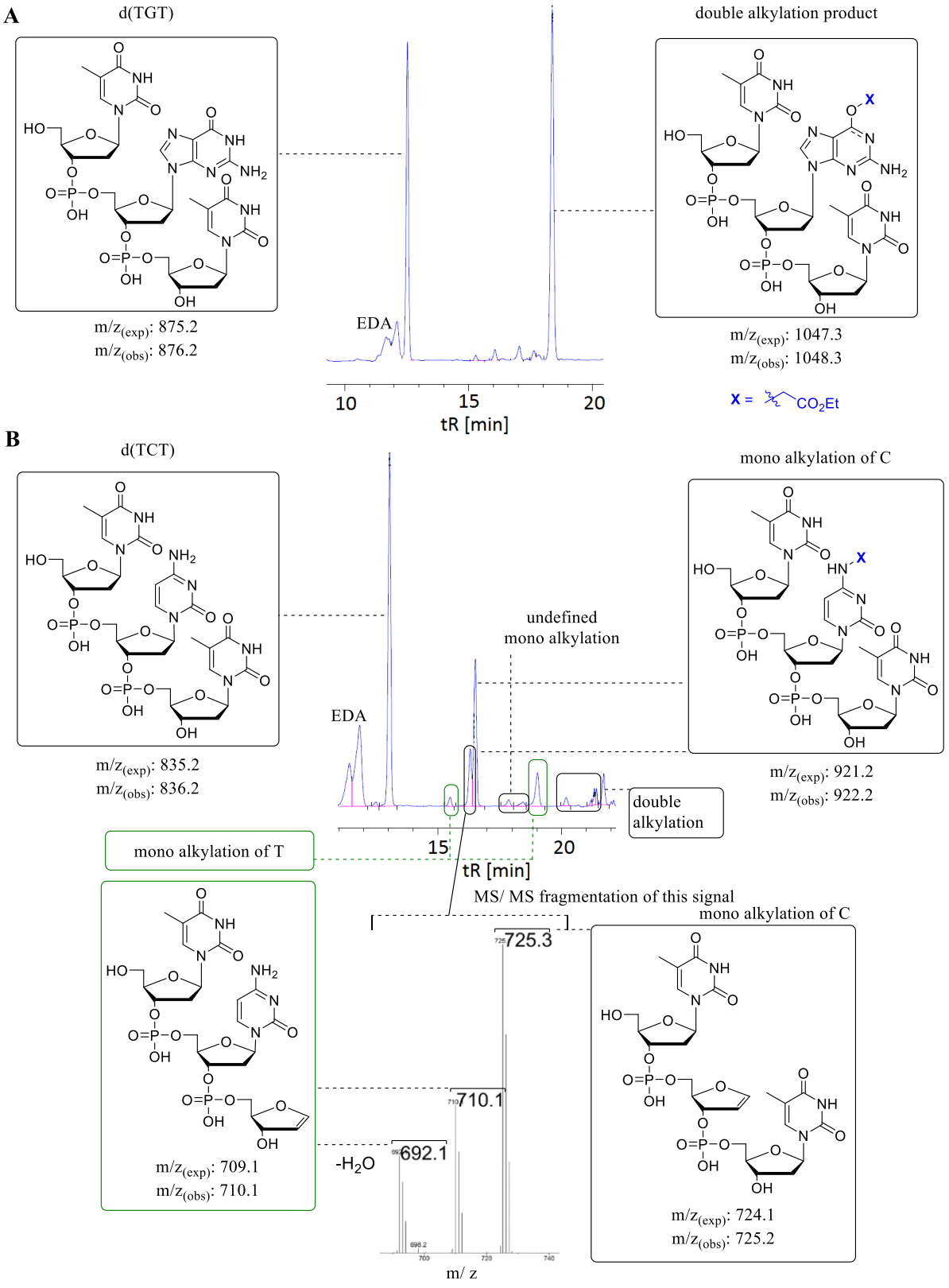


Figure 59. HPLC analysis of d(TGT) (A) and d(TCT) (B) alkylation reaction with 1033 catalyzed by 103 following the general procedure (Table 6, Entry 8 and 12).

The different catalytic activity of the [(NHC)CuCl] complexes **103** and **105** is presented in Table 1 by comparing d(TCT) alkylation with EDA (Table 6: Entry 12 and 13). The modification of d(TCT) was much more efficient catalyzed by **103**. Already after one hour 52% of the ssDNA was converted compared to 8 % conversion using the ammonium functionalized catalyst **105**. The main electronic differences between these two catalysts is the saturation in the NHC skeleton (Figure 60, highlighted with a blue box). The water solubilizing groups (highlighted in green) should not have a big influence on the electron density around the copper center. From a synthetic point of view catalyst **103** is more stable towards oxygen then **105** but more investigations are needed to study their different reactivity towards ssDNA alkylation. In detail, synthesizing their corresponding saturated **156** and unsaturated **157** partners (Figure 60) and testing them towards oligonucleotide modification would reveal whether the NHC skeleton is responsible for the different reactivity of these catalysts **103** and **105**.

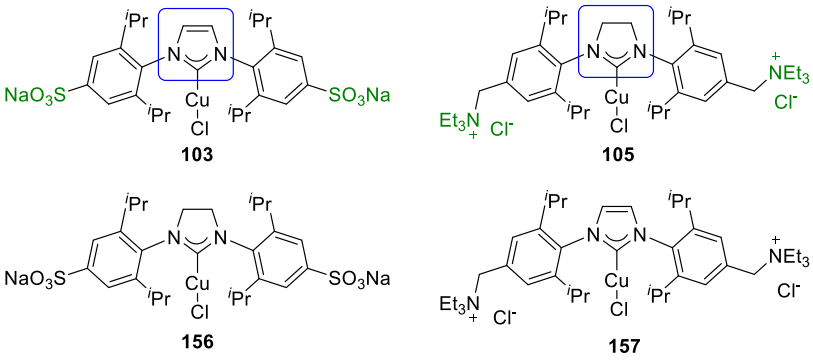
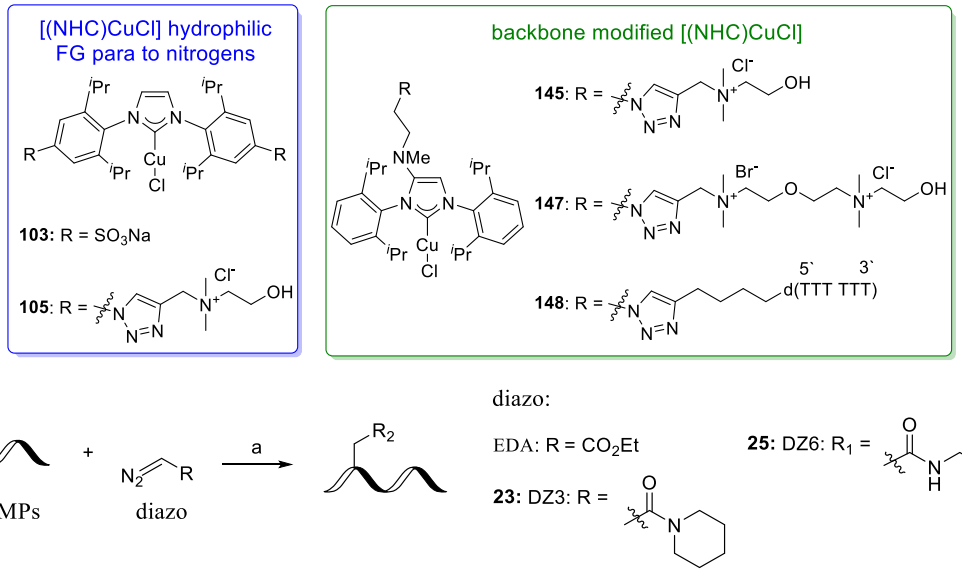


Figure 60. Differences in electronic properties of the [(NHC)CuCl] complexes **103** and **105**. Structures of corresponding complexes **156** and **157**.

The [(NHC)CuCl] catalysts **103** was successful implemented in nucleic acid alkylation and a diazo structure dependent regioselectivity was observed. The accurate alkylation site of guanine in short oligonucleotides, monophosphates and triphosphates with EDA was determined with *in situ* formed Cu(I) (see chapter 2.5.2) and revealed to be the O^6 position. All reactions shown in Table 6 catalyzed by **103** were carried out with a stock solution of the Cu(I) source in water. This stock solution was stored at -20°C but still oxidation to Cu(II) occurred after several days. To investigate the catalyst stability, more water soluble [(NHC)CuCl] complexes were synthesized (see section 3.4). Finally, a backbone modification was the key in finding a storable water soluble [(NHC)CuCl] complex with activity in nucleic acid alkylation.

Table 7. Different catalytic reactivity of [(NHC)CuCl] as catalyst in GMP and dGMP alkylation



Entry	catalyst	MP/ ssDNA/ssRNA	Diazo	Co-solvent [%v/v]	Time [h]	Conversion MP/ oligo [%]
1	**103**	GMP	EDA[a]	DMSO, 20	1/ 1.5	80/ 84
2	**103**[b]	d(ATGC)	**23**[a]	DMSO, 20	3	32
3	**106**	d(ATGC)	**23** [a]	DMSO, 20	1.5	71
4	**106**	d(ATGC)	**23** [c]	Dioxane, 20	1.5	-
5	**106**	d(ATGC)	**25** [a]	DMSO, 20	3	53
6	**106**	11-mer	**25** [a]	DMSO, 20	3	33
7	**145**[d]	GMP	**23**[a]	DMSO, 20	1.5	94
8	**145** [e]	GMP	**23** [e]	DME, 40	0.5	95
9	**147**	dGMP	EDA[a]	DMSO, 20	1.5/ 3.5	40/ 76
10	**147**	dGMP	**23** [a]	DMSO, 20	0.5	94
11	**148**	dGMP	**23** [a]	DMSO, 20	0.5	85
12	**147**	13-mer	**23** [a]	DMSO, 20	2.5	19
13	**147**	UUUUGUUUU	DZ3[a]	DMSO, 20	6.5	-

(a) 1 mM [(NHC)CuCl] catalyst, 5 mM ssDNA/MP, 50 mM diazo, 100 mM MES buffer pH 6, H_2O, room temperature. [a] Stock solution in DMSO. [b] 10 mol% catalyst loading. [c] Stock solution in dioxane. [d] Stock suspension in H_2O. [e] Stock solution in DME = dimethoxyethane. 11-mer = d(TTTTGTTTTCC), 13-mer = d(TATTATGTTTTAA).

The alkylation potential of the successful synthesized [(NHC)CuCl] complexes **103**, **106**, **145**, **147** and **148** is summarized in Table 7. Three diazo substrates EDA and the DAA **23** and **25** were used in this context as the carbene source. The effect of the diazo compound structure in carbene insertions was already discussed

in chapter 2, reminding that compound **23** typically gave the most efficient reactions. Here, we present the different catalytic activity of the [(NHC)CuCl] complexes **103**, **106**, **145**, **147** and **148** towards nucleic acid alkylation, with complex **106** displaying the highest activity. Already after 1.5 hours d(ATGC) was efficiently alkylated (Table 7, Entry 3). This conversion number was the highest observed so far for d(ATGC) alkylation with **23**. Unfortunately, the stability of compound **106** towards oxidation was low. Comparable to catalyst **103**, the catalyst stock solutions of **106** became inactive in a short time.

As already mentioned backbone modification of the NHC ligands enhanced the final Cu(I) complex stability. This skeleton functionalization yielded catalysts **144**, **147** and **148**, which could be reused as a stock solution for nucleic acid alkylation reactions. The introduction of one ammonium functionality in the skeleton was insufficient to give good water solubility. Nevertheless, compound **145** showed high catalytic activity towards GMP alkylation in a heterogenous system (Table 7, Entry 7) or homogenous reaction mixture using dimethyxyethane (DME) as co-solvent (Table 7, Entry 8). Finally, the click reaction of a di-ammonium tag gave the required water solubility to the [(NHC)CuCl] complex. Guanine alkylation catalyzed by **147** was slower (Table 7, Entry 9) compared to a non skeleton functionalized catalyst **103**. An additional two hours reaction time was necessary to reach a similar conversion number of the monophosphate. However, catalyst **147** could be stored as a stock solution in water and re-used for oligonucleotide alkylation. The potential of further tailoring of the [(NHC)CuCl] complexes is demonstrated by successful dGMP alkylation catalyzed by the T-mer modified Cu(I) species **148** (Table 7, Entry 11). Tandem mass fragmentation of all monoalkylation products isolated after carbene transfer reaction catalyzed by **103**, **106**, **145**, **147** and **148** showed the expected ions for guanine modification (two representative examples: Figure 62 and Figure 63). Further investigation toward ribonucleic acid alkylation in a reductant free Cu(I) catalytic system was done. The test substrate UUUUGUUUU was treated with DAA **23** in the presence of **147** (Table 7, Entry 13). HPLC analysis of the reaction mixture showed a small new signal after 6.5 hours (Figure 61B, highlighted with a box) but the expected mass for the ssRNA alkylation product could not be detected. While monoribonucleotide react efficiently (GMP), the short ssRNA did not prove to be a viable substrate for carbene insertion. Unfortunately, the purity of the commercial RNA is low (Figure 61B). These impurities could have an influence on the system and diminish the catalytic activity as we have seen previously with the coordinating ligand THPTA (Chapter 2.5.3).

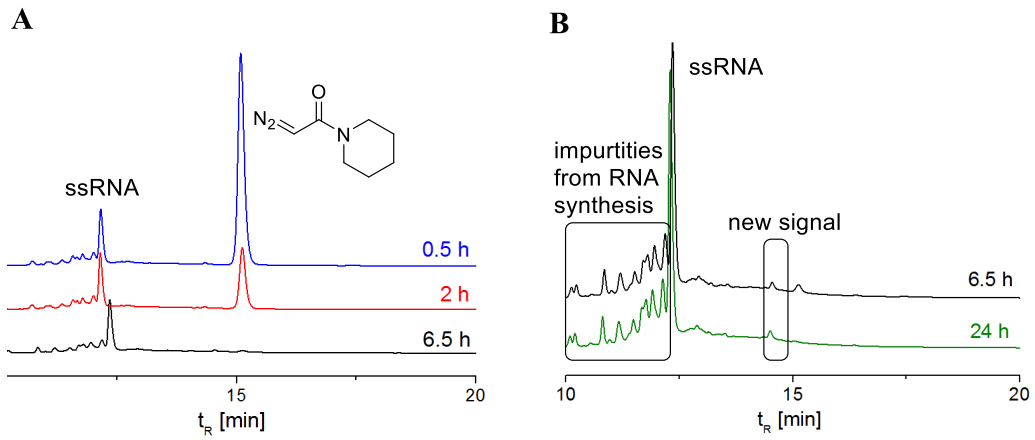


Figure 61. HPLC analysis of UUUUGUUUU alkylation reaction with **23** catalyzed by **147** following the general procedure (Table 7, Entry 13). (A) HPLC analysis of the reaction mixture after 0.5 h (blue), 2 h (red) and 6.5 h (black). (B) HPLC analysis of the reaction mixture after 6.5 (black) and 24 h (green).

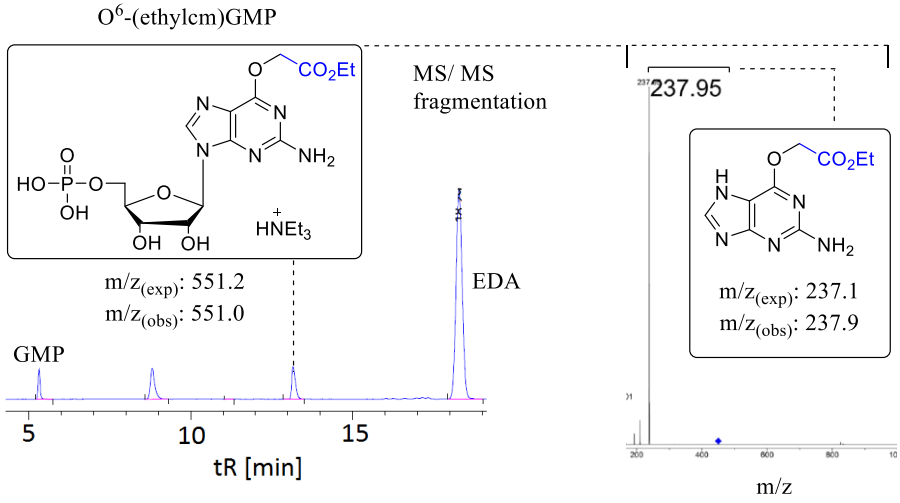


Figure 62. HPLC-MS analysis as representative examples of guanine modification. GMP alkylation with EDA catalyzed by **103** (Table 7, Entry 1).

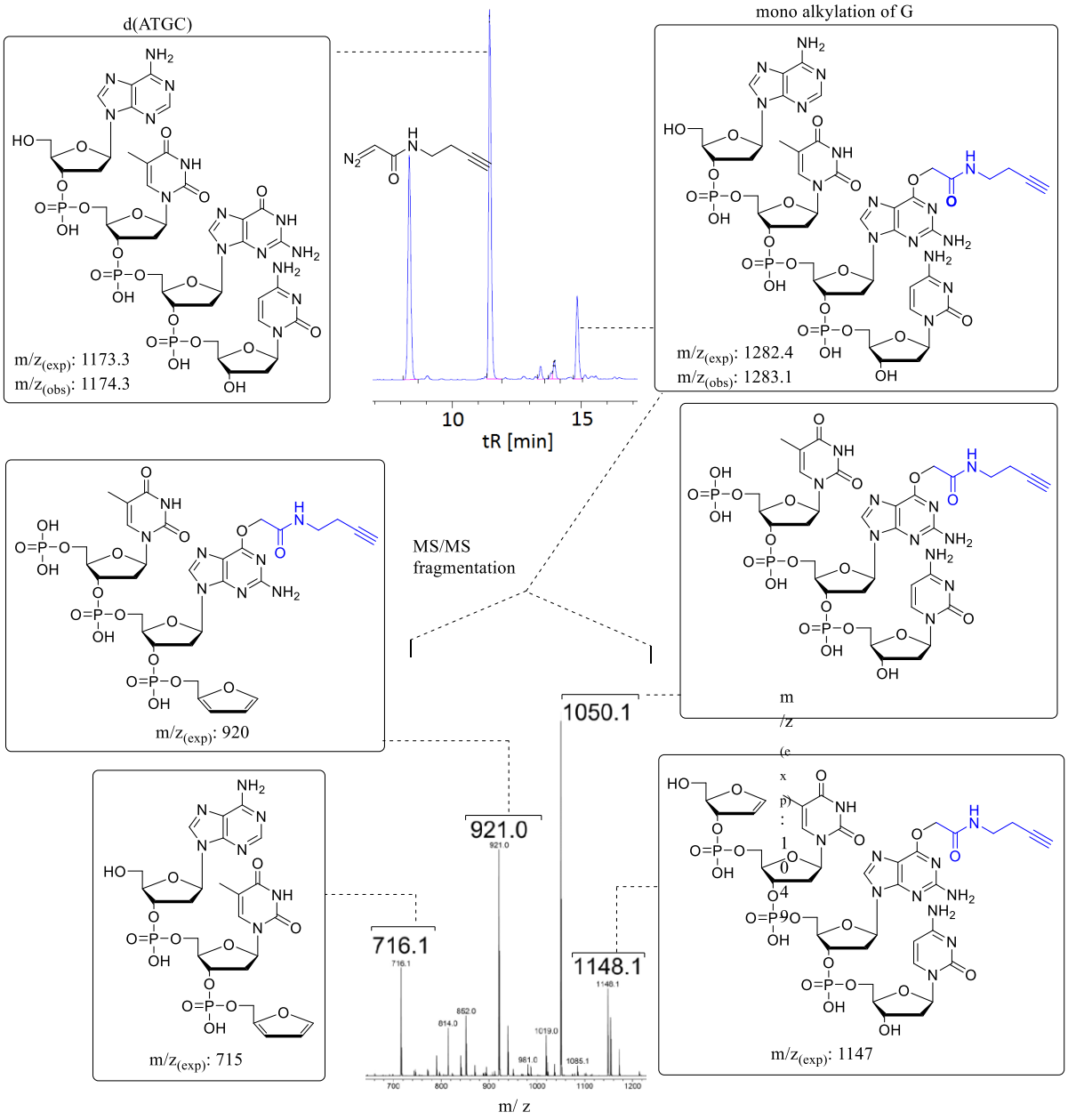


Figure 63. HPLC-MS analysis as representative examples of guanine modification: d(ATGC) alkylation with **25** catalyzed by **106** (Table 2, Entry 5).

The limitation of this methodology is presented by the carbene insertion reaction of medium sized oligonucleotides. Already nucleic acids with 11 to 13 bases showed lower conversion of the ssDNA (Table 2, Entry 6 and 12). For example modification of the 11-mer with the alkyne functionalized diazo **25** yielded 33 % conversion to the mono alkylation product after three hours. Modification of the ssDNA with higher adenine content gave 19 % conversion after 2.5h, resulting in two mono alkylation products (Figure 64). The drop in conversion is likely related to sequestering of copper in the oligonucleotides.

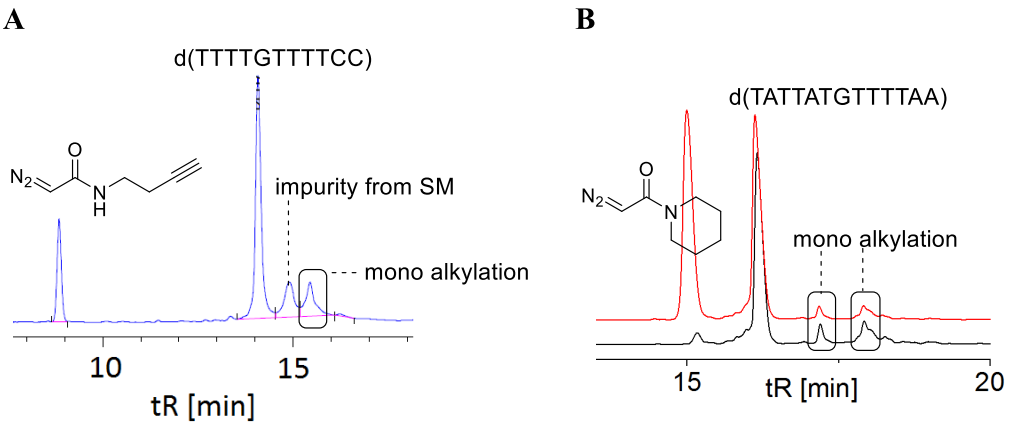


Figure 64. (A) HPLC analysis of d(TTTTGTTTTCC) alkylation reaction with **25** catalyzed by **106** following the general procedure (Table 7, Entry 6). (B) HPLC analysis of d(TATTATGTTTTAA) alkylation reaction with **23** catalyzed by **147** following the general procedure (Table 2, Entry 12).

3.6 Confirmation of O^6-G modification in larger oligonucleotides

3.6.1 Principle of templated nucleic acid modification catalyzed by covalent bound NHC-Cu(I) system

Our method for O^6-G modification in ssDNA catalyzed by *in situ* formed Cu(I) or stabilized as NHC complex becomes more inefficient the longer the oligonucleotide is. Therefore, selective modification of longer ssDNAs still remains a challenge. A template directed strategy could overcome this barrier. The idea was to synthesize a [(NHC)CuX] complex, functionalized with an aldehyde or azide group, to form a covalent bond to a guiding sequence via oxime or triazole, respectively (Figure 65). Linkers with different length can be installed between the catalyst and the guiding sequence to ensure flexibility of the system. The guiding sequence is designed to be complementary to the target sequence (10 base pairs). Hybridization of the sequences would then bring copper(I) close to a specific guanine. Final addition of the carbene transfer reagent, the diazo substrate, would result in selective guanine alkylation.

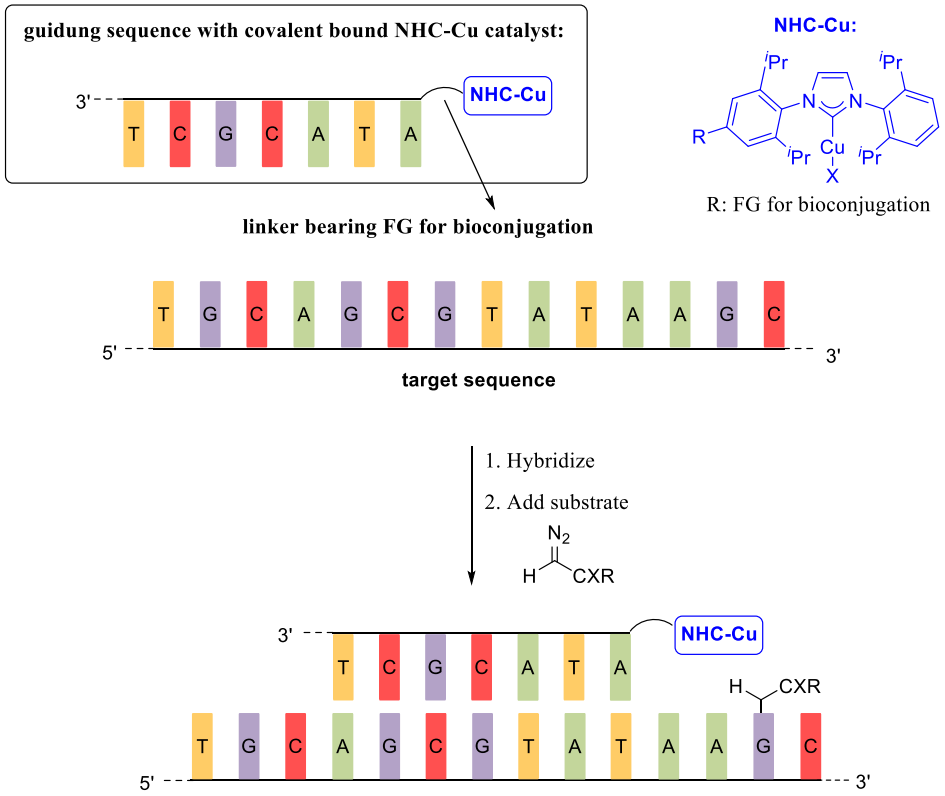
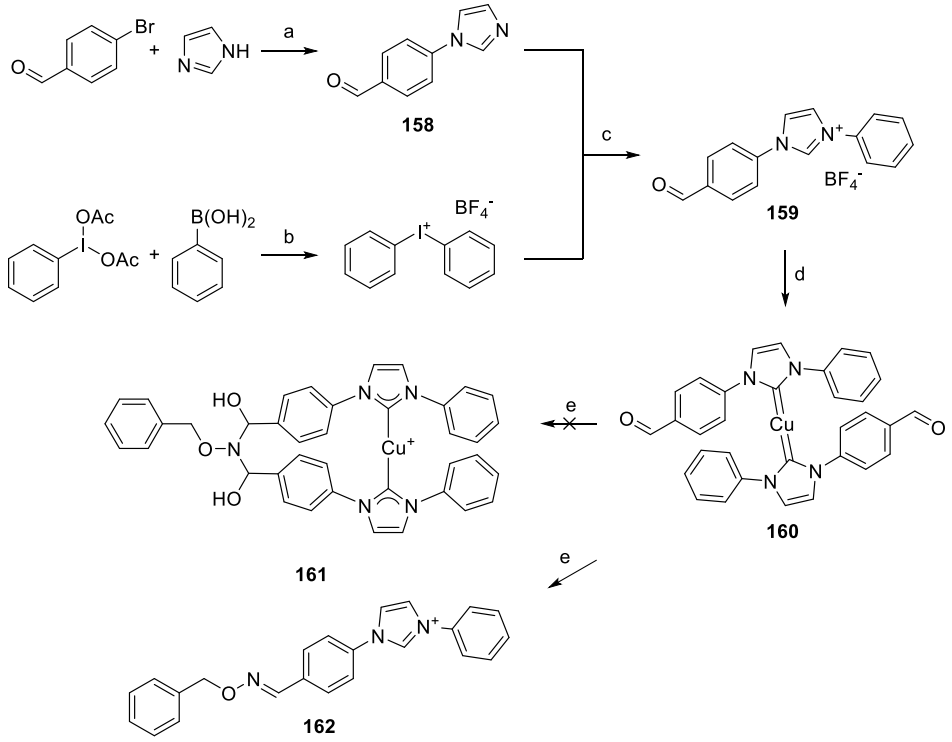


Figure 65. Principle of templated directed oligonucleotide modification

The synthesis of an aldehyde functionalized NHC was reported by Gao and coworkers in 2013.[222] According to the literature direct aryl quaternization of the aldehyde substituted imidazole **158** with the dibenzyl iodonium salt yielded the aldehyde modified NHC ligand **159** (Scheme 32). Complexation with

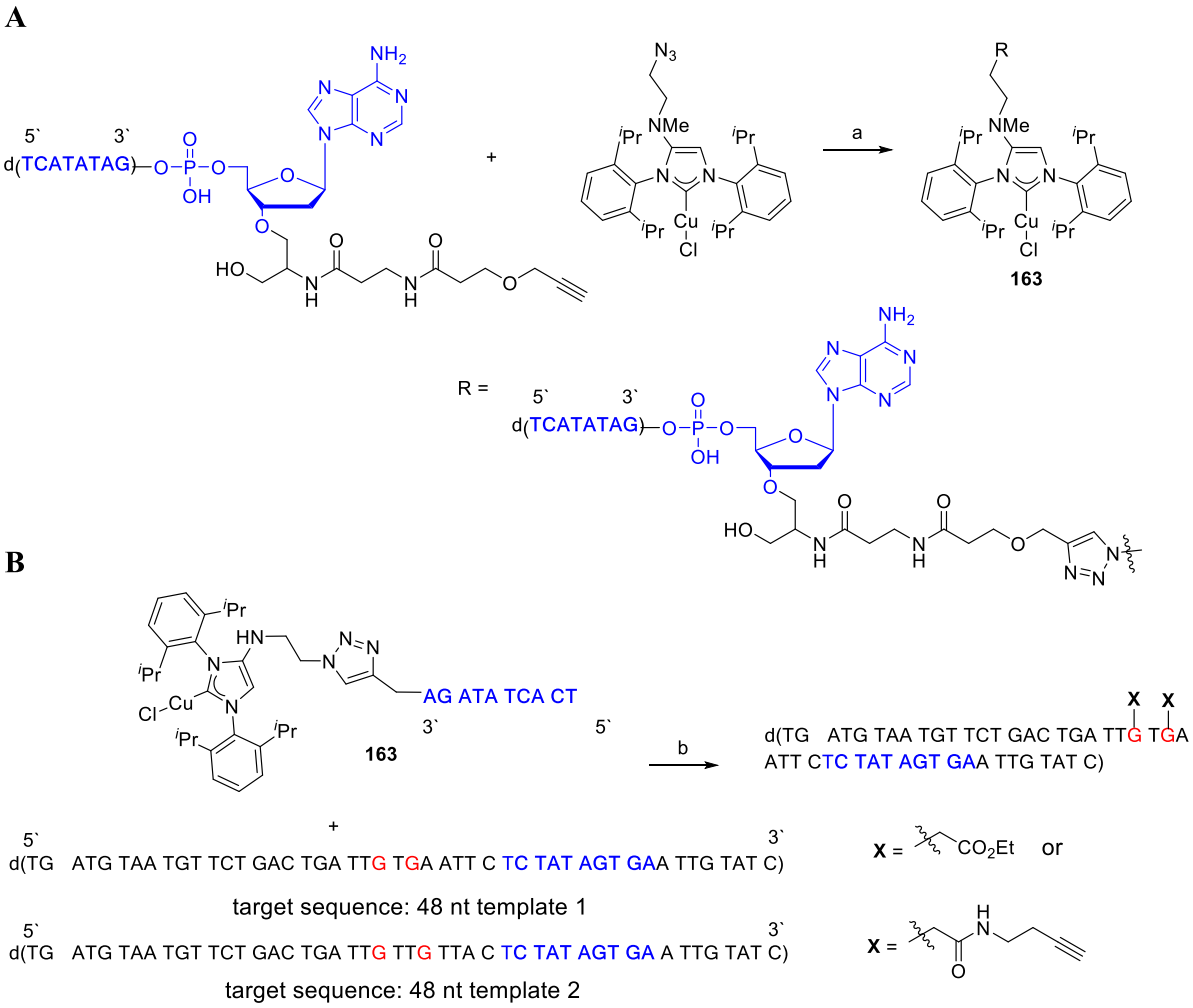
Cu(I) was successful but the dimeric [(NHC)$_2$Cu] catalyst **160** was formed. A test reaction with phenyl hydroxylamine showed that the complex stability was too low to form the corresponding hemiaminal **161**. Instead only oxime formation to the free ligand **162** was detected.



Scheme 32. Synthesis of aldehyde conjugated NHC-Cu(I) complex 24. (a) CuI, Cs_2CO_3, DMF, 120°C, 40 h (54 %). (b) $BF_3 \cdot OEt_2$, CH_2Cl_2, reflux 1 h (66 %). (c) $Cu(OAc)_2 \cdot H_2O$, DMF, 100°C, 4 h (55 %). (d) aq. NH_3, CuCl, H_2O, rt - 60°C, on. (e) benzylhydroxylamine, H_2O, MeCN, rt, 24 h.

As reported in section 3.5 a skeleton modification of the NHC ligand introduced more stability to the NHC Cu(I) bond. Therefore, the skeleton modified [(NHC)CuCl] complex **144** was clicked by Dr. Dnyanesh Rasale to a commercial ssDNA 10-mer bearing an alkyne functionality at the 3`end (Scheme £mA). Then he tested the resulting guiding catalyst **163** towards template directed oligonucleotide modification using two different 48 nucleotides (nt) ssDNA templates and two different diazo substrates, EDA and the alkyne functionalized diazoacetamide **25** (Scheme 33B). Ethanol precipitation of the ssDNAs from the alkylation reaction mixture and reinjection to HPLC showed a small shift of the compounds (Figure 66A and B). MALDI analysis of the reaction mixtures showed a higher mass than the starting templates 1 and 2. Unfortunately, the precision of the MALDI is not high enough to validate whether mono or double alkylation occurred. In the next step a 68 nt template was modified guided by catalyst **163** and the carbene source **25** (Scheme 34). After 6 hours the DNAs were precipitated with ethanol and further functionalized via click reaction with Sulfo-Cyn5-azide. Afterwards a better separation of the signals was possible in the

HPLC (Figure 66C). Additionally, final analysis by gel electrophoresis confirmed the modification of the 68 nt template. Further investigations are ongoing by Dr. Dnyanesh Rasale to show whether this alkylation occurred template directed.



Scheme 33. Templated directed oligonucleotide modification. (A) Synthesis of guiding sequence **163** (B) Alkylation reaction of 48 nt sequence 1 or 2 with EDA or **25**, repectively, catalyzed by the guiding sequence **163**. (a) $CuSO_4$, sodium ascorbate, DMSO, H_2O, room temperature, 2h. (b) EDA or **25**, MES buffer pH 6, 31°C, 6h.

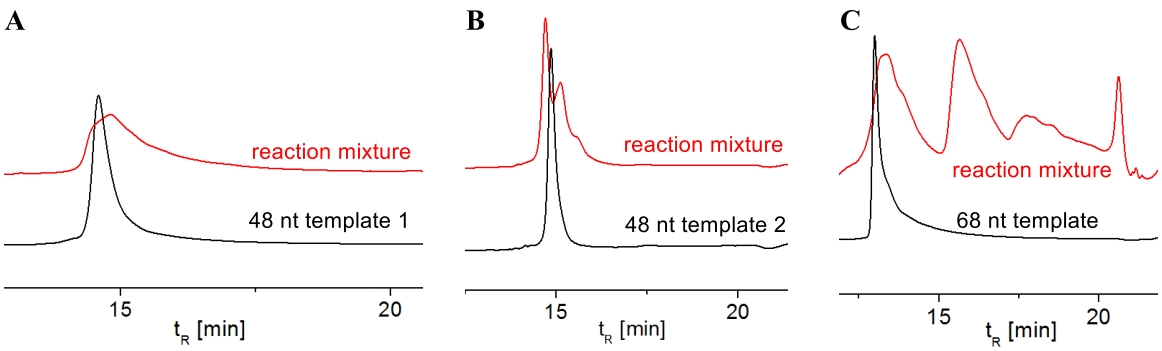
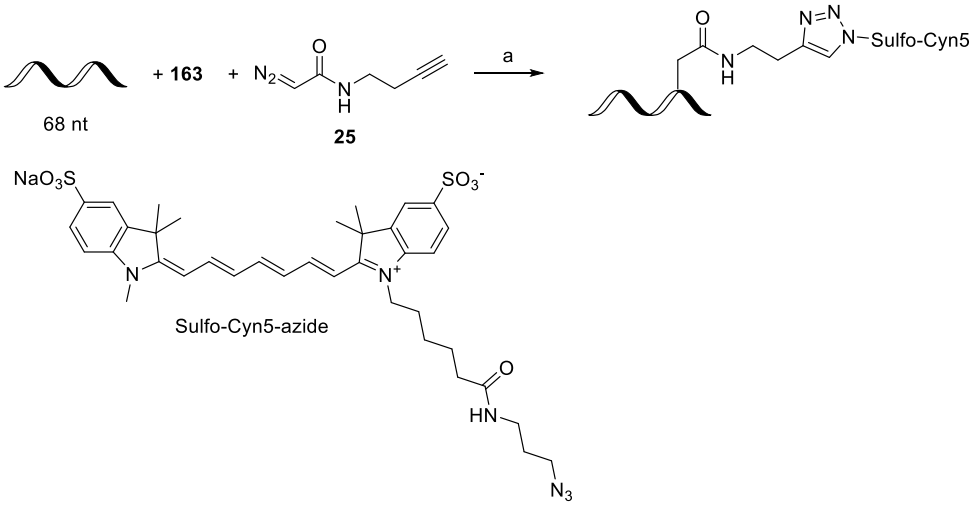


Figure 66. HPLC analysis of the templated directed oligonucleotide alkylation reactions. (A) HPLC analysis of reaction mixture after EtOH precipitation: template 1 with **25** catalyzed by **163** (red); HPLC analysis of template 1 (black). (B) HPLC analysis of reaction mixture after EtOH precipitation: template 2 with EDA cataylzed by **163** (red); HPLC analysis of template 2 (black). (C) HPLC analysis of alkylation reaction mixture after Sulfo-Cy5-azide treatment (red); HPLC analysis of 68 nt template (black).



Scheme 34. Templated directed oligonucleotide alkylation using a 68 nt template and diazo acetamide **25** guided by the catalyst **163** and subsequent functionalization of the alkylation product via click reaction with Sulfo-Cy5-azide. (a) MES buffer pH 6, 31°C, 6 h, precipitation of the DNA with EtOH then Sulfo-Cy5-azide, $CuSO_4$, ascorbate, DMSO, H_2O, room temperature, 4 h.

3.7 Conclusion and future direction

A reducing agent free approach for copper(I) catalyzed carbene insertion reactions targeting nucleic acids has been developed. The method allows an efficient alkylation of short ssDNAs with α-diazo carbonyl compounds maintaining the selectivity for the O^6 position in guanine. The catalyst copper(I) is stabilized as a NHC complex and utilized in an aqueous medium at ambient temperatures.

Several water soluble [(NHC)CuCl] complexes were synthesized and showed to be active for DNA alkylation. The structural design of the NHC ligand highly influences the resulting reactivity and stability of the copper(I) catalyst. A skeleton functionalization of the NHC heterocycle with σ- donors such as tertiary amines, introduces higher stability to the NHC-Cu bond. Most importantly, this skeleton alteration provides access to further functionalization. A broad substrate scope is post-synthetic attachable to the NHC-Cu catalyst to fine-tune its polarity.

Preliminary experiments for a template directed alkylation method for longer oligonucleotide modification were also described. To execute our approach a post-synthetic functionalization of an [(NHC)CuCl] catalyst was essential. Bio-conjugation of a short ssDNA via click reaction to the [(NHC)CuCl] skeleton provides an efficient and easy synthesis of the guiding catalyst.

Future work on NHC-Cu(I) system should address several questions:

(a) Template directed alkylation of longer oligonucleotide maintaining the selectivity for O^6-G adducts. This requires a detectable difference between the alkylation products and the unmodified oligonucleotide and could be realized by:
   - Labelling with specific tags (e.g. biotin and Sulfo-Cy5 dye)
   - digestion to shorter oligonucleotides or nucleotides using enzymes without affecting the modification site
   - Primer extension experiments with enzymes which do not tolerate the modification. This would lead to an extension stop at the modification sites and give the corresponding truncated fragments. Separation of these fragments by length with electrophoresis would give hints about the modification sites.

(b) Diazo-substrate variation to investigate whether the alkylation selectivity could be changed
   - diazo substrates with alkyl groups in the α-position
   - electron withdrawing groups in the α-position
   - or silyl functionalization in the α-position

(c) An in-depth study on the electronic properties of the [(NHC)CuCl] complexes to improve modification rate.

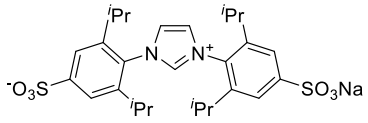
3.8 Experimental

3.8.1 General Experimental Information

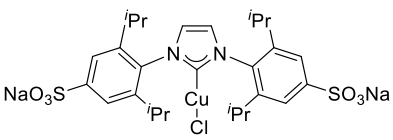
All reagents and solvents used were of analytical grade. Buffers were prepared with ultrapure water. All chemicals were purchased from Sigma-Aldrich, Alfa Aesar or Acros and used as received. Solid-phase oligonucleotide synthesis was carried out on an Expedite 8909 nucleic acid synthesis system (PerCeptiveBiosystems) on 1µmol CPG columns using standard phosphoramidite chemistry with 0.3 M 5-benzylthio-1-H-tetrazole as activator. Purification was carried out on a preparative Shimadzu UFLC system with a Gemini-NX 5 µm C18 21.2 x 250 mm (Phenomenex) column using 100 mM triethylammonium acetate (TEAA pH 7.25)/acetonitrile (MeCN) gradients as mobile phase. Elution was carried out at a flow rate of 20 mL/min monitored at 254 nm using **Method A:** 0-15 % MeCN in 30 min, 15-95 % MeCN in 10 min, 95 % MeCN for 5 min for short ssDNA and **Method B**: 0-12% MeCN in 25 min, 12-16 % MeCN in 10 min, 16-95 % MeCN in 5 min, 95 % MeCN for 5 min for longer ssDNA. Alkylation reactions were analyzed on an analytical Shimadzu UFLC system or Agilent 1100 LC system with an Eclipse XDB-C8 4.6 x 150 mm (Agilent) column using 100 mM TEAA/MeCN gradient. Elution was carried out at a flow rate of 1 mL/min using **Method C**: 0-16 % MeCN in 18 min, 16-80 % MeCN in 5 min, 80 % MeCN for 3 min with peak detection at 254 nm. Purification was carried out on the preperative Shimadzu UFLC system as described above using **Method D**: 0-16 % MeCN in 20 min, 16-20 % MeCN in 5 min, 20-95 % in 1 min, 95 % MeCN for 5 min. Aqueous product fractions were freeze dried on a Christ Alpha 2-4 LDplus flask lyophilizer at 0.1 mbar. ^{1}H, ^{13}C and ^{31}P-NMR, HMQC and HMBC spectra were acquired on a BrukerAvance (400, 500 or 600 MHz proton frequency) spectrometer at 298.15 K. Chemical shifts relative to TMS were referenced to the solvent's residual peak and are reported in ppm. ESI MS spectra were measured on a Bruker Esquire3000plus mass spectrometer by direct injection in positive polarity of the ion trap detector.

3.8.2 Chemical synthesis

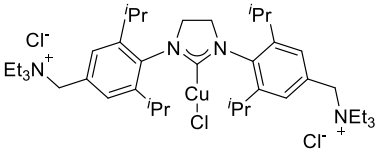
Sulfonate functionalized NHC ligand (101). Compound **101** was synthesized according to a published procedure.[205] White solid (330 mg, 95 %). The analytical data were in agreement with the literature.



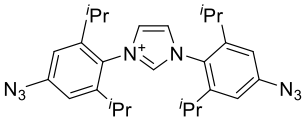
Sulfonate functionalized [(NHC)CuCl] complex (103). A suspension of immidazolium salt **101** (100 mg, 159 µM) and sodium terpentoxide (191 µL 2.5 M in THF) were stirred for 60 min at room temperature. Then fresh CuCl (15.7 mg, 159 µM) was added and the mixture was stirred over nigth at room temperature. The solvent was evaporated under reduced pressure and the residue was purified by prep RP-HPLC using method A. The corresponding fractions were lyophilized to yield **103** as a white solid. ^{1}H-NMR (400 MHz, D_2O) δ 7.58 (s, 4H), 7.33 (s, 2H), 2.30 (dt, *J* = 13.5, 6.6 Hz, 2H), 0.95 (d, *J* = 6.7 Hz, 12H), 0.77 (d, *J* = 6.8 Hz, 3H). ^{13}C-NMR (101 MHz, D_2O) δ/ppm: 146.72, 144.81, 136.83, 125.53, 121.90, 28.83, 23.57, 22.74. HRMS (ESI): $C_{54}H_{68}CuN_4Na_5O_{12}S_4^{2+}$ *calcd*.: 635.1245, *found*.: 635.1255.



Ammonium functionalized [(NHC)CuCl] complex (105). Compound **105** was synthesized according to a published procedure.[212] White solid (30 mg, 69 %). The analytical data were in agreement with the literature.



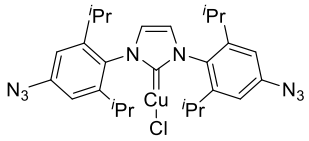
Azide functionalized NHC ligand (111). Compound **111** was synthesized according to a published procedure.[217] White solid (327 mg, 74 %). The analytical data were in agreement with the literature.



General procedure for copper complexation via transmetalation reaction

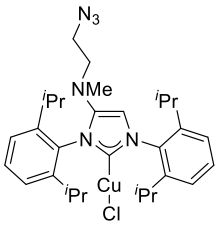
The NHC ligand, Ag_2O and LiCl in THF (degased) were stirred at 70°C for 2 h, filtered over Hyflo and washed with DCM (degased). The solvent was evaporated under reduced pressure and the mixture re-dissolved in DCM and degassed for 10 min by bubbling nitrogen through. Fresh CuCl was added and the reaction mixture stirred over night at 37°C. After filtration over Hyflo washed with DCM (degased) and evaporated under reduced pressure.

Azide functionalized [NHC)CuCl] complex (112). According to the general procedure compound **111**



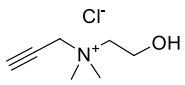
(68.6 mg, 135 µM), Ag_2O (68.6 mg, 135 µM), (68.6 mg, 135 µM), THF (3.5 mL), CuCl (68.6 mg, 135 µM) and CH_2Cl_2 (3.5 mL) were reacted and purified to yield **112** as a brown solid (52.9 mg, 69 %). ^{1}H-NMR (400 MHz, MeOD) δ 7.61 (s, 2H), 7.02 (s, 4H), 2.58 (dt, *J* = 13.6, 6.8 Hz, 4H), 1.27 (d, *J* = 6.9 Hz, 12H), 1.24 (d, *J* = 6.9 Hz, 6H). HRMS (ESI): $C_{54}H_{68}CuN_4Na_5O_{12}S_4^{2+}$ *calcd*.: 635.1245, *found*.: 635.1255.

Azide functionalized [(NHC)CuCl] in the skeleton (144). According to the general procedure compound



142 (105 mg, 168 µM), Ag_2O (23.43 mg, 102 µM), LiCl (35.7 mg, 842 µM), THF (3.5 mL), CuCl (25.0 mg, 253 µM) and CH_2Cl_2 (3.5 mL) were reacted and purified to yield **144** as a brown solid (77 mg, 78 %). ^{1}H-NMR (400 MHz, $CDCl_3$) δ/ppm: 7.48 (dt, *J* = 19.7, 7.8 Hz, 2H), 7.34 – 7.26 (m, 4H), 6.59 – 6.55 (s, 1H), 2.97 – 2.89 (m, 2H), 2.89 – 2.82 (m, 2H), 2.71 (s, 3H), 2.53 (dt, *J* = 13.7, 6.9 Hz, 2H), 2.64 (dt, *J* = 13.6, 6.8 Hz, 2H), 1.37 – 1.21 (m, 24H). ^{13}C-NMR (101 MHz, $CDCl_3$) δ/ppm: 146.04, 145.61, 144.84, 134.97, 132.26, 130.70, 130.37, 124.59, 124.14, 108.61, 53.64, 49.07, 40.83, 29.00, 28.69, 26.04, 25.01, 24.00, 22.76. HRMS (ESI): $C_{32}H_{45}CuN_7^+$ *calcd*.: 590.3027, *found*.: 590.3035.

Ammonium functionalized terminal alkyne (114). A solution of *N*,*N*-dimethylaminoethanol (1.00 g, 11



mM) in toluene (1.5mL) was added dropwise to a solution of propargyl chloride (70% w/w in toluene, 1.25 mL, 11 mmol) in toluene (1.5mL). The reaction was stirred for one day at room temperature. The obtained precipitate was filtered and washed with pentane to yield **1043** as white solid (1.17 g, 64%). ^{1}H-NMR (400 MHz, Methanol-d4) δ/ppm: 4.43 (d, *J* = 2.5 Hz, 2H), 4.05 – 3.99 (m, 2H), 3.62 – 3.57 (m, 2H), 3.55 (t, *J* = 2.5 Hz, 1H), 3.27 (s, 6H). ^{13}C-NMR (101 MHz, $CDCl_3$) δ/ppm: x. HRMS (ESI): $C_7H_{14}NO^+$ *calcd*.: 128.1070, *found*.: 635.1255.

General procedure for alkylation of imidazole. To a solution of of bromoalkyl species in DMF, the imidazole compound was added and the reaction mixture was heated up to reflux for 3h. The imidazolium salt was triturated with EtOAc. The mixture was filtered and washed with EtOAc, dried and purified by flash RP-column chromatography using method x. The corresponding fractions were lyophilized to yield the target imidazolium salt.

Sodium 2,2'-(1H-imidazole-3-ium-1,3-diyl)bis(ethane-1-sulfonate) (115). According to the general procedure 1H-imidazole (200 mg, 2.94 mM), sodium(bromoethyl)sulfonate (1.24 g, 5.88 mM) in DMF (7 mL) were reacted and purified to yield **115** as white solid (431 mg, 40%). ^{1}H-NMR (400 MHz, D_2O) δ/ppm: 8.97 (t, *J* = 1.7 Hz, 1H), 7.60 (d, *J* = 1.7 Hz, 2H), 4.64 (t, *J* = 6.5 Hz, 4H), 3.46 (t, *J* = 6.5 Hz, 4H). ^{13}C-NMR (101 MHz, D_2O) δ/ppm: 136.79, 122.64, 49.83, 45.22. HR-MS (ESI): $C_7H_{11}N_2Na_2O_6S_2^+$ *calcd*.: 328.9848, *found*: 328.9848.

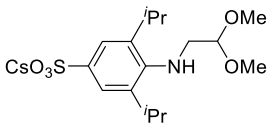
2-(1-methyl-1H-imidazol-3-ium-3-yl)ethane-1-sulfonate (116). According to the general procedure 1-methyl-imidazole (400 mg, 4.82 mM), sodium(bromoethyl)sulfonate (1.02 g, 4.82 mM) in DMF (8 mL) were reacted and purified to yield **116** as white solid (579 mg, 44%). ^{1}H-NMR (400 MHz, D_2O) δ/ppm: 8.81 (d, *J* = 0.8 Hz, 1H), 7.57 (t, *J* = 1.9 Hz, 1H), 7.46 (t, *J* = 1.8 Hz, 1H), 4.67 – 4.59 (m, 2H), 3.92 (d, *J* = 0.6 Hz, 3H), 3.49 – 3.41 (m, 2H). ^{13}C-NMR (101 MHz, D_2O) δ/ppm: 136.69, 123.60, 122.37, 49.84, 45.05, 35.72. (ESI): $C_6H_{10}N_2NaO_3S^+$ *calcd*.: 213.0304, *found*: 213.0308.

3-(2-hydroxyethyl)-1-methyl-1H-imidazol-3-ium bromide (117). According to the general procedure 1-methyl-imidazole (500 mg, 6.03 mM), bromo ethanol (753 mg, 6.03 mM) in DMF (12 mL) were reacted and purified to yield **£18** as white solid (909 mg, 73%). ^{1}H-NMR (400 MHz, D_2O) δ/ppm: 8.68 (s, 1H), 7.44 (t, *J* = 1.9 Hz, 1H), 7.38 (t, *J* = 1.8 Hz, 1H), 4.25 (t, *J* = 5.0 Hz, 2H), 3.88 – 3.84 (m, 2H), 3.83 (s, 3H). ^{13}C-NMR (101 MHz, D_2O) δ/ppm: 136.64, 123.90, 122.83, 60.16, 51.98, 36.53. (ESI): $C_6H_{11}N_2O^+$ *calcd*.: 127.0866, *found*: 127.0867.

"Der Mensch erfand die Atombombe, doch keine Maus der Welt würde eine Mausefalle konstruieren."

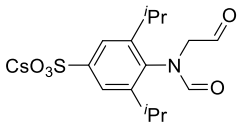
Albert Einstein (1879 - 1955)

Cesium (2,2-dimethoxyethyl)amino)-2,6-diisopropylbenzenesulfonate (121). To a solution of



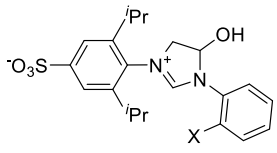
compound **120** (3.30 g, 12.8 mmol) in DMF (30 mL), 2-bromo-1,1-dimethoxyethane (2.15 mL, 17.7 mmol) and Cs_2CO_3 (8.36 g, 25.7 mmol) was added and the reaction mixture was stirred for 16.5 h at 160°C. After that the solvent was evaporated under reduced pressure and the residue was purified by flash RP-column chromatography (H_2O/ MeCN) and to obtain **121** as a yellowish solid (2.39 g, 39%). ^{1}H-NMR (400 MHz, D_2O) δ/ppm: 7.59 (s, 2H), 4.60 (t, *J* = 5.4 Hz, 1H), 3.44 (s, 6H), 3.24 (dt, *J* = 14.1, 7.0 Hz, 2H), 3.15 (d, *J* = 5.3 Hz, 2H), 1.6 (d, *J* = 6.8 Hz, 12H). ^{13}C-NMR (101 MHz, D_2O) δ/ppm: 144.30, 142.30, 137.77, 121.20, 103.72, 54.67, 51.79, 27.61, 23.17. HRMS (ESI): $C_{16}H_{27}CsNO_5S^+$ *calcd*.: 478.0659, *found*: 478.0659.

Cesium 2,6-diisopropyl-N-(2-oxoethyl)formamido)benzenesulfonate (122). Acetic anhydride (925 µL,



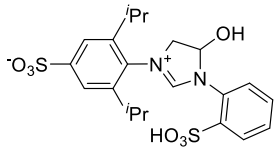
9.71 mM) and formic acid (899 µL, 22.6 mM) were stirred at room temperature for 2 h. Compound **121** (1.55g, 3.25 mM) was placed in a pressure tube, CH_2Cl_2 (20 mL) and the previous prepared mixture from acetic anhydride and acetic acid was added. The reaction was stirred at 45°C for 6 h and the solvent was evaporated under reduced pressure. To the residue formic acid (8 mL) was added and the reaction was stirred for 24 h. The formic acid was evaporated via azeotropic coevaporation with toluene (3x 10 mL). The residue was purified by RP-column chromatography using (H_2O/ MeCN) to obtain **122** as white solid (756 mg, 51 %). ^{1}H-NMR (400 MHz, DMSO) δ/ppm: 9.61 (s, 1H), 8.16 (s, 1H), 7.48 (s, 2H), 4.27 (d, *J* = 1.0 Hz, 2H), 3.09 (dt, *J* = 13.5, 6.8 Hz, 2H), 1.15 (dd, *J* = 7.9, 6.7 Hz, 6H). ^{13}C-NMR (101 MHz, DMSO) δ/ppm: 196.9, 163.73, 146.63, 135.50, 121.37, 57.40, 27.68, 24.30, 23.68. HR-MS (ESI): $C_{15}H_{20}NO_5S^-$ *calcd*.: 326.1068, *found*: 326.1066.

General procedure for cyclization of chelating ligands. To a solution of **122** in acetic anhydride and



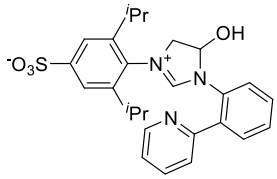
HBF4*OEt_2 was added and after 15 minutes a white precipitate was observed. Et_2O was added and the mixture was centrifuged at 4400 g for 5 min. The supernatant was carefully discarded and the pellet was re-suspended in DMF. The aniline species was added and the reaction mixture was stirred over night at room temperature. The desired product was triturated with Et_2O and purified by flash RP-column chromatography (H_2O/ MeCN).

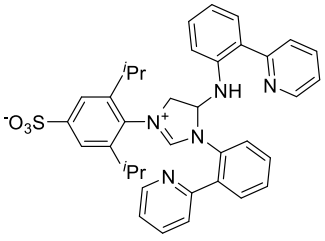
4-(5-hydroxy-1-(2-sulfophenyl)-4,5-dihydro-1H-imidazol-3-ium-3-yl)-3,5-



diisopropylbenzenesulfonate (124). According to the general procedure compound **122** (527 mg, 1.15 mM), acetic anhydride (2 mL), HBF4*OEt_2 (375 µL, 1.38 mM) and 2-aminobenzenesulfonic acid (228 mg, 1.32 mM) in DMF (2 mL) were reacted and purified (two columns were necessary) to yield **124** as white solid (529 mg, 19 %.). ^{1}H-NMR (400 MHz, D_2O) δ/ppm: 8.84 (d, *J* = 1.2 Hz, 1H), 7.96 – 7.86 (m, 1H), 7.65 (s, 2H), 7.62 – 7.58 (m, 2H), 7.53 – 7.48 (m, 1H), 6.21 (dd, *J* = 8.1, 3.1 Hz, 1H), 4.53 (ddd, *J* = 13.9, 8.0, 1.4 Hz, 1H), 4.08 (dd, *J* = 14.0, 3.1 Hz, 1H), 3.09 – 3.02 (m, 1H), 2.99 (dt, *J* = 13.6, 6.8 Hz, 1H), 1.22 (dd, *J* = 22.1, 6.8 Hz, 6H), 1.12 (dd, J = 25.2, 6.8 Hz, 6H). ^{13}C-NMR (101 MHz, D_2O) δ/ppm: 158.46, 148.58, 147.88, 145.07, 140.52, 132.49, 131.60, 131.22, 130.73, 129.30, 128.24, 122.06, 121.83, 86.84, 60.25, 28.51, 28.17, 23.75, 23.28, 22.65. HR-MS (ESI): $C_{21}H_{25}N_2O_7S_2^-$ *calcd*.: 481.1109, *found*: 481.1111.

4-(5-hydroxy-1-(2-(pyridin-2-yl)phenyl)-4,5-dihydro-1H-imidazol-3-ium-3-yl)-3,5-



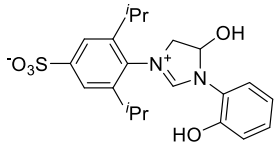


125b

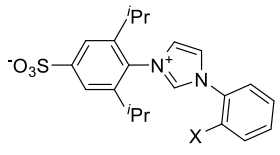
diisopropylbenzenesulfonate (125).

According to the general procedure compound **122** (250 mg, 544 µM), acetic anhydride (2 mL), HBF4*OEt_2 (178 µL, 653 µM) and 2-(pyridin-2-yl)aniline (139 mg, 816 µM) in DMF (2 mL) were reacted and purified to yield **125** and the side product **125b**. Separation of the two compounds by flash RP-column was unsuccessful. The mixture was used for the next step without further purification.

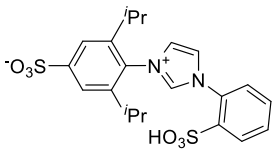
3-(2,6-diisopropyl-4-sulfophenyl)-5-hydroxy-1-(2-hydroxyphenyl)-4,5-dihydro-1H-imidazol-3-ium



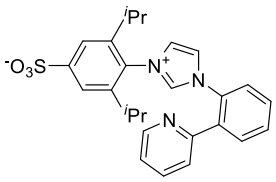
(126). According to the general procedure compound **122** (235 mg, 512 mM), acetic anhydride (2 mL), HBF4*OEt_2 (167 µL, 614 µM) and 2-aminophenol (118 mg, 1.08 mM) in DMF (2 mL) were reacted and purified to yield **126** as white solid (215 mg, 33 %.). ^{1}H-NMR (400 MHz, D_2O) δ/ppm: 10.71 (s, 1H), 9.42 (s, 1H), 8.08 (d, *J* = 6.7 Hz, 1H), 7.56 (dd, *J* = 6.8, 1.8 Hz, 2H), 7.47 (dd, *J* = 8.0, 1.6 Hz, 1H), 7.29 (td, *J* = 7.9, 7.5, 1.6 Hz, 1H), 7.07 (dd, *J* = 8.4, 1.3 Hz, 1H), 6.98 (td, *J* = 7.7, 1.3 Hz, 1H), 6.34 (ddd, *J* = 8.1, 6.6, 3.4 Hz, 1H), 4.61 – 4.48 (m, 1H), 3.94 (dd, *J* = 13.5, 3.4 Hz, 1H), 3.06 (dt, *J* = 13.5, 6.7 Hz, 1H), 2.95 (p, *J* = 6.7 Hz, 1H), 1.38 – 1.12 (m, 12H). ^{13}C-NMR (101 MHz, DMSO) δ/ppm: 158.46, 151.56, 150.61, 145.98, 145.81, 130.06, 129.72, 125.59, 121.64, 121.58, 119.49, 116.52, 85.88, 60.19, 27.86, 27.77, 24.79, 24.40, 23.85, 23.28. HR-MS (ESI): $C_{21}H_{27}N_2O_5S^+$ *calcd*.: 419.1635, *found*: 419.1632.

General procedure for water elimination. To a solution of the corresponding hemiaminal in DMF,  HBF_4*OEt_2 was added and the reaction mixture was stirred at 120°C for 4 d. The solvent was evaporated under reduced pressure and the residue was purified twice by flash RP-column chromatography (H_2O/ MeCN).

3,5-diisopropyl-4-(1-(2-sulfophenyl)-1H-imidazol-3-ium-3-yl)benzenesulfonate (127). According to the general procedure compound **125** (92.2 mg, 191 µM) and HBF_4*OEt_2 (236 µL, 956 µM) in DMF (2 mL) were reacted and purified to yield **127** as white solid (43 mg, 71 %). ^{1}H NMR (400 MHz, D_2O) δ/ppm: 9.44 (d, *J* = 1.4 Hz, 1H), 8.16 – 8.07 (m, 1H), 7.97 (t, *J* = 1.7 Hz, 1H), 7.90 (t, *J* = 2.0 Hz, 1H), 7.82 (s, 2H), 7.82 – 7.77 (m, 2H), 7.74 – 7.69 (m, 1H), 2.63 (dt, *J* = 13.5, 6.8 Hz, 2H), 1.21 (t, *J* = 7.2 Hz, 12H). ^{13}C-NMR (101 MHz, D_2O) δ/ppm: 147.17, 145.56, 138.99, 138.66, 132.72, 132.19, 131.93, 130.58, 128.64, 128.47, 125.89, 124.23, 121.66, 28.40, 23.09, 23.06. HRMS (ESI): $C_{21}H_{23}N_2O_6S_2^-$ *calcd.*: 463.1003, *found*: 463.1004.



3,5-diisopropyl-4-(1-(2-(pyridin-2-yl)phenyl)-1H-imidazol-3-ium-3-yl)benzenesulfonate (128).



According to the general procedure the mixture of **125** and **125b** (175 mg, 365 µM) and HBF_4*OEt_2 (236 µL, 956 µM) in DMF (2 mL) were reacted and purified to yield **128** as white solid (51 mg, 31 %). ^{1}H NMR (400 MHz, D_2O) δ/ppm: δ 9.88 (t, J = 1.6 Hz, 1H), 8.45 (ddd, J = 4.8, 1.9, 1.0 Hz, 1H), 8.21 (dt, *J* = 13.4, 1.8 Hz, 2H), 8.03 – 7.94 (m, 2H), 7.91 – 7.83 (m, 2H), 7.82 – 7.77 (m, 1H), 7.74 (dt, *J* = 7.8, 1.0 Hz, 1H), 7.60 (s, 2H), 7.48 (ddd, *J* = 7.6, 4.8, 1.1 Hz, 1H), 2.28 (dt, *J* = 13.5, 6.7 Hz, 2H), 1.17 (d, *J* = 6.8 Hz, 6H), 1.09 (d, *J* = 6.7 Hz, 6H) . HRMS (ESI): $C_{26}H_{28}N_3O_3S^-$ *calcd.*: 462.1846, *found*: 462.1851.

3-(2,6-diisopropyl-4-sulfophenyl)-1-(2-hydroxyphenyl)-1H-imidazol-3-ium (129). According to the general procedure compound **126** (63.2 mg, 151 µM) and HBF_4*OEt_2 (337 µL, 1.36 mM) in DMF (2 mL) were reacted and purified to yield **129** as white solid (43 mg, 71 %). ^{1}H NMR (400 MHz, DMSO) δ/ppm: 11.03 (s, 1H), 9.98 (d, *J* = 1.5 Hz, 1H), 8.38 (d, *J* = 1.8 Hz, 1H), 8.32 (t, *J* = 1.6 Hz, 1H), 7.771 (dd, *J* = 8.1, 1.6 Hz, 1H), 7.64 (s, 2H), 7.50 – 7.41 (m, 1H), 7.16 (d, *J* = 8.2 Hz, 1H), 7.08 (t, *J* = 7.6 Hz, 1H), 2.42 (p, *J* = 6.7 Hz, 2H), 1.18 (t, *J* = 6.8 Hz, 12H). ^{13}C-NMR (101 MHz, D_2O) δ/ppm: 151.27, 150.74, 144.70, 139.01, 131.61, 130.29, 126.18, 124.81, 124.32, 122.14, 121.29, 119.66, 116.88, 28.14, 23.77, 23.62. HRMS (ESI): $C_{21}H_{23}N_2O_4S^-$ *calcd*.: 399.1384, *found*: 399.1387.

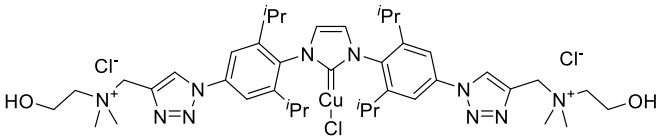
Ammonium functionalized amine (152). To a solution of 2-(Dimethylamino)ethyl ether (200 µL, 1.05 mmol) in acetonitrile (1.5 mL) chloro ethanol (70.3 µL, 1.05 mmol) was added and the mixture was stirred at 85°C over night. The mixture was filtered and the solvent evaporated to yield **152** as white solid (120 mg, 48%). ^{1}H-NMR (250 MHz, MeOD) δ/ppm: 4.01 (ddd, *J* = 6.5, 5.2, 2.7 Hz, 2H), 3.92 (dq, *J* = 7.3, 2.7 Hz, 2H), 3.74 – 3.67 (m, 2H), 3.64 (t, *J* = 5.5 Hz, 2H), 3.61 – 3.52 (m, 2H), 3.24 (s, 3H), 2.60 (t, *J* = 5.5 Hz, 2H), 2.32 (s, 3H). ^{13}C-NMR (151 MHz, D_2O) δ/ppm: 178.04, 148.66, 137.64, 136.13, 135.29, 127.50, 124.41, 117.04, 65.15, 59.11, 55.42, 50.92, 28.85, 23.75, 22.59. HRMS (ESI): $C_{41}H_{62}Cl_3CuN_{10}O_2^{2+}$ *calcd*.: 412.2015, *found*.: 412.2021.

Ammonium functionalized amine (153). To a suspension of compound **152** (120 mg, 498 µmol) in acetonitrile (1.5 mL) propargylbromide (56.7 µL, 748 µmol) was added and the mixture was stirred at 85°C over night. The mixture was filtered and washed with Et_2O, dried to yield **1045** as white solid (111 mg, 62%). ^{1}H-NMR (400 MHz, MeOD) δ/ppm: 4.48 (d, *J* = 2.3 Hz, 2H), 4.08 – 3.98 (m, 6H), 3.84 – 3.75 (m, 4H), 3.64 – 3.60 (m, 2H), 3.59 (t, *J* = 2.6 Hz, 1H), 3.29 (s, 6H), 3.27 (s, 6H). ^{13}C-NMR (101 MHz, MeOD) δ/ppm: 83.21, 72.60, 67.74, 65.95, 65.71, 64.65, 56.90, 56.31, 53.15, 52.12. HRMS (ESI): $C_{13}H_{28}N_2O_2^{2+}$ *calcd*.: 244.2140, *found*.: 412.2021.

1-(4-formylphenyl)-3-phenyl-1H-imidazol-3-ium (159). The NHC **159** was synthesized according to a published procedure.[222] White solid (212 mg, 55 %). The analytical data were in agreement with the literature.

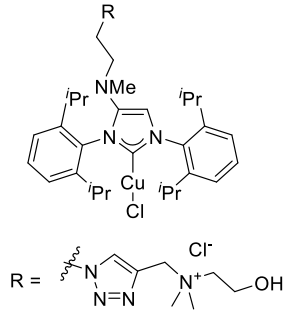
General procedure for bio-conjugation of the skeleton modified [(NHC)CuCl] complex via click reaction. The azide functionalized [(NHC)CuCl] complex, terminal alkyne, $CuSO_4$ and sodium ascorbate were stirred for 2 h at 55°C in a DMSO/water mixture (1:1, degassed before by bubbling nitrogen through the solution). The mixture was purified by prep RP-HPLC using method A. The corresponding fraction were lyophilized (three times, the fraction were degassed before freezing) to yield the functionalized [(NHC)CuCl] complex.

Ammonium functionalized [(NHC)CuCl] complex (106). According to the general procedure compound



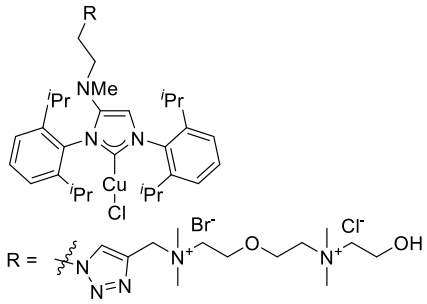
111 (29.8 mg, 52 µM), alkyne **114** (16.7 mg, 102 µM), $CuSO_4$ (19.59 mg, 78 µM) and sodium ascorbate (15.6 mg, 78 µM) were reacted and purified to yield the target complex **106** as white solid. (27 mg, 58 %). ^{1}H-NMR (600 MHz, D_2O) δ/ppm: 8.96 (s, 2H), 7.79 (s, 4H), 7.66 (s, 2H), 4.95 – 4.85 (s, 2H), 4.25 – 4.10 (m, 2H), 3.68 – 3.51 (m, 2H), 3.26 – 3.22 (s, 7H), 3.21 – 3.14 (q, J = 7.3 Hz, 1H), 1.25 – 1.20 (d, J = 7.1 Hz, 12H). ^{13}C-NMR (151 MHz, D_2O) δ/ppm: 178.04, 148.66, 137.64, 136.13, 135.29, 127.50, 124.41, 117.04, 65.15, 59.11, 55.42, 50.92, 28.85, 23.75, 22.59. HRMS (ESI): $C_{41}H_{62}C_{l3}CuN_{10}O_2^{2+}$ *calcd*.: 412.2015, *found*.: 412.2021.

Ammonium functionalized [(NHC)CuCl] complex (145). According to the general procedure compound



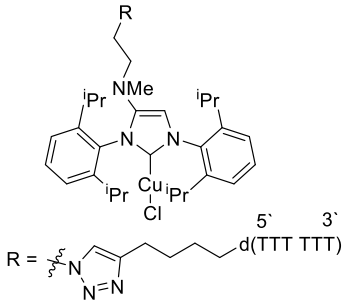
144 (53.8 mg, 99.0 µM), alkyne **114** (15.4 mg, 94.1 µM), $CuSO_4$ (24.7 mg, 99.0 µM), sodium ascorbate (19.6 mg, 99.0 µM) were reacted and purified to yield **145** as white solid (36.7 mg, 49 %). ^{1}H-NMR (400 MHz, MeOD) δ/ppm: 8.02 (s, 1H), 7.59 (t, *J* = 7.6 Hz, 1H), 7.51 (t, *J* = 7.8 Hz, 1H), 7.41 (d, *J* = 7.7 Hz, 2H), 7.36 (d, *J* = 7.8 Hz, 2H), 7.21 (s, 1H), 4.71 (s, 2H), 4.26 (t, *J* = 6.7 Hz, 2H), 4.19 – 4.02 (m, 2H), 3.45 (m, 2H), 3.14 (s, 6H), 2.73 (s, 3H), 2.72 – 2.64 (m, 2H), 2.56 (p, *J* = 6.7 Hz, 2H), 1.34 – 1.21 (m, 24H). ^{13}C-NMR (101 MHz, MeOD) δ/ppm 147.54, 147.09, 146.59, 136.56, 131.90, 131.49, 125.82, 125.23, 112.30, 66.39, 60.54, 57.02, 55.96, 51.82, 41.14, 30.20, 29.91, 26.64, 25.30, 24.36, 23.09. HRMS (ESI): $C_{37}H_{56}ClCuN_7O^+$ *calcd*.: 712.3525, *found:* 712.354.

Ammonium functionalized [(NHC)CuCl] complex (147). According to the general procedure compound



144 (30.1 mg, 51.0 µM), alkyne **153** (17.6 mg, 49.1 µM), $CuSO_4$ (19.25 mg, 77.0 µM), sodium ascorbate (15.3 mg, 77.0 µM) were reacted and purified to yield **147** as white solid (36.7 mg, 49 %). ^{1}H-NMR (500 MHz, D_2O) δ/ppm: 7.84 (s, 1H), 7.56 (dt, *J* = 13.1, 7.8 Hz, 2H), 7.42 (d, *J* = 7.7 Hz, 2H), 7.33 (d, *J* = 7.8 Hz, 2H), 7.28 (s, 1H), 4.60 (s, 2H), 4.46 (t, *J* = 5.9 Hz, 2H), 4.13 – 3.98 (m, 7H), 3.75 (t, *J* = 4.5 Hz, 2H), 3.61-3.60 (m, 4H), 3.23 (s, 6H), 3.20 (d, *J* = 6.3 Hz, 2H), 3.10 (s, 6H), 2.79 (s, 3H), 2.79 (p, *J* = 6.8 Hz, 2H), 2.59 (p, *J* = 6.8 Hz, 2H), 1.29 – 1.08 (m, 24H). ^{13}C-NMR (151 MHz, D_2O) δ/ppm: 174.37, 146.38, 146.20, 145.57, 134.97, 134.36, 131.90, 130.46, 130.37, 128.57, 124.29, 124.09, 112.28, 66.18, 64.42, 64.35, 64.25, 63.41, 59.40, 55.29, 55.18, 52.10, 50.59, 47.38, 39.52, 28.44, 28.30, 25.39, 24.02, 23.00, 21.73. HR-MS (ESI): $C_{43}H_{70}ClCuN_8O_2^{2+}$ *calcd*.: 414.2298; *found*: 414.2299.

d(TTTTTT) functionalized [(NHC)CuCl] complex (148). 5`hexynyl modified d(TTTTTT) (10 µL, 19.7



mmol stock solution in H_2O) was placed in a 1.5 mL eppendorf tube (1.5 mL), DMSO (85 µL, degased), H_2O (50 µL, degased), compound **144** (15 µL, 15 mM in degased DMSO) and sodium ascorbate (20 µL, 50 mM stock solution in H_2O) were added and the mixture was degased for 30 seconds. The $CuSO_4$ stock solution (20 µL, 50 mM stock solution in H_2O) was added and the mixture was stirred at rt for 4h. Then NaOAC (20 µL, 3 M stock solution in H_2O) and EtOH (550 µL HPLC grade, degased) were added and the mixture was placed for 30 min at -20°C. After centrifuging at 4°C for 15 min the supernatent was removed and the obtained pellet washed with EtOH (550 µL HPLC grade, degased) centrifuged again at 4°C for 15 min, the supernatent was removed, H_2O (degased) was added and the solution was lyophilized to yield the target complex **148** as white solid (50 µL a 797 µM, 27 %). HPLC and MALDI-TOf analysis of complex **148** are presented in Figure 67.

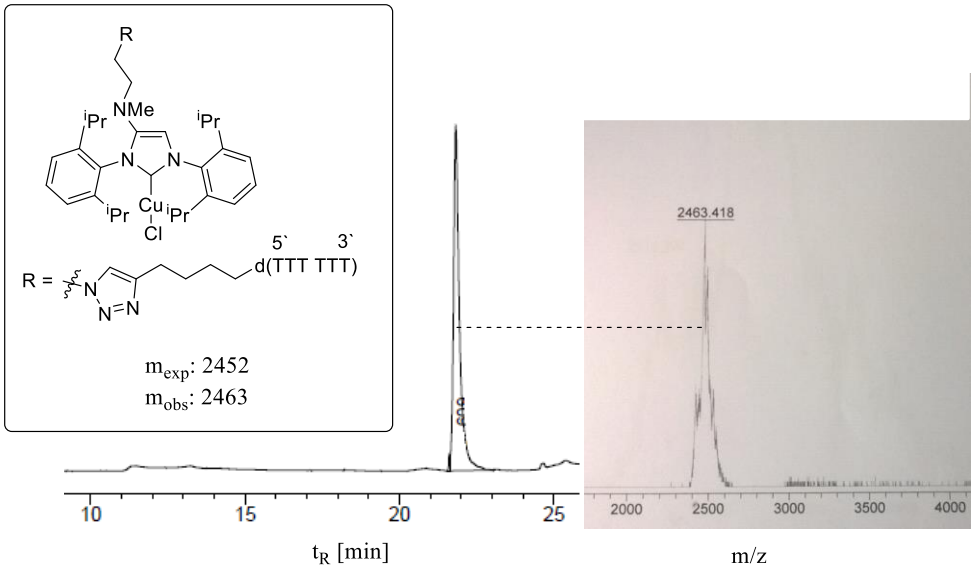


Figure 67. HPLC and MALDI-Tof analysis of NHC-Cu complex **148**.

References

1. Bird, A. P., *Nature* **1986,** *321* (6067), 209-213.
2. Carell, T.; Brandmayr, C.; Hienzsch, A.; Muller, M.; Pearson, D.; Reiter, V.; Thoma, I.; Thumbs, P.; Wagner, M., *Angew Chem Int Edit* **2012,** *51* (29), 7110-7131.
3. Opalinska, J. B.; Gewirtz, A. M., *Nat Rev Drug Discov* **2002,** *1* (7), 503-514.
4. Nikolova, E. N.; Kim, E.; Wise, A. A.; O'Brien, P. J.; Andricioaei, I.; Al-Hashimi, H. M., *Nature* **2011,** *470* (7335), 498-U84.
5. Alseth, I.; Dalhus, B.; Bjoras, M., *Curr. Opin. Genet Dev* **2014,** *26*, 116-123.
6. Izatt, R. M.; Christensen, J. J.; Rytting, J. H., *Chem Rev* **1971,** *71* (5), 439-+.
7. Legault, P.; Pardi, A., *J Am Chem Soc* **1997,** *119* (28), 6621-6628.
8. Gillingham, D.; Geigle, S.; von Lilienfeld, O. A., *Chem Soc Rev* **2016,** *45* (9), 2637-2655.
9. Lilley, D. M. J., *Philos T R Soc B* **2011,** *366* (1580), 2910-2917.
10. Sharmeen, L.; Kuo, M. Y. P.; Dintergottlieb, G.; Taylor, J., *J Virol* **1988,** *62* (8), 2674-2679.
11. Gong, B.; Chen, J. H.; Chase, E.; Chadalavada, D. M.; Yajima, R.; Golden, B. L.; Bevilacqua, P. C.; Carey, P. R., *J Am Chem Soc* **2007,** *129* (43), 13335-13342.
12. Lu, X. F.; Heilman, J. M.; Blans, P.; Fishbein, J. C., *Chem Res Toxicol* **2005,** *18* (9), 1462-1470.
13. Parrish, J. P.; Kastrinsky, D. B.; Wolkenberg, S. E.; Igarashi, Y.; Boger, D. L., *J Am Chem Soc* **2003,** *125* (36), 10971-10976.
14. Boger, D. L.; Garbaccio, R. M., *Bioorgan Med Chem* **1997,** *5* (2), 263-276.
15. Boger, D. L.; Garbaccio, R. M., *Accounts Chem Res.* **1999,** *32* (12), 1043-1052.
16. Tomasz, M., Mitomycin C: Small, fast and deadly (but very selective) (vol 2, pg 575, 1995). *Chem Biol* **1995,** *2* (12), 865-865.
17. Bass, P. D.; Gubler, D. A.; Judd, T. C.; Williams, R. M., *Chem Rev* **2013,** *113* (8), 6816-6863.
18. Moore, H. W., Bioactivation as a Model for Drug Design Bioreductive Alkylation. *Science* **1977,** *197* (4303), 527-532.
19. Turner, A. B., Quinone Methides. *Q Rev Chem Soc* **1964,** *18* (4), 347-360.
20. Bizanek, R.; Mcguinness, B. F.; Nakanishi, K.; Tomasz, M., *Biochemistry-Us* **1992,** *31* (12), 3084-3091.
21. Tomasz, M.; Chowdary, D.; Lipman, R.; Shimotakahara, S.; Veiro, D.; Walker, V.; Verdine, G. L.,. *P Natl Acad Sci USA* **1986,** *83* (18), 6702-6706.
22. Tomasz, M.; Lipman, R.; Chowdary, D.; Pawlak, J.; Verdine, G. L.; Nakanishi, K., *Science* **1987,** *235* (4793), 1204-1208.
23. Williams, J. H.; Phillips, T. D.; Jolly, P. E.; Stiles, J. K.; Jolly, C. M.; Aggarwal, D., *Am J Clin Nutr* **2004,** *80* (5), 1106-1122.
24. Essigmann, J. M.; Croy, R. G.; Nadzan, A. M.; Busby, W. F.; Reinhold, V. N.; Buchi, G.; Wogan, G. N., *P Natl Acad Sci USA* **1977,** *74* (5), 1870-1874.
25. Coleman, R. S.; Perez, R. J.; Burk, C. H.; Navarro, A., *J Am Chem Soc* **2002,** *124* (44), 13008-13017.
26. Armstrong, R. W.; Salvati, M. E.; Nguyen, M., *J Am Chem Soc* **1992,** *114* (8), 3144-3145.
27. Gates, K. S., *Chem Res Toxicol* **2000,** *13* (10), 953-956.
28. Zang, H.; Gates, K. S., *Chem Res Toxicol* **2003,** *16* (12), 1539-1546.
29. Hara, M.; Asano, K.; Kawamoto, I.; Takiguchi, T.; Katsumata, S.; Takahashi, K. I.; Nakano, H., *J Antibiot* **1989,** *42* (12), 1768-1774.
30. Pommier, Y.; Kohlhagen, G.; Bailly, C.; Waring, M.; Mazumder, A.; Kohn, K. W.,*Biochemistry-Us* **1996,** *35* (41), 13303-13309.
31. Kanne, D.; Straub, K.; Rapoport, H.; Hearst, J. E., *Biochemistry-Us* **1982,** *21* (5), 861-871.
32. Sastry, S. S.; Spielmann, H. P.; Dwyer, T. J.; Wemmer, D. E.; Hearst, J. E., *J Photoch Photobio B* **1992,** *14* (1-2), 65-79.
33. Brogan, B. L.; Zic, J. A.; Kinney, M. C.; Hu, J. Y.; Hamilton, K. S.; Greer, J. P., *J Am Acad Dermatol* **2003,** *49* (2), 223-228.

34. Gupta, A. K.; Anderson, T. F., Psoralen Photochemotherapy. *J Am Acad Dermatol* **1987,** *17* (5), 703-734.
35. Nawrat, C. C.; Moody, C. J., Natural products containing a diazo group. *Natural Product Reports* **2011,** *28* (8), 1426-1444.
36. Herzon, S. B.; Woo, C. M., *Natural Product Reports* **2012,** *29* (1), 87-118.
37. Colis, L. C.; Woo, C. M.; Hegan, D. C.; Li, Z. W.; Glazer, P. M.; Herzon, S. B., *Nat Chem* **2014,** *6* (6), 504-510.
38. Catane, R.; Vonhoff, D. D.; Glaubiger, D. L.; Muggia, F. M., *Cancer Treat Rep* **1979,** *63* (6), 1033-1038.
39. Cervantes-Madrid, D.; Romero, Y.; Duenas-Gonzalez, A., *Biomed Res Int* **2015**.
40. Ellison, R. R.; Karnofsky, D. A.; Sternberg, S. S.; Murphy, M. L.; Burchenal, J. H, *Cancer* **1954,** *7* (4), 801-814.
41. Katoh, Y.; Maekawa, M.; Sano, Y., *Mutat Res-Genet Tox* **1995,** *342* (1-2), 37-41.
42. Lilja, H. S.; Hyde, E.; Longnecker, D. S.; Yager, J. D., *Cancer Res* **1977,** *37* (11), 3925-3931.
43. Lilja, H. S.; Longnecker, D. S.; Curphey, T. J.; Daniel, D. S.; Adams, W. E., *Cancer Lett* **1981,** *12* (1-2), 139-146.
44. B. Sedgwick *Carcinogenesis* **1997,** *18* (8), 1561-1567.
45. Lewin, M. H.; Bailey, N.; Bandaletova, T.; Bowman, R.; Cross, A. J.; Pollock, J.; Shuker, D. E. G.; Bingham, S. A., *Cancer Res* **2006,** *66* (3), 1859-1865.
46. Yamamoto, H.; Uchigata, Y.; Okamoto, H., *Nature* **1981,** *294* (5838), 284-286.
47. Mcgarrity, J. F.; Smyth, T., *J Am Chem Soc* **1980,** *102* (24), 7303-7308.
48. Lenzen, S., Oxidative stress: the vulnerable beta-cell. *Biochem Soc T* **2008,** *36*, 343-347.
49. Szkudelski, T., *Physiol Res* **2001,** *50* (6), 537-546.
50. Singer, B., *Nature* **1976,** *264* (5584), 333-339.
51. Neog, B.; Sinha, S.; Bhattacharyya, P. K., *Comput Theor Chem* **2013,** *1018*, 19-25.
52. Kohn, K. W.; Hartley, J. A.; Mattes, W. B., *Nucleic Acids Research* **1987,** *15* (24), 10531-10549.
53. Wheeler, G. P., *Cancer Res* **1962,** *22* (6), 651-&.
54. Lawley, P. D.; Phillips, D. H., *Mutat Res-Fund Mol M* **1996,** *355* (1-2), 13-40.
55. Pesonen, M.; Vahakangas, K.; Halme, M.; Vanninen, P.; Seulanto, H.; Hemmila, M.; Pasanen, M.; Kuitunen, T., *Front Pharmacol* **2010,** *1*.
56. Rajski, S. R.; Williams, R. M., *Chem Rev* **1998,** *98* (8), 2723-2795.
57. Fisher, B.; Sherman, B.; Rockette, H.; Redmond, C.; Margolese, R.; Fisher, E. R., *Cancer* **1979,** *44* (3), 847-857.
58. Polavarapu, A.; Stillabower, J. A.; Stubblefield, S. G. W.; Taylor, W. M.; Baik, M. H., *J Org Chem* **2012,** *77* (14), 5914-5921.
59. Newlands, E. S.; Stevens, M. F. G.; Wedge, S. R.; Wheelhouse, R. T.; Brock, C., *Cancer Treat Rev* **1997,** *23* (1), 35-61.
60. Margison, G. P.; Santibanez Koref, M. F.; Povey, A. C., *Mutagenesis* **2002,** *17* (6), 483-7.
61. Ramirez, Y. P.; Weatherbee, J. L.; Wheelhouse, R. T.; Ross, A. H., *Pharmaceuticals (Basel)* **2013,** *6* (12), 1475-506.
62. Chen, F. X.; Bodell, W. J.; Liang, G. N.; Gold, B., *Chem Res Toxicol* **1996,** *9* (1), 208-214.
63. Conway, N. E.; Mclaughlin, L. W., *Bioconjugate Chem* **1991,** *2* (6), 452-457.
64. Weisbrod, S. H.; Marx, A., *Chem Commun* **2008,** (44), 5675-5685.
65. Berndl, S.; Herzig, N.; Kele, P.; Lachmann, D.; Li, X. H.; Wolfbeis, O. S.; Wagenknecht, H. A., *Bioconjugate Chem* **2009,** *20* (3), 558-564.
66. Yamana, K.; Zako, H.; Asazuma, K.; Iwase, R.; Nakano, H.; Murakami, A., *Angew Chem Int Edit* **2001,** *40* (6), 1104-+.
67. Capek, P.; Cahova, H.; Pohl, R.; Hocek, M.; Gloeckner, C.; Marx, A., *Chem-Eur J* **2007,** *13* (21), 6196-6203.
68. Gierlich, J.; Gutsmiedl, K.; Gramlich, P. M. E.; Schmidt, A.; Burley, G. A.; Carell, T., *Chem-Eur J* **2007,** *13* (34), 9486-9494.

69. Seela, F.; Zulauf, M., *Chem-Eur J* **1998,** *4* (9), 1781-1790.
70. Rublack, N.; Springstubbe, D.; Nguyen, H.; Appel, B.; Muller, S., *4th European Conference on Chemistry for Life Sciences* **2011**, 71-74.
71. Cahova, H.; Panattoni, A.; Kielkowski, P.; Fanfrlik, J.; Hocek, M., *Acs Chem Biol* **2016,** *11* (11), 3165-3171.
72. Hocek, M., *J Org Chem* **2014,** *79* (21), 9914-9921.
73. Cahova, H.; Pohl, R.; Bednarova, L.; Novakova, K.; Cvacka, J.; Hocek, M., *Org Biomol Chem* **2008,** *6* (20), 3657-3660.
74. Beaucage, S. L.; Caruthers, M. H., *Tetrahedron Lett* **1981,** *22* (20), 1859-1862.
75. Tian, J. D.; Ma, K. S.; Saaem, I., Advancing high-throughput gene synthesis technology. *Mol Biosyst* **2009,** *5* (7), 714-722.
76. Damoiseaux, R.; Keppler, A.; Johnsson, K., *Chembiochem* **2001,** *2* (4), 285-287.
77. Pegg, A. E., Repair of O-6-alkylguanine by alkyltransferases. *Mutat Res-Rev Mutat* **2000,** *462* (2-3), 83-100.
78. Kusmierek, J. T.; Singer, B., *Biochemistry-Us* **1982,** *21* (22), 5723-5728.
79. Jin, S. X.; Miduturu, C. V.; McKinney, D. C.; Silverman, S. K., *J Org Chem* **2005,** *70* (11), 4284-4299.
80. Ramsay, N.; Jemth, A. S.; Brown, A.; Crampton, N.; Dear, P.; Holliger, P., *J Am Chem Soc* **2010,** *132* (14), 5096-5104.
81. Staiger, N.; Marx, A., *Chembiochem* **2010,** *11* (14), 1963-1966.
82. Langer, P. R.; Waldrop, A. A.; Ward, D. C., *Natl Acad Sci-Biol* **1981,** *78* (11), 6633-6637.
83. Thum, O.; Jager, S.; Famulok, M., *Angew Chem Int Edit* **2001,** *40* (21), 3990-+.
84. Kuwahara, M.; Nagashima, J.; Hasegawa, M.; Tamura, T.; Kitagata, R.; Hanawa, K.; Hososhima, S.; Kasamatsu, T.; Ozaki, H.; Sawai, H., *Nucleic Acids Research* **2006,** *34* (19), 5383-5394.
85. Kuwahara, M.; Hanawa, K.; Ohsawa, K.; Kitagata, R.; Ozaki, H.; Sawai, H., *Bioorgan Med Chem* **2006,** *14* (8), 2518-2526.
86. Di Giusto, D. A.; Wlassoff, W. A.; Giesebrecht, S.; Gooding, J. J.; King, G. C., *J Am Chem Soc* **2004,** *126* (13), 4120-4121.
87. Matsui, M.; Nishiyama, Y.; Ueji, S. I.; Ebara, Y., *Bioorg Med Chem Lett* **2007,** *17* (2), 456-460.
88. Gramlich, P. M. E.; Wirges, C. T.; Gierlich, J.; Carell, T., *Org Lett* **2008,** *10* (2), 249-251.
89. Burley, G. A.; Gierlich, J.; Mofid, M. R.; Nir, H.; Tal, S.; Eichen, Y.; Carell, T., *J Am Chem Soc* **2006,** *128* (5), 1398-1399.
90. Kaufmann, G. F.; Meijler, M. M.; Sun, C. Z.; Chen, D. W.; Kujawa, D. P.; Mee, J. M.; Hoffman, T. Z.; Wirsching, P.; Lerner, R. A.; Janda, K. D., *Angew Chem Int Edit* **2005,** *44* (14), 2144-2148.
91. Jager, S.; Rasched, G.; Kornreich-Leshem, H.; Engeser, M.; Thum, O.; Famulok, M., *J Am Chem Soc* **2005,** *127* (43), 15071-15082.
92. Hirao, I., *Biotechniques* **2006,** *40* (6), 711-+.
93. Kimoto, M.; Endo, M.; Mitsui, T.; Okuni, T.; Hirao, I.; Yokoyama, S., *Chem Biol* **2004,** *11* (1), 47-55.
94. Hirao, I.; Kimoto, M.; Mitsui, T.; Fujiwara, T.; Kawai, R.; Sato, A.; Harada, Y.; Yokoyama, S., *Nat Methods* **2006,** *3* (9), 729-735.
95. Kawai, R.; Kimoto, M.; Ikeda, S.; Mitsui, T.; Endo, M.; Yokoyama, S.; Hirao, L., *J Am Chem Soc* **2005,** *127* (49), 17286-17295.
96. Moriyama, K.; Kimoto, M.; Mitsui, T.; Yokoyama, S.; Hirao, I., *Nucleic Acids Research* **2005,** *33* (15).
97. Yoshikawa, M.; Kato, T.; Takenishi, T., *Tetrahedron Lett* **1967,** (50), 5065-+.
98. Hollenstein, M., *Molecules* **2012,** *17* (11), 13569-13591.
99. Thoresen, L. H.; Jiao, G. S.; Haaland, W. C.; Metzker, M. L.; Burgess, K., Rigid, *Chem-Eur J* **2003,** *9* (19), 4603-4610.
100. Hocek, M.; Fojta, M., *Org Biomol Chem* **2008,** *6* (13), 2233-2241.
101. Hocek, M.; Fojta, M., *Chem Soc Rev* **2011,** *40* (12), 5802-5814.

102. Kuwahara, M.; Sugimoto, N., *Molecules* **2010,** *15* (8), 5423-5444.
103. Sawai, H.; Ozaki, A. N.; Satoh, F.; Ohbayashi, T.; Masud, M. M.; Ozaki, H., *Chem Commun* **2001,** (24), 2604-2605.
104. Kawasaki, T.; Nagatsugi, F.; Ali, M. M.; Maeda, M.; Sugiyama, K.; Hori, K.; Sasaki, S., *J Org Chem* **2005,** *70* (1), 14-23.
105. Zhou, Q. B.; Rokita, S. E., *P Natl Acad Sci USA* **2003,** *100* (26), 15452-15457.
106. Liu, Y.; Rokita, S. E., *Abstr Pap Am Chem S* **2009,** *238*.
107. Onizuka, K.; Taniguchi, Y.; Sasaki, S., *Bioconjugate Chem* **2009,** *20* (4), 799-803.
108. Pljevaljcic, G.; Pignot, M.; Weinhold, E., *J Am Chem Soc* **2003,** *125* (12), 3486-3492.
109. Holstein, J. M.; Schulz, D.; Rentmeister, A., *Chem Commun* **2014,** *50* (34), 4478-4481.
110. Schulz, D.; Holstein, J. M.; Rentmeister, A., *Angew Chem Int Edit* **2013,** *52* (30), 7874-7878.
111. Holstein, J. M.; Stummer, D.; Rentmeister, A., *Chem Sci* **2015,** *6* (2), 1362-1369.
112. Squires, J. E.; Patel, H. R.; Nousch, M.; Sibbritt, T.; Humphreys, D. T.; Parker, B. J.; Suter, C. M.; Preiss, T., *Nucleic Acids Research* **2012,** *40* (11), 5023-5033.
113. Motorin, Y.; Lyko, F.; Helm, M., *Nucleic Acids Research* **2010,** *38* (5), 1415-1430.
114. Frommer, M.; Mcdonald, L. E.; Millar, D. S.; Collis, C. M.; Watt, F.; Grigg, G. W.; Molloy, P. L.; Paul, C. L., *P Natl Acad Sci USA* **1992,** *89* (5), 1827-1831.
115. Clark, S. J.; Harrison, J.; Paul, C. L.; Frommer, M., *Nucleic Acids Research* **1994,** *22* (15), 2990-2997.
116. Grunau, C.; Renault, E.; Rosenthal, A.; Roizes, G., *Nucleic Acids Research* **2001,** *29* (1), 270-274.
117. Henderson, I. R.; Jacobsen, S. E., *Nature* **2007,** *447* (7143), 418-424.
118. Lister, R.; O'Malley, R. C.; Tonti-Filippini, J.; Gregory, B. D.; Berry, C. C.; Millar, A. H.; Ecker, J. R., *Cell* **2008,** *133* (3), 523-536.
119. Paul, F.; Patt, J.; Hartwig, J. F., *J Am Chem Soc* **1994,** *116* (13), 5969-5970.
120. Guram, A. S.; Buchwald, S. L., *J Am Chem Soc* **1994,** *116* (17), 7901-7902.
121. Casanova, R.; Reichstein, T., *Helv Chim Acta* **1950,** *33* (2), 417-422.
122. Yates, P., *J Am Chem Soc* **1952,** *74* (21), 5376-5381.
123. Salzmann, T. N.; Ratcliffe, R. W.; Christensen, B. G.; Bouffard, F. A., *J Am Chem Soc* **1980,** *102* (19), 6161-6163.
124. Maas, G., *Angew Chem Int Edit* **2009,** *48* (44), 8186-8195.
125. Rajagopalan, T. G.; Stein, W. H.; Moore, S., *J Biol Chem* **1966,** *241* (18), 4295-+.
126. Bachmann, S.; Fielenbach, D.; Jorgensen, K. A., *Org Biomol Chem* **2004,** *2* (20), 3044-3049.
127. Maier, T. C.; Fu, G. C., *J Am Chem Soc* **2006,** *128* (14), 4594-4595.
128. Lee, E. C.; Fu, G. C., *J Am Chem Soc* **2007,** *129* (40), 12066-+.
129. Salomon, R. G.; Kochi, J. K., *J Am Chem Soc* **1973,** *95* (10), 3300-3310.
130. Anciaux, A. J.; Hubert, A. J.; Noels, A. F.; Petiniot, N.; Teyssie, P., *J Org Chem* **1980,** *45* (4), 695-702.
131. Shishkov, I. V.; Rominger, F.; Hofmann, P., *Organometallics* **2009,** *28* (4), 1049-1059.
132. Fraile, J. M.; Garcia, J. I.; Martinez-Merino, V.; Mayoral, J. A.; Salvatella, L., *J Am Chem Soc* **2001,** *123* (31), 7616-7625.
133. Liang, Y.; Zhou, H. L.; Yu, Z. X., *Am Chem Soc* **2009,** *131* (49), 17783-17785.
134. Antos, J. M.; Francis, M. B., *J Am Chem Soc* **2004,** *126* (33), 10256-10257.
135. Antos, J. M.; McFarland, J. M.; Iavarone, A. T.; Francis, M. B., *J Am Chem Soc* **2009,** *131* (17), 6301-6308.
136. Popp, B. V.; Ball, Z. T., *J Am Chem Soc* **2010,** *132* (19), 6660-+.
137. Popp, B. V.; Ball, Z. T., *Chem Sci* **2011,** *2* (4), 690-695.
138. Ho, C. M.; Zhang, J. L.; Zhou, C. Y.; Chan, O. Y.; Yan, J. J.; Zhang, F. Y.; Huang, J. S.; Che, C. M., A *J Am Chem Soc* **2010,** *132* (6), 1886-1894.
139. Tishinov, K.; Schmidt, K.; Haussinger, D.; Gillingham, D. G., *Angew Chem Int Edit* **2012,** *51* (48), 12000-12004.
140. Tishinov, K.; Fei, N.; Gillingham, D., *Chem Sci* **2013,** *4* (12), 4401-4406.

141. Fry, S. C., *Biochem J* **1998,** *332*, 507-515.
142. Baker, W. L.; Goode, J.; Cooper, L., *Mikrochim Acta* **1992,** *106* (3-6), 143-152.
143. Toma, T.; Shimokawa, J.; Fukuyama, T., *Org Lett* **2007,** *9* (16), 3195-3197.
144. Labiuk, S. L.; Delbaere, L. T. J.; Lee, J. S., *J Biol Inorg Chem* **2003,** *8* (7), 715-720.
145. Gao, Y. G.; Sriram, M.; Wang, A. H. J., *Nucleic Acids Research* **1993,** *21* (17), 4093-4101.
146. Prutz, W. A.; Butler, J.; Land, E. J., *J Radiat Biol* **1990,** *58* (2), 215-234.
147. Johnson, D. K.; Stevenson, M. J.; Almadidy, Z. A.; Jenkins, S. E.; Wilcox, D. E.; Grossoehme, N. E., *Dalton T* **2015,** *44* (37), 16494-16505.
148. Geierstanger, B. H.; Kagawa, T. F.; Chen, S. L.; Quigley, G. J.; Ho, P. S., *J Biol Chem* **1991,** *266* (30), 20185-20191.
149. Oikawa, S.; Kawanishi, S., *Biochemistry-Us* **1996,** *35* (14), 4584-4590.
150. Hong, V.; Presolski, S. I.; Ma, C.; Finn, M. G., *Angew Chem Int Edit* **2009,** *48* (52), 9879-9883.
151. Hong, V.; Steinmetz, N. F.; Manchester, M.; Finn, M. G., *Bioconjugate Chem* **2010,** *21* (10), 1912-1916.
152. Presolski, S. I.; Hong, V.; Cho, S. H.; Finn, M. G., *J Am Chem Soc* **2010,** *132* (41), 14570-14576.
153. Liu, H.; Xu-Welliver, M.; Pegg, A. E., *Mutat Res* **2000,** *452* (1), 1-10.
154. Liem, L. K.; Wong, C. W.; Lim, A.; Li, B. F., *J Mol Biol* **1993,** *231* (4), 950-9.
155. Bender, K.; Federwisch, M.; Loggen, U.; Nehls, P.; Rajewsky, M. F., *Nucleic Acids Res* **1996,** *24* (11), 2087-94.
156. Shuker, D. E.; Margison, G. P., *Cancer Res* **1997,** *57* (3), 366-9.
157. Nawrat, C. C.; Moody, C. J., *Nat Prod Rep* **2011,** *28* (8), 1426-44.
158. Katoh, Y.; Maekawa, M.; Sano, Y., *Mutat Res* **1995,** *342* (1-2), 37-41.
159. Rasimas, J. J.; Kar, S. R.; Pegg, A. E.; Fried, M. G., *J Biol Chem* **2007,** *282* (5), 3357-66.
160. Abdu, K.; Aiertza, M. K.; Wilkinson, O. J.; Grasby, J. A.; Senthong, P.; Povey, A. C.; Margison, G. P.; Williams, D. M., *Chem Commun* **2012,** *48* (91), 11214-11216.
161. Moore, S. A.; Shuker, D. E. G., *J Labelled Compd Rad* **2011,** *54* (14), 855-858.
162. Shibata, T.; Glynn, N.; McMurry, T. B. H.; McElhinney, R. S.; Margison, G. P.; Williams, D. M., *Nucleic Acids Research* **2006,** *34* (6), 1884-1891.
163. Millington, C. L.; Watson, A. J.; Marriott, A. S.; Margison, G. P.; Povey, A. C.; Williams, D. M., *Nucleos Nucleot Nucl* **2012,** *31* (4), 328-338.
164. Gloeckner, C.; Sauter, K. B. M.; Marx, A., *Angew Chem Int Edit* **2007,** *46* (17), 3115-3117.
165. Obeid, S.; Schnur, A.; Gloeckner, C.; Blatter, N.; Welte, W.; Diederichs, K.; Marx, A., *Chembiochem* **2011,** *12* (10), 1574-1580.
166. Wyss, L. A.; Nilforoushan, A.; Eichenseher, F.; Suter, U.; Blatter, N.; Marx, A.; Sturla, S. J., Specific *J Am Chem Soc* **2015,** *137* (1), 30-33.
167. Velazquez, H. D.; Verpoort, F., *Chem Soc Rev* **2012,** *41* (21), 7032-7060.
168. Gstottmayr, C. W. K.; Bohm, V. P. W.; Herdtweck, E.; Grosche, M.; Herrmann, W. A., *Angew Chem Int Edit* **2002,** *41* (8), 1363-1365.
169. Marko, I. E.; Sterin, S.; Buisine, O.; Mignani, G.; Branlard, P.; Tinant, B.; Declercq, J. P., *Science* **2002,** *298* (5591), 204-206.
170. Lee, H. M.; Jiang, T.; Stevens, E. D.; Nolan, S. P., *Organometallics* **2001,** *20* (6), 1255-1258.
171. Grasa, G. A.; Viciu, M. S.; Huang, J. K.; Nolan, S. P., *J Org Chem* **2001,** *66* (23), 7729-7737.
172. Cesar, V.; Bellemin-Laponnaz, S.; Gade, L. H., *Chem Soc Rev* **2004,** *33* (9), 619-636.
173. Wanzlick, H. W., *Angew Chem Int Edit* **1962,** *74* (4), 129-&.
174. Wanzlick, H. W.; Schonher.Hj, *Angew Chem Int Edit* **1968,** *7* (2), 141-&.
175. Arduengo, A. J.; Harlow, R. L.; Kline, M., *J Am Chem Soc* **1991,** *113* (1), 361-363.
176. Hu, X. L.; Tang, Y. J.; Gantzel, P.; Meyer, K., *Organometallics* **2003,** *22* (4), 612-614.
177. Nemcsok, D.; Wichmann, K.; Frenking, G., *Organometallics* **2004,** *23* (15), 3640-3646.
178. Jacobsen, H.; Correa, A.; Poater, A.; Costabile, C.; Cavallo, L., *Coordin Chem Rev* **2009,** *253* (5-6), 87-703.

179. Chianese, A. R.; Kovacevic, A.; Zeglis, B. M.; Faller, J. W.; Crabtree, R. H., *Organometallics* **2004,** *23* (10), 2461-2468.
180. Herrmann, W. A.; Schutz, J.; Frey, G. D.; *Organometallics* **2006,** *25* (10), 2437-2448.
181. Dorta, R.; Stevens, E. D.; Scott, N. M.; Costabile, C.; Cavallo, L.; Hoff, C. D.; Nolan, S. P., *J Am Chem Soc* **2005,** *127* (8), 2485-2495.
182. Kelly, R. A.; Clavier, H.; Giudice, S.; Scott, N. M.; Stevens, E. D.; Bordner, J.; Samardjiev, I.; Hoff, C. D.; Cavallo, L.; Nolan, S. P., *Organometallics* **2008,** *27* (2), 202-210.
183. Arduengo, A. J.; Dias, H. V. R.; Calabrese, J. C.; Davidson, F., *Organometallics* **1993,** *12* (9), 3405-3409.
184. Raubenheimer, H. G.; Cronje, S.; Vanrooyen, P. H.; Olivier, P. J.; Toerien, J. G., *Ang Chem Intern Edit* **1994,** *33* (6), 672-673.
185. Egbert, J. D.; Cazin, C. S. J.; Nolan, S. P., *Catal Sci Technol* **2013,** *3* (4), 912-926.
186. Teyssot, M. L.; Jarrousse, A. S.; Manin, M.; Chevry, A.; Roche, S.; Norre, F.; Beaudoin, C.; Morel, L.; Boyer, D.; Mahiou, R.; Gautier, A., *Dalton T* **2009,** (35), 6894-6902.
187. Fraser, P. K.; Woodward, S., *Tetrahedron Lett* **2001,** *42* (14), 2747-2749.
188. Jurkauskas, V.; Sadighi, J. P.; Buchwald, S. L., *Org Lett* **2003,** *5* (14), 2417-2420.
189. Kaur, H.; Zinn, F. K.; Stevens, E. D.; Nolan, S. P., *Organometallics* **2004,** *23* (5), 1157-1160.
190. Diez-Gonzalez, S.; Kaur, H.; Zinn, F. K.; Stevens, E. D.; Nolan, S. P., *J Org Chem* **2005,** *70* (12), 4784-4796.
191. Zheng, S. Z.; Li, F. W.; Liu, J. M.; Xia, C. G., *Tetrahedron Lett* **2007,** *48* (33), 5883-5886.
192. Kacprzynski, M. A.; May, T. L.; Kazane, S. A.; Hoveyda, A. H., *Angew Chem Int Edit* **2007,** *46* (24), 4554-4558.
193. Lazreg, F.; Slawin, A. M. Z.; Cazin, C. S. J., *Organometallics* **2012,** *31* (22), 7969-7975.
194. Fructos, M. R.; Belderrain, T. R.; Nicasio, M. C.; Nolan, S. P.; Kaur, H.; Diaz-Requejo, M. M.; Perez, P. J., *J Am Chem Soc* **2004,** *126* (35), 10846-10847.
195. Fructos, M. R.; de Fremont, P.; Nolan, S. P.; Diaz-Requejo, M. M.; Perez, P. J., *Organometallics* **2006,** *25* (9), 2237-2241.
196. Xu, Q.; Appella, D. H., *Org Lett* **2008,** *10* (7), 1497-1500.
197. Herrmann, W. A.; Goossen, L. J.; Spiegler, M., *J Organomet Chem* **1997,** *547* (2), 357-366.
198. Moore, L. R.; Cooks, S. M.; Anderson, M. S.; Schanz, H. J.; Griffin, S. T.; Rogers, R. D.; Kirk, M. C.; Shaughnessy, K. H., O*rganometallics* **2006,** *25* (21), 5151-5158.
199. Papini, G.; Pellei, M.; Lobbia, G. G.; Burini, A.; Santini, C., *Dalton T* **2009,** (35), 6985-6990.
200. Mesnager, J.; Lammel, P.; Jeanneau, E.; Pinel, C., *Appl Catal a-Gen* **2009,** *368* (1-2), 22-28.
201. Almassy, A.; Nagy, C. E.; Benyei, A. C.; Joo, F., *Organometallics* **2010,** *29* (11), 2484-2490.
202. Azua, A.; Mata, J. A.; Peris, E., *Organometallics* **2011,** *30* (20), 5532-5536.
203. Azua, A.; Sanz, S.; Peris, E., *Chem-Eur J* **2011,** *17* (14), 3963-3967.
204. Gillingham, D. G.; Hoveyda, A. H., *Angew Chem Int Edit* **2007,** *46* (21), 3860-3864.
205. Fleckenstein, C.; Roy, S.; Leuthausser, S.; Plenio, H., *Chem Commun* **2007,** (27), 2870-2872.
206. Roy, S.; Plenio, H., *Adv Synth Catal* **2010,** *352* (6), 1014-1022.
207. Czegeni, C. E.; Papp, G.; Katho, A.; Joo, F., *J Mol Catal a-Chem* **2011,** *340* (1-2), 1-8.
208. Ozdemir, I.; Yigit, B.; Cetinkaya, B.; Ulku, D.; Tahir, M. N.; Arici, C., *J Organomet Chem* **2001,** *633* (1-2), 27-32.
209. Jordan, J. P.; Grubbs, R. H., *Angew Chem Int Edit* **2007,** *46* (27), 5152-5155.
210. Kolb, H. C.; Finn, M. G.; Sharpless, K. B., *Angew Chem Int Edit* **2001,** *40* (11), 2004-+.
211. Rostovtsev, V. V.; Green, L. G.; Fokin, V. V.; Sharpless, K. B., *Angew Chem Int Edit* **2002,** *41* (14), 2596-+.
212. Wang, W. L.; Wu, J. L.; Xia, C. H.; Li, F. W., *Green Chem* **2011,** *13* (12), 3440-3445.
213. Gaulier, C.; Hospital, A.; Legeret, B.; Delmas, A. F.; Aucagne, V.; Cisnetti, F.; Gautier, A., *Chem Commun* **2012,** *48* (33), 4005-4007.
214. Enders, D.; Breuer, K.; Raabe, G.; Runsink, J.; Teles, J. H.; Melder, J. P.; Ebel, K.; Brode, S., *Ange Chem Int Edit* **1995,** *34* (9), 1021-1023.

215. Schaper, L. A.; Hock, S. J.; Herrmann, W. A.; Kuhn, F. E., A*ngew Chem Int Edit* **2013,** *52* (1), 270-289.
216. Velazquez, H. D.; Garcia, Y. R.; Vandichel, M.; Madder, A.; Verpoort, F., *Org Biomol Chem* **2014,** *12* (46), 9350-9356.
217. Gibard, C.; Avignant, D.; Cisnetti, F.; Gautier, A., *Organometallics* **2012,** *31* (22), 7902-7908.
218. Furstner, A.; Alcarazo, M.; Cesar, V.; Lehmann, C. W., *Chem Commun* **2006,** (20), 2176-2178.
219. Scholl, M.; Ding, S.; Lee, C. W.; Grubbs, R. H., *Org Lett* **1999,** *1* (6), 953-956.
220. Nelson, D. J.; Nolan, S. P., C*hem Soc Rev* **2013,** *42* (16), 6723-53.
221. Zhang, Y.; Cesar, V.; Storch, G.; Lugan, N.; Lavigne, G., *Angew Chem Int Edit* **2014,** *53* (25), 6482-6486.
222. Lv, T. Y.; Wang, Z.; You, J. S.; Lan, J. B.; Gao, G., J *Org Chem* **2013,** *78* (11), 5723-5730.