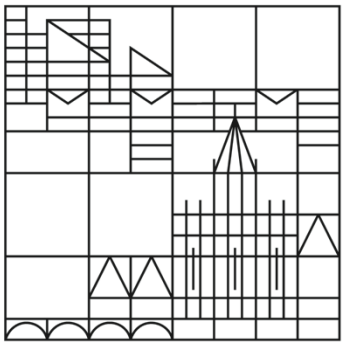
New Approaches for Detection of Epigenetic Markers in DNA and RNA

Dissertation submitted for the degree of
Doctor of Natural Sciences
(Dr. rer. nat.)

Presented by
Janina Carolin von Watzdorf

at the

Universität Konstanz



Faculty of Science

Department of Chemistry

Date of the oral examination: 13.01.2017

First referee: Prof. Dr. Andreas Marx

Second referee: Prof. Dr. Valentin Wittmann

Konstanzer Online-Publikations-System (KOPS)
URL: http://nbn-resolving.de/urn:nbn:de:bsz:352-0-395342

This work was prepared from 2012 to 2016 in the group of Prof. Andreas Marx (Chair of Organic and Cellular Chemistry) at the University of Konstanz, Faculty of Chemistry.

Parts of this work are published in:

[1] Janina von Watzdorf,* Kim Leitner,* Prof. Dr. Andreas Marx, „Modified Nucleotides for Discrimination between Cytosine and the Epigenetic Marker 5-Methylcytosine", Angew. Chem. Int. Ed. **2016**, *55*, 3229-3232.

Janina von Watzdorf, Prof. Dr. Andreas Marx, „6-Substitued-2-Aminopurine-2'-Deoxyribonucleoside-5´-Triphosphates that Trace Cytosine Methylation", ChemBioChem, **2016**, *17*, published online: 06.07.2016.

Dr. Claudia Huber,* Janina von Watzdorf,* Prof. Dr. Andreas Marx, "5-Methylcytosine-sensitive variants of Thermococcus kodakaraensis DNA polymerase", Nucleic Acids Res., **2016**, *44*, 9881 - 9890.

* contributed equally

Danksagung

In erster Line möchte ich mich herzlich bei Herrn Prof. Dr. Andreas Marx für die Aufnahme in seine Arbeitsgruppe und die Vergabe eines sehr vielseitigen und interessanten Themas bedanken. Ganz besonders möchte ich mich für die sehr fördernde, wie auch motivierende Betreuung und die außerordentliche Unterstützung in jeglicher Richtung bedanken. Ebenso möchte ich mich in diesem Zusammenhang für das in mich gesetzte Vertrauen und die vielen Freiheiten zur selbstständigen Bearbeitung und Gestaltung meines Themas bedanken.

An dieser Stelle möchte ich auch Herrn Prof. Dr. Valentin Wittmann für die Übernahme des Zweitgutachtens und Herrn Prof. Dr. Martin Scheffner für die Übernahme des Prüfungsvorsitzes danken.

Allen ehemaligen und aktuellen Mitgliedern der AG Marx danke ich für die tolle Arbeitsatmosphäre und ständige Hilfsbereitschaft. Im Besonderen möchte ich mich hierbei bei Kim Leitner für die sehr gute Zusammenarbeit bedanken. Ebenso danke ich meinen direkten Laborkollegen Stephan Hacker, Holger Bußkamp, Jana Balintová und Alexander Finke für die schöne Zeit.

Des weiteren danke ich allen Studenten, die meine Arbeit mit ihren Praktika, Bachelorarbeiten und Masterarbeiten unterstützt haben.

Bei Kim Leitner, Sarah Wallrodt, Eugenia Hoffmann, Odin Keiper, Daniel Rösner und Karin Betz möchte ich für das Korrekturlesen dieser Arbeit danken.

Allen Freunden, die ich an dieser Stelle nicht alle namentlich erwähnen kann, danke ich für die moralische Unterstützung, so wie für die schöne Zeit, die mich sämtlichen Laborstress vergessen ließ. Ein besonders großer Dank geht hierbei vor allem an Eugenia Hoffmann und Daniela Verga, die immer ein offenes Ohr für mich haben und mich zum Lachen bringen!

Mein ganz besonderer Dank gilt aber vor allem meiner Familie! Meinen Eltern möchte ich nicht nur für die finanzielle Unterstützung danken, die mir mein Studium erst ermöglicht hat, sondern vor allem auch für ihre mentale Unterstützung. Meinen Eltern, aber auch meinen Geschwistern, danke ich besonders dafür, dass sie jederzeit für mich da waren und immer an mich geglaubt haben. Vielen Dank - ohne euch wäre all dies niemals möglich gewesen!

Table of Contents

Danksagung ..... **V**

Table of Contents ..... **VI**

List of Abbreviations ..... **VIII**

1. Introduction **10**

1.1. DNA Modifications 11

1.1.1. 5-Methyl-2´- deoxycytosine (5mC) *11*

1.1.2. 5-Hydroxymethyl-2´- deoxycytosine (5hmC) *12*

1.2. Detection of DNA Modifications 13

1.2.1. Detection of 5mC *13*

Bisulfite Sequencing (BS-Seq.) 13

Methylation Specific PCR (MSP) 15

Methylation-Sensitive Single Nucleotide Primer Extension (Ms-SNuPe) 16

Bisulfite Pyrosequencing (PS) 16

Maxam-Gilbert Sequencing 16

Precipitation of Methylated DNA 17

Methylation Sensitive Restriction Enzymes 17

1.2.2. Detection of 5hmC *17*

TAB-Sequencing 17

oxBS-Sequencing 18

1.3. RNA modifications 18

1.3.1. Pseudouridine *19*

1.3.2. 2´-O-Methyl-Cytosine *20*

1.4. Real-Time DNA Sequencing from Single DNA Polymerase Molecules (SMRT) 21

1.5. Capillary Electrophoresis 22

1.6. DNA Polymerases 23

1.6.1. Chemical Mechanism of the Nucleotidyl Transfer *24*

1.6.2. DNA Polymerase Selectivity *25*

1.6.3. Kinetic Analysis of DNA Polymerase Action *25*

1.6.4. DNA Polymerase Families *26*

A-Family DNA Polymerases 26

B-Family DNA Polymerases 26

1.7. Modified Nucleotides 27

1.7.1. 5´-Triphosphate Synthesis *27*

2. Aim of This Work **30**

3. Results and Discussion **32**

3.1. Initial Screening for Discrimination between C and 5mC 32

3.1.1 Synthesis of modified nucleotides *33*

3.1.2. Screening of modified nucleotides to discriminate 5mC *35*

3.2. 6 - Modified dGTP Derivatives for the Detection of 5mC 37

3.2.1. Synthesis of 6-modified dGTP Derivatives *38*

3.2.1.1. O^6-Alkyl-dGTP 38

3.2.1.2. 6-Amino-dGTP 38

3.2.1.3. 6-Thioalkyl-dGTP 40
3.2.2 Screening of Modified Nucleotides for 5mC Detection 42
3.2.2.1. Testing KlenTaq DNA Polymerase for Incorporation of Modified Nucleotides 42
3.2.2.2. Testing KOD exo$^-$ DNA Polymerase for Incorporation of Modified Nucleotides 44
3.3. Capillary Gel Electrophoresis 48
3.3.1. Discrimination of 5mC by Emplyoing Modified Nucleotides and 9°North DNA Polymerase 50
3.4. Selectivity Studies for Incorporation of Modified Nucleotides 53
3.5. Doubly Modified dGTP Derivatives for Improved Discrimination of 5mC 56
3.5.1. Synthesis of 8-Br-O^6-Alkyl-dGTP 56
3.5.2. Incorporation Studies of Doubly Modified Nucleotides 57
3.6. Screening of KOD exo$^-$ Libraries Regarding Discrimination of 5mC 59
3.7 Detection of 5hmC 69
3.7.1 Discrimination of 5hmC Employing KlenTaq 69
3.7.2 Discrimination of 5hmC Employing KOD exo$^-$ 70
3.7.3 Discrimination of 5hmC Employing 9 °North exo$^-$ 72
3.8. Detection of RNA Modifications 75
3.8.1 Discrimination of 2´-O-Methyl-C by the Application of Modified Nucleotides75
3.8.2 Discrimination of Pseudouridine by the Application of Modified Nucleotides77
3.8.2.1. Kinetics for Incorporation of Modified Nucleotides 80
3.8.3. Ψ Detection Assay 81
3.8.3.1. RT-KTq2 F667Y for Detection of Pseudouridine 83
3.8.3.2. Real-Time PCR Experiments 84
3.8.3.3. Experiments employing RNA from extracts 88

4. Summary and Outlook 93

5. Zusammenfassung 97

6 Experimental Part 101
6.1 Chemical Synthesis 101
6.2. Molecular Biology/ Biochemical Methods 125

7. References 129

8. Appendix 135
8.1 NMR
8.2 HR-MS
8.3 Primer Extension Experiments

List of Abbreviations

5mC	5-methylcytosine
5hmC	5-hydroxymethylcytosine
5fC	5-formylcytosine
5caC	5-carboxylcytosine
A	adenine
abs.	absolute
ATP	adenosine triphosphate
BER	base excision repair
br s	broad singlet
C	cytosine
CE	capillary electrophoresis
cm	centimeter
d	doublet
d_6-DMSO	deuterated dimethyl sulfoxide
dd	doublet of doublets
DCM	dichloromethane
ddd	doublet of doublets of doublets
dGTP	2´-deoxyguanosine triphosphate
DMAP	4-dimethylaminopyridine
DMF	*N*,*N*-dimethylformamide
DMSO	dimethyl sulfoxide
dt	doublet of triplets
dq	doublet of quartets
DNA	deoxyribonucleic acid
dNTP	2´-deoxy nucleoside triphosphate
EDC	1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide
eq.	equivalent
FPLC	fast protein liquid chromatography
G	guanine
g	gram
h	heptet
h	hours
HPLC	high-performance liquid chromatography
HR-ESI-MS	high resolution electron spray mass spectrometry
Hz	Hertz
kg	kilogram
KO*t*Bu	potassium *tert*-butoxide
l	liter
M	molar
m	multiplet
min	minute
mM	millimolar
MHz	megahertz

mg	milligram
ml	milliliter
mmol	millimole
N	normal
NBS	*N*-bromosuccinimide (1-bromo-2,5-pyrrolidinedione)
nm	nanometer
nM	nanomolar
NMP	1-Methyl-2-pyrrolidone
NMR	nuclear magnetic resonance
P	pentet
PAGE	polyacrylamine gel electrophoresis
PNK	polynucleotide kinase
ppm	parts per million
q	quartet
RP	reversed phase
s	Singlet
SMRT	single molecule real time sequencing
T	tymine
t	triplet
TDA-1	tris[2-(2-methoxyethoxy)ethyl]amine
TEA	triethylamine
TEAB	triethyl ammonium bicarbonate buffer
TET	Ten eleven translocation
TLC	thin layer chromatography
TMP	trimethyl phosphate
U	uracile
UV	Ultraviolet
ZMV	zero mode waveguide
[v/v]	percent by volume
µl	microliter
µM	micromolar
µm	micrometer
µmol	micromole

1. Introduction

Due to the diverse functions nucleic acids have to conduct, they carry a wide range of different chemical modifications (see **Figure 1**). Modified nucleotides are employed to implement essential influences in various cellular processes in eukaryotic organisms.[2] The roles that modified nucleobases may play in a variety of cellular processes show the importance that nature places on structural diversity.[3] The diversity established by the four canonical nucleobases adenine (A), cytosine (C), guanine (G) and thymine (T) or uracil (U) is not sufficient to ensure the complex functions and regulation processes performed by nucleic acids.[3] Nucleic acid modifications therefore provide additional layers of complexity on biological regulation. Contrary to previous opinions that those modifications are static and only play fine-tuning functions, recent results point to a rather dynamic regulation.[2]

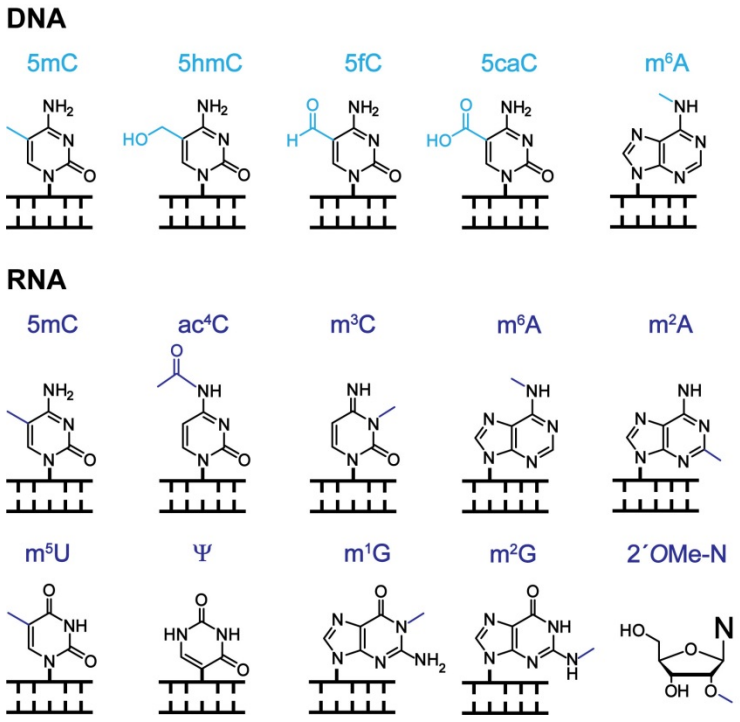


Figure 1: Selected modifications found in nucleic acids.

Our understanding of nucleic acid modifications has expanded over the past few decades, nevertheless the precise functions and regulations stayed elusive due to the lack of sufficient detection methods. Therefore, much research still needs to be done. To fully understand how these modified nucleobases are controlled and what roles they fulfil, improved detection techniques are required.

1.1. DNA Modifications

DNA methylation has been identified as key player in the epigenetic regulation of gene expression. 5-Methyl-2´-deoxycytosine (5mC), known as most common mark of DNA methylation, plays such a distinct role in cellular processes impacting development and gene expression, that it is considered to be the "5th base" and has been studied for decades.[4] In 2009, two groups simultaneously reported the discovery of an oxidised version of 5mC: 5-hydroxymethyl-2´-deoxycytosine (5hmC). They could not only show that 5hmC is present in mammalian cells, but could prove that 5hmC was generated by oxidation of 5mC by the family of ten-eleven-translocation (TET) oxygenases. Since then, many results did not only point at 5hmC as intermediate in 5mC demethylation pathways, but as additional epigenetic marker as well.

1.1.1. 5-Methyl-2´- deoxycytosine (5mC)

The existence of 5mC was first reported in 1951 by Wyatt *et al.* but its precise function as important factor in gene regulation stayed elusive for decades. Now, cytosine methylation has been identified as an important factor in gene regulation in mammalian cells, as many results suggested the repressive nature of 5mC.[5]

The activity of so called writer enzymes, mammalian methyltransferases, was already detected early on in 1968.[6] But not until two decades later, the first DNA methyltransferase (Dnmt1) was purified and its function studied in more detail.[7]

The occurrence of DNA methylation is generally associated with the repression of gene expression. Early studies showed that the application of 5-azacytidine, which acts inhibitory on DNA methylation, led to a reactivation of silenced genes.[5c, 8] Later studies employing *dnmt1* knockout mice further revealed that the loss of methylation resulted in the reactivation of several naturally silenced genes.[5b] Furthermore, numerous proteins that recognise 5mC (MBD1, MDD2, MBD4, MeCP2) were found and identified to be involved in 5mC-dependent transcriptional repression.[9] The discovery and characterisation of those 5mC readers led to a more profound understanding of the regulatory effect of cytosine methylation in DNA. The repression of gene expression was thereby suggested to be caused by the specific binding of the reader proteins. Thus, it could be shown that the presence of DNA methylation in the promoter region is directly connected to repression of transcription.[10] In contrast to the repressive effect of 5mC in the promoter region, DNA methylation in the gene body was identified to show positive correlation with gene expression, which further highlights the diverse functions of 5mC in gene expression.[11] The mentioned functions of 5mC in transcription regulation typically synergise with various histone modifications, as the different enzymes, known to be related to cytosine methylation or its interactions, network with various histone marks or histone modifying enzymes.[2] As 5mC was found to be strongly connected to gene expression, cell-type specific variations in methylation patterns were identified.[11e] Hence, tissue-specific genes were found to be highly methylated in most tissue cells but undermethylated in their tissue of expression.[12] Additionally, the promoter region of housekeeping genes were found to be constitutively unmethylated in every tissue.[12]

Above mentioned regulatory effects of methylation led to the suggestion that this process has to be very dynamic and reversible.[13] Although the writer and reader proteins of 5mC were found, the identity of eraser enzymes remained undiscovered. In 2009 the methylcytosine dioxygenases called ten-eleven translocation (TET) proteins were identified to oxidise 5mC to 5hmC.[14] Further studies demonstrated that the action of TET enzymes does not only lead to oxidation of 5mC to 5hmC, but results in the higher oxidised species 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) as well.[15] An active demethylation process can thereby be performed as both 5fC and 5caC are recognised and excised by the human thymine DNA glycosylase (TDG). Subsequent base excision repair (BER) replaces the modified cytosine derivative with its unmodified counterpart.[15a, 16] Additionally, cell division may also lead to passive dilution of the oxidised 5mC variants to the unmethylated stage.[17]

Due to the close relationship between cytosine methylation and levels of gene expression in eukaryotic genomes, 5mC has been connected to embryonic stem cell differentiation, genomic imprinting, cellular development and carcinogenesis.[14b, 14c, 18] The level of epigenetic cytosine-methylation needs to be precisely regulated since changes of the occurrence or distribution of 5mC have been shown to lead to severe genetic malfunctions.[19] Some human cancers have for example been associated with aberrant TET activity.[20] Reduced 5hmC abundance due to downregulation of TET activity was additionally observed during tumor progression.[20] Due to its effects on tumorigenesis, 5mC possesses an important link to human health. In addition, detection of 5mC can serve as important biomarker for diagnostics as well as disease therapy.

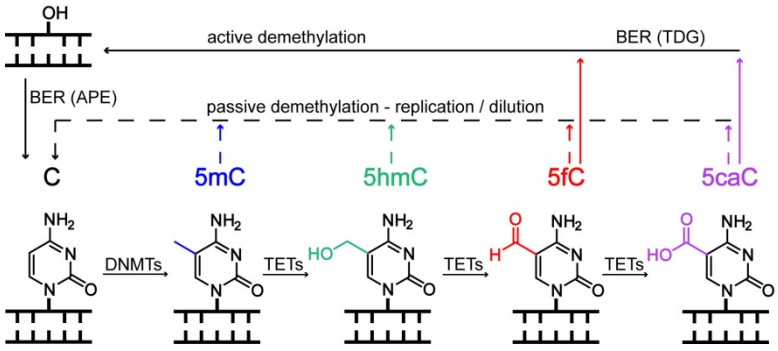


Figure 2: Reversible cytosine methylation in DNA. Demethylation mechanisms based on oxidation of 5mC. Solid line: glycosylase / BER mechanism, which proceeds over abasic site through the thymine DNA glycosylase (TDG). The action of the apyrimidinic endonuclease (APE) results in removal of the abasic site, followed by insertion of a new dC. Derived from Chen *et al.*[2]

1.1.2. 5-Hydroxymethyl-2´- deoxycytosine (5hmC)

As discussed above, methylation of cytosine at the C5 position in DNA is a crucial epigenetic process.[21] Different oxidation products of 5mC were detected and proposed to be involved in active demethylation of cytosine. However, recent discovery of 5-hydroxymehtyl-2´-deoxycytosine (5hmC) in mammalian tissue DNA suggests that 5hmC plays an important biological role with potential medical significance.[22] 5hmC has been found to be widespread in many tissues and cell types, although with varying levels of abundance (< 0.1 - 0.7 % of all cytosines).[23] As the levels of 5hmC have been found to be relatively stable within many tissues and in embryonic stem cells, the question arises if 5hmC

could have additional functions besides serving as intermediate in active demethylation pathways.[24] It has been shown that 5hmC levels are comparatively high in embryonic stem cells and remarkably reduced following differentiation into embryoid bodies.[14c, 25] Additionally, 5hmC has been found to be enriched at binding sites for pluripotency-associated transcription factors. Binding of those proteins to DNA is hindered by 5mC; therefore 5hmC may relieve the repressive effect of 5mC, further indicating its regulatory effect.[26] The suggested link between 5hmC and pluripotency highlights the importance to identify and discriminate between the different oxidation levels of 5mC.

1.2. Detection of DNA Modifications

As mentioned above, DNA modifications are known since decades but their precise functions remained elusive due to the lack of appropriate detection methods. However, a lot of research was done - especially on the field of 5mC - and many different methods for detection are available. Even if many approaches for sequencing or detection of different epigenetic marks were published, they all hold several disadvantages. Therefore, new sequencing methods are required to enable deeper insights into the role of those modifications. Due to its important link to human health, simple and easy accessible detection methods are required to allow site-specific detection of epigenetic marks for the application in personalised medicine.

1.2.1. Detection of 5mC

Different concepts for the discrimination between cytosine and 5mC have been described and rely on affinity enrichment,[27] endonuclease digestion,[28] nanopore sequencing,[29] specific interactions of proteins with 5mC[30] or different chemical behaviour concerning redox reactivity[31] or selective deamination of C using sodium bisulfite.[5a]

Bisulfite Sequencing (BS-Seq.)

Bisulfite sequencing is a genomic sequencing method that provides positive identification of 5mC with single nucleotide resolution and has become routine for the genome wide detection of 5mC.[32] This method relays on the comparison of two different sequencing runs before and after treatment of the respective DNA with sodium bisulfite. As the method is based on the selective bisulfite mediated deamination of C to uracil (U) in the presence of 5mC, the sites of cytosine methylation can be revealed by comparison of those two sequencing runs (see **Figure 3**).

Hayatsu and Shapiro reported already in 1970 the specific deamination of cytosine derivatives in the presence of sodium bisulfite on single nucleosides.[32-33] Due to the susceptibility of the ring system to nucleophilic attacks they could show that sodium bisulfite reversibly adds to the 5, 6 double bond of the pyrimidine nucleobases. Treatment of cytosine with sodium bisulfite under acidic conditions leads to a sulfonated adduct. However, bisulfite does not react further to displace the amino group. Adjusting the pH to basic conditions converts the cytosine adduct to uracil upon elimination. It was further discovered that the deamination of 5mC to thymine (T) via the corresponding sulfonate adduct is

nearly two orders of magnitude slower than for the unmethylated cytosine. BS-Seq. was following established taking advantage of this rate difference during deamination (see **Figure 3**).

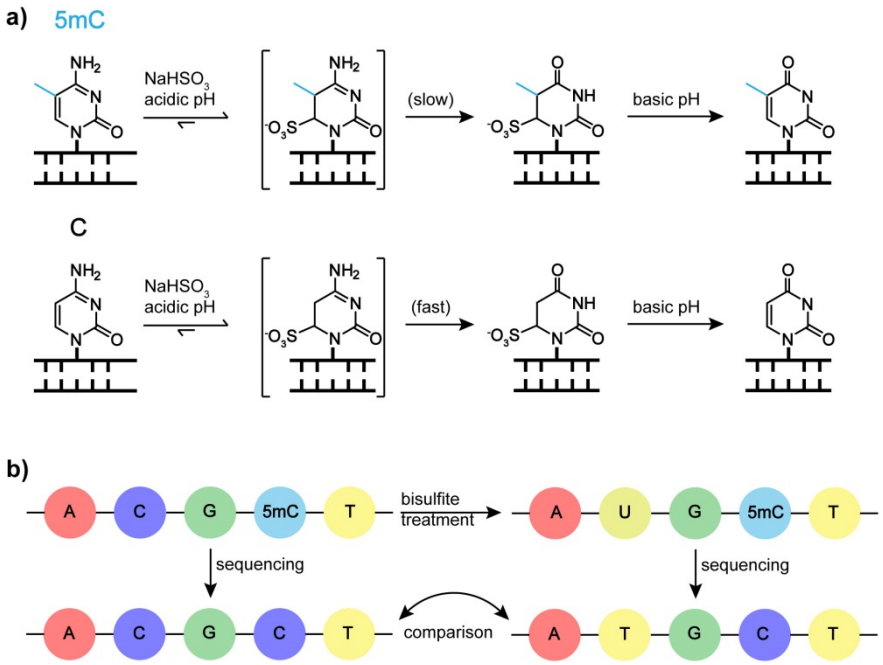


Figure 3: Bisulfite-sequencing (BS-Seq.). **a)** Bisulfite mediated deamination of C and 5mC; **b)** bisulfite-sequencing. Conventional sequencing approaches read 5mC as C. Bisulfite treatment of the genomic DNA converts C to U, while 5mC remains unreacted. After sequencing, C will therefore be read as U, while 5mC will be sequenced as C. All positions read as C, indicate a 5mC. Comparison of both sequencing runs reveals the positions of the unmethylated Cs. Modified after Miller.[4]

To use this difference in bisulfite mediated deamination for genomic sequencing of 5mC residues, total genomic DNA is fully denatured and treated with a freshly prepared sodium bisulfite solution under conditions, such that C is selectively converted to U, but 5mC remains unreacted. After incubation of the denatured DNA with sodium bisulfite, unreacted bisulfite needs to be removed by dialysis and the pH of the solution has to be adjusted to basic conditions upon addition of sodium hydroxide to achieve stoichiometric deamination of all sulfonated cytosine adducts in presence of the corresponding 5mC adducts.

The second part of this procedure involves PCR amplification of any region of interest in the bisulfite reacted DNA to yield a fragment in which U (formerly C) and T residues have been amplified as T and only 5mC residues haven been amplified as C. Due to bisulfite mediated conversion of C to U, the bisulfite reaction yields products in which opposite strands are no longer complementary. 5mC will be identified by a subsequent sequencing run, as all positions which are still read as C represent the positions of 5mC. Comparison of the output of conventional sequencing methods and bisulfite sequencing reveals the location of unmethylated C in the DNA strand.[5a] Exact methylation maps can readily be established by comparison of the sequencing outputs before and after bisulfite treatment.[5a, 34]

Despite the potency this method offers for genome wide 5mC detection with single-base resolution, BS-Seq. possesses several drawbacks.

The key factor for successful 5mC detection is based in the initial denaturation, as this step is critical for both PCR yield and full deamination. Especially CpG-rich sequences are often found to escape complete denaturation, resulting in incomplete conversion of the DNA, therefore simulating too high methylation rates.[35] Specific and stoichiometric deamination can be affected by several factors. It was shown that the bisulfite solution needs to be prepared fresh immediately before use. Additionally, it needs to be taken care that reaction temperature as well as time is precisely controlled, as longer incubation times and higher temperatures lead to increased rates of DNA strand brakes. The removal of remaining bisulfite and controlled conditions during alkali treatment proved to be essential to ensure that no deamination of 5mC occurs. The last and perhaps most critical step is the design of appropriate primers for the PCR. Those primers should be designed to favour amplification of fully bisulfite-converted DNA from a mixture which may also contain partially converted oligomers. This requires the design of more than one primer to every target sequence. In addition, PCR amplification needs to be optimised for every DNA sequence of interest.[34]

Those drawbacks require precise reaction control of all steps of BS-Seq. and quite big amounts of sample material, as two sequencing runs are required for comparison and acidic bisulfite treatment results in the loss of 95 % of genomic DNA.[36] Additionally, BS-Seq. is time consuming and tedious as two sequencing runs need to be compared.

However, many DNA methylation analysis techniques were established based on the bisulfite modification reaction, as bisulfite treatment can be performed on a large spectrum of DNA from different origins and amounts.[35]

Methylation Specific PCR (MSP)

For MSP separate primers need to be designed, either specific for methylated or unmethylated DNA, including the DNA region of interest. All unmethylated cytosines will be converted to U by bisulfite treatment, while 5mC stays unconverted. Subsequent to bisulfite treatment, two different PCR runs need to be performed each employing one primer either specific for C or 5mC. The results can be visualised by agarose gel electrophoresis and comparison between both PCR runs clearly points to all cytosine residues being methylated.[37]

This method provides a simple and easy way for the methylation analysis of a small number of 5mC loci. However, several drawbacks hinder this interesting tool from broader application. As described above, bisulfite mediated conversion of CpG-rich sequences is often deficient, leading to an error-prone output. Therefore, appropriate controls (fully methylated and unmethylated DNA) need to be performed. Additionally, the exact annealing temperature has to be determined for every sequence context, as low stringency can result in primer mispairing.[35] When using PCR primers spanning more than one possible methylation site, the effect of every single site on primer binding and MSP-outcome needs to be evaluated. Additionally, no quantitative data can be obtained reliably; as only qualitative detection of methylation can be achieved the methylation threshold needs to be determined for every site under investigation.[37] Therefore, this technique offers an interesting tool for site specific qualitative 5mC detection, but proves to be tedious if varying methylation sites have to be studied as the experimental setup needs to be optimised for every single investigated methylation site.

Methylation-Sensitive Single Nucleotide Primer Extension (Ms-SNuPe)

Ms-SNuPe is another site-specific detection method taking advantage of bisulfite-mediated conversion of C to U in the presence of 5mC. Genomic DNA is treated under conditions described for BS-Seq. and consecutively amplified by PCR using gene specific primers. The resulting products are purified by gel electrophoresis and a primer is annealed adjacent to the methylation site of interest. The DNA polymerase-catalysed incorporation of radioactively labelled dCTP or dTTP using the resulting primer template complex is analysed. The amount of radioactivity incorporated into DNA by applying dTTP in the presence of the DNA polymerase correlates with the amount of unmethylated C, while the amount of integrated radioactivity in the presence of dCTP corresponds to the amount of methylated DNA.[38] Despite the simple way for site-specific 5mC detection provided by this method, this approach holds the same drawbacks as described above.

Bisulfite Pyrosequencing (PS)

Bisulfite converted DNA regions of interest are amplified by PCR. The respective products are isolated and purified with the help of biotinylated amplification primers. A sequencing primer is annealed with the amplified DNA and a defined mixture of different enzymes is added to visualise incorporation. The nucleotides are sequentially added to the reaction mixture according to the known sequence of the DNA to analyse. Pyrophosphate is released during the incorporation process. The released pyrophosphate is enzymatically processed to enable a luciferase dependent reaction. The intensity of the emitted light signal is recorded by a camera. The signal intensity is thereby correlated with the number of base repeats in the DNA sequence. Remaining dNTPs are following degraded to dNMPs by the action of an apyrase. The next nucleotide can subsequently be added to monitor its incorporation efficiency. When reaching a possible 5mC site, dTTP and dCTP have to be added consecutively. The signal ratio obtained from these injections correlates with the amount of methylation at this locus.[35] This approach holds the advantage that whole DNA regions can be examined and no radioactive or gel-based detection methods are required. However, several disadvantages deriving from bisulfite-mediated conversion of C to U remain.

Maxam-Gilbert Sequencing

All techniques described above are based on the bisulfite mediated conversion of C to U in the presence of 5mC. On the contrary, this technique relies on chemical reactions cleaving the DNA in a base specific manner under defined conditions.[39] Incubation of the DNA with dimethylsulfate and subsequent treatment with piperidine results in cleavage at guanine or adenine residues. Hydrazine treatment followed by piperidine treatment induces strand breaks at cytosine or uracil residues. The DNA fragments can subsequently be amplified and radiolabelled, followed by separation via electrophoresis and visualisation through autoradiography.[40] As hydrazine shows reduced reactivity to 5mC compared to C and U, the Maxam-Gilbert sequencing reaction can be applied for 5mC detection.[41] Even if this method circumvents the need to perform bisulfite-mediated conversion, the hydrazine-dependent cleavage of genomic DNA holds similar disadvantages.

Precipitation of Methylated DNA

DNA needs to be fragmented to sizes of 200 - 1500 bp using mechanical or enzymatic methods. In contrast to already described methods, the DNA will not be altered or cleaved but detected by 5mC recognising proteins. Complexes of DNA and 5mC binding proteins are captured by standard precipitation techniques. After unspecific bound DNA was removed by stringent washing steps, the bound DNA can be eluted from the protein complexes. In contrast to the methods described above, no chemical reaction leading to a modified DNA-molecule is required. However, sequence bias can result from possible sequence preferences during DNA binding.

Methylation Sensitive Restriction Enzymes

Restriction endonucleases have found widespread application in molecular biology, as they have been shown to recognise and cleave DNA target sequences in a specified manner.[40] Interestingly, the cleavage activity of those enzymes does not only depend on the sequence but also on DNA modifications in their recognition sequence.[35] Therefore, a restriction endonuclease sensitive to 5mC in a CpG in its target sequence can be used for 5mC detection. This method provides a robust and easy way for site-specific 5mC detection. However, the dependence of this approach on the availability of specific recognition sequences does not allow widespread application.[35]

1.2.2. Detection of 5hmC

New methods for epigenetic sequencing are needed after the discovery of 5hmC, as bisulfite sequencing alone is not capable to distinguish between both modifications.[4] Hayatsu already reported in 1979 that 5-methylenesulfonate, formed during treatment of 5hmC with sodium bisulfite, undergoes deamination even more slowly than 5mC and is therefore also read as C when amplified and sequenced.[42] Therefore, He[43] and Balasubramanian[44] established modified BS-Seq. protocols that provide base-pair resolution of 5hmC.[4] Both methods introduce an additional step leading to selective chemical transformations prior to bisulfite treatment.

TAB-Sequencing

It has already been reported before that 5caC behaves like C during bisulfite-mediated conversion, meaning it will be read as T in subsequent sequencing (see **Figure 4 b**).[15a] In this approach He *et al.* use the ability of Tet1 to oxidise 5mC all the way to 5caC in the presence of C.[43] They envisioned that they could identify the 5hmC loci by comparison of three different sequencing runs. Through conventional BS-Seq. they can identify all sites possessing the epigenetic marks 5mC or 5hmC. To distinguish between 5mC and 5hmC an additional BS-Seq. run is required after 5mC is selectively converted to 5caC. Therefore, 5hmC needs to be protected to circumvent oxidation of 5hmC to 5caC during Tet1 treatment. During TET-assisted BS-Seq. (TAB-Sequencing), they use the enzyme β-glucosyl tranferase (βGT) to selectively convert 5hmC to β-glucosyl-5-hydroxymethylcytosine (5gmC). After protection of 5hmC as 5gmC, the DNA is treated with excess Tet1 to oxidise all 5mC residues to 5caC. Subsequent BS-Seq. converts all C and 5caC bases (formerly 5mC) to U while

5gmC remains unaffected. After amplification (5gmC amplifies to C and U to T) and sequencing, all sites that are read as C represent 5hmC loci (see **Figure 4a**).

oxBS-Sequencing

The approach developed by Balasubramanian *et al.*[44] is conceptually similar to He´s method but delivers a different output. While TAB-Seq. uses the selective oxidation of 5mC to 5caC, oxidative BS-Seq. utilises the oxidation of 5hmC to 5fC in the presence of 5mC with $KRuO_4$. Similar to 5caC, 5fC undergoes bisulfite-mediated deformylative deamination to yield U and is amplified and sequenced as T. Therefore, genomic DNA is treated with $KRuO_4$ to oxidise all 5hmC residues to 5fC, while 5mC stays unreactive. After subsequent BS - Seq. and comparison between the different sequencing runs, all 5mC and 5hmC loci can be distinguished as 5mC will be read as C and 5hmC as T (see **Figure 4c**).

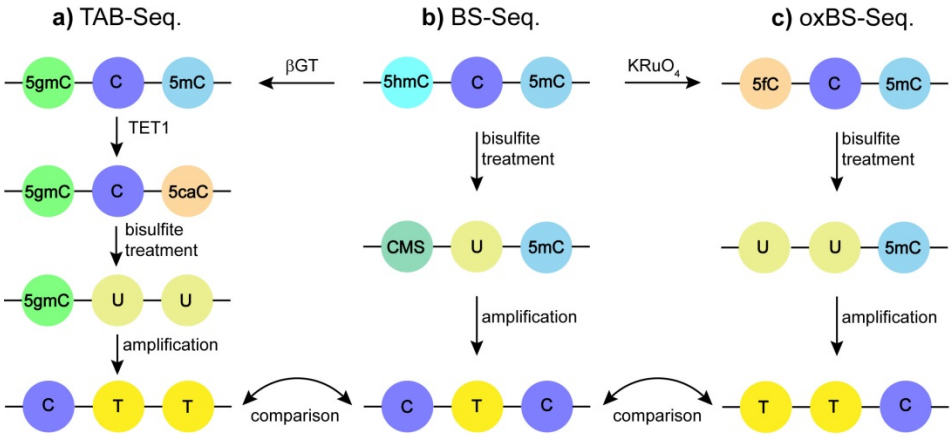


Figure 4: Comparison of BS-Seq. (**b**), TAB-Seq. (**a**) and oxBS-Seq. (**c**). Figure derived from Schüler.[4]

While both detection methods generate different outputs, they ultimately yield the same information by comparison of three different sequencing runs.[4] Both methods can readily be used for 5hmC detection. Nevertheless, both methods possess several drawbacks. The main disadvantages rely on the need of three different sequencing runs, as comparison of those is tedious and error-prone. Additionally, methods based on bisulfite-mediated conversion own all disadvantages already described for BS-Seq. (see **Bisulfite Sequencing (BS-Seq.)**). Therefore, new sequencing methods are required for detection of the oxidised epigenetic marker 5hmC as well.

1.3. RNA modifications

Modified nucleotides are not only found as regulatory elements in DNA, but even more widespread in RNA.[2] Chemical alterations are thereby not restricted to modifications at the nucleobase. In fact, modifications are known to occur at a variety of positions all over the four canonical nucleotides.[3] Modified nucleotides are suspected to be present in all RNA species, although the highest diversity

has been detected in tRNA. Those modified nucleotides play important roles in stabilisation of RNA structures, fine-tuning of numerous interactions between RNA itself or RNA-binding partners or protection against nucleolytic degradation.[45] While some RNA modifications are known for decades, their occurrence and distribution and thereby their function still remain elusive.[3] Modifications of RNA were divided into three categories in regard to their suggested function in the cellular environment. (1) Modifications which add an additional, dynamic layer of information on top of the primary sequence, as mRNA modifications. (2) Modifications that refine molecular recognition by expanding the RNA vocabulary, such as modifications in the decoding region of tRNA. (3) Modifications which are able to tune RNA biogenesis by enforcing certain RNA structures, as modifications in rRNA and snRNA.[46] The diverse functions of RNA modifications explain the huge diversity of their occurrence. Much research needs to be done in this dynamic and versatile field. Therefore, new detection methods are required for many RNA modifications.

1.3.1. Pseudouridine

Psuedouridine (Ψ) was the first posttranscriptional modification discovered in RNA and is one of the most abundant, as it is highly conserved across species and is present in a wide range of cellular RNAs.[47] Ψ-Synthetases can catalyse the base-specific isomerisation of uridine (U) to Ψ by a very unusual transformation.[47b] Initially, the nucleobase gets liberated by cleavage of the glycosidic bond. The nucleobase can then be rotated 180° along the N3-C6 axis. Eventually, the nucleobase gets connected to the sugar moiety by establishing a new bond between C5 of the nucleobase and C1´ of the sugar (see **Figure 5a**).[47b] Thus, in Ψ the nucleobase is connected to the sugar moiety via a rigid C-C bond. Just as Ψ has an unaltered Watson-Crick base pairing moiety, but possesses an additional hydrogen bond donor at its non-Watson-Crick site (see **Figure 5a**). Due to those structural changes, Ψ can alter the RNA secondary structure by increasing base stacking, improving base pairing via the non-Watson-Crick edge and rigidifying the sugar-phosphate backbone.[48] Thus, chemical and physical properties of RNA molecules can be varied, if Ψ gets incorporated. Unlike other posttranscriptional modifications, the isomerisation from U to Ψ seems to be irreversible, since the glycosidic bond is converted into a more stable C-C bond. The irreversibility of this transformation suggests distinct roles of Ψ.[49] Pseudouridinylation in mRNA has been suggested to perform regulatory roles in mRNA metabolism, as conversion of U to Ψ, which affects stability of mRNA, seems to be tuned in response to environmental changes and stresses.[50] Despite its obvious importance, distribution and function of Ψ are largely unexplored due to missing detection methods.[49] The first transcriptome-wide maps of Ψ were only published in 2014 employing a selective chemical-labelling approach.[50] As already reported for the detection of other nucleic acid modifications, Ψ can be discriminated against U in RNA by its specific chemical properties.[51] RNA fragments (100 - 300 nt) are incubated with CMCT (1-cyclohexyl-(2-morpholinoethyl)carbodiimide) followed by alkaline treatment. The carbodiimid moiety of CMCT thereby modifies N1 of all G residues as well as N3 of U and N1 and N3 of Ψ (see **Figure 5b**). Subsequent alkaline treatment removes all modifications, except those positioned at the N3 position of Ψ. This very bulky modification acts as barrier to reverse transcription, allowing the discrimination between U and Ψ.[52]

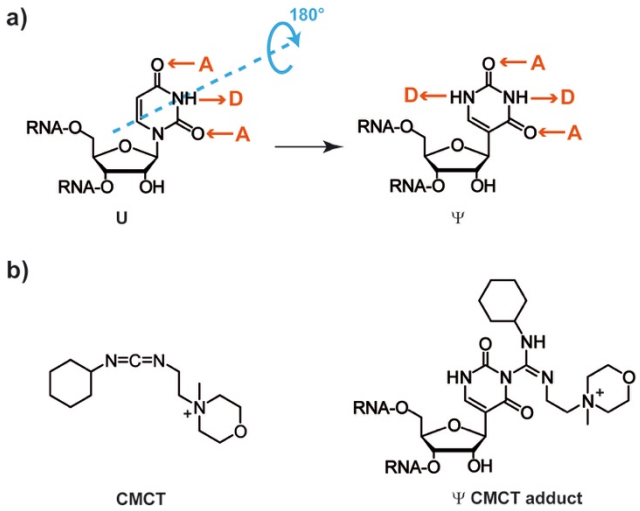


Figure 5: Structure and detection of Pseudouridine. **a)** Schematic representation of U to Ψ isomerisation; **b)** structures of CMCT and the Ψ-CMCT adduct.

Despite its utility, this detection method possesses all drawbacks which were already discussed for detection methods, using chemical modifications reactions prior to sequencing.

1.3.2. 2´-*O*-Methyl-Cytosine

Another widespread modification of RNA is methylation of the 2´-OH group of the ribose moiety, which is present in all major classes of eukaryotic RNA.[53] It has been shown that this modification occurs predominantly around functionally important regions.[54] This observation suggests that 2´-*O*-methylation contributes in maintaining ribosome functions. Additionally, 2´-*O*-methylated nucleotides have been found within the cap structure of mRNAs and are involved in differentiation between self and non-self RNAs.[55] Detection of 2´-*O*-methylation is important to get deeper insights into the mechanistic and roles of this posttranscriptional modification. Especially, as RNA was recently connected to several diseases due to its regulatory function.[56] In spite of intensifying studies in this field, sufficient detection methods are still missing. Many methods were presented to sense 2´-*O*-methylation based on reverse transcription or RNase H treatment. But they all hold the disadvantage to be technical demanding, laborious and deliver ambiguous results.[53]

One direct method for the site-specific detection of 2´-*O*-methylation was recently reported by our group.[53] A KlenTaq DNA polymerase variant with reverse transcriptase activity (RT-KTq2) was shown to be hindered by the presence of 2´-*O*-methylation in RNA. This feature of the DNA polymerase mutant was exploited to establish a methylation-sensitive qRT-PCR based assay. Thereby, the potential of DNA polymerases to act discriminatory against 2´-*O*-methylated nucleotides could be proven.[53] Even if this method is a great progress, as it is an easy detection assay, simple methods for whole sequencing approaches are still missing and needed.

1.4. Real-Time DNA Sequencing from Single DNA Polymerase Molecules (SMRT)

Sanger sequencing exploits the ability of some DNA polymerases to incorporate 3´-deoxynucleotides, which act chain-terminating.[57] Subsequent electrophoretic methods help to reveal the sequence of the investigated DNA strand. While this method relies on the low error rate of DNA polymerases, it does not exploit the potential of the used DNA polymerases for high catalytic rates or processivity.[58] Sanger sequencing is therefore time-consuming and read lengths are restricted to comparatively short DNA strands. Additionally, read out of those experiments is tedious and time-consuming as the number of samples that can be analysed in parallel is restricted, therefore limiting the scope of analysis. By increasing the speed as well as the length of individual sequencing reads, sequencing can be accelerated and costs can be reduced to reveal large-scale genomic complexity.[59]

Thus, Pacific Biosciences developed a method for real-time sequencing of single DNA molecules (SMRT).[60] This approach is based on the template-based DNA polymerisation activity of DNA polymerases, utilising its intrinsic characteristics regarding speed, fidelity and processivity with fluorescence-labelled nucleotides.

For this purpose, single DNA polymerase molecules are immobilised at the bottom of zero-mode waveguide (ZMW) nanostructure arrays. Those arrays consist of holes with a diameter of ~100 nm in a ~100 nm thick metal film which is deposited on a transparent substrate (see **Figure 6a**). Thereby, each ZMW becomes a nanophotonic visualisation chamber with a detection volume of just ~100 zeptoliters. This small volume enables the observation of a single nucleotide incorporation event despite the background of relatively high concentrations of diffusing fluorescence-labelled nucleotides.[60]

Nucleobase-labelled nucleotides are poorly incorporated in consecutive positions and would lead to high levels of background, as the fluorescent dyes cannot be cleaved of during the incorporation process. In contrast, by linking the fluorophore to the terminal phosphate moiety, phosphodiester bond formation during DNA polymerase catalysed incorporation results in the liberation of the dye. Thereby, a quantitative replacement of all natural dNTPs by the modified nucleotides can be achieved, as the fluorescent dye will be cleaved as part of the incorporation process, leaving behind a natural non-fluorescent DNA. For sequencing, each of the four nucleobases is labelled with a distinct fluorescent dye to enable discrimination between the different bases. During incorporation, the DNA polymerase holds the correct nucleotide including its colour coded fluorophore for several milliseconds in the ZMW, producing a fluorescent signal which can be detected. The corresponding fluorophore is thereby orders of magnitude longer in the ZMW than the average time associated with diffusing nucleotides (2 - 10 µs) or noncanonical sampling (< 1ms). Thus, SMRT allows real-time observation of DNA synthesis by employing the ability of DNA polymerases to perform uninterrupted template-directed synthesis using four distinguishable fluorescently labelled dNTPs (see **Figure 6**).[60]

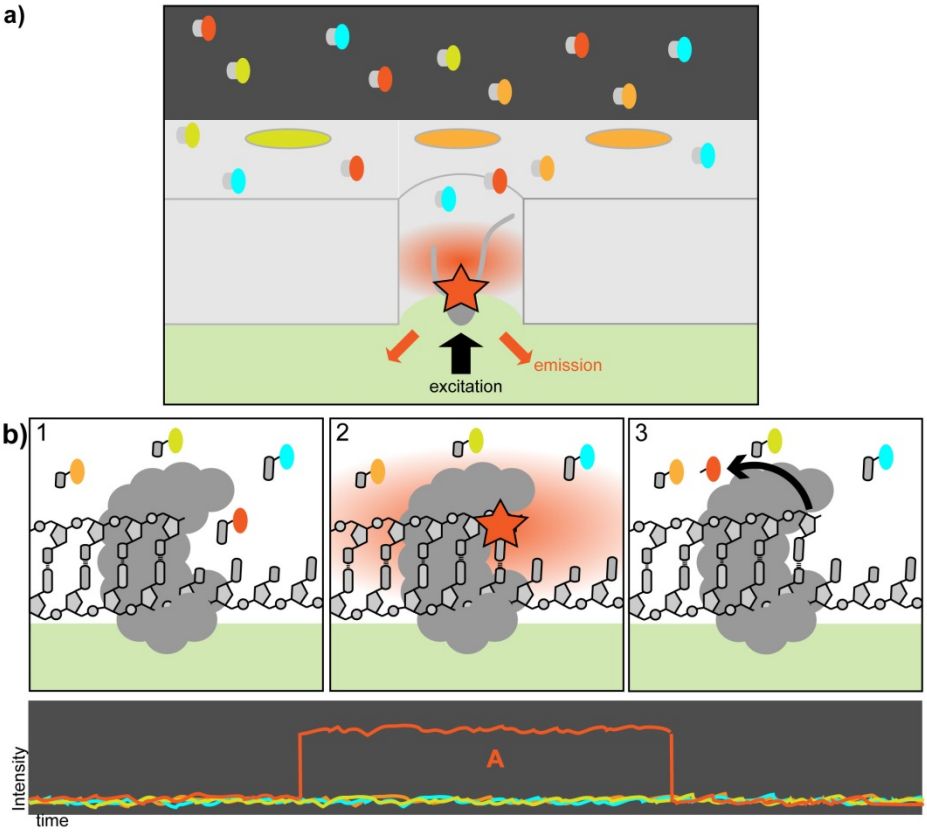


Figure 6: Principle of SMRT. **a)** experimental geometry; **b)** schematics of reaction steps involved in SMRT with a corresponding expected time trace of measured fluorescence intensity. Figure after Eid et *al.*[60]

1.5. Capillary Electrophoresis

General methods to study DNA polymerases usually detect enzyme activity by the measurement of a fluorescent or radioactive signal.[61] Polyacrylamide gel electrophoresis (PAGE) is widely used to analyse those fluorescently or radioactively labelled experiments. Despite its utility, analysis of those experiments by denaturing PAGE is time consuming and tedious. Since the number of samples that can be analysed on one gel is restricted, PAGE analysis is relatively inefficient and therefore limits the scope of enzyme analysis. Additionally, quantification of primer extension bands on PAGE gels is elaborate and requires manual scanning and analysis.[62]

Capillary gel electrophoresis (CE) is a new, alternative method, which already replaced the usage of sequencing gels in fluorescent Sanger DNA sequencing and therefore accelerated high-throughput sequencing of the human genome.[63] Analogous to separation using PAGE gels, CE separates fluorescently labelled nucleic acids according to their size and charge as they migrate through a polymer filled capillary.[62] After the CE samples are applied electrokinetically, high voltage electrophoresis allows single base resolution.[62]

Detection is achieved by laser excitation. Several different fluorescently labelled primers, possessing well separated excitation and emission spectra, can be detected in parallel allowing multiplexing. In addition, sample loading and data acquisition is automated and rapid, allowing several samples to be analysed in a short time.

1.6. DNA Polymerases

DNA polymerases catalyse the template-mediated incorporation of nucleotides into a growing primer strand in 5´- 3´ direction. Those enzymes perform repetitive cycles of nucleotide binding, base-pairing, nucleotidyl transfer, pyrophosphate release and movement to the next templating position. They adopt several different conformations, like an open and a closed state.[64] The enzyme action can be summarised by a general scheme. In principle, all enzymatic steps, shown in **Figure 7** are reversible.

In the first step (**1**), the DNA polymerase binds to the DNA primer/template complex resulting in a binary complex. Upon binding of an incoming dNTP, the open ternary complex is formed in step **2**. Nucleotide binding triggers the conformational change between the open and the closed ternary complex (step **3**), which is followed by the nucleotidyl transfer that takes place in step four (**4**). During processing of dNTPs, the highly energetic phosphorous anhydride bond is cleaved and pyrophosphate is liberated.[64] This cleavage takes place in a two-metal-ion mechanism, which will be described in more detail (**1.6.1. Chemical Mechanism of the Nucleotidyl Transfer**) and represents the driving force of this reaction. In step five (**5**) the DNA polymerase relaxes to its initial conformation followed by pyrophosphate release (step **6**).[65] The DNA polymerase may continue the synthesis reaction after translocation (step **7**), dissociate from the extended primer template complex to bind another DNA substrate (step **8**)[66] or initiate its exonuclease activity (step **9**).[64, 66-67]

Several kinetic studies exploring modified nucleotides tried to illuminate the question of the rate limiting step.[68] Despite the fact that those studies shed light on the complex mechanism, they failed to reveal one step as rate limiting for all different classes of DNA polymerases. In general, steps which lead to synthesis that is in conformity with Watson-Crick rules are found to be faster, generating kinetic selectivity of the reaction. As mentioned before, all steps were found to be reversible. Even hydrolysis of the anhydrous bond during nucleotidyl transfer can be reversed by pyrophosphorolysis, which generates dNTP by degradation of the primer strand. Therefore, dNTP concentration plays an important part in shifting the equilibrium of this process to one or the other direction.[69] This circumstance can be used in biotechnological applications, if incorporation of artificial nucleotides should be facilitated. The addition of pyrophosphatase results in direct hydrolysis of the released pyrophosphate and thereby hinders the degradation of the primer strand. With this approach incorporation of the artificial nucleoside-monophosphates can be improved.[70]

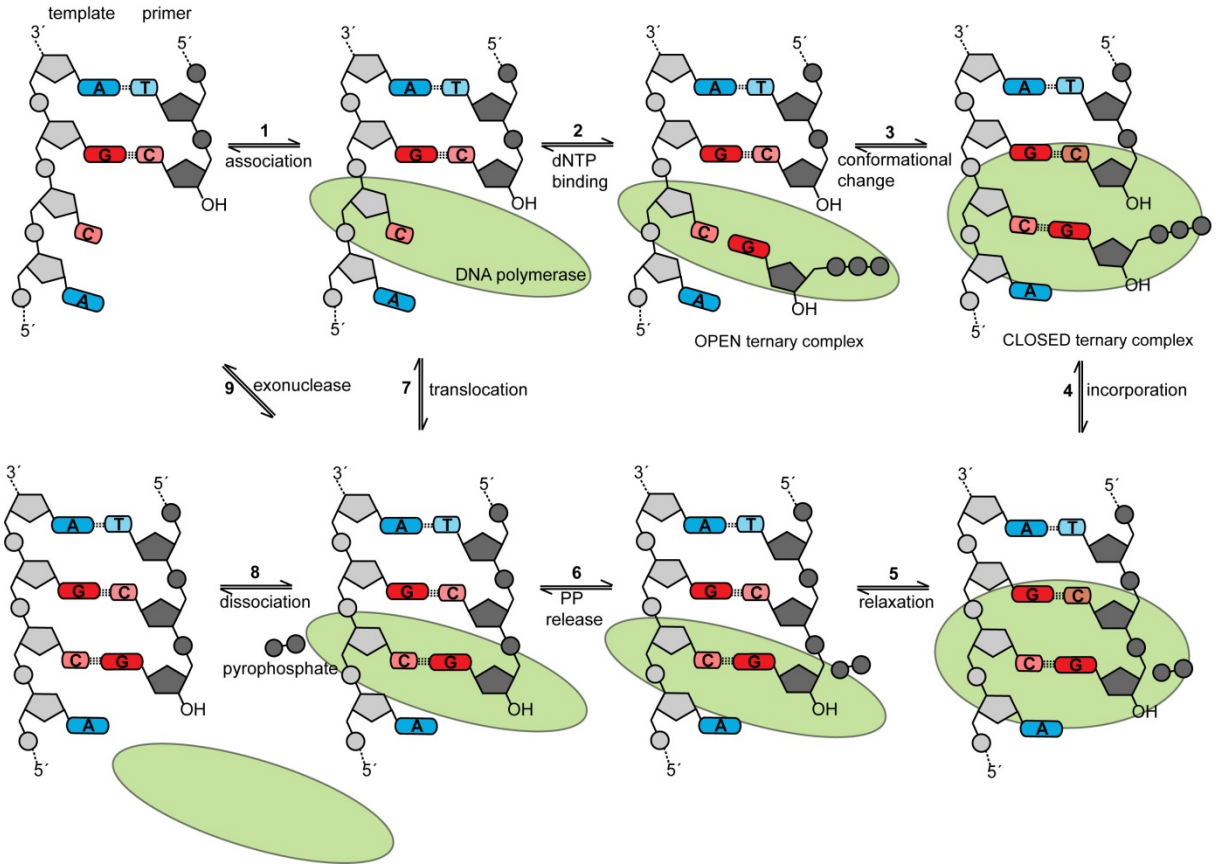


Figure 7: schematic representation of DNA polymerase catalysed DNA synthesis.

1.6.1. Chemical Mechanism of the Nucleotidyl Transfer

The chemistry required to elongate the DNA is a phosphoryl transfer reaction which proceeds in two distinct steps.[71] It is well established that a nucleophilic attack of the 3´-OH group of the primer end on the α-phosphorous of the incoming nucleoside-5´-triphosphate leads to hydrolysis of the triphosphate moiety. The incoming dNTP is accompanied by two metal ions (usually Mg^{2+}), which are coordinated within the active centre of the DNA polymerase between the phosphates of the nucleotide and two aspartic acid residues that are widely conserved among DNA and RNA polymerases.[71,72] Hydrolysis of the triphosphate chain by a two-metal-ion mechanism results in formation of a phosphodiester bond under the release of pyrophosphate. Thereby, a basic amino acid residue near the sugar moiety of the incoming dNTP abstracts the proton from the 3´-OH group to generate a more reactive nucleophile.[72] Metal ion A is surrounded by two water molecules and facilitates the 3´-oxygen attack on the α-phosphate by additionally activating the 3´-OH of the primer by lowering its pKa (see **Figure 8**). The metal ion contacts the 3´-OH of the primer that is perfectly positioned for an in-line attack. Subsequently, the electron-rich 3´-oxygen attacks the α-phosphate, creating a trigonal-bipyramidal pentacoordinated transition state that is stabilised through coordination of metal ion B with the oxygens of the β-γ-phosphate groups, thereby assisting leaving of the pyrophosphate.[73] This attack results in the inversion of the α-phosphate stereochemistry.[74] Once the reaction is complete, the pyrophosphate and the metal ions dissociate and the DNA must translate and rotate relative to the

DNA polymerase. Thereby, the newly formed primer terminus is correctly positioned in the primer binding site for another catalytic cycle.[75]

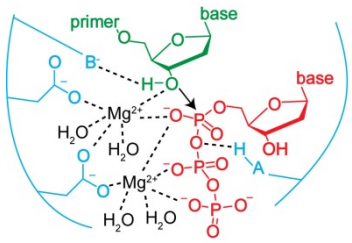


Figure 8: Extended two-metal-ion mechanism of the DNA polymerase catalysed nucleotidyl transfer. Figure was designed according to Castro *et al.*.[73a] Green: 3´-primer end, red: incoming dNTP, blue: DNA polymerase.

1.6.2. DNA Polymerase Selectivity

The DNA polymerase can discriminate against incorporation of non-canonical nucleotides at different steps. Generally, steps that lead to incorporation obeying Watson-Crick rules are faster, resulting in kinetic selectivity of this reaction. In a first step, binding of correct dNTPs occurs with higher affinity, than binding of mispaired substrates.[76]

The first step leading to selection of the correct nucleotide occurs during binding of the incoming dNTP and thereby generating the open ternary complex as the DNA polymerase can discriminate against binding of a non-canonical dNTP (step **2**).[76] The subsequent conformational transformation of the DNA polymerase to form the closed ternary complex facilitates an “induced fit” mechanism. During adjustment of the finger-domain, a binding pocket is formed which is shape complementary to a correct dNTP.[76] This tight fit of the nascent base pair to the active site pocket contributes to the selection of the correct dNTP since misaligned intermediates disrupt the geometry of the active site and hinders the chemical transfer reactions (step **3**).[76] The correct positioning of the dNTP in the active centre of the DNA polymerase is important to enable the phosphoryl-transfer occurring during incorporation of the correct dNMP.

1.6.3. Kinetic Analysis of DNA Polymerase Action

Two different methods are available for analysing kinetic data of DNA polymerase activity.[77] One method includes steady-state kinetic experiments conducted with varying dNTP concentrations. The concentration of DNA and dNTP substrates is maintained in molar excess versus the concentration of the employed enzyme to allow the DNA polymerase to incorporate only one nucleotide per primer/template complex. It has to be ensured that these experiments are conducted under “single-completed-hit-conditions”. Those conditions are mathematically fulfilled, if the primer extension rate is kept under 20 %.[79] Such experiments reveal the substrate dependence (K_M) of the catalytic turnover rate, as well as the maximum catalytic turnover rate (k_{cat} or v_{Max}) as described by Michaelis and Menten.[77] Means, those experiments reveal kinetic parameters over the whole reaction cycle, including binding and dissociation of the DNA polymerase from the primer/template complex.

Additionally, relatively low amounts of DNA polymerase are required and no specialised equipment is needed to obtain data.[64]

In contrast, pre-steady-state measurements are performed under “single-turnover-conditions” in regard to dNTP and the primer/template complex. These experiments allow the investigation of the polymerisation reaction on a millisecond time scale. Thereby, the dissociation constant (K_D) as well as the maximum turnover rate (k_{pol}) can be determined.[78] This includes examination of dNTP binding rates, conformational changes and formation of the phosphodiester bond.[79]

1.6.4. DNA Polymerase Families

Based on sequence homology and structural similarity, DNA polymerases can be divided into seven families (A, B, C, D, X, Y and RT).[80] Crystallographic studies revealed a characteristic right-hand shaped structure consisting of finger, thumb and palm sub-domains.[81]

A-Family DNA Polymerases

Family A DNA polymerases can be divided into replicative and repair enzymes.[82] The three repair enzymes DNA polymerase I from *E.coli*, *Thermus aquaticus* (Taq) and *Bacillus stearothermophilus* (Bst) are the most prominent members of family A DNA polymerases.[83] The DNA polymerase I possesses three functional activities that are located in three different domains. The N-terminal domain harbours the 5´-3´endonuclease activity, the central domain contains the 3´-5´exonuclease activity, which is not functional in all members, and the C-terminal domain, consisting of finger, thumb and palm subdomains, bears the DNA polymerase activity.[82] Klenow *et al.* could show in 1970 that an N-terminally cleaved fragment of the DNA polymerase I, named Klenow Fragment, maintains its 3´-5´exonuclease as well as its DNA polymerase activity while losing its 5´-3´ endonuclease domain.[84] The KlenTaq DNA polymerase is the N-terminally truncated form (amino acids 293-832) of the Taq (*Thermus aquaticus*) DNA polymerase, missing the 5´-3´endonuclease domain. The large fragment of the Taq DNA polymerase is the orthologue of the DNA polymerase I from *E.coli*, which is involved in nucleotide excision repair and in processing of Okazaki fragments during lagging strand synthesis in replication.[82] As the Taq DNA polymerase lacks the 3´-5´ exonuclease domain, the KlenTaq DNA polymerase only consists of the C-terminal domain associated with the DNA polymerase activity.[84-85] As the DNA polymerase KlenTaq derives from thermophile bacteria that tolerate high temperatures, it possesses a temperature optimum of 75-80 °C.[82, 86] Therefore, it exhibits a great potential for the application in various biochemical experiments as PCR.[86]

B-Family DNA Polymerases

The family B contains the DNA polymerase II from *E.coli*, eukaryotic replicative DNA polymerases, as well as DNA polymerases encoded on plasmids of mitochondria, various fungi and plants, viral and archeabacterial DNA polymerases and DNA polymerases of bacteriophages.[82] As members of this sequence family possess a distinct 3´-5´ exonuclease activity, they are mainly involved in DNA replication, carrying out processive DNA synthesis and are able to correct errors.[82] Archeal B-family DNA polymerases have been shown to be particular suitable for incorporation of modified nucleotides.

It could be shown that several B-family DNA polymerases are capable to accept a wide range of modifications, as dNTP analogues modified at the sugar[87] or nucleobase[88] moiety were successfully processed, in contrast to the employment of A-family DNA polymerases. So far, the reason for those diverging efficiencies in incorporation of modified nucleotides could not be elucidated. Different B-family DNA polymerases are widely used for biochemical approaches, as enzymes from *Thermococcus kodakarensis* (KOD), *Thermococcus* species 9° North or *Pyrococcus furiosus* (Pfu) to just mention a few.[89]

1.7. Modified Nucleotides

Due to their widespread applications in medicine, biology, chemistry, biochemistry and material science,[90] the chemistry of modified nucleosides, nucleotides or oligonucleotides continues to be a rapidly developing field. Modified nucleotides have been established for the investigation of many biochemical processes and therefore enhanced our current understanding.[90] Nucleotide analogues can not only be used to improve our understanding of many cellular processes, but can even be employed for medical applications, as the modified nucleotides can compete against their natural counterparts. The potential of different derivatives to be used for the treatment of various diseases has already been proven by their application as antiviral and anticancer drugs.[91] Thus, efficient methods for the synthesis of modified nucleotides are required, as they are of general interest and widespread importance.

1.7.1. 5´-Triphosphate Synthesis

Due to the great importance of phosphorylated biological molecules, several different phosphorylation methods were developed. The method established by Eckstein et *al.*[92] takes advantage of the multifunctionality of salicyl phosphorochloridite (see **Figure 9a**). Thereby, they developed a facile synthesis route for 5´-*O*-triphosphates, which could be employed for the generation of 5´-*O*-(1-thiotriphosphates) as well. They used 3´-OH protected nucleosides, which were reacted with salicylphosphorodichloridite to generate intermediate **I** as diastereomeric mixture. Treatment of **I** with pyrophosphate resulted in the formation of a cyclic phosphorous (III) species (**II**), which could be oxidised with iodine/water leading to the desired 5´-*O*-triphosphate in up to 72 % yield.[92] The nucleoside cyclic triphosphate resulting from oxidation of compound **II** cannot be detected due to immediate hydrolysis in the aqueous conditions used during oxidation. Alternatively, intermediate **II** can be reacted with sulphur to yield the corresponding 5´-*O*-(1-thiotriphosphate).[92] Despite the advantage to have one method which can result in the generation of 5´-*O*-triphosphates and 5´-*O*-(1-thiotriphosphates), this method requires the protection of the 3´-OH group thereby complicating its application. Since the 5´-OH group of nucleosides is more reactive than its 3´-OH group, selective protection of the 3´-OH group is rather challenging. To achieve selective 3´-OH protection, nucleosides need to undergo several selective protection and deprotection procedures.

Thus, Huang *et al.*[93] established a protection-free variant of the described method by generation of a mild and selective phosphitylating reagent that differentiates the different functionalities present in the nucleoside (see **Figure 9b**). Therefore, they took advantage of the multifunctionality and high reactivity of salicyl phosphorochloridite.[92, 94] By treatment of the salicyl derivative with pyrophosphate, they managed to generate a weak phosphitylating reagent **IV**. The selective phosphitylation reagent **IV** can be generated *in situ* and can selectively react with the 5´-hydroxyl group of nucleosides without the need for protection groups on the sugar or nucleobases.[93]

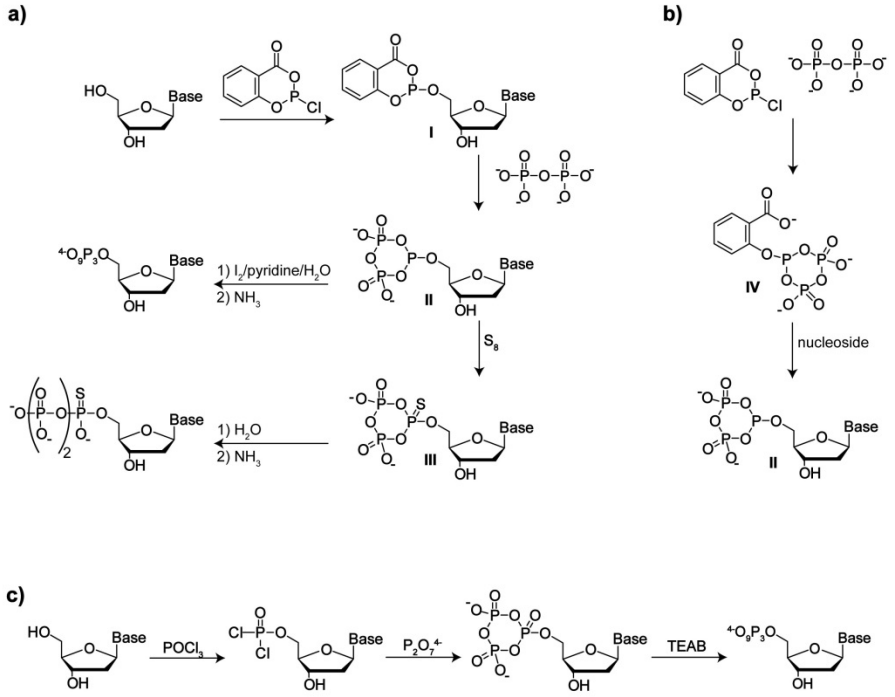


Figure 9: Methods for 5´-*O*-triphosphorylation. **a)** Synthesis according Eckstein, **b)** according Huang, **c)** according Yoshikawa and Kovács.

Another protection-free strategy for synthesis of 5´-*O*-triphosphates employs highly reactive phosphorous oxychloride ($POCl_3$) as phosphitylation reagent (see **Figure 9 c**). Trimethylphosphate is used as solvent to reduce the reactivity of $POCl_3$, therefore limiting possible side reactions. During this method, dichlorophosphate **V** is primarily generated, which can be seen as equivalent to an activated monophosphate species. Subsequent treatment with pyrophosphate yields formation of the desired 5´-*O*-triphosphate in a protection-free one pot synthesis.[93] Yoshikawa *et al.* reported in 1969 the reaction of unprotected nucleosides with $POCl_3$ in trialkyl phosphate solvents mainly leading to 5´-phosphorodichloridate **V**.[95] In situ hydrolysis results in the formation of nucleoside 5´-monophosphates.[96] The direct transformation of 5´-phosphorodichloridates to nucleoside 5´-triphosphates is obtained through treatment with an excess of tri-*n*-butylammonium pyrophosphate in DMF under anhydrous conditions followed by basic or neutral hydrolysis.[92] The occurrence of a highly reactive trimetaphosphate intermediate **VI** formed by intramolecular condensation could be proven by ^{31}P NMR analysis.

Two decades later, Kovács *et al.*[97] reported that hydrogen chloride, formed during hydrolysis of $POCl_3$, resulted in several side reactions if modified nucleosides containing unsaturated side chains were transformed. As the very reactive nature of unsaturated side chains is known in acidic conditions, they performed the 5´-*O*-triphosphorylation reaction in presence of a base. During their studies, proton sponge (1,8-bis(dimethylamino)naphthalene) proved to be the best suited base as it accelerated the reaction significantly. Due to its steric effects, proton sponge is known as very strong base with weak nucleophilic character.[97] Thus, they could show that even modified nucleosides with highly reactive unsaturated side chains could be converted to the corresponding 5´-*O*-triphosphates in the presence of proton sponge.[97] Besides the drawback by usage of the very reactive $POCl_3$ this method holds the disadvantage that different phosphorylated derivatives are obtained which are challenging to separate during purification. Nevertheless, this method holds a great potential as various modified nucleosides can be converted and different 5´-*O*-phosphorylated (mono-, di- and triphosphate) species can be obtained.

2. Aim of This Work

As several modifications in nucleic acids can be related to human diseases, new methods are strongly needed for their detection. Nowadays, several detection methods are known and rely mainly on selective chemical modifications[5a, 31] or the ability of different enzymes to distinguish between modified nucleotides and their unmodified counterparts.[28, 30] However, those technologies hold several disadvantages. In particular detection systems, which require chemical modifications prior to sequencing proved to be time-consuming, tedious and error-prone.[34-36] Especially for the application in personalised medicine convenient methods are required that allow multiplexing.

The aim of this thesis was to establish new approaches for the detection of the epigenetic markers 5mC and 5hmC without the need to perform modification reactions prior to sequencing. No DNA polymerase was known to be able to directly discriminate C against 5mC or 5hmC without the use of mismatched primers as direct sequencing of those epigenetic markers is rather challenging due to their unaltered Watson-Crick face.[116] DNA polymerases are known to discriminate against incorporation of non-canonical nucleotides very efficiently.[76] Nevertheless, different DNA polymerases are known to accept modified nucleotides with reduced efficiencies compared to their natural counterparts.[70] The aim of this thesis was to take advantage of the discrimination machinery of DNA polymerases as well as the ability of those DNA polymerases to incorporate modified nucleotides. By the introduction of various modifications at different sites of the nucleobase moiety of dGTP, dNTP derivatives which enable the DNA polymerase to discriminate between C and the epigenetic markers 5mC and 5hmC should be found. By enhancing the size and the steric hindrance of the introduced modification, a dNTP analogue should be found that is still accepted by the DNA polymerase but leads to a steric rearrangement of the primer/template complex in the active site of the enzyme in a way that the DNA polymerase will sense the presence or absence of the small methyl-group in 5mC. Following, a novel assay should be established that allows sensing of nucleic acid modifications without the need to perform modification reactions prior to sequencing. For this purpose, a tool box of variously modified nucleotides should be synthesised. Subsequently, the latter should be tested in combination with different DNA polymerases to find a combination of enzyme and nucleotide analogue, which leads to diverging incorporation efficiencies opposite C and the epigenetic markers 5mC and 5hmC in DNA polymerase catalysed reactions.

As modified nucleotides are not only known to be present in DNA, but can be found in RNA as well,[2] the generated tool box should be applied in combination with the KlenTaq variant RT-KTq2. This variant is known to exhibit reverse transcriptase activity,[125] which could be used to sense RNA modifications by incorporation of modified nucleotides.

Once identified, the most promising combinations of DNA polymerase and modified nucleotide should be further studied by the measurement of steady-state kinetics and exploited for the establishment of new detection approaches.

A second approach aimed at finding and characterising DNA polymerase variants with increased discrimination in incorporating dGMP opposite 5mC compared to C. If variants possessing this ability

would be discovered, they would provide a valuable tool for the application in DNA methylation profiling. Thereby, numerous DNA polymerase libraries, generated by site-directed mutagenesis, should be screened to find the most promising DNA polymerase variants regarding discrimination behaviour between C and 5mC. Those variants should be identified, characterised and later on be tested towards their ability to be used in SMRT (single molecule real time) sequencing approaches.

3. Results and Discussion

3.1. Initial Screening for Discrimination between C and 5mC

For initial screening experiments, several purine - based derivatives were tested to find those structural elements that lead to the most pronounced discrimination behaviour between C and 5mC. Different positions all over the molecule were chosen for introduction of modifications to evaluate, which sites are most promising for further derivatisation.

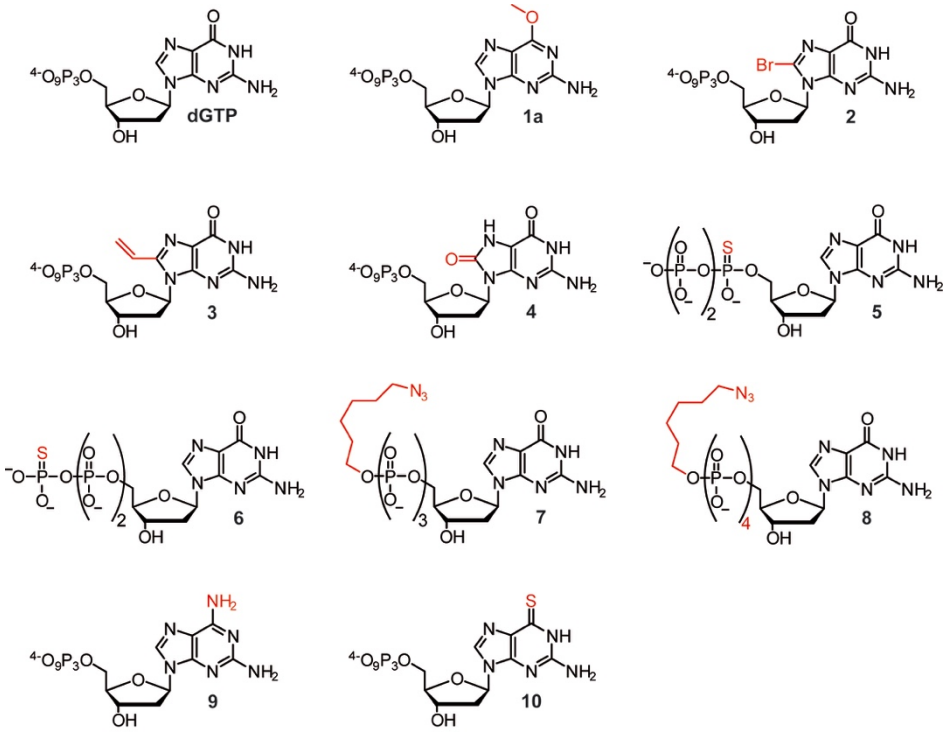


Figure 10: Modified nucleotides used for initial screening experiments. Nucleotides dGTP, **4**, **5** and **6** are commercially available, nucleotides **1a**, **2**, **3** and **7** - **10** were synthesized as described in **3.1.1 Synthesis**.

Positions 6 (**1a**, **9** + **10**) and 8 (**2** - **4**) were chosen for modification, since those sites are easily accessible for derivatisation and these modified nucleotides are known to be accepted by different DNA polymerases.[98] Additionally, nucleotides modified at the phosphate chain (**5** - **8**) were selected, as several dNTP analogues modified at this position were already successfully utilized in combination with DNA polymerases for sequencing approaches.[99] Modification at the terminal phosphate holds the advantage that this modification is cleaved during the incorporation process, leaving behind a natural DNA. It was reported, that a linker extending the triphosphate moiety by the introduction of additional phosphates proved beneficial for improving incorporation efficiencies.[100] Therefore, commercially available dGTP analogues (**4**, **5**, **6**) and several synthesized derivatives (**1a**, **2**, **3**, **7** - **10**) were tested in comparison to the unmodified dGTP (see **Figure 10**) employing different DNA polymerases.

3.1.1 Synthesis of modified nucleotides

The O^6-alkylated dGTP derivative **1a** was synthesized starting from commercially available 2´-deoxyguanosine. After protection of both 3´- and 5´-OH-groups, the acetylated[101] nucleoside **11** was following chlorinated in position 6 to obtain nucleoside **12** in a decent yield of 63 %.[102] The methoxide modification was introduced by nucleophilic displacement of the chlorine and simultaneous deprotection of the OH-groups in 86 % yield. The corresponding nucleotide **1a** was synthesized by triphosphorylation of nucleoside **13a** in 8 % yield (see **Figure 11**).[97] This synthesis strategy was used for further alkoxide derivatisation as described in **3.2.1.1. *O6-Alkyl-dGTP***.

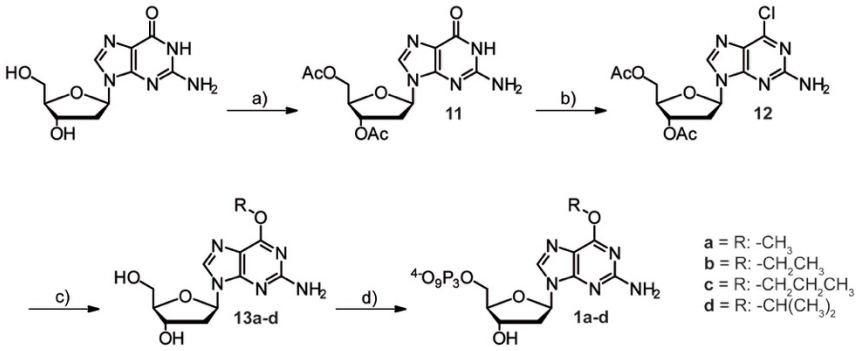


Figure 11: Synthesis of modified nucleotides **1a** - **d**. a) Ac_2O, TEA, DMA, CH_3CN, 3 h, rt, yield: quant.; b) 1: CH_3CN, tetraethylammonium chloride, *N,N*-dimethylaniline, $POCl_3$, 10 min, 0 °C; 2: 10 min, rt; 3: 15 min, reflux; 4: H_2O, 20 min, 0°C, yield: 63 %; c) 1 M NaOR in ROH, 16 h, reflux, yield: 78 - 91 %; d) 1: proton sponge, $POCl_3$, TMP, 30 min, 0°C; 2: $(Bu_3NH)_2H_2P_2O_7$, Bu_3N, DMF, 30 min, rt. 3: 0.1 M TEAB, 10 min, rt; yield: 8 - 35 %.

The dGTP derivatives **2** and **3** were modified in position 8, starting from commercially available 2´-deoxyguanosine and subsequently converted to the corresponding 5´-triphosphates after known procedures[97, 98b] in good yields. For this, 2´-deoxyguanosine was brominated in position 8 by treatment with *N*-bromosuccinimide to yield compound **14** in an excellent yield of 98 %.[98b] Compound **14** was following 5´-triphosphorylated after the procedure of Yoshikawa, modified by Kovacs[97] to yield nucleotide **2** in 20 % yield. For synthesis of nucleoside **15**, the brominated nucleoside **14** was submitted to a Stille - coupling,[98b] followed by 5´-triphosphorylation to obtain nucleotide **3** in a yield of 13 % (see **Figure 12**).

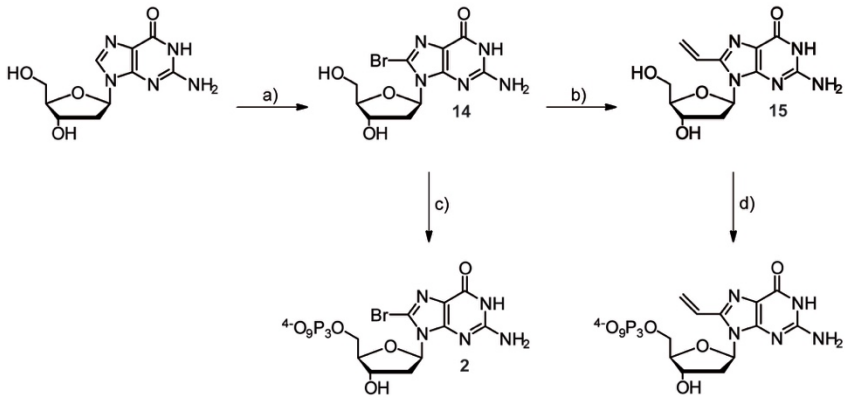


Figure 12: Synthesis of modified nucleotides **2** and **3**. a) NBS, H_2O/CH_3CN, 1 h, rt; yield: 98 %; b) $P(PPh_3)_4$, tributyl(vinyl)stannane, NMP, 2 h, 110 °C; yield: 83 %; c) + d) 1: proton sponge, $POCl_3$, TMP, 30 min, 0 °C; 2: $(Bu_3NH)_2H_2P_2O_7$, *n*Bu_3N, DMF, 30 min, rt. 3: 0.1 M TEAB, 10 min, rt; yield: 16 - 20 %.

Compounds **7** and **8** were synthesized starting from dGTP following known procedures.[103] The commercially available sodium salt of dGTP was first subjected to ion-exchange using a CHELEX column, to change the counter - ion from sodium to tetrabutylammonium to enhance the solubility of the nucleotide in organic solvents, in particular in DMF (see **6.1.68. General procedure E**). For alkylation, the tetrabutylammonium salt of dGTP was incubated with 1-azido-6-bromohexane under anhydrous conditions overnight and purified by ion-exchange and RP-HPLC to obtain the γ-phosphate modified nucleotide **7** in 10 % yield. Since position 1 offered an additional possible site for alkylation, 2D-NMR experiments were recorded to prove the modification to be located at the γ-phosphate. The δ-phosphate-modified tetraphosphate was synthesized by coupling of dGTP with 1-azido-6-hexylphosphate. The activated tetrabutylammonium salt of dGTP was reacted with 1-azido-6-hexylphosphate under anhydrous conditions for several days. After subsequent purification by FPLC and RP-HPLC, the modified tetraphosphate **8** could be obtained in 18 % yield (see **Figure 13**).[103]

Nucleotides **9** and **10** were synthesised starting from the commercially available nucleoside derivatives by 5´-triphosphorylation in 11 - 23 % yield.

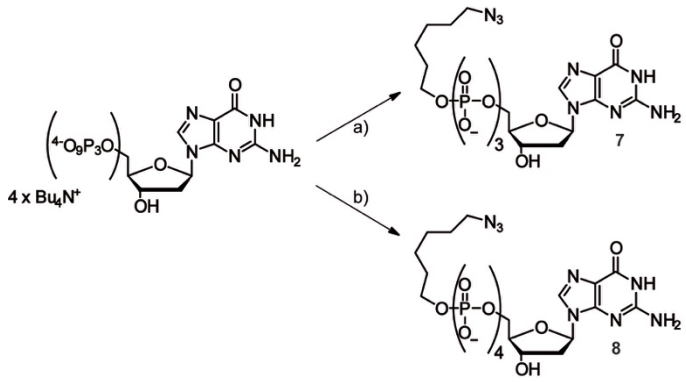


Figure 13: Synthesis of modified nucleotides **7** and **8**. a) 1-Azido-6-bromohexane, DMF, 16 h, rt; yield: 10 %; b) 1: EDC, DMF, 3 h, rt; 2: MeOH, 2 h, rt; 3: 1-azido-6-hexylphosphate, 3 days, 40 °C, yield: 18 %.

3.1.2. Screening of modified nucleotides to discriminate 5mC

All modified nucleotides depicted in **Figure 10** were tested towards their potential to be used in 5mC detection due to diverging incorporation efficiencies opposite C or the epigenetic marker 5mC by different DNA polymerases. Two DNA polymerases belonging to two different sequence families were chosen for previous screening experiments: the DNA polymerases KlenTaq, belonging to sequence family A and KOD exo$^-$, member of sequence family B.[104] Archeal DNA polymerases (as KOD exo$^-$) are known to have the ability to detect U in DNA by a "read ahead" mechanism.[105] Detection occurs at a site distant from the nucleotide incorporation site. As U differs from T only by an additional methylation in position 5, T can be seen as 5mU. Thus, attempts to identify combinations of modified nucleotides and DNA polymerase with increased discrimination between C and 5mC seem promising.

The unmodified dGTP was tested in comparison to the modified nucleotides **1** - **10** in radioactive single-nucleotide incorporation primer extension experiments employing both mentioned DNA polymerases. Analysis was afterwards performed by denaturing PAGE analysis and visualisation by autoradiography. A non-elongated primer band can be seen as reference on the left side of every gel picture and incorporation of the respective nucleotide by a DNA polymerase can be monitored by the occurrence of a second band, which can be detected above the band resulting from remaining primer due to reduced migration mobility. Incorporation efficiencies were determined by terms of the integrated gel band intensities, as % incorporation = $100*(I_{extension})/(I_{primer}+I_{extension})$. Discrimination ratios were subsequently determined by calculating the quotient of % incorporation opposite C and % incorporation opposite 5mC.

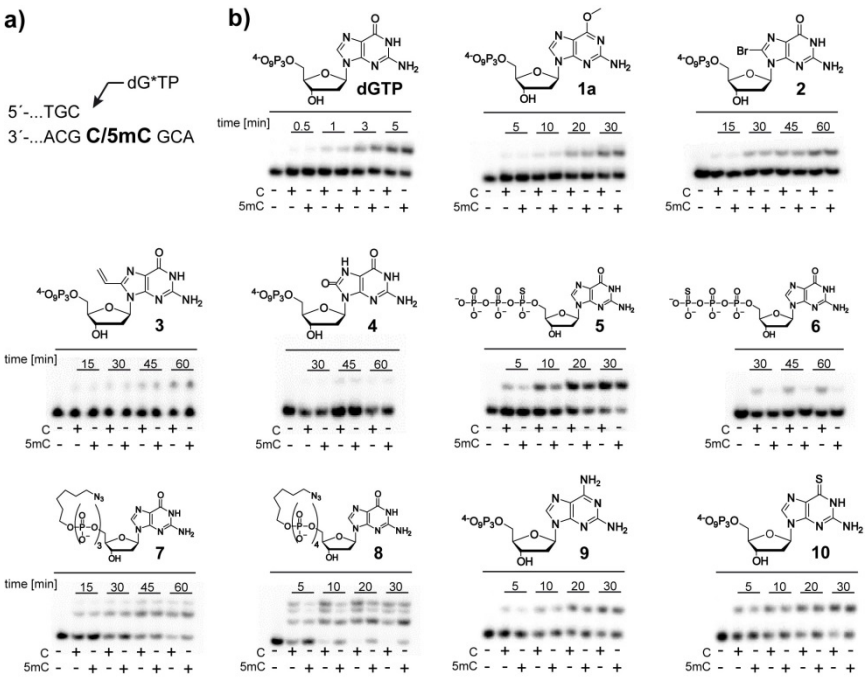


Figure 14: Screening of dGTP and modified nucleotides using KlenTaq. **a)** Partial primer/template sequence used; **b)** PAGE analysis of single-nucleotide incorporation primer extension experiments of dGMP and nucleotides **1a** - **10** opposite a template containing C in comparison to a template containing 5mC employing the DNA polymerase KlenTaq. 50 µM dGTP/dG*TP and 5 nM KlenTaq were used; reactions were stopped after indicated time points.

The incorporation of dGMP or the modified nucleotides **1** - **10** by the DNA polymerase KlenTaq does not show promising results (see **Figure 14**). All modified nucleotides are accepted by this A-family DNA polymerase, but incorporation efficiencies were notably decreased in comparison to the unmodified dGMP (see **Figure 14**). Those nucleotides, modified in position 8 (**2** - **4**), are processed with very low efficiencies, compared to the unmodified dGTP. Especially application of nucleotide **4** shows barely any incorporation under the chosen conditions. In addition, no promising differences in incorporation efficiencies between processing of those nucleotides opposite C in comparison to 5mC can be observed. Anyhow, nucleotide **6** shows favoured incorporation opposite C; but the overall turnover of this nucleotide by the KlenTaq DNA polymerase is very low. Another nucleotide showing some discriminating effects, nucleotide **8**, was not further considered as well, since incorporation of this δ-phosphate modified tetraphosphate leads to several different incorporation bands, making interpretation difficult (see **Figure 14**). We suggest the reason for those multiple incorporation bands to be an additional hydrolysis mechanism of nucleotide **8**. If cleavage occurs between ß- and γ-phosphate, instead of cleavage between α- and ß-phosphate, alkylated pyrophosphate would be released during incorporation and the driving force of the incorporation would therefore be maintained. Hence, I suggest that a second incorporation band derives from incorporation of dGDP in addition to incorporation of dGMP.

In contrast, testing those nucleotides in combination with the B-family DNA polymerase KOD exo$^-$ shows already some discrimination for incorporation of dGMP opposite C and 5mC, as dGTP is processed with slightly higher efficiency opposite C, than opposite 5mC (see **Figure 15**). This difference in incorporation efficiencies can even be further increased by the application of the modified nucleotides **1** - **10**. Again, processing of the modified nucleotides shows decreased efficiencies in comparison to dGTP, but still most modified nucleotides are accepted with good incorporation efficiencies. Again, those nucleotides modified in position 8 (**2** - **4**) are poorly accepted. For nucleotide **4**, as well as for the γ-thiophosphate modified dGTP derivative **6**, no notable incorporation opposite C or 5mC can be observed after 60 min under the chosen conditions. Since those nucleotides are very poorly incorporated by both DNA polymerases, they were not considered for further experiments. As already observed in the experiments employing KlenTaq, processing of nucleotide **8** led to an incorporation pattern on the PAGE gels, which is difficult to interpret due to three different extension bands. Despite the decreased incorporation efficiencies for the 8 modified nucleotides **2** and **3**, increased discrimination between C and 5mC can be observed for both nucleotides, while incorporation opposite C is favoured over incorporation opposite 5mC. Nucleotide **5** shows decent incorporation efficiencies with slightly favoured incorporation opposite C in comparison to 5mC.

The most promising results can be achieved by processing nucleotides **1a**, **7**, **9** and **10**. All four nucleotides are processed opposite C with remarkably higher efficiencies than opposite 5mC and the incorporation efficiencies opposite C are just slightly decreased compared to the unmodified dGMP. Therefore, position 6 of dGTP was chosen for further derivatisation. This position seems most promising for the desired application, since nucleotides modified in this position are well accepted by both DNA polymerases and processing of those nucleotides employing the DNA polymerase KOD exo$^-$ leads to the most pronounced differences in incorporation efficiencies (see **Figure 15**).

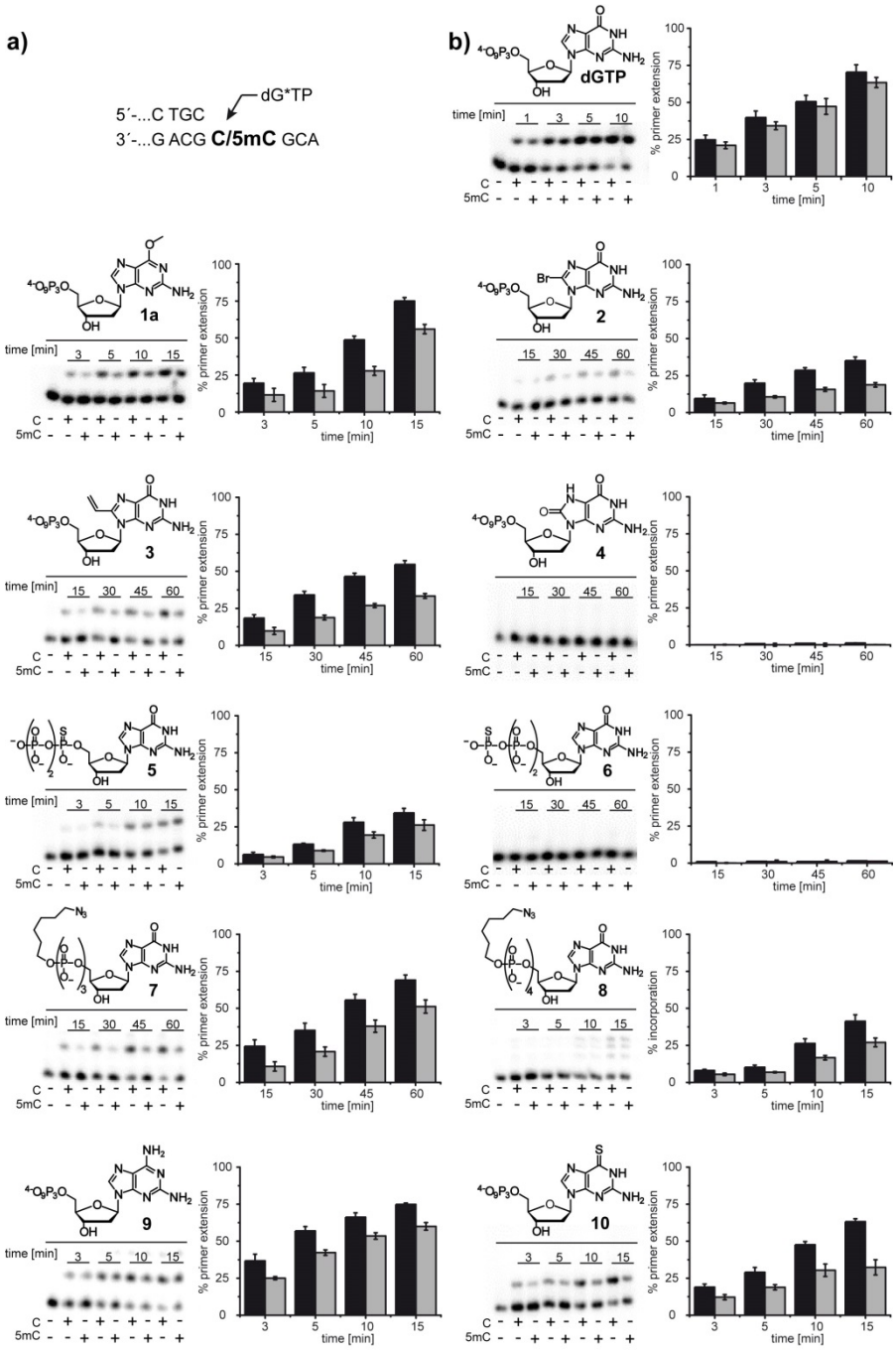


Figure 15: Screening of dGTP and modified nucleotides using KOD exo$^-$. **a)** Partial primer / template sequence used; **b)** PAGE analysis and quantitative evaluation of single-nucleotide incorporation primer extension experiments of dGMP and nucleotides **1a** - **10** opposite a template containing C (black) in comparison to a template containing 5mC (grey) employing the DNA polymerase KOD exo$^-$. 50 µM dGTP or dG*TP and 5 nM KOD exo$^-$ were used; reactions were stopped after indicated time points. Experiments were done at least in triplicates.

3.2. 6 - Modified dGTP Derivatives for the Detection of 5mC

After testing different purine-based 2´-deoxy-nucleotides (see **Figure 10**) for their ability to sense 5mC in DNA polymerase-catalysed reactions, I found that modification of dGTP in position 6 was most

promising. dGTP analogues, modified at this position are processed with remarkably different efficiencies opposite C than opposite 5mC by the DNA polymerase KOD exo$^-$ (see **Figure 15**). In order to investigate, if this discrimination is more general and extendable to other modifications, I decided to synthesise different dGTP derivatives that are modified at position 6 and explore their potential to enhance the observed discrimination. In addition, modification at position 6 can clarify, if interruption of the Watson-Crick face of dGMP interferes with DNA polymerase-catalysed incorporation opposite C or 5mC.

3.2.1. Synthesis of 6-modified dGTP Derivatives

As dGTP derivatives modified in position 6 were identified to be most promising for the application in 5mC detection, I decided to synthesise a tool box of nucleotides, variously modified in position 6. Therefore, O^6-alkyl-dGTP derivatives **1b** - **1d** were synthesized bearing alkyl modifications with increasing size and steric hindrance (see **Figure 11**).

Additionally, 6-amino-dGTP derivatives were chosen as targets with various modifications. The synthesis of 6-amino modified nucleotides allows a broader spectrum of introduced modifications, as the amino group in position 6 can not only be modified with one alkyl chain, but with two. Therefore, the effect of differently sized modifications can be studied in addition to the effect of different H-bonding properties of the amino group, as those will change due to the introduction of a second modification. Bearing only one modification (**9a** - **g**), the 6-amino group will act as H-bonding donor. In contrast, after introduction of a second alkyl chain (**9h** - **o**), the amino group can serve as H-bond acceptor, bearing therefore an H-bonding pattern, which is closer to the one of the unmodified dGTP (see **Figure 16**).

Initial screening experiments pointed as well to the potential of nucleotide **10** for 5mC detection (see **Figure 15**). The exchange of the oxygen atom in position 6 by a sulphur atom shows already an increase in discrimination behaviour. Since the alkylation of oxygen in position 6 in nucleotide **1a** led as well to a promising compound for incorporation studies, we thought about combining both modifications, to further increase the observed differences in incorporation efficiencies. Therefore, a third class of in position 6 modified dGTP derivatives was envisioned as target: nucleotides **10a** + **b** (see **Figure 18**).

3.2.1.1. *O*^{6}-Alkyl-dGTP

O^6-alkylated dGTP derivatives **1b** - **1d** were synthesised analogously as described for nucleotide **1a** (see **Figure 11**). The different alkoxide groups were introduced by reaction of the chlorinated nucleoside **12** with the respective alkoxide solutions. The obtained nucleosides **13b** - **13d** were then converted to the corresponding nucleotides **1b** - **1d** as described before (see **3.1.1 Synthesis**).

3.2.1.2. 6-Amino-dGTP

At first, the 6-amino-dGTP derivatives **9a** - **9o** were synthesised similarly to the O^6-alkylated dGTP derivatives **1a** - **1d**. The nucleoside precursor **12** was reacted with the aqueous solutions of the different amines, which should be introduced in position 6. Due to simultaneous deprotection of the

acetyl groups under basic conditions, this procedure resulted in the corresponding nucleosides **16a**, **h**, **i** and **l** which could be converted to the particular nucleotides **9a**, **h**, **i** and **l** (see **Figure 16a**).

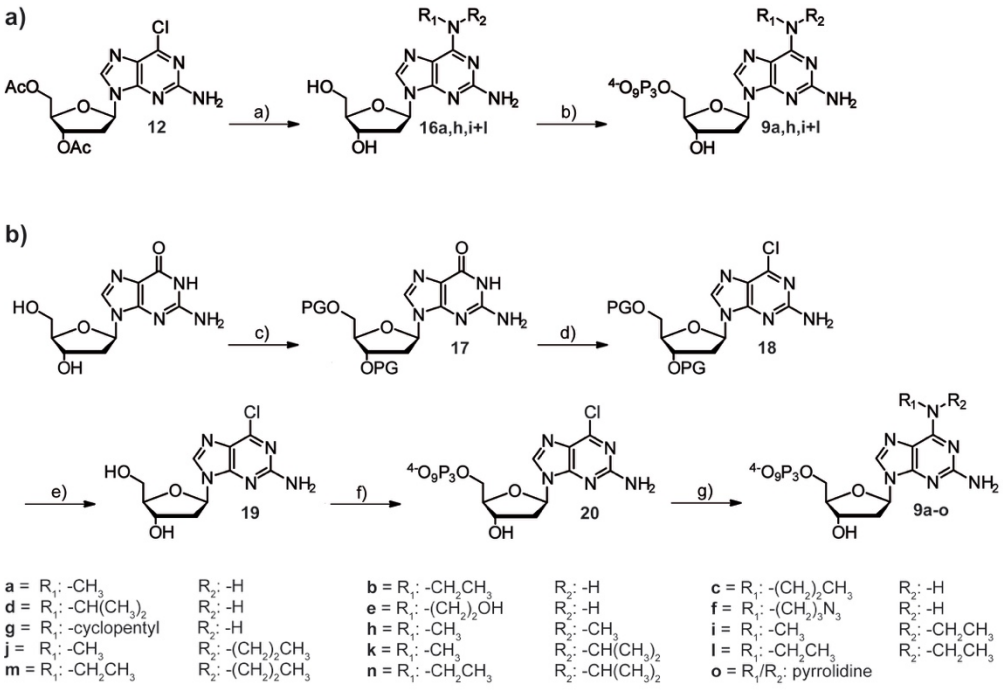


Figure 16: **a)** Synthesis of modified nucleotides **9a** - **9o**. a) Aqueous NR_1R_2, 16 h, rt, yield: 71 - 97 %, b) 1: proton sponge, $POCl_3$, TMP, 30 min, 0°C; 2: $(Bu_3NH)_2H_2P_2O_7$, *n*Bu_3N, DMF, 30 min, rt. 3: 0.1 M TEAB, 10 min, rt, yield: 32 - 42 %; **b)** Shortened synthesis route; c) TBDMS-Cl, imidazole, DMF, rt, 20 h, yield: 79 %; d) 1: CH_3CN, tetraethylammonium chloride, *N,N*-dimethylaniline, $POCl_3$, 10 min, 0 °C; 2: 10 min, rt; 3: 15 min, reflux; 4: H_2O, 20 min, 0°C, yield: 37 %; e) triethylamine trihydrofluoride, THF, 16 h, rt, yield: 93 %; f) 1: proton sponge, $POCl_3$, TMP, 30 min, 0 °C; 2: $(Bu_3NH)_2H_2P_2O_7$, Bu_3N, DMF, 30 min, rt. 3: 0.1 M TEAB, 10 min, rt, yield: 13 %; g) aqueous NR_1R_2, 16 h, rt, yield: 27 - 97 %.

Triphosphorylation reactions are tedious and time-consuming. Therefore, I thought about a different strategy to synthesise those compounds. To circumvent the need to perform a 5´-triphophorylation reaction for every modification, I decided to introduce the modifications on the triphosphate level. For that purpose we synthesised a nucleotide precursor **20** (see **Figure 16b**) bearing chlorine in position 6. Therefore, nucleoside **19** which could be converted to nucleotide **20** by 5´-triphosphorylation was synthesised. Deprotection of nucleoside **12** proved to be challenging since deprotection under nucleophilic conditions led to the partial loss of the chlorine in position 6 by nucleophilic replacement. For that reason, I decided to change the protecting strategy for the 3´- and 5´-OH-groups used during chlorination. I was looking for an OH-protecting group that is stable under the acidic conditions used during chlorination and can be cleaved under non-nucleophilic conditions. Therefore, I thought to use TBDMS as protecting group, since TBDMS is supposed to be stable under acidic conditions and is easily cleaved by using the fluoride anion, which does not affect the chlorine in position 6. Thus, I started from commercially available 2´-deoxyguanosine, which was protected using TBDMS-Cl after known procedures in 79 % yield.[106] Subsequent chlorination was performed using the same conditions as used before for the acetylated species with a yield of 37 %, followed by deprotection in a very good yield of 93 %. The obtained nucleoside **19** was subjected to 5´-triphosphorylation yielding

the desired nucleotide precursor **20** with 13 %. This nucleotide precursor **20** could be treated with the aqueous solutions of the respective amines to obtain all 6-amino modified dGTP derivatives **9a** - **9o** after purification through ion-exchange chromatography and RP-HPLC in yields ranging from 18 - 97 % depending on nucleophilicity and steric hindrance of the used amines.

Even if the nucleotides **9a** - **o** needed to be purified by ion-exchange chromatography as well as RP-HPLC to get rid of all remaining amines, the synthesis route could be shortened and simplified since the need to perform several triphosphorylation reactions was avoided. With synthesis of nucleotide **20**, we could show the potential of this nucleotide to act as precursor for the introduction of several modifications in position 6 by nucleophilic replacement. Nucleotide **20** is even commercially available, enhancing its scope to serve as nucleotide precursor for several modified dGTP analogues.

3.2.1.3. 6-Thioalkyl-dGTP

As I could present nucleotide **20** as precursor for the introduction of several modifications in position 6, I incubated nucleotide **20** with the particular thiols under basic conditions at room temperature as well as under reflux for several hours. Unfortunately, I could not observe any generated product. Neither by TLC nor by mass spectrometry was it possible to detect the desired product. TLC analysis suggested explicitly that the nucleotide precursor was degenerated by this treatment, resulting in an undefined mixture. I suggested that the limited stability of the triphosphate chain was the problem. Therefore, I changed the procedure and tried to introduce the modification on the nucleoside level. Unfortunately, switching to the nucleoside precursor **12** did not change the outcome of the reaction. Again, the desired product could not be detected by any method and the reaction resulted in a crude reaction mixture with no defined product.

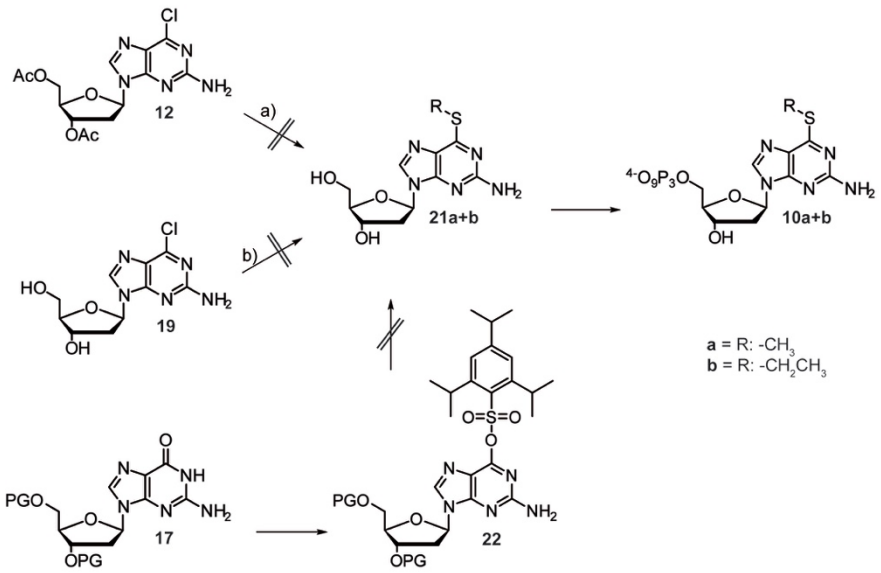


Figure 17: Failed synthesis of nucleotides **10a** and **b**. a) NaSMe or EtSH + KO*t*Bu, DMF, rt or reflux, 16 h, yield: no product obtained; b) NaSMe or EtSH + KO*t*Bu, DMF, rt or reflux, 16 h, yield: no product obtained.

To make sure that no thioester, formed during deprotection of the acetyl-groups by the thiol reagent, could lead to undesired side-reactions resulting in this undefined reaction mixture, I changed to the unprotected nucleoside **19** as precursor. Since no deprotection could occur during incubation with the thiol, the amount of possible side reactions was limited. Unfortunately no defined product could be detected and no formation of nucleosides **21a**+**b** could be observed (see **Figure 17**).

Next, I thought about using a different leaving group than chlorine, facilitating the nucleophilic substitution. Due to the stability and solubility of 2´-deoxyguanosine, only few leaving groups are known for position 6.[107] One of them is tosylate-related. However, literature reports[107] that soft nucleophiles like thiols would rather attack the sulphur, therefore leading back to the unmodified nucleoside **17**.

Finally, I decided to introduce the thio-modifications on the nucleobase level followed by glycosylation and 5´-triphosphorylation.[108] Reaction of 6-chloro-2-amino-purine with the respective thiols led to the modified nucleobases **23a** and **23b** in around 55 % yields. Those nucleobases were glycosylated to yield the nucleosides **24a** and **24b**, which were following subjected to deprotection and 5´-triphosphorylation.

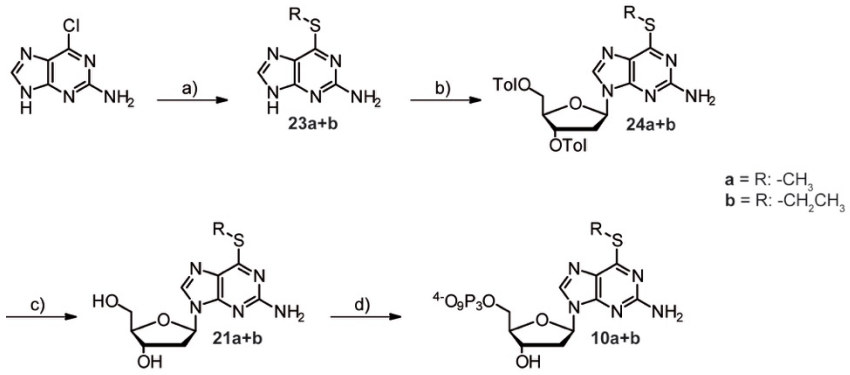


Figure 18: Synthesis of nucleotides **10a** + **b**. a) **23a** Sodium methoxide, DMF, rt, 16 h, yield: 55 %; **23b** potassium *tert*-butoxide, ethanethiol, reflux, 16 h yield: 49 %; b) 1: NaH (60 % in mineral oil), CH_3CN, rt, 30 min; 2: 1-chloro-2-deoxy-3,5-di-*O*-toluoyl-α-D-ribofuranose, rt, 20 h; yield: 57 - 64 % c) 7 N ammonia in methanol, 4 °C, 16 h, yield: 61 - 99 %; d) 1: proton sponge, $POCl_3$, TMP, 30 min, 0 °C; 2: $(Bu_3NH)_2H_2P_2O_7$, Bu_3N, DMF, 30 min, rt. 3: 0.1 M TEAB, 10 min, rt, yield: 28 %.

Since it has already been described in 1960 by Hoffer[109] that N^9-substituted purines are formed as main products during the nucleobase-anion glycosylation of the respective purines with 2-deoxy-3,5-di-*O*-(4-toluoyl-α-D-*erythro*-pentofuranosyl)chloride, glycosylation was performed according to Seela *et al.*[110] using their findings. Even if the occurrence of the N^7-isomer as side product is known, this procedure was used as the undesired isomer was expected to occur as minor by-product. Seela and co-workers could show before, that the ratio of the N^7- and N^9-isomers will be shifted towards the undesired N^7-isomer by usage of 6-methoxy-2-amino-purine instead of 6-chloro-2-amino-purine.[111] They reported a comparative study of this glycosylation reaction using various 6-alkoxy-2-amino-purines. They carried out the reaction at room temperature using acetonitrile as solvent, powdered

potassium hydroxide as base and TDA-1 as catalyst.[112] The two isomeric glycosylation products were separated by flash chromatography. They could show, that even if the ratio of the undesired N^7-isomer was increased by usage of 6-alkoxy-2-aminopurines, the N^9-isomer was still favoured if very short alkyl-chains were used.[110] Therefore, I decided to use this glycosylation method for synthesis of the nucleoside derivatives **24a** and **24b** with short thio-alkyl modifications. Unfortunately, glycosylation under those conditions did not show the favoured occurrence of the N^9-isomer. The reaction yielded a 1:1 mixture from both isomers, which could not be separated sufficiently by chromatographic methods. Thus, I changed glycosylation conditions to the usage of sodium hydride as base in acetonitrile, without any catalyst.[113] In this procedure the N^9-isomer was clearly favoured and finally led to the modified nucleosides **24a** and **24b** in 57 - 64 % yield (see **Figure 18**).

Deprotection of the toluoyl-groups was first carried out with diluted sodium methoxide solution, as reported by Seela *et al.*.[110] This treatment did not only result in liberation of both OH-groups but in replacement of the thio-alkyl modifications by methoxide as well. Therefore, I changed the deprotection procedure and used 7 N ammonia in methanol at 4 °C for 16 h.[108] These milder conditions led to cleavage of the OH-protecting groups but did not affect the introduced modifications in position 6. Using those conditions, I could obtain the nucleosides **21a** and **21b** in 61 - 99 % yield. Both nucleosides were subsequently subjected to 5´-triphosphorylation reactions, which finally resulted in the desired thio-modified nucleotides **10a** and **10b** in 28 % yield.

3.2.2 Screening of Modified Nucleotides for 5mC Detection

Now, having a toolbox of variously modified nucleotides in hands, the nucleotides were tested towards their action on 5mC. Since the B-family DNA polymerase KOD exo$^-$ showed the most promising trends towards diverging incorporation efficiencies opposite C in comparison to the incorporation opposite 5mC, I decided to employ a second DNA polymerase, belonging to the sequence family B: the DNA polymerase 9°North exo$^-$.

3.2.2.1. Testing KlenTaq DNA Polymerase for Incorporation of Modified Nucleotides

As already indicated by previous experiments (see **Figure 14**), no significant discrimination between C and 5mC can be observed by employing the DNA polymerase KlenTaq in combination with any of the modified nucleotides (see **Figure 19**).

The results shown in **Figure 19** elucidate that incorporation efficiencies of the modified dGMP derivatives decrease considerably with increasing size or steric hindrance of the introduced modification. Even if incorporation efficiencies of all modified nucleotides are decreased in comparison to the unmodified dGTP, processing of nucleotide **1a** is notably more efficient than processing of nucleotide **1d** possessing the steric most demanding alkoxy modification *iso*propyl. The same can be observed for the amino modified analogues **9a** - **9o**. Nucleotide **9**, bearing no alkyl modification at the 6-amino group, is processed with good efficiencies, while incorporation efficiencies decrease notably by introduction of further modifications. Therefore, the lowest incorporation rates can be observed employing the nucleotides **9n** and **9o**, as both nucleotides possess steric demanding modifications.

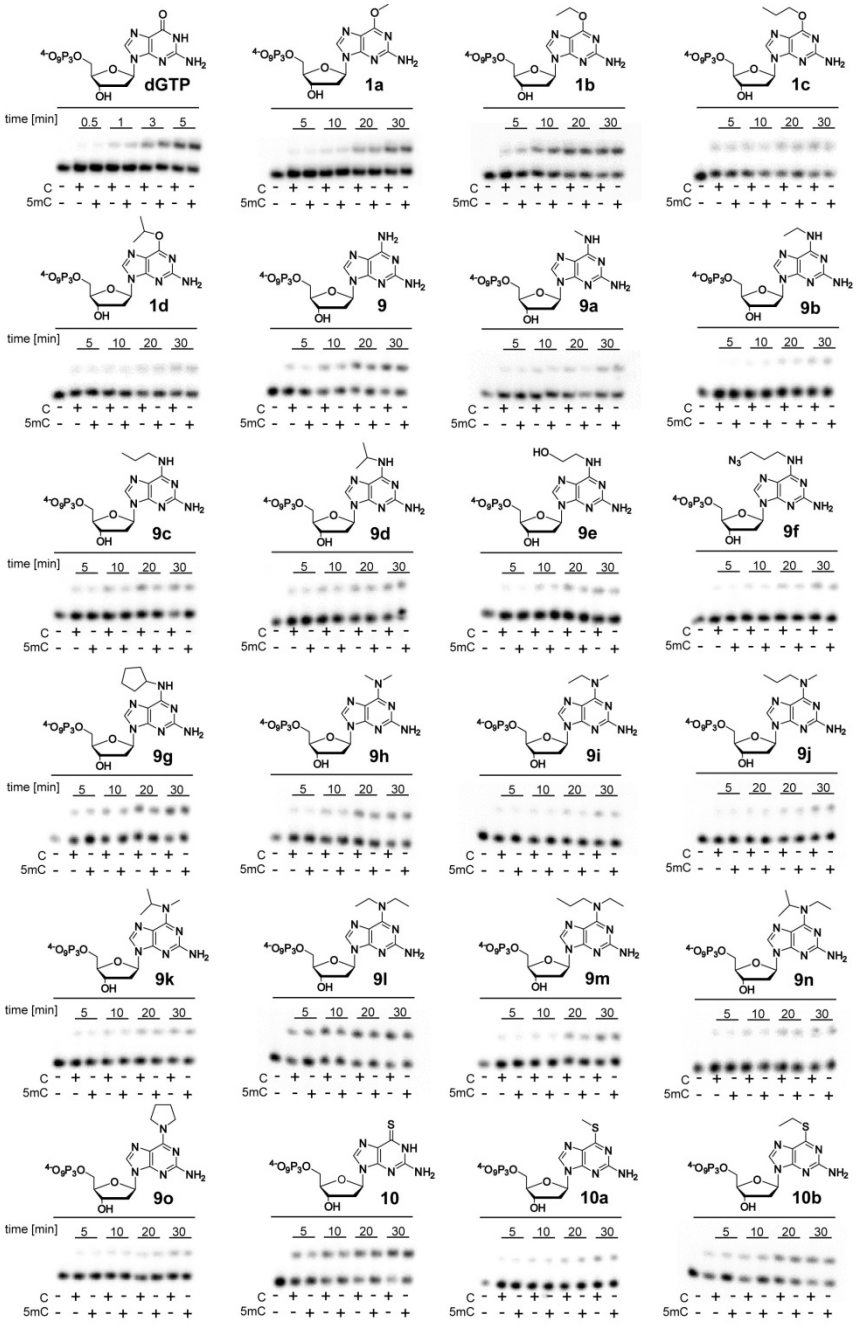


Figure 19: PAGE analysis of single-nucleotide incorporation primer extension experiments of dGMP and nucleotides **1a** - **d**, **9**, **9a** - **9o**, **10**, **10a** + **b** opposite a template containing C in comparison to a template containing 5mC employing the DNA polymerase KlenTaq. 50 µM dGTP or dG*TP and 5 nM KlenTaq were used; reactions were stopped after indicated time points.

The same holds true for the 6-thio modified nucleotides. While nucleotide **10** is processed with decent yields, incorporation efficiencies decrease due to the introduction of additional alkyl modifications in nucleotides **10a** and **10b**.

3.2.2.2. Testing KOD exo$^-$ DNA Polymerase for Incorporation of Modified Nucleotides

Furthermore, the in position 6 modified dGTP derivatives were tested towards their potentials to be used for 5mC detection by employing the B-family DNA polymerase KOD exo$^-$ (see **Figure 20**). To ensure quantitative evaluation of those single-nucleotide incorporation primer extension experiments, all experiments were done at least in triplicates. **Figure 20** summarises all observed incorporation efficiencies opposite C or 5mC as well as the discrimination ratio, determined by calculating the quotient of % incorporation opposite C and % incorporation opposite 5mC. Colour coding highlights those nucleotides with the most efficient incorporation by KOD exo$^-$ (blue) as well as those leading to the highest discrimination (red).

In contrast, utilising the sequence-family B DNA polymerase KOD exo$^-$ I found that this DNA polymerase is capable to incorporate all modified nucleotides with only slightly decreased incorporation efficiencies compared to the unmodified dGMP. As already observed before, the incorporation efficiencies decrease with increased steric hindrance of the introduced modification. Therefore, nucleotide **1a** is incorporated with a higher efficiency than nucleotide **1d**. The same tendency can be observed for the amino modified nucleotides **9a** - **9o** as well as for the thio modified nucleotides **10**, **10a** and **10b**.

Interestingly, I observed the tendency that all modified nucleotides are more efficiently processed opposite C than opposite 5mC by KOD exo$^-$ (see **Figure 20**). As mentioned before, already by processing of the unmodified dGTP, discrimination can be observed with favoured incorporation opposite C. This discrimination improves by the usage of modified nucleotides. The best discrimination ratios were measured during incorporation of nucleotides **1b**, **9b**, **9g** and **9o**. By employing KOD exo$^-$ with the unmodified dGTP a discrimination ratio of 1.36 was found. This discrimination increases remarkably, if the oxygen in position 6 is alkylated. Nucleotide **1a** already shows improved discrimination compared to dGTP, but the introduction of the bigger ethyl-group in nucleotide **1b** leads to an even higher discrimination. Hence, the discrimination ratio observed during processing of dGTP (1.36) can be improved to a factor of 3.16 for nucleotide **1b**. The employment of larger modifications does not further improve the discrimination ratio. So, discrimination decreases for nucleotides **1c** and **1d**. Similar effects can be observed for the amino modified nucleotides **9a** - **9d**. Again, the discrimination increases with introduction of a methylamino- (**9a**) or ethylamino-group (**9b**) in comparison to 6-amino-dGTP **9**. This effect vanishes after introduction of more bulky modifications (**9c** - **9f**). Interestingly, the difference in incorporation efficiencies opposite C and 5mC increases further by employment of tertiary amines (**9h** - **9o**), as the highest discrimination can be observed by the most rigid, sterically hindered modification pyrrolidine **9o**. In this case, a threefold higher incorporation can be observed opposite C, compared to 5mC. Improved discrimination for 6-thio dGTP **10** was observed before. Additional modification by alkylation (**10a**+**b**) improves this discrimination further, but still discrimination detected by processing of nucleotide **1b** is best (see **Figure 20**).

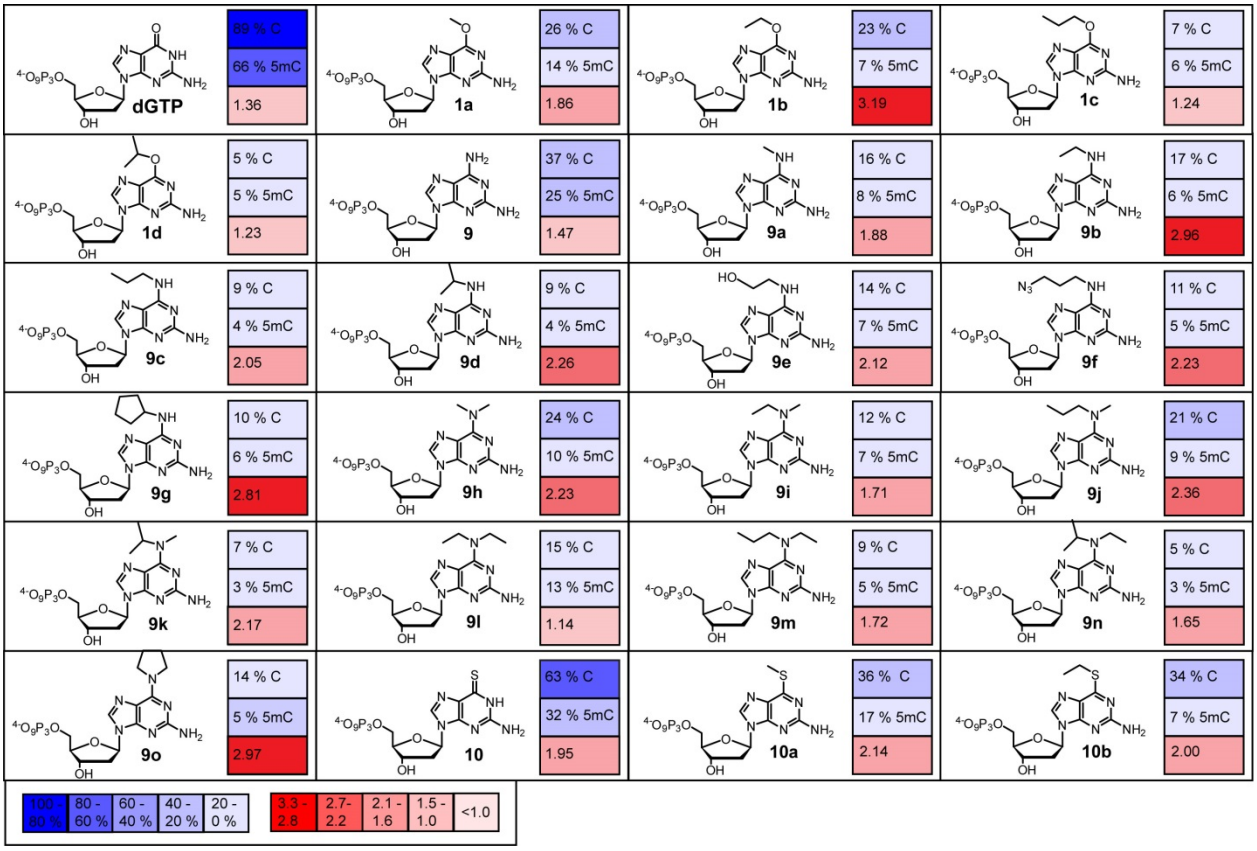


Figure 20: Structures of modified dG*TP analogues, including % incorporation opposite C or 5mC employing DNA polymerase KOD exo$^-$ in single-nucleotide incorporation primer extension reactions. 50 µM dGTP or dG*TP and 5 nM KOD exo$^-$ were used; reactions were stopped after 10 min. Experiments were done at least in triplicates. Arithmetic mean is given; errors are given in the appendix.

3.3.2.2.1. Kinetics for Incorporation of *O*^{6}-Alkyl-dGTP Derivatives by KOD exo$^-$

To further investigate these findings, I determined steady-state kinetics[114] for incorporation of the nucleotides dGMP and **1a** - **1d** opposite C and 5mC (see **Table 1**). Those nucleotides were chosen for further studies, as they proved most promising in previous experiments in combination with KOD exo$^-$. Comparison of the catalytic efficiencies (k_{cat}/K_M) observed for processing of dGTP and the 6-alkyl modified nucleotides opposite C and 5mC in the template strand confirms all tendencies observed in the above described primer extension experiments. The incorporation efficiencies of all modified nucleotides decrease with increased steric hindrance of the modifications, as the catalytic efficiencies decrease sequentially from 1.5 ± 0.1 $s^{-1}M^{-1}$ observed during processing of dGTP to 0.031±0.005 $s^{-1}M^{-1}$ for **1d** (see **Table 1** + **Figure 21**).

However, the ratio of the catalytic efficiencies observed during processing of the respective nucleotide opposite C in comparison to the incorporation opposite the epigenetic marker 5mC varies. Unmodified dGTP is processed opposite C with 1.4 fold higher catalytic efficiency compared to the incorporation opposite 5mC. This discrimination ratio for nucleotide incorporation opposite C in comparison to the incorporation opposite 5mC increases for nucleotide **1a** to a factor of 2.6 and for nucleotide **1b** even further to 4.2 (see **Table 1**).

Table 1: Steady-state kinetic analysis of single-nucleotide incorporation of dGMP and modified nucleotides **1a** - **1d** opposite C or 5mC employing DNA polymerase KOD exo$^-$. The ratio was calculated by the quotient of k_{cat}/K_M opposite C and k_{cat}/K_M opposite 5mC.

nucleotide	template	k_{cat} [s^{-1}][a]	K_M [µM][a]	k_{cat}/K_M [$s^{-1}µM^{-1}$][a]	ratio
dGTP	C	5.9 ± 0.1	4.0 ± 0.3	1.5 ± 0.1	1.36
	5mC	3.5 ± 0.1	3.3 ± 0.3	1.1 ± 0.1	
1a	C	3.338 ± 0.001	20.7 ± 2.6	0.16 ± 0.02	2.58
	5mC	1.12 ± 0.6	18.2 ± 1.6	0.062 ± 0.010	
1b	C	2.27 ± 0.06	15.8 ± 2.0	0.14 ± 0.02	4.24
	5mC	0.91 ± 0.06	27.4 ± 4.0	0.033 ± 0.007	
1c	C	2.55 ± 0.04	24.1 ± 3.1	0.105 ± 0.015	1.08
	5mC	1.47 ± 0.02	15.2 ± 1.4	0.097 ± 0.011	
1d	C	2.44 ± 0.14	78.8 ± 8.8	0.031 ± 0.005	1.48
	5mC	1.19 ± 0.09	56.0 ± 7.7	0.021 ± 0.004	

[a] Data points derive from triplicates. ± describes SD.
Sequence primer: 5´-d(CGAAATGATCCCATCCAGCTGC)-3´

When increasing the steric bulk of the nucleotide modification, a decline of the ratio to 1.1 for nucleotide **1c** and 1.5 for **1d** can be observed. Having a closer look at the kinetic data depicted in **Table 1**, it can be seen, that discrimination between C and 5mC is mainly based on differences in k_{cat}. For incorporation of all nucleotides, k_{cat} is higher for processing opposite C than 5mC indicating more efficient incorporation opposite the template containing C, although K_M was as well higher for the incorporation opposite C. Solely for processing of nucleotide **1b** a lower K_M can be observed for the incorporation opposite C. Hence, k_{cat}/K_M shows the best discrimination in case of **1b**, proving **1b** to be the most promising nucleotide in combination with KOD exo$^-$.

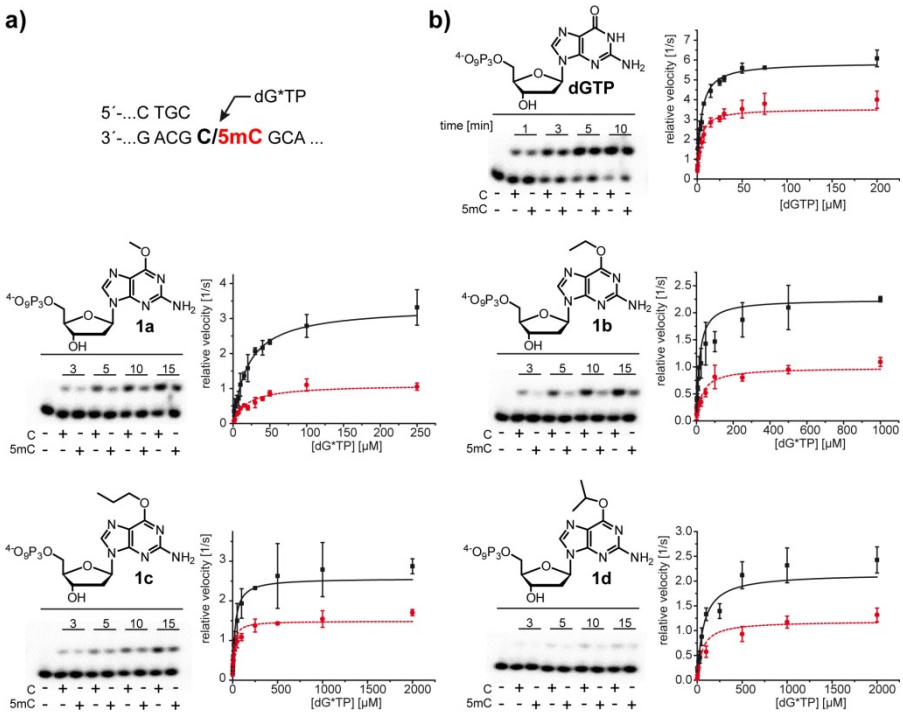


Figure 21: Kinetic studies of incorporation of dGMP and **1a-d** by KOD exo$^-$. **a)** Partial primer / template sequence used; **b)** left: chemical structure and PAGE analysis of single-nucleotide incorporation primer extension experiments of dGMP and nucleotides **1a-d** employing KOD exo$^-$. 50 µM dGTP or dG*TP and 5 nM KOD exo$^-$ were used; reactions were stopped after indicated time points; right: steady-state kinetics of single-nucleotide incorporation of dGMP or dG*MP opposite C (black solid line) or 5mC (red dashed line). Experiments were done at least in triplicates.

3.3. Capillary Gel Electrophoresis

As working with radioactivity has obvious disadvantages, I switched to fluorescently labelled primers for single-nucleotide incorporation primer extension experiments. Despite its utility, analysis of those experiments by denaturing PAGE has several disadvantages (see **1.5. Capillary Electrophoresis**). Those drawbacks can be overcome if the DNA polymerase catalysed extension of fluorescently labelled primers is analysed using CE. Thereby, the throughput can be dramatically increased, as CE offers the possibility for multiplexing. Since this method is by default used for Sanger DNA sequencing, microsatellite analysis or single nucleotide polymorphism analysis,[62, 115] it was necessary to adopt this method to meet my requirements.[137]

First, separation of short, fluorescently labelled DNA oligonucleotides was optimised. Therefore, two FAM labelled oligomers (21mer and 22mer) were purchased. Conditions applied during separation were modified and adapted until single base resolution of the oligomer mixtures was ensured (**Figure 22**). Run parameters established for this separation are listed in **6.2.9. Capillary Electrophoresis**

Capillary Electrophoresis separates fluorescently labelled oligonucleotides by size and charge as they migrate through a polymer filled capillary. Single nucleotide incorporation primer extension experiments are conducted as described (6.2.6. Primer Extension Assay[136]). Instead of radioactively labelled primers, FAM labelled ones are employed and reactions are stopped by addition of stop solution without bromophenol blue and xylene cyanol (80 % (v/v) formamide, 20 mM EDTA). CE samples are prepared by mixing 10 µl LIZ size standard diluted 1:80 in HiDi formamide (Applied Biosystems) and 10 µl fluorescently labelled reaction mixture (after addition of stop solution diluted 1:10 with HiDi formamide, to dilute EDTA). If multiple reactions are to be analysed in parallel, those are pooled in one well prior to injection. The reaction sample and size standard are injected electrokinetically into a 50 cm capillary array filled with Performance Optimised Polymer (POP6). High voltage electrophoresis (15 kV) over 180 s ensures single base resolution. Typical run parameters are depicted in **Table 4.**

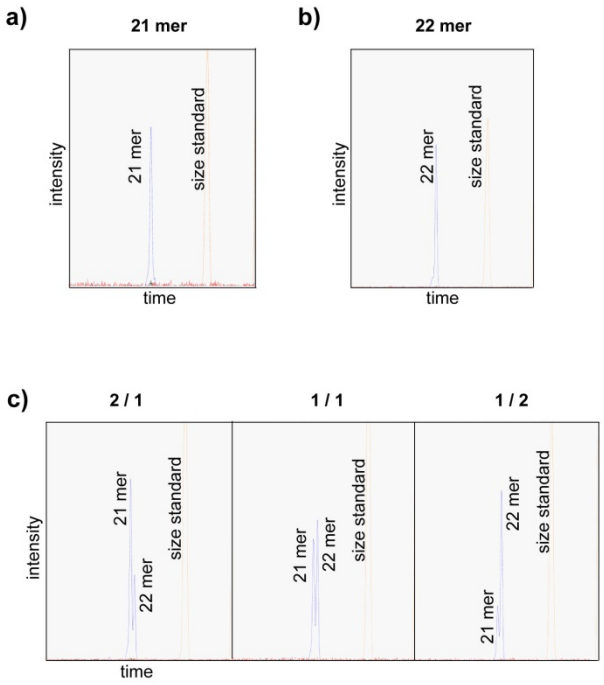


Figure 22: Separation of fluorescent primers using CE. **a)** Migration behaviour of a FAM labelled 21mer (DNA) analysed by capillary electrophoresis; **b)** migration behaviour of a FAM labelled 22mer (DNA) analysed by capillary electrophoresis; **c)** migration comparison of different mixtures (2/1; 1/1; 1/2) of a 21mer and 22mer (DNA) to prove sufficient resolution for analysis of single - nucleotide incorporation primer extension reactions. orange: size standard (LIZ 120), blue: FAM labeled primers.

The established assay was applied for single-nucleotide incorporation primer extension experiments. Therefore, processing of dGTP by the DNA polymerase 9°North exo$^-$ was examined by the usage of fluorescently labelled primers and analysis by CE (see **Figure 23 a** - **c**). Radioactively labelled primer extension experiments and separation by PAGE (see **Figure 23d**) was performed additionally to allow comparison between both methods.

Comparison of the quantitative evaluation of both analysis methods showed that the results are very similar. Therefore, evaluation of incorporation of the modified nucleotides by DNA polymerase 9°North exo$^-$ will be examined by CE. To take full advantage of the possibilities for automating supplied by employing CE, the described assay was adapted to allow the analysis of several reactions in one capillary and one run. In this approach single-nucleotide incorporation primer extension experiments are analysed. Thus, several reactions can be separated during one run in one capillary, if those extension reactions are performed using fluorescently labelled oligomers of different sizes. As depicted in **Figure 23e** sufficient separation could be ensured, if those primers were designed with a size difference of 8 nucleotides varying in size from 21 to 55 nucleotides.

With this experimental setup, it is possible to analyse 5 reactions in one run and one capillary in parallel enhancing the through put. The principles of this assay can be applied to further scale up the through put. Processed primers can be labelled using different dye-labels having well separated excitation and emission spectra (e.g. FAM, NED, VIC and PET). Additionally, longer primers can be used for analysis. Simultaneous analysis of multiple substrates or reaction conditions can therefore be enabled by multiplexing oligonucleotide design by size and fluorescent dye.

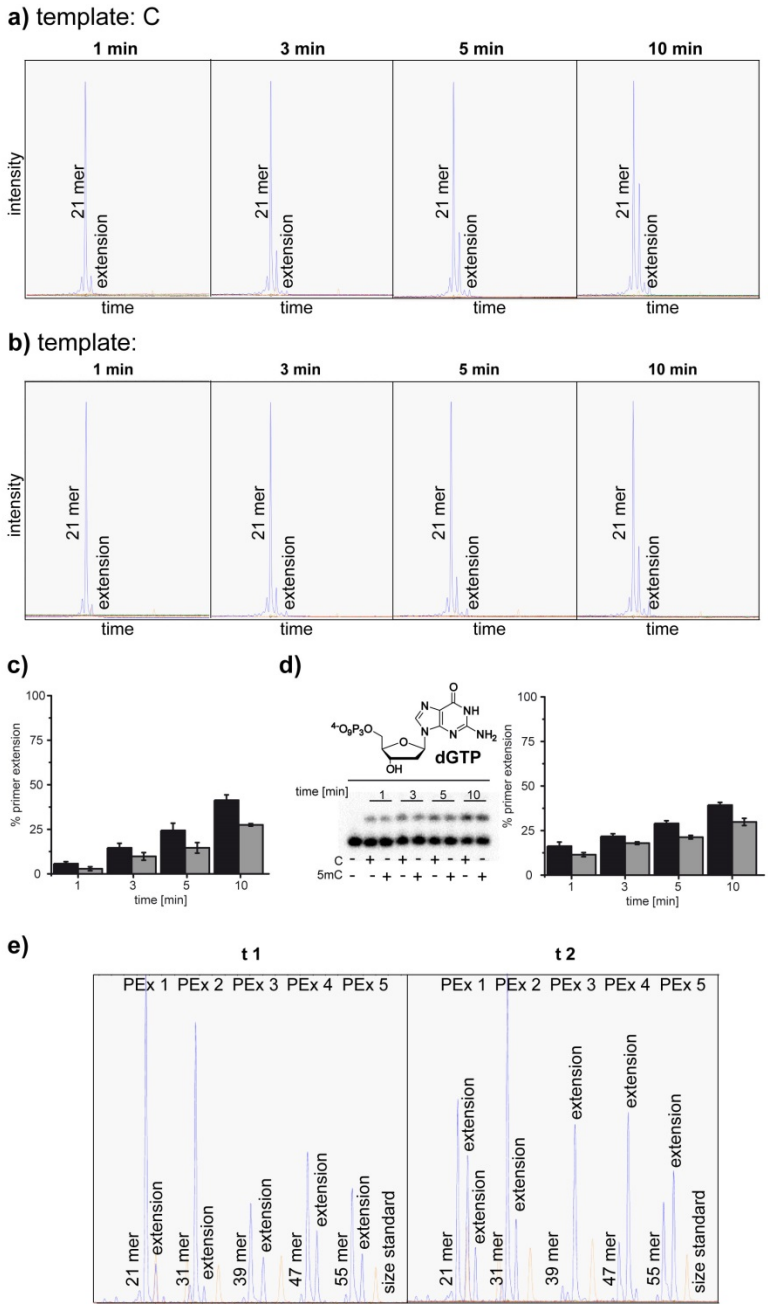


Figure 23: Capillary electrophoretic and PAGE analysis of single-nucleotide incorporation primer extension experiments of dGMP employing DNA polymerase 9°North exo$^-$. **a)** Opposite a template containing C; **b)** opposite a template containing C; **c)** quantitative analysis of single-nucleotide primer experiments depicted in **a** (black) and **b** (grey); **d)** PAGE analysis and quantitative evaluation of single-nucleotide incorporation primer extension experiments of dGMP employing DNA polymerase 9°North exo$^-$; **e)** capillary electrophoretic analysis of 5 different single-nucleotide incorporation experiments employing different lengths of FAM labelled primers in parallel. 50 µM dGTP and 10 nM 9°North exo$^-$ were used; reactions were stopped after indicated time points. Experiments were done at least in triplicates.

3.3.1. Discrimination of 5mC by Emplyoing Modified Nucleotides and 9°North DNA Polymerase

Testing all in position 6 modified dGTP derivatives for incorporation opposite C and 5mC employing the second B family DNA polymerase 9°North, it can be observed that all nucleotides are successfully

processed (see **Figure 24**). As already observed before, incorporation efficiencies of the dGMP analogues decrease by increasing size of the introduced modification. Additionally, incorporation opposite C is favoured for all nucleotides over incorporation opposite 5mC. Again, discrimination can already be observed for processing of the unmodified dGTP. This discrimination ratio enhances by the usage of some of the modified nucleotides. In contrast to the results obtained by employing the DNA polymerase KOD exo$^-$, the alkylation of dGTP in position 6 (**1a** - **1d**) does not improve discrimination behaviour. Instead, the amino-modified nucleotides (**9a** - **9o**) show remarkable improvement in differences in incorporation efficiencies using 9° North exo$^-$. The highest discrimination rates can be observed by employing nucleotides **9b**, **9c** or **9n** (see **Figure 24** and **Figure 25**).

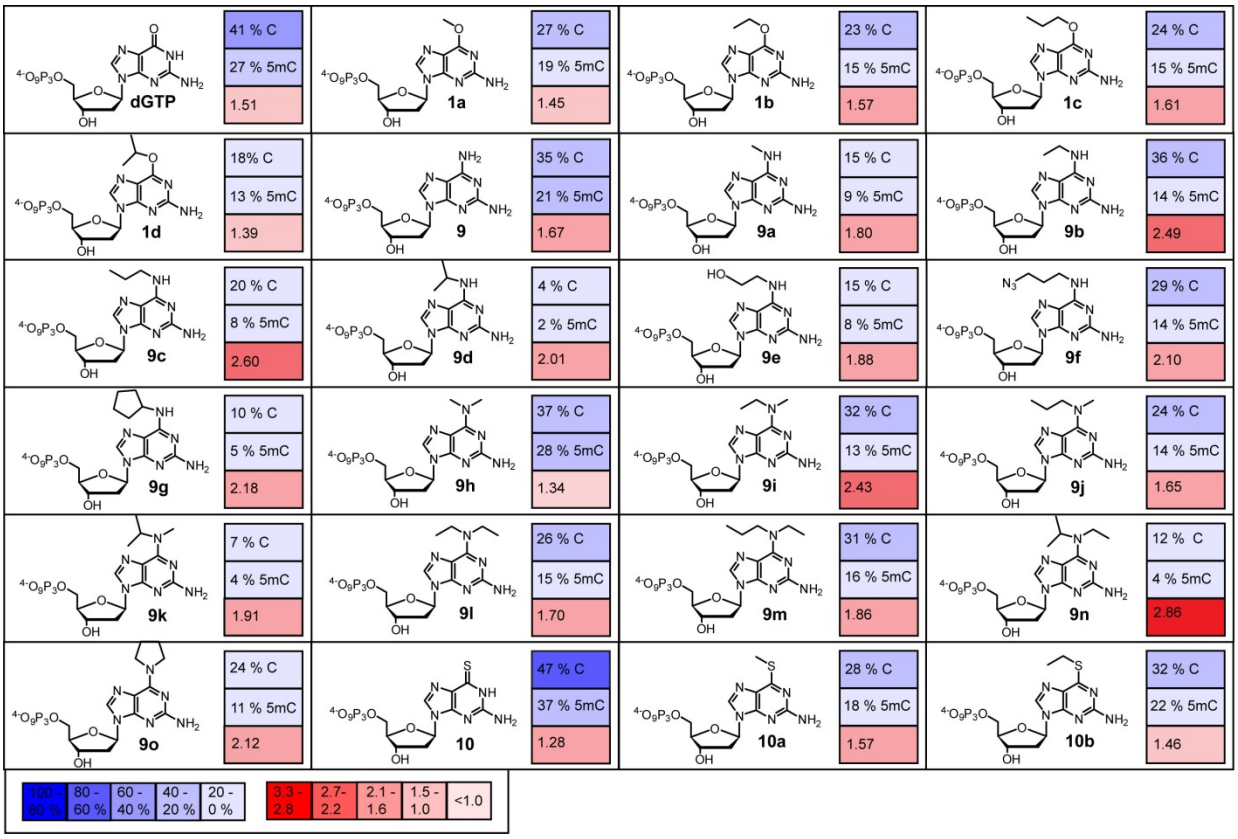


Figure 24: Structures of modified dG*TP analogues including % incorporation opposite C or 5mC employing DNA polymerase 9°North exo$^-$ in single-nucleotide primer extension experiments. 50µM dGTP or dG*TP and 10 nM 9°North exo$^-$ were used, reactions were stopped after 10 min. Experiments were done at least in triplicates. Arithmetic mean is given; errors are given in the appendix.

But still, more pronounced discrimination behaviour can be observed by employing the DNA polymerase KOD exo$^-$ in combination with nucleotide **1b**. Even if the discrimination ratio is not further improved by usage of the second B family DNA polymerase 9°North exo$^-$, we proved the tendencies observed in previous studies: this DNA polymerase was capable to incorporate all in position 6 modified nucleotides. Additionally, discrimination can already be observed by incorporation of unmodified dGMP and this discrimination can be further improved by processing of modified nucleotides.

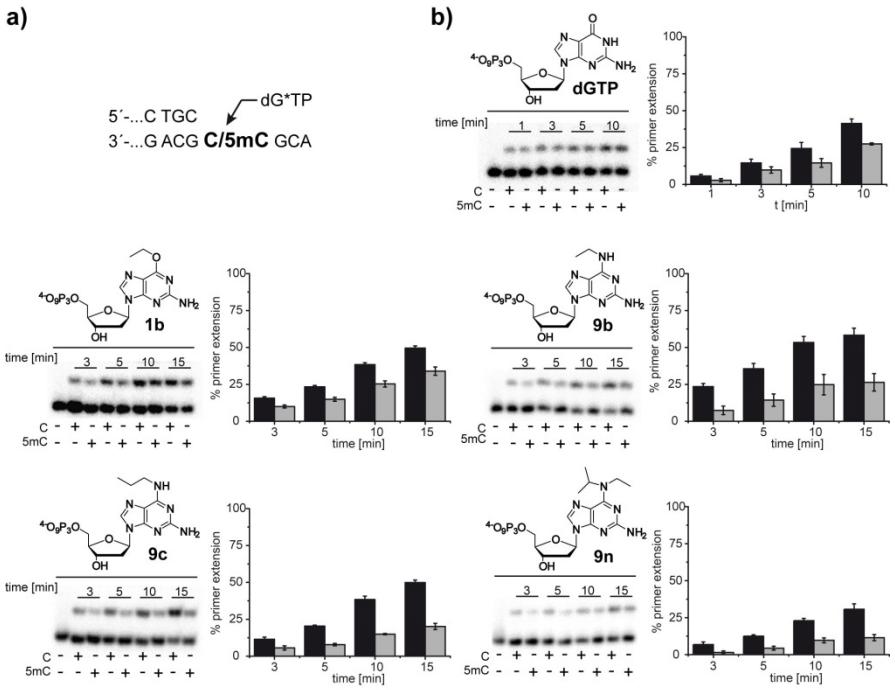
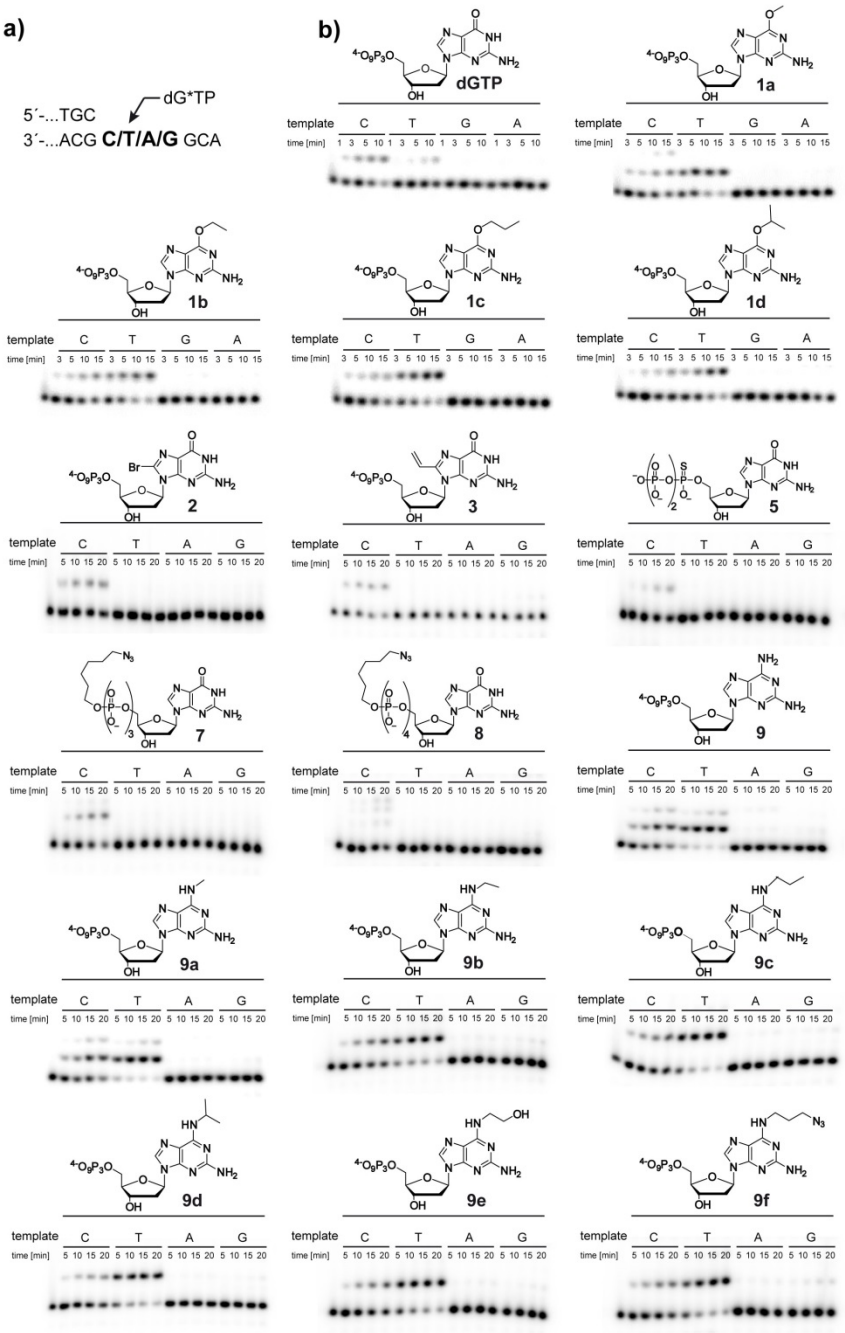


Figure 25: Incorporation experiments of modified nucleotides leading to most pronounced differences in incorporation efficiencies employing DNA polymerase 9° North exo⁻. **a)** Partial primer / template sequence used. **b)** PAGE analysis and quantitative evaluation of single-nucleotide incorporation primer extension experiments of dGMP and nucleotides **1b**, **9b**, **9c** and **9n** opposite a template containing C (black) in comparison to a template containing 5mC (grey) employing the DNA polymerase 9°North exo⁻. 50 µM dGTP or dG*TP and 10 nM 9°North exo⁻ were used; reactions were stopped after indicated time points. Experiments were done at least in triplicates.

3.4. Selectivity Studies for Incorporation of Modified Nucleotides

Single-nucleotide incorporation primer extension experiments employing modified nucleotides and the B family DNA polymerase KOD exo$^-$ showed promising discrimination behaviour, which could be proven by steady-state kinetics (see **Figure 21** and **Table 1**). To further investigate the potential of this approach to be used for 5mC sequencing, selectivity studies needed to be performed. To determine, if this approach is suited for sequencing applications, it had to be ensured, that incorporation of those modified nucleotides opposite C or 5mC is favoured over incorporation opposite T, A and G. Therefore, single-nucleotide incorporation primer extension experiments were performed employing all modified nucleotides in combination with the DNA polymerase KOD exo$^-$ with templates containing C, T, A or G (see **Figure 26**). As nucleotides **4** and **6** do not lead to significant incorporation under the chosen conditions, those nucleotides were not considered.

Unfortunately, it can be seen, that incorporation of all in position 6 modified nucleotides (**1a** - **1d**, **9**, **9a** - **9o**, **10a** and **10b**) is favoured opposite T. Therefore, 5mC detection employing those nucleotides needs to be performed site specific[1] or requires two sequencing runs, as known for bisulfite-sequencing. If the nucleotides, modified at position 8 (**2** + **3**) were used as substrate for the DNA polymerase-catalysed reaction, no incorporation opposite T can be observed under the chosen conditions. The same holds true for processing of the phosphate-chain modified nucleotides (**5**, **7** + **8**). As favoured incorporation opposite C and 5mC can be observed for those nucleotides, while no processing opposite T is detectable, selectivity of those nucleotides for C is ensured. Those findings will be further employed in the experiments described in **3.5. Doubly Modified dGTP Derivatives**.


a)
dG*TP
5´-...TGC
3´-...ACG C/T/A/G GCA
b)
dGTP
1a
1b
1c
1d
2
3
5
7
8
9
9a
9b
9c
9d
9e
9f
template
time [min]
C T G A
C T A G

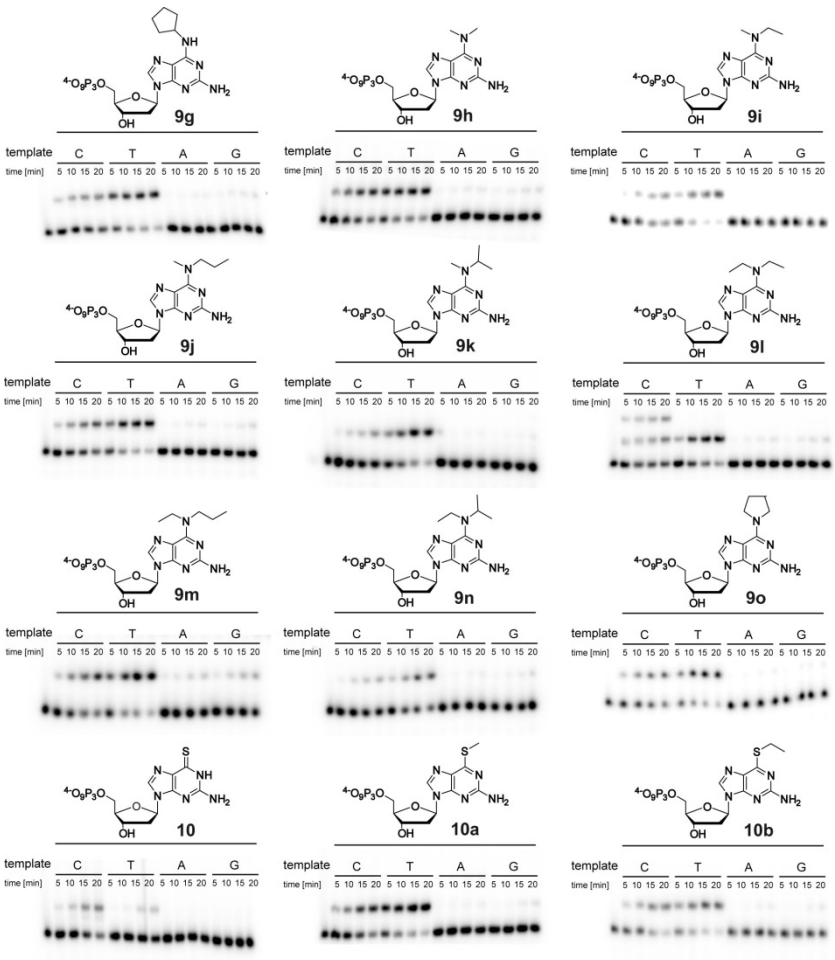


Figure 26: PAGE analysis of single-nucleotide incorporation primer extension experiments of dGMP and modified nucleotides **1** - **10** opposite a template containing C, T, G or A employing DNA polymerase KOD exo$^-$. 50 µM dGTP or dG*TP and 5 nM KOD exo$^-$ were used; reactions were stopped after indicated time points.

3.5. Doubly Modified dGTP Derivatives for Improved Discrimination of 5mC

Those nucleotides, not modified at the Watson Crick base pairing site, but modified at position 8 or the phosphate-chain as well as nucleotide **10**, show selectivity for incorporation opposite C (see **Figure 26**). Employing those nucleotides in combination with the DNA polymerase KOD exo$^-$, discrimination can be observed (see **Figure 15**). The best discrimination of those nucleotides with preserved selectivity can be detected for nucleotide **2**. Since this discrimination is not sufficient for application in 5mC detection, two different modifications should be combined in one molecule. Hence, it was decided to synthesise nucleotides **30a** + **b**, thereby combining the modifications of nucleotides **1a** or **1b** and **2**, since nucleotide **1b** showed the best results for discrimination between C and 5mC and nucleotide **2** showed discrimination while preserving selectivity. Combining the modifications of nucleotides **1a** and **10** in synthesising nucleotide **10a** showed already improvement of discrimination (see **Figure 20**).

3.5.1. Synthesis of 8-Br-*O*^{6}-Alkyl-dGTP

The doubly modified dGTP derivatives **30a**+**b** (see **Figure 27**) were synthesised starting from commercially available 2´-deoxyguanosine, which was brominated in position 8 as described before (see **3.1.1 Synthesis**). Chlorination in position 6 after bromination in position 8 proved to be challenging. Due to the altered electrochemical surrounding,[118] chlorination in position 6 could not be achieved. One possible reason is that halogenation of the nucleobase is known to enhance the reactivity of the exocyclic amino-group, thus leading to possible side reactions. Hence, we decided to use a Mitsunobu reaction for ether formation. Mitsunobu reactions are known to transform OH-groups into different functional groups like esters, ethers or amines.[138] Therefore, the remaining sugar OH-groups needed to be protected to avoid side reactions during the Mitsunobu transformation.

Synthesis of nucleotides **30a** and **30b** was started by bromination of 2´-deoxyguanosine, as described before. Next, the OH-groups of the brominated nucleoside **14** needed to be masked. It was decided to protect those hydroxyl-groups by reaction with 1,3-dichloro-1,1,3,3-tetra-*iso*-propyl-disiloxane. This protecting group is known to be stable under the applied reaction conditions and is supposed to be cleaved smoothly by reaction with triethylamine trihydrofluoride.[131] Thus, it was possible to obtain the nucleoside precursor **27**, which was used for Mitsunobu reaction, in 85 % yield. The subsequent modification reaction was performed analogously to known procedures,[119] whereby both modifications (methoxy and ethoxy) could be introduced to yield nucleosides **28a**+**b** each in 57 %. Deprotection of the OH-groups was carried out as describe before in good yields. Last, the nucleosides **29a**+**b** were transformed into the corresponding 5´-triphosphates by standard procedures in 7 - 9 % yield (see **Figure 27**).

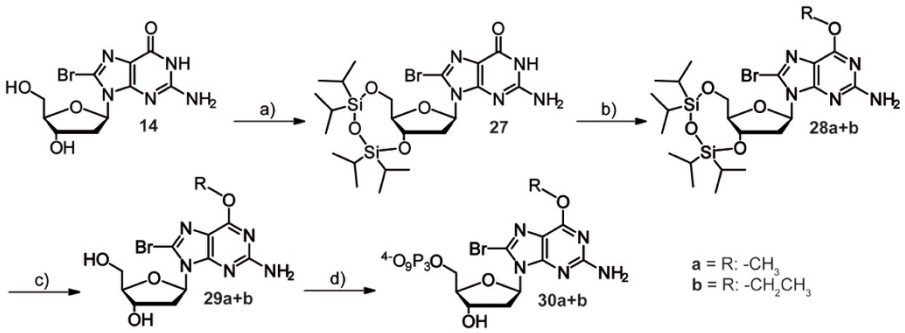


Figure 27: Synthesis of doubly modified nucleotides. a) 1,3-Dichloro-1,1,3,3-tetra-*iso*-propyl-disiloxane, pyridine, rt, 16 h, yield: 87 %; b) 1: PPh_3, 1,4-dioxane, ROH, diethyl azodicarboxylate, 0 °C, 10 min; 2: rt, 1h, yield: 57 % c) triethylamine trihydrofluoride, THF, rt, 16 h, yield: 80 % - quantitative; d) 1: proton sponge, $POCl_3$, TMP, 30 min, 0 °C; 2: $(Bu_3NH)_2H_2P_2O_7$, *n*Bu_3N, DMF, 30 min, rt. 3: 0.1 M TEAB, 10 min, rt, yield: 7 - 9 %.

3.5.2. Incorporation Studies of Doubly Modified Nucleotides

The incorporation of doubly modified nucleotides **30a**+**b** was investigated employing the DNA polymerase KOD exo$^-$ (see **Figure 28**). As shown in **Figure 28**, those doubly modified nucleotides were incorporated with reduced efficiencies. Therefore, DNA polymerase concentration was increased to 20 nM to achieve noteworthy incorporation. Processing of nucleotide **30a** opposite templates coding for C or 5mC does not show any promising discrimination effects. Employing nucleotide **30b**, discrimination, which is comparable to the discrimination achieved by processing of nucleotide **2**, can be observed. But since incorporation leads to several extension bands on the PAGE gel, evaluation of these bands is difficult. To investigate, whether selectivity of those modified nucleotides is maintained, single-nucleotide incorporation primer extension experiments were performed opposite a template coding for C, T, A or G (see **Figure 28 c**).

Incorporation experiments of nucleotide **30b** opposite C, T, A and G clearly point out that selectivity for C is lost due to the combination of both modifications. Under the chosen conditions, nucleotide **30b** is preferably incorporated opposite G instead of C. Therefore, combination of two of the most promising modifications in one molecule does not lead to any promising results. Neither discrimination, nor selectivity of the synthesised nucleotides seem to be improved.

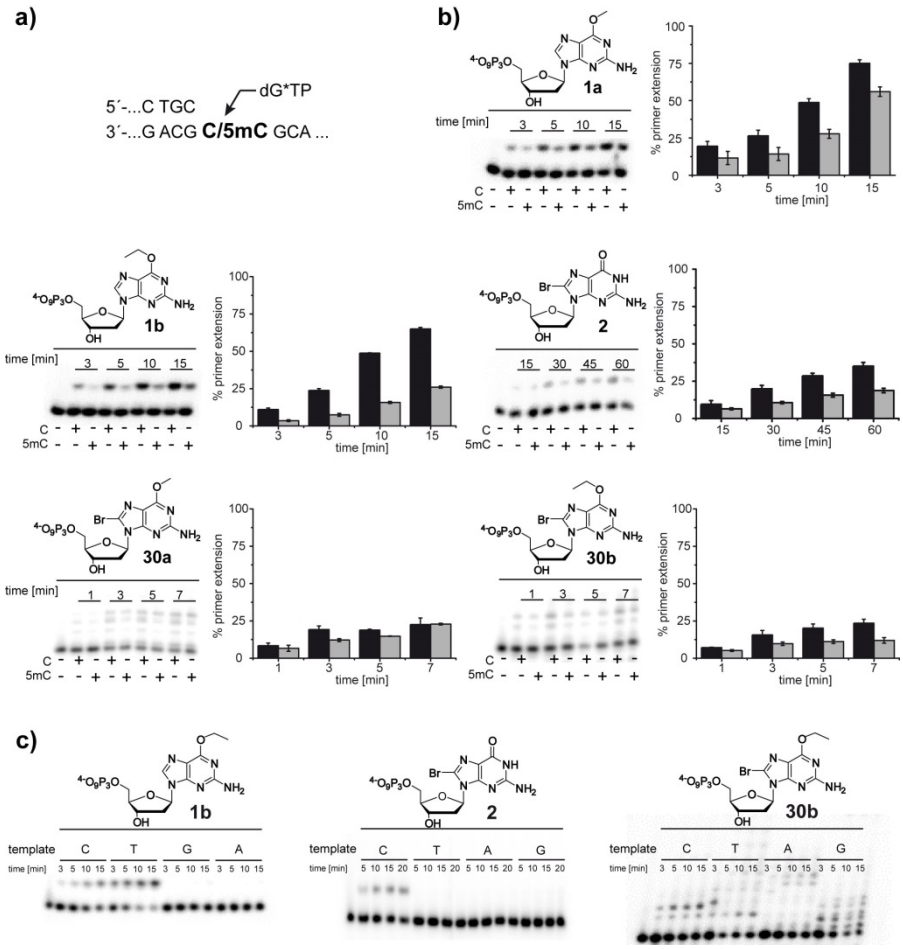


Figure 28: Incorporation experiments of nucleotides **1a**, **1b**, **2**, **30a** and **30b**. a**)** Partial primer / template sequence used. **b)** PAGE analysis and quantitative evaluation of single-nucleotide incorporation primer extension experiments of dGMP and nucleotides **1a**, **1b**, **2**, **30a** and **30b** opposite a template containing C (black) in comparison to a template containing 5mC (grey) employing the DNA polymerase KOD exo$^-$. **c)** PAGE analysis of single-nucleotide incorporation primer extension experiments of nucleotides **1b**, **2** and **30b** opposite a template containing C, T, A or G. 50 µM dGTP or dG*TP and 5 nM (nucleotides **1a**, **1b**, **2**) or 20 nM (nucleotides **30a** and **30b**) KOD exo$^-$ were used; reactions were stopped after indicated time points. Experiments were done at least in triplicates.

3.6. Screening of KOD exo$^-$ Libraries Regarding Discrimination of 5mC

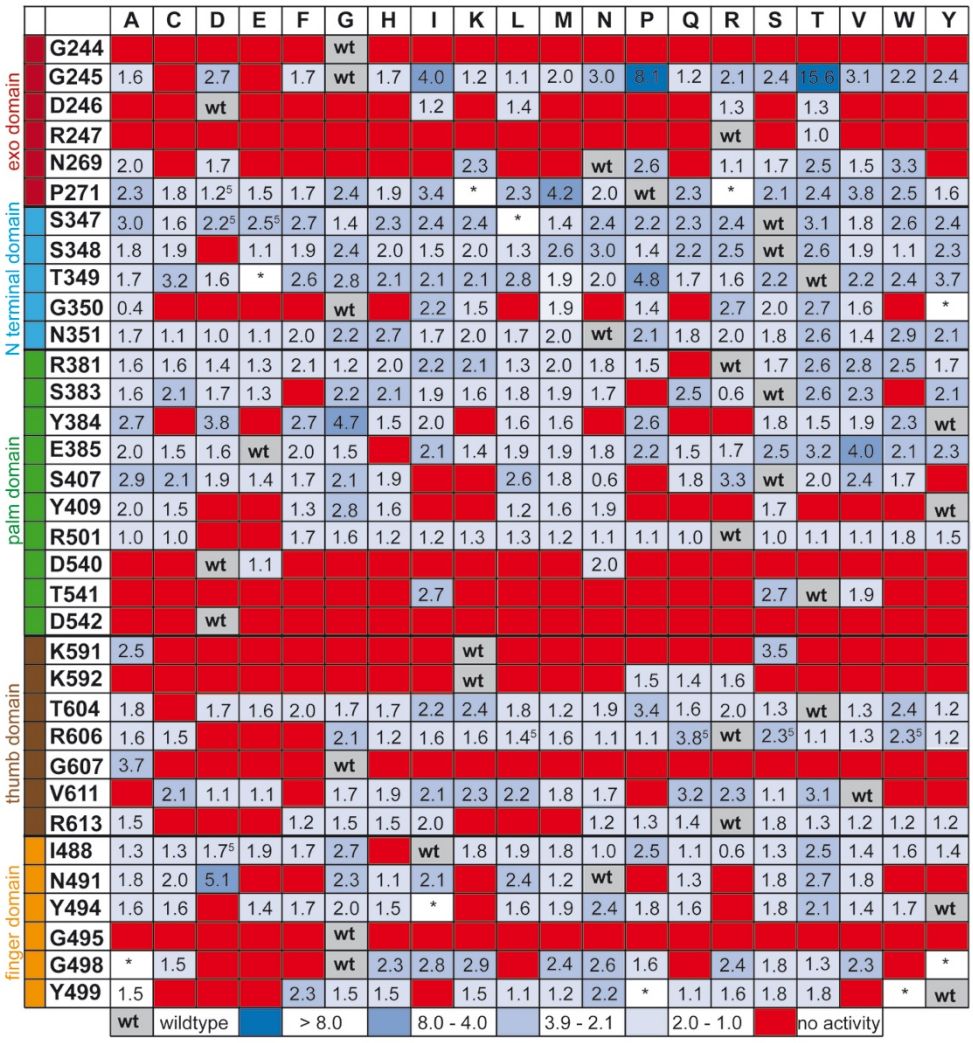
In previous experiments, the scope of modified nucleotides for 5mC detection was demonstrated. Nevertheless, its limitations were discovered as well. I could show that modified nucleotides are processed opposite C and 5mC with remarkably diverging incorporation efficiencies, therefore enabling sensing of this epigenetic marker. Though, I had to realise as well that many modified nucleotides are not only efficiently incorporated opposite C, but some of them with even higher efficiencies opposite T. The lost selectivity prohibits the application of this discrimination for sequencing approaches via one single sequencing run. Nevertheless, the applicability of those findings for site-specific 5mC detection could be demonstrated in further studies together with Kim Leitner.[1] Additionally, it is possible to exploit those effects for 5mC sequencing via two independent sequencing runs, as it was already shown for BS-Seq. This approach still bears the advantage, that no previous modification step of the genomic DNA would be required.

Since I already exploited the potential of modified nucleotides, the next step is modifying the used DNA polymerase. I aimed at identifying mutants with increased discrimination in incorporating dGMP opposite 5mC compared to C. As DNA polymerases normally replicate 5mC without difficulty and the methyl group in 5mC does not hinder Watson-Crick base pairing, this is a challenging task. However, archeal DNA polymerases like the KOD DNA polymerase possess the ability to detect U in DNA.[105] As U differs from T only by an additional methylation in 5 position, T can be seen as 5mU. Additionally, the B family DNA polymerase KOD exo$^-$ was verified before to be most promising for those approaches. KOD exo$^-$ DNA polymerase variants with an increased ability for 5mC sensing would provide a valuable tool for the application for DNA methylation profiling.

Therefore, it was decided to generate libraries of site specifically mutated KOD exo$^-$. Claudia Huber and Martina Adam generated 34 libraries by mutating several sites in different domains of the DNA polymerase, which look most promising in respective of showing possible interaction sites with the substrates in the active site. In order to achieve all 19 possible mutants at every investigated site, they introduced defined mutations at the respective sites via site-directed mutagenesis. In the following, those libraries were tested by Claudia Huber and Martina Adam for activity in real-time PCR to distinguish between active and inactive mutants. Only those mutants showing a threshold-crossing point (Ct value) of less than 30 cycles were considered as active and investigated in further screening experiments.

To reveal the most promising mutant for 5mC detection, I performed single-nucleotide incorporation primer extension experiments employing the cleared bacterial lysates of those mutants which were identified as active. Those mutants were applied in processing two different nucleotides: dGTP and 8-modified nucleotide **3**. Nucleotide **3** was chosen for further investigations, because this nucleotide led to decent discrimination ratios while maintaining selectivity of the DNA polymerase for C (see **Figure 15** and **Figure 26**). As previous experiments pointed to decreased selectivity between C and T, I employed incorporation opposite three different templates containing C, 5mC and T to ensure that

selectivity of the mutated DNA polymerases opposite C was maintained. Discrimination ratios were determined after 10 min of incubation time and are summarised in **Figure 29** and **Figure 30**.



	A	C	D	E	F	G	H	I	K	L	M	N	P	Q	R	S	T	V	W	Y
G244						wt														
G245	1.6		2.7		1.7	wt	1.7	4.0	1.2	1.1	2.0	3.0	8.1	1.2	2.1	2.4	15.6	3.1	2.2	2.4
D246			wt					1.2		1.4					1.3		1.3			
R247															wt		1.0			
N269	2.0		1.7						2.3			wt	2.6		1.1	1.7	2.5	1.5	3.3	
P271	2.3	1.8	1.2[5]	1.5	1.7	2.4	1.9	3.4	*	2.3	4.2	2.0	wt	2.3	*	2.1	2.4	3.8	2.5	1.6
S347	3.0	1.6	2.2[5]	2.5[5]	2.7	1.4	2.3	2.4	2.4	*	1.4	2.4	2.2	2.3	2.4	wt	3.1	1.8	2.6	2.4
S348	1.8	1.9		1.1	1.9	2.4	2.0	1.5	2.0	1.3	2.6	3.0	1.4	2.2	2.5	wt	2.6	1.9	1.1	2.3
T349	1.7	3.2	1.6	*	2.6	2.8	2.1	2.1	2.1	2.8	1.9	2.0	4.8	1.7	1.6	2.2	wt	2.2	2.4	3.7
G350	0.4					wt		2.2	1.5		1.9		1.4		2.7	2.0	2.7	1.6		*
N351	1.7	1.1	1.0	1.1	2.0	2.2	2.7	1.7	2.0	1.7	2.0	wt	2.1	1.8	2.0	1.8	2.6	1.4	2.9	2.1
R381	1.6	1.6	1.4	1.3	2.1	1.2	2.0	2.2	2.1	1.3	2.0	1.8	1.5		wt	1.7	2.6	2.8	2.5	1.7
S383	1.6	2.1	1.7	1.3		2.2	2.1	1.9	1.6	1.8	1.9	1.7		2.5	0.6	wt	2.6	2.3		2.1
Y384	2.7		3.8		2.7	4.7	1.5	2.0		1.6	1.6		2.6			1.8	1.5	1.9	2.3	wt
E385	2.0	1.5	1.6	wt	2.0	1.5		2.1	1.4	1.9	1.9	1.8	2.2	1.5	1.7	2.5	3.2	4.0	2.1	2.3
S407	2.9	2.1	1.9	1.4	1.7	2.1	1.9			2.6	1.8	0.6		1.8	3.3	wt	2.0	2.4	1.7	
Y409	2.0	1.5			1.3	2.8	1.6			1.2	1.6	1.9				1.7				wt
R501	1.0	1.0			1.7	1.6	1.2	1.2	1.3	1.3	1.2	1.1	1.1	1.0	wt	1.0	1.1	1.1	1.8	1.5
D540			wt	1.1								2.0								
T541								2.7								2.7	wt	1.9		
D542			wt																	
K591	2.5								wt							3.5				
K592									wt				1.5	1.4	1.6					
T604	1.8		1.7	1.6	2.0	1.7	1.7	2.2	2.4	1.8	1.2	1.9	3.4	1.6	2.0	1.3	wt	1.3	2.4	1.2
R606	1.6	1.5				2.1	1.2	1.6	1.6	1.4[5]	1.6	1.1	1.1	3.8[5]	wt	2.3[5]	1.1	1.3	2.3[5]	1.2
G607	3.7					wt														
V611		2.1	1.1	1.1		1.7	1.9	2.1	2.3	2.2	1.8	1.7		3.2	2.3	1.1	3.1	wt		
R613	1.5				1.2	1.5	1.5	2.0				1.2	1.3	1.4	wt	1.8	1.3	1.2	1.2	1.2
I488	1.3	1.3	1.7[5]	1.9	1.7	2.7		wt	1.8	1.9	1.8	1.0	2.5	1.1	0.6	1.3	2.5	1.4	1.6	1.4
N491	1.8	2.0	5.1			2.3	1.1	2.1		2.4	1.2	wt		1.3		1.8	2.7	1.8		
Y494	1.6	1.6		1.4	1.7	2.0	1.5	*		1.6	1.9	2.4	1.8	1.6		1.8	2.1	1.4	1.7	wt
G495						wt														
G498	*	1.5				wt	2.3	2.8	2.9		2.4	2.6	1.6		2.4	1.8	1.3	2.3		*
Y499	1.5				2.3	1.5	1.5		1.5	1.1	1.2	2.2	*	1.1	1.6	1.8	1.8		*	wt

Figure 29: Discrimination chart of generated KOD exo$^-$ DNA polymerase mutants employing single-nucleotide incorporation of dGMP. Mutated sites are listed on the left according to colour coded protein domain. Amino acid exchanges are depicted in the top lane. Colour coding indicates discrimination ratio of mutants in bacterial lysates, reactions were stopped after 10 min (except x^5; reactions were stopped after 5 min); mutants marked with an asterix (*) could not be obtained during cloning and are missing from the library. 100 µM dGTP and 1.5 µl cleared bacterial lysate (dilution 1/50) was used per 15 µl reaction.

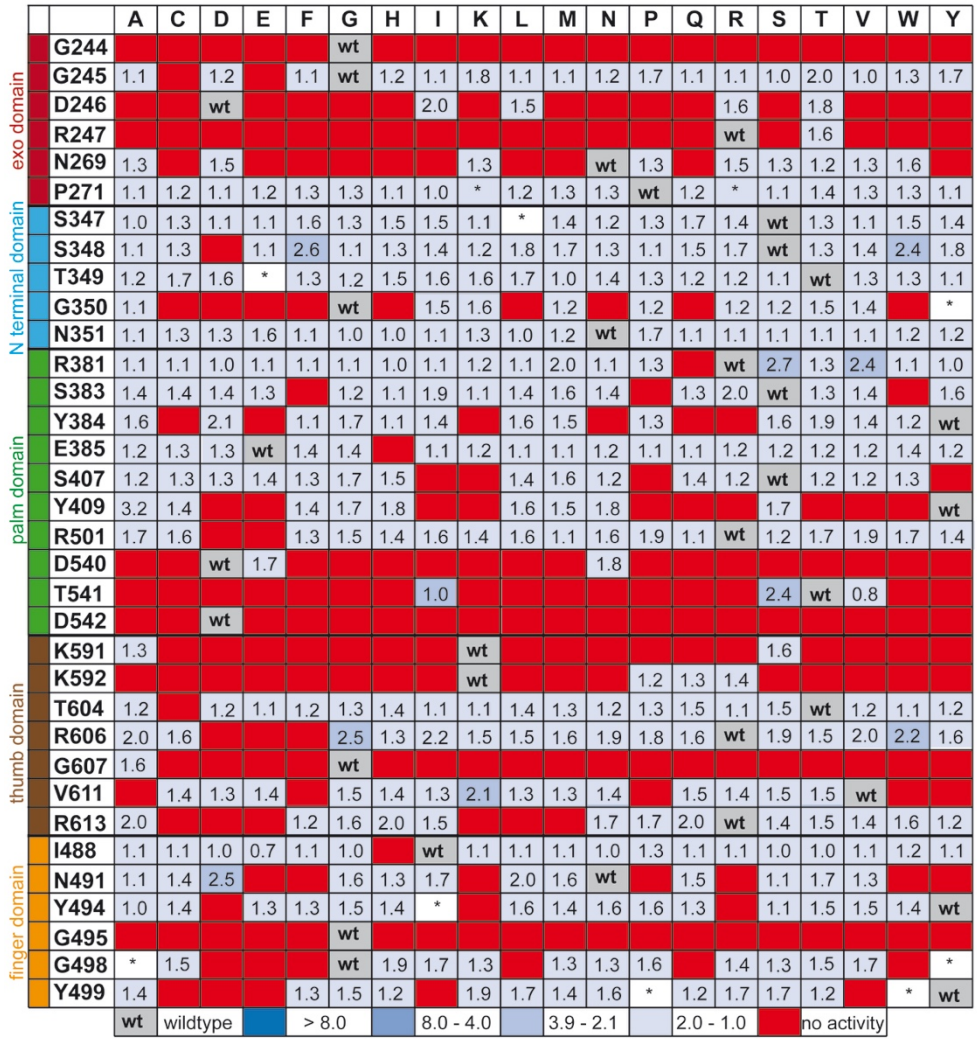


Figure 30: Discrimination chart of generated KOD exo$^-$ DNA polymerase mutants employing single-nucleotide incorporation of **3**. Mutated sites are listed on the left according to colour coded protein domain. Amino acid exchanges are depicted in the top lane. Colour coding indicates discrimination ratio of mutants in bacterial lysates, reactions were stopped after 10 min; mutants marked with an asterix (*) could not be obtained during cloning and are missing from the library. 100 µM **3** and 1.5 µl cleared bacterial lysate (dilution 1/1) was used per 15 µl reaction.

The depicted results show that mutation at some sites leads exclusively to variants with remarkable reduced activity, despite expected protein content of the cleared bacterial lysates. Additionally, it can be observed, that some mutations lead to increased differences in incorporation efficiencies of dGMP opposite C or 5mC, while selectivity for favoured processing opposite C over T was ensured (respective gels are depicted in the appendix). For some mutants, no increase in discrimination can be observed. During those investigations, it became obvious that employment of nucleotide **3** does not lead to a remarkable improvement in discrimination behaviour (see **Figure 29** and **Figure 30**), whereas efficiencies in progressing nucleotide **3** are clearly decreased in comparison to the usage of dGTP. Therefore, it was decided to focus on the employment of DNA polymerase KOD exo$^-$ mutants with the unmodified nucleotide dGTP. The most promising KOD exo$^-$ variant with regard to discrimination behaviour as well as selectivity was picked. The first semi-quantitative experiments clearly point to site G245 as the most promising site for further investigation. I focused on G245 KOD

exo⁻ variants with the highest discrimination: G245D, G245I, G245N, G245P, G245S, G245T, G245V and G245Y (see **Figure 29**), which were expressed and purified for further experiments.

Single-nucleotide incorporation primer extension experiments were subsequently performed employing the purified KOD exo⁻ mutants in comparison to the wild type with dGTP and templates containing C or 5mC. Those investigations prove enhanced discrimination for all enzymes except for the mutant G245Y (see **Figure 31**). Discrimination ratios between 1.4 (wild type) and 3.2 (G245I) were determined, thereby verifying a remarkable improvement of discrimination. Again, selectivity of those mutants for the incorporation of dGMP opposite C in comparison to T, A and G was ensured (see **Figure 33**). Those findings show the great potential of generating DNA polymerase variants to further enhance their capabilities for desired applications.

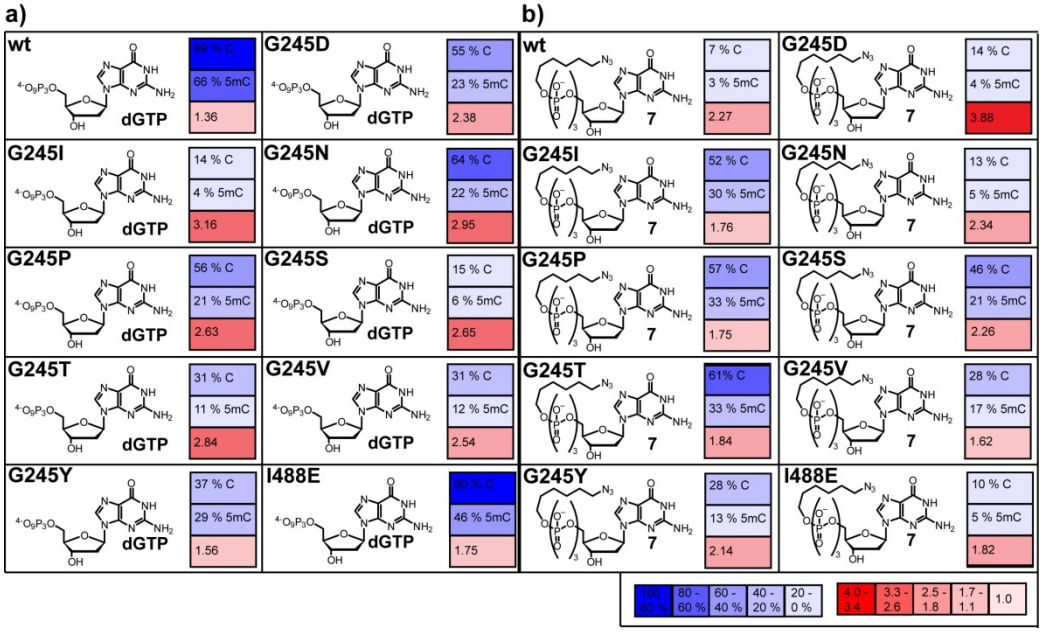


Figure 31: Structures of **a**) dGTP and modified **b**) dG*TP analogues **7** including % incorporation opposite C or 5mC employing DNA polymerase KOD exo⁻ wt and mutants in single-nucleotide primer extension experiments. 50µM dGTP or **7** and 5 nM KOD exo⁻ were used, reactions were stopped after 15 min. Discrimination ratios were determined by calculating the quotient of % incorporation opposite C and % incorporation opposite 5mC. Experiments were done at least in triplicates. Arithmetic mean is given; errors are given in the appendix.

Nucleotides that are fluorescently labelled at the terminal phosphate are required for next-generation sequencing methods, as SMRT.[60] Therefore, we tested nucleotide **7** in addition to the unmodified dGTP. Nucleotide **7** bears an azide, connected to the terminal γ-phosphate. This functional group can later on be used as attachment site for a fluorescent dye, as the azide can be reduced under mild conditions and the resulting amine exploited in active-ester chemistry (see **Figure 34**).

To ensure that the following quantitative primer extension reactions are representative, all experiments were done in replicates (see **Figure 31** and **Figure 32**). As it can be seen in **Figure 31** and **Figure 32**, mutation of the DNA polymerase leads to decreased efficiencies in processing the unmodified dGTP. In contrast, incorporation efficiency of nucleotide **7** was improved by employing the KOD exo⁻ mutants

in comparison to the wild type. Incorporation experiments of the most promising KOD exo$^-$ are depicted in **Figure 32** and **Figure 33**.

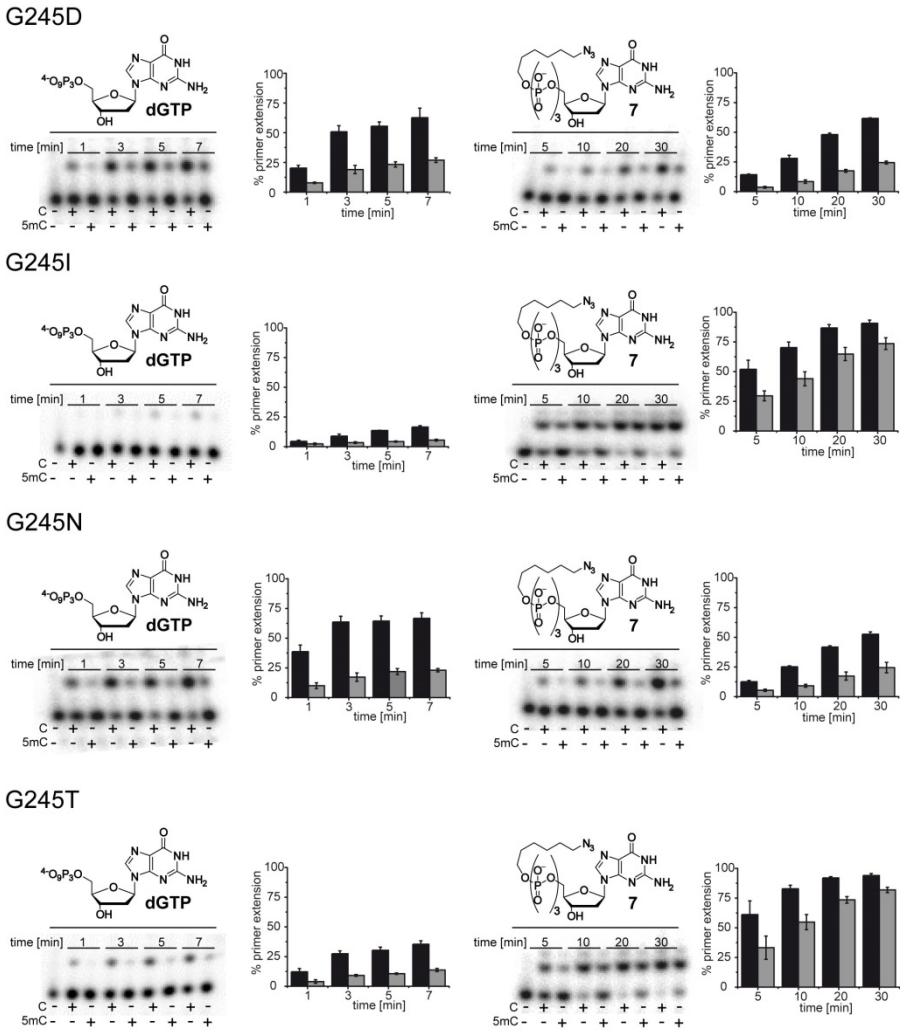


Figure 32: PAGE analysis and quantitative evaluation of single-nucleotide incorporation primer extension experiments of dGMP and nucleotides **7** opposite a template containing C (black) in comparison to a template containing 5mC (grey) employing DNA polymerase KOD exo$^-$ mutants 50 µM dGTP or **7** and 5 nM KOD exo$^-$ were used; reactions were stopped after indicated time points. Experiments were done at least in triplicates.

Even if incorporation efficiencies of dGTP decreased, discrimination between C and 5mC increased and could be enhanced to a considerable factor of 3.16 for the mutant G245I. But employing the mutant G245I with nucleotide **7** shows decreased discrimination. Best discrimination during incorporation of **7** can be observed for mutant G245D. By usage of nucleotide **7** with DNA polymerase variants, discrimination enhanced further to a factor of 3.88 (see **Figure 31** and **Figure 32**).

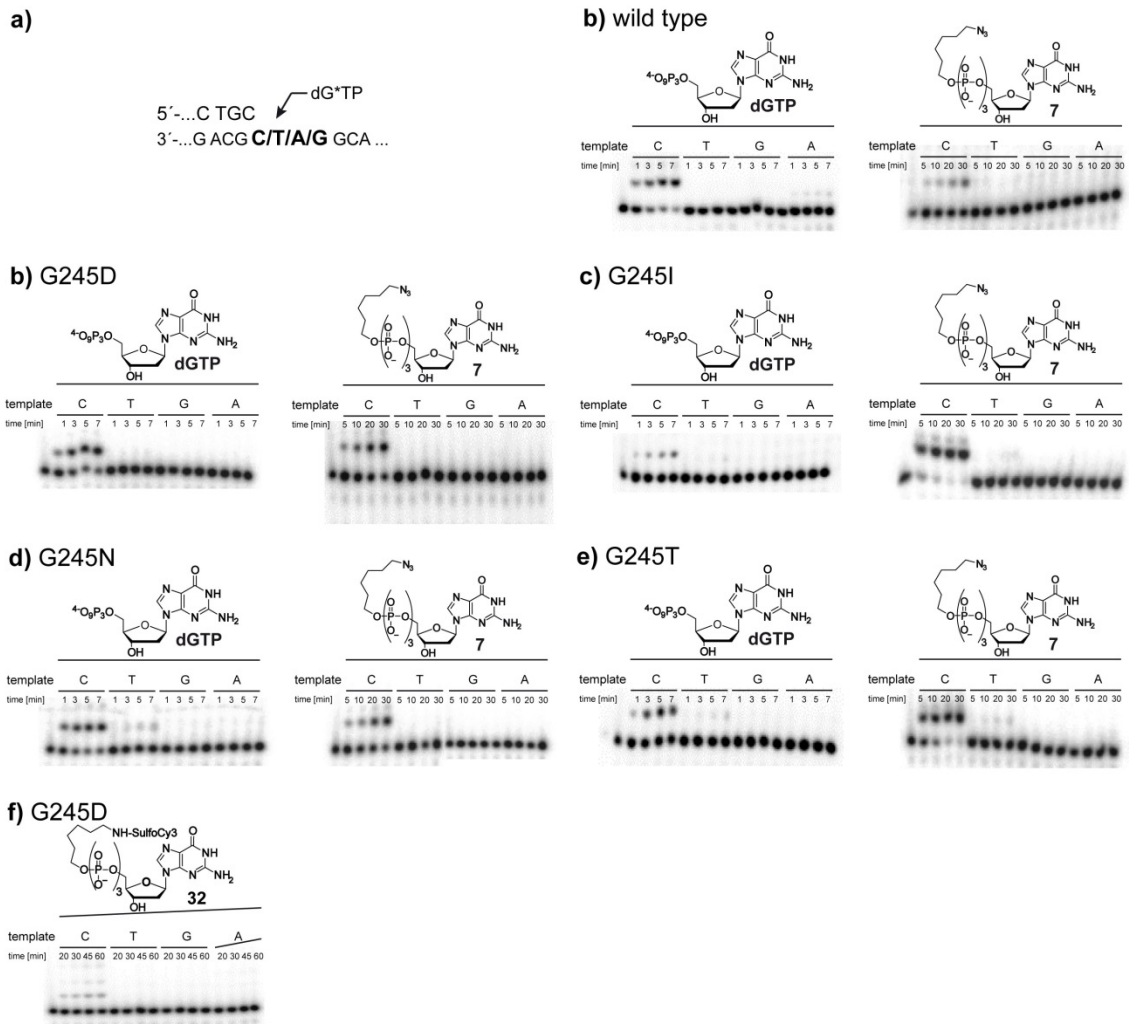


Figure 33: **a)** partial primer/template complex; PAGE analysis of single-nucleotide incorporation primer extension experiments using dGTP, nucleotide **7** and **32** (**f**) opposite a template containing C, T, A or G employing purified KOD exo$^-$ DNA polymerase mutants: **b)** wild type, **c)** G245D, **d)** G245I, **e)** G245N; 5nM DNA polymerase was used for processing dGTP, 20 nM enzyme was employed for processing nucleotides **7** and **32**, reactions were stopped after indicated time points.

Fluorescently labelled nucleotides are required for applications in SMRT sequencing.[60] The fluorescent analogue **32** was therefore synthesised starting from nucleotide **7** (see **Figure 34**). As mentioned above, nucleotide **7** was reduced under mild conditions in 77 % yield. The resulting nucleotide **31** was in the following converted to nucleotide **32** by coupling with the NHS ester of the fluorescent dye sylfo-Cy3.

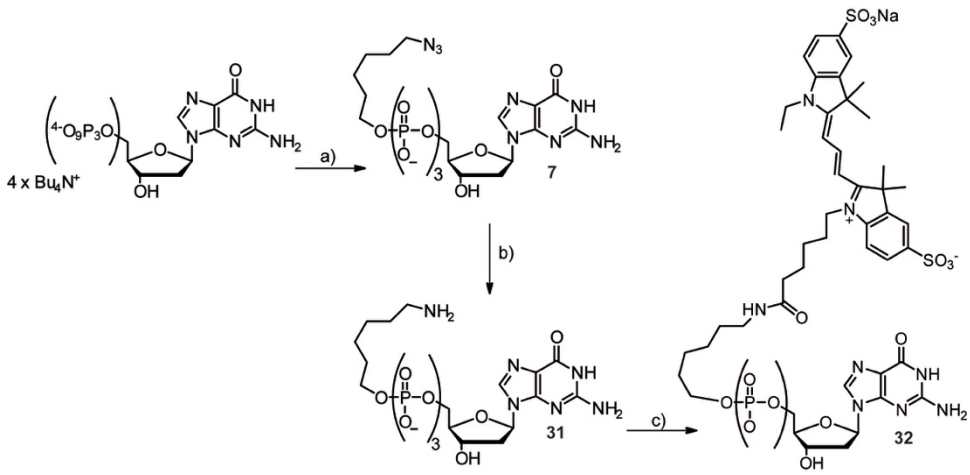


Figure 34: Synthesis of dye modified nucleotide **32**. a) 1-Azido-6-bromohexane, DMF, 16 h, rt; yield: 10 %; b) tris(2-carboxyethyl)phosphine, H_2O/ MeOH/ TEA 3/4/2,3 days, rt; yield: 77 %; **c)** sulfo-Cy3-NHS ester, 0.1 M $NaHCO_3$, 16 h, rt; yield: 77 %.

The unmodified dGTP as well as nucleotides **7** and **32** were further employed in DNA polymerase catalysed experiments. Comparative studies of processing those nucleotides opposite C or 5mC by the DNA polymerase KOD exo$^-$ and its mutant G245D show favoured incorporation of all nucleotides opposite C.

As already observed before, incorporation of unmodified dGMP results in higher discrimination if the DNA polymerase KOD exo$^-$ was mutated by replacing glycine in position 245 by aspartic acid (see **Figure 35**). Just as more efficient incorporation of nucleotide **7** by the DNA polymerase KOD exo$^-$ G245D mutant was shown before. The incorporation efficiency of nucleotide **32** by the KOD exo$^-$ mutant G245D decreases further in comparison to processing dGTP or nucleotide **7**, but discrimination is maintained (see **Figure 35c**).

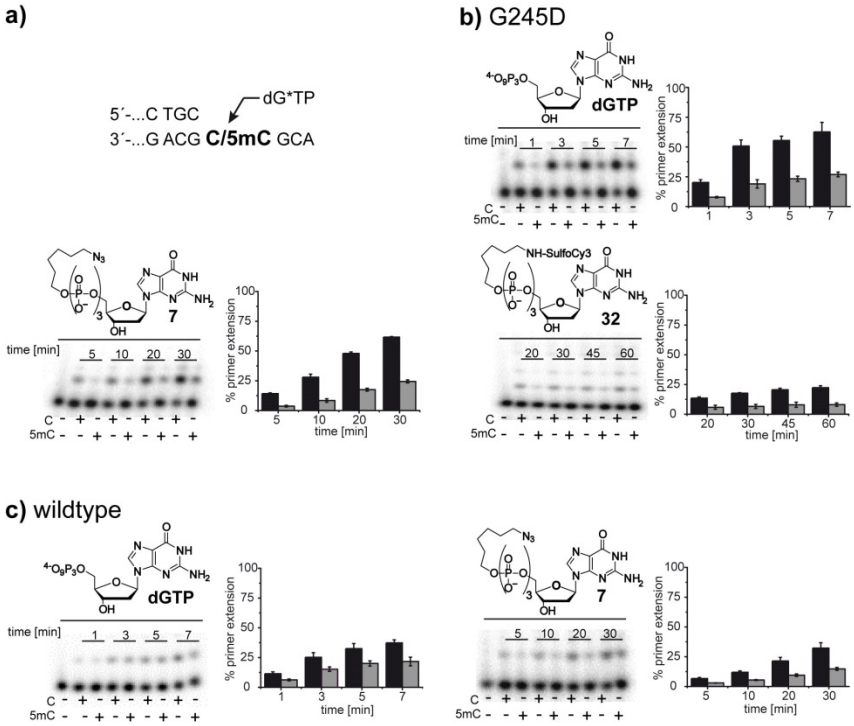


Figure 35: Primer extension experiments employing nucleotides dGTP, **7** and **32** and the DNA polymerase variants KOD exo$^-$ and G245D. **a)** Partial primer / template sequence used. **b)** PAGE analysis and quantitative evaluation of single-nucleotide incorporation primer extension experiments of **b**) dGMP and nucleotide **7** opposite a template containing C (black) in comparison to a template containing 5mC (grey) employing the DNA polymerase KOD exo$^-$; and **c)** of dGMP and nucleotides **7** and **32** employing the DNA polymerase KOD exo$^-$ mutant G245D. 50 µM dGTP or dG*TP and 2.5 nM of KOD exo$^-$ and 5 nM G245D was used for incorporation of dGMP; 20 nM for nucleotides **7** and **32**. Reactions were stopped after indicated time points. Experiments were done at least in triplicates.

To further prove those findings, steady-state kinetics were determined for processing dGTP and the fluorescently labelled nucleotide **32** (see **Table 2** and **Figure 36**) by the DNA polymerases KOD exo$^-$ and its mutant G245D. Results are listed in **Table 2**. Comparison of the catalytic efficiencies (k_{cat}/K_M) observed during processing of dGTP opposite C and 5mC confirms the tendencies observed in quantitative primer extension experiments (see **Figure 35**). Incorporation efficiencies decrease with usage of the KOD exo$^-$ G245D from 1.5 $s^{-1}\mu M^{-1}$ to 0.5 $s^{-1}\mu M^{-1}$. However, the ratio of the catalytic efficiencies observed during processing opposite C in comparison to processing opposite 5mC increases from 1.4 to a factor of 1.9 (see **Table 2**). If dGTP was replaced by nucleotide **32** a remarkable decline of the catalytic efficiency can be observed to a value of 0.007 $s^{-1}\mu M^{-1}$, demonstrating decreased incorporation efficiency. But again, the ratio between incorporation opposite C and 5mC increases to a factor of 2.4.

Table 2: Steady-state kinetic analysis of single-nucleotide incorporation of dGMP and modified nucleotide **32** opposite C or 5mC employing DNA polymerase KOD exo$^-$ wt or mutant G245D. The ratio was calculated by the quotient of k_{cat}/K_M opposite C and k_{cat}/K_M opposite 5mC.

polymerase	nucleotide	template	k_{cat} [s^{-1}] [a]	K_M [μM] [a]	k_{cat}/K_M [$s^{-1}\mu M^{-1}$] [a]	ratio
wt	**dGTP**	C	5.9 ± 0.1	4.0 ± 0.3	1.5 ± 0.1	1.36
		5mC	3.5 ± 0.1	3.3 ± 0.3	1.1 ± 0.1	
G245D	**dGTP**	C	10.9 ± 1.2	20.8 ± 6.2	0.53 ± 0.06	1.92
		5mC	2.7 ± 0.3	9.6 ± 1.9	0.276 ± 0.041	
G245D	**32**	C	2.42 ± 0.22	344.0 ± 49.6	0.0070 ± 0.0006	2.41
		5mC	0.585 ± 0.043	203.8 ± 32.2	0.0029 ± 0.0002	

[a] Data points derive from triplicates. ± describes SD.
Sequence primer: 5´-d(CGAAATGATCCCATCCAGCTGC)-3´

Having a closer look at the kinetic data, it can be seen, that discrimination between C and 5mC is mainly based on differences in k_{cat}. For usage of dGTP as well as nucleotide **32**, k_{cat} is higher for the incorporation opposite C than 5mC, indicating more efficient incorporation opposite the unmodified C, than opposite its modified counterpart.

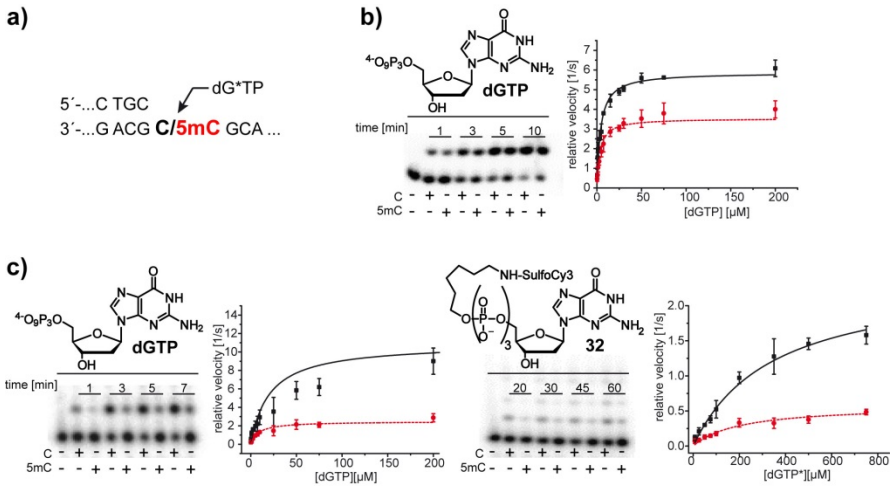


Figure 36: Kinetic evaluation of processing nucleotides dGTP and **32** by the DNA polymerases KOD exo$^-$ and G245D. **a)** Partial primer / template sequence used. **b)** left: chemical structure and PAGE analysis of single-nucleotide incorporation primer extension experiments of dGMP employing KOD exo$^-$. right: steady-state kinetics of single-nucleotide incorporation of dGMP opposite C (black solid line) or 5mC (red dashed line). **c)** as b) employing KOD exo$^-$ G245D with dGMP and **32**. Experiments were done at least in triplicates. Left: 50 μM dGTP or **32** and 5 nM polymerase were used for dGTP or 20 nM for **32**; reactions were stopped after indicated time points.

Nevertheless, K_M was as well higher for the incorporation opposite C than 5mC. That means, that the concentration of the employed nucleotide can be chosen in a way, that the effect of K_M can be overruled, thereby increasing the discrimination as it will be determined solely by k_{cat}. By this, the

discrimination for the fluorescently labelled nucleotide **32** and the KOD exo$^-$ mutant G245D can be increased to a remarkable factor of 4.1, calculated by the quotient of k_{cat} opposite C and 5mC.

3.7 Detection of 5hmC

5mC has generally been associated with gene repression playing an important role in development, aging and disease.[12, 121] It has been regarded as a stable, highly heritable mark and the mechanisms controlling its dynamics still remain elusive. The discovery of the 5mC dioxygenase ten-eleven translocation 1 (TET1), which was found to oxidise 5mC to 5-hydroxymethylcytosine (5hmC), shed new light on the dynamic nature of 5mC as 5hmC may be part of an active demethylation process.[14c] Moreover, 5hmC may not only be an intermediate in the active removal of 5mC. Recent studies suggest that 5hmC may act additionally as an epigenetic modification itself, thus adding another layer of complexity to the network of epigenetic regulation.[26a, 122]
To further understand the role of 5hmC in active demethylation or as epigenetic mark, new techniques for the detection of 5hmC are required. Therefore, we suggested employing our toolbox of variously modified nucleotides to the detection of 5hmC as well.

3.7.1 Discrimination of 5hmC Employing KlenTaq

Employing all modified nucleotides opposite C and both epigenetic markers 5mC and 5hmC with the DNA polymerase KlenTaq, no remarkable difference can be observed (see **Figure 37**). The same tendencies as observed before can be seen during this KlenTaq catalysed reactions. Incorporation efficiencies decrease remarkable during incorporation of modified nucleotides in comparison to processing the unmodified dGTP. Additionally, incorporation efficiencies decrease if nucleotides that possess sterically more demanding modifications are employed.

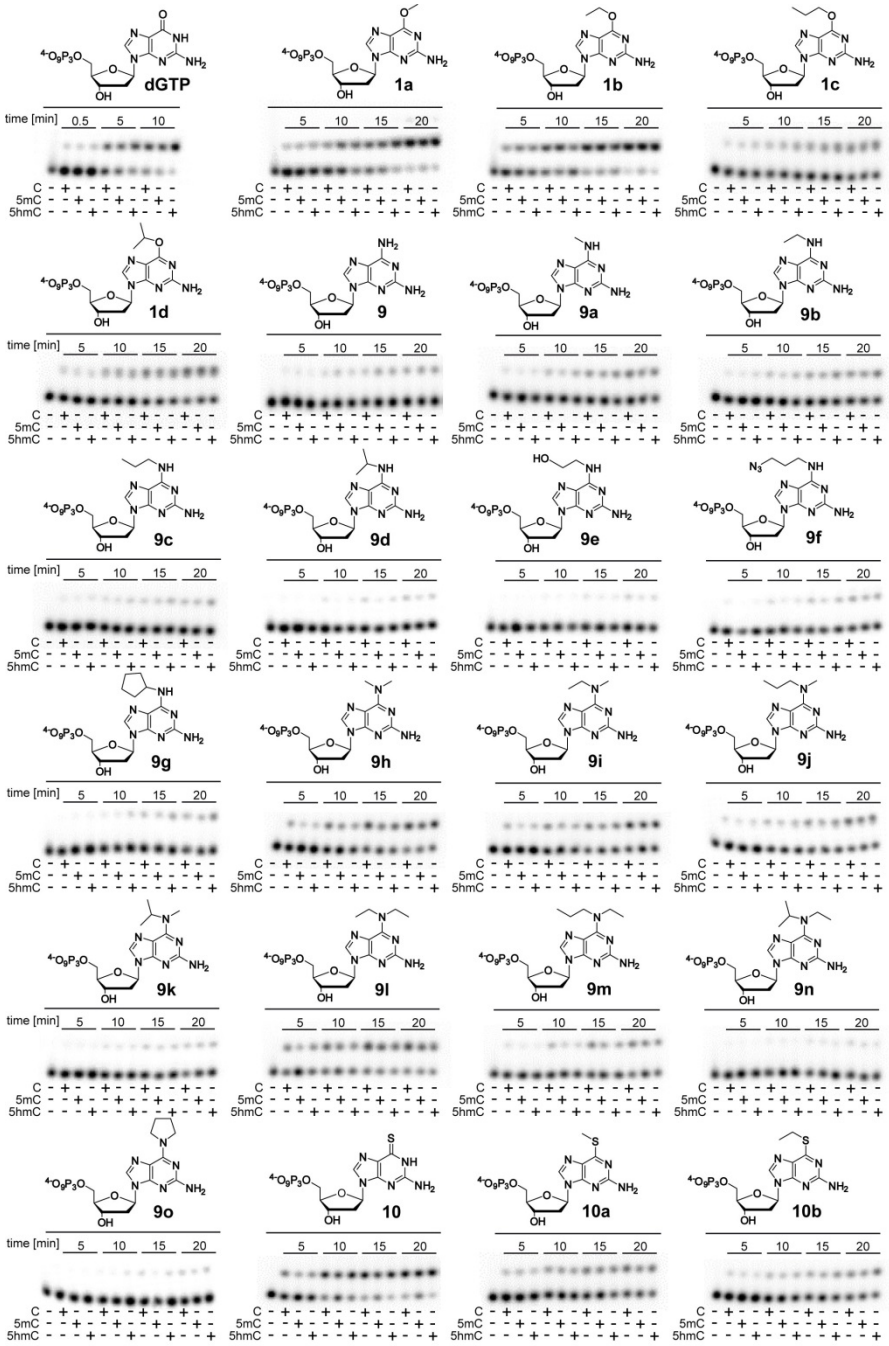


Figure 37: PAGE analysis of single-nucleotide incorporation primer extension experiments of dGMP and nucleotides **1a**-**d**, **9**, **9a**-**9o**, **10**, **10a**+**b** opposite a template containing C, 5mC or 5hmC employing DNA polymerase KlenTaq. 50 µM dGTP or dG*TP and 5 nM KlenTaq were used; reactions were stopped after indicated time points.

3.7.2 Discrimination of 5hmC Employing KOD exo$^-$

Again, employing the B-family DNA polymerase KOD exo$^-$ proved to show more promising trends in regard of discrimination ratios. Incorporation of dGMP or its modified analogues is more efficient opposite C as well as opposite 5mC than incorporation opposite 5hmC. Hence, already for incorporation of unmodified dGMP, favoured incorporation opposite C and 5mC can be observed in comparison to incorporation opposite 5hmC (see **Figure 38**). This tendency could be maintained for

usage of all modified nucleotides. Additional, diverging incorporation efficiencies between 5mC and 5hmC were detected as well.

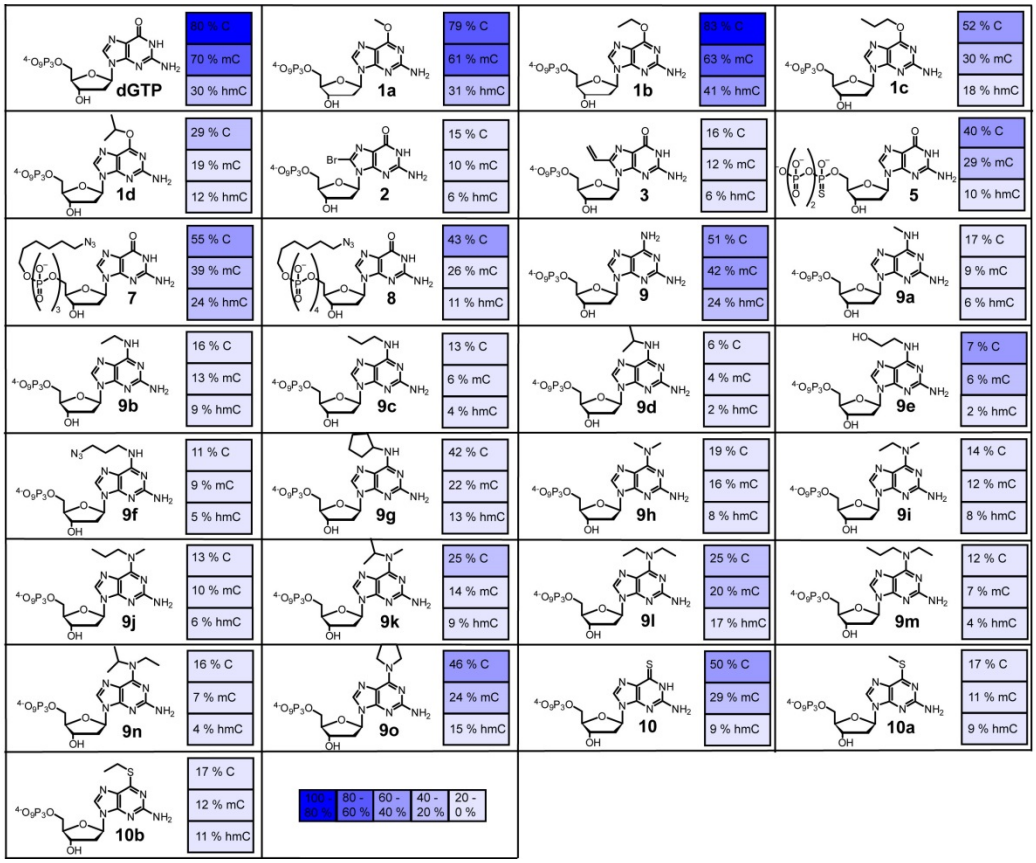


Figure 38: Structures of dGTP and modified dG*TP analogues **1** - **10** including % incorporation opposite C, 5mC or 5hmC employing DNA polymerase KOD exo$^-$ in single-nucleotide primer extension experiments. 50µM dGTP or dG*TP and 5 nM KOD exo$^-$ were used, reactions were stopped after 15 min. Experiments were done at least in triplicates. Arithmetic mean is given; errors are given in the appendix.

As summarised in **Figure 38**, different discrimination ratios can be observed between C and 5hmC, as well as between 5mC and 5hmC. Surprisingly, decent discrimination between C or 5mC and 5hmC could be detected for employing the unmodified dGTP with KOD exo$^-$. However, as shown before, little, but no sufficient discrimination can be found between C and 5mC. Employing the modified nucleotides with all three templates, increased discriminations can be observed between C and 5mC, as already shown in previous experiments. Discrimination between C and 5hmC is slightly enhanced as well. But, comparing incorporation efficiencies opposite 5mC with those opposite 5hmC, strong variations in discrimination behaviour can be observed (see **Figure 38** and **Figure 39**). The most promising nucleotides regarding detection of 5hmC are shown in **Figure 39**.

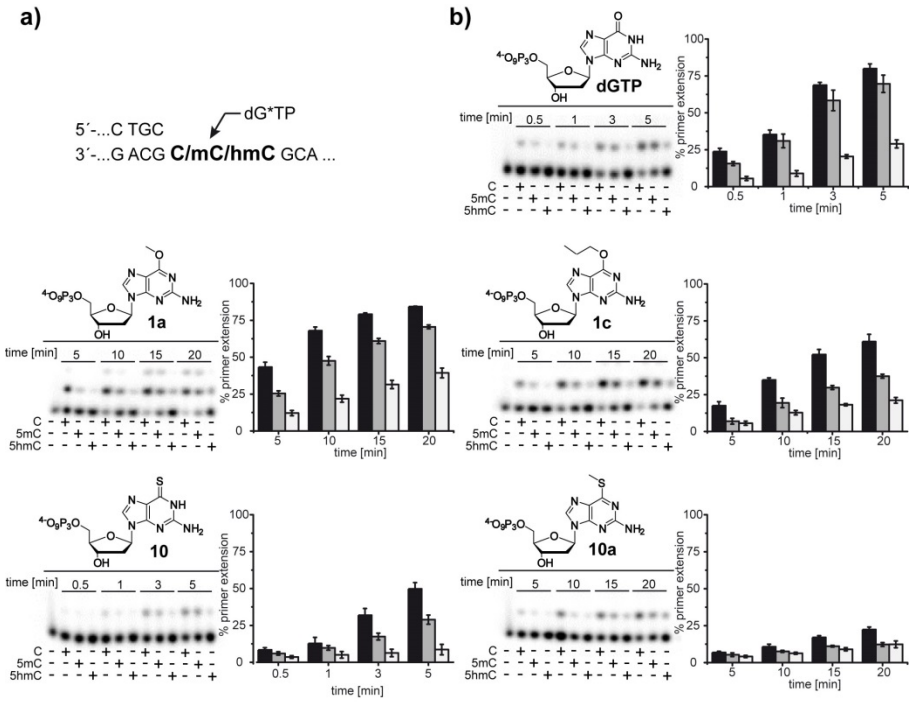


Figure 39: Incorporation studies of nucleotides dGTP, **1a**, **1c**, **10** and **10a** opposite C, 5mC and 5hmC employing KOD exo$^-$. **a)** Partial primer / template sequence used. **b)** PAGE analysis and quantitative evaluation of single-nucleotide incorporation primer extension experiments of dGMP and nucleotides **1a**, **1c**, **10** and **10a** opposite a template containing C (black) in comparison to a template containing 5mC (grey) or 5hmC (white) employing the DNA polymerase KOD exo$^-$. 50 µM dGTP or dG*TP and 5 nM KOD exo$^-$ was used. Reactions were stopped after indicated time points. Experiments were done at least in triplicates.

Some nucleotides show little, if any discrimination between 5mC and 5hmC, as observed for nucleotide **10a**. But notable discrimination can be observed between C and both epigenetic markers, allowing detection of epigenetic modifications in general. In contrast, application of the unmodified dGTP allows sensing of 5hmC, due to diverging incorporation efficiencies between 5hmC and C as well as 5mC.

Incorporation of other nucleotides, for example **1a**, **1c** and **10** shows remarkable discrimination between all three templates. Especially nucleotide **10** shows increased discrimination between C and 5hmC with a factor of 5.6.

3.7.3 Discrimination of 5hmC Employing 9 °North exo$^-$

Since remarkable discrimination was observed by employing all modified nucleotides in combination with the B family DNA polymerase KOD exo$^-$ opposite three different templates containing C, 5mC or 5hmC, it was suggested to investigate if those trends can be observed for employing the second B family DNA polymerase 9 °North exo$^-$ as well. As it can be seen in **Figure 40**, comparable trends as described for KOD exo$^-$ can be observed by employing 9 °North exo$^-$. Incorporation efficiencies decrease for processing all modified nucleotides, as those efficiencies decrease with increasing size and steric hindrance of the introduced modifications. Additionally, favoured incorporation opposite C in comparison to the incorporation opposite 5mC and 5hmC can be seen.

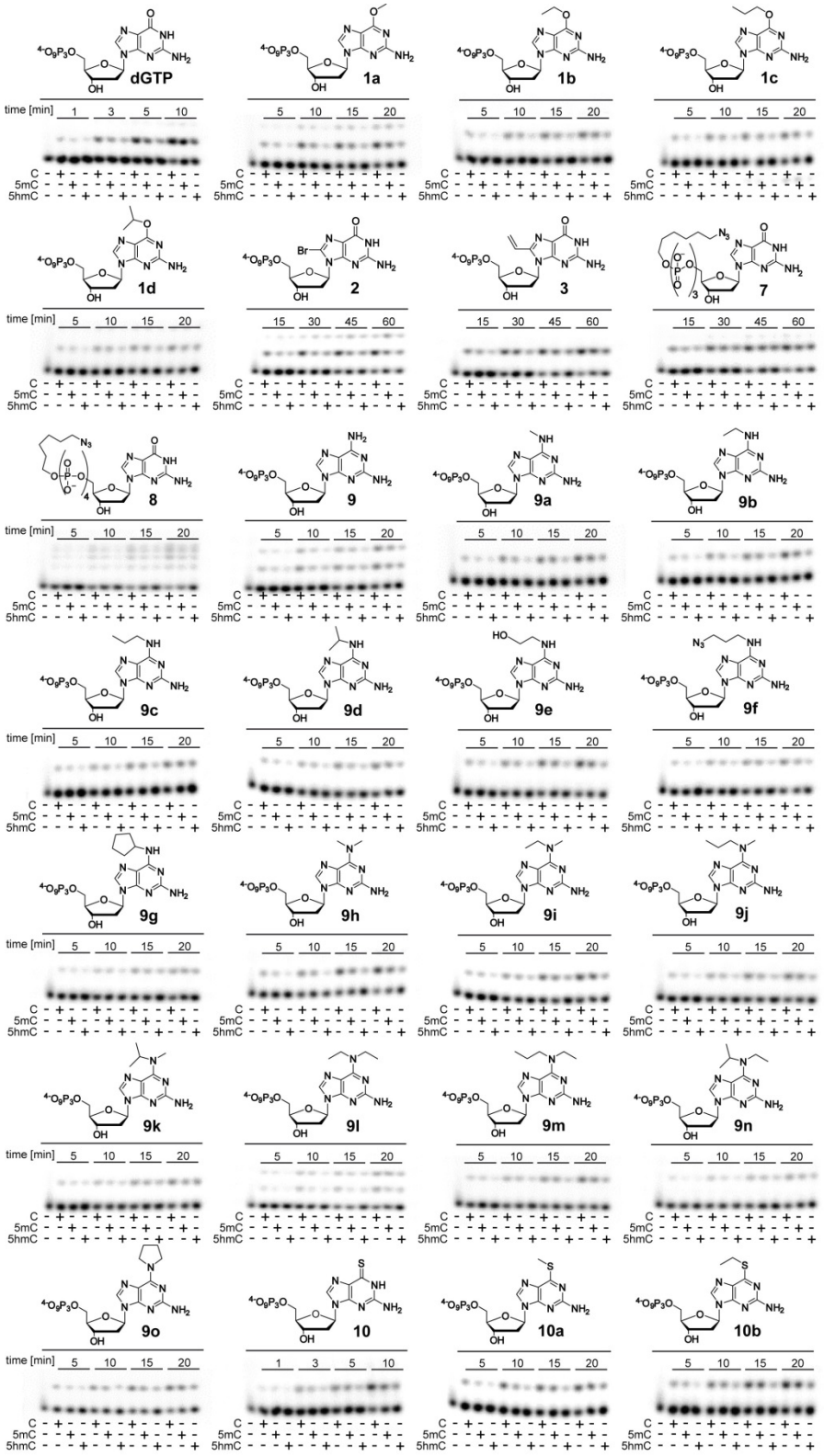


Figure 40: PAGE analysis of single-nucleotide incorporation primer extension experiments of dGMP and nucleotides **1a** - **d**, **2**, **3**, **7** - **9**, **9a** - **9o**, **10**, **10a** + **b** opposite a template containing C, 5mC or 5hmC employing DNA polymerase 9° North exo$^-$. 50 µM dGTP or dG*TP and 5 nM 9° North exo$^-$ were used; reactions were stopped after indicated time points.

Again, already usage of unmodified dGTP leads to considerable discrimination between C and 5hmC. This discrimination increases slightly, when modified nucleotides are employed. Incorporation efficiency of the employed nucleotides opposite 5mC can be placed between the respective efficiencies of processing opposite C and 5hmC.

All trends, observed for employing the DNA polymerase KOD exo$^-$ with the modified nucleotides opposite C, 5mC and 5hmC can be verified by the usage of 9 °North exo$^-$. Still, discrimination behaviour is not enhanced by the application of the DNA polymerase 9 °North exo$^-$.

3.8. Detection of RNA Modifications

The occurrence of modified nucleotides is not just known for DNA, but plays important roles in RNA as well. Cellular RNA contains over 150 chemically altered nucleotides, which are usually formed by enzymatic posttranscriptional modifications.[123] Those modified nucleotides play important roles in stabilisation of RNA structures, fine-tuning of numerous interactions between RNA itself or RNA-binding partners or protection against nucleolytic degradation.[45] As those modifications of RNA are typically of low abundance, they frequently go unnoticed by standard detection methods.[124] Therefore, new detection methods are required to detect different RNA modifications in a simple manner.

Nina Blatter established a KlenTaq mutant (RT-KTq2), which can extend DNA primers utilising a RNA template.[125] Having a whole tool box of variously modified nucleotides in hand, it was decided to employ those nucleotides in combination with this mentioned KlenTaq variant for the detection of RNA modifications, as I could already prove the potential offered by employing purine-based derivatives for the discrimination between modified and unmodified nucleotides.

3.8.1 Discrimination of 2´-*O*-Methyl-C by the Application of Modified Nucleotides

Methylation of the 2´-OH-group of the ribose moiety is one of the most abundant RNA modifications.[126] Even if its occurrence and association with heritable diseases and cancer are well studied, the detailed function of 2´-*O*-methylation in RNA is not well understood.[45, 127] Since new detection methods are required to further study the function of this posttranscriptional RNA modification, I screened the modified nucleotides for their ability for discrimination through diverging incorporation efficiencies opposite C and the modified nucleotide 2´-*O*Me-C by the DNA polymerase RT-KTq2.

Employing all modified nucleotides in combination with the KlenTaq mutant RT-KTq2 in single-nucleotide incorporation primer extension experiments showed, as already observed before, that the A family DNA polymerase is accepting modified nucleotides with remarkably decreased incorporation efficiencies compared to dGTP (see **Figure 41**). Even if the DNA polymerase concentration was increased to 100 nM, no processing of nucleotides **6** and **9a** - **9o** can be observed. But all of the other dGMP analogues can be successfully incorporated under the used conditions. In contrast to the above described studies, diverging incorporation efficiencies opposite C and its modified counterpart could be observed employing DNA polymerase RT-KTq2 (see **Figure 41** and **Figure 42**).

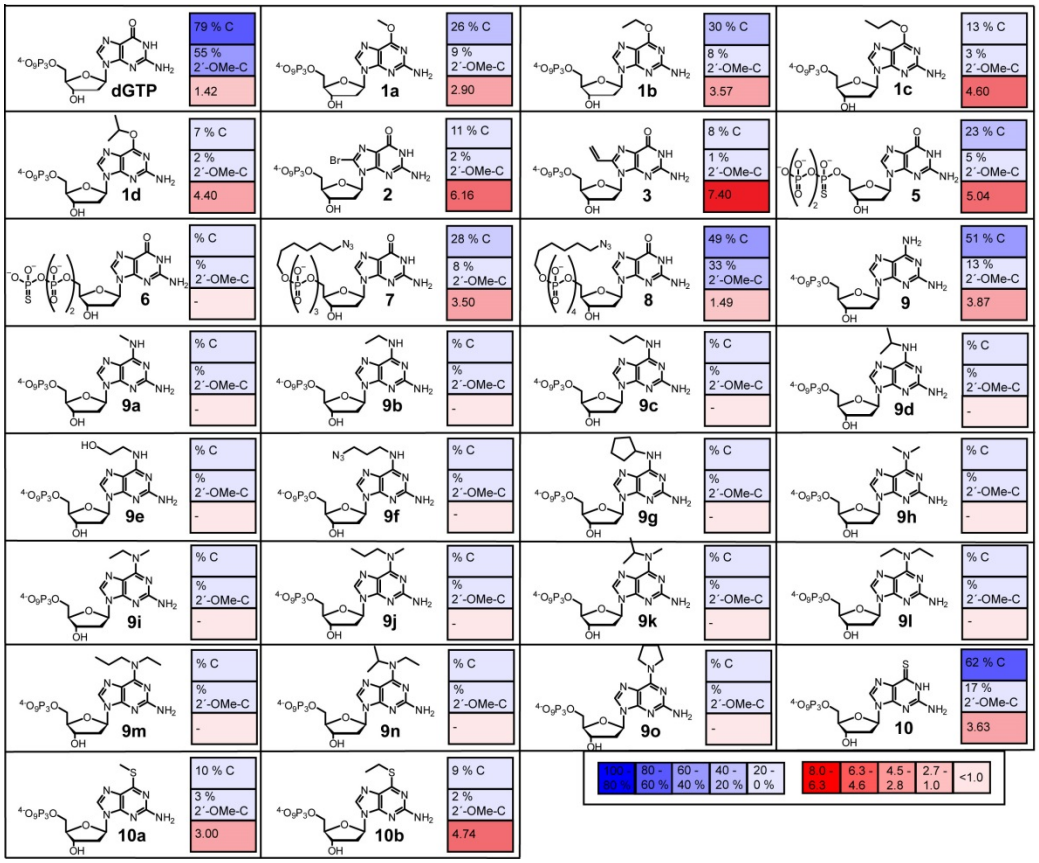


Figure 41: Structures of dGTP and modified dG*TP analogues including % incorporation opposite C or 2´-OMe-C employing DNA polymerase mutant RT-KTq2 in single-nucleotide primer extension experiments. 50µM dGTP or dG*TP and 20 nM RT-KTq2 were used, reactions were stopped after 15 min. Discrimination ratios were determined by calculating the quotient of % incorporation opposite C and % incorporation opposite 2´-OMe-C. Experiments were done at least in triplicates. Arithmetic mean is given; errors are given in the appendix.

Interestingly, already incorporation of unmodified dGMP is more efficient opposite C than opposite 2´-OMe-C resulting in a discrimination factor of 1.42. This discrimination can be remarkable enhanced by processing modified nucleotides. Incorporation of the O^6-alkylated dGMP derivatives improves discrimination further to a factor of 4.60 for nucleotide **1c**. Processing of the thio modified nucleotides **10**, **10a**+**10b** leads to even higher discrimination ratios up to 4.74 for nucleotide **10b** (see **Figure 41** + **Figure 42**). Incorporation of nucleotide **5** improves discrimination to a ratio of 5.04. But this discrimination can still be enhanced further, if the 8-modified nucleotides **2** and **3** are employed. Nucleotide **2** leads to a discrimination of 6.16 while nucleotide **3** shows a discrimination factor of 7.40. Again, all nucleotides are preferably processed opposite C. The most promising nucleotides for 2´-OMe-C detection are depicted in **Figure 42**.

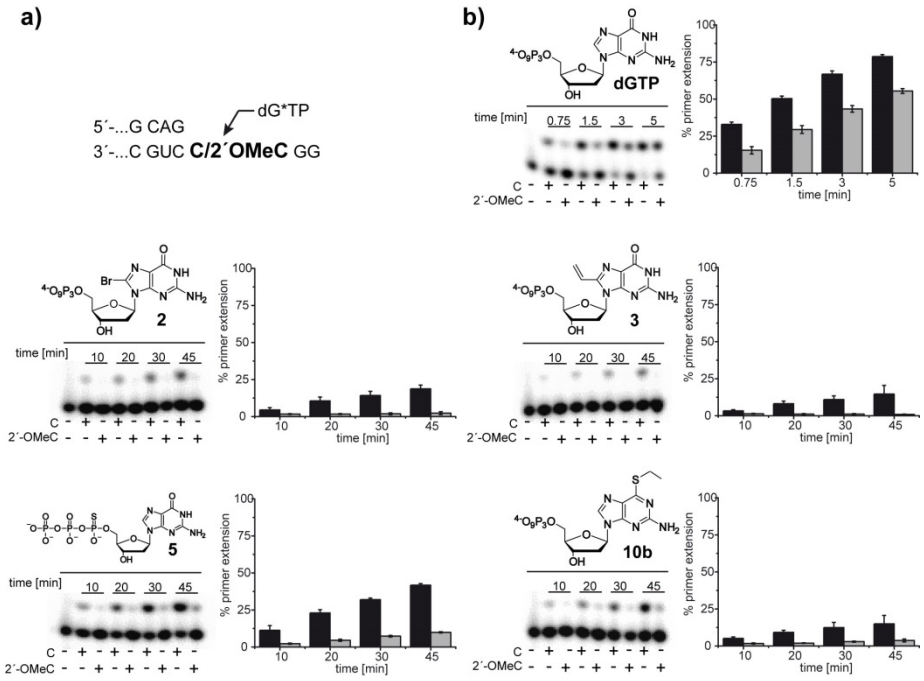


Figure 42: RT-KTq2-catalysed incorporation of dGTP and nucleotides **2**, **3**, **5** and **10b** opposite C and 2´-OMe-C. **a)** Partial primer / template sequence used. **b)** PAGE analysis and quantitative evaluation of single-nucleotide incorporation primer extension experiments of dGMP and nucleotides **2**, **3**, **5** and **10b** opposite a RNA template containing C (black) in comparison to a RNA template containing 2´-OMe-C (grey) employing the DNA polymerase RT-KTq2. 50 µM dGTP or dG*TP and 20 nM RT-KTq2 were used. Reactions were stopped after indicated time points. Experiments were done at least in triplicates.

3.8.2 Discrimination of Pseudouridine by the Application of Modified Nucleotides

Pseudouridine (Ψ) is the most abundant RNA modification.[50b] Despite its frequent occurrence in a wide range of cellular RNAs, the roles of single Ψ modifications have mainly stayed elusive.[50b] Nevertheless, many studies have suggested connections between RNA pseudouridylation and human diseases.[49] Robust and simple tools to detect the occurrence and distribution of Ψ are needed to evaluate global Ψ dynamics and changes to reveal the underlying mechanisms.

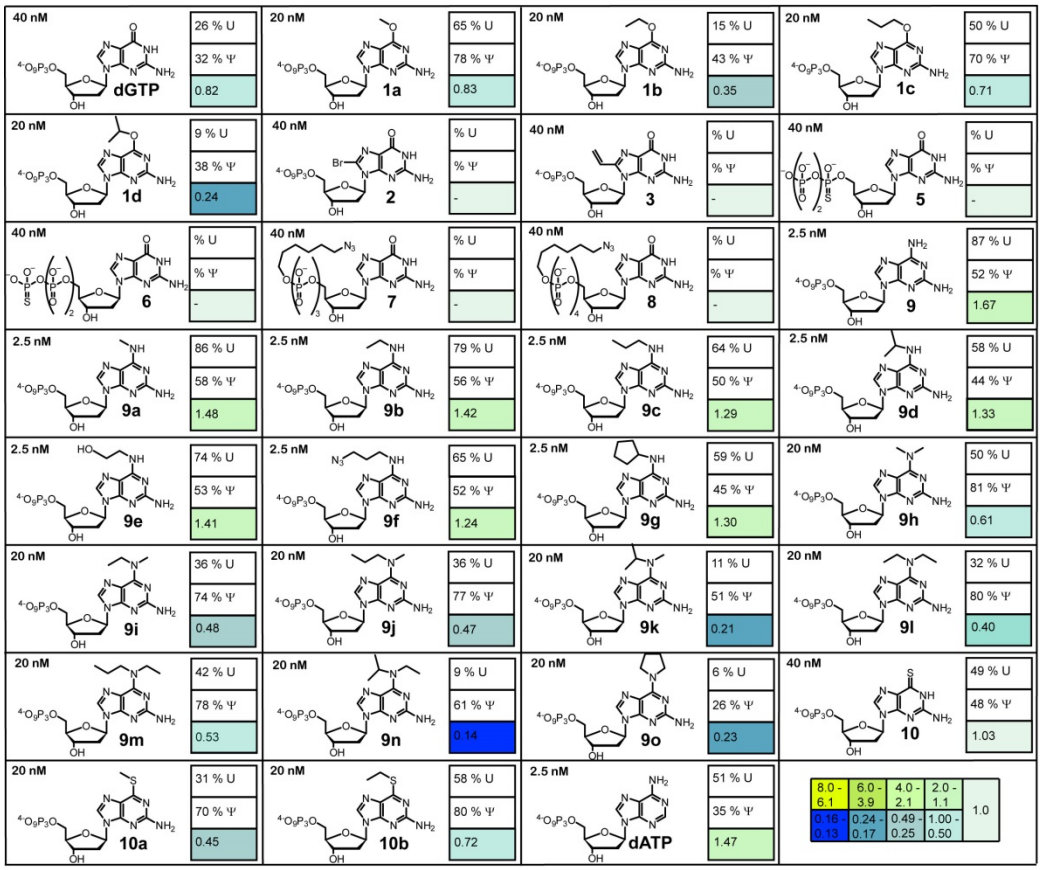


Figure 43: Structures of dGTP and modified dG*TP analogues including % incorporation opposite U or Pseudouridine (Ψ) employing DNA polymerase mutant RT-KTq2 in single-nucleotide primer extension experiments. 50µM dGTP or dG*TP were used, DNA polymerase concentrations are indicated in the left corner of each box, reactions were stopped after 15 min. Discrimination ratios were determined by calculating the quotient of % incorporation opposite U and % incorporation opposite Ψ. Experiments were done at least in triplicates. Arithmetic mean is given; errors are given in the appendix.

As already observed before, in position 6 modified purines are preferably incorporated opposite T (see **Figure 26**). Therefore, it was decided to employ incorporation of modified nucleotides by the DNA polymerase KlenTaq variant RT-KTq2 for discrimination between U and Ψ as well.

As expected, nucleotides possessing an altered Watson-Crick basepairing site (**1a** - **1d**, **9**, **9a** - **9o**, **10**, **10a** - **10b**), are incorporated with higher efficiencies as the unmodified dGTP or those nucleotides, modified at the 8-position (**2** + **3**) or the phosphate moiety (**5** - **8**) (see **Figure 43**). Incorporation efficiencies vary widely between the different modified nucleotides. Therefore, different DNA polymerase concentrations were employed for those incorporation experiments, as indicated in **Figure 43**. Still, even after increasing the RT-KTq2 concentration up to 100 nM, no considerable incorporation of nucleotides **2**, **3** and **5** - **8** can be observed. As expected, the most efficient incorporation can be seen for dAMP and nucleotides **9** and **9a** - **9d**. These findings can be explained, since those nucleotides resemble the structure of dATP, which is the natural counterpart of U.

Though, unexpectedly the incorporation efficiencies of the modified nucleotides **9** and **9a** - **9d** are not decreased in comparison to the unmodified dATP, as shown before for the A family DNA polymerase KlenTaq.

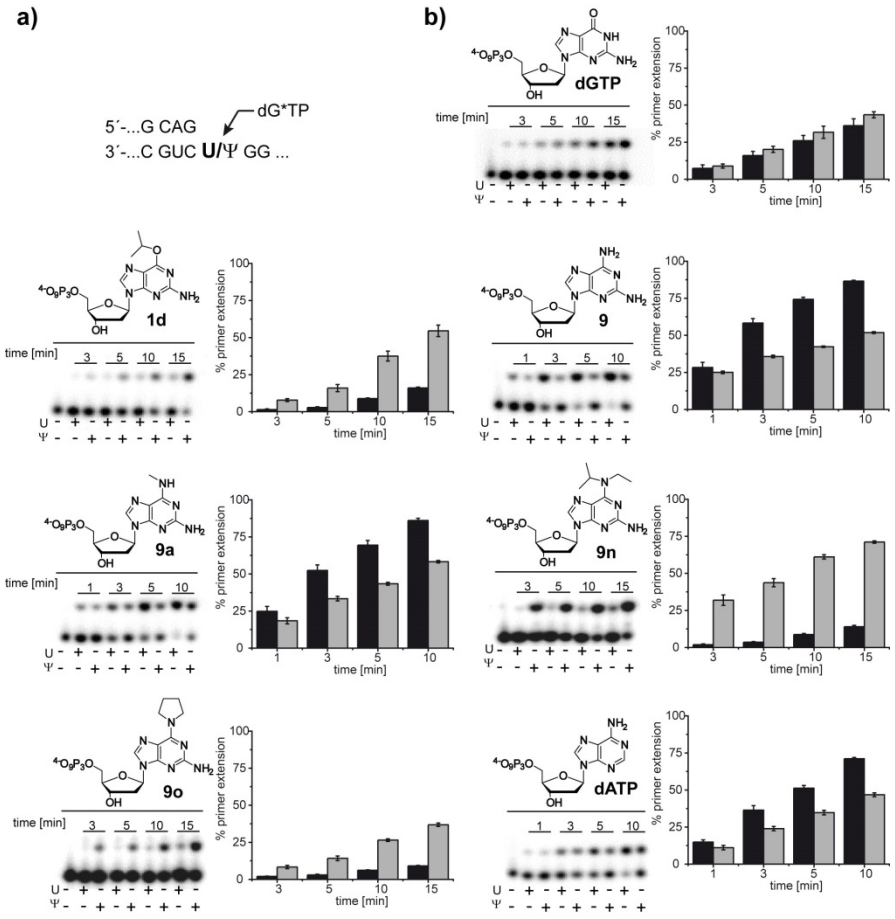


Figure 44: RT-KTq2-catalysed incorporation of dGTP and nucleotides **1d**, **9**, **9a**, **9n**, **9o** and **dATP** opposite U and Ψ. **a)** Partial primer / template sequence used. **b)** PAGE analysis and quantitative evaluation of single-nucleotide incorporation primer extension experiments of dGMP and nucleotides **1d**, **9**, **9a**, **9n**, **9o** and **dATP** opposite a RNA template containing U (black) in comparison to a RNA template containing Ψ (grey) employing the DNA polymerase mutant KlenTaq RT-KTq2. 50 µM dGTP or dG*TP and 40 nM (dGTP), 20 nM (**1d**, **9n**, **9o**) or 2.5 nM (**9**, **9a**, **dATP**) RT-KTq2 were used. Reactions were stopped after indicated time points. Experiments were done at least in triplicates.

Again, it can be observed that the efficiency in processing the modified nucleotides decreases with increasing size of the employed modification. For dATP and the modified nucleotides **9** and **9a** - **9d** discrimination with favoured incorporation opposite the unmodified U can be observed (discrimination ratio above 1.0). This discrimination ratio of around 1.5 during processing dATP cannot be enhanced by the usage of modified nucleotides (see **Figure 43** and **Figure 44**).

Increasing DNA polymerase concentration leads to considerable incorporation of dGMP and the modified nucleotides **1a** - **1d**, **9h** - **9o** and **10a** - **10b**. Surprisingly, usage of those nucleotides leads to a favoured incorporation opposite the RNA modification Ψ (discrimination ratio lower than 1.0). This is the first time discrimination was reversed and favoured processing opposite a modified nucleotide was observed during this studies. Additionally, this experiment shows remarkable discriminating effects between U and Ψ. Processing dGTP shows already distinct discrimination with a ratio of 0.82. This discrimination can be enhanced by O^6- alkylation of dGTP. Usage of **1a** shows comparable discrimination as achieved by employment of dGTP, but increasing the length of the alkyl chain to

ethyl, two fold higher discrimination (0.35 for **1b**) can be observed. Interestingly, this discrimination decreases for the incorporation of **1c** to increase dramatically for processing the *iso*propyl containing nucleotide **1d** to a factor of 0.25.

Employing the thio-modified analogues shows as well discrimination for both alkylated nucleotides **10a** + **10b** with preferred incorporation opposite Ψ.

The best discrimination can be observed by incorporation of the tertiary amines **9h** - **9o**. It can be seen that the discrimination ratio increases with usage of sterically more demanding modifications, as incorporation efficiencies decrease. A sequentially increased discrimination for processing nucleotides **9h** - **9k** from 0.61 to 0.21 can be observed as the sterically hindrance of one of the two alkyl chains, attached to the amino group in position 6, increases. The same tendency can be observed for nucleotides **9l** - **9o** (see **Figure 43** and **Figure 44**). The best discrimination can be detected for nucleotide **9n** with an amazing discrimination ratio of 0.14. This means, nucleotide **9n** is processed by the DNA polymerase mutant RT-KTq2 with almost 7 fold higher efficiency opposite Ψ than opposite C. Since it was shown before (see 3.8.1 Discrimination of 2´-*O*-Methyl-C), that nucleotide **9n** is not incorporated opposite C, this nucleotide provides the potential to be used for Ψ sequencing approaches in combination with RT-KTq2.

3.8.2.1. Kinetics for Incorporation of Modified Nucleotides

To further investigate those findings, steady-state kinetics[114] were determined for incorporation of the nucleotides dAMP and **9n** opposite U and Ψ (**Table 3**).

Table 3: Steady-state kinetic analysis of single-nucleotide incorporation of dAMP and modified nucleotide **9n** opposite U or Ψ employing DNA polymerase mutant RT-KTq2. The ratio was calculated by the quotient of k_{cat}/K_M opposite U and k_{cat}/K_M opposite Ψ.

nucleotide	template	k_{cat} [s^{-1}][a]	K_M [µM][a]	k_{cat}/K_M [$s^{-1}µM^{-1}$][a]	ratio
dATP	U	7.1 ± 0.3	7.2 ± 1.0	0.98 ± 0.18	1.849
	Ψ	4.1 ± 0.1	7.7 ± 0.8	0.53 ± 0.07	
9n	U	0.47 ± 0.02	74.8 ± 8.1	0.0064 ± 0.0009	0.067
	Ψ	5.84 ± 0.3	61.7 ± 9.3	0.095 ± 0.019	

[a] Data points derive from triplicates. ± describes SD.
Sequence primer: 5´-d(ACTACAAGCCCCAAAAGCAG)-3´

Comparison of the Michaelis Constant K_M shows increased concentrations for processing of **9n** in comparison to dATP. But, comparable results of K_M for incorporation of dAMP or **9n** opposite U in comparison to Ψ can be observed. Hence, the observed discriminating effects need to originate from differences in k_{cat}. Indeed, the catalytic constants k_{cat} for the conversion of both nucleotides opposite U and Ψ diverge. For usage of dATP, a higher k_{cat} value was observed for incorporation opposite U in comparison to the incorporation opposite Ψ. By employing nucleotide **9n** a reversed tendency can be seen.

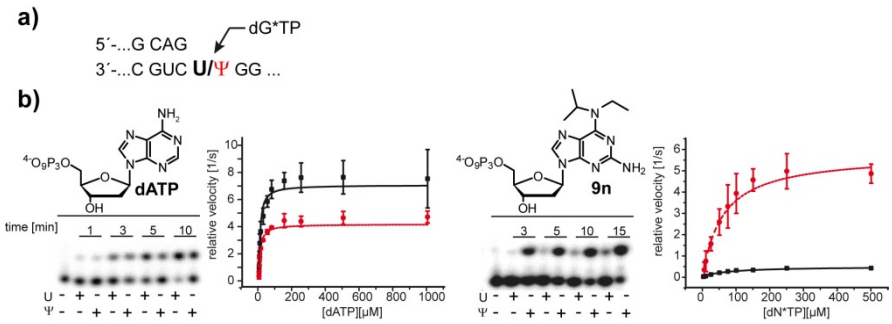


Figure 45: Kinetic evaluation of RT-KTq2-catalysed processing of dATP and **9n** opposite U and Ψ. **a)** Partial primer / template sequence used. **b)** left: chemical structure and PAGE analysis of single-nucleotide incorporation primer extension experiments of dAMP or **9n** employing DNA polymerase mutant RT-KTq2. 50 µM dGTP and 2.5 nM (dATP) or 20 nM (**9n**) DNA polymerase were used; reactions were stopped after indicated time points. right: steady-state kinetics of single-nucleotide incorporation of dAMP or **9n** opposite U (black solid line) or Ψ (red dashed line). Experiments were done at least in triplicates.

Comparison of the catalytic efficiencies (k_{cat}/K_M) observed for processing dATP and the modified nucleotide **9n** opposite U and Ψ in the template strand confirms all tendencies observed in the above described primer extension experiments. The incorporation efficiency for **9n** decreases considerably, as the catalytic efficiency decreases to 0.0064 $s^{-1}M^{-1}$, in comparison to 0.98 $s^{-1}M^{-1}$ observed for the usage of dATP (see **Table 3** and **Figure 45**).

However, the ratio of the catalytic efficiencies observed during processing the respective nucleotide opposite U in comparison to the incorporation opposite Ψ varies. Unmodified dATP is processed opposite U with 1.8 fold higher catalytic efficiency than opposite Ψ. Employing nucleotide **9n**, it can be seen that this nucleotide is incorporated opposite Ψ with 15 fold higher catalytic efficiency compared to the incorporation opposite U.

As those investigations verify all trends observed in the described primer extension reaction and prove remarkable discrimination between U and Ψ, if nucleotide **9n** is used, a novel assay for Ψ detection was suggested.

3.8.3. Ψ Detection Assay

To employ this very promising discrimination in a novel assay for Ψ detection, the approach depicted in **Figure 46** was suggested.

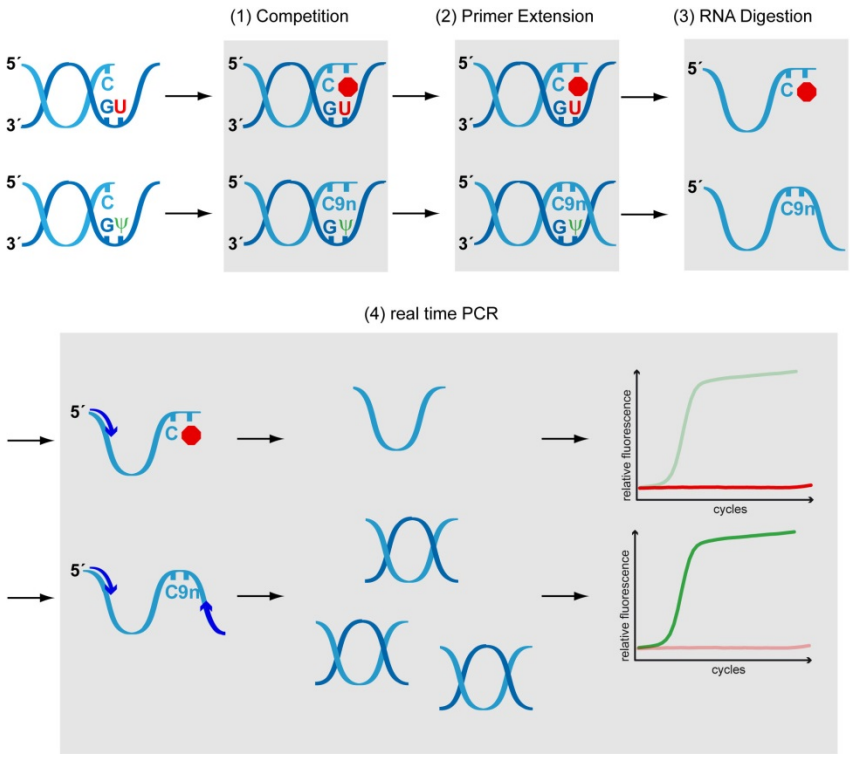


Figure 46: Suggested 4 step approach for positive Ψ detection. (1) Competitive incorporation of ddAMP and **9n** opposite U or Ψ leads to favoured incorporation of ddAMP opposite U and **9n** opposite Ψ. (2) Subsequent primer extension employing natural dNTPs leads solely to extension of primer opposite Ψ as primer paired opposite U was blocked by the extension by ddAMP. (3) RNA digestion removes all RNA template. (4) Extended DNA primer can be used as template for real time PCR.

As reverted discrimination was observed by more efficient incorporation of nucleotide **9n** opposite Ψ than opposite U, a positive four - step detection method should be established taking advantage of the favoured incorporation opposite Ψ. First, a primer and template construct was designed in a manner that the first incorporation will occur opposite U (red) or Ψ (green). For sequence design, different sequence contexts, derived from human 18S rRNA, which are known to contain Ψ-modifications, were chosen. The first step of the depicted approach (**competition**) consists of a competitive incorporation of nucleotide **9n** opposite Ψ and ddATP, which acts as chain terminator, opposite U. This incorporation step is followed by a second **primer extension** step. During this second step, a dNTP mix (dATP, dTTP, dCTP, dGTP) will be added to allow the DNA polymerase to extend all primer molecules, which are not blocked by the incorporation of ddAMP. Therefore, all primers paired opposite Ψ will be elongated, as nucleotide **9n** was incorporated in the first step. In case the primer was paired with a template containing U, no primer extension will be possible, as ddAMP was incorporated, lacking the 3´-OH-group, which is required to extend the respective primer. The third step involves **RNA digestion** of the analysed template. To make sure, subsequent real-time PCR is selective for the elongated primer, all of the RNA templates need to be degraded. The last step consists of a **real-time PCR**. PCR primers were designed in a way that one primer will be complementary to the extended region of step 2. This set up will ensure that real-time PCR can only result in exponential amplification, if the respective primer was extended during the second step. If the used template contained the unmodified U, no primer extension occurred and real-time PCR can only

result in linear amplification, as only one primer can bind to the non-elongated primer, which serves as template during this last step. That will result in a curve in real-time PCR, which is characteristic for exponential amplification if incorporation during the first step occurred opposite Ψ (see **Figure 46**, green curve). If the template contains the unmodified U instead of Ψ, no such curve will be observed (see **Figure 46**, red curve). Thus, I established a novel approach for a positive detection method, as an amplification curve during real-time PCR only can be seen, if Ψ was present in the template.

3.8.3.1. RT-KTq2 F667Y for Detection of Pseudouridine

As pointed out before, ddAMP needs to be incorporated opposite U to perform the depicted assay. The A-family DNA polymerase KlenTaq is known to lack this ability.[128] As reported 1995 by Tabor *et al.* substitution of tyrosine for phenylalanine at residue 667 of the Taq DNA polymerase dramatically increases the ability of this DNA polymerase to incorporate ddNMP.[128] Therefore, it was decided to introduce this mutation in the KlenTaq mutant RT-KTq2 to enhance its ability to incorporate ddAMP. Those experiments were performed by Konrad Bergen.

Having the mutant RT-KTq2 F667Y in hand, discrimination behaviour of this mutated enzyme between C and Ψ was tested again. Therefore, nucleotides dATP, ddATP and **9n** were employed in single-incorporation primer extension experiments opposite Ψ and its unmodified counterpart U (see **Figure 47**).

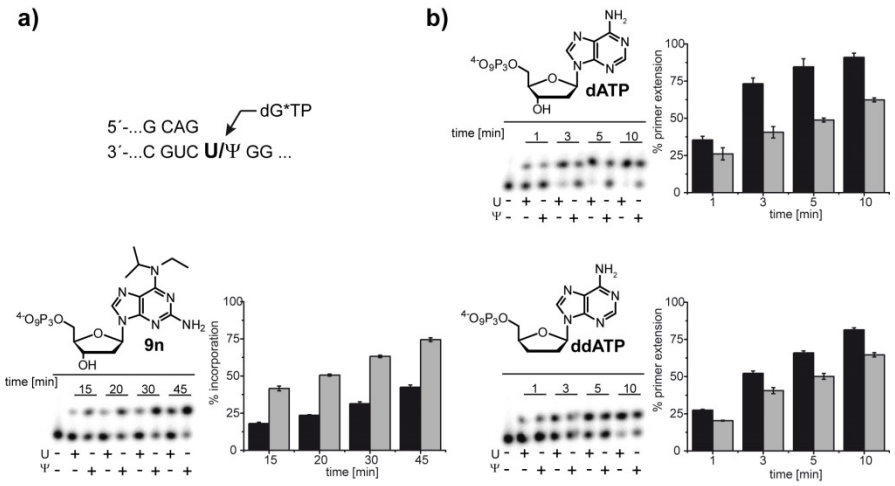


Figure 47: Processing of dATP, **9n** and ddATP by the KlenTaq variant RT-KTq2 F667Y opposite U and Ψ. **a)** Partial primer / template sequence used. **b)** PAGE analysis and quantitative evaluation of single-nucleotide incorporation primer extension experiments of dAMP, **9n** and ddAMP opposite a RNA template containing U (black) in comparison to a RNA template containing Ψ (grey) employing the DNA polymerase variant KlenTaq RT-KTq2 F667Y. 50 µM dN*TP and 2.5 nM (dATP, ddATP) or 50 nM (**9n**) RT-KTq2 F667Y were used. Reactions were stopped after indicated time points. Experiments were done at least in triplicates.

For processing nucleotide dATP and **9n** the same tendencies can be observed as before. While dATP shows favoured processing opposite U, **9n** is preferably incorporated opposite Ψ. Albeit those tendencies are maintained, the discrimination ratio for processing nucleotide **9n** decreases remarkably.

Due to its exponential amplification, real-time PCR needs only very little template for successful detection. Therefore, I investigated if, despite reduced discrimination, it is still possible to find conditions that can be used for the above described approach. Thus, competition experiments were performed (see **Figure 48**). A mix from two different nucleotides - in this case ddATP and **9n** - is employed for single-nucleotide incorporation experiments with the respective DNA polymerase. Due to the different sizes of the used nucleotides, the extended primers will show diverging migration behaviour during separation by PAGE allowing the analysis of which nucleotide is incorporated to which extend. Different ratios between ddATP and **9n** were chosen for further investigations. As shown in **Figure 48** both nucleotides are incorporated to the same extent opposite U, if a ratio of 1/693 (ddATP/**9n**) is used. Equal incorporation opposite Ψ can be achieved by the usage of a ratio of 1/125. This data suggest that even if discrimination is reduced due to the introduced mutation, discrimination between U and Ψ can still be exploited for the depicted assay.

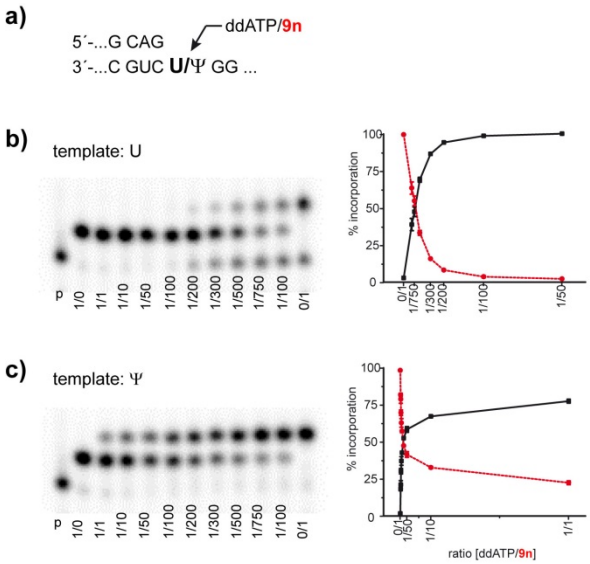


Figure 48: Competition experiments. **a)** Partial primer / template sequence used. **b)** + **c)** PAGE analysis and quantitative evaluation of single-nucleotide incorporation primer extension experiments of mixtures of ddATP (black solid line) and **9n** (red dashed line) opposite a RNA template containing U (b) or Ψ (c). Ratios of ddATP and **9n** as indicated. Reactions were stopped after 60 min. 50 µM of dN*TP mixture and 20 nM DNA polymerase were used. Experiments were done at least in triplicates.

3.8.3.2. Real-Time PCR Experiments

Initial experiments were performed employing artificial RNA templates containing U or Ψ. Primer and the respective template were annealed in 1 x reaction buffer. To ensure, that the envisioned approach is working, initial control experiments were performed. The first step (**competition**) was therefore performed by incubating the DNA polymerase RT-KTq2 F667Y with the annealed primer/template complex and either ddATP or **9n**. Therefore, the respective reactions can serve as positive (**9n** incorporation opposite Ψ) or negative (ddAMP incorporation opposite U) controls. After initial primer extension for 1 h at 55 °C, a mix of all natural dNTPs was added for primer extension (step 2: **primer extension**). Thus, primer paired with the template containing U will not be extended,

while the primer which is paired opposite Ψ will be elongated. The analysed RNA template needs to be hydrolysed in the next step to ensure that the subsequent PCR reaction is selective for the extended DNA primer.

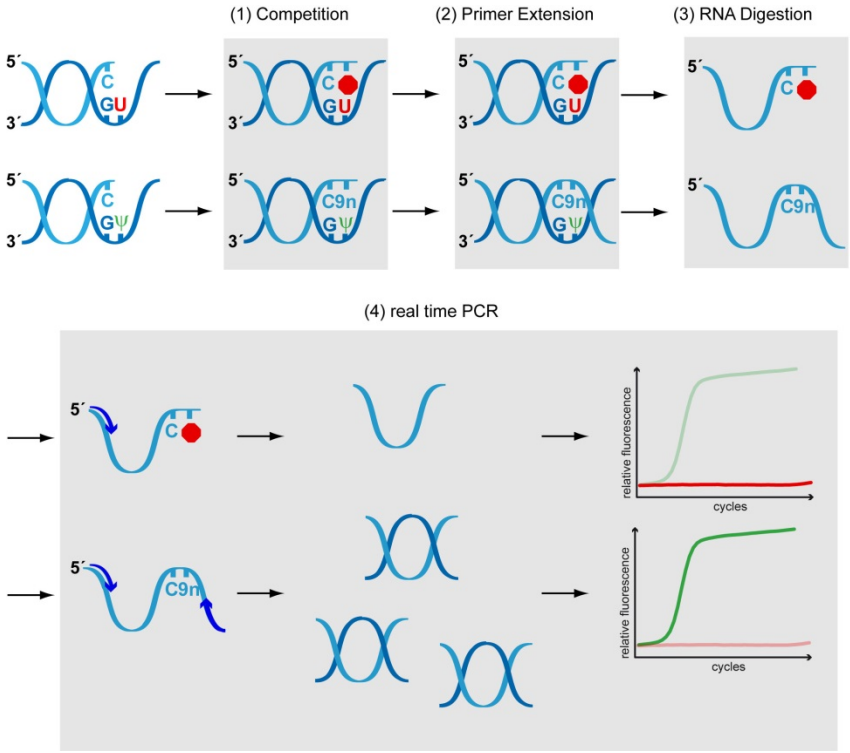


Figure 49: Suggested 4 step approach for positive Ψ detection. (1) Competitive incorporation of ddAMP and **9n** opposite U or Ψ leads to favoured incorporation of ddAMP opposite U and **9n** opposite Ψ. (2) Subsequent primer extension employing natural dNTPs leads solely to extension of primer opposite Ψ as primer paired opposite U was blocked by the extension by ddAMP. (3) RNA digestion removes all RNA template. (4) Extended DNA primer can be used as template for real time PCR.

Several possibilities are known to digest RNA templates, as required in step 3 (**RNA digestion**). Since additional purification steps should be avoided to simplify Ψ-detection, it was decided to use two RNase enzymes for RNA digestion: RNase H and RNase I_f. RNase H will degrade only the RNA in RNA:DNA hybrids,[139] while RNase I_f is an RNA endonuclease, which will cleave at RNA dinucleotide bonds with a clear preference for single-stranded RNA over double-stranded RNA.[140] The advantage of those enzymes lies in their selectivity for RNA. No DNA oligomers are degraded whereby the reactions do not require purification. Another common method for RNA degradation is its treatment with bases.[141] Treatment with a NaOH solution will lead to RNA hydrolysis. Under basic conditions, the 2´-OH-group can act as nucleophile attacking the adjacent phosphorous in the phosphodiester bond of the sugar-phosphate backbone of the RNA.[142] Even if this approach offers a very simple and easy way for RNA degradation, the reactions need to be purified by size exclusion. This additional purification step would be necessary to adjust the pH of the reactions for the subsequent amplification step. Since additional purification steps will add additional layers of complexity to our assay, we decided to degrade the RNA template in an enzymatic manner.

After RNA digestion, the extended primer derived from the previous steps is used as template for the subsequent **real-time PCR** analysis. Real-time PCR is performed according to standard procedures.[143] Sybr Green is added as fluorescent dye and an aptamer is added to block the DNA polymerase.[144] This aptamer melts during the first heating step in real-time PCR whereby the DNA polymerase regains its activity.[144]

Curves derived from real-time PCR analysis are depicted in **Figure 50**. Unfortunately, analysis of the described control experiments did not lead to delayed amplification, if the template containing U was employed in combination with ddATP (see **Figure 50 a**). Water was used instead of template as negative control for real-time PCR (H_2O). Additional negative control was performed by the addition of water instead of dNTPs in steps 1 and 2 of this assay (neg). Equally, a positive control was performed, if natural dNTPs were added during steps 1 and 2 (pos) instead of addition of modified nucleotides during step 1. Both negative controls (H_2O and neg) show strongly delayed amplification curves, and no product band can be detected on the agarose gel (amplification stopped after 20 cycles). Positive control (pos) shows amplification occurring around 10 cycles and a nice band on the corresponding agarose gel. Therefore, it can be concluded that real-time PCR is working perfectly. Since the second negative control (U), derived from incubation of the template containing U with ddATP during the first step, does not show delayed amplification in comparison to the second positive control (Ψ), it was assumed, that the first step of this assay needs improvement. Analysis via agarose gel shows distinct bands for both control experiments as well. Intensity of those bands is comparable to the band derived from the positive control (pos).

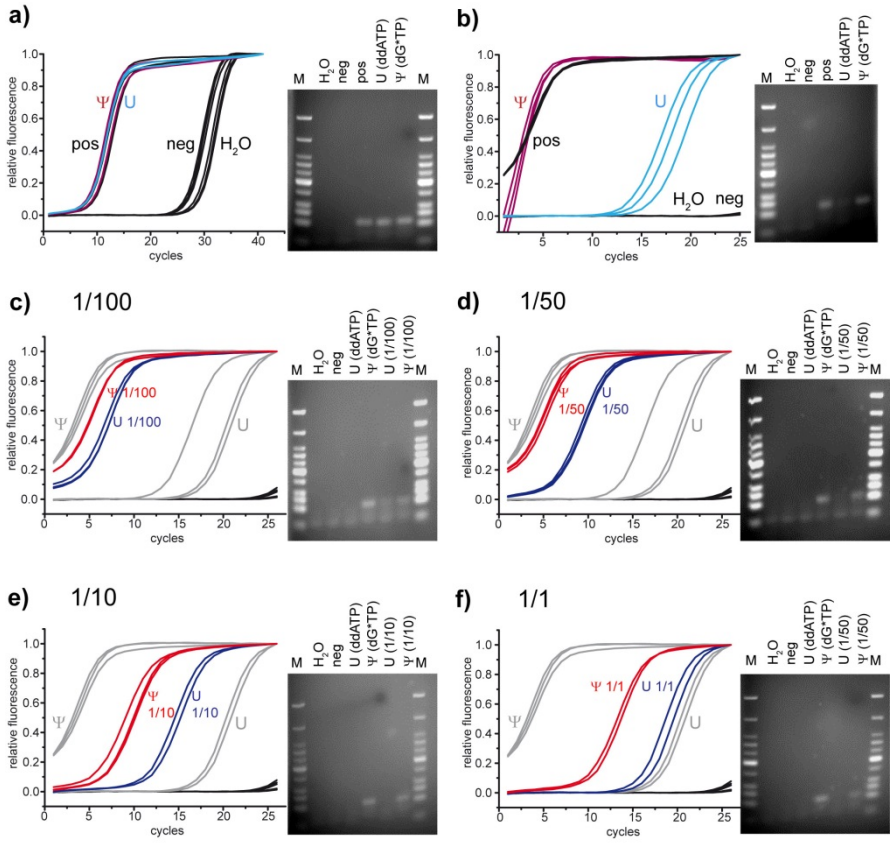


Figure 50: RT-PCR experiments employing RT-KTq2 F667Y in the described assay; left: real-time PCR analysis, right: agarose gel analysis. **a)** Reaction was performed without inorganic pyrophosphatase **b)** reaction including inorganic pyrophosphatase **c)** - **f)** different ratios between ddATP/**9n** were employed during **competition** step.

To improve discrimination between positive (Ψ) and negative control (U) addition of a thermostable inorganic phyrophosphatase was suggested. This enzyme will hydrolyse all pyrophosphate molecules derived from incorporation of ddAMP. By degradation of free pyrophosphate it should be ensured that the incorporated ddAMP or **9n** cannot be removed again by the DNA polymerase. Otherwise the assay was performed as described, including all described control experiments. After including pyrophosphatase in the first and second steps, real-time PCR analysis results in remarkably delayed amplification, if the primer was blocked by incorporation of ddATP opposite U (U) (see **Figure 50b**). By this approach a difference in C_t values of 17.3 ± 2.7 cycles can be detected between both templates. This difference is even visible, if amplification was stopped after 10 cycles and reactions were analysed on agarose gels (see **Figure 50b** right).

To be able to distinguish between U and Ψ, both templates were employed with a mixture of ddATP and **9n** in step 1 (**competition**). Previous competition experiments (see **Figure 48**) indicated that no remarkable incorporation of nucleotide **9n** opposite U occurs if a ratio of 1/100 (ddATP/**9n**) was employed. Hence, different ratios of ddATP and **9n** were applied in the **competition** step. If a ratio of 1/100 was applied during the first step, real-time PCR resulted in curves that differ 2.1 ± 0.3 cycles (see **Figure 50c**). Usage of a ratio of 1/50 improves that difference to a ΔCt = 5.0 ± 0.6 (see **Figure 50d**). This difference is enhanced further to ΔCt = 7.1 ± 2.7 by the usage of a 1/10 ratio (see **Figure 50e**). The best discrimination of ΔCt = 8.2 ± 2.9 is detected, if a ratio of ddATP and nucleotide **9n** of

1/1 was employed. In that case, no difference can be detected, whether the template containing U was incubated during step 1 solely with ddATP (U) or with a 1/1 mixture of ddATP and **9n** (U 1/1). That observation suggests that during the **competition** step only ddATP will be incorporated opposite U and further adjustment of the ratio between both nucleotides will not improve the discrimination. The corresponding agarose gels verify the results obtained during real-time PCR. Reactions were stopped between 4 and 20 cycles and analysed via agarose gels.

3.8.3.3. Experiments employing RNA from extracts

I could show that the envisioned approach can be employed for Ψ - detection, as about eight cylces delayed amplification can be detected during real-time PCR analysis, if an artificial template containing U was employed with DNA polymerase RT-KTq2 F667Y and an 1/1 ratio between ddATP and nucleotide **9n**. To verify, that this approach can even be applied to RNA extracts, radioactively labelled single-nucleotide incorporation primer extension reactions were performed. As mentioned above, different sequence contexts derived from human 18S rRNA were chosen, which are known to contain the Ψ modification. *In vitro* transcribed RNA was generated as unmodified control. A second control was established as additional sequence contexts were chosen which are known to contain the unmodified C. Thereby it was ensured, that no differences in concentration between the RNA extract and the *in vitro* transcribed control RNA resulted in delayed amplification during real-time PCR.

Unfortunately, primer extension reactions show that incorporation of the modified nucleotide **9n** on an RNA extract is difficult (see **Figure 51a**). Employing the DNA polymerase RT-KTq2 and its mutant RT-KTq2 F667Y, incorporation of dAMP or ddAMP respectively can be observed. In contrast, processing nucleotide **9n** by both enzymes does not lead to considerable incorporation by any enzyme under the chosen conditions. Nucleotide **9n** bears quite big and steric demanding modifications, therefore incorporation was seen before to be hindered as incorporation efficiencies decrease remarkably in comparison to the unmodified dAMP. The concentration of both enzymes was increased from 100 nM to 500 nM final concentration to improve the observed incorporation (see **Figure 51b**). Incorporation efficiencies of dAMP/ddAMP are enhanced by both DNA polymerases. But incorporation of nucleotide **9n** can only be observed by the mutant RT-KTq2 F667Y. As this variant is needed for the described approach, I focused on RT-KTq2 F667Y for further primer extension experiments.

As adequate incorporation of nucleotide **9n** was observed for artificial templates, the problem was suspected to be the rather big RNA molecule or its complex secondary structure. Hence, I focused on methods to break the RNA into smaller pieces to facilitate incorporation. To make sure, no sequence bias result in those difficulties in incorporation, two different sequence contexts were investigated in parallel (see **Figure 51b** and **c**). As described before, different methods are known to degrade RNA, as the RNA molecule is rather instable. The difficulty is to achieve only partial hydrolysis of the RNA. Enzymes, which are known, to break RNA will degrade the oligomer mostly to very small pieces - some even to the monomer. Therefore, it was decided to employ five different methods to facilitate incorporation: 1: mechanical degradation; 2: basic cleavage; 3: additional oligonucleotide, which is designed to pair in an adjacent loop and therefore supposed to melt the secondary RNA structure; 4:

enzymatic degradation by usage of a fragmentase - reaction is stopped by addition of EDTA; 5: enzymatic degradation by fragmentase - stopped through heat denaturation.

1: Mechanical degradation was performed using a nebuliser kit from Invitrogen. The reaction mixture was diluted and the reactions were performed in bigger scale to allow application of this tool. Additionally, handling does not allow the analysis of several reactions in row, since mechanical degradation cannot be parallelised.

2: As discussed above, RNA can be degraded under basic conditions. The length of the resulting RNA fragments is regulated by incubation time as the reaction can be stopped, by pH adjustment. Even if this approach is rather simple and easy to handle, the treated reaction mixtures need purification, since otherwise high salt content hinders DNA polymerase activity. Again, purification via size exclusion is tedious and adds an additional layer of complexity to our approach.

3: This approach does not intent to break the RNA itself, but to perturb its complex secondary structure. If the RNA in the investigated region was single stranded, DNA polymerase catalysed incorporation of ddAMP or **9n** will be facilitated. Therefore, I designed an oligomer, which is supposed to hybridise adjacent regions. Since the oligomer was designed in a manner that the corresponding melting point was higher as for the respective secondary RNA structure, this hybridisation was supposed to unwind the RNA and therefore facilitate primer binding.

4: Reaction handling of enzymatic degradation by usage of ds fragmentase (NEB) was simple and easy. Many reactions can be performed in parallel. The enzyme reaction consists out of two enzymes and needs to be stopped after a defined time point, to ensure that the RNA is not degraded to a high extend. The reaction is stopped by complexation of magnesium ions upon addition of 0.5 M EDTA solution. To enable activity of the DNA polymerase in the subsequent PCR reaction, the reaction mixture needs to be heated to denature the employed fragmenting enzymes and afterwards purified by size exclusion via spin columns. Thereby, EDTA gets removed and PCR buffer can be added for the following PCR reaction.

5: Purification of every reaction mixture is not convenient, as it is time consuming and does not allow parallelisation, making this assay more complicated. Thus, the fragmentase reaction should be stopped by heat denaturation. No purification step will be required, thereby simplifying reaction handling.

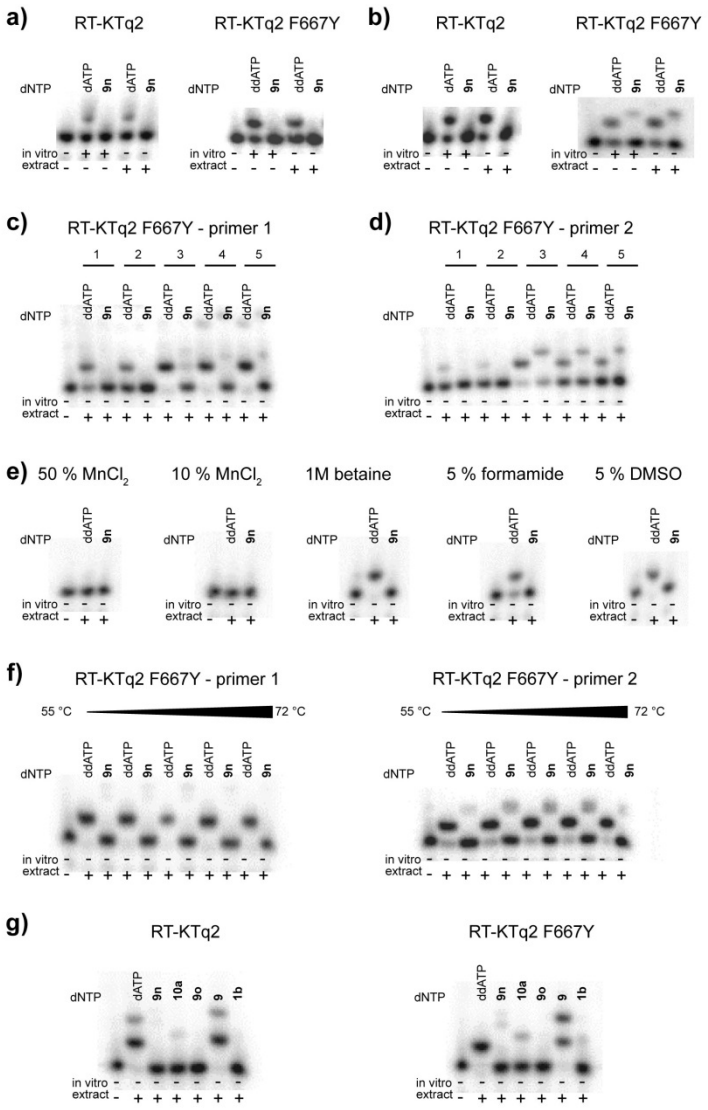


Figure 51: PAGE analysis of single-nucleotide incorporation primer extension experiments employing RNA extracts or *in vitro* transcribed RNA; **a)** single-nucleotide primer extension experiments of dAMP or ddAMP and **9n** opposite U (in vitro transcribed RNA) or Ψ (RNA extract) employing 100 nM RT-KTq2 and RT-KTq2 F667Y **b)** as a) employing 500 nM enzyme; **c)** as b) employing RT-KTq2 F667Y and primer 1 but template was treated prior to reaction: 1: mechanical degradation, 2: basic cleavage, 3: additional oligomer, 4: fragmentase stopped by EDTA, 5: fragementase stopped by heat denaturation; **d)** as c) employing primer 2; **e)** as d) 3, PCR enhancers were added as indicated; **f)** as d) 3 employing primers 1 and 2, different temperature were used as indicated, reactions were performed upon cycling (50 x 10 s 95 °C, 60 s at indicated temperature) **g)** single-nucleotide primer extension experiments of dAMP or ddAMP, **9n**, **10a**, **9o**, **9** and **1b** opposite U (*in vitro* transcribed RNA) or Ψ (RNA extract) employing 500 nM RT-KTq2 and RT-KTq2 F667Y, final nucleotide concentration of 100 µM was used and reactions were stopped after 1 h.

In general, it was observed that primer 1 leads to more efficient incorporation of ddAMP than primer 2, but primer 2 resulted in higher incorporation efficiencies of nucleotide **9n** as almost no processing of this nucleotide can be observed for the primer deriving from sequence context 1 (see **Figure 51c** and **d**). Therefore, primer 1 was not further considered.

The different ways to facilitate incorporation on RNA extracts led to different outcomes. Employing mechanical degradation (1) results in no improvement of incorporation efficiencies. As reaction handling proved to be inconvenient as well, this approach was not further followed. Basic

fragmentation of RNA (2) was tedious due to a required purification step and does not increase incorporation remarkably. Employment of an additional oligomer (3) to melt the secondary structure, leads to enhanced incorporation efficiencies of both nucleotides. Using the fragmentase enzyme mix for enzymatic degradation of the RNA extract, no improvement can be observed regarding incorporation of ddATP or **9n**, if EDTA was used for reaction termination (4) or the reaction was stopped by heat denaturation (5).

As the best incorporation efficiency for nucleotide **9n** can be detected due to addition of an additional oligomer (3) and this method is convenient regarding reaction handling, this approach was further investigated. Even if processing of nucleotide **9n** is improved, incorporation efficiency still is not sufficient for application in Ψ detection. Therefore, a further attempt aimed at the addition of so called "PCR enhancers".[145] Manganese-ions are known to enhance the ability of DNA polymerases to accept artificial substrates. But upon partial replacement of magnesium chloride by manganese chloride (50 % or 10%), no incorporation of ddATP or **9n** is observed anymore (see **Figure 51e**). Addition of betaine, formamide or DMSO results in all cases in more efficient incorporation of ddAMP. But incorporation of **9n** is no longer detectable. Taken together, addition of PCR enhancers does not improve incorporation of **9n** at all.

In a last attempt, it was investigated, if variation of the reaction temperature can influence processing **9n** in a positive way. In addition to different temperatures, it was decided to test if multiple cycles of denaturing and annealing will enhance incorporation efficiencies (see **Figure 51g**). Again, employing primer 1 shows efficient incorporation of ddAMP but no incorporation of **9n**. If primer 2 was applied, complete extension of the primer can be observed for processing ddATP. The incorporation efficiency of nucleotide **9n** is increased at higher reaction temperatures, but unfortunately incorporation is not sufficient for Ψ detection. As nucleotide **9n** possesses big and steric demanding modifications, nucleotides bearing smaller modifications, but still leading to decent discrimination between U and Ψ were tested. As it can be seen in **Figure 51g**, incorporation of modified nucleotides by employing the RNA extract proves to be very challenging. Even those nucleotides as **9b** and **10a**, bearing small modifications are barely incorporated. Only nucleotide **9** is incorporated with high efficiencies, but unfortunately this molecule is not suited for application in our Ψ detection assay, as it shows the same tendencies in discrimination as ddAMP.

It could be shown that the depicted approach is working for Ψ detection, if artificial templates are employed. Switching to RNA extracts, no sufficient incorporation of the modified nucleotide is observed. It was assumed that the distinctly lower incorporation efficiency of **9n** opposite U or Ψ compared to a Watson-Crick basepair (G-C) has a bigger impact in case of long natural RNA than in case of the short artificial templates used in the initial assay. In order to optimise the application an improved overall incorporation efficiency is desirable. To address this issue, structural data were inspected. Assuming that the **9n** - Ψ basepair adopts the same orientation as the G-ddCTP pair in the available crystal structure of the KlenTaq variant RT-KTq2 complexed with a RNA/DNA hybrid (PDB ID: 4BWM) we simply superposed a **9n** - Ψ pair in the model building program Coot. Albeit for a more realistic view experimental structural studies are necessary, the model should deliver a first idea about potential interactions in the active site that might influence the incorporation reaction. One amino acid residue clearly attracts attention (see **Figure 52**). Threonine at site 664 is close to the alkyl

modifications of the nucleotide **9n** and may decrease incorporation efficiency by avoiding full closure of the O-helix. Further approaches will therefore aim at mutating the employed DNA polymerase RT-KTq2 F667Y to exchange this amino acid into e.g. alanine or serine thereby hopefully enhancing incorporation of nucleotide **9n**.

To further investigate the fundamental mechanisms leading to the described discrimination between U and Ψ, it would be interesting to exchange the amino acid residue arginine 677 as well. This residue seems to be in close proximity to the N1-amino group of Ψ allowing interactions. No interaction between the C5 of uridine and Arg677 can be formed in contrast to the possible interactions with N1 of Ψ. Therefore, it would be interesting to investigate discrimination behaviour if arginine 677 was exchanged by smaller amino acid residues, or amino residues that can form stronger hydrogen bonds with N1 of Ψ. Thereby, the underlying mechanism resulting in this pronounced discrimination can be studied and discrimination could even be enhanced further.

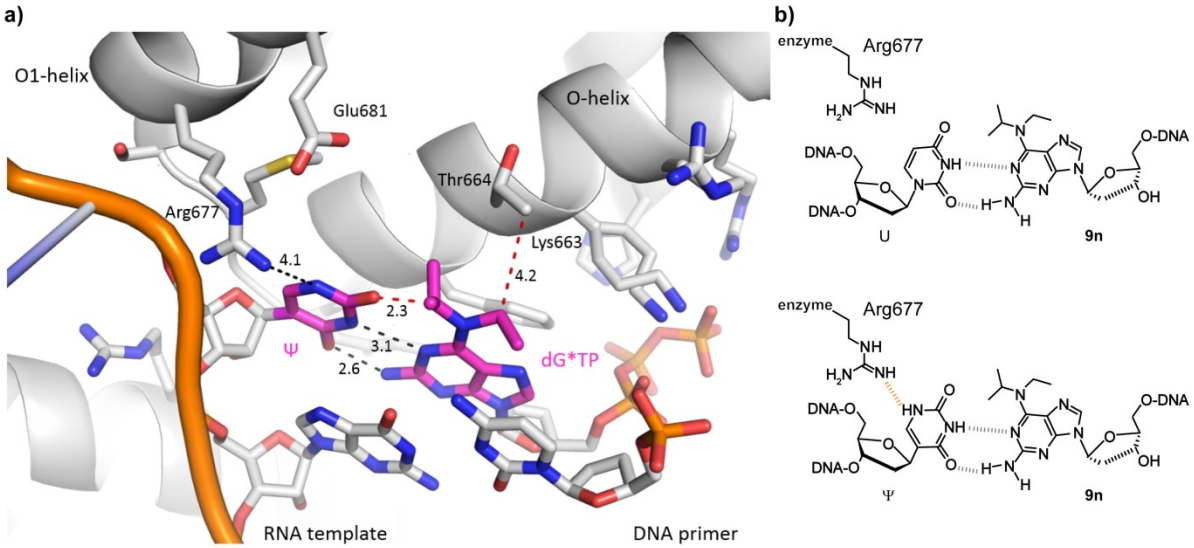


Figure 52: Structural analysis of possible interactions during incorporation. **a)** Model of a Ψ - **9n** pair in the active site of KlenTaq DNA Polymerase mutant RT-KTq2 complexed with a RNA/DNA hybrid (PDB ID: 4BWM). The nucleobases Ψ and **9n** (shown in pink) were manually superposed to the the nucleobases of the G-ddCTP base pair in the published crystal structure to get an idea about potential clashes/interactions in the active site. Possible interactions are shown by black dashed lines and clashes by red dashed lines. Distances are given in Å; **b)** chemical structures of U, Ψ, **9n** and Arg667. Top: possible H-bonding interactions between U and **9n**; bottom: possible interactions between Ψ, **9n** and Arg677.

4. Summary and Outlook

To summarise, several purine-based 2´-deoxy-nucleotides were tested for their ability to sense 5mC in DNA polymerase-catalysed reactions. Purine derivatives, modified in position 6 were thereby identified as most promising for further derivatisation. Several 6-modified dGTP analogues bearing altered H-bonding properties as well as differently sized modifications with varying steric demands were following synthesised and analysed.

Having a variously modified nucleotide toolbox in hand, those nucleotides were tested towards their discrimination behaviour between C and 5mC during incorporation by different DNA polymerases. All modified nucleotides were accepted and incorporated by all tested DNA polymerases (KlenTaq, KOD exo$^-$ and 9°North exo$^-$), even if modified nucleotides were processed with reduced efficiencies. Employing the KlenTaq DNA polymerase failed to show any promising differences in incorporation efficiencies between C and 5mC. In contrast, both sequence family B DNA polymerases KOD exo$^-$ and 9° North exo$^-$ delivered promising results regarding favoured incorporation of dGMP and modified nucleotides opposite C compared to 5mC. In general, increased discrimination between C and 5mC was found by enhancing the bulkiness of the introduced 6-modification to a certain degree, while incorporation efficiencies decreased. These findings strongly suggest that sensing 5mC is a general phenomenon by this class of nucleotides in combination with B family DNA polymerases and is not restricted to particular 6-modifications of dGTP since it applies to various functionalities.
The best discrimination between C and 5mC could be identified employing the DNA polymerase KOD exo$^-$ with O^6-ethyl-dGTP **1b**. Those results were verified by steady-state kinetics, showing a 4.2 fold more efficient incorporation opposite C than 5mC. Subsequent selectivity studies revealed lost selectivity of nucleotide **1b** between C and T. Even if 6-modified nucleotides cannot easily be employed for sequencing applications, due to their lost selectivity for C, those experiments clearly show that O^6-modified nucleotides can be used to sense the presence or absence of 5mC in the template strand.

As working with radioactivity has obvious disadvantages and analysis via denaturing PAGE limits the scope of enzyme analysis,[62, 115] a new method was established by the usage of fluorescently labelled primers and analysis via capillary electrophoresis. By optimisation of run parameters, single base resolution was achieved and a robust method could be presented, which allows multiplexing and therefore analysis of over hundred reactions in parallel. This method was used for the investigation of all modified nucleotides towards their potential to be used in 5mC detection employing the second B family DNA polymerase 9 °North exo$^-$. Multiplexing can even be enhanced further by employing longer oligonucleotides as primers. Additionally, processed primers can be labelled using different dye-labels having well separated excitation and emission spectra (e.g. FAM, NED, VIC and PET). Simultaneous analysis of multiple substrates or reaction conditions can therefore be enabled by multiplexing oligonucleotide design by size and fluorescent dye.

Remarkable differences in incorporation efficiencies were observed employing differently modified nucleotides. The best effects regarding discrimination were observed by the application of 6-modified nucleotides. Unfortunately, those nucleotides showed lost selectivity between incorporation opposite C and T. Another class of derivatives, 8-modified dGTP analogues did not show as promising trends regarding diverging incorporation efficiencies. But discrimination with favoured incorporation opposite C was detected, while their selectivity was maintained. Therefore, it was envisioned that combining those modifications will maintain selectivity while enhancing discrimination behaviour. After synthesis of two in position 6 and 8 doubly modified nucleotides, those were tested towards their behaviour in DNA polymerase catalysed reactions. Even if nucleotides were incorporated by KOD exo$^-$, selectivity between C and G was lost and discrimination decreased in comparison to the 6-modified derivative.

Due to the lost selectivity of the discriminating nucleotides, establishment of one direct method to sense 5mC can only be realised for site directed detection.[1] Since the potential of modified nucleotides was already exploited, the focus was put on modifying the used DNA polymerases as next step. I aimed at identifying mutants with increased discrimination in incorporating dGMP opposite 5mC compared to C. As DNA polymerases normally replicate 5mC without difficulty and the methyl group in 5mC does not hinder Watson-Crick base pairing, this is a challenging task. Therefore, Claudia Huber and Martina Adam generated 34 different libraries by mutating several sites in different domains of the DNA polymerase KOD exo$^-$. In order to achieve all 19 possible mutants at every investigated site, they introduced defined mutations at the respective sites via site-directed mutagenesis. After those mutants were tested for activity, they were investigated for their enhanced ability to discriminate between C and 5mC in the template strand by the usage of dGTP and nucleotide 8-vinyl-dGTP **3**. The site G245 was identified as most promising for further investigations. Processing of nucleotide **3** did not show enhanced discrimination compared to dGTP, but remarkably reduced incorporation efficiencies. Therefore, it was decided to focus on application of dGTP in further experiments. After studying incorporation efficiencies of dGMP opposite C and 5mC and determining steady-state kinetics,[114] the mutant G245D was verified to lead to the most pronounced differences in incorporation efficiencies. The selectivity of this mutant was proven to be maintained. For the application in sequencing approaches, it was decided to synthesise a nucleotide, bearing a fluorescent dye attached to the terminal γ-phosphate (**32**). The behaviour of this nucleotide in DNA polymerase-catalysed reactions was studied via steady-state kinetics. These studies revealed a 4.1 fold difference in k_{cat} with favoured processing of this nucleotide opposite C. Nevertheless, it could be observed that employment of the modified nucleotide **32** slowed the DNA polymerase variants remarkably. To enable application of those effects in SMRT sequencing approaches,[60] the efficiency of processing nucleotide **32** by the KOD exo$^-$ variant needs to be improved. It has been shown in literature[100] that the introduction of additional phosphate groups into the phosphate chain of nucleotides dramatically increases the ability of DNA polymerases to process nucleotides, modified at the terminal phosphate. Therefore, derivatives of nucleotide **32** need to be synthesised which bear additional phosphate groups between the linker and the sugar moiety.

As studies regarding 5hmC intensified, recent data suggested that 5hmC may not only be an intermediate in the active removal of 5mC. Moreover, 5hmC may act additionally as an epigenetic modification itself, thus adding another layer of complexity to the network of epigenetic regulations.[24,25] Having a variously modified nucleotide toolbox in hands, those nucleotides were tested towards their action on 5hmC as well. Same tendencies were observed, as described for 5mC detection. All nucleotides were incorporated by all DNA polymerases, albeit with decreased incorporation efficiencies. Employing KlenTaq DNA polymerase, no notable differences in incorporation efficiencies were observed between incorporation opposite C, 5mC and 5hmC. On the opposite, employing both sequence family B DNA polymerases, reduced incorporation efficiencies were observed for all nucleotides, including dGMP, opposite 5hmC. Already remarkable discrimination was observed between C and 5hmC for processing of dGTP employing KOD exo$^-$. Discrimination between 5mC and 5hmC varied stongly between the employed mofdified nucleotides.

The occurrence of modified nucleotides is not just known for DNA, but plays important roles in RNA as well, as cellular RNA contains over 150 chemically altered nucleotides, which are usually formed by enzymatic posttranskriptional modifications.[2, 3] Since those modifications of RNA are typically of low abundance, they frequently are unnoticed by standard detection methods.[46] Nina Blatter established a KlenTaq mutant (RT-KTq2),[125] which could extend DNA primers utilising a RNA template. Having a whole tool box of variously modified nucleotides in hand, it was decided to employ those nucleotides in combination with this mentioned KlenTaq mutant for the detection of RNA modifications, as the potential offered by employing purine based derivatives for the discrimination between modified and unmodified nucleotides was already proven.
Methylation of the 2´-OH group of the ribose moiety is one of the most abundant RNA modifications.[126] To investigate, whether the modified nucleotides can be exploited for RNA modifications as well by combining them with the RT-KTq2, all modified nucleotides were screened for 2´-OMe-C detection. As observed before for KlenTaq, incorporation efficiencies of the employed nucleotides decreased considerably with increasing size of the modification. Even enhancing the concentration of this KlenTaq variant RT-KTq2 incorporation of all nucleotides could not be achieved. Interestingly, already incorporation of unmodified dGMP was more efficient opposite C than opposite 2´-OMe-C resulting in a discrimination factor of 1.42. This discrimination could be remarkably enhanced by processing modified nucleotides. The highest discrimination was observed employing nucleotide **3**, as **3** was processed with 7.40 fold higher efficiency opposite C than opposite 2´-OMe-C. Again, all nucleotides were preferably processed opposite C. This observed discrimination effects need to be proven by the measurement of steady-state kinetics and if verified should be exploited in 2´-OMe-C detection.

The occurrence of another important RNA modification - pseudouridine (Ψ) - was recently connected to human diseases.[49] Therefore, robust and simple tools to detect the occurrence and distribution of Ψ are needed to evaluate global Ψ dynamics and changes to reveal the underlying mechanisms. As already discovered before, in position 6 modified purines were preferably incorporated opposite T. Therefore, it was decided to employ incorporation of modified nucleotides by

the DNA polymerase RT-KTq2 for discrimination between U and Ψ as well. As already observed for incorporation opposite C and 2´-OMe-C, it was not possible to achieve incorporation for all modified nucleotides. But those nucleotides, being incorporated showed very promising trends regarding discrimination behaviour. Nucleotides, resembling the structure of dATP, showed favoured incorporation opposite U, while other nucleotides showed remarkably increased incorporation opposite Ψ. This was the first time reverse discrimination could be observed by favoured incorporation opposite a modified nucleotide. Verified by steady-state kinetic data, a doubly 6-amino modified nucleotide **9n** could be presented which was incorporated opposite Ψ with 15 fold higher catalytic efficiency compared to the incorporation opposite U. To employ this very promising discrimination in Ψ detection, a novel assay was presented. I could successfully demonstrate the potential of this assay, as remarkable discrimination between U and Ψ was observed employing artificial RNA templates. To adopt this assay to human RNA extracts, further investigations need to be done. First of all, the incorporation efficiency of nucleotide **9n** on RNA extracts by RT-KTq2 needs to be improved. Inspection of structural data pointed to two amino acids that should be mutated. The steric demanding modifications of **9n** might hinder efficient incorporation due to the close proximity of Thr664. By exchanging this amino acid by smaller residues, incorporation efficiencies could be improved allowing efficient processing of **9n** on RNA extracts. To further investigate the fundamental mechanisms leading to the described discrimination between U and Ψ, it would be interesting to exchange the amino acid residue arginine 677. This residue seems to be in close proximity to the N1-amino group of Ψ allowing interactions. Therefore, it would be interesting to investigate discrimination behaviour if arginine 677 was exchanged by smaller amino acid residues, or amino residues that can form stronger hydrogen bonds with N1 of Ψ. Thereby, the underlying mechanism resulting in this pronounced discrimination can be studied and discrimination could even be enhanced further.

Herewith I have developed a tool box to examine incorporation efficiencies opposite different modifications in nucleic acids. As the biological impact of different modified nucleotides in DNA and RNA is currently intensively studied, a great potential can be seen in applying this approach to the detection of the different modifications in nucleic acids.

5. Zusammenfassung

Zusammenfassend wurden zahlreiche Purin - basierte 2´-Deoxynucleotide auf ihre Anwendbarkeit für die Detektion von 5mC getestet, wodurch Position 6 als am vielversprechendsten für weitere Modifizierungen identifiziert wurde. Daraufhin wurde eine Vielzahl an 6-modifizierten dGTP Analoga synthetisiert, die sowohl variierende H-Brückenbindungs-Eigenschaften, als auch verschieden große Modifikationen mit unterschiedlichen sterischen Ansprüchen aufweisen.

Da nun eine Vielzahl an verschiedenartig modifizierten Nukleotiden zur Verfügung stand, wurden diese in Bezug auf ihr Diskriminierungsverhalten zwischen C und 5mC getestet. Eine Diskriminierung sollte durch den Einbau dieser Nukleotide mit Hilfe von verschiedenen DNA-Polymerasen zustande kommen. Alle modifizierten Nukleotide wurde von allen getesteten DNA-Polymerasen (KlenTaq, KOD exo$^-$ und 9° Nord exo$^-$) akzeptiert, auch wenn die modifizierten Derivate mit verminderter Effizienz umgesetzt wurden. Die Anwendung der KlenTaq DNA-Polymerase führte zu keinen nennenswerten Unterschieden der Einbaueffizienzen gegenüber C und 5mC. Im Gegensatz dazu, konnten für beide B-Familien DNA-Polymerasen vielversprechende Tendenzen mit bevorzugtem Einbau gegenüber C beobachtet werden. Im Allgemeinen wurde eine erhöhte Diskriminierung erreicht, wenn die Sperrigkeit der in Position 6 eingeführten Modifikationen bis zu einem gewissen Grad erhöht wurde, auch wenn die Einbaueffizienzen gleichzeitig sanken. Diese Entdeckungen legen nahe, dass das Erkennen von 5mC ein generelles Phänomen ist, das nicht an eine bestimmte Modifikation oder eine bestimmte DNA-Polymerase gebunden ist.
Die beste Diskriminierung zwischen C und 5mC konnte durch die DNA-Polymerase KOD exo$^-$ und das Nukleotid O^6-Ethyl-dGTP **1b** erzielt werden. Diese Ergebnisse wurden durch steady-state Kinetiken bestätigt, die einen 4.2 fach effizienteren Einbau gegenüber C als gegenüber 5mC zeigten. Nachfolgende Selektivitäts-Studien offenbarten einen Selektivitäts-Verlust von allen in Position 6 modifizierten Nukleotiden zwischen C und T. Dadurch wird die Anwendung dieser Ergebnisse für eine Sequenzierung durch einen einzigen Lauf verhindert. Diese Effekte können jedoch in der positions-spezifischen Detektion von 5mC Anwendung finden.[1]

Da die Arbeit mit Radioaktivität offensichtliche Nachteile birgt und die Analyse von Primer-Verlängerungs-Reaktionen über denaturierende PAGE den Durchsatz erheblich beeinträchtigt, haben wir durch den Einsatz von Fluoreszenz-markierten Primern und Kapillarelektrophorese eine neue Methode entwickelt. Durch die Optimierung der Lauf-Parameter konnten wir eine Einzel-Nukleotid-Auflösung erreichen und somit eine robuste Methode präsentieren, die die gleichzeitige Analyse von über hundert Reaktionen ermöglicht. Diese Methode wurde genutzt um alle modifizierten Nukleotide in Kombination mit der DNA Polymerase 9° Nord exo$^-$ für den Einsatz zur 5mC Detektion zu testen.

Beachtliche Unterschiede in Einbaueffizienzen gegenüber C und 5mC konnten durch den Einsatz von modifizierten Nukleotiden erreicht werden. Die besten Effekte konnten hierbei durch die in Position 6 modifizierten Nukleotide erzielt werden. Diese zeigten jedoch keine Selektivität zwischen C und T. Eine andere Klasse an modifizierten Nukleotiden - modifiziert in Position 8 - zeigt weniger

ausgeprägte Trends in Bezug auf Diskriminierung, weist jedoch Selektivität für C auf. Daher wurde beschlossen, diese Modifikationen zu kombinieren, um Nukleotide zu erhalten, die unter Erhalt ihrer Selektivität für C eine erhöhte Diskriminierung zwischen C und 5mC aufweisen. Nach Synthese dieser in Position 6 und 8 doppelt modifizierten Nukleotide, wurden diese auf ihr Einbauverhalten durch DNA-Polymerasen untersucht. Hierbei musste jedoch festgestellt werden, dass diese Nukleotide von KOD exo$^-$ zwar eingebaut wurden, jedoch weder erhöhte Diskriminierung, noch Selektivität für C aufwiesen.

Durch den Verlust der Selektivität der diskriminierenden Nukleotide, können sie ausschließlich zur positions-spezifischen Detektion von 5mC eingesetzt werden. Daher wurde im nächsten Schritt ein Fokus auf die Modifikation der verwendeten DNA-Polymerase gelegt. Hierzu sollten DNA-Polymerasen Mutanten identifiziert werden, die eine erhöhte Diskriminierung zwischen C und 5mC aufweisen. Da DNA-Polymerasen 5mC normalerweise ohne Probleme replizieren können und die Methylgruppe in Position 5 nicht mit der Watson-Crick Basenpaarung interferiert, ist dies jedoch eine herausfordernde Aufgabe. Claudia Huber und Martina Adam haben hierzu 34 verschiedene KOD exo$^-$ Bibliotheken generiert, in dem sie verschiedene Positionen in unterschiedlichen Domänen über „site-directed“ Mutagenese variiert haben. Nachdem diese Mutanten auf ihre Aktivität getestet worden waren, wurden ihre Diskriminierung zwischen C und 5mC untersucht. Hierfür wurden dGTP und 8-vinyl-dGTP **3** eingesetzt. Die Position G245 wurde als vielversprechendste Stelle gefunden. Durch die Umsetzung von **3** konnte keine verbesserte Diskriminierung erzielt werden. Die Einbaueffizienzen sanken jedoch deutlich. Daher wurde der Fokus im Folgenden auf die Umsetzung von dGTP gelegt. Nachdem steady-state Kinetiken für den Einbau von dGMP gegenüber von C und 5mC durchgeführt wurden, konnte die Mutante G245D als Variante mit den besten Diskriminierungs-Eigenschaften bestätigt werden. Zusätzlich konnte gezeigt werden, dass die Selektivität dieser Mutante für C erhalten geblieben war. Um diese Effekte für Sequenzierungen anwenden zu können, wurde ein weiteres Nukleotid synthetisiert, an dessen terminalem Phosphat ein Fluoreszenz-Farbstoff angebracht wurde (**32**). Das Verhalten dieses Nukleotides in DNA-Polymerasen-katalysierten Experimenten wurde wiederum anhand von steady-state Kinetiken untersucht. Diese Studien haben gezeigt, dass Nukleotid **32** von G245D mit einer 4.1 fach höheren katalytischen Effizienz (k_{cat}) gegenüber C als gegenüber 5mC eingebaut wird. Allerding wurde ebenfalls gezeigt, dass der Einsatz des modifizierten Nukleotides **32** die DNA Polymerasen deutlich verlangsamt hat. Um die Anwendung dieser Ergebnisse in SMRT Sequenzierungen zu ermöglichen, muss die Effizienz des Umsatzes des Nukleotids **32** durch die KOD exo$^-$ Variante verbessert werden. In der Literatur wurde bereits gezeigt, dass eine zusätzliche Phosphatgruppe in der Phosphatkette des Nukleotids die Fähigkeit der DNA-Polymerasen diese Nukleotide umzusetzen erheblich gesteigert wird. Aus diesem Grund sollten in der Zukunft weitere Derivate des Nukleotids **32** mit zusätzlichen Phosphatgruppen synthetisiert werden.

Aktuelle Forschungen legen die Vermutung nahe, dass 5hmC nicht nur als Intermediat in der Demethylierung von 5mC, sondern auch selbst als epigenetischer Marker dient. Da wir eine Vielzahl an modifizierten Nukleotiden synthetisiert haben, sollten diese ebenfalls für die Detektion von 5hmC getestet werden. Für die Diskriminierung zwischen C oder 5mC und 5hmC wurden ähnliche

Tendenzen beobachtet. Alle Nukleotide wurden von allen DNA-Polymerasen eingebaut, wenn auch mit verminderter Effizienz. Durch die DNA-Polymerase KlenTaq konnte keine bedeutende Diskriminierung beobachtet werden. Im Gegensatz dazu wurden wiederum deutliche Diskriminierungen zwischen C, 5mC und 5hmC beobachtet, wenn die B-Familien DNA-Polymerasen KOD exo$^-$ und 9 °Nord exo$^-$ eingesetzt wurden. Hierbei konnte gezeigt werden, dass alle Nukleotide - einschließlich dGMP - deutlich schlechter gegenüber 5hmC als gegenüber von C eingebaut wurden.

Modifizierte Nukleotide sind nicht nur in DNA sondern auch in RNA bekannt. Über 150 verschiedene Modifikationen konnten bereits in zellulärer RNA nachgewiesen werden, wobei diese teilweise sehr geringe Konzentrationen aufweisen. Dies erschwert deren Detektion erheblich. Nina Blatter entwickelte eine KlenTaq Mutante (RT-KTq2), die DNA ausgehend von einem RNA Templat synthetisieren kann. Daher entschieden wir, die Vielzahl an modifizierten Nukleotiden in Kombination mit dieser Mutante zur Detektion von RNA-Modifikationen einzusetzen.
Die Methylierung der 2´-OH Gruppe der Ribose ist eine der weitverbreitetsten RNA-Modifikationen. Daher haben wir alle modifizierten Nukleotide für die Anwendung zur Detektion von RNA-Modifikationen in Verbindung mit der KlenTaq Mutante RT-KTq2 getestet. Wie wir bereits beobachten konnten, nahm die Einbaueffizienz der modifizierten Nukleotide mit zunehmender Größe der Modifikationen ab. Manche Nukleotide konnten auch nach deutlicher Erhöhung der DNA-Polymerasen Konzentration nicht eingebaut werden. Interessanterweise konnte bereits für das unmodifizierte dGMP ein 1.42 höherer Einbau gegenüber C als gegenüber 2´-OMe-C beobachtet werden. Die höchste Diskriminierung (Faktor 7.40) konnte allerdings durch den Einbau des Nukleotids **3** erzielt werden. Wiederum wurden alle Nukleotide bevorzugt gegenüber dem unmodifizierten C eingebaut.

Eine weitere wichtige RNA-Modifikation, Pseudouridin Ψ, konnte mittlerweile mit verschiedenen menschlichen Krankheiten assoziiert werden. Daher sind auch zur Detektion dieser Modifikationen einfache und robuste Methoden von Nöten. Da bereits erkannt wurde, dass in 6-Position modifizierte Nukleotide bevorzugt gegenüber von T eingebaut werden, wurden diese Nukleotide auch in Bezug auf die Detektion von Ψ untersucht. Nicht alle modifizierten Nukleotide konnten von RT-KTq2 gegenüber von U oder Ψ eingebaut werden. Die Nukleotide, die eingebaut wurden zeigten jedoch sehr interessante Effekte. Derivate, welche der Struktur von dATP und dessen H-Brücken Basenpaarungsmuster mehr glichen als dem von dGTP zeigten einen bevorzugten Einbau gegenüber von U. Andere Nukleotide hingegen bewiesen einen bevorzugten Umsatz gegenüber von Ψ. Dies ist somit das erste Mal, dass ein effektiverer Einbau eines Nukleotides gegenüber einer Nukleinsäure-Modifikation und somit eine invertierte Diskriminierung beobacht werden konnte. Nachdem diese Effekte durch steady-state Kinetiken verifiziert werden konnten, erwies sich Nukleotid **9n** mit einem 15-fach effektiverem Umsatz gegenüber Ψ als gegenüber U als vielversprechendstes Derivat. Diese beachtliche Diskriminierung konnte im Folgenden für die Entwicklung einer neuartigen Detektionsmethode auf artifizieller RNA eingesetzt werden. Um diese sehr viel versprechende Methode auf RNA Extrakte übertragen zu können, müssen jedoch noch einige Untersuchungen angestellt werden, da Nukleotid **9n** nicht gegenüber von RNA Extrakten als Templat umgesetzt werden konnte.

Zusammenfassend wurde eine Vielzahl an verschiedenartig modifizierten Nukleotiden synthetisiert und ihre Einbaueffizienzen durch DNA Polymerasen gegenüber unterschiedlicher Nukleinsäure-Modifikationen getestet. Da die biologische Auswirkung dieser Nukleinsäure-Modifikationen zurzeit intensiv untersucht wird, sehe ich ein sehr großes Potential in der Anwendung dieses Ansatzes zur Detektion von modifizierten Nukleotiden in DNA, wie auch in RNA.

6 Experimental Part

6.1 Chemical Synthesis

6.1.1. General Information

Solvents and reagents were purchased from Sigma-Aldrich, Fluka, Acros or Carbosynth and were used without further purification. Dry solvents were purchased from Sigma-Aldrich. Mixtures of solvents were declared as percent by volume [v/v]. DNA polymerases KlenTaq and KOD exo$^-$ were expressed and purified as described before.[130] T4 polynucleotide kinase (PNK) was purchased from New England Biolabs. [γ-^{32}P]-ATP was purchased from Hartmann Analytics and natural dNTPs from Roche. Reactions were conducted with exclusion of air and moisture as needed. Thin layer chromatography (TLC) was performed using silica gel 60 F254 aluminium plates from Merck. Spots were visualized under UV-light or by staining. Preparative flash chromatography was carried out using silica gel G 60 (40-63 µm, Merck) with a pressure of 0.3 bar. NMR spectra were recorded on a Bruker Avance III 400 MHz spectrometer and a Bruker AVIII 600 MHz spectrometer. ^{1}H, ^{32}P and ^{13}C chemical shifts are reported relative to the residual solvent peak and are given in ppm (δ). Data are reported as follows: chemical shift (multiplicity (singlet (s), broad singlet (br s), doublet (d), doublet of doublets (dd), doublet of doublet of doublets (ddd), triplet (t), doublet of triplets (dt), quartet (q), doublet of quartets (dq), pentet (p), heptet (h), multiplet (m)), coupling constants [Hz], integration, assignment). HR-ESI-MS spectra were recorded on a Bruker Daltronics microTOF II in positive or negative mode.

6.1.2. Buffers for Chemical Synthesis

Triethylammonium bicarbonate buffer (TEAB buffer)

700 ml of triethylamine (5.0 mol) were added to 3.5 l of water. 5 kg of dry ice were evaporated and the carbon dioxide passed through the solution. If the pH was between 7.5 and 8.0 the pH was adjusted to 7.5 using acetic acid. If the pH was higher than 8.0 more dry ice was evaporated and the carbon dioxide passed through the solution. The solution was filled with water to 5 l to yield 1 M TEAB buffer. The buffer was diluted to 0.1 M with water as needed.

Triethylammonium acetate buffer (TEAA buffer)

57 ml of acetic acid (1.0 mol) and 139 ml of triethylamine (1.0 mol) were added to 700 ml of water. After the solution has cooled to room temperature, the pH was adjusted to 7.0 and the solution filled to 1 l to yield 1 M TEAA buffer. The buffer was diluted with water to 50 mM as needed.

6.1.3. Bis(tributylammonium)-pyrophosphate[92]

Sodium pyrophosphate dibasic (2.3 g, 10 mmol, 1 eq.) was dissolved in 75 ml water and applied to a washed Amberlite IR-120 column (3 cm diameter, 20 cm length, pH 5). All flow through was dropped into an ice-cold solution of Bu_3N (4.75 ml, 20 mmol, 2 eq.) in 40 ml ethanol. The column was washed with water till the pH of the flow through was pH 5. The solution was concentrated *in vacuo* and co-evaporated with 10 ml ethanol twice and three times with 10 ml abs. DMF. The residue was diluted to a total volume of 20 ml with abs. DMF to yield a 0.5 M solution of bis(tributylammonium)-pyrophosphate in DMF.

6.1.4. 8-Bromo-2´-deoxyguanosine 14 [98b]

2´-Deoxyguanosine (4.0 g, 14.9 mmol, 1.0 eq.) was dissolved in a mixture from acetonitrile (160 ml) and water (40 ml). *N*-bromosuccinimide (4.3 g, 23.95 mmol, 1.6 eq.) was added and the reaction mixture was stirred for 1 h at room temperature and subsequently evaporated to dryness. The crude residue was suspended in acetone, filtered and dried *in vacuo*. Yield: 98 % (5.05 g, 14.60 mmol). ^{1}H-NMR (400 MHz, DMSO-d_6): 10.78 (br s, 1H, -NH-), 6.47 (br s, 2H, -NH_2), 6.15 (dd, *J* = 6.7 Hz, 7.9 Hz, 1H, H-1´), 5.24 (d, *J* = 4.3 Hz, 1H, 3´-OH), 4.85 (br s, 1H, 5´-OH), 4.40 (dt, *J* = 3.2 Hz, 6.4 Hz, 1 H, H-3´), 3.80 (dt, *J* = 3.0 Hz, 5.6 Hz, 1H, H-4´), 3.64 – 3.60 (m, 1H, H-5´a), 3.52 – 3.48 (m, 1H, H-5´b), 3.16 (ddd, *J* = 6.4 Hz, 8.0 Hz, 13.1 Hz, 1H, H-2´a), 2.10 (ddd, *J* = 3.0 Hz, 6.8 Hz, 13.3 Hz, 1H, H-2´b) ppm. ^{13}C-NMR (100 MHz, DMSO-d_6): 155.8 (C-6), 153.7 (C-2), 152.4 (C-4), 121.0 (C-8), 117.9 (C-5), 88.3 (C-4´), 85.5 (C-1´), 71.4 (C-3´), 62.3 (C-5´), 36.8 (C-2´) ppm. HR-ESI-MS $[M-1H]^-$: m/z calculated: 343.9989, m/z found: 344.0025.

6.1.5. 8-Vinyl-2´-deoxyguanosine 15 [98b]

8-Bromo-2´-deoxyguanosine (1.0 g, 2.9 mmol, 1.0 eq.), tetrakis(triphenylphosphine) palladium (0) (167 mg, 0.144 mmol, 0.05 eq.) and tributyl(vinyl)stannane (1.01 g, 3.18 mmol, 1.1 eq.) were dissolved in *N*-methyl-2-pyrrolidone (7 ml) and stirred for 2 h at 110°C. The reaction mixture was following poured in 100 ml diethyl ether and the crude residue was collected by filtration and purified using column flash chromatography with methylene chloride and 10 % methanol. Yield: 83 % (697 mg, 2.38 mmol). ^{1}H-NMR (400 MHz, DMSO-d_6): 10.66 (br s, 1H, -NH-), 7.01 (dd, *J* = 11.0 Hz, 17.0 Hz, 1H, vinyl-H), 6.45 (br s, 2H, -NH_2), 6.26 (dd, *J* = 6.1 Hz, 8.8 Hz, 1H, H-1´), 6.11 (dd, *J* = 2.2 Hz, 17.1 Hz, 1H, vinyl-H), 5.40 (dd, *J* = 2.0 Hz, 11.0 Hz, 1H, vinyl-H), 5.27 (d, *J* = 4.2 Hz, 1H, 3´-OH), 5.00 (t, *J* = 5.4 Hz, 1H, 5´-OH), 4.37 (dq, *J* = 2.9 Hz, 6.3 Hz, 1 H, H-3´), 3.81 – 3.76 (m, 1H, H-4´), 3.66 - 3.52 (m, 2H, H-5´a/b), 2.63 (ddd, *J* = 6.6 Hz, 8.9 Hz, 13.0 Hz, 1H, H-2´a), 2.06 (ddd, *J* = 2.4 Hz, 6.1 Hz, 13.1 Hz, 1H, H-2´b) ppm. ^{13}C-NMR (100 MHz, DMSO-d_6): 157.0 (C-6), 153.8 (C-2), 151.9 (C-4), 144.4 (C-8), 125.4 (C-5), 119.6 (vinyl-C), 116.7 (vinyl-C), 87.7 (C-4´), 82.9 (C-1´), 71.0 (C-3´), 62.0 (C-5´), 40.0 (C-2´) ppm. HR-ESI-MS $[M-1H]^-$: m/z calculated: 292.1051, m/z found: 292.1061.

6.1.6. 3´, 5´-Di-*O*-acetyl-2´-deoxyguanosine 11 [102]

2´-Deoxyguanosine (5.0 g, 19.0 mmol, 1.0 eq.), DMAP (0.229 g, 1.9 mmol, 0.1 eq.) and triethylamine (7.3 ml, 57.0 mmol, 3.0 eq.) were dissolved in acetonitrile (100 ml). Acetic anhydride (5.39 ml, 57.0 mmol, 3.0 eq.) was added and the reaction mixture was stirred for 3 h at room temperature. The reaction was stopped by addition of 1 ml methanol and the solution was concentrated to dryness. The remaining white solid was co-evaporated using acetonitrile and washed with a mixture of ethanol and diethyl ether (1/1) to yield a white solid in quantitative yield (6.7 g, 19.0 mmol). ^{1}H-NMR (400 MHz, DMSO-d_6): 10.65 (br s, 1H, -NH-), 7.91 (s, 1H, H-8), 6.48 (br s, 2H, -NH_2), 6.13 (dd, *J* = 5.9 Hz, 8.6 Hz, 1H, H-1´), 5.29 (dt, *J* = 2.0 Hz, 6.1 Hz, 1H, H-3´), 4.29 - 4.24 (m, 1H, H-5´a), 4.21 - 4.16 (m, 2H, H-4´, H-5´b), 2.91 (ddd, *J* = 6.3 Hz, 8.7 Hz, 14.6 Hz, 1H, H-2´a), 2.45 (ddd, *J* = 2.0 Hz, 5.9 Hz, 14.0 Hz, 1H, H-2´b), 2.08 (s, 3H, -CH_3), 2.03 (s, 3H, -CH_3) ppm. ^{13}C-NMR (100 MHz, DMSO-d_6): 170.6 (-**C**(O)CH_3), 170.5 (-**C**(O)CH_3), 157.1 (C-6), 154.2 (C-2), 151.6 (C-4), 135.6 (C-8), 117.3 (C-5), 83.1 (C-1´), 81.9 (C-4´), 75.0 (C-3´), 64.1 (C-5´), 35.9 (C-2´), 21.3 (-CH_3), 21.0 (-CH_3) ppm. HR-ESI-MS $[M+1H]^+$: m/z calculated: 352.1252, m/z found: 352.1235.

6.1.7. 3´, 5´-Di-*O*-acetyl-6-deoxy-6-chloro-2´-deoxyguanosine 12 [102]

To a suspension of 3´,5´-di-*O*-acetyl-2´-deoxyguanosine (4.0 g, 11.4 mmol, 1.0 eq.) in acetonitrile (60.0 ml) were added tetraethylammonium chloride (2.83 g, 17.0 mmol, 1.5 eq.) and *N,N*-dimethylaniline (8.7 ml, 68.3 mmol, 6.0 eq.). The reaction mixture was cooled to 0°C and phosphoryl chloride (6.2 ml, 68.3 mmol, 6.0 eq.) was added drop wise. The reaction mixture was stirred at room temperature for 10 min. The mixture was heated to reflux for 15 min in a preheated oil bath, afterwards immediately cooled with an ice bath and quickly concentrated to dryness. The remaining phosphoryl chloride was slowly hydrolysed by addition of 30 ml of ice water under cooling. The solution was stirred for 20 min and extracted with ethyl acetate. The combined organic layers were washed with a saturated solution of sodium hydrogencarbonate and dried over $MgSO_4$, concentrated *in vacuo* and the crude residue was purified by column flash chromatography using methylene chloride with 2 % methanol yielding a white foam in 63 % yield (2.66 g, 7.2 mmol). ^{1}H-NMR (400 MHz, $CDCl_3$): 7.91 (s, 1H, H-8), 6.28 (dd, *J* = 6.2 Hz, 7.6 Hz, 1H, H-1´), 5.43 (dt, *J* = 2.3 Hz, 6.0 Hz, 1H, H-3´), 5.16 (br s, 2H, -NH_2), 4.46 (dd, *J* = 6.1 Hz, 13.4 Hz, H-5´a), 4.39 – 4.34 (m, 2H, H-4´, H-5´b), 2.97 (ddd, *J* = 6.4 Hz, 7.8 Hz, 14.2 Hz, 1H, H-2´a), 2.56 (ddd, *J* = 2.6 Hz, 6.1 Hz, 14.2 Hz, 1H, H-2´b), 2.14 (s, 3H, -C(O)-C**H$_3$**), 2.08 (s, 3H, -C(O)-C**H$_3$**) ppm. ^{13}C-NMR (100 MHz, $CDCl_3$): 170.6 (-**C**(O)CH_3), 170.2 (-**C**(O)CH_3), 159.0 (C-2), 153.1 (C-6), 151.8 (C-4), 140.4 (C-8), 126.0 (C-5), 84.8 (C-1´), 82.5 (C-4´), 74.5 (C-3´), 63.7 (C-5´), 36.9 (C-2´), 20.9 (-CH_3), 20.8 (-CH_3) ppm. HR-ESI-MS $[M+1H]^+$: m/z calculated: 370.091, m/z found: 370.090.

6.1.8. 3´,5´-Di-*O*-*tert*-butyldimethylsilyl-2´-deoxyguanosine 17 [106]

2´-Deoxyguanosine (4.0 g, 15.0 mmol, 1.0 eq.), *tert*-butyldimethylsilyl chloride (7.0 g, 46.4 mmol, 3.0 eq.) and imidazole (6.3 g, 92.6 mmol, 6.0 eq.) were dissolved in 15 ml abs. DMF and stirred at

room temperature for 25 h. After 10 min precipitation occurred and 30 ml ethanol were added after 25 h. The suspension was stored at -20 °C for 10 h and the product was isolated by filtration. The white precipitate was purified by washing with cold ethanol and dried *in vacuo*. Yield: 79 % (6.1 g, 11.9 mmol). ^{1}H-NMR (400 MHz, DMSO-d_6): 10.60 (br s, 1H, -NH-), 7.87 (s, 1H, H-8), 6.46 (br s, 2H, -NH_2), 6.10 (dd, *J* = 6.0 Hz, 7.7 Hz, 1H, H-1´), 4.49 (dt, *J* = 3.0 Hz, 5.6 Hz, 1 H, H-3´), 3.81 (ddd, *J* = 2.7 Hz, 4.5 Hz, 5.7 Hz, 1H, H-4´), 3.70 (dd, *J* = 5.8 Hz, 11.0 Hz, 1H, H-5´a), 3.64 (dd, *J* = 4.5 Hz, 11.0 Hz, 1H, H-5´b), 2.64 (ddd, *J* = 5.6 Hz, 7.7 Hz, 13.3 Hz, 1H, H-2´a), 2.23 (ddd, *J* = 3.2 Hz, 6.0 Hz, 13.2 Hz, 1H, H-2´b), 0.89 (s, 9H, -C(CH_3)$_3$), 0.87 (s, 9H, -C(CH_3)$_3$), 0.10 (s, 6H, 2 x –CH_3), 0.04 (s, 3H, -CH_3), 0.04 (s, 3H, –CH_3) ppm. ^{13}C-NMR (100 MHz, DMSO-d_6): 157.1 (C-6), 154.1 (C-2), 151.4 (C-4), 135.2 (C-8), 117.1 (C-5), 87.4 (C-4´), 82.5 (C-1´), 72.6 (C-3´), 63.1 (C-5´), 40.0 (C-2´), 26.3 (-**C**(CH_3)$_3$), 26.2 (-**C**(CH_3)$_3$), 18.5 (-C(**C**H_3)$_3$), 18.2 (-C(**C**H_3)$_3$), -4.26 (-CH_3), -4.47 (-CH_3), -4.98 (-CH_3) , -5.03 (-CH_3) ppm. HR-ESI-MS $[M+1H]^+$: m/z calculated: 496.2770, m/z found: 496.2801.

6.1.9. 3´,5´-Di-*O-tert*-butyldimethylsilyl-6-chloro-2´-deoxyguanosine 18 [102]

To a suspension of 3´,5´-di-*O-tert*-butyldimethylsilyl-2´-deoxyguanosine (2.82 g, 5.7 mmol, 1.0 eq.) in acetonitrile (40.0 ml) were added tetraethylammonium chloride (1.42 g, 8.5 mmol, 1.5 eq.) and *N,N*-dimethylaniline (4.3 ml, 34.2 mmol, 6.0 eq.). The reaction mixture was cooled to 0 °C and phosphoryl chloride (3.2 ml, 34.2 mmol, 6.0 eq.) was added drop wise. The reaction mixture was stirred at room temperature for 10 min. The mixture was heated to reflux for 15 min in a preheated oil bath, afterwards immediately cooled with an ice bath and quickly concentrated to dryness. The remaining phosphoryl chloride was slowly hydrolysed by addition of 30 ml of ice water under cooling. The solution was stirred for 20 min and extracted with ethyl acetate. The combined organic layers were washed with a saturated solution of sodium hydrogencarbonate and dried over $MgSO_4$, concentrated *in vacuo* and the crude residue was purified by column flash chromatography using methylene chloride with 2 % methanol yielding a white foam in 37 % yield (1.1 g, 2.1 mmol). ^{1}H-NMR (400 MHz, DMSO-d_6): 8.29 (s, 1H, H-8), 6.94 (br s, 2H, -NH_2), 6.22 (t, *J* = 6.7 Hz, 1H, H-1´), 4.53 (dt, *J* = 3.4 Hz, 6.1 Hz, 1H, H-3´), 3.85 - 3.82 (m, 1H, H-4´), 3.73 (dd, *J* = 5.7 Hz, 11.1 Hz, 1H, H-5´a), 3.65 (dd, *J* = 4.5 Hz, 11.1 Hz, 1H, H-5´b), 2.77 (ddd, *J* = 5.7 Hz, 7.2 Hz, 13.1 Hz, 1H, H-2´a), 2.29 (ddd, *J* = 3.7 Hz, 6.2 Hz, 13.3 Hz, 1H, H-2´b), 0.88 (s, 9H, -C(CH_3)$_3$), 0.84 (s, 9H, -C(CH_3)$_3$), 0.10 (s, 6H, 2 x –CH_3), 0.02 (s, 3H, -CH_3), 0.02 (s, 3H, –CH_3) ppm. ^{13}C-NMR (100 MHz, DMSO-d_6): 160.1 (C-6), 154.0 (C-4), 149.0 (C-2), 140.3 (C-8), 124.0 (C-5), 87.4 (C-4´), 83.1 (C-1´), 72.3 (C-3´), 63.0 (C-5´), 38.9 (C-2´), 26.2 (-**C**(CH_3)$_3$), 26.1 (-**C**(CH_3)$_3$), 18.4 (-C(**C**H_3)$_3$), 18.2 (-C(**C**H_3)$_3$), -4.3 (-CH_3), -4.5 (-CH_3), -5.0 (-CH_3) , -5.1 (-CH_3) ppm. HR-ESI-MS $[M+1H]^+$: m/z calculated: 514.2431, m/z found: 514.2399.

6.1.10. 6-Chloro-2´-deoxyguanosine 19 [131]

3´,5´-Di-*O-tert*-butyldimethylsilyl-6-chloro-2´-deoxyguanosine (1.0 g, 1.9 mmol, 1 eq.) was dissolved in abs. THF (10 ml) at room temperature. Triethylamine trihydrofluoride (196 µL, 1.2 mmol, 1.7 eq.) was added and the reaction was stirred at room temperature for 16 h. Following, the solution

was concentrated *in vacuo* and the crude residue was purified by column flash chromatography using methylene chloride with up to 10 % methanol to yield 93 % product (504.0 mg, 1.7 mmol). ^{1}H-NMR (400 MHz, DMSO-d_6): 8.34 (s, 1H, H-8), 6.94 (br s, 2H, -NH_2), 6.22 (t, *J* = 6.7 Hz, 1H, H-1´), 5.28 (d, *J* = 4.1 Hz, 1H, 3´-OH), 4.92 (t, *J* = 5.5 Hz, 1H, 5´-OH), 4.37 (dq, *J* = 3.5 Hz, 6.7 Hz, 1H, H-3´), 3.83 (dt, *J* = 2.9 Hz, 4.6 Hz, 1H, H-4´), 3.58 (dt, *J* = 5.1 Hz, 11.7 Hz, 1H, H-5´a), 3.51 (dt, *J* = 5.0 Hz, 11.7 Hz, 1H, H-5´b), 2.61 (ddd, *J* = 5.8 Hz, 7.4 Hz, 13.2 Hz, 1H, H-2´a), 2.26 (ddd, *J* = 3.4 Hz, 6.2 Hz, 13.2 Hz, 1H, H-2´b) ppm. ^{13}C-NMR (100 MHz, DMSO-d_6): 160.2 (C-6), 154.1 (C-4), 149.9 (C-2), 141.5 (C-8), 124.0 (C-5), 88.2 (C-4´), 83.5 (C-1´), 71.0 (C-3´), 62.0 (C-5´), 40.0 (C-2´) ppm. HR-ESI-MS $[M+1H]^+$: m/z calculated: 286.0701, m/z found: 286.0682.

6.1.11. General procedure A:

2´,3´-Di-*O*-acetyl-6-deoxy-6-chloro-2´-deoxyguanosine (1.0 eq.) was dissolved in freshly prepared 1 M solution of the respective sodium alkoxide (30.0 eq.) in the corresponding alcohol and stirred under reflux overnight. After cooling to room temperature, the reaction mixture was concentrated to dryness, dissolved in water and the pH was adjusted to pH 7 using acetic acid. The aqueous solution was extracted with ethyl acetate and the organic layer was concentrated *in vacuo*. The remaining crude residue was further purified by column flash chromatography using methylene chloride with 4 % methanol. The products were obtained as colourless foams.

6.1.12. O^6-Methyl-2´-deoxyguanosine 13a

Yield: 87 % (264 mg, 0.9 mmol). ^{1}H-NMR (400 MHz, DMSO-d_6): 8.08 (s, 1H, H-8), 6.43 (br s, 2H, -NH_2), 6.21 (dd, *J* = 6.1 Hz, 7.7 Hz, H-1´), 5.26 (d, *J* = 4.0 Hz, 3´-OH), 4.98 (t, *J* = 5.5 Hz, 1H, 5´-OH), 4.35 (dq, *J* = 3.0 Hz, 6.0 Hz, 1H, H-3´), 3.96 (s, 3H, -O-CH_3), 3.82 (dt, *J* = 2.8 Hz, 4.5 Hz, 1H, H-4´), 3.57 (dt, *J* = 5.1 Hz, 11.8 Hz, 1H, H-5´a), 3.50 (ddd, *J* = 4.6 Hz, 5.5 Hz, 11.6 Hz, 1H, H-5´b), 2.58 (ddd, *J* = 5.9 Hz, 7.9 Hz, 13.2 Hz, 1H, H-2´a), 2.21 (ddd, *J* = 3.0 Hz, 6.1 Hz, 13.1 Hz, 1H, H-2´b) ppm. ^{13}C-NMR (100 MHz, DMSO-d_6): 161.1 (C-6), 160.2 (C-2), 154.2 (C-4), 138.2 (C-8), 114.4 (C-5), 88.1 (C-4´), 83.2 (C-1´), 71.2 (C-3´), 62.3 (C-5´), 53.6 (-O-CH_3), 40.0 (C-2´) ppm. HR-ESI-MS $[M+1H]^+$: m/z calculated: 282.1197, m/z found: 282.1188.

6.1.13. O^6-Ethyl-2´-deoxyguanosine 13b

Yield: 83 % (265 mg, 0.9 mmol). ^{1}H-NMR (400 MHz, DMSO-d_6): 8.07 (s, 1H, H-8), 6.38 (br s, 2H, -NH_2), 6.21 (dd, *J* = 6.1 Hz, 7.8 Hz, 1H, H-1´), 5.25 (d, *J* = 4.0 Hz, 1H, 3´-OH), 4.98 (t, *J* = 5.7 Hz, 1H, 5´-OH), 4.45 (q, *J* = 7.1 Hz, 2H, -O-CH_2-), 4.35 (dq, *J* = 3.1 Hz, 6.1 Hz, 1H, H-3´), 3.82 (dt, *J* = 2.9 Hz, 4.4 Hz, 1H, H-4´), 3.57 (dt, *J* = 5.1 Hz, 11.7 Hz, 1H, H-5´a), 3.50 (ddd, *J* = 4.4 Hz, 5.8 Hz, 11.7 Hz, 1H, H-5´b), 2.58 (ddd, *J* = 3.1 Hz, 5.8 Hz, 7.8 Hz, 1H, H-2´a), 2.20 (ddd, *J* = 3.0 Hz, 6.0 Hz, 13.1 Hz, 1H, H-2´b), 1.35 (t, *J* = 7.1 Hz, 3H, -CH_3) ppm. ^{13}C-NMR (100 MHz, DMSO-d_6): 160.8 (C-6), 160.2 (C-2),

154.3 (C-4), 138.1 (C-8), 114.4 (C-5), 88.1 (C-4´), 83.2 (C-1´), 71.2 (C-3´), 62.2 (C-5´), 62.0 (-O-CH_2-), 39.5 (C-2´), 15.0 (-CH_3) ppm. HR-ESI-MS $[M+1H]^+$: m/z calculated: 296.1353, m/z found: 296.1342.

6.1.14. O^6-Propyl-2´-deoxyguanosine 13c

Yield: 78 % (260 mg, 0.8 mmol). ^{1}H-NMR (600 MHz, DMSO-d_6): 8.07 (s, 1H, H-8), 6.39 (br s, 2H, -NH_2), 6.21 (dd, *J* = 6.2 Hz, 7.6 Hz, 1H, H-1´), 5.27 (d, *J* = 3.8 Hz, 1H, 3´-OH), 4.99 (t, *J* = 5.4 Hz, 1H, 5´-OH), 4.36 - 4.34 (m, 3H, H-3´, -O-CH_2-), 3.82 (dt, *J* = 2.6 Hz, 4.4 Hz, 1H, H-4´), 3.58 - 3.55 (m, 1H, H-5´a), 3.52 - 3.48 (m, 1H, H-5´b), 2.58 (ddd, *J* = 5.7 Hz, 7.9 Hz, 13.4 Hz, 1H, H-2´a), 2.20 (ddd, *J* = 2.9 Hz, 5.9 Hz, 13.1 Hz, 1H, H-2´b), 1.77 (q, *J* = 7.1 Hz, 2H, -C**H**$_2$-CH_3), 0.97 (t, *J* = 7.4 Hz, 3H, -CH_3) ppm. ^{13}C-NMR (150 MHz, DMSO-d_6): 160.9 (C-6), 160.2 (C-4), 154.3 (C-2), 138.0 (C-8), 114.4 (C-5), 88.0 (C-4´), 83.2 (C-1´), 71.2 (C-3´), 67.6 (-O-CH_2-), 62.2 (C-5´), 40.0 (C-2´), 22.3 (-O-CH_2-**C**H_2-), 10.8 (-CH_3) ppm. HR-ESI-MS $[M+1H]^+$: m/z calculated: 310.1524, m/z found: 310.1500.

6.1.15. O^6-*Iso*-propyl-2´-deoxyguanosine 13d

Yield: 91 % (304 mg, 1.0 mmol). ^{1}H-NMR (400 MHz, DMSO-d_6): 8.05 (s, 1H, H-8), 6.34 (br s, 2H, -NH_2), 6.20 (dd, *J* = 6.5 Hz, 7.9 Hz, 1H, H-1´), 5.48 (h, *J* = 6.1 Hz, 1H, -O-CH-), 5.25 (d, *J* = 4.0 Hz, 1H, 3´-OH), 4.98 (t, *J* = 5.7 Hz, 1H, 5´-OH), 4.35 (dq, *J* = 2.9 Hz, 6.0 Hz, 1H, H-3´), 3.82 (dt, *J* = 2.7 Hz, 4.5 Hz, 1H, H-4´), 3.57 (dt, *J* = 5.0 Hz, 11.7 Hz, 1H, H-5´a), 3.50 (ddd, *J* = 4.4 Hz, 5.8 Hz, 11.7 Hz, 1H, H-5´b), 2.57 (ddd, *J* = 5.7 Hz, 7.9 Hz, 13.3 Hz, 1H, H-2´a), 2.20 (ddd, *J* = 2.9 Hz, 5.9 Hz, 13.1 Hz, 1H, H-2´b), 1.34 (d, *J* = 2.5 Hz, 3H, -CH_3), 1.33 (d, *J* = 2.5 Hz, 3H, -CH_3) ppm. ^{13}C-NMR (100 MHz, DMSO-d_6): 160.4 (C-6), 160.2 (C-2), 154.3 (C-4), 138.0 (C-8), 114.6 (C-5), 88.1 (C-4´), 83.3 (C-1´), 71.3 (C-3´), 68.7 (-O-CH-), 62.2 (C-5´), 40.0 (C-2´), 22.4 (-CH_3), 22.4 (-CH_3) ppm. HR-ESI-MS $[M+1H]^+$: m/z calculated: 310.1510, m/z found: 310.1500.

6.1.16. General procedure B:[132]

2´,3´-Di-*O*-acetyl-6-deoxy-6-chloro-2´-deoxyguanosine (1.0 eq.) was dissolved in an aqueous solution of the respective amine and stirred at room temperature overnight. The reaction mixture was concentrated *in vacuo* and the pH was adjusted to pH 7 using acetic acid. The aqueous solution was extracted with ethyl acetate and the organic layer was concentrated to dryness. The remaining crude residue was further purified by column flash chromatography using methylene chloride with 4 % methanol. The products were obtained as colourless foams.

6.1.17. 6-Methylamino-2´deoxyguanosine 16a

Yield: 77 % (175 mg, 0.624 mmol). ^{1}H-NMR (400 MHz, DMSO-d_6): 7.89 (s, 1H, H-8), 7.22 (br s, 1H, -NH-), 6.17 (dd, *J* =5.9 Hz, 8.1 Hz, 1H, H-1´), 5.79 (br s, 2H, -NH_2), 5.24 (t, *J* = 5.4 Hz, 1H, 5´-OH),

5.22 (d, *J* = 3.9 Hz, 1H, 3´-OH), 4.34 (dq, *J* = 2.8 Hz, 5.6 Hz, 1H, H-3´), 3.83 (dt, *J* = 2.4 Hz, 4.1 Hz, 1H, H-4´), 3.58 (dt, *J* = 4.2 Hz, 12.0 Hz, 1H, H-5´a), 3.53 - 3.47 (m, 1H, H-5´b), 2.88 (br s, 3H, -CH_3), 2.58 (ddd, *J* = 5.6 Hz, 8.3 Hz, 13.5 Hz, 1H, H-2´a), 2.16 (ddd, *J* = 2.6 Hz, 5.7 Hz, 13.0 Hz, 1H, H-2´b) ppm. ^{13}C-NMR (100 MHz, DMSO-d_6): 160.4 (C-6), 155.7 (C-2), 150.7 (C-4), 135.7 (C-8), 114.2 (C-5), 88.1 (C-4´), 83.5 (C-1´), 71.5 (C-3´), 62.4 (C-5´), 39.9 (C-2´), 31.1 (-CH_3) ppm. HR-ESI-MS $[M+1H]^+$: m/z calculated: 281.1357, m/z found: 281.1346.

6.1.18. 6-Dimethylamino-2´-deoxyguanosine 16h

Yield: 97 % (309 mg, 1.05 mmol). ^{1}H-NMR (400 MHz, DMSO-d_6): 7.93 (s, 1H, H-8), 6.19 (dd, *J* = 6.0 Hz, 8.0 Hz, 1H, H-1´), 5.78 (br s, 2H, -NH_2), 5.23 (d, *J* = 3.9 Hz, 1H, 3´-OH), 5.16 (t, *J* = 5.6 Hz, 1H, 5´-OH), 4.34 (dq, *J* = 2.8 Hz, 5.8 Hz, 1H, H-3´), 3.82 (dt, *J* = 2.7 Hz, 4.2 Hz, 1H, H-4´), 3.57 (dt, *J* = 4.7 Hz, 11.7 Hz, 1H, H-5´a), 3.50 (ddd, *J* = 4.2 Hz, 6.0 Hz, 11.9 Hz, 1H, H-5´b), 3.36 (br s, 6H, 2 x –CH_3), 2.55 (ddd, *J* = 5.8 Hz, 8.2 Hz, 13.4 Hz, 1H, H-2´a), 2.17 (ddd, *J* = 2.8 Hz, 6.0 Hz, 13.0 Hz, 1H, H-2´b) ppm. ^{13}C-NMR (100 MHz, DMSO-d_6): 159.8 (C-6), 155.2 (C-2), 152.7 (C-4), 135.1 (C-8), 114.3 (C-5), 88.1 (C-4´), 83.3 (C-1´), 71.4 (C-3´), 62.4 (C-5´), 40.1 (C-2´), 37.7 (-CH_3) ppm. HR-ESI-MS $[M+1H]^+$: m/z calculated: 295.1513, m/z found: 295.1497.

6.1.19. 6-Methylethylamino-2´-deoxyguanosine 16i

Yield: 91 % (380 mg, 1.23 mmol). ^{1}H-NMR (400 MHz, $CDCl_3$): 7.92 (s, 1H, H-8), 6.19 (dd, *J* = 6.0 Hz, 8.0 Hz, 1H, H-1´), 5.76 (s, 2H, -NH_2), 5.23 (d, *J* = 3.9 Hz, 1H, 3´-OH), 5.17 (t, *J* = 5.5 Hz, 1H, 5´-OH), 4.35 (dq, *J* = 2.9 Hz, 5.9 Hz, 1H, H-3´), 3.98 (br s, 2H, -N-CH_2-), 3.82 (dt, *J* = 2.5 Hz, 4.4 Hz, 1H, H-4´), 3.57 (dt, *J* = 4.6 Hz, 11.8 Hz, 1H, H-5´a), 3.50 (ddd, *J* = 4.1 Hz, 6.0 Hz, 11.7 Hz, 1H, H-5´a), 3.26 (br s, 3H, -N-CH_3), 2.56 (ddd, *J* = 5.8 Hz, 8.3 Hz, 13.7 Hz, H-2´a), 2.17 (ddd, *J* = 2.8 Hz, 6.0 Hz, 13.0 Hz, 1H, H-2´b), 1.12 (t, *J* = 6.9 Hz, 3H, -CH_2-C**H**$_3$) ppm. ^{13}C-NMR (100 MHz, $CDCl_3$): 159.8 (C-2), 154.4 (C-6), 152.6 (C-4), 135.1 (C-8), 114.1 (C-5), 88.1 (C-4´), 83.3 (C-1´), 71.3 (C-3´), 62.4 (C-5´), 44.5 (-N-CH_2-), 40.0 (C-2´), 35.3 (N-CH_3), 12.1 (-CH_3) ppm. HR-ESI-MS $[M+1H]^+$: m/z calculated: 309.1670, m/z found: 309.1670.

6.1.20. 6-Diethylamino-2´-deoxyguanosine 16l

Yield: 71 % (310 mg, 0.96 mmol). ^{1}H-NMR (400 MHz, $CDCl_3$): 7.47 (s, 1H, H-8), 6.17 (dd, *J* = 5.7 Hz, 9.6 Hz, 1H, H-1´), 4.72 - 4.68 (m, 3H, H-3´, -NH_2), 4.17 - 4.15 (m, 1H, H-4´), 3.95 (dd, 1H, *J* = 1.5 Hz, 12.7 Hz, H-5´a), 3.87 (br s, 4H, 2 x -CH_2-), 3.74 (dd, 1H, *J* = 1.6 Hz, 12.6 Hz, H-5´b), 3.04 (ddd, *J* = 5.2 Hz, 9.6 Hz, 14.7 Hz, 1H, H-2´a), 2.21 - 2.16 (m, 1H, H-2´b), 1.21 (t, *J* = 7.0 Hz, 6H, 2 x -CH_3) ppm. ^{13}C-NMR (100 MHz, $CDCl_3$): 158.3 (C-6), 154.3 (C-2), 150.6 (C-4), 135.9 (C-8), 116.1 (C-5), 89.3 (C-4´), 87.6 (C-1´), 73.5 (C-3´), 63.7 (C-5´), 42.9 (2 x -CH_2), 40.0 (C-2´), 13.5 (2 x -CH_3) ppm. HR-ESI-MS $[M+1H]^+$: m/z calculated: 323.1826, m/z found: 323.1826.

6.1.21. 6-Thio-2´-deoxyguanosine[133]

2´-Deoxyguanosine (0.5 g, 2.0 mmol, 1.0 eq.) was dried by co-evaporation with pyridine, suspended in dry pyridine (40 ml) and cooled with an ice bath under nitrogen atmosphere. Trifluoroacetic anhydride (2.3 ml, 16.0 mmol, 8.0 eq.) was added drop wise at 0°C and the suspension was stirred for 40 min at room temperature. Following, a suspension of sodium hydrogen sulfide (3.4 g, 60.0 mmol, 30.0 eq.) in anhydrous DMF (60 ml) was added in portions and the suspension was stirred for further 24 h. The reaction mixture was poured into 100 ml of 0.16 M ammonium bicarbonate solution under vigorous stirring. Following, the mixture was concentrated to dryness and the residue was triturated with methanol and filtered. The filtrate was concentrated to dryness and purified by column flash chromatography using dichloromethane with 10 % methanol. The product was precipitated from the resulting brown oil by addition of diethyl ether and collected by filtration in 42 % (240 mg, 0.85 mmol) yield. ^{1}H-NMR (400 MHz, DMSO-d_6): 11.93 (br s, 1H, -NH-), 8.10 (s, 1H, H-8), 6.80 (br s, 2H, -NH_2), 6.11 (dd, *J* = 6.3 Hz, 7.4 Hz, 1H, H-1´), 5.27 (d, *J* = 4.0 Hz, 1H, 3´-OH), 4.92 (t, *J* = 5.5 Hz, 1H, 5´-OH), 4.34 (dq, *J* = 3.4 Hz, 4.4 Hz, 1H, H-3´), 3.81 (dt, *J* = 2.9 Hz, 4.5 Hz, 1H, H-4´), 3.56 (ddd, *J* = 4.7 Hz, 5.6 Hz, 11.7 Hz, 1H, H-5´a), 3.50 (ddd, *J* = 4.5 Hz, 5.4 Hz, 11.7 Hz, 1H, H-5´b), 2.55 - 2.48 (m, 1H, H-2´a), 2.21 (dd, *J* = 3.2 Hz, 5.9 Hz, 13.9 Hz, 1H, H-2´b) ppm. ^{13}C-NMR (100 MHz, DMSO-d_6): 175.6 (C-6), 153.5 (C-2), 147.9 (C-4), 138.7 (C-8), 128.8 (C-5), 88.2 (C-4´), 83.1 (C-1´), 71.1 (C-3´), 62.1 (C-5´), 40.0 (C-2´) ppm. HR-ESI-MS $[M+1H]^+$: m/z calculated: 284.0812, m/z found: 284.0814.

6.1.22. 2-Amino-6-methylthiopurin 23a [113b]

Sodium methoxide (2.1 g, 30.0 mmol, 10.0 eq.) and 2-amino-6-chloropurine (0.5 g, 3.0 mmol, 1.0 eq.) were dissolved in abs. DMF (20 ml) and stirred at room temperature overnight. After 16 h, the suspension was concentrated *in vacuo* and the remaining solid was dissolved in 10 ml water. The aqueous solution was neutralized using acetic acid and extracted with ethyl acetate. The organic layer was dried over $MgSO_4$, concentrated to dryness and the crude residue was purified using column flash-chromatography (DCM/MeOH 10/1) to yield 55 % (0.3 g, 1.7 mmol). ^{1}H-NMR (400 MHz, DMSO-d_6): 12.48 (br s, 1H, -NH-), 7.87 (s, 1H, H-8), 6.31 (br s, 2H, -NH_2), 2.56 (s, 3H, -S-CH_3) ppm. ^{13}C-NMR (100 MHz, DMSO-d_6): 160.1 (C-6), 159.6 (C-2), 151.9 (C-4), 139.1 (C-8), 124.4 (C-5), 11.2 (-S-CH_3) ppm. HR-ESI-MS $[M+1H]^+$: m/z calculated: 180.0338, m/z found: 180.0347.

6.1.23. 2-Amino-6-ethylthiopurin 23b [113b]

Potassium *tert*-butoxide (2.7 g, 24.0 mmol, 10.0 eq.) was dissolved in 20 ml abs. DMF. Ethanethiole (1.7 ml, 24.0 mmol, 10.o eq.) was added and the reaction mixture was stirred at room temperature for 30 min. Following 2-amino-6-chloropurine (0.4 g, 2.4 mmol, 1.0 eq.) was added and the reaction mixture was refluxed overnight in a sealed vessel. After 16 h, the suspension was concentrated *in vacuo* and the remaining solid was dissolved in 10 ml water. The aqueous solution was neutralized using acetic acid and extracted with ethyl acetate. The organic layer was dried over

$MgSO_4$, concentrated to dryness and the crude residue was purified using column flash-chromatography using ethyl acetate as solvent to yield 49 % (0.23 g, 1.17 mmol). ^{1}H-NMR (400 MHz, DMSO-d_6): 12.47 (br s, 1H, H-9), 7.87 (s, 1H, H-8), 6.29 (br s, 2H, -NH_2), 3.25 (q, *J* = 7.3 Hz, 2H, -S-CH_2-), 1.32 (t, *J* = 7.3 Hz, 3H, -CH_3) ppm. ^{13}C-NMR (100 MHz, DMSO-d_6): 160.1 (C-6), 159.3 (C-2), 152.1 (C-4), 139.1 (C-8), 124.3 (C-5), 22.4 (-S-CH_2-), 15.6 (-CH_3) ppm. HR-ESI-MS $[M+1H]^+$: m/z calculated: 196.0651, m/z found: 196.0674.

6.1.24. 3´,5´-Di-*O*-toluoyl-6-methylthio-2´-deoxyguanosine 24a [134]

2-Amino-6-methylthiopurin (220 mg, 1.20 mmol, 1.0 eq.) and sodium hydride (60% in mineral oil) (51 mg, 1.34 mmol, 1.1 eq.) were suspended in 30 ml abs. CH_3CN and stirred for 30 min at room temperature. After 30 min 1-chloro-2-deoxy-3,5-di-*O*-toluoyl-α-D-ribofuranose (521 mg, 1.34 mmol, 1.1 eq.) was added and the suspension was stirred at room temperature for 20 hours. The suspension was filtered and the filtrate was concentrated *in vacuo*. The crude residue was purified by column flash chromatography (*n*-hexane/ ethyl acetate 3/1) to yield 64 % (0.41 g, 0.77 mmol). ^{1}H-NMR (400 MHz, $CDCl_3$): 7.96 (d, *J* = 8.1 Hz, 2H, $H_{toluoyl}$), 7.90 (d, *J* = 8.2 Hz, 2H, $H_{toluoyl}$), 7.77 (s, 1H, H-8), 7.27 (d, *J* = 8.0 Hz, 2H, $H_{toluoyl}$), 7.22 (d, *J* = 8.0 Hz, 2H, $H_{toluoyl}$), 6.38 (dd, *J* = 5.9 Hz, 8.3 Hz, 1H, H-1´), 5.79 (dt, *J* = 2.2 Hz, 6.2 Hz, 1H, H-3´), 4.98 (s, 2H, -NH_2), 4.81 (dd, *J* = 4.1 Hz, 11.4 Hz, 1H, H-5´a), 4.67 - 4.60 (m, 2H, H-4´, H-5´b), 3.15 (ddd, *J* = 6.2 Hz, 8.4 Hz, 14.4 Hz, 1H, H-2´a), 2.71 (ddd, *J* = 2.1 Hz, 5.9 Hz, 14.1 Hz, 1H, H-2´b), 2.61 (s, 3H, -S-CH_3), 2.44 (s, 3H, -CH_3), 2.40 (s, 3H, -CH_3) ppm. ^{13}C-NMR (100 MHz, $CDCl_3$): 166.3 (-C(O)O-), 165.9 (-C(O)O-), 162.5 (C-2), 158.9 (C-4), 149.9 (C-6), 144.5 (-**C**-CH_3), 144.1 (-**C**-CH_3), 138.0 (C-8), 129.8 (2 x $C_{toluoyl}$), 129.6 (2 x $C_{toluoyl}$), 129.3 (2 x $C_{toluoyl}$), 129.2 (2 x $C_{toluoyl}$), 126.7 (-**C**-C(O)O-), 126.5 (-**C**-C(O)O-), 126.3 (C-5), 84.5 (C-1´), 82.7 (C-4´), 75.2 (C-3´), 64.0 (C-5´), 37.0 (C-2´), 21.7 (-CH_3), 21.6(-CH_3), 11.5 (-S-CH_3) ppm. HR-ESI-MS $[M+1H]^+$: m/z calculated: 534.1806, m/z found: 534.1764.

6.1.25. 6-Methylthio-2´-deoxyguanosine 21a [108]

3´,5´-Di-O-toluoyl-6-methylthio-2´-deoxyguanosine (370 mg, 0.75 mmol, 1.0 eq.) was dissolved in 50 ml 7 N ammonia in methanol and stirred at 4 °C for 16 h. Following the solution was concentrated to dryness and the crude residue was purified by column flash chromatography using *n*-hexane and ethyl acetate (3/1) to yield 99 % (0.22 g, 0.74 mmol). ^{1}H-NMR (400 MHz, $CDCl_3$): 7.66 (s, 1H, H-8), 6.50 (br s, 1H, -NH_2 a), 6.23 (dd, *J* = 5.6 Hz, 9.5 Hz, 1H, H-1´), 5.36 (br s, 1H, -NH_2 b), 4.99 (br s, 2H, 3´-OH, 5´-OH), 4.76 (d, *J* = 5.1 Hz, 1H, H-3´), 4.21 - 4.20 (m, 1H, H-4´), 3.98 (dd, *J* = 1.8 Hz, 12.8 Hz, 1H, H-5´-a), 3.77 (d, *J* = 12.6 Hz, 1H, H-5´b), 3.03 (ddd, *J* = 5.2 Hz, 9.6 Hz, 14.6 Hz, 1H, H-2´a), 2.61 (s, 3H, S-CH_3), 2.26 (dd, *J* = 5.6 Hz, 13.6 Hz, 1H, H-2´b) ppm. ^{13}C-NMR (100 MHz, $CDCl_3$): 159.1 (C-6), 153.2 (C-4), 143.7 (C-2), 134.4 (C-8), 122.3 (C-5), 84.5 (C-4´), 82.2 (C-1´), 68.7 (C-3´), 58.6 (C-5´), 35.8 (C-2´), 6.8 (-S-CH_3) ppm. HR-ESI-MS $[M+1H]^+$: m/z calculated: 298.0968, m/z found: 298.0947.

6.1.26. 3´,5´-Di-*O*-toluoyl-6-ethylthio-2´-deoxyguanosine 24b [134]

2-Amino-6-ethylthiopurin (270 mg, 1.38 mmol, 1.0 eq.) and sodium hydride (60% in mineral oil) (58 mg, 1.52 mmol, 1.1 eq.) were suspended in 30 ml abs. CH_3CN and stirred for 30 min at room temperature. After 30 min 1-chloro-2-deoxy-3,5-di-*O*-toluoyl-α-D-ribofuranose (591 mg, 1.52 mmol, 1.1 eq.) was added and the suspension was stirred at room temperature for 20 hours. The suspension was filtered and the filtrate was concentrated *in vacuo*. The crude residue was purified by column flash chromatography (*n*-hexane/ ethyl acetate 3/1) to yield 57 % (0.43 g, 0.79 mmol). ^{1}H-NMR (400 MHz, $CDCl_3$): 7.97 - 7.95 (m, 2H, $H_{toluoyl}$), 7.91 - 7.89 (m, 2H, $H_{toluoyl}$), 7.76 (s, 1H, H-8), 7.28 (d, *J* = 8.3 Hz, 2H, $H_{toluoyl}$), 7.22 (d, *J* = 8.1 Hz, 2H, $H_{toluoyl}$), 6.38 (dd, *J* = 5.9 Hz, 8.4 Hz, 1H, H-1´), 5.79 (dt, *J* = 2.1 Hz, 6.1 Hz, 1H, H-3´), 4.94 (br s, 2H, -NH_2), 4.81 (dd, *J* = 4.1 Hz, 11.4 Hz, 1H, H-5´a), 4.67 - 4.60 (m, 2H, H-4´, H-5´b), 3.28 (q, *J* = 7.4 Hz, 2H, -S-CH_2-), 3.16 (ddd, *J* = 6.2 Hz, 8.4 Hz, 14.4 Hz, 1H, H-2´-a), 2.70 (ddd, *J* = 5.9 Hz, 14.2 Hz, 2.1 Hz, 1H, H-2´b), 2.44 (s, 3H, -CH_3), 2.40 (s, 3H, -CH_3), 1.40 (t, *J* = 7.4 Hz, 3H, -S-CH_2-C**H**$_3$) ppm. ^{13}C-NMR (100 MHz, $CDCl_3$): 166.3 (-C(O)O-), 165.9 (-C(O)O-), 162.2 (C-6), 158.9 (C-2), 150.2 (C-4), 144.4 ($C_{toluoyl}$), 144.1 ($C_{toluoyl}$), 138.0 (C-8), 129.8 (2 x $C_{toluoyl}$), 129.7 (2 x $C_{toluoyl}$), 129.3 (2 x $C_{toluoyl}$), 129.3 (2 x $C_{toluoyl}$), 126.8 (-C(O)O-**C**-), 126.5 (-C(O)O-**C**-), 126.3 (C-5), 84.5 (C-1´), 82.6 (C-4´), 75.3 (C-3´), 64.0 (C-5´), 37.1 (C-2´), 23.0 (-S-CH_2-), 21.8 (-$CH_{3\ toluoyl}$), 21.7 (-$CH_{3\ toluoyl}$), 14.8 (-CH_3) ppm.

6.1.27. 6-Ethylthio-2´-deoxyguanosine 21b [108]

3´,5´-Di-O-toluoyl-6-ethylthio-2´-deoxyguanosine (430 mg, 0.79 mmol, 1.0 eq.) was dissolved in 50 ml 7 N ammonia in methanol and stirred at 4°C for 16 h. Following the solution was concentrated to dryness and the crude residue was purified by column flash chromatography using *n*-hexane and ethyl acetate (3/1) to yield 61 % (0.15 g, 0.48 mmol). ^{1}H-NMR (400 MHz, DMSO-d_6): 8.14 (s, 1H, H-8), 6.48 (br s, 2H, -NH_2), 6.21 (dd, *J* = 6.0 Hz, 7.7 Hz, 1H, H-1´), 5.27 (d, *J* = 4.0 Hz, 1H, 3´-OH), 4.96 (t, *J* = 5.5 Hz, 1H, 5´-OH), 4.35 (dq, *J* = 2.9 Hz, 6.3 Hz, 1H, H-3´), 3.82 (dt, *J* = 2.7 Hz, 4.6 Hz, 1H, H-4´), 3.57 (dt, *J* = 5.1 Hz, 11.7 Hz, 1H, H-5´a), 3.50 (m, 1H, H-5´b), 3.25 (q, *J* = 7.3 Hz, 2H, -S-CH_2-), 2.59 (ddd, *J* = 5.8 Hz, 7.7 Hz, 13.3 Hz, 1H, H-2´a), 2.22 (ddd, *J* = 3.1 Hz, 6.0 Hz, 13.1 Hz, 1H, H-2´b), 1.32 (t, *J* = 7.3 Hz, 3H, -CH_3) ppm. ^{13}C-NMR (100 MHz, DMSO-d_6): 160.2 (C-6), 159.9 (C-2), 151.0 (C-4), 139.0 (C-8), 124.7 (C-5), 88.1 (C-4´), 83.2 (C-1´), 71.1 (C-3´), 62.1 (C-5´), 40.0 (C-2´), 22.2 (-S-CH_2-), 15.5 (-CH_3) ppm. HR-ESI-MS $[M+1H]^+$: m/z calculated: 312.1125, m/z found: 312.1101.

6.1.28. 8-Bromo-3´,5´-O-(1,1,3,3-tetrakis(-*iso*-propyl)-1,3-disiloxanediyl) -2´-deoxyguanosine 27 [119]

8-Bromo-2´-deoxyguanosine (2.9 g, 8.37 mmol, 1.0 eq.) was resuspended in dry pyridine (50 ml) and cooled to 0°C. 1,3-Dichloro-1,1,3,3-tetra-*iso*-propyl-disiloxane (2.9 g, 9.18 mmol, 1.1 eq.) was added dropwise, the reaction mixture was allowed to warm to room temperature and the reaction was stirred over night. The pyridine was removed *in vacuo*, the residue was poured into water and

following extracted with chloroform. The combined organic layers were successively washed with 10 % of HCl, saturated sodium bicarbonate and brine. After drying of the organic layers over sodium sulfate, the solution was evaporated to dryness and purified using column flash-chromatography, eluting with 20 % *n*-hexane in ethyl acetate. Yield: 87 % (4.28 g, 7.28 mmol). ^{1}H-NMR (400 MHz, DMSO-d_6): 10.89 (br s, 1H, -NH-), 6.28 (br s, 2H, -NH_2), 6.17 (dd, *J* = 5.0 Hz, 8.4 Hz, 1H, H-1´), 4.89 (dt, *J* = 5.9 Hz, 8.4 Hz, 1H, H-3´), 3.94 (m, 1H, H-5´a), 3.82 (m, 2H, H-4´, H-5´b), 3.15 (ddd, *J* = 5.2 Hz, 8.6 Hz, 13.6 Hz, 1H, H-2´a), 2.40 (ddd, *J* = 6.2 Hz, 8.5 Hz, 14.1 Hz, 1H, H-2´b), 1.03 (m, 28 H, 4 x –Si-*i*Pr) ppm. ^{13}C-NMR (100 MHz, DMSO-d_6): 155.9 (C-6), 154.0 (C-2), 152.4 (C-4), 119.9 (C-8), 117.6 (C-5), 85.0 (C-4´), 83.0 (C-1´), 73.1 (C-3´), 63.4 (C-5´), 37.7 (C-2´), 17.8 (-CH_3), 17.7 (3 x –CH_3), 17.6 (-CH_3), 17.5 (-CH_3), 17.4 (-CH_3), 17.3 (-CH_3), 13.2 (-**C**(-CH_3)$_2$), 13.1 (-**C**(-CH_3)$_2$), 12.8 (-**C**(-CH_3)$_2$), 12.5 (-**C**(-CH_3)$_2$) ppm. HR-ESI-MS $[M-1H]^-$: m/z calculated: 586.1511, m/z found: 586.1549.

6.1.29. 8-Bromo-3´,5´-*O*-(1,1,3,3-tetrakis(isopropyl)-1,3-disiloxanediyl)-*O*^{6}-methyl-2´-deoxyguanosine 28a [119]

Triphenylphosphine (285 mg, 1.08 mmol, 1.1 eq.) and 8-Bromo-3´,5´-*O*-(1,1,3,3-tetrakis(isopropyl)-1,3-disiloxanediyl)-2´-deoxyguanosine (0.58 g, 0.98 mmol, 1.0 eq.) were dried *in vacuo* for 30 min, dissolved in abs. 1,4-dioxane (10 ml) and cooled to 0°C using an ice bath. Methanol (63 µl, 1.08 mmol, 1.1 eq.) was added followed by dropwise addition of diethyl azodicarboxylate (40 % solution in toluene, 431 µl, 1.08 mmol, 1.1 eq.). The reaction mixture was allowed to warm to room temperature and was stirred for 1 h. The solvents were removed under reduced pressure and the crude residue was purified using column flash chromatography eluting with 10 % ethyl acetate in *n*-hexane. Yield: 57 % (345 mg, 0.56 mmol). ^{1}H-NMR (400 MHz, DMSO-d_6): 6.22 (br s, 2H, -NH_2), 6.20 (dd, *J* = 4.1 Hz, 8.2 Hz, 1H, H-1´), 5.02 (dt, *J* = 2.3 Hz, 7.1 Hz, 1H, H-3´), 3.95 (s, 3H, -O-CH_3), 3.91 (m, 1H, H-5´a), 3.79 (m, 2H, H-4´, H-5´b), 3.24 (ddd, *J* = 4.1 Hz, 8.3 Hz, 13.3 Hz, 1H, H-2´a), 2.47 (ddd, *J* = 6.9 Hz, 8.7 Hz, 15.7 Hz, 1H, H-2´b), 1.03 (m, 28 H, 4 x –Si-*i*Pr) ppm. ^{13}C-NMR (100 MHz, DMSO-d_6): 159.9 (C-6, C-2), 154.6 (C-4), 124.4 (C-8), 114.8 (C-5), 84.9 (C-4´), 83.5 (C-1´), 73.0 (C-3´), 63.4 (C-5´), 53.9 (-O-CH_3), 37.4 (C-2´), 17.8 (-CH_3), 17.7 (-CH_3), 17.7 (-CH_3), 17.6 (-CH_3), 17.6 (-CH_3), 17.5 (-CH_3), 17.4 (-CH_3), 17.3 (-CH_3), 13.3 (-**C**(-CH_3)$_2$), 13.1 (-**C**(-CH_3)$_2$), 12.8 (-**C**(-CH_3)$_2$), 12.5 (-**C**(-CH_3)$_2$) ppm. HR-ESI-MS $[M+1H]^+$: m/z calculated: 602.1824, m/z found: 602.1796.

6.1.30. 8-Bromo-3´,5´-*O*-(1,1,3,3-tetrakis(isopropyl)-1,3-disiloxanediyl)-*O*^{6}-ethyl-2´-deoxyguanosine 28b [119]

Triphenylphosphine (285 mg, 1.08 mmol, 1.1 eq.) and 8-Bromo-3´,5´-*O*-(1,1,3,3-tetrakis(isopropyl)-1,3-disiloxanediyl)-2´-deoxyguanosine (0.58 g, 0.98 mmol, 1.0 eq.) were dried *in vacuo* for 30 min, dissolved in abs. 1,4-dioxane (10 ml) and cooled to 0°C using an ice bath. Ethanol (63 µl, 1.08 mmol, 1.1 eq.) was added followed by dropwise addition of diethyl azodicarboxylate (40 % solution in toluene, 431 µl, 1.08 mmol, 1.1 eq.). The reaction mixture was allowed to warm to room temperature and was stirred for 1 h. The solvents were removed under reduced pressure and the crude residue

was purified using column flash chromatography eluting with 10 % ethyl acetate in *n*-hexane. Yield: 57 % (345 mg, 0.56 mmol). ^{1}H-NMR (400 MHz, DMSO-d_6): 6.20 (dd, *J* = 4.1 Hz, 8.6 Hz, 1H, H-1´), 6.18 (br s, 2H, -NH_2), 5.03 (m, 1H, H-3´), 4.44 (q, *J* = 7.1 Hz, 2H, -O-CH_2-), 3.90 (m, 1H, H-5´a), 3.79 (m, 2H, H-4´, H-5´b), 3.24 (ddd, *J* = 4.1 Hz, 8.3 Hz, 13.5 Hz, 1H, H-2´a), 2.45 (m, 1H, H-2´b), 1.34 (t, *J* = 7.1 Hz, 3H, -O-CH_2-C**H**$_3$), 1.04 (m, 28 H, 4 x –Si-*i*Pr) ppm. ^{13}C-NMR (100 MHz, DMSO-d_6): 159.9 (C-6), 159.8 (C-2), 154.6 (C-4), 124.3 (C-8), 114.8 (C-5), 84.9 (C-4´), 83.5 (C-1´), 73.1 (C-3´), 63.4 (C-5´), 62.3 (-O-CH_2-), 37.4 (C-2´), 17.8 (-CH_3), 17.7 (-CH_3), 17.7 (-CH_3), 17.6 (-CH_3), 17.6 (-CH_3), 17.5 (-CH_3), 17.4 (-CH_3), 17.3 (-CH_3), 14.9 (-O-CH_2-**C**H_3), 13.3 (-**C**(-CH_3)$_2$), 13.1 (-**C**(-CH_3)$_2$), 12.8 (-**C**(-CH_3)$_2$), 12.5 (-**C**(-CH_3)$_2$) ppm. HR-ESI-MS $[M+1H]^+$: m/z calculated: 616.1981, m/z found: 616.1935.

6.1.31. 8-Bromo-*O*^{6}-methyl-2´-deoxyguanosine 29a [131]

8-Bromo-3´,5´-*O*-(1,1,3,3-tetrakis(isopropyl)-1,3-disiloxanediyl)-*O*^{6}-methyl-2´-deoxy-guanosine (180 mg, 0.30 mmol, 1.0 eq.) was dissolved in abs. THF (5 ml). Triethylamine trihydrofluoride (102 µl, 0.63 mmol, 2.1 eq.) was added and the reaction mixture was stirred over night at room temperature. The solution was concentrated under reduced pressure and the crude residue was purified using column flash chromatography eluting with 1% methanol in methylene chloride. Yield: 80 % (86 mg, 0.23 mmol). ^{1}H-NMR (400 MHz, DMSO-d_6): 6.52 (br s, 2H, -NH_2), 6.21 (dd, *J* = 6.6 Hz, 7.9 Hz, 1H, H-1´), 5.26 (d, *J* = 4.2 Hz, 1H, 3´-OH), 4.93 (dd, *J* = 5.3 Hz, 6.5 Hz, 1H, 5´-OH), 4.43 (dt, *J* = 3.2 Hz, 6.4 Hz, 1H, H-3´), 3.82 (dt, *J* = 2.8 Hz, 5.5 Hz, 1H, H-4´), 3.63 (ddd, *J* = 5.3 Hz, 6.2, Hz, 11.6 Hz, 1H, H-5´a), 3.49 (ddd, *J* = 5.7 Hz, 6.6 Hz, 11.9 Hz, 1H, H-5´b), 3.28 (m, 1H, H-2´a), 2.12 (ddd, *J* = 2.9 Hz, 6.8 Hz, 13.2 Hz, 1H, H-2´b) ppm. HR-ESI-MS $[M+1H]^+$: m/z calculated: 360.0302, m/z found: 360.0293.

6.1.32. 8-Bromo-*O*^{6}-ethyl-2´-deoxyguanosine 29b [131]

8-Bromo-3´,5´-*O*-(1,1,3,3-tetrakis(isopropyl)-1,3-disiloxanediyl)-*O*^{6}-ethyl-2´-deoxyguanosine (580 mg, 0.94 mmol, 1.0 eq.) was dissolved in abs. THF (10 ml). Triethylamine trihydrofluoride (323 µl, 1.97 mmol, 2.1 eq.) was added and the reaction mixture was stirred over night at room temperature. The solution was concentrated under reduced pressure and the crude residue was purified using column flash chromatography eluting with 1% methanol in methylene chloride. Yield: quantitative (352 mg, 0.94 mmol). ^{1}H-NMR (400 MHz, DMSO-d_6): 6.20 (dd, *J* = 6.6 Hz, 8.0 Hz, 1H, H-1´), 4.97 (m, 1H, 3´-OH), 4.59 (dd, *J* = 2.4 Hz, 5.7 Hz, 1H, 5´-OH), 4.48 (m, 1H, H-3´), 4.43 (q, *J* = 7.1 Hz, 2H, -O-CH_2-), 3.83 (dt, *J* = 2.8 Hz, 5.5 Hz, 1H, H-4´), 3.65 (m, 1H, H-5´a), 3.49 (ddd, *J* = 2.1 Hz, 4.4 Hz, 11.6 Hz, 1H, H-5´b), 3.27 (m, 1H, H-2´a), 2.12 (ddd, *J* = 2.9 Hz, 6.7 Hz, 13.2 Hz, 1H, H-2´b), 1.34 (t, *J* = 5.3 Hz, 3H, -CH_3) ppm. HR-ESI-MS $[M+1H]^+$: m/z calculated: 374.0458, m/z found: 374.0431.

6.1.33. 6-azido-hexylphosphate[97, 135]

6-Bromo-hexanol (1.67 g, 9.2 mmol, 1.0 eq.) and sodium azide (1.2 g, 18.4 mmol, 2.0 eq.) were dissolved in abs. DMF (50 ml). The reaction mixture was stirred over night at 80°C. After 16 h the

solution was cooled to room temperature and poured in 200 ml water. The aqueous solution was extracted with ethyl acetate three times. The organic layer was washed with brine twice, dried over $MgSO_4$ and concentrated to dryness. The crude residue was used for the phosphorylation without further purification.

Phosphorous oxychloride (855 µl, 9.2 mmol, 1.0 eq.) was dissolved in abs. THF (30 ml) and cooled to 0°C using an ice bath. Triethylamine (1.28 ml, 9.2 mmol, 1.0 eq.) and 6-azido-hexanol were added and the resulting suspension was stirred at 0°C. After 4 hours the reaction was quenched using 0.1 M TEAB (40 ml) and stirred at room temperature for 30 min. Following the solution was concentrated *in vacuo*. The crude residue was dissolved in 1 M NaOH and washed with ethyl ether three times. The pH of the aqueous layer was adjusted to pH 1 using hydrochloric acid and extracted with ethyl ether three times. The combined organic layers were dried over $MgSO_4$ and concentrated to dryness. Yield: 78 % (1.60 g, 7.18 mmol). ^{1}H-NMR (400 MHz, DMSO-d_6): 9.73 (br s, 2H, -OP(O)(OH)$_2$), 4.02 (q, J = 6.8 Hz, 2H, -C**H**$_2$-OPO$_3$H$_2$), 3.26 (t, J = 6.9 Hz, 2H, -CH$_2$-N$_3$), 1.68 (p, J = 6.7 Hz, 2H, -C**H$_2$**-CH$_2$-OPO$_3$H$_2$), 1.60 (p, J = 6.9 Hz, 2H, -C**H$_2$**-CH$_2$-CH$_2$-OPO$_3$H$_2$), 1.39 (m, 4H, -CH$_2$-CH$_2$-CH$_2$-N$_3$) ppm. ^{31}P-NMR (162 MHz, DMSO-d_6): 0.93 (s, 1P) ppm. ^{13}C-NMR (100 MHz, DMSO-d_6): 65.8 (-**C**H$_2$-OPO$_3$H$_2$), 51.3 (-CH$_2$-N$_3$), 28.7 (-**C**H$_2$-CH$_2$-OPO$_3$H$_2$), 26.2 (-**C**H$_2$-CH$_2$-CH$_2$-OPO$_3$H$_2$), 24.9 (-**C**H$_2$-**C**H$_2$-CH$_2$-N$_3$) ppm. HR-ESI-MS [M-1H]$^-$: m/z calculated: 222.0649, m/z found: 222.0672.

6.1.34. General procedure C:[97]

Typical reaction scales range from 40 mg to 120 mg of starting nucleoside. The respective nucleoside (1.0 eq.) and proton sponge (*N,N,N´,N´*-tetramethyl-1,8-naphthalenediamine) (1.5 eq.) were dried *in vacuo,* dissolved in dry trimethyl phosphate (1 ml per 20 mg of starting nucleoside) at room temperature and cooled to 0 °C. Phosphorous oxychloride (1.2 eq.) was added drop wise at 0 °C and the mixture was stirred under nitrogen atmosphere. After 30 min TLC showed complete conversion of starting material and 0.5 M solution of $(Bu_3NH)_2H_2P_2O_7$ in anhydrous DMF (5.0 eq.) and *n*Bu$_3$N (10.0 eq.) were added simultaneously to the mixture. The mixture was warmed to room temperature and stirred for 30 min. Then 0.1 M aqueous triethyl ammonium bicarbonate buffer (pH 7.5) (10 ml) was added and the mixture was stirred for further 30 min. The aqueous layer was washed with ethyl acetate several times to remove trimethyl phosphate and then concentrated to dryness. The residue was dissolved in water and purified by ion exchange chromatography (DEAE Sephadex A25, buffer A: 0.1 M TEAB, buffer B: 1 M TEAB, linear gradient: 0 % B to 100 % B) and further purified using reversed phase (RP)-HPLC (Nucleodur RP 18-ec, buffer A: 50 mM TEAA, buffer B: acetonitrile, linear gradient: 5 % B to 100 % B). The triphosphates were concentrated to dryness. To get rid of the triethylammonia acetate the residues were dissolved in water and freeze dried several times.

6.1.35. 8-Bromo-2´-deoxyguanosine-5´-*O*-triphosphate 2

Yield: 20 % (57.0 µmol). ^{1}H-NMR (400 MHz, MeOD): 6.31 (t, J = 7.1 Hz, 1H, H-1´), 4.85 (dt, J = 3.0 Hz, 6.2 Hz, 1H, H-3´), 4.50 - 4.38 (m, 1H, H-4´), 4.19 - 4.09 (m, 2H, H-5´a/b), 3.43 (dt, J = 6.8 Hz, 13.6

Hz, 1H, H-2´a), 2.22 (ddd, *J* = 3.3 Hz, 6.9 Hz, 13.5 Hz, 1H, H-2´b) ppm, ^{31}P-NMR (162 MHz, MeOD): -10.37 (d, *J* = 21.1 Hz, 1P, α-P), -11.00 (d, *J* = 21.3 Hz, 1P, γ-P), -23.6 (t, *J* = 21.3 Hz, 1P β-P) ppm. HR-ESI-MS $[M-1H]^-$: m/z calculated: 583.8990, m/z found: 583.8994. ε = 10100 L $mol^{-1}cm^{-1}$ (λ= 253 nm).

6.1.36. 8-Vinyl-2´-deoxyguanosine-5´-*O*-triphosphate 3

Yield: 16 % (56.0 µmol). ^{1}H-NMR (400 MHz, D_2O): 6.90 (dd, *J* = 11.3 Hz, 17.2 Hz, 1H, vinyl-H), 6.35 (t, *J* = 7.3 Hz, 1H, H-1´), 6.27 - 6.16 (m, 1H, vinyl-H), 5.74 - 5.69 (m, 1H, vinyl-H), 4.83 - 4.79 (m, 1H, H-3´), 4.31 (ddd, *J* = 3.8 Hz, 6.5 Hz, 9.6 Hz, 1H, H-5´a), 4.26 - 4.13 (m, 2H, H-4´, H-5´b), 3.10 - 2.99 (m, 1H, H-2´a), 2.33 (ddd, *J* = 3.5 Hz, 6.9 Hz, 14.0 Hz, 1H, H-2´b) ppm. ^{31}P-NMR (162 MHz, D_2O): -10.76 (d, *J* = 20.0 Hz, 1P, α-P), -11.41 (d, *J* = 20.2 Hz, 1P, γ-P), -23.4 (t, *J* = 19.8 Hz, 1P β-P) ppm. HR-ESI-MS $[M-1H]^-$: m/z calculated: 532.0041, m/z found: 532.0023. ε = 11010 L $mol^{-1}cm^{-1}$ (λ= 282 nm).

6.1.37. O^6-Methyl-2´-deoxyguanosine-5´-*O*-triphosphate 1a

Yield: 8 % (41.0 µmol). ^{1}H-NMR (400 MHz, D_2O): 8.37 (s, 1H, H-8), 6.40 (t, *J* = 6.8 Hz, 1H, H-1´), 4.81 - 4.79 (m, 1H, H-3´), 4.31 - 4.28 (m, 1H, H-4´), 4.24 - 4.18 (m, 2H, H-5´a/b), 4.10 (s, 3H, $-CH_3$), 2.87 - 2.80 (m, 1H, H-2´a), 2.58 (ddd, *J* = 3.5 Hz, 6.2 Hz, 14.0 Hz, 1H, H-2´b) ppm. ^{13}C-NMR (100 MHz, D_2O): 161.2 (C-6), 160.1 (C-2), 152.5 (C-4), 138.5 (C-8), 113.1 (C-5), 85.7 (C-4´), 83.7 (C-1´), 71.0 (C-3´), 65.4 (C-5´), 54.5 ($-O-CH_3$), 38.7 (C-2´) ppm. ^{31}P-NMR (162 MHz, D_2O): -9.85 (d, *J* = 17.5 Hz, 1P, α-P), -11.38 (d, *J* = 19.6 Hz, 1P, γ-P), -23.1 (m, 1P β-P) ppm. HR-ESI-MS $[M-1H]^-$: m/z calculated: 520.0037, m/z found: 520.0041. ε = 10100 L $mol^{-1}cm^{-1}$ (λ= 280 nm).

6.1.38. O^6-Ethyl-2´-deoxyguanosine-5´-*O*-triphosphate 1b

Yield: 22 % (48.2 µmol). ^{1}H-NMR (400 MHz, D_2O): 8.24 (s, 1H, H-8), 6.39 (dd, *J* = 6.5 Hz, 7.4 Hz, 1H, H-1´), 4.83 - 4.80 (m, 1H, H-3´), 4.55 (q, *J* = 7.1 Hz, 2 H, $-C\mathbf{H}_2-CH_3$), 4.30 - 4.27 (m, 1H, H-4´), 4.26 - 4.21 (m, 1H, H-5´a), 4.17 (ddd, *J* = 4.0 Hz, 5.4 Hz, 11.4 Hz, 1H, H-5´b), 2.85 (ddd, *J* = 6.5 Hz, 7.4 Hz, 14.0 Hz, 1H, H-2´a), 2.55 (ddd, *J* = 3.4 Hz, 6.3 Hz, 14.0 Hz, 1H, H-2´b), 1.46 (t, *J* = 7.1 Hz, 3H, $-CH_3$) ppm. ^{13}C-NMR (100 MHz, D_2O): 160.7 (C-6), 160.0 (C-2), 152.1 (C-4), 138.3 (C-8), 112.7 (C-5), 85.8 (C-4´), 83.9 (C-1´), 71.0 (C-3´), 65.4 (C-5´), 64.0 ($-O-CH_2-$), 38.8 (C-2´), 13.6 ($-CH_3$) ppm. ^{31}P-NMR (162 MHz, D_2O): -10.00 (d, *J* = 19.4 Hz, 1P, α-P), -11.34 (d, *J* = 19.1 Hz, 1P, γ-P), -23.1 (t, *J* = 19.9 Hz, β-P) ppm. HR-ESI-MS $[M-1H]^-$: m/z calculated: 534.0198, m/z found: 534.0181. ε = 8900 L $mol^{-1}cm^{-1}$ (λ= 281 nm).

6.1.39. O^6-Propyl-2´deoxyguanosine-5´-*O*-triphosphate 1c

Yield: 19 % (50.1 µmol). ^{1}H-NMR (400 MHz, D_2O): 8.23 (s, 1H, H-8), 6.35 (t, *J* = 6.9 Hz, 1H, H-1´), 4.78 - 4.76 (m, 1H, H-3´), 4.40 (t, *J* = 6.6 Hz, 2H, -O-CH_2-), 4.28 - 4.24 (m, 1H, H-4´), 4.23 - 4.13 (m, 2H, H-5´a/b), 2.83 (ddd, *J* = 6.3 Hz, 8.0 Hz, 13.8 Hz, 1H, H-2´a), 2.55 (ddd, *J* = 3.3 Hz, 6.2 Hz, 14.0 Hz, 1H, H-2´b), 1.83 (q, *J* = 7.1 Hz, 2H, -C**H$_2$**-CH_3), 1.03 (t, *J* = 7.4 Hz, 3H, -CH_3) ppm. ^{13}C-NMR (100 MHz, D_2O): 161.2 (C-6), 160.0 (C-2), 152.9 (C-4), 138.5 (C-8), 113.8 (C-5), 85.6 (C-4´), 83.4 (C-1´), 71.2 (C-3´), 69.4 (-O-CH_2), 65.5 (C-5´), 38.6 (C-2´), 21.7 (-**C**H_2-CH_3), 9.65 (-CH_3) ppm. ^{31}P-NMR (162 MHz, D_2O): -11.00 (d, *J* = 20.0 Hz, 1P, α-P), -11.45 (d, *J* = 20.3 Hz, 1P, γ-P), -23.34 (t, *J* = 20.1 Hz, β-P) ppm. HR-ESI-MS [M-1H]$^-$: m/z calculated: 548.0354, m/z found: 548.0352. ε = 9400 L mol^{-1}cm^{-1} (λ= 247 nm).

6.1.40. O^6-*Iso*-propyl-2´-deoxyguanosine-5´-*O*-triphosphate 1d

Yield: 35 % (89.7 µmol). ^{1}H-NMR (400 MHz, D_2O): 8.23 (s, 1H, H-8), 6.35 (t, *J* = 6.9 Hz, 1H, H-1´), 5.43 (h, *J* = 6.2 Hz, -O-C**H**-$(CH_3)_2$), 4.78 - 4.76 (m, 1H, H-3´), 4.28 - 4.24 (m, 1H, H-4´), 4.21 - 4.12 (m, 2H, H-5´a/b), 2.81 (ddd, *J* = 6.1 Hz, 7.6 Hz, 13.8 Hz, 1H, H-2´a), 2.54 (ddd, *J* = 3.3 Hz, 6.3 Hz, 14.0 Hz, 1H, H-2´b), 1.40 (d, *J* = 6.6 Hz, 3H, -CH_3), 1.39 (d, *J* = 6.6 Hz, 3H, -CH_3) ppm. ^{13}C-NMR (100 MHz, D_2O): 160.6 (C-6), 160.0 (C-2), 152.9 (C-4), 138.5 (C-8), 114.0 (C-5), 85.6 (C-4´), 83.4 (C-1´), 71.5 (C-3´), 71.1 (-O-**C**H-), 65.5 (C-5), 38.6 (C-2´), 21.1 (-CH_3) ppm. ^{31}P-NMR (162 MHz, D_2O): -11.00 (d, *J* = 20.0 Hz, 1P, α-P), -11.47 (d, *J* = 19.2 Hz, 1P, γ-P), -23.4 (t, *J* = 19.9 Hz, β-P) ppm. HR-ESI-MS [M-1H]$^-$: m/z calculated: 548.0354, m/z found: 548.0351. ε = 9400 L mol^{-1}cm^{-1} (λ= 247 nm).

6.1.41. 6-Chloro-2´-deoxyguanosine-5´-*O*-triphosphate 20

Yield: 13 % (178.0 µmol). ^{1}H-NMR (400 MHz, D_2O): 8.35 (s, 1H, H-8), 6.32 (t, *J* = 6.8 Hz, 1H, H-1´), 4.78 - 4.75 (m, 1H, H-3´), 4.25 - 4.22 (m, 1H, H-4´), 4.18 - 4.15 (m, 2H, H-5´a/b), 2.84 (ddd, *J* = 6.2 Hz, 7.5 Hz, 13.7 Hz, 1H, H-2´a), 2.55 (ddd, *J* = 3.6 Hz, 6.4 Hz, 14.0 Hz, 1H, H-2´b) ppm. ^{13}C-NMR (100 MHz, D_2O): 159.4 (C-6), 152.8 (C-2), 150.2 (C-4), 142.1 (C-8), 123.5 (C-5), 85.7 (C-4´), 83.7 (C-1´), 70.8 (C-3´), 65.5 (C-5´), 38.1 (C-2´) ppm. ^{31}P-NMR (162 MHz, D_2O): -10.86 (d, *J* = 20.1 Hz, 1P, α-P), -11.53 (d, *J* = 20.3 Hz, 1P, γ-P), -23.5 (t, *J* = 20.2 Hz, β-P) ppm. HR-ESI-MS [M-1H]$^-$: m/z calculated 523.9535, m/z found: 523.9525. ε = 9 800 L mol^{-1}cm^{-1} (λ= 306 nm).

6.1.42. 6-Thio-2´-deoxyguanosine-5´-*O*-triphosphate 10

Yield: 23 % (47.8 µmol). ^{1}H-NMR (400 MHz, D_2O): 8.24 (s, 1H, H-8), 6.32 (t, *J* = 6.5 Hz, 1H, H-1´), 4.91 - 4.84 (m, 1H, H-3´), 3.30 - 4.25 (m, 1H, H-4´), 3.23 -3.19 (m, 2H, H-5´a/b), 2.87 - 2.80 (m, 1H, H-2´a), 2.53 (ddd, *J* = 3.4 Hz, 6.4 Hz, 14.5 Hz, 1H, H-2´b) ppm. ^{13}C-NMR (100 MHz, D_2O): 173.5 (C-6), 153.3 (C-2), 147.7 (C-4), 140.4 (C-8), 128.3 (C-5), 85.7 (C-4´), 83.8 (C-1´), 71.1 (C-3´), 65.4 (C-5´), 38.4 (C-2´) ppm. ^{31}P-NMR (162 MHz, D_2O): -9.97 (d, J = 19.8 Hz, 1P, α-P), -11.24 (d, J = 19.7 Hz, 1P,

γ-P), -22.85 (t, J = 19.8 Hz, 1P, β-P) ppm. HR-ESI-MS [M-1H]⁻: m/z calculated 521.9610, m/z found: 521.9677. ε = 24 800 L $mol^{-1}cm^{-1}$ (λ= 341 nm).

6.1.43. 6-Thiomethyl-2´-deoxyguanosine-5´-*O*-triphosphate 10a

Yield: 28 % (57.5 µmol). ^{1}H-NMR (400 MHz, D_2O): 8.26 (s, 1H, H-8), 6.36 (dd, *J* = 7.5 Hz, 6.3 Hz, 1H, H-1´), 4.85 - 4.80 (m, 1H, H-3´), 4.29 - 4.24 (m, 1H, H-4´), 4.22 - 4.13 (m, 2H, H-5´a/b), 2.85 (ddd, *J* = 6.1 Hz, 7.6 Hz, 13.8 Hz, 1H, H-2´a), 2.59 (s, 3H, -CH_3), 2.57 (ddd, *J* = 3.4 Hz, 6.2 Hz, 14.0 Hz, 1H, H-2´b) ppm. ^{13}C-NMR (100 MHz, D_2O): 162.5 (C-6), 159.5 (C-2), 129.3 (C-4), 139.3 (C-8), 124.0 (C-5), 85.6 (C-4´), 83.2 (C-1´), 71.1 (C-3´), 65.5 (C-5´), 38.5 (C-2´), 11.1 (-S-CH_3) ppm. ^{31}P-NMR (162 MHz, D_2O): -11.60 (d, J = 20.0 Hz, 1P, α-P), -11.46 (d, *J* = 20.2 Hz, 1P, γ-P), -23.34 (t, J = 20.1 Hz, 1P, β-P) ppm. HR-ESI-MS [M-1H]⁻: m/z calculated 535.9784, m/z found: 535.9802. ε = 11 500 L $mol^{-1}cm^{-1}$ (λ= 309 nm).

6.1.44. 6-Thioethyl-2´-deoxyguanosine-5´-*O*-triphosphate 10b

Yield: 28 % (90.0 µmol). ^{1}H-NMR (400 MHz, D_2O): 8.22 (s, 1H, H-8), 6.31 (t, *J* = 6.9 Hz, 1H, H-1´), 4.72 - 4.63 (m, 1H, H-3´), 4.26 - 4.21 (m, 1H, H-5´a), 4.18 - 4.10 (m, 2H, H-4´, H-5´b), 3.26 - 3.17 (m, 2H, -S-CH_2-), 2.81 (dt, *J* = 6.8 Hz, 13.8 Hz, 1H, H-2´a), 2.55 (ddd, *J* = 3.3 Hz, 6.3 Hz, 14.2 Hz, 1H, H-2´b), 1.33 (t, *J* = 7.4 Hz, 1H, H-2´b) ppm. ^{13}C-NMR (100 MHz, D_2O): 158.3 (C-6), 155.7 (C-2), 145.7 (C-4), 135.6 (C-8), 120.3 (C-5), 82.0 (C-4´), 79.5 (C-1´), 67.5 (C-3´), 61.9 (C-5´), 34.8 (C-2´), 19.4 (-S-CH_2-), 10.1 (-CH_3) ppm. ^{31}P-NMR (162 MHz, D_2O): -10.0 - -11.1(m, 1P, α-P), -11.1 – -12.0 (m, 1P, γ-P), -20.5 – -22.5 (m, 1P, β-P) ppm. HR-ESI-MS [M-1H]⁻: m/z calculated 549.9958, m/z found: 549.9957. ε = 11 500 L $mol^{-1}cm^{-1}$ (λ= 309 nm).

6.1.45. 6-Amino-2´-deoxyguanosine-5´-*O*-triphosphate 9

Yield: 11 % (41.7 µmol). ^{1}H-NMR (400 MHz, D_2O): 8.11 (s, 1H, H-8), 6.22 (t, *J* = 6.8 Hz, 1H, H-1´), 4.73 (dt, *J* = 3.3 Hz, 6.3 Hz, 1H, H-3´), 4.24 - 4.20 (m, 1H, H-4´), 4.18 - 4.14 (m, 2H, H-5´a/b), 2.69 (dt, *J* = 7.0 Hz, 13.5 Hz, 1H, H-2´a), 2.49 (ddd, *J* = 3.5 Hz, 6.2 Hz, 13.9 Hz, 1H, H-2´b) ppm. ^{13}C-NMR (100 MHz, D_2O): 157.6 (C-6), 153.5 (C-2), 150.7 (C-4), 137.9 (C-8), 112.0 (C-5), 85.6 (C-4´), 83.2 (C-1´), 71.0 (C-3´), 65.5 (C-5´), 38.9 (C-2´) ppm. ^{31}P-NMR (162 MHz, D_2O): -10.51 (d, *J* = 19.7 Hz, 1P, α-P), -11.39 (d, *J* = 18.9 Hz, 1P, γ-P), -23.2 (t, *J* = 19.5 Hz, β-P) ppm. HR-ESI-MS [M-1H]⁻: m/z calculated 505.0034, m/z found: 505.0045. ε = 8 743 L $mol^{-1}cm^{-1}$ (λ= 279 nm).

6.1.46. 6-Methylamino-2´-deoxyguanosine-5´-*O*-triphosphate 9a

Yield: 42 % (152.5 µmol), ^{1}H-NMR (400 MHz, D_2O): 8.09 (s, 1H, H-8), 6.27 (t, *J* = 6.9 Hz, 1H, H-1´), 4.77 (dq, *J* = 3.3 Hz, 6.3 Hz, 1H, H-3´), 4.26 - 4.23 (m, 1H, H-4´), 4.19 - 4.12 (m, 2H, H-5´a/b),

3.02 (br s, 3H, -CH_3) 2.76 (ddd, *J* = 6.1 Hz, 7.6 Hz, 13.8 Hz, 1H, H-2´a), 2.53 (ddd, *J* = 3.5 Hz, 6.2 Hz, 13.9 Hz, 1H, H-2´b) ppm. ^{13}C-NMR (100 MHz, D_2O): 159.6 (C-6), 154.9 (C-2), 149.1 (C-4), 136.5 (C-8), 112.8 (C-5), 85.4 (C-4´), 82.9 (C-1´), 71.0 (C-3´), 65.5 (C-5), 38.7 (C-2´), 23.0 (-CH_3) ppm. ^{31}P-NMR (162 MHz, D_2O): -9.72 (d, *J* = 19.3 Hz, 1P, γ-P), -11.28 (d, *J* = 19.3 Hz, 1P, α-P), -22.84 (t, *J* = 19.8 Hz, β-P) ppm. HR-ESI-MS [M-1H]$^-$: m/z calculated 519.0201, m/z found: 519.0189. ε = 8 743 L mol^{-1}cm^{-1} (λ= 279 nm).

6.1.47. 6-Dimethylamino-2´-deoxyguanosine-5´-*O*-triphosphate 9h

Yield: 32 % (65.4 µmol). ^{1}H-NMR (400 MHz, D_2O): 8.07 (s, 1H, H-8), 6.27 (t, *J* = 6.9 Hz, 1H, H-1´), 4.70 - 4.65 (m, 1H, H-3´), 4.27 - 4.14 (m, 3H, H-4´,H-5´a/b), 3.30 (br s, 6H, 2x -CH_3), 2.74 (dt, *J* = 6.9 Hz, 13.8 Hz, 1H, H-2´a), 2.53 (ddd, *J* = 3.3 Hz, 6.2 Hz, 13.9 Hz, 1H, H-2´b) ppm. ^{13}C-NMR (100 MHz, D_2O): 157.3 (C-6), 153.2 (C-2), 150.2 (C-4), 135.9 (C-8), 112.9 (C-5), 85.5 (C-4´), 83.2 (C-1´), 71.0 (C-3´), 65.5 (C-5´), 38.8 (C-2´, 2x –CH_3) ppm. ^{31}P-NMR (162 MHz, D_2O): -10.59 (d, *J* = 19.9 Hz, 1P, α-P), -11.46 (d, *J* = 19.9 Hz, 1P, γ-P), -23.25 (t, *J* = 19.8 Hz, β-P) ppm. HR-ESI-MS [M-1H]$^-$: m/z calculated 533.0358, m/z found: 533.0344. ε = 8 743 L mol^{-1}cm^{-1} (λ= 279 nm).

6.1.48. 6-Methylethylamino-2´-deoxyguanosine-5´-*O*-triphosphate 9i

Yield: 37 % (0.12 mmol). ^{1}H-NMR (400 MHz, D_2O): 8.03 (s, 1H, H-8), 6.24 (dd, *J* = 6.2 Hz, 7.5 Hz, 1H, H-1´), 4.73 (dt, *J* = 3.5 Hz, 6.4 Hz, H-3´), 4.23 - 4.19 (m, 1H, H-5´a), 4.18 - 4.10 (m, 2H, H-4´, H-5´b), 3.75 - 3.64 (m, 2H, -N-CH_2-), 3.09 - 3.05 (m, 3H, -N-CH_3), 2.64 (ddd, *J* = 6.1 Hz, 7.7 Hz, 13.7 Hz, 1H, H-2´a), 2.42 (ddd, *J* = 3.3 Hz, 6.2 Hz, 13.9 Hz, 1H, H-2´b), 1.05 (t, J = 7.1 Hz, 3H, -CH_3) ppm. ^{13}C-NMR (100 MHz, $CDCl_3$): 158.5 (C-2), 153.3 (C-6), 150.8 (C-4), 135.6 (C-8), 112.8 (C-5), 88.5 (C-4´), 82.9 (C-1´), 71.1 (C-3´), 65.5(C-5´), 45.6 (-N-CH_2-), 38.7 (C-2´), 35.9 (-N-CH_3), 12.1 (CH_3) ppm. ^{31}P-NMR (162 MHz, D_2O): -10.10 (m, 1P, α-P), -11.42 (d, *J* = 19.5 Hz, 1P, γ-P), -22.88 (m, 1P, β-P) ppm. HR-ESI-MS [M-1H]$^-$: m/z calculated 547.0514, m/z found: 547.0499. ε = 8 743 L mol^{-1}cm^{-1} (λ= 279 nm).

6.1.49. 6-Diethylamino-2´-deoxyguanosine-5´-*O*-triphosphate 9l

Yield: 32 % (99.7 µmol). ^{1}H-NMR (400 MHz, D_2O): 8.05 (s, 1H, H-8), 6.27 (dd, *J* = 6.2 Hz, 7.6 Hz, 1H, H-1´), 4.75 (dq, *J* = 2.8 Hz, 3.3 Hz, 1H, H-3´), 4.28 - 4.10 (m, 3H, H-4´,H-5´a/b), 3.79 - 3.69 (m, 4H, N-$(CH_2)_2$-), 2.76 - 2.69 (m, 1H, H-2´a), 2.50 (ddd, *J* = 3.4 Hz, 6.3 Hz, 13.9 Hz, 1H, H-2´b), 1.16 (t, *J* = 7.0 Hz, 6H, 2x -CH_3) ppm. ^{13}C-NMR (100 MHz, D_2O): 158.7 (C-6), 153.0 (C-2), 150.8 (C-4), 135.5 (C-8), 112.8 (C-5), 85.4 (C-4´), 83.0 (C-1´), 70.9 (C-3´), 65.4 (C-5´), 43.2 (2x -CH_2-), 38.7 (C-2´), 16.7 (2x -CH_3) ppm. ^{31}P-NMR (162 MHz, D_2O): -10.02 (m, 1P, α-P), -11.40 (d, *J* = 19.6 Hz, 1P, γ-P), -22.89 (t, *J* = 19.7 Hz, β-P) ppm. HR-ESI-MS [M-1H]$^-$: m/z calculated 561.0671, m/z found: 561.0673. ε = 8 743 L mol^{-1}cm^{-1} (λ= 279 nm).

6.1.50. 8-Bromo-*O*^{6}-methyl-2´-deoxyguanosine-5´-*O*-triphosphate 30a

Yield: 9 % (4.4 µmol). ^{1}H-NMR (400 MHz, D_2O): 6.47 (m, 1H, H-1´), 4.92 (dq, *J* = 2.5 Hz, 3.1 Hz, 1H, H-3´), 4.36 (ddd, *J* = 2.4 Hz, 5.2 Hz, 10.1 Hz, 1H, H-5´a), 4.28 - 4.17 (m, 2H, H-4´, H-5´b), 4.08 (s, 3H, -O-CH_3), 3.39 (dt, *J* = 7.1 Hz, 14.2 Hz, 1H, H-2´a), 2.44 (ddd, *J* = 3.9 Hz, 7.2 Hz, 14.1 Hz, 1H, H-2´b) ppm. ^{31}P-NMR (162 MHz, D_2O): -10.87 (d, *J* = 20.0 Hz, 1P, α-P), -11.16 (d, *J* = 19.9 Hz, 1P, γ-P), -23.21 (t, *J* = 20.0 Hz, β-P) ppm. HR-ESI-MS [M-1H]$^-$: m/z calculated 597.9146, m/z found: 597.9173. ε = 10 100 L mol^{-1}cm^{-1} (λ= 252 nm).

6.1.51. 8-Bromo-*O*^{6}-ethyl-2´-deoxyguanosine-5´-*O*-triphosphate 30b

Yield: 7 % (4.2 µmol). ^{1}H-NMR (400 MHz, D_2O): 6.47 (t, *J* = 7.1Hz, 1H, H-1´), 4.94 - 4.90 (m, 1H, H-3´), 4.59 - 4.51 (m, 2H, -O-CH_2-), 4.41 - 4.29 (m, 1H, H-5´a), 4.28 - 4.14 (m, 2H, H-4´, H-5´b), 3.38 (dt, *J* = 7.1 Hz, 14.3 Hz, 1H, H-2´a), 2.43 (ddd, *J* = 3.8 Hz, 7.2 Hz, 14.2 Hz, 1H, H-2´b) ppm. ^{31}P-NMR (162 MHz, D_2O): -10.84 (m, 1P, α-P), -11.13 (d, *J* = 20.8 Hz, 1P, γ-P), -23.16 (m, 1P β-P) ppm. HR-ESI-MS [M-1H]$^-$: m/z calculated 611.9303, m/z found: 661.9338. ε = 10 100 L mol^{-1}cm^{-1} (λ= 252 nm).

6.1.52. General procedure D:

The respective amine or the aqueous solution of the respective amine (100.0 eq.) was added to a 10 mM solution of 2´-deoxy-6-chloro-5´-*O*-triphosphate (10 µmol) and shaken overnight at room temperature. The reaction mixture was concentrated to dryness. The residue was dissolved in water and purified by ion exchange chromatography (DEAE Sephadex A25, buffer A: 0.1 M TEAB, buffer B: 1 M TEAB, linear gradient: 0 % B to 100 % B) and further purified using reversed phase (RP)-HPLC (Nucleodur RP 18-ec, buffer A: 50 mM TEAA, buffer B: acetonitrile, linear gradient: 5 % B to 100 % B). The triphosphates were concentrated to dryness. To get rid of the triethylammonia acetate the residues were dissolved in water and freeze dried several times.

6.1.53. 6-Methylamino-2´-deoxyguanosine-5´-*O*-triphosphate 9a

Yield: 92 % (9.2 µmol), ^{1}H-NMR (400 MHz, D_2O): 8.09 (s, 1H, H-8), 6.27 (t, *J* = 6.9 Hz, 1H, H-1´), 4.77 (dq, *J* = 3.3 Hz, 6.3 Hz, 1H, H-3´), 4.26 - 4.23 (m, 1H, H-4´), 4.19 - 4.12 (m, 2H, H-5´a/b), 3.02 (br s, 3H, -CH_3) 2.76 (ddd, *J* = 6.1 Hz, 7.6 Hz, 13.8 Hz, 1H, H-2´a), 2.53 (ddd, *J* = 3.5 Hz, 6.2 Hz, 13.9 Hz, 1H, H-2´b) ppm. ^{13}C-NMR (100 MHz, D_2O): 159.6 (C-6), 154.9 (C-2), 149.1 (C-4), 136.5 (C-8), 112.8 (C-5), 85.4 (C-4´), 82.9 (C-1´), 71.0 (C-3´), 65.5 (C-5´), 38.7 (C-2´), 23.0 (-CH_3) ppm. ^{31}P-NMR (162 MHz, D_2O): -9.72 (d, *J* = 19.3 Hz, 1P, γ-P), -11.28 (d, *J* = 19.3 Hz, 1P, α-P), -22.84 (t, *J* = 19.8 Hz, β-P) ppm. HR-ESI-MS [M-1H]$^-$: m/z calculated 519.0201, m/z found: 519.0189. ε = 8 743 L mol^{-1}cm^{-1} (λ= 279 nm).

6.1.54. 6-Ethylamino-2´-deoxyguanosine-5´-*O*-triphosphate 9b

Yield: 87 % (8.7 µmol). ^{1}H-NMR (400 MHz, D_2O): 8.18 (s, 1H, H-8), 6.33 (t, *J* = 6.8 Hz, 1H, H-1´), 4.86 - 4.80 (m, 1H, H-3´), 4.33 - 4.27 (m, 1H, H-5´a), 4.27 - 4.22 (m, 2H, H-4´, H-5´b), 3.62 (br s, 2H, -NH-C**H**$_2$-), 2.75 (dt, *J* = 6.7 Hz, 13.8 Hz, H-2´a), 2.54 (ddd, *J* = 3.4 Hz, 6.2 Hz, 14.0 Hz, H-2´b), 1.31 (t, *J* = 7.2 Hz, 3H, -CH_3) ppm. ^{31}P-NMR (162 MHz, D_2O): -10.72 - -10.82 (m, 1P, α-P), -11.19 (d, *J* = 19.6 Hz, 1P, γ-P), -22.60 - - 22.55 (t, *J* = 19.7 Hz, β-P) ppm. HR-ESI-MS $[M-1H]^-$: m/z calculated 533.0358, m/z found: 533.0433. ε = 8 743 L mol^{-1}cm^{-1} (λ= 279 nm).

6.1.55. 6-Propylamino-2´-deoxyguanosine-5´-*O*-triphosphate 9c

Yield: 63 % (6.3 µmol). ^{1}H-NMR (400 MHz, D_2O): 8.17 (s, 1H, H-8), 6.34 (t, *J* = 6.9 Hz, 1H, H-1´), 4.90 - 4.82 (m, 1H, H-3´), 4.30 (dt, *J* = 3.2 Hz, 5.5 Hz, 1H, H-5´a), 4.26 - 4.16 (m, 2H, H-4´, H-5´b), 3.54 (br s, 2H, -NH-C**H**$_2$-), 2.77 (dt, *J* = 6.8 Hz, 13.7 Hz, 1H, H-2´a), 2.54 (ddd, *J* = 3.3 Hz, 6.3 Hz, 14.0 Hz, 1H, H-2´b), 1.71 (h, *J* = 7.3 Hz, 2H, -NH-CH_2-C**H**$_2$-), 1.00 (t, *J* = 7.4 Hz, 3H, -CH_3) ppm. ^{31}P-NMR (162 MHz, D_2O): -10.48 - -10.95 (m, 1P, α-P), -11.2 (d, *J* = 20.2 Hz, 1P, γ-P), -22.89 - -23.42 (m, β-P) ppm. HR-ESI-MS $[M-1H]^-$: m/z calculated 547.0514, m/z found: 547.0584. ε = 8 743 L mol^{-1}cm^{-1} (λ= 279 nm).

6.1.56. 6-*Iso*-propylamino-2´-deoxyguanosine-5´-*O*-triphosphate 9d

Yield: 43 % (4.3 µmol). ^{1}H-NMR (400 MHz, D_2O): 8.16 (s, 1H, H-8), 6.40 - 6.34 (m, 1H, H-1´), 4.87 - 4.80 (m, 1H, H-3´), 4.42 - 4.33 (m, 1H, H-5´a), 4.31 - 4.13 (m, 3H, -NH-C**H**-, H-4´, H-5´b), 2.80 (dt, *J* = 6.7 Hz, 13.8 Hz, 1H, H-2´a), 2.54 (ddd, *J* = 3.1 Hz, 6.1 Hz, 14.1 Hz, 1H, H-2´b), 1.32 - 1.30 (m, 6H, 2 x -CH_3) ppm. ^{31}P-NMR (162 MHz, D_2O): -10.42 - -10.76 (m, 1P, α-P), -11.18 - -11.35 (m, 1P, γ-P), -23.00 - -23.42 (m, β-P) ppm. HR-ESI-MS $[M-1H]^-$: m/z calculated 547.0514, m/z found: 547.0584. ε = 8 743 L mol^{-1}cm^{-1} (λ= 279 nm).

6.1.57. 6-Hydroxy-1-ethanamino-2´-deoxyguanosine-5´-*O*-triphosphate 9e

Yield: 52 % (5.2 µmol). ^{1}H-NMR (400 MHz, D_2O): 8.19 (s, 1H, H-8), 6.36 (t, *J* = 6.9 Hz, 1H, H-1´), 4.71 - 4.66 (m, 1H, H-3´), 4.32 - 4.29 (m, 1H, H-5´a), 4.26 - 4.20 (m, 2H, H-4´, H-5´b), 3.87 - 3.85 (m, 2H, -C**H$_2$**-OH), 3.82 - 3.75 (m, 2H, -NH-C**H$_2$**-),2.78 (dt, *J* = 6.9 Hz, 13.8 Hz, 1H, H-2´a), 2.55 (ddd, *J* = 3.3 Hz, 6.2 Hz, 14.0 Hz, 1H, H-2´b) ppm. ^{31}P-NMR (162 MHz, D_2O): -10.64 - -10.86 (m, 1P, α-P), -11.21 (d, *J* = 19.5 Hz, 1P, γ-P), -23.00 - -23.33 (m, β-P) ppm. HR-ESI-MS $[M-1H]^-$: m/z calculated 549.0307, m/z found: 549.0375. ε = 8 743 L mol^{-1}cm^{-1} (λ= 279 nm).

6.1.58. 6-Azido-1-propanamino-2´-deoxyguanosine-5´-*O*-triphosphate 9f

Yield: 86 % (8.6 µmol). ^{1}H-NMR (400 MHz, D_2O): 8.17 (s, 1H, H-8), 6.35 (t, *J* = 7.0 Hz, H-3´), 4.72 - 4.63 (m, 1H, H-3´), 4.30 - 4.28 (m, 1H, H-5´-a), 4.26 - 4.18 (m, 2H, H-4´, H-5´b), 3.73 - 3.63 (m, 2H, -NH-C**H$_2$**-), 3.50 (t, *J* = 6.6 Hz, 2H, -CH_2-N_3), 2.77 (dt, *J* = 7.0 Hz, 14.0 Hz, 1H, H-2´a), 2.54 (ddd, *J* = 3.2 Hz, 6.1 Hz, 14.0 Hz, 1H, H-2´b), 1.99 (p, *J* = 6.7 Hz, 2H, -NH-CH_2-C**H$_2$**-) ppm. ^{31}P-NMR (162 MHz, D_2O): -10.63 - - 10.88 (m, 1P, α-P), -11.15 - - 11.31 (m, 1P, γ-P), - 22.95 - -22.45 (t, *J* = 19.7 Hz, β-P) ppm. HR-ESI-MS [M-1H]$^-$: m/z calculated 588.0528, m/z found: 588.0534. ε = 8 743 L mol^{-1}cm^{-1} (λ= 279 nm).

6.1.59. 6-Cyclohexylamino-2´-deoxyguanosine-5´-*O*-triphosphate 9g

Yield: 97 % (9.7 µmol). ^{1}H-NMR (400 MHz, D_2O): 8.16 (s, 1H, H-8), 6.38 - 6.34 (m, 1H, H-1´), 4.54 - 4.31 (m, 1H, H-3´), 4.31 - 4.28 (m, 1H, H-5´a), 4.27 - 4.13 (m, 2H, H-4´, H-5´b), 2.79 (dt, *J* = 6.4 Hz, 13.7 Hz, 1H, H-2´a), 2.54 (ddd, *J* = 3.0 Hz, 6.1 Hz, 14.1 Hz, 1H, H-2´b), 1.82 - 1.63 (m, 9H, $H_{cyclopentyl}$) ppm. ^{31}P-NMR (162 MHz, D_2O): -10.23 - -10.80 (m, 1P, α-P), -11.00 - -11.30 (m, 1P, γ-P), -22.80 - -22.67 (m, β-P) ppm. HR-ESI-MS [M-1H]$^-$: m/z calculated 573.0671, m/z found: 573.0714. ε = 8 743 L mol^{-1}cm^{-1} (λ= 279 nm).

6.1.60. 6-Dimethylamino-2´-deoxyguanosine-5´-*O*-triphosphate 9h

Yield: 85 % (8.5 µmol). ^{1}H-NMR (400 MHz, D_2O): 8.07 (s, 1H, H-8), 6.27 (t, *J* = 6.9 Hz, 1H, H-1´), 4.70 - 4.65 (m, 1H, H-3´), 4.27 - 4.14 (m, 3H, H-4´,H-5´a/b), 3.30 (br s, 6H, 2x -CH_3), 2.74 (dt, *J* = 6.9 Hz, 13.8 Hz, 1H, H-2´a), 2.53 (ddd, *J* = 3.3 Hz, 6.2 Hz, 13.9 Hz, 1H, H-2´b) ppm. ^{13}C-NMR (100 MHz, D_2O): 157.3 (C-6), 153.2 (C-2), 150.2 (C-4), 135.9 (C-8), 112.9 (C-5), 85.5 (C-4´), 83.2 (C-1´), 71.0 (C-3´), 65.5 (C-5´), 38.8 (C-2´, 2x -CH_3) ppm. ^{31}P-NMR (162 MHz, D_2O): -10.59 (d, *J* = 19.9 Hz, 1P, α-P), -11.46 (d, *J* = 19.9 Hz, 1P, γ-P), -23.25 (t, *J* = 19.8 Hz, β-P) ppm. HR-ESI-MS [M-1H]$^-$: m/z calculated 533.0358, m/z found: 533.0344. ε = 8 743 L mol^{-1}cm^{-1} (λ= 279 nm).

6.1.61. 6-Methylethylamino-2´-deoxyguanosine-5´-*O*-triphosphate 9i

Yield: 73 % (7.3 µmol). ^{1}H-NMR (400 MHz, D_2O): 8.03 (s, 1H, H-8), 6.24 (dd, *J* = 6.2 Hz, 7.5 Hz, 1H, H-1´), 4.73 (dt, *J* = 3.5 Hz, 6.4 Hz, H-3´), 4.23 - 4.19 (m, 1H, H-5´a), 4.18 - 4.10 (m, 2H, H-4´, H-5´b), 3.75 - 3.64 (m, 2H, -N-CH_2-), 3.09 - 3.05 (m, 3H, -N-CH_3), 2.64 (ddd, *J* = 6.1 Hz, 7.7 Hz, 13.7 Hz, 1H, H-2´a), 2.42 (ddd, *J* = 3.3 Hz, 6.2 Hz, 13.9 Hz, 1H, H-2´b), 1.05 (t, J = 7.1 Hz, 3H, -CH_3) ppm. ^{13}C-NMR (100 MHz, $CDCl_3$): 158.5 (C-2), 153.3 (C-6), 150.8 (C-4), 135.6 (C-8), 112.8 (C-5), 88.5 (C-4´), 82.9 (C-1´), 71.1 (C-3´), 65.5 (C-5´), 45.6 (-N-CH_2-), 38.7 (C-2´), 35.9 (-N-CH_3), 12.1 (CH_3) ppm. ^{31}P-NMR (162 MHz, D_2O): -10.10 (m, 1P, α-P), -11.42 (d, *J* = 19.5 Hz, 1P, γ-P), -22.88 (m, 1P, β-P) ppm. HR-ESI-MS [M-1H]$^-$: m/z calculated 547.0514, m/z found: 547.0499. ε = 8 743 L mol^{-1}cm^{-1} (λ= 279 nm).

6.1.62. 6-Methylpropylamino-2´-deoxyguanosine-5´-*O*-triphosphate 9j

Yield: 67 % (6.7 µmol). ^{1}H-NMR (400 MHz, D_2O): 8.15 (s, 1H, H-8), 6.40 - 6.36 (m, 1H, H-1´), 4.74 - 4.68 (m, 1H, H-3´), 4.32 - 4.17 (m, 3H, H-4´, H-5´a/b), 3.94 (br s, 2H, -N-CH_2-), 3.36 (brs, 3H, -N-CH_3), 2.82 - 2.78 (m, 1H, H-2´a), 2.57 - 2.53 (m, 1H, H-2´b), 1.73 (h, *J* = 7.5 Hz, 2H, -N-CH_2-C**H$_2$**-), 0.92 (t, *J* = 7.3 Hz, 3H, -CH_2-CH_3) ppm. ^{31}P-NMR (162 MHz, D_2O): -9.47 - - 11.93 (m, 2P, α-P, γ-P), -22.10 - -23.33 (m, Hz, β-P). HR-ESI-MS [M-1H]$^-$: m/z calculated 561.0671, m/z found: 561.0681. ε = 8 743 L mol^{-1}cm^{-1} (λ= 279 nm).

6.1.63. 6-Methyl-*iso*-propylamino-2´-deoxyguanosine-5´-*O*-triphosphate 9k

Yield: 34 % (3.4 µmol). ^{1}H-NMR (400 MHz, d_4-MeOD): 8.18 (s, 1H, H-8), 6.49 (t, *J* = 6.8 Hz, 1H, H-1´), 4.87 (dt, *J* = 3.4 Hz, 6.4 Hz, 1H, H-3´), 4.42 - 4.40 (m, 2H, H-4´, H-5´a), 4.39 - 4.26 (m, 1H, H-5´b), 3.49 (p, *J* = 1.6 Hz, 1H, -N-C**H**-(CH_3)$_2$), 3.3.7 (br s, 3H, -N-CH_3), 2.93 (dt, *J* = 6.7 Hz, 13.5 Hz, 1H, H-2´a), 2.54 (ddd, *J* = 3.6 H, 6.3 Hz, 13.4 Hz, 1H, H-2´b), 1.41 (d, *J* = 6.8 Hz, 6H, -N-CH-(C**H$_3$**)$_2$) ppm. ^{31}P-NMR (162 MHz, d_4-MeOD): -10.44 (d, *J* = 22.7 Hz, 1P, α-P), -11.26 (d, *J* = 20.4 Hz, 1P, γ-P), -23.71 (t, *J* = 21.2 Hz, β-P) ppm. HR-ESI-MS [M-1H]$^-$: m/z calculated 561.0671, m/z found: 561.0710. ε = 8 743 L mol^{-1}cm^{-1} (λ= 279 nm).

6.1.64. 6-Diethylamino-2´-deoxyguanosine-5´-*O*-triphosphate 9l

Yield: 65 % (6.5 µmol). ^{1}H-NMR (400 MHz, D_2O): 8.05 (s, 1H, H-8), 6.27 (dd, *J* = 6.2 Hz, 7.6 Hz, 1H, H-1´), 4.75 (dq, *J* = 2.8 Hz, 3.3 Hz, 1H, H-3´), 4.28 - 4.10 (m, 3H, H-4´,H-5´a/b), 3.79 - 3.69 (m, 4H, N-(CH_2)$_2$-), 2.76 - 2.69 (m, 1H, H-2´a), 2.50 (ddd, *J* = 3.4 Hz, 6.3 Hz, 13.9 Hz, 1H, H-2´b), 1.16 (t, *J* = 7.0 Hz, 6H, 2x -CH_3) ppm. ^{13}C-NMR (100 MHz, D_2O): 158.7 (C-6), 153.0 (C-2), 150.8 (C-4), 135.5 (C-8), 112.8 (C-5), 85.4 (C-4´), 83.0 (C-1´), 70.9 (C-3´), 65.4 (C-5´), 43.2 (2x -CH_2-), 38.7 (C-2´), 16.7 (2x -CH_3) ppm. ^{31}P-NMR (162 MHz, D_2O): -10.02 (m, 1P, α-P), -11.40 (d, *J* = 19.6 Hz, 1P, γ-P), -22.89 (t, *J* = 19.7 Hz, β-P) ppm. HR-ESI-MS [M-1H]$^-$: m/z calculated 561.0671, m/z found: 561.0673. ε = 8 743 L mol^{-1}cm^{-1} (λ= 279 nm).

6.1.65. 6-Ethylpropylamino-2´-deoxyguanosine-5´-*O*-triphosphate 9m

Yield: 48 % (4.8 µmol). ^{1}H-NMR (400 MHz, d_4-MeOD): 7.97 (s, 1H, H-8), 6.29 (t, *J* = 6.8 Hz, 1H, H-1´), 4.69 (dt, *J* = 3.3 Hz, 6.2 Hz, 1H, H-3´), 4.24 - 4.14 (m, 2H, H-4´, H-5´a), 4.11 - 4.08 (m, 1H, H-5´b), 3.89 (br s, 2H, -N-CH_2-), 3.82 (br s, 2H, -N-CH_2-), 2.73 (dt, *J* = 6.4 Hz, 13.6 Hz, 1H, H-2´a), 2.34 (ddd, *J* = 3.5 Hz, 6.2 Hz, 13.5 Hz, 1H, H-2´b), 1.69 (h, *J* = 7.4 Hz, 2H, -N-CH_2-C**H$_2$**-), 1.21 (t, *J* = 7.0 Hz, 3H, -N-CH_2-C**H$_3$**), 0.93 (t, *J* = 7.4 Hz, 3H, -N-CH_2-CH_2-C**H$_3$**) ppm. ^{31}P-NMR (162 MHz, d_4-MeOD): -10.41 (d, *J* = 21.3 Hz, 1P, α-P), -11.19 (d, *J* = 20.5 Hz, 1P, γ-P), -23.64 (t, *J* = 21.5 Hz, β-P) ppm. HR-ESI-MS [M-1H]$^-$: m/z calculated 575.0827, m/z found: 575.0835. ε = 8 743 L mol^{-1}cm^{-1} (λ= 279 nm).

6.1.66. 6-Ethyl-*iso*-propylamino-2´-deoxyguanosine-5´-*O*-triphosphate 9n

Yield: 27 % (2.7 µmol). ^{1}H-NMR (400 MHz, D_2O): 8.07 (s, 1H, H-8), 6.29 (t, *J* = 6.8 Hz, 1H, H-1´), 5.27 (m, 1H, -N-CH-), 4.75 (m, 1H, H-3´), 4.24 - 4.11 (m, 3H, H-4´, H-5´a/b), 3.60 (m, 2H, -N-CH_2-), 2.71 (dt, *J* = 6.8 Hz, 13.7 Hz, 1H, H-2´a), 2.49 (ddd, *J* = 3.3 Hz, 6.4 Hz, 14.2 Hz, 1H, H-2´b), 1.17 (d, *J* = 6.8 Hz, 6H, -N-CH-($C\mathbf{H_3}$)$_2$), 1.13 (t, *J* = 6.3 Hz, 3H, -N-CH_2-$C\mathbf{H_3}$) ppm. ^{31}P-NMR (162 MHz, D_2O): -10.27 (m, 1P, α-P), -11.14 (m, 1P, γ-P), -22.93 (m, 1P, β-P) ppm. HR-ESI-MS [M-1H]$^-$: m/z calculated 575.0827, m/z found: 575.0826. ε = 8 743 L mol^{-1}cm^{-1} (λ= 279 nm).

6.1.67. 6-Pyrrolidine-2´-deoxyguanosine-5´-*O*-triphosphate 9o

Yield: 18 % (1.8 µmol). ^{1}H-NMR (400 MHz, D_2O): 8.16 (s, 1H, H-8), 6.36 (t, *J* = 6.9 Hz, 1H, H-1´), 4.86 - 4.80 (m, 1H, H-3´), 4.31 - 4.27 (m, 1H, H-5´a), 4.25 - 4.17 (m, 2H, H-4´, H-5´b), 4.07 - 3.55 (m, 4H, $H_{Pyrrolidine}$), 2.80 (dt, *J* = 6.8 Hz, 13.8 Hz, 1H, H-2´a), 2.54 (ddd, *J* = 3.4 Hz, 6.3 Hz, 14.0 Hz, 1H, H-2´a), 5.11 - 0.201 (m, 4H, $H_{Pyrrolidine}$) ppm. ^{31}P-NMR (162 MHz, D_2O): -10.56 (d, *J* = 19.7 Hz, 1P, α-P), -11.26 (d, *J* = 19.9 Hz, 1P, γ-P), -23.11 (t, *J* = 18.6 Hz, β-P) ppm. HR-ESI-MS [M-1H]$^-$: m/z calculated 559.0514, m/z found: 559.0541. ε = 8 743 L mol^{-1}cm^{-1} (λ= 279 nm).

6.1.68. General procedure E

Typical reactions range from 30 µmol to 400 µl of nucleotide. Tetrabutyl ammonium bromide (10 g, 1 mmol) in 100 ml water was applied to an ion exchange column (20 ml, CHELEX 100 ion exchange resin). The column was washed with 500 ml water and the acetate free triphosphate in 10 ml water was applied to the column. The column was again washed with 500 ml water. The fractions containing the nucleotide were pooled and evaporated to dryness. The residue was coevaporated with abs. DMF thre times.

6.1.69. γ-Azidohexyl-2´-deoxyguanosine-5´-*O*-triphosphate 7[103]

1-azido-6-bromohexane (23 mg, 0.11 mmol, 3.0 eq.) and the tetrabutylammonium salt of dGTP (37 µmol, 1.0 eq.) were each dissolved in 1 ml abs. DMF and stored over molecular sieve under nitrogen atmosphere overnight in separate flasks. The two solutions were following combined under nitrogen atmosphere and stored at room temperature over night. The mixture was subsequently dried *in vacuo* and the residue was dissolved in water to perform purification by ion exchange chromatography followed by purification by RP-HPLC and repeated freeze drying. Yield: 10 % (3.5 µmol). ^{1}H-NMR (400 MHz, MeOD-d_5): 8.02 (s, 1H, H-8), 6.24 (t, *J* = 6.8 Hz, 1H, H-1´), 4.78 (dt, J = 3.2 Hz, 6.2 Hz, 1H, H-3´), 4.33 - 4.25 (m, 1H, H-5´a), 4.22 - 4.11 (m, 2H, H-4´, H-5´b), 4.01 (q, *J* = 6.5 Hz, 2H, -PO_4-CH_2-), 3.26 (t, *J* = 7.0 Hz, 2H, -CH_2-N_3), 2.87 (dt, *J* = 6.7 Hz, 13.5 Hz, 1H, H-2´a), 2.32 (ddd, *J* = 3.5 Hz, 6.3 Hz, 13.3 Hz, 1H, H-2´a), 1.64 (p, *J* = 6.5 Hz, 2H, H-linker), 1.57 (p, *J* = 6.9 Hz, 2H, H-linker), 1.43 - 1.37 (m, 4H, H-linker) ppm. ^{31}P-NMR (162 MHz, D_2O): -10.45 (dt, *J* = 6.2 Hz, 17.8 Hz, 1P, α-P), -10.80

(d, *J* = 6.3 Hz, 18.0 Hz, 1P, γ-P), -21.86 (t, *J* = 17.3 Hz, β-P). HR-ESI-MS [M-1H]$^-$: m/z calculated: 631.0838. m/z found: 631.0831. ε = 8 743 L mol^{-1}cm^{-1} (λ= 279 nm).

6.1.70. γ-Aminohexyl-2´-deoxyguanosine-5´-*O*-triphosphate 31[103]

γ-Azidohexyl-2´-deoxyguanosine-5´-*O*-triphosphate **7** (12.4 µmol, 1 eq.) was dissolved in a mixture of 3 ml water, 4 ml methanol and 2 ml triethylamine. Tris(2-carboxyethyl)phosphine (17.8 mg, 62.0 µmol, 5 eq.) was added solubilized in 1 ml water and the reaction was stirred at room temperature for 3 days. The solution was concentrated *in vacuo* and the crude residue was purified by RP-HPLC followed by repeated freeze drying. Yield: 77 % (9.6 µmol). ^{1}H-NMR (400 MHz, D_2O): 8.11 (s, 1H, H-8), 6.32 (t, *J* = 7.0 Hz, 1H, H-1´), 4.84 – 4.81 (m, 1H, H-3´), 4.28 – 4.25 (m, 1H, H-4´), 4.22 – 4.19 (m, 2H, H-5´a/b), 3.95 (q, *J* = 6.7 Hz, 2H, -PO_4-CH_2-), 2.98 (t, *J* = 7.4 Hz, 2H, -CH_2-NH_2), 2.87 (ddd, *J* = 6.2 Hz, 7.7 Hz, 13.9 Hz, 1H, H-2´a), 2.52 (ddd, *J* = 3.4 Hz, 6.4 Hz, 14.0 Hz, 1H, H-2´a), 1.64 – 1.56 (m, 4H, H-linker), 1.35 - 1.31 (m, 4H, H-linker) ppm. ^{31}P (162 MHz, D_2O): -10.94 (d, *J* = 20.0 Hz, 1P, α-P), -11.48 (d, *J* = 19.2 Hz, 1P, γ-P), -23.37 (t, *J* = 19.6 Hz, β-P) ppm. HR-ESI-MS [M-1H]$^-$: m/z calculated: 605.0933.m/z found: 605.0971. ε = 8 743 L mol^{-1}cm^{-1} (λ= 279 nm).

6.1.71. γ-(Sulfo-Cy3-amidohexyl)-2´-deoxyguanosine-5´-*O*-triphosphate 32

The coupling reaction was carried out in 900 µl water and 100 µl 1 M $NaHCO_3$ solution (pH 8.7). γ-Aminohexyl-2´-deoxyguanosine-5´-*O*-triphosphate (9.6 µmol, 1.0 eq.) was dissolved in that mixture and the sulfo-Cy3-NHS ester (8.5 mg, 23.0 µmol, 1.2 eq.) was added as DMF solution (200 µl). The reaction mixture was stirred overnight at room temperature. The solution was following concentrated to dryness and purified by ion exchange followed by RP-HPLC. The residue was freeze dried several times to get rid of remaining triethyl ammonia acetate. Yield: 77 % (7.4 µmol). ^{1}H-NMR (400 MHz, D_2O): 8.48 (t, *J* = 13.4 Hz, 1H, H-b), 8.07 (s, 1H, H-8), 7.94 - 7.83 (m, 4H, 2 x H-B, 2 x H-C), 7.39 (t, *J* = 9.0 Hz, 2H, 2 x H-A), 6.41 (d, *J* = 7.5 Hz, 1H, H-a), 6.38 (d, *J* = 7.3 Hz, 1H, H-c), 6.24 (t, *J* = 6.9 Hz, 1H, H-1´), 4.74 - 4.71 (m, 1H, H-3´), 4.20 - 4.08 (m, 5H, H-4´, H-5´a/b, -N-C**H**$_2$-CH_3), 3.88 (q, *J* = 6.7 Hz, 2H, -PO_4-CH_2-), 3.41 (q, *J* = 7.3 Hz, 2H, H-V), 3.01 (t, *J* = 7.2 Hz, 2H, -CH_2-NH-), 2.72 (dt, *J* = 7.1 Hz, 13.8 Hz, 1H, H-2´a), 2.43 (ddd, *J* = 3.4 Hz, 6.4 Hz, 14.1 Hz, 1H, H-2´a), 2.22 (t, *J* = 7.0 Hz, 2H, H-I), 1.85 - 1.91 (m, 2H, H-IV), 1.73 (s, 6H, 2x -CH_3), 1.72 (s, 6H, 2x -CH_3), 1.70 - 1.66 (m, 2H, H-II), 1.64 - 1.56 (m, 4H, H-linker), 1.58 - 1.53 (m, 2H, H-III), 1.47 (t, *J* = 7.4 Hz, 3H, N-CH_2-C**H**$_3$), 1.35 - 1.31 (m, 4H, H-linker) ppm. ^{31}P-NMR (162 MHz, D_2O): -10.47 (d, *J* = 19.0 Hz, 1P, α-P), -11.19 (d, *J* = 18.5 Hz, 1P, γ-P), -22.58 (t, *J* = 18.5 Hz, β-P) ppm. HR-ESI-MS [M-1H]$^-$: m/z calculated: 1217.2897 m/z found: 1217.2774. ε = 162 000 L mol^{-1}cm^{-1} (λ= 563 nm).

6.1.72. δ-Azidohexyl-2´-deoxyguanosine-5´-*O*-triphosphate 8[103]

1-azido-6-hexylphosphate (155 mg, 0.7 mmol, 10.0 eq.) and the tetrabutylammonium salt of dGTP (70 µmol, 1 eq.) were each dissolved in 1 ml abs. DMF and stored over molecular sieve under

nitrogen atmosphere overnight in separate flasks. 1-ethyl-3-(3-dimethylaminopropyl)carbodiimid (EDC (hydrochloride), 40 mg, 0.21 mmol, 3 eq.) was added to the nucleotide solution and incubated at room temperature under nitrogen atmosphere for 3 h. Methanol (23 µl, 0.21 mmol, 3 eq.) was added to quench the remaining EDC and the mixture was stored under nitrogen atmosphere for 2 h. The solution of 6-azido-hexylphosphate in abs. DMF and dry triethylamine (194 µl, 1.4 mmol, 20.0 eq.) were added to the solution and the mixture was incubated at 40°C for 3 days. The reaction mixture was subsequently concentrated *in vacuo* and the residue was dissolved in water to perform purification by ion exchange chromatography followed by purification by RP-HPLC and repeated freeze drying. Yield: 18 % (12.5 µmol). ^{1}H-NMR (400 MHz, MeOD-d_5): 7.98 (s, 1H, H-8), 6.23 (t, *J* = 7.0 Hz, 1H, H-1´), 4.82 - 4.79 (m, 1H, H-3´), 4.36 (dt, *J* = 5.8 Hz, 11.2 Hz, 1H, H-5´a), 4.28 - 4.21 (m, 1H, H-4´), 4.17 - 4.14 (m, 1H, H-5´b), 4.03 (q, *J* = 6.5 Hz, 2H, -PO_4-CH_2-), 3.26 (t, *J* = 7.0 Hz, 2H, -CH_2-N_3), 2.92 (dt, *J* = 6.7 Hz, 13.5 Hz, 1H, H-2´a), 2.31 (ddd, *J* = 3.1 Hz, 6.4 Hz, 13.6 Hz, 1H, H-2´a), 1.65 (p, *J* = 7.1 Hz, 2H, H-linker), 1.57 (p, *J* = 6.9 Hz, 2H, H-linker), 1.45 - 1.37 (m, 4H, H-linker) ppm. ^{31}P-NMR (162 MHz, D_2O): -11.15 (d, *J* = 15.6 Hz, 1P, α-P), -11.48 (d, *J* = 16.1 Hz, 1P, γ-P), -23.11 (t, *J* = 15.6 Hz, β-P) ppm. HR-ESI-MS $[M-1H]^-$: m/z calculated: 711.0501 m/z found: 711.0487. ε = 8 743 L $mol^{-1}cm^{-1}$ (λ= 279 nm).

6.2. Molecular Biology/ Biochemical Methods

6.2.1. Analytical Denaturing PAGE

Denaturing polyacrylamide gels (12 % or 15 %) were prepared by polymerization of a solution of bisacrylamide/acrylamide (12 % or 15 %) and urea (8.3 M) in TBE buffer using peroxodisulfate (APS, 0.08 %) and *N,N,N´,N´*-tetramethylethylene diamine (TEMED, 0.04 %) for initiation of the polymerization. Following, the solution was filled in a sequencing gel chamber (Bio-Rad) and left for polymerization for at least 30 min. After addition of TBE buffer (1 x) to the electrophoresis unit, the gel was prewarmed by electrophoresis at 100 W for 20 min before samples were applied to the gel. After electrophoresis at 100 W for approximately 2 h, the gel was transferred to Whatman filter paper and dried at 80 °C under reduced pressure using a gel dryer (Model 583, Bio-Rad). The dried gel was exposed to an imager screen overnight and read out was performed using a molecular imager (FX, Bio-Rad). Quantification was done using the Bio-Rad software Image Lab.ink.

6.2.2. Preparative Denaturing PAGE

Preparative denaturing polyacrylaminde gels were used for purification of DNA oligonucleotides. PAGE-gels were prepared (8-12 %) according to the DNA fragment's size. Gels were prepared as described before for analytical gels (6.2.1. Analytical Denaturing PAGE). The solution was applied to the gel chamber (Roth) with a final gel thickness of 1.5 mm. DNA oligonucleotides were supplied with 1x denaturing PAGE loading Dye. DNA fragments were separated in 1 x TBE buffer by applying 100 W at up to 50°C. The desired DNA fragments were following visualized by UV light (254 nm) using TLC plated coated with a fluorescence indicator and excised. Subsequently the gel was mechanically shredded, water was added and the suspension was incubated at 50°C over night. The oligonucleotide solution was separated from remaining gel pieces by filtration through silanised glass-fibres wool. DNA was following ethanol precipitated, dissolved in ddH_2O and concentrations were determined by absorption at 260 nm.

6.2.3 Ethanol Precipitation

The DNA containing solution was mixed with 1/10 volume of 3 M sodium acetate solution (pH 5.2) and 2.5 volumes of ice-cold 100 % ethanol. The resulting mixture was incubated at -20°C over night and following centrifuged for 30 min at 4°C (20.000 x g). The supernatant was discarded and the remaining pellet washed with 70 % ice-cold ethanol. After, the mixture was centrifuged at 4°C for 30 min (20.000 x g) and the supernatant was discarded. This washing/centrifugation step was repeated. Following, the resulting DNA pellet was dried *in vacuo*, dissolved in ddH_20 and stored at -20°C.

6.2.4 DNA or RNA Concentration Determination

DNA or RNA concentration was determined by employing the Nanodrop ND1000. 2 µL were loaded onto the pedestals of the instrument and the absorption of the DNA or RNA sample was measured at 260 nm with water or the respective buffer as reference. Concentrations were calculated after Lambert Beers Law using the particular extinction coefficients.

6.2.5. 5´-Radioactive Labelling[88a]

DNA oligonucleotide primers were radioactively labelled at the 5´-terminus by usage of T4-PNK and [γ-^{32}P]-ATP. The reactions contained 0.4 µM primer, 1 x PNK reaction buffer, 0.8 µCi/µl [γ-^{32}P]-ATP and 0.4 U/µl T4 PNK in a total volume of 50 µl. The reaction mixture was incubated at 37 °C for 1 h. The labelling reaction was stopped after 1 h by denaturation of the enzyme for 2 min at 95 °C. Buffer and excess [γ-^{32}P]-ATP were removed by gelfitration (MicroSpin Sephadex G-25). Addition of 20 µl of unlabelled primer (10 µM) led to a final concentration of 3 µM of diluted radioactive labelled primer.

6.2.6. Primer Extension Assay[136]

A mixture of 150 nM of a [γ-^{32}P]-labeled primer and 200 nM of the respective template in 1x reaction buffer was heated to 95 °C for 2 min and subsequently cooled to 4 °C for annealing. The DNA polymerase was added and the reaction was started by addition of 50 µM of the respective dNTP at 55 °C. Reactions (10 – 20 µl) were stopped after the desired incubation time by addition of 2 µl of the reaction mixture to 10 µl stop solution (80 % (v/v) formamide, 20 mM EDTA, 0.25 % (w/v) bromophenol blue, 0.25 % (w/v) xylene cyanol) and analysed by 12 % or 15 % denaturing PAGE. Visualization was performed by phosphorimaging.

6.2.7. Enzyme Kinetics[114]

Steady-state kinetics of KOD exo$^-$ and KlenTaq were measured under single completed hit conditions. Single-nucleotide incorporation for the respective dNTPs in complex with the respective templates was performed as described, analysed by 12 % or 15 % denaturating PAGE and visualized by phosphorimaging. Concentrations of the DNA polymerases were chosen, that less than 20 % of the applied primer was extended under the used time points. The rate of single-nucleotide incorporation was determined at various dNTP concentrations for different incubation times varying from 10 s – 120 s. The amount of extended primer was plotted against incubation time for each examined dNTP concentration. For kinetic analysis, the reaction velocities divided through DNA polymerase concentrations were plotted against the employed dNTP concentrations. Using OriginPro8, experimental data was fit to a hyperbolic equation [velocity] = v_{Max}[dNTP]/(K_M+[dNTP]) to determine K_M and k_{cat}.

6.2.8. Library Expression and Screening

KOD exo$^-$ polymerase variants of the employed library were expressed in 96-well plates. Cells were grown in 70 µl LB-media per well supplemented with 100 mg/mL carbenicillin by inoculation with 5 µL of a 384-deepwell library stock. Cells were grown at 37°C for 2 h and mildly shaken. Protein expression was induced by addition of IPTG (0.5 mM final concentration). After protein expression at 37 °C for 2.5 h, cells were harvested by centrifugation at 4000 x g, for 20 min at 4 °C. Pellets were lysed using lysozyme at a concentration of 0.1 mg/mL at 37°C for 20 min in 1x *KOD* Reaction Buffer (200 µL per well). After heat-denaturation at 75 °C for 45 min bacterial cell debris were removed by centrifugation (4000 x g, 60 min at 4 °C). The supernatant was used for primer extension studies.

6.2.9. Capillary Electrophoresis

Capillary Electrophoresis separates fluorescently labelled oligonucleotides by size and charge as they migrate through a polymer filled capillary. Single nucleotide incorporation primer extension experiments are conducted as described (6.2.6. Primer Extension Assay[136]). Instead of radioactively labelled primers, FAM labelled ones are employed and reactions are stopped by addition of stop solution without bromophenol blue and xylene cyanol (80 % (v/v) formamide, 20 mM EDTA). CE samples are prepared by mixing 10 µl LIZ size standard diluted 1:80 in HiDi formamide (Applied Biosystems) and 10 µl fluorescently labelled reaction mixture (after addition of stop solution diluted 1:10 with HiDi formamide, to dilute EDTA). If multiple reactions are to be analysed in parallel, those are pooled in one well prior to injection. The reaction sample and size standard are injected electrokinetically into a 50 cm capillary array filled with Performance Optimised Polymer (POP6). High voltage electrophoresis (15 kV) over 180 s ensures single base resolution. Typical run parameters are depicted in **Table 4**.

Table 4 Typical run parameters for the Genetic Analyser 3500 analysing single-nucleotide incorporation primer extension experiments.

polymer	POP6
capillary length	50 cm
dye set	G5
application type	fragment
oven temperature	60 °C
buffer temperature	35 °C
pre run voltage	15 kV
pre run time	180 s
injection voltage	2 kV
injection time	10 s
run voltage	15 kV
run time	4080 s
size standard	120 LIZ

Migration of the fluorescently labelled primers and size standard fragments are recorded by a CCD camera. All reactions were separated using a 3500 Genetic Analyzer (Applied Biosystems) and fluorescent peaks were analysed using Gene Mapper version 5 (Applied Biosystems).

6.2.10. Assay RNA

A mixture of 150 nM of primer and 200 nM of the respective RNA template in 1x reaction buffer was heated to 95 °C for 2 min and subsequently cooled to 4 °C for annealing. The DNA polymerase was added and the reaction was started by addition of 50 µM of the respective mixture of nucleotide **9n** and ddATP at 55 °C. After 30 min incubation, 1 µl of 1mM dNTP mix (dATP, dCTP, dGTP, dTTP) was added and the reaction was incubated for 60 min at 55 °C. RNA template was digested by addition of 2 U/µl RNase If and 0.02 U/µl RNase H. The reaction was incubated at 37°C for 1 h. The reaction mixture was subsequently subjected to real-time PCR.

Therefore, 1 µl of the respective reaction mixture was used as template. 1 x Taq master mix (myPols), 200 nM of a forward and reverse primer, 0.5 x Sybr Green and 50 nM aptamer were applied for the real-time PCR reaction. Following PCR cycle was employed: initial denaturation for 2 min at 95 °C; varying number of cycles (10 - 50) of denaturation for 15 s at 95°C, annealing for 30 s at 62.5°C and amplification for 6 s at 72 °C. High resolution melting was conducted as last step

7. References

[1] J. v. Watzdorf, K. Leitner, A. Marx, *Angew. Chem. Int. Ed. Engl.* **2016**, *55*, 3229 - 3232.
[2] K. Chen, B. S. Zhao, C. He, *Cell chem. biol.* **2016**, *23*, 74-85.
[3] T. Carell, C. Brandmayr, A. Hienzsch, M. Muller, D. Pearson, V. Reiter, I. Thoma, P. Thumbs, M. Wagner, *Angew. Chem. Int. Ed. Engl.* **2012**, *51*, 7110-7131.
[4] P. Schuler, A. K. Miller, *Angew. Chem. Int. Ed. Engl.* **2012**, *51*, 10704-10707.
[5] a) M. Frommer, L. E. McDonald, D. S. Millar, C. M. Collis, F. Watt, G. W. Grigg, P. L. Molloy, C. L. Paul, *Proc. Natl. Acad. Sci. U S A* **1992**, *89*, 1827-1831;
b) E. Li, C. Beard, R. Jaenisch, *Nature* **1993**, *366*, 362-365;
c) D. W. Clough, L. M. Kunkel, R. L. Davidson, *Science* **1982**, *216*, 70-73.
[6] F. Kalousek, N. R. Morris, *The J. Biol. Chem.* **1968**, *243*, 2440-2442.
[7] a) T. H. Bestor, V. M. Ingram, *Proc. Natl. Acad. Sci. U S A* **1983**, *80*, 5559-5563;
b) D. Jahner, H. Stuhlmann, C. L. Stewart, K. Harbers, J. Lohler, I. Simon, R. Jaenisch, *Nature* **1982**, *298*, 623-628.
[8] a) P. A. Jones, S. M. Taylor, *Cell* **1980**, *20*, 85-93;
b) T. Mohandas, R. S. Sparkes, L. J. Shapiro, *Science* **1981**, *211*, 393-396.
[9] A. P. Bird, A. P. Wolffe, *Cell* **1999**, *99*, 451-454.
[10] a) J. Boyes, A. Bird, *Cell* **1991**, *64*, 1123-1134;
b) S. M. Iguchi-Ariga, W. Schaffner, *Genes Dev.* **1989**, *3*, 612-619;
c) I. Kovesdi, R. Reichel, J. R. Nevins, *Proc. Natl. Acad. Sci. U S A* **1987**, *84*, 2180-2184;
d) F. Watt, P. L. Molloy, *Genes Dev.* **1988**, *2*, 1136-1143.
[11] a) D. Aran, G. Toperoff, M. Rosenberg, A. Hellman, *Hum. Mol. Genet.* **2011**, *20*, 670-680;
b) M. P. Ball, J. B. Li, Y. Gao, J. H. Lee, E. M. LeProust, I. H. Park, B. Xie, G. Q. Daley, G. M. Church, *Nature biotechnology* **2009**, *27*, 361-368;
c) A. Hellman, A. Chess, *Science* **2007**, *315*, 1141-1143;
d) L. Laurent, E. Wong, G. Li, T. Huynh, A. Tsirigos, C. T. Ong, H. M. Low, K. W. Kin Sung, I. Rigoutsos, J. Loring, C. L. Wei, *Genome Res.* **2010**, *20*, 320-331;
e) R. Lister, M. Pelizzola, R. H. Dowen, R. D. Hawkins, G. Hon, J. Tonti-Filippini, J. R. Nery, L. Lee, Z. Ye, Q. M. Ngo, L. Edsall, J. Antosiewicz-Bourget, R. Stewart, V. Ruotti, A. H. Millar, J. A. Thomson, B. Ren, J. R. Ecker, *Nature* **2009**, *462*, 315-322;
f) T. A. Rauch, X. Wu, X. Zhong, A. D. Riggs, G. P. Pfeifer, *Proc. Nat.l Acad. Sc.i U S A* **2009**, *106*, 671-678;
g) Y. J. Shann, C. Cheng, C. H. Chiao, D. T. Chen, P. H. Li, M. T. Hsu, *Genome Res.* **2008**, *18*, 791-801;
h) X. Yang, H. Han, D. D. De Carvalho, F. D. Lay, P. A. Jones, G. Liang, *Cancer Cell* **2014**, *26*, 577-590.
[12] M. M. Suzuki, A. Bird, *Nat. Rev. Genet.* **2008**, *9*, 465-476.
[13] a) W. Mayer, A. Niveleau, J. Walter, R. Fundele, T. Haaf, *Nature* **2000**, *403*, 501-502;
b) J. Oswald, S. Engemann, N. Lane, W. Mayer, A. Olek, R. Fundele, W. Dean, W. Reik, J. Walter, *Cur. Biol.* **2000**, *10*, 475-478;
c) S. C. Wu, Y. Zhang, *Nat. Rev. Mol. Cell Biol.* **2010**, *11*, 607-620.
[14] a) S. Ito, A. C. D'Alessio, O. V. Taranova, K. Hong, L. C. Sowers, Y. Zhang, *Nature* **2010**, *466*, 1129-1133;
b) S. Kriaucionis, N. Heintz, *Science* **2009**, *324*, 929-930;
c) M. Tahiliani, K. P. Koh, Y. Shen, W. A. Pastor, H. Bandukwala, Y. Brudno, S. Agarwal, L. M. Iyer, D. R. Liu, L. Aravind, A. Rao, *Science* **2009**, *324*, 930-935.
[15] a) Y. F. He, B. Z. Li, Z. Li, P. Liu, Y. Wang, Q. Tang, J. Ding, Y. Jia, Z. Chen, L. Li, Y. Sun, X. Li, Q. Dai, C. X. Song, K. Zhang, C. He, G. L. Xu, *Science* **2011**, *333*, 1303-1307;
b) S. Ito, L. Shen, Q. Dai, S. C. Wu, L. B. Collins, J. A. Swenberg, C. He, Y. Zhang, *Science* **2011**, *333*, 1300-1303;
c) T. Pfaffeneder, B. Hackner, M. Truss, M. Munzel, M. Muller, C. A. Deiml, C. Hagemeier, T. Carell, *Angew. Chem. Int. Ed. Engl.* **2011**, *50*, 7008-7012.
[16] A. Maiti, A. C. Drohat, *The J. Biol. Chem.* **2011**, *286*, 35334-35338.
[17] A. Inoue, Y. Zhang, *Science* **2011**, *334*, 194.
[18] M. Munzel, D. Globisch, T. Carell, *Angew. Chem. Int. Ed. Engl.* **2011**, *50*, 6460-6468.
[19] J. P. Thomson, P. J. Skene, J. Selfridge, T. Clouaire, J. Guy, S. Webb, A. R. Kerr, A. Deaton, R. Andrews, K. D. James, D. J. Turner, R. Illingworth, A. Bird, *Nature* **2010**, *464*, 1082-1086.

[20] a) C. James, F. Mazurier, S. Dupont, R. Chaligne, I. Lamrissi-Garcia, M. Tulliez, E. Lippert, F. X. Mahon, J. M. Pasquet, G. Etienne, F. Delhommeau, S. Giraudier, W. Vainchenker, H. de Verneuil, *Blood* **2008**, *112*, 2429-2438;
b) M. Ko, Y. Huang, A. M. Jankowska, U. J. Pape, M. Tahiliani, H. S. Bandukwala, J. An, E. D. Lamperti, K. P. Koh, R. Ganetzky, X. S. Liu, L. Aravind, S. Agarwal, J. P. Maciejewski, A. Rao, *Nature* **2010**, *468*, 839-843;
c) C. G. Lian, Y. Xu, C. Ceol, F. Wu, A. Larson, K. Dresser, W. Xu, L. Tan, Y. Hu, Q. Zhan, C. W. Lee, D. Hu, B. Q. Lian, S. Kleffel, Y. Yang, J. Neiswender, A. J. Khorasani, R. Fang, C. Lezcano, L. M. Duncan, R. A. Scolyer, J. F. Thompson, H. Kakavand, Y. Houvras, L. I. Zon, M. C. Mihm, Jr., U. B. Kaiser, T. Schatton, B. A. Woda, G. F. Murphy, Y. G. Shi, *Cell* **2012**, *150*, 1135-1146;
d) C. Liu, L. Liu, X. Chen, J. Shen, J. Shan, Y. Xu, Z. Yang, L. Wu, F. Xia, P. Bie, Y. Cui, X. W. Bian, C. Qian, *PloS one* **2013**, *8*, e62828.
[21] C. S. Nabel, S. A. Manning, R. M. Kohli, *ACS Chem. Biol.* **2012**, *7*, 20-30.
[22] a) R. J. Klose, A. P. Bird, *Trends Biochem. Sci.* **2006**, *31*, 89-97;
b) L. Lercher, M. A. McDonough, A. H. El-Sagheer, A. Thalhammer, S. Kriaucionis, T. Brown, C. J. Schofield, *Chem. Com.* **2014**, *50*, 1794-1796.
[23] a) D. Globisch, M. Munzel, M. Muller, S. Michalakis, M. Wagner, S. Koch, T. Bruckl, M. Biel, T. Carell, *PloS one* **2010**, *5*, e15367;
b) M. Munzel, D. Globisch, T. Bruckl, M. Wagner, V. Welzmiller, S. Michalakis, M. Muller, M. Biel, T. Carell, *Angew. Chem. Int. Ed. Engl.* **2010**, *49*, 5375-5377;
c) C. X. Song, K. E. Szulwach, Y. Fu, Q. Dai, C. Yi, X. Li, Y. Li, C. H. Chen, W. Zhang, X. Jian, J. Wang, L. Zhang, T. J. Looney, B. Zhang, L. A. Godley, L. M. Hicks, B. T. Lahn, P. Jin, C. He, *Nat. Biotechnol.* **2011**, *29*, 68-72;
d) A. Szwagierczak, S. Bultmann, C. S. Schmidt, F. Spada, H. Leonhardt, *Nucleic Acids Res.* **2010**, *38*, e181.
[24] L. Tan, Y. G. Shi, *Development* **2012**, *139*, 1895-1902.
[25] M. R. Branco, G. Ficz, W. Reik, *Nat. Rev. Genet.* **2012**, *13*, 7-13.
[26] a) G. Ficz, M. R. Branco, S. Seisenberger, F. Santos, F. Krueger, T. A. Hore, C. J. Marques, S. Andrews, W. Reik, *Nature* **2011**, *473*, 398-402;
b) H. Wu, A. C. D'Alessio, S. Ito, Z. Wang, K. Cui, K. Zhao, Y. E. Sun, Y. Zhang, *Genes Dev.* **2011**, *25*, 679-684.
[27] S. H. Cross, J. A. Charlton, X. Nan, A. P. Bird, *Nat. Gen.* **1994**, *6*, 236-244.
[28] J. Kaput, T. W. Sneider, *Nucleic Acids Res.* **1979**, *7*, 2303-2322.
[29] W. W. Li, L. Gong, H. Bayley, *Angew. Chem. Int. Ed. Engl.* **2013**, *52*, 4350-4355.
[30] G. Kubik, M. J. Schmidt, J. E. Penner, D. Summerer, *Angew. Chem. Int. Ed. Engl.* **2014**, *53*, 6002-6006.
[31] a) T. Wang, T. Hong, T. Tang, Q. Zhai, X. Xing, W. Mao, X. Zheng, L. Xu, J. Wu, X. Weng, S. Wang, T. Tian, B. Yuan, B. Huang, L. Zhuang, X. Zhou, *J. Am. Chem. Soc.* **2013**, *135*, 1240-1243;
b) S. Bareyt, T. Carell, *Angew. Chem. Int. Ed. Engl.* **2008**, *47*, 181-184.
[32] H. Hayatsu, Y. Wataya, K. Kazushige, *J. Am. Chem. Soc.* **1970**, *92*, 724-726.
[33] R. Shapiro, R. E. Servis, M. Welcher, *J. Am. Chem. Soc.* **1970**, *92*, 422-&.
[34] S. J. Clark, J. Harrison, C. L. Paul, M. Frommer, *Nucleic Acids Res.* **1994**, *22*, 2990-2997.
[35] O. Ammerpohl, J. I. Martin-Subero, J. Richter, I. Vater, R. Siebert, *Biochim. Biophy. Acta* **2009**, *1790*, 847-862.
[36] C. Grunau, S. J. Clark, A. Rosenthal, *Nucleic Acids Res.* **2001**, *29*, E65-65.
[37] J. G. Herman, J. R. Graff, S. Myohanen, B. D. Nelkin, S. B. Baylin, *Proc. Natl. Acad. Sci. U S A* **1996**, *93*, 9821-9826.
[38] a) M. L. Gonzalgo, P. A. Jones, *Nucleic Acids Res.* **1997**, *25*, 2529-2531;
b) M. L. Gonzalgo, P. A. Jones, *Methods* **2002**, *27*, 128-133.
[39] A. M. Maxam, W. Gilbert, *Proc. Natl. Acad. Sci U S A* **1977**, *74*, 560-564.
[40] A. Schmitz, M. Short, O. Ammerpohl, C. Asbrand, J. Nickel, R. Renkawitz, *J. Biol. Chem.* **1997**, *272*, 20850-20856.
[41] H. Ohmori, J. I. Tomizawa, A. M. Maxam, *Nucleic Acids Res.* **1978**, *5*, 1479-1485.
[42] H. Hayatsu, M. Shiragami, *Biochemistry* **1979**, *18*, 632-637.
[43] M. Yu, G. C. Hon, K. E. Szulwach, C. X. Song, L. Zhang, A. Kim, X. Li, Q. Dai, Y. Shen, B. Park, J. H. Min, P. Jin, B. Ren, C. He, *Cell* **2012**, *149*, 1368-1380.
[44] M. J. Booth, M. R. Branco, G. Ficz, D. Oxley, F. Krueger, W. Reik, S. Balasubramanian, *Science* **2012**, *336*, 934-937.
[45] J. P. Bachellerie, J. Cavaille, A. Huttenhofer, *Biochimie* **2002**, *84*, 775-790.
[46] X. Wang, C. He, *Mol. Cell* **2014**, *56*, 5-12.

[47] a) W. E. Cohn, E. Volkin, *Nature* **1951**, *167*, 483-484; bJ. Ge, Y. T. Yu, *Trends Biochem. Sci.* **2013**, *38*, 210-218.
[48] M. Charette, M. W. Gray, *IUBMB life* **2000**, *49*, 341-351.
[49] B. S. Zhao, C. He, *Cell Res.* **2015**, *25*, 153-154.
[50] a) T. M. Carlile, M. F. Rojas-Duran, B. Zinshteyn, H. Shin, K. M. Bartoli, W. V. Gilbert, *Nature* **2014**, *515*, 143-146;
b) S. Schwartz, D. A. Bernstein, M. R. Mumbach, M. Jovanovic, R. H. Herbst, B. X. Leon-Ricardo, J. M. Engreitz, M. Guttman, R. Satija, E. S. Lander, G. Fink, A. Regev, *Cell* **2014**, *159*, 148-162.
[51] A. F. Lovejoy, D. P. Riordan, P. O. Brown, *PloS one* **2014**, *9*, e110799.
[52] J. Ofengand, M. Del Campo, Y. Kaya, *Methods* **2001**, *25*, 365-373.
[53] J. Aschenbrenner, A. Marx, *Nucleic Acids Res.* **2016**, *44*, 3495-3502.
[54] W. A. Decatur, M. J. Fournier, *Trends Biochem. Sci.* **2002**, *27*, 344-351.
[55] S. Daffis, K. J. Szretter, J. Schriewer, J. Li, S. Youn, J. Errett, T. Y. Lin, S. Schneller, R. Zust, H. Dong, V. Thiel, G. C. Sen, V. Fensterl, W. B. Klimstra, T. C. Pierson, R. M. Buller, M. Gale, Jr., P. Y. Shi, M. S. Diamond, *Nature* **2010**, *468*, 452-456.
[56] Q. Zou, J. Li, Q. Hong, Z. Lin, Y. Wu, H. Shi, Y. Ju, *BioMed Res. Int.* **2015**, *2015*, 810514.
[57] F. Sanger, S. Nicklen, A. R. Coulson, *Proc. Natl. Acad. Sci. U S A* **1977**, *74*, 5463-5467.
[58] a) L. Blanco, A. Bernad, J. M. Lazaro, G. Martin, C. Garmendia, M. Salas, *J. Biol. Chem.* **1989**, *264*, 8935-8940;
b) S. Tabor, H. E. Huber, C. C. Richardson, J. Biol. Chem. **1987**, *262*, 16212-16223.
[59] a) C. Feschotte, E. J. Pritham, *Annu. Rev. Genet.* **2007**, *41*, 331-368;
b) E. Tuzun, A. J. Sharp, J. A. Bailey, R. Kaul, V. A. Morrison, L. M. Pertz, E. Haugen, H. Hayden, D. Albertson, D. Pinkel, M. V. Olson, E. E. Eichler, *Nat. Genet.* **2005**, *37*, 727-732.
[60] J. Eid, A. Fehr, J. Gray, K. Luong, J. Lyle, G. Otto, P. Peluso, D. Rank, P. Baybayan, B. Bettman, A. Bibillo, K. Bjornson, B. Chaudhuri, F. Christians, R. Cicero, S. Clark, R. Dalal, A. Dewinter, J. Dixon, M. Foquet, A. Gaertner, P. Hardenbol, C. Heiner, K. Hester, D. Holden, G. Kearns, X. Kong, R. Kuse, Y. Lacroix, S. Lin, P. Lundquist, C. Ma, P. Marks, M. Maxham, D. Murphy, I. Park, T. Pham, M. Phillips, J. Roy, R. Sebra, G. Shen, J. Sorenson, A. Tomaney, K. Travers, M. Trulson, J. Vieceli, J. Wegener, D. Wu, A. Yang, D. Zaccarin, P. Zhao, F. Zhong, J. Korlach, S. Turner, *Science* **2009**, *323*, 133-138.
[61] a) E. Viennois, F. Chen, H. Laroui, M. T. Baker, D. Merlin, *BMC Res. notes* **2013**, *6*, 360;
b) M. D. Driscoll, J. Rentergent, S. Hay, *The FEBS journal* **2014**, *281*, 2042-2050.
[62] L. Greenough, K. M. Schermerhorn, L. Mazzola, J. Bybee, D. Rivizzigno, E. Cantin, B. E. Slatko, A. F. Gardner, *Nucleic Acids Res.* **2016**, *44*, e15.
[63] a) E. S. Lander, L. M. Linton, B. Birren, C. Nusbaum, M. C. Zody, J. Baldwin, K. Devon, K. Dewar, M. Doyle, W. FitzHugh, R. Funke, D. Gage, K. Harris, A. Heaford, J. Howland, L. Kann, J. Lehoczky, R. LeVine, P. McEwan, K. McKernan, J. Meldrim, J. P. Mesirov, C. Miranda, W. Morris, J. Naylor, C. Raymond, M. Rosetti, R. Santos, A. Sheridan, C. Sougnez, Y. Stange-Thomann, N. Stojanovic, A. Subramanian, D. Wyman, J. Rogers, J. Sulston, R. Ainscough, S. Beck, D. Bentley, J. Burton, C. Clee, N. Carter, A. Coulson, R. Deadman, P. Deloukas, A. Dunham, I. Dunham, R. Durbin, L. French, D. Grafham, S. Gregory, T. Hubbard, S. Humphray, A. Hunt, M. Jones, C. Lloyd, A. McMurray, L. Matthews, S. Mercer, S. Milne, J. C. Mullikin, A. Mungall, R. Plumb, M. Ross, R. Shownkeen, S. Sims, R. H. Waterston, R. K. Wilson, L. W. Hillier, J. D. McPherson, M. A. Marra, E. R. Mardis, L. A. Fulton, A. T. Chinwalla, K. H. Pepin, W. R. Gish, S. L. Chissoe, M. C. Wendl, K. D. Delehaunty, T. L. Miner, A. Delehaunty, J. B. Kramer, L. L. Cook, R. S. Fulton, D. L. Johnson, P. J. Minx, S. W. Clifton, T. Hawkins, E. Branscomb, P. Predki, P. Richardson, S. Wenning, T. Slezak, N. Doggett, J. F. Cheng, A. Olsen, S. Lucas, C. Elkin, E. Uberbacher, M. Frazier, et al., *Nature* **2001**, *409*, 860-921;
b) J. C. Venter, M. D. Adams, E. W. Myers, P. W. Li, R. J. Mural, G. G. Sutton, H. O. Smith, M. Yandell, C. A. Evans, R. A. Holt, J. D. Gocayne, P. Amanatides, R. M. Ballew, D. H. Huson, J. R. Wortman, Q. Zhang, C. D. Kodira, X. H. Zheng, L. Chen, M. Skupski, G. Subramanian, P. D. Thomas, J. Zhang, G. L. Gabor Miklos, C. Nelson, S. Broder, A. G. Clark, J. Nadeau, V. A. McKusick, N. Zinder, A. J. Levine, R. J. Roberts, M. Simon, C. Slayman, M. Hunkapiller, R. Bolanos, A. Delcher, I. Dew, D. Fasulo, M. Flanigan, L. Florea, A. Halpern, S. Hannenhalli, S. Kravitz, S. Levy, C. Mobarry, K. Reinert, K. Remington, J. Abu-Threideh, E. Beasley, K. Biddick, V. Bonazzi, R. Brandon, M. Cargill, I. Chandramouliswaran, R. Charlab, K. Chaturvedi, Z. Deng, V. Di Francesco, P. Dunn, K. Eilbeck, C. Evangelista, A. E. Gabrielian, W. Gan, W. Ge, F. Gong, Z. Gu, P. Guan, T. J. Heiman, M. E. Higgins, R. R. Ji, Z. Ke, K. A. Ketchum, Z. Lai, Y. Lei, Z. Li, J. Li, Y. Liang, X. Lin, F. Lu, G. V. Merkulov, N. Milshina, H. M. Moore, A. K. Naik, V. A. Narayan, B. Neelam, D. Nusskern, D. B. Rusch, S. Salzberg, W.

Shao, B. Shue, J. Sun, Z. Wang, A. Wang, X. Wang, J. Wang, M. Wei, R. Wides, C. Xiao, C. Yan, et al., *Science* **2001**, *291*, 1304-1351.
[64] A. J. Berdis, *Chem. Rev.* **2009**, *109*, 2862-2879.
[65] T. A. Steitz, *Harvey lectures* **1997**, *93*, 75-93.
[66] K. A. Johnson, *Annu. Rev. Biochem.* **1993**, *62*, 685-713.
[67] M. J. Donlin, S. S. Patel, K. A. Johnson, *Biochemistry* **1991**, *30*, 538-546.
[68] a) R. D. Kuchta, P. Benkovic, S. J. Benkovic, *Biochemistry* **1988**, *27*, 6716-6725;
b) M. W. Frey, L. C. Sowers, D. P. Millar, S. J. Benkovic, *Biochemistry* **1995**, *34*, 9185-9192.
[69] a) P. L. Boyer, H. Q. Gao, P. K. Clark, S. G. Sarafianos, E. Arnold, S. H. Hughes, *J. Virol.* **2001**, *75*, 6321-6328;
b) P. R. Meyer, S. E. Matsuura, A. G. So, W. A. Scott, *Proc. Natl. Acad. Sci. U S A* **1998**, *95*, 13471-13476.
[70] M. Xiao, A. Phong, K. L. Lum, R. A. Greene, P. R. Buzby, P. Y. Kwok, *Genome Res.* **2004**, *14*, 1749-1755.
[71] J. R. Knowles, *Annu.Rev. Biochem.* **1980**, *49*, 877-919.
[72] T. A. Steitz, *J. Biol. Chem.* **1999**, *274*, 17395-17398.
[73] a) C. Castro, E. D. Smidansky, J. J. Arnold, K. R. Maksimchuk, I. Moustafa, A. Uchida, M. Gotte, W. Konigsberg, C. E. Cameron, *Nat. Struct. Mol. Biol.* **2009**, *16*, 212-218;
b) L. S. Beese, T. A. Steitz, *The EMBO journal* **1991**, *10*, 25-33;
c) T. A. Steitz, *Cur. Opin. Struct. Biol.* **1993**, *3*, 31-38.
[74] a) M. Garcia-Diaz, K. Bebenek, J. M. Krahn, L. C. Pedersen, T. A. Kunkel, *DNA repair* **2007**, *6*, 1333-1340;
b) G. A. Cisneros, L. Perera, M. Garcia-Diaz, K. Bebenek, T. A. Kunkel, L. G. Pedersen, *DNA repair* **2008**, *7*, 1824-1834.
[75] C. A. Brautigam, T. A. Steitz, *Cur. Opin. Struct. Biol.* **1998**, *8*, 54-63.
[76] a)F. J. Bollum, *Adv. Exp. Med. Biol.* **1982**, *145*, 1-11;
b) O. Dominguez, J. F. Ruiz, T. Lain de Lera, M. Garcia-Diaz, M. A. Gonzalez, T. Kirchhoff, A. C. Martinez, A. Bernad, L. Blanco, *The EMBO journal* **2000**, *19*, 1731-1742;
c) I. Wong, S. S. Patel, K. A. Johnson, *Biochemistry* **1991**, *30*, 526-537.
[77] a) T. E. Barman, S. R. Bellamy, H. Gutfreund, S. E. Halford, C. Lionne, *Cell. Mol. Life Sci.* **2006**, *63*, 2571-2583;
b) K. A. Johnson, *Cur. Opin. Biotechnol.* **1998**, *9*, 87-89.
[78] F. R. Bryant, K. A. Johnson, S. J. Benkovic, *Biochemistry* **1983**, *22*, 3537-3546.
[79] S. Creighton, L. B. Bloom, M. F. Goodman, *Methods in Enzym.* **1995**, *262*, 232-256.
[80] J. Filee, P. Forterre, T. Sen-Lin, J. Laurent, *J. Mol. Evol.* **2002**, *54*, 763-773.
[81] a)vS. Korolev, M. Nayal, W. M. Barnes, E. Di Cera, G. Waksman, *Proc. Natl. Acad. Sci. U S A* **1995**, *92*, 9264-9268;
b) Y. Li, Y. Kong, S. Korolev, G. Waksman, *Protein science : a publication of the Protein Society* **1998**, *7*, 1116-1123.
[82] P. J. Rothwell, V. Mitaksov, G. Waksman, *Mol. Cell* **2005**, *19*, 345-355.
[83] M. J. Bessman, A. Kornberg, I. R. Lehman, E. S. Simms, *Biochimica et Biophysica Acta* **1956**, *21*, 197-198.
[84] H. Klenow, K. Overgaard-Hansen, S. A. Patkar, *European journal of biochemistry / FEBS* **1971**, *22*, 371-381.
[85] a) P. Setlow, D. Brutlag, A. Kornberg, *J. Biol. Chem.* **1972**, *247*, 224-231;
b) W. M. Barnes, *Gene* **1992**, *112*, 29-35.
[86] a) Y. Li, S. Korolev, G. Waksman, *The EMBO journal* **1998**, *17*, 7514-7525;
b) Y. Li, G. Waksman, *Protein science : a publication of the Protein Society* **2001**, *10*, 1225-1233.
[87] a) R. N. Veedu, B. Vester, J. Wengel, *Chembiochem* **2007**, *8*, 490-492;
b) R. N. Veedu, B. Vester, J. Wengel, *J. Am. Chem. Soc.* **2008**, *130*, 8124-8125;
c) M. Strerath, J. Gaster, D. Summerer, A. Marx, *Chembiochem* **2004**, *5*, 333-339.
[88] a) A. Baccaro, A. L. Steck, A. Marx, *Angew. Chem. Int. Ed. Engl.* **2012**, *51*, 254-257;
b) G. F. Kaufmann, M. M. Meijler, C. Sun, D. W. Chen, D. P. Kujawa, J. M. Mee, T. Z. Hoffman, P. Wirsching, R. A. Lerner, K. D. Janda, *Angew. Chem. Int. Ed. Engl.* **2005**, *44*, 2144-2148;
c) O. Thum, S. Jager, M. Famulok, *Angew. Chem. Int. Ed. Engl.* **2001**, *40*, 3990-3993;
d) H. A. Held, S. A. Benner, *Nucleic Acids Res.* **2002**, *30*, 3857-3869;
e) G. A. Burley, J. Gierlich, M. R. Mofid, H. Nir, S. Tal, Y. Eichen, T. Carell, *J. Am. Chem. Soc.* **2006**, *128*, 1398-1399;
f) J. Gierlich, K. Gutsmiedl, P. M. Gramlich, A. Schmidt, G. A. Burley, T. Carell, *Chemistry* **2007**, *13*, 9486-9494.

[89] a) A. C. Rodriguez, H. W. Park, C. Mao, L. S. Beese, *J. Mol. Biol.* **2000**, *299*, 447-462;
b) H. Hashimoto, M. Nishioka, S. Fujiwara, M. Takagi, T. Imanaka, T. Inoue, Y. Kai, *J. Mol. Biol.* **2001**, *306*, 469-477;
c) T. Kuroita, H. Matsumura, N. Yokota, M. Kitabayashi, H. Hashimoto, T. Inoue, T. Imanaka, Y. Kai, *J. Mol. Biol.* **2005**, *351*, 291-298.
[90] S. Moran, R. X. Ren, S. Rumney, E. T. Kool, *J. Am. Chem. Soc.* **1997**, *119*, 2056-2057.
[91] L. Guan, G. W. van der Heijden, A. Bortvin, M. M. Greenberg, *Chembiochem* **2011**, *12*, 2184-2190.
[92] J. Ludwig, F. Eckstein, *J. Org. Chem.* **1989**, *54*, 631-635.
[93] J. Caton-Williams, M. Smith, N. Carrasco, Z. Huang, *Organic letters* **2011**, *13*, 4156-4159.
[94] K. He, K. W. Porter, A. Hasan, Briley, Jd, B. R. Shaw, *Nucleic Acids Res.* **1999**, *27*, 1788-1794.
[95] M. Yoshikawa, T. Kato, Takenish.T, *B. Chem. Soc. Jpn.* **1969**, *42*, 3505.
[96] D. E. Hoard, D. G. Ott, *J. Am. Chem. Soc.* **1965**, *87*, 1785-1788.
[97] T. Kovacs, L. Otvos, *Tetrahedron Lett.* **1988**, *29*, 4525-4528.
[98] a) A. Sassa, W. A. Beard, D. D. Shock, S. H. Wilson, *JoVE* **2013**, e50695;
b) B. Holzberger, J. Strohmeier, V. Siegmund, U. Diederichsen, A. Marx, *Bioorg. Med. Chem. Lett.* **2012**, *22*, 3136-3139.
[99] a) D. Chatterji, V. Gopal, *Methods Enzymol.* **1996**, *274*, 456-478;
b) J. G. Schlageck, M. Baughman, L. R. Yarbrough, *J. Biol. Chem.* **1979**, *254*, 12074-12077;
c) L. R. Yarbrough, J. G. Schlageck, M. Baughman, *J. Biol. Chem.* **1979**, *254*, 12069-12073.
[100] a) S. Kumar, A. Sood, J. Wegener, P. J. Finn, S. Nampalli, J. R. Nelson, A. Sekher, P. Mitsis, J. Macklin, C. W. Fuller, *Nucleosides, Nucleotides & Nucleic Acids* **2005**, *24*, 401-408;
b) A. Sood, S. Kumar, S. Nampalli, J. R. Nelson, J. Macklin, C. W. Fuller, *J. Am. Chem. Soc.* **2005**, *127*, 2394-2395.
[101] V. Nair, F. Zhang, X. Ma, E. Bonsu, *Nucleosides, Nucleotides & Nucleic Acids* **2009**, *28*, 408-423.
[102] Y. Wang, D. Rosner, M. Grzywa, A. Marx, *Angew. Chem. Int. Ed. Engl.* **2014**, *53*, 8159-8162.
[103] S. M. Hacker, M. Mex, A. Marx, *J. Org. Chem.* **2012**, *77*, 10450-10454.
[104] J. Ito, D. K. Braithwaite, *Nucleic Acids Res.* **1991**, *19*, 4045-4057.
[105] T. T. Richardson, L. Gilroy, Y. Ishino, B. A. Connolly, G. Henneke, *Nucleic Acids Res.* **2013**, *41*, 4207-4218.
[106] J. S. Woo, S. T. Sigurdsson, P. B. Hopkins, *J. Am. Chem. Soc.* **1993**, *115*, 3407-3415.
[107] Z. Janeba, P. Francom, M. J. Robins, *J. Org. Chem.* **2003**, *68*, 989-992.
[108] N. B. Hanna, B. K. Bhattacharya, R. K. Robins, T. L. Avery, G. R. Revankar, *J. Med. Chem.* **1994**, *37*, 177-183.
[109] M. Hoffer, *Chemische Berichte* **1960**, *93*, 2777 - 2781.
[110] F. Seela, P. Leonard, *Helv. Chim. Acta* **1997**, *80*, 1301-1318.
[111] a) F. Seela, P. Leonard, *Helv. Chim. Acta* **1996**, *79*, 477-487;
b) R. W. Balsiger, J. A. Montgomery, *J. Org. Chem.* **1960**, *25*, 1573-1575.
[112] a) F. Seela, B. Westermann, U. Bindig, *J. Chem. Soc. Perk. T 1* **1988**, 697-702;
b) H. D. Winkeler, F. Seela, *J. Org. Chem.* **1983**, *48*, 3119-3122.
[113] a) R. P. Hodge, C. K. Brush, C. M. Harris, T. M. Harris, *J. Org. Chem.* **1991**, *56*, 1553-1564;
b) L. K. Huang, Y. C. Cherng, Y. R. Cheng, J. P. Jang, Y. L. Chao, Y. J. Cherng, *Tetrahedron* **2007**, *63*, 5323-5327.
[114] a) J. Petruska, M. F. Goodman, M. S. Boosalis, L. C. Sowers, C. Cheong, I. Tinoco, Jr., *Proc. Natl. Acad. Sci. U S A* **1988**, *85*, 6252-6256;
b) M. S. Boosalis, J. Petruska, M. F. Goodman, *J. Biol. Chem.* **1987**, *262*, 14689-14696;
c) S. Creighton, M. M. Huang, H. Cai, N. Arnheim, M. F. Goodman, *J. Biol. Chem.* **1992**, *267*, 2633-2639.
[115] K. R. Mitchelson, *Methods Mol. Biol.* **2001**, *162*, 3-26.
[116] J. Aschenbrenner, M. Drum, H. Topal, M. Wieland, A. Marx, *Angew. Chem. Int. Ed. Engl.* **2014**, *53*, 8154-8158.
[117] a) G. S. Ti, B. L. Gaffney, R. A. Jones, *J. Am. Chem. Soc.* **1982**, *104*, 1316-1319;
b) H. Lang, M. Gottlieb, M. Schwarz, S. Farkas, B. S. Schulz, F. Himmelsbach, R. Charubala, W. Pfleiderer, *Helv. Chim. Acta* **1999**, *82*, 2172-2185.
[118] A. Laayoun, C. Coulombeau, J. F. Constant, M. Berger, J. Cadet, J. Lhomme, *Int. J. Rad. Biol.* **1994**, *66*, 259-266.
[119] C. E. Elmquist, J. S. Stover, Z. Wang, C. J. Rizzo, *J. Am. Chem. Soc.* **2004**, *126*, 11189-11201.
[120] L. K. Clayton, M. F. Goodman, E. W. Branscomb, D. J. Galas, *J. Biol. Chem.* **1979**, *254*, 1902-1912.

[121] a) T. H. Bestor, A. Coxon, *Current biology* **1993**, *3*, 384-386;
b) A. Bird, *Genes De.* **2002**, *16*, 6-21;
c) H. Cedar, Y. Bergman, *Nature Rev. Genet.* **2009**, *10*, 295-304.

[122] Y. Xu, F. Wu, L. Tan, L. Kong, L. Xiong, J. Deng, A. J. Barbera, L. Zheng, H. Zhang, S. Huang, J. Min, T. Nicholson, T. Chen, G. Xu, Y. Shi, K. Zhang, Y. G. Shi, *Mol. Cell* **2011**, *42*, 451-464.

[123] P. A. Limbach, P. F. Crain, J. A. McCloskey, *Nucleic Acid Res.* **1994**, *22*, 2183-2196.

[124] A. W. Kellner, *Anais da Academia Brasileira de Ciencias* **2010**, *82*, 799-800.

[125] N. Blatter, K. Bergen, O. Nolte, W. Welte, K. Diederichs, J. Mayer, M. Wieland, A. Marx, *Angew. Chem. Int. Ed. Engl.* **2013**, *52*, 11935-11939.

[126] a) D. Piekna-Przybylska, W. A. Decatur, M. J. Fournier, *Nucleic Acids Res.* **2008**, *36*, D178-183;
b) B. E. Maden, *Progr. Nucl. Acid Res. Mol. Biol.* **1990**, *39*, 241-303.

[127] a) B. Gonzales, D. Henning, R. B. So, J. Dixon, M. J. Dixon, B. C. Valdez, *Hum. Mol. Genet.* **2005**, *14*, 2035-2043;
b) G. A. Stepanov, J. A. Filippova, A. B. Komissarov, E. V. Kuligina, V. A. Richter, D. V. Semenov, *BioMed Res. Int.* **2015**, *2015*, 206849.

[128] S. Tabor, C. C. Richardson, *Proc. Natl. Acad. Sci. U S A* **1995**, *92*, 6339-6343.

[129] a) H. Heyn, M. Esteller, *Nature Rev. Genet.* **2012**, *13*, 679-692;
b) J. Sandoval, H. Heyn, S. Moran, J. Serra-Musach, M. A. Pujana, M. Bibikova, M. Esteller, *Epigenetics* **2011**, *6*, 692-702;
c) M. Bibikova, B. Barnes, C. Tsan, V. Ho, B. Klotzle, J. M. Le, D. Delano, L. Zhang, G. P. Schroth, K. L. Gunderson, J. B. Fan, R. Shen, *Genomics* **2011**, *98*, 288-295;
d) S. Dedeurwaerder, M. Defrance, E. Calonne, H. Denis, C. Sotiriou, F. Fuks, *Epigenomics* **2011**, *3*, 771-784.

[130] B. Holzberger, S. Obeid, W. Welte, K. Diederichs, A. Marx, *Chem. Sci.* **2012**, *3*, 2924-2931.

[131] K. Yamada, Y. Hattori, T. Inde, T. Kanamori, A. Ohkubo, K. Seio, M. Sekine, *Bioorg. Med. Chem. Lett.* **2013**, *23*, 776-778.

[132] J. Polanski, F. Zouhiri, L. Jeanson, D. Desmaele, J. d'Angelo, J. F. Mouscadet, R. Gieleciak, J. Gasteiger, M. Le Bret, *J. Med. Chem.* **2002**, *45*, 4647-4654.

[133] P. P. Kung, R. A. Jones, *Tetrahedron Lett.* **1991**, *32*, 3919-3922.

[134] K. Ramasamy, N. Imamura, R. K. Robins, G. R. Revankar, *J. Heterocyc. Chem.* **1988**, *25*, 1893-1898.

[135] H. Li, A. C. Fahrenbach, A. Coskun, Z. X. Zhu, G. Barin, Y. L. Zhao, Y. Y. Botros, J. P. Sauvage, J. F. Stoddart, *Angew. Chem. Int. Ed. Engl.* **2011**, *50*, 6782-6788.

[136] K. Bergen, A. L. Steck, S. Strutt, A. Baccaro, W. Welte, K. Diederichs, A. Marx, *J. Am. Chem. Soc.* **2012**, *134*, 11840-11843.

[137] V. Dolnik, *J. Biochem. Biophys. Methods,* **1999**, *41*, 103-119.

[138] a) O. Mitsonobu, *Synthesis,* **1981**, 1-20;
b) O. Mitsonobu, Y. Yanada, *Bull. Chem. Soc. Jpn.,* **1967**, *40*, 2380-2382

[139] J.F. Davis, Z. Hostomosk, S.R. Jordan, D.A. Matthews, *Science,* **1991**, *252,* 88-95.

[140] P.F. Spahr, B. R. Hollingworth, *J. Biol. Chem.,* **1961**, *236,* 823-831.

[141] D. Voet, J. Voet*,* **2001**, *Biochemistry (4. Ed.)*, New York, J. Wiley & Sons, 85

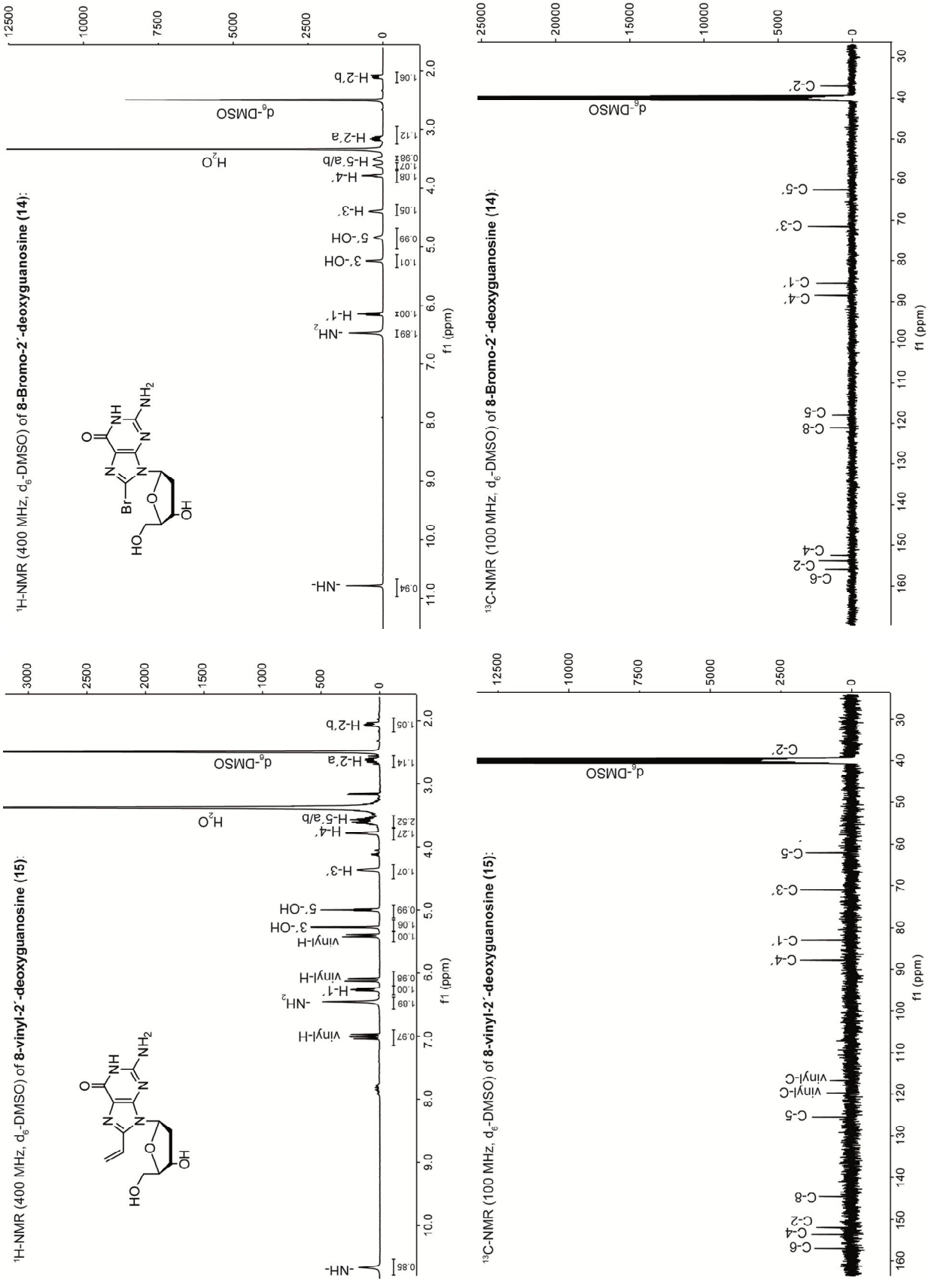
[142] D. Elliot, M. Ladomery*,* **2011**, *Molecular Biology of RNA (1. Ed.)*, Oxford University Press, 34 - 64

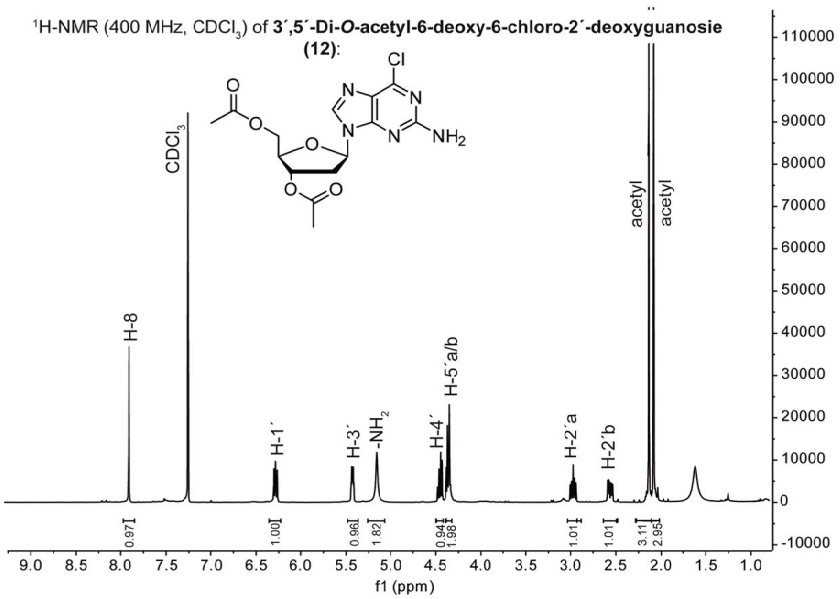
[143] W. Rychlik, W. J. Spencer, R. E. Rhoads*, Nucleic Acids Res.,* **1990**, *18*, 6409 - 6412.

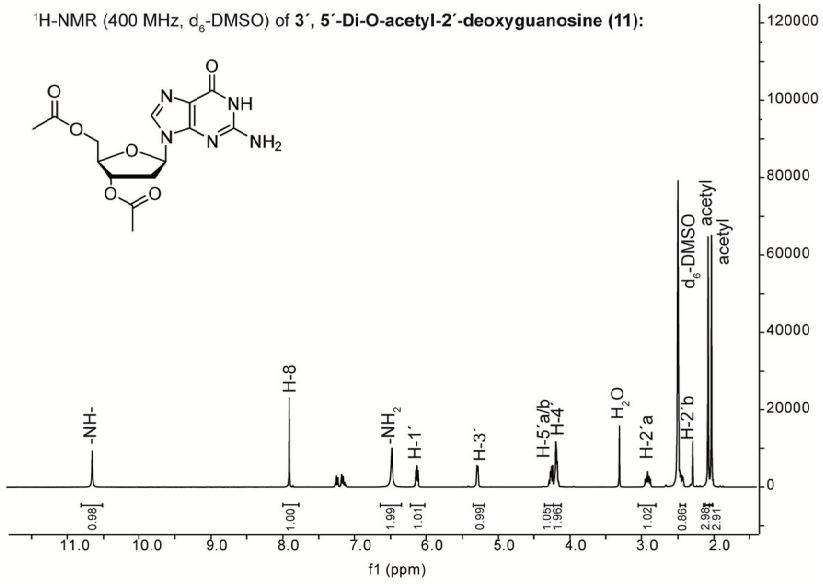
[144] a) Y. Lin, D.S. Jayasena, *J. Mol. Biol.,* **1997**, *271*, 100 - 111;
b) D.E. Kellogg, I. Rybalkin, S. Chen, N. Mukhamedov, T. Vlasik, P.D: Siebert, A. Chendik, *Biotechniques,* **1994**, *16*, 1134 - 1137.

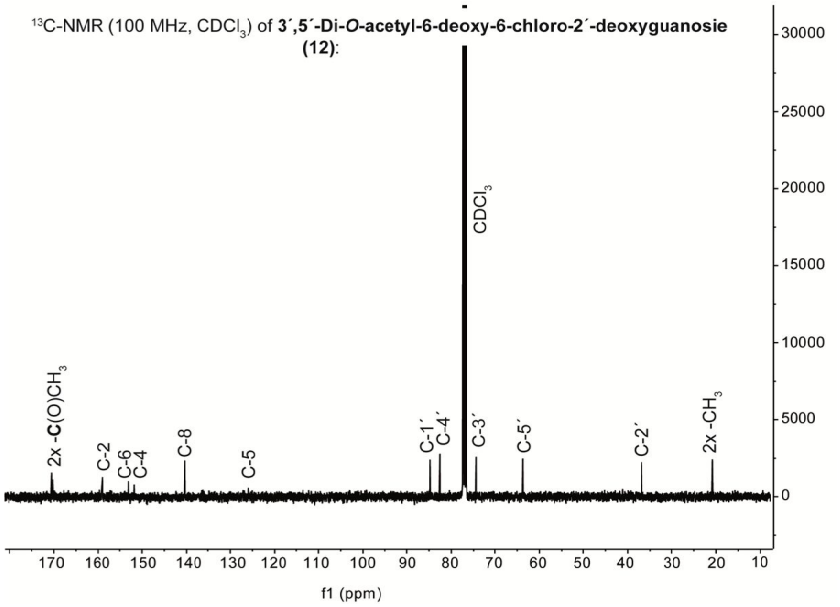
8. Appendix

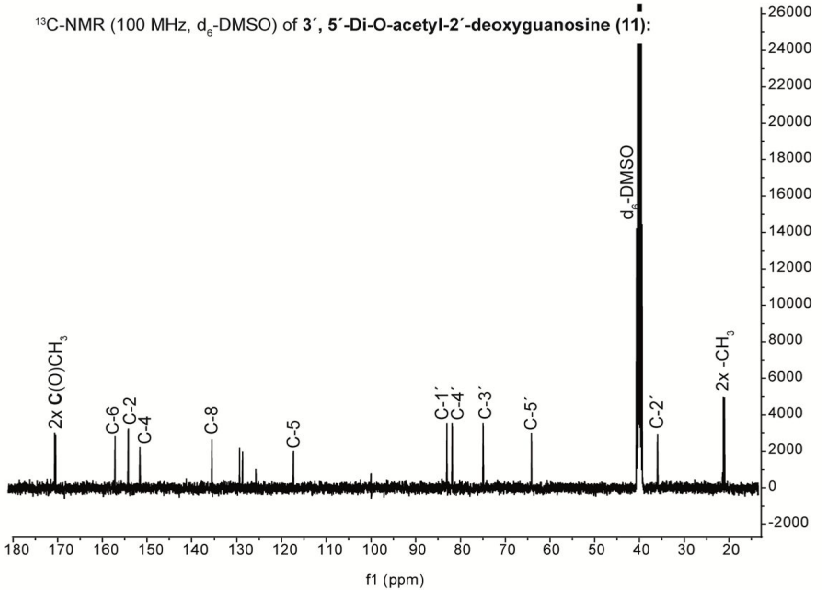
8.1 NMR



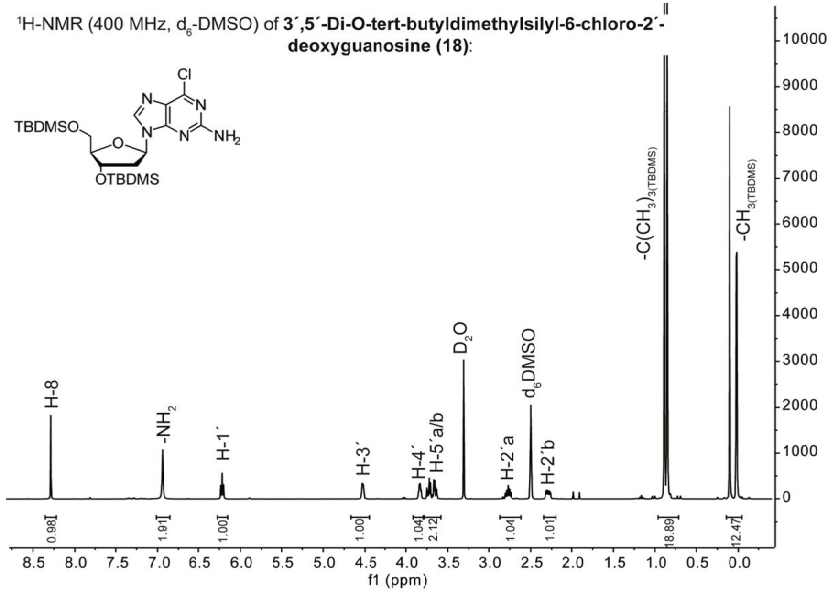

^{1}H-NMR (400 MHz, $CDCl_3$) of **3´,5´-Di-*O*-acetyl-6-deoxy-6-chloro-2´-deoxyguanosie (12)**:
H-8
$CDCl_3$
H-1´
H-3´
$-NH_2$
H-4´
H-5´a/b
H-2´a
H-2´b
acetyl
acetyl
f1 (ppm)


^{1}H-NMR (400 MHz, d_6-DMSO) of **3´, 5´-Di-O-acetyl-2´-deoxyguanosine (11)**:
-NH-
H-8
$-NH_2$
H-1´
H-3´
H-5´a/b
H-4´
H_2O
H-2´a
H-2´b
d_6-DMSO
acetyl
acetyl
f1 (ppm)

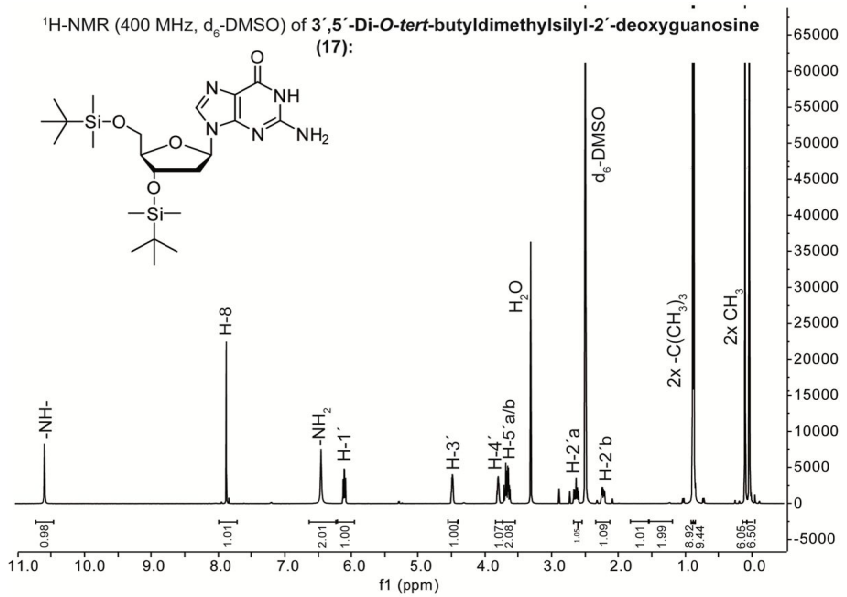

^{13}C-NMR (100 MHz, $CDCl_3$) of **3´,5´-Di-*O*-acetyl-6-deoxy-6-chloro-2´-deoxyguanosie (12)**:
2x -**C**(O)CH_3
C-2
C-6
C-4
C-8
C-5
C-1´
C-4´
$CDCl_3$
C-3´
C-5´
C-2´
2x -CH_3
f1 (ppm)


^{13}C-NMR (100 MHz, d_6-DMSO) of **3´, 5´-Di-O-acetyl-2´-deoxyguanosine (11)**:
2x **C**(O)CH_3
C-6
C-2
C-4
C-8
C-5
C-1´
C-4´
C-3´
C-5´
d_6-DMSO
C-2´
2x -CH_3
f1 (ppm)

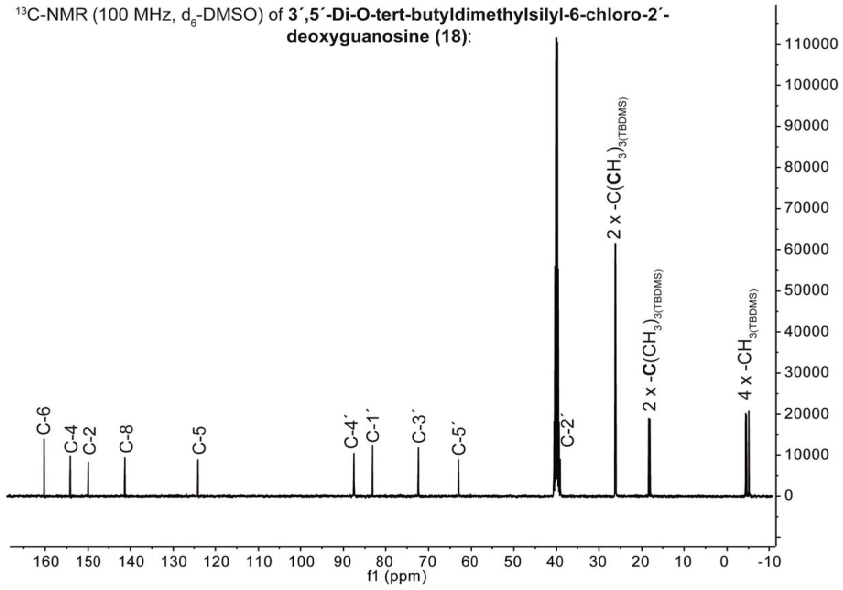
^{1}H-NMR (400 MHz, d_6-DMSO) of **3´,5´-Di-O-tert-butyldimethylsilyl-6-chloro-2´-deoxyguanosine (18)**:



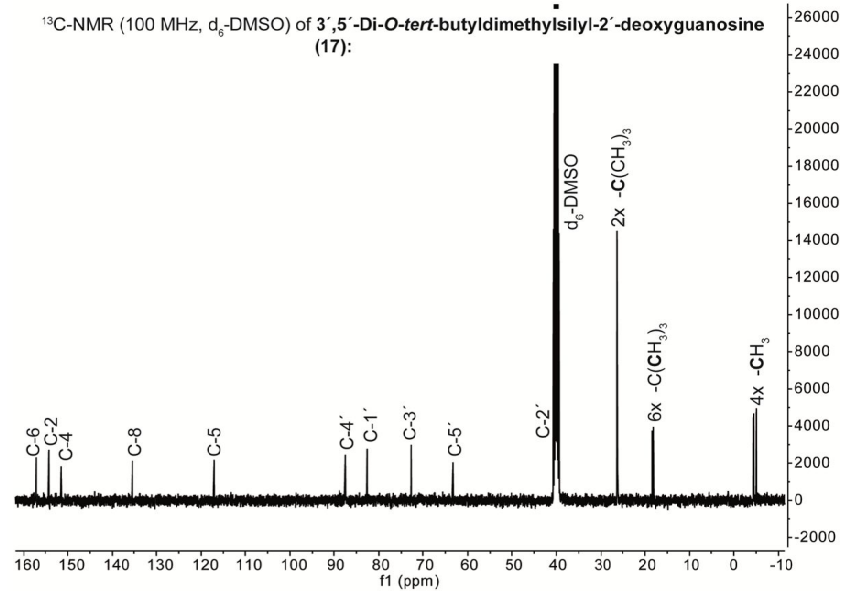
^{1}H-NMR (400 MHz, d_6-DMSO) of **3´,5´-Di-*O-tert*-butyldimethylsilyl-2´-deoxyguanosine (17)**:

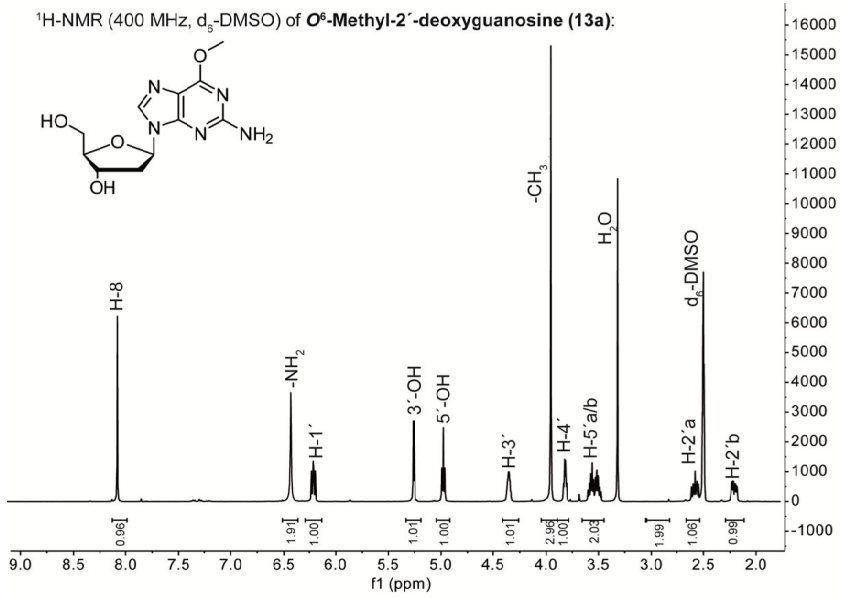


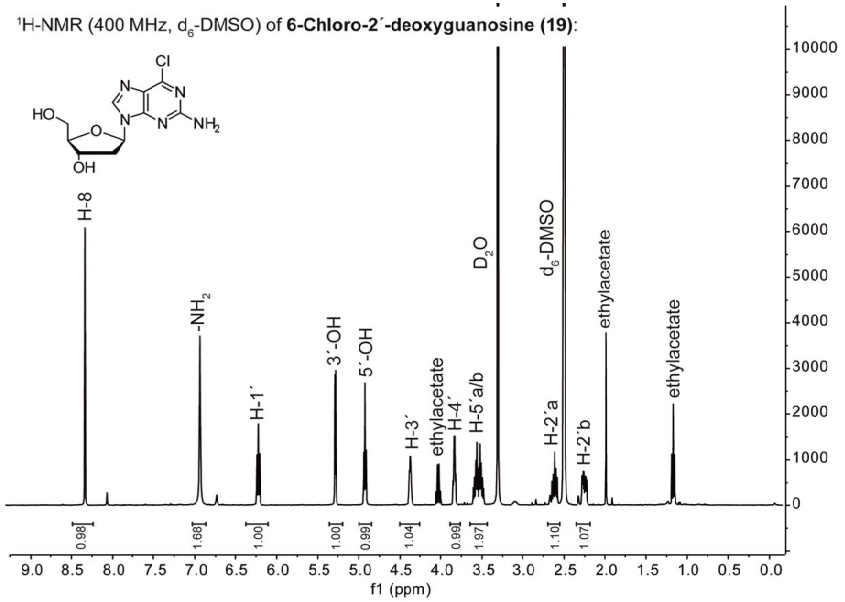
^{13}C-NMR (100 MHz, d_6-DMSO) of **3´,5´-Di-O-tert-butyldimethylsilyl-6-chloro-2´-deoxyguanosine (18)**:

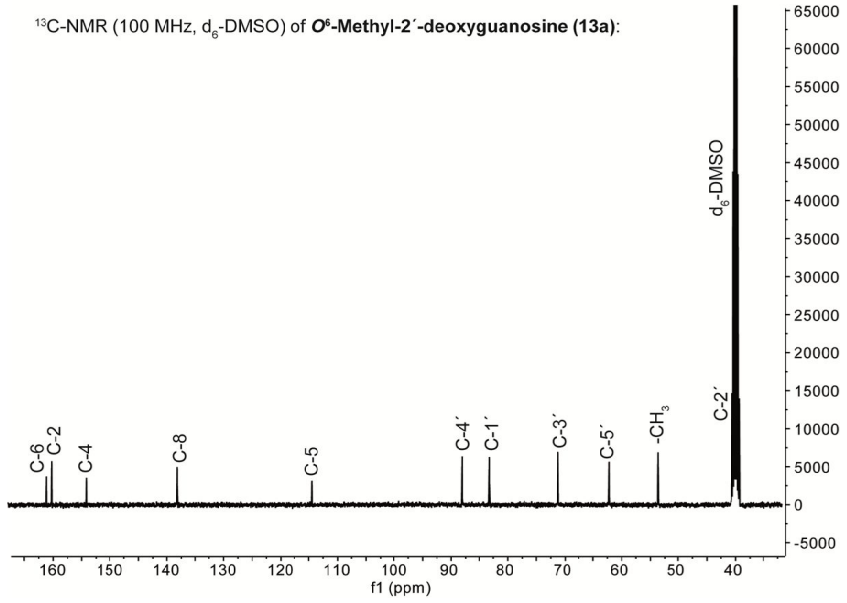


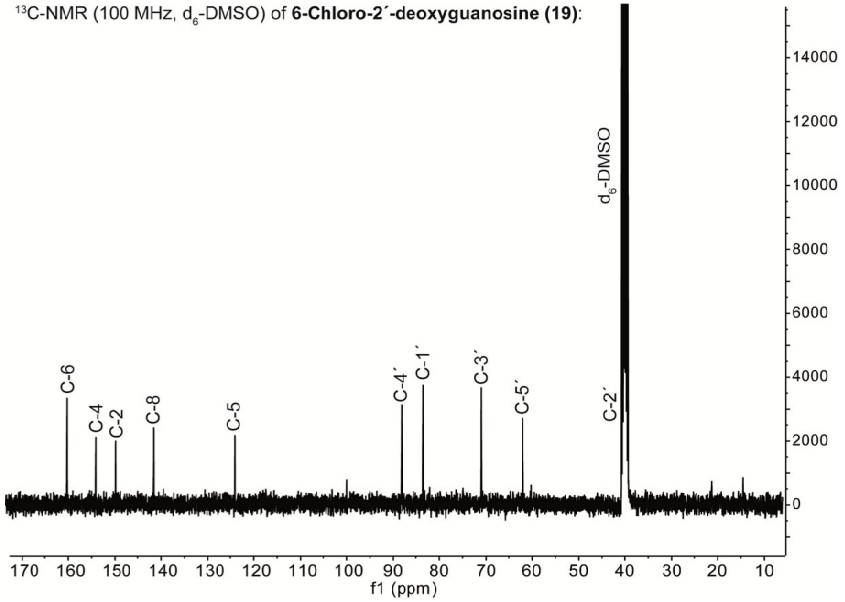
^{13}C-NMR (100 MHz, d_6-DMSO) of **3´,5´-Di-*O-tert*-butyldimethylsilyl-2´-deoxyguanosine (17)**:

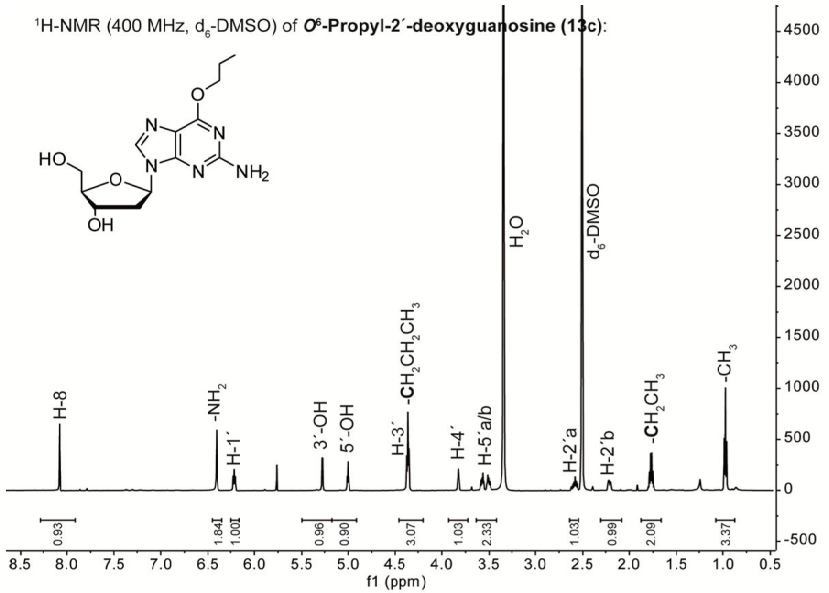


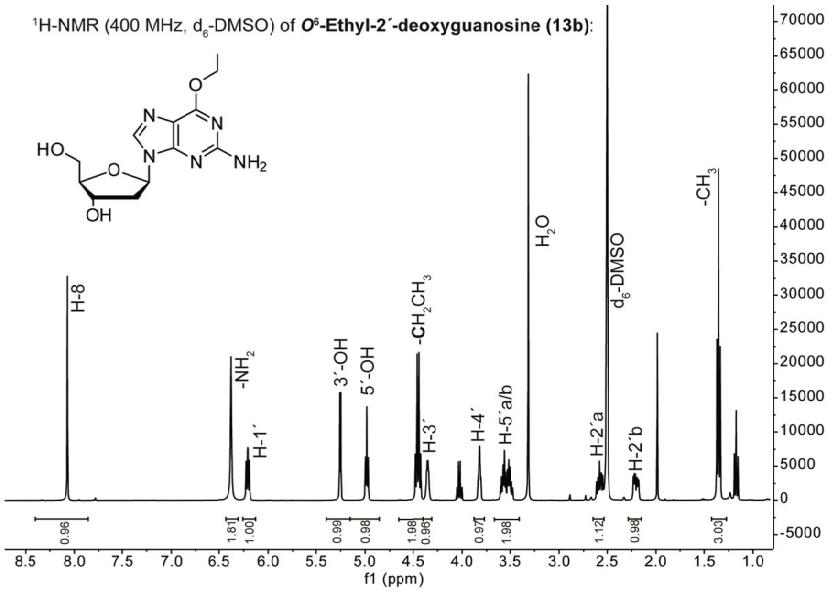

^{1}H-NMR (400 MHz, d$_6$-DMSO) of **O^6-Methyl-2´-deoxyguanosine (13a)**:
H-8
-NH$_2$
H-1´
3´-OH
5´-OH
H-3´
-CH$_3$
H-4´
H-5´a/b
H$_2$O
d$_6$-DMSO
H-2´a
H-2´b
f1 (ppm)

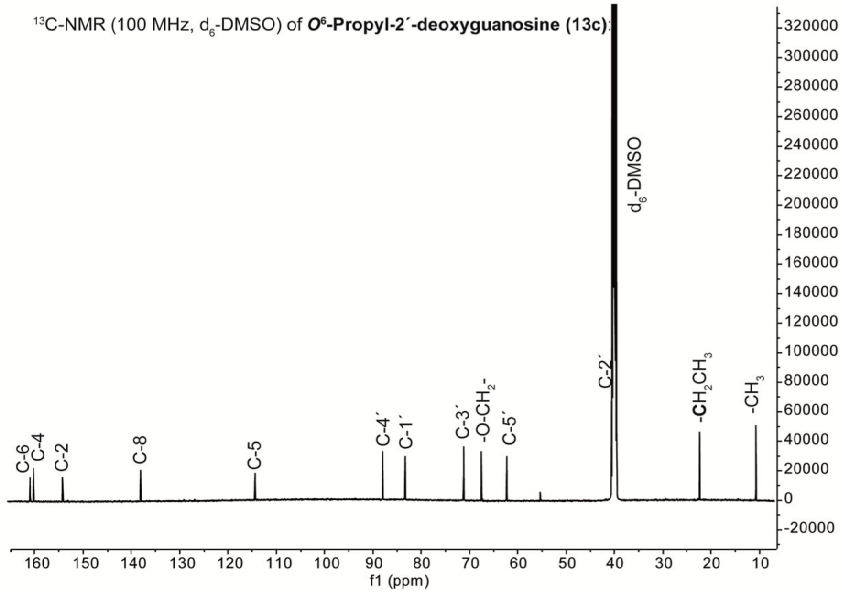

^{1}H-NMR (400 MHz, d$_6$-DMSO) of **6-Chloro-2´-deoxyguanosine (19)**:
H-8
-NH$_2$
H-1´
3´-OH
5´-OH
H-3´
ethylacetate
H-4´
H-5´a/b
D$_2$O
d$_6$-DMSO
H-2´a
H-2´b
ethylacetate
ethylacetate
f1 (ppm)

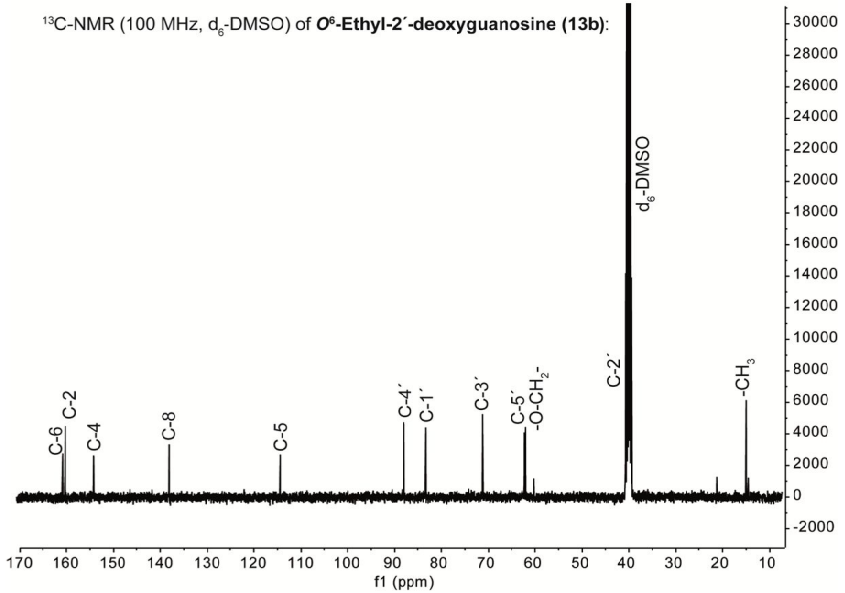

^{13}C-NMR (100 MHz, d$_6$-DMSO) of **O^6-Methyl-2´-deoxyguanosine (13a)**:
C-6
C-2
C-4
C-8
C-5
C-4´
C-1´
C-3´
C-5´
-CH$_3$
d$_6$-DMSO
C-2´
f1 (ppm)

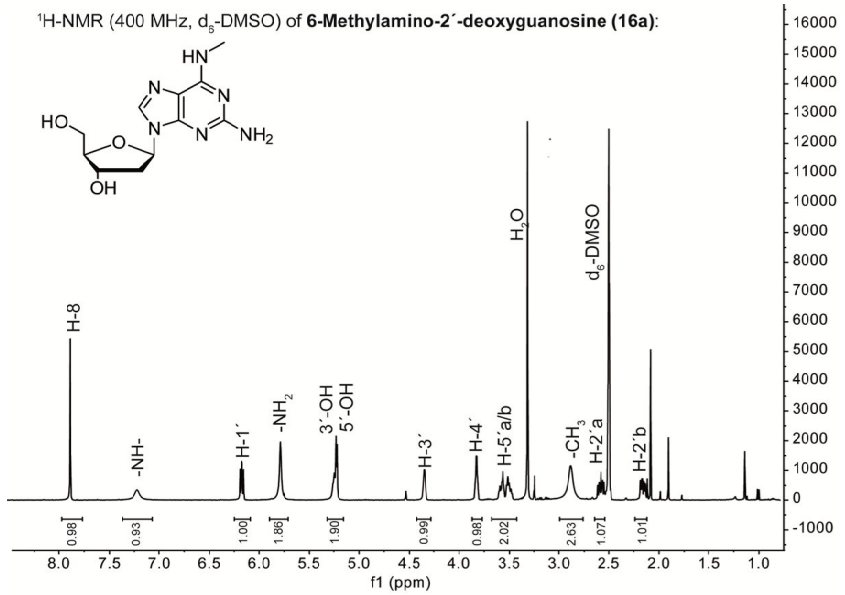

^{13}C-NMR (100 MHz, d$_6$-DMSO) of **6-Chloro-2´-deoxyguanosine (19)**:
C-6
C-4
C-2
C-8
C-5
C-4´
C-1´
C-3´
C-5´
d$_6$-DMSO
C-2´
f1 (ppm)

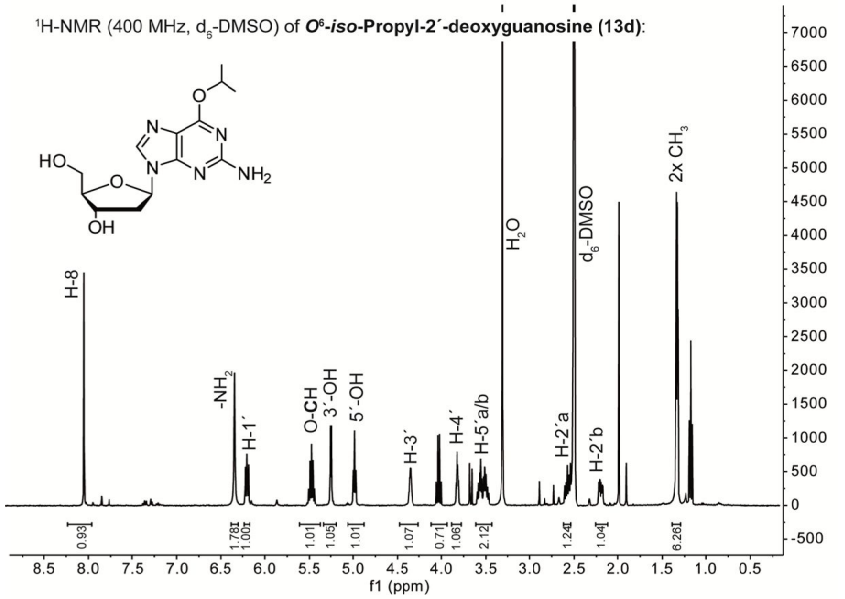

^{1}H-NMR (400 MHz, d$_6$-DMSO) of **O^6-Propyl-2´-deoxyguanosine (13c)**:
HO
O
N
N
N
N
NH$_2$
OH
H-8
-NH$_2$
H-1´
3´-OH
5´-OH
H-3´
-CH$_2$CH$_2$CH$_3$
H-4´
H-5´a/b
H$_2$O
d$_6$-DMSO
H-2´a
H-2´b
-CH$_2$CH$_3$
-CH$_3$
f1 (ppm)

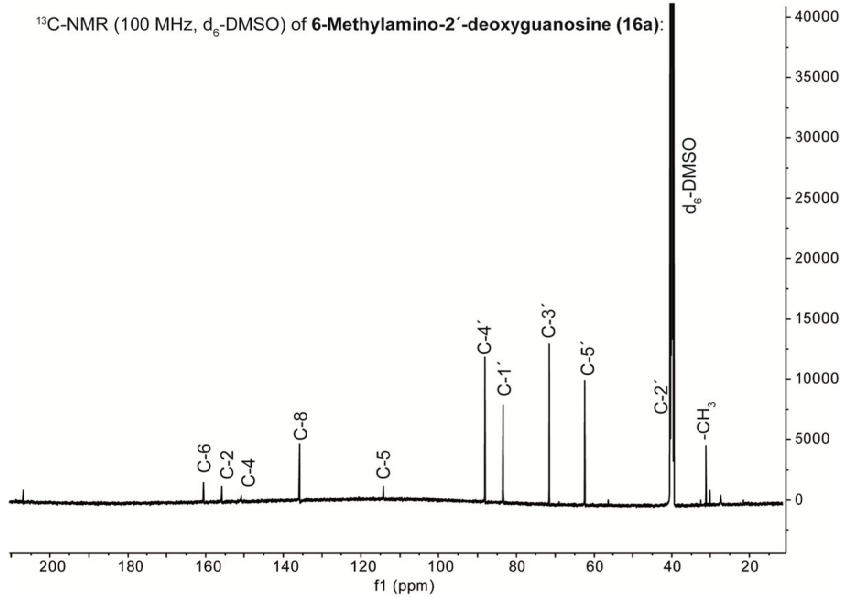

^{1}H-NMR (400 MHz, d$_6$-DMSO) of **O^6-Ethyl-2´-deoxyguanosine (13b)**:
HO
O
N
N
N
N
NH$_2$
OH
H-8
-NH$_2$
H-1´
3´-OH
5´-OH
-CH$_2$CH$_3$
H-3´
H-4´
H-5´a/b
H$_2$O
H-2´a
d$_6$-DMSO
H-2´b
-CH$_3$
f1 (ppm)

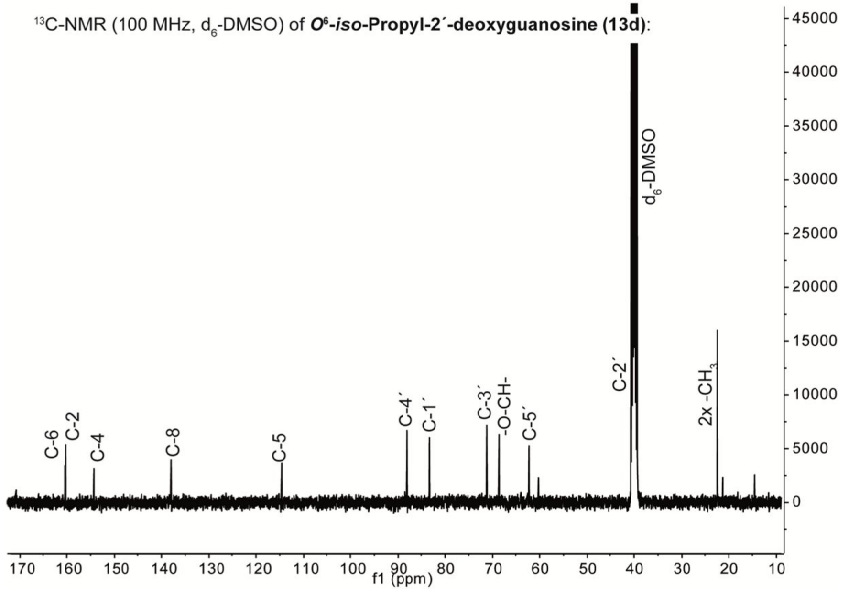

^{13}C-NMR (100 MHz, d$_6$-DMSO) of **O^6-Propyl-2´-deoxyguanosine (13c)**:
C-6
C-4
C-2
C-8
C-5
C-4´
C-1´
C-3´
-O-CH$_2$-
C-5´
C-2´
d$_6$-DMSO
-CH$_2$CH$_3$
-CH$_3$
f1 (ppm)

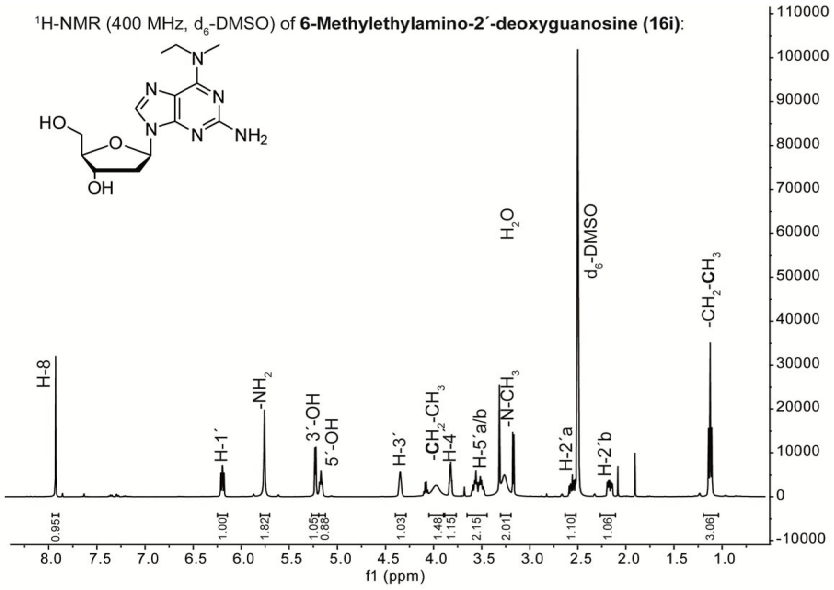

^{13}C-NMR (100 MHz, d$_6$-DMSO) of **O^6-Ethyl-2´-deoxyguanosine (13b)**:
C-6
C-2
C-4
C-8
C-5
C-4´
C-1´
C-3´
C-5´
-O-CH$_2$-
C-2´
d$_6$-DMSO
-CH$_3$
f1 (ppm)

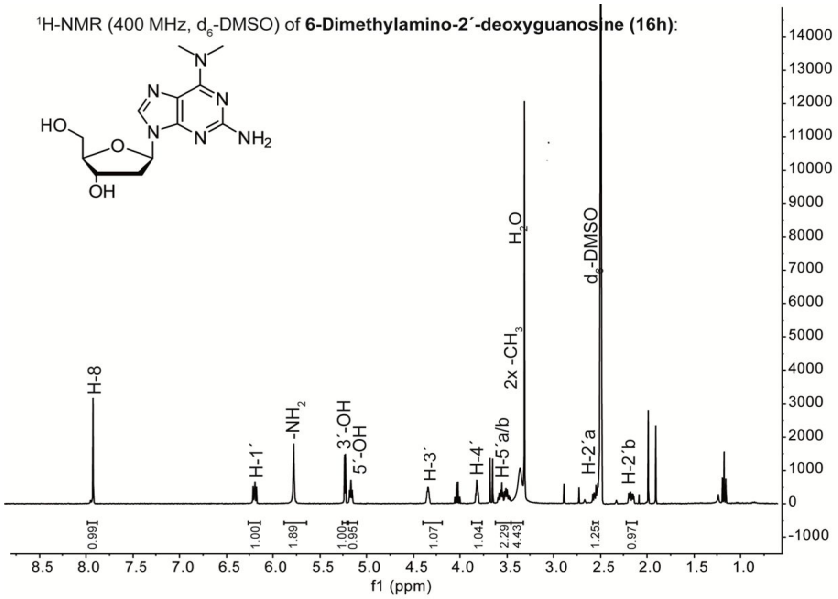

^{1}H-NMR (400 MHz, d_5-DMSO) of **6-Methylamino-2´-deoxyguanosine (16a)**:
HN
HO
OH
NH$_2$
H-8
-NH-
H-1´
-NH$_2$
3´-OH
5´-OH
H-3´
H-4´
H-5´a/b
H$_2$O
-CH$_3$
H-2´a
d$_6$-DMSO
H-2´b
f1 (ppm)

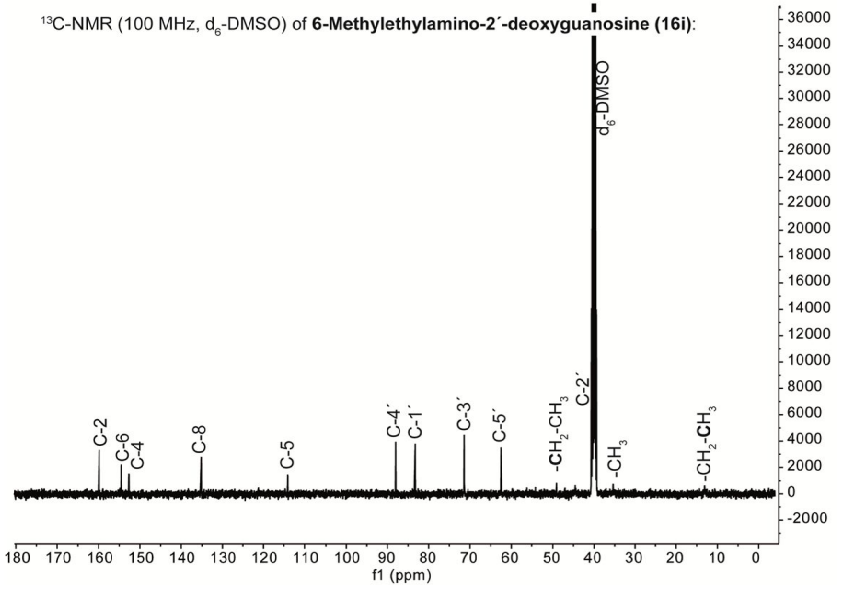

^{1}H-NMR (400 MHz, d_5-DMSO) of ***O*^{6}-*iso*-Propyl-2´-deoxyguanosine (13d)**:
HO
OH
NH$_2$
H-8
-NH$_2$
H-1´
O-CH
3´-OH
5´-OH
H-3´
H-4´
H-5´a/b
H$_2$O
H-2´a
d$_6$-DMSO
H-2´b
2x CH$_3$
f1 (ppm)

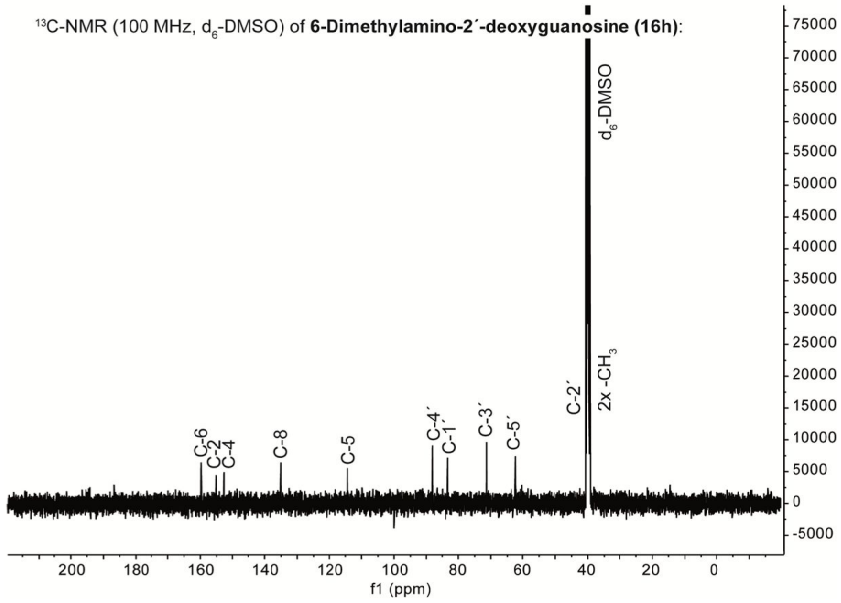

^{13}C-NMR (100 MHz, d_6-DMSO) of **6-Methylamino-2´-deoxyguanosine (16a)**:
C-6
C-2
C-4
C-8
C-5
C-4´
C-1´
C-3´
C-5´
C-2´
d$_6$-DMSO
-CH$_3$
f1 (ppm)


^{13}C-NMR (100 MHz, d_6-DMSO) of ***O*^{6}-*iso*-Propyl-2´-deoxyguanosine (13d)**:
C-6
C-2
C-4
C-8
C-5
C-4´
C-1´
C-3´
-O-CH-
C-5´
C-2´
d$_6$-DMSO
2x -CH$_3$
f1 (ppm)

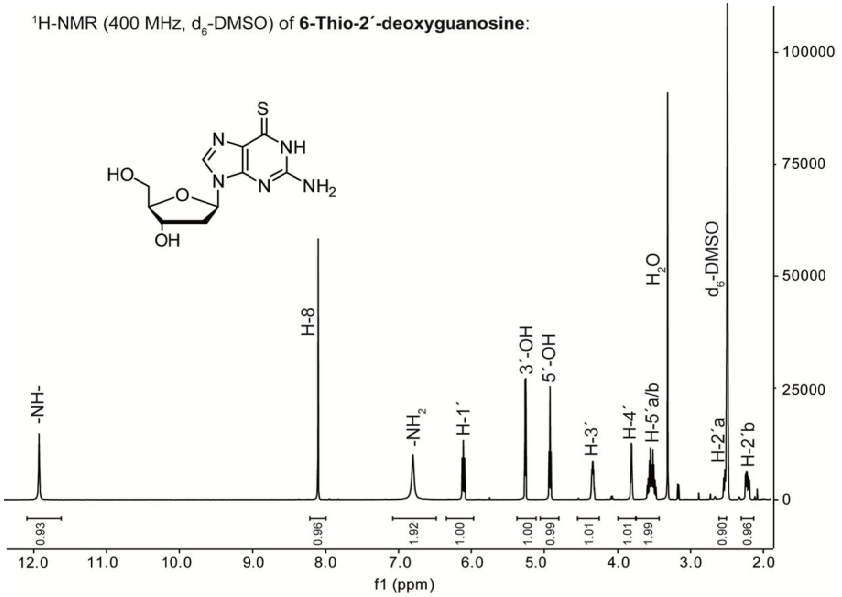

^{1}H-NMR (400 MHz, d_6-DMSO) of **6-Methylethylamino-2´-deoxyguanosine (16i)**:
HO
OH
N
NH$_2$
H-8
-NH$_2$
H-1´
3´-OH
5´-OH
H-3´
-CH$_2$-CH$_3$
H-4´
H-5´a/b
-N-CH$_3$
H$_2$O
H-2´a
d$_6$-DMSO
H-2´b
-CH$_2$-CH$_3$
f1 (ppm)


^{1}H-NMR (400 MHz, d_6-DMSO) of **6-Dimethylamino-2´-deoxyguanosine (16h)**:
HO
OH
N
NH$_2$
H-8
H-1´
-NH$_2$
3´-OH
5´-OH
H-3´
H-4´
H-5´a/b
2x -CH$_3$
H$_2$O
d$_6$-DMSO
H-2´a
H-2´b
f1 (ppm)

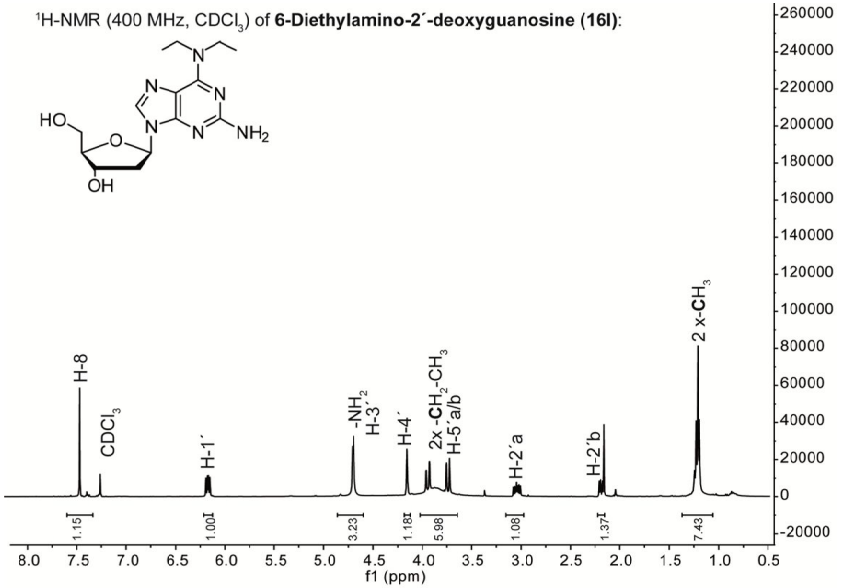

^{13}C-NMR (100 MHz, d_6-DMSO) of **6-Methylethylamino-2´-deoxyguanosine (16i)**:
d$_6$-DMSO
C-2
C-6
C-4
C-8
C-5
C-4´
C-1´
C-3´
C-5´
-CH$_2$-CH$_3$
C-2´
-CH$_3$
-CH$_2$-CH$_3$
f1 (ppm)


^{13}C-NMR (100 MHz, d_6-DMSO) of **6-Dimethylamino-2´-deoxyguanosine (16h)**:
d$_6$-DMSO
C-6
C-2
C-4
C-8
C-5
C-4´
C-1´
C-3´
C-5´
C-2´
2x -CH$_3$
f1 (ppm)

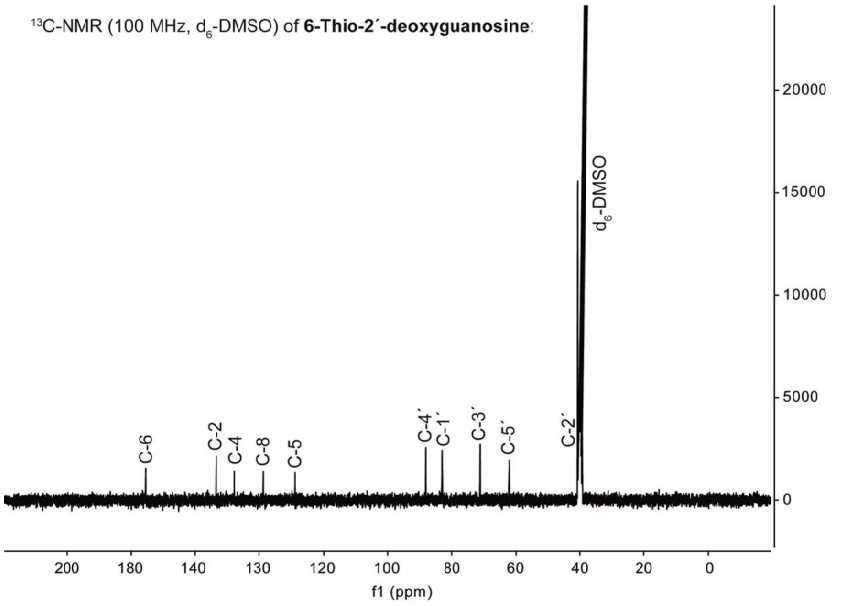
^{1}H-NMR (400 MHz, d_6-DMSO) of **6-Thio-2´-deoxyguanosine**:



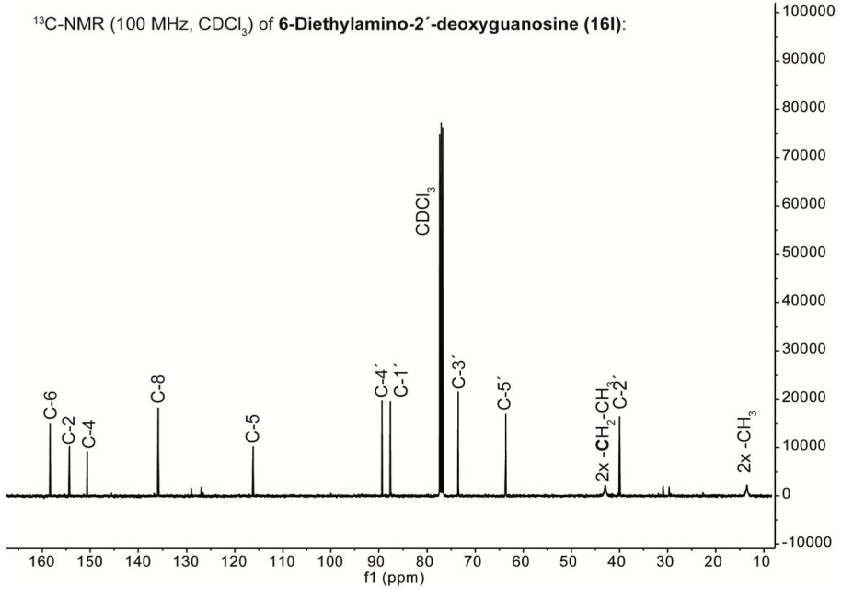
^{1}H-NMR (400 MHz, $CDCl_3$) of **6-Diethylamino-2´-deoxyguanosine (16l)**:

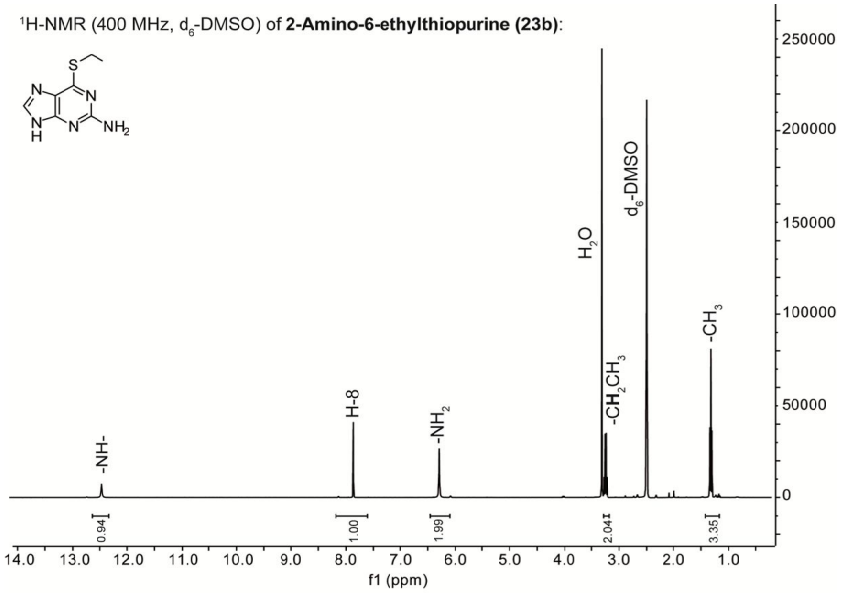


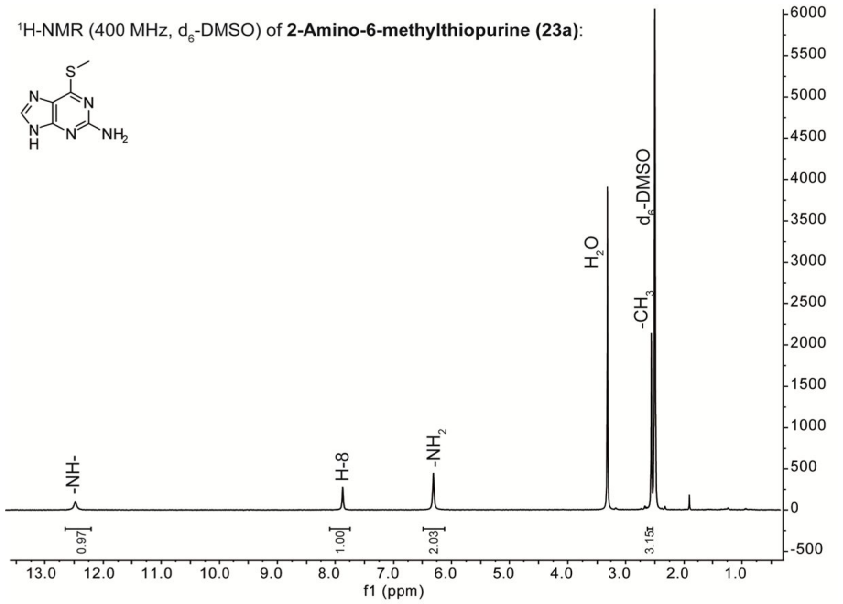
^{13}C-NMR (100 MHz, d_6-DMSO) of **6-Thio-2´-deoxyguanosine**:

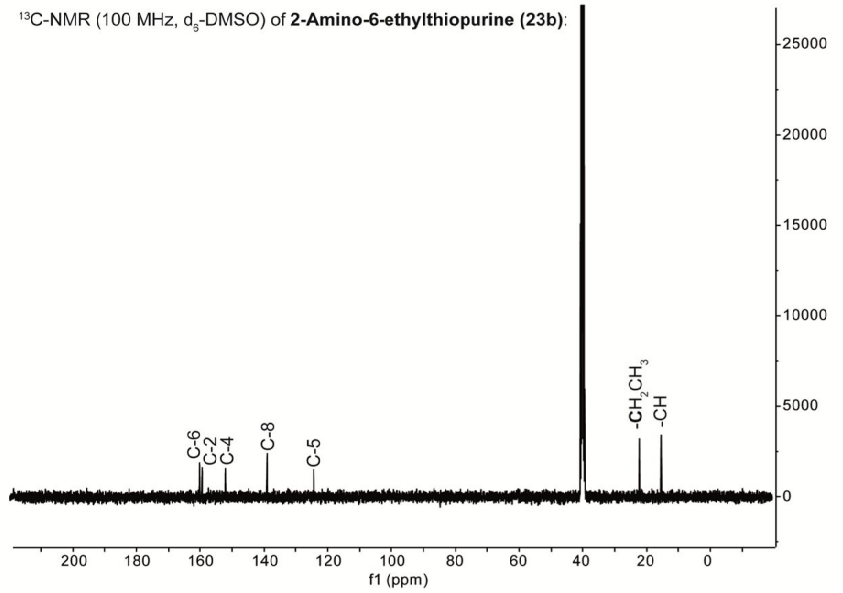


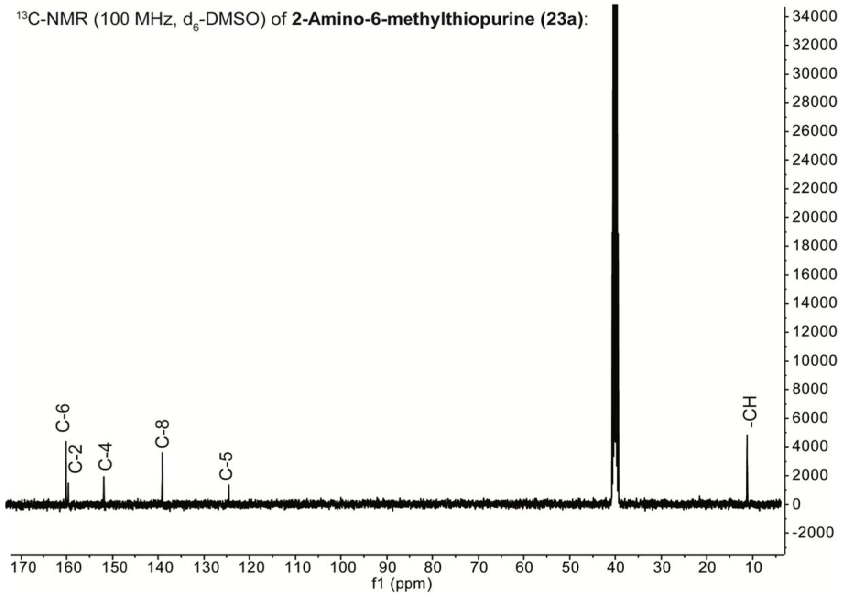
^{13}C-NMR (100 MHz, $CDCl_3$) of **6-Diethylamino-2´-deoxyguanosine (16l)**:



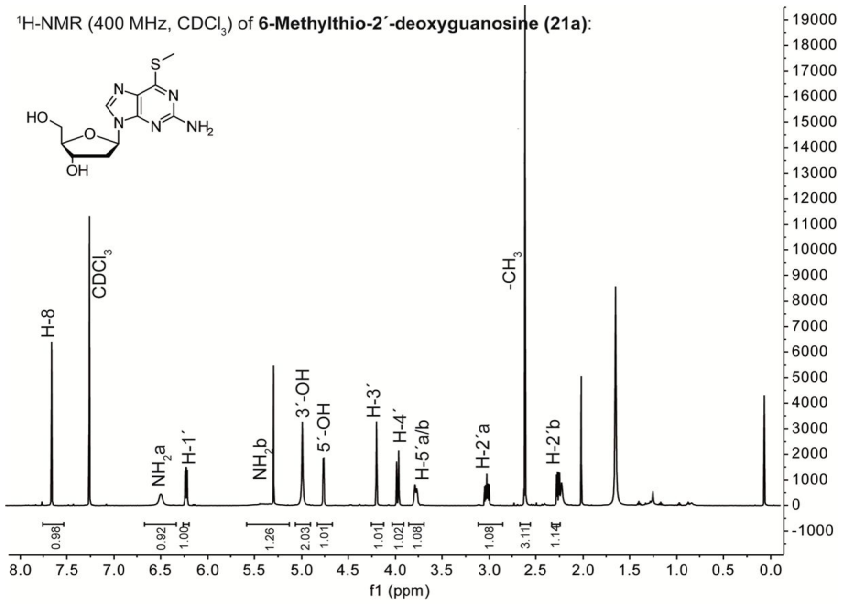

^{1}H-NMR (400 MHz, d_6-DMSO) of **2-Amino-6-ethylthiopurine (23b)**:
-NH-
H-8
-NH$_2$
H$_2$O
-CH$_2$CH$_3$
d_6-DMSO
-CH$_3$
0.94
1.00
1.99
2.04
3.35
f1 (ppm)


^{1}H-NMR (400 MHz, d_6-DMSO) of **2-Amino-6-methylthiopurine (23a)**:
-NH-
H-8
-NH$_2$
H$_2$O
-CH$_3$
d_6-DMSO
0.97
1.00
2.03
3.15
f1 (ppm)

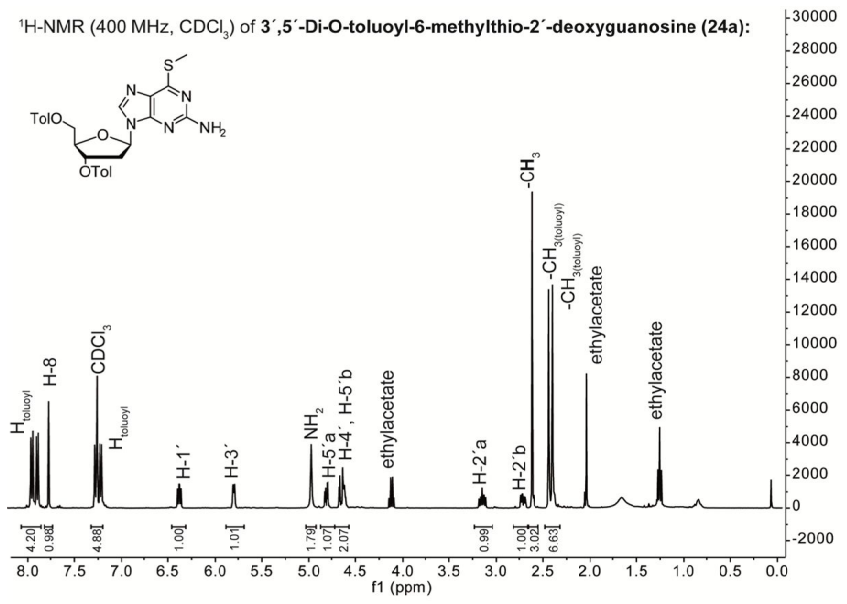

^{13}C-NMR (100 MHz, d_6-DMSO) of **2-Amino-6-ethylthiopurine (23b)**:
C-6
C-2
C-4
C-8
C-5
-CH$_2$CH$_3$
-CH
f1 (ppm)


^{13}C-NMR (100 MHz, d_6-DMSO) of **2-Amino-6-methylthiopurine (23a)**:
C-6
C-2
C-4
C-8
C-5
-CH
f1 (ppm)

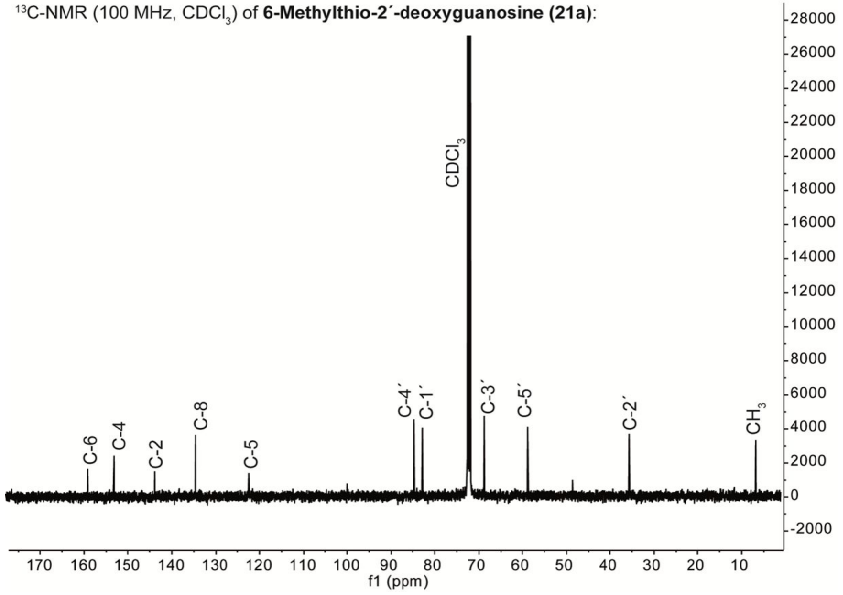
^{1}H-NMR (400 MHz, $CDCl_3$) of **6-Methylthio-2´-deoxyguanosine (21a)**:



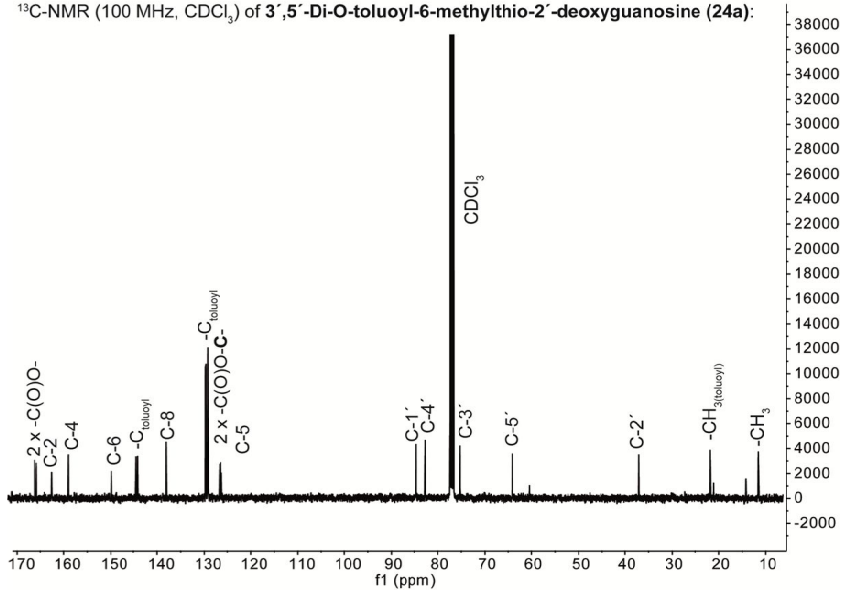
^{1}H-NMR (400 MHz, $CDCl_3$) of **3´,5´-Di-O-toluoyl-6-methylthio-2´-deoxyguanosine (24a):**

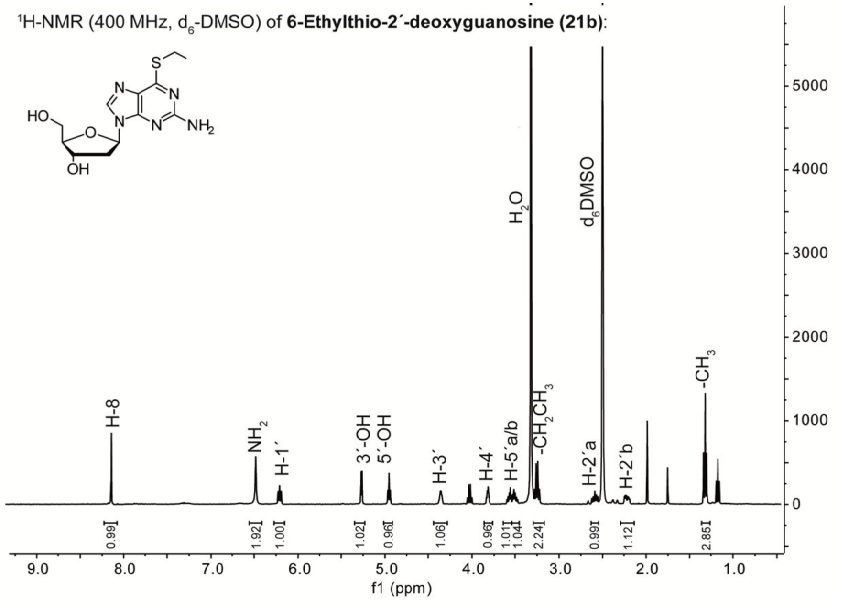


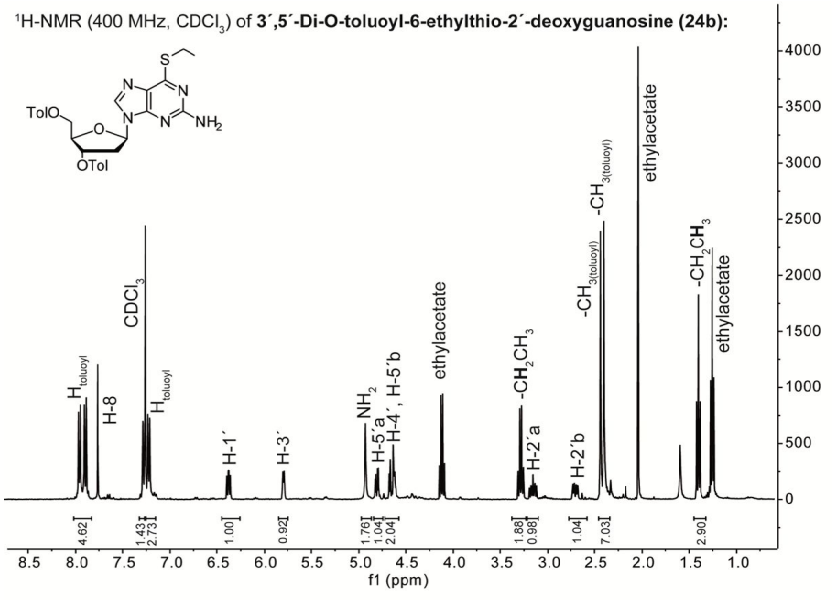
^{13}C-NMR (100 MHz, $CDCl_3$) of **6-Methylthio-2´-deoxyguanosine (21a)**:

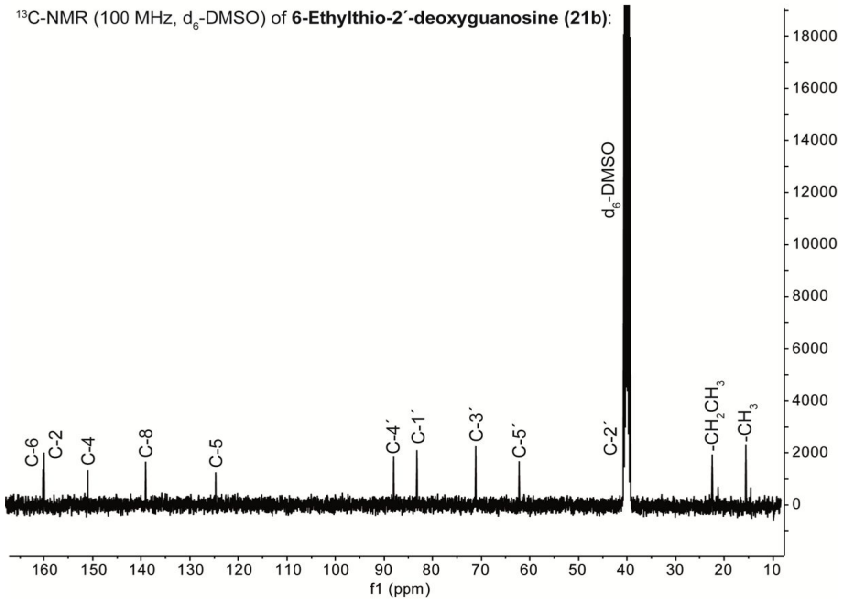


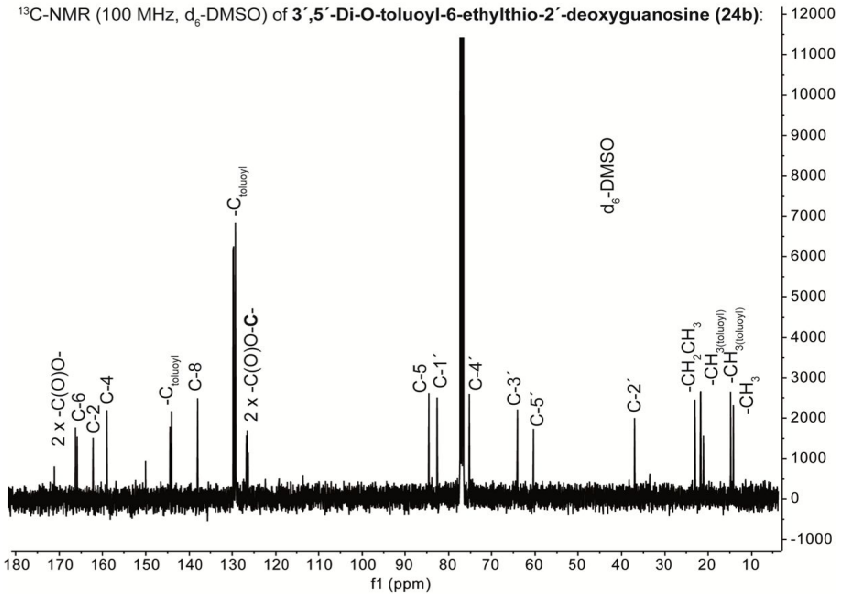
^{13}C-NMR (100 MHz, $CDCl_3$) of **3´,5´-Di-O-toluoyl-6-methylthio-2´-deoxyguanosine (24a)**:

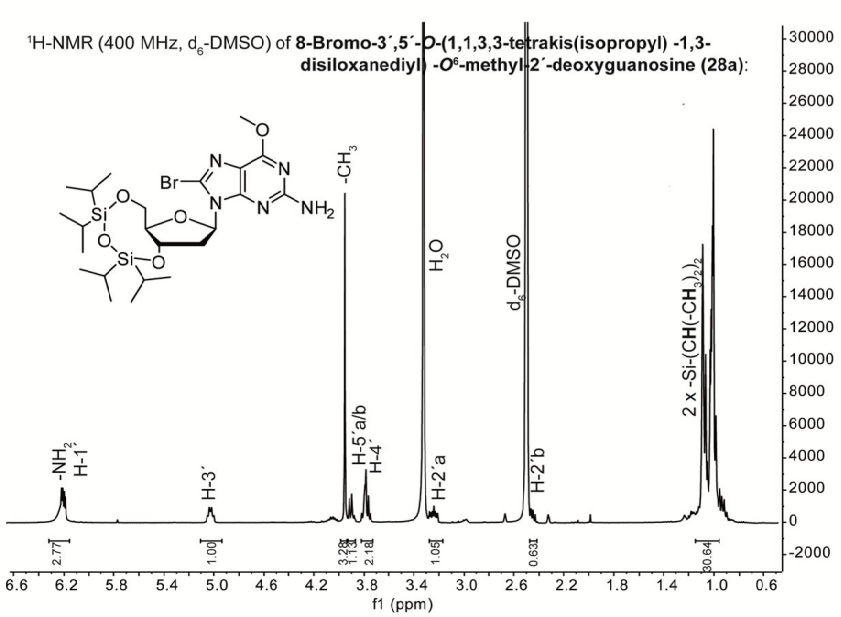


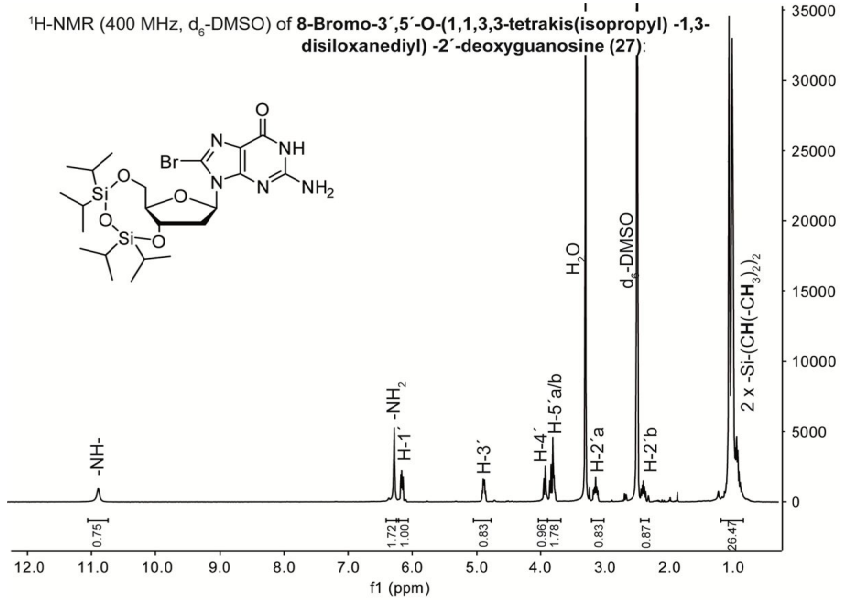

^{1}H-NMR (400 MHz, d_6-DMSO) of **6-Ethylthio-2´-deoxyguanosine (21b)**:
HO
OH
O
N
N
N
N
S
NH$_2$
H-8
NH$_2$
H-1´
3´-OH
5´-OH
H-3´
H-4´
H-5´a/b
-CH$_2$CH$_3$
H$_2$O
d$_6$DMSO
H-2´a
H-2´b
-CH$_3$
0.99
1.92
1.00
1.02
0.96
1.06
0.96
1.01
1.04
2.24
0.99
1.12
2.85
f1 (ppm)

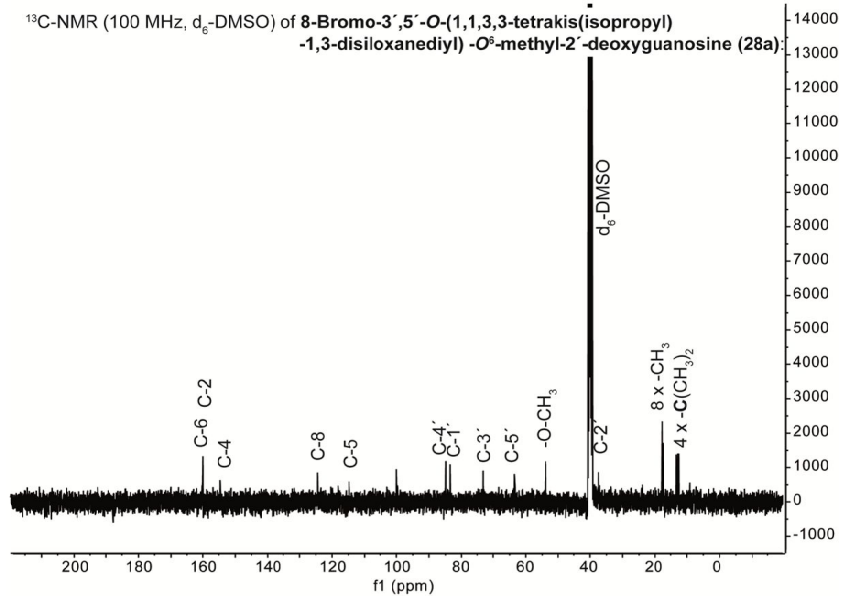

^{1}H-NMR (400 MHz, $CDCl_3$) of **3´,5´-Di-O-toluoyl-6-ethylthio-2´-deoxyguanosine (24b):**
TolO
OTol
O
N
N
N
N
S
NH$_2$
H$_{toluoyl}$
H-8
CDCl$_3$
H$_{toluoyl}$
H-1´
H-3´
NH$_2$
H-5´a
H-4´, H-5´b
ethylacetate
-CH$_2$CH$_3$
H-2´a
H-2´b
-CH$_{3(toluoyl)}$
ethylacetate
-CH$_2$CH$_3$
ethylacetate
4.62
1.43
2.73
1.00
0.92
1.76
1.04
2.04
1.88
0.98
1.04
7.03
2.90
f1 (ppm)

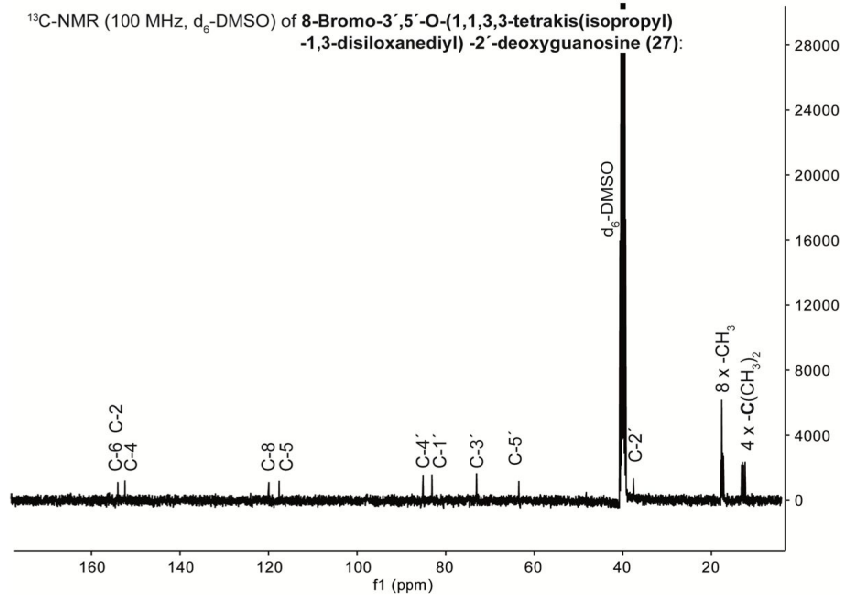

^{13}C-NMR (100 MHz, d_6-DMSO) of **6-Ethylthio-2´-deoxyguanosine (21b)**:
C-6
C-2
C-4
C-8
C-5
C-4´
C-1´
C-3´
C-5´
d$_6$-DMSO
C-2´
-CH$_2$CH$_3$
-CH$_3$
f1 (ppm)

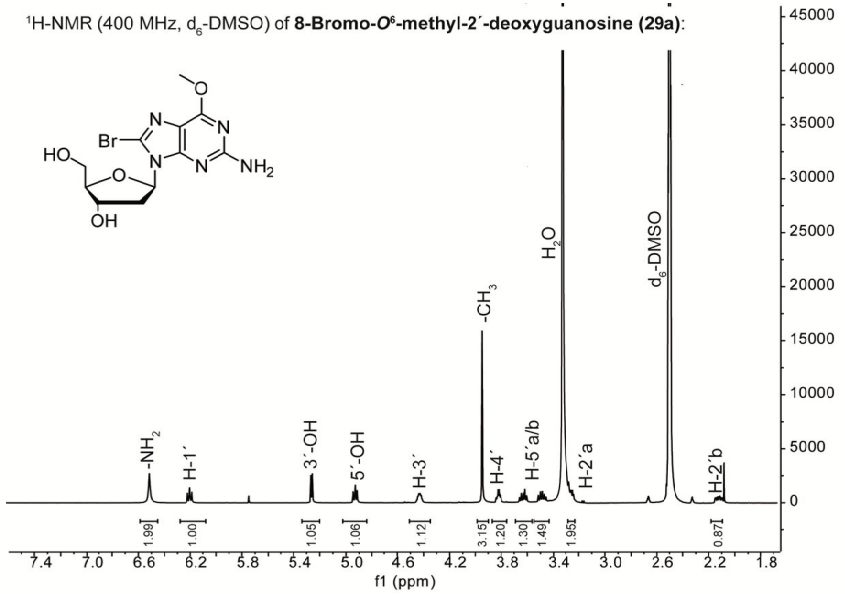

^{13}C-NMR (100 MHz, d_6-DMSO) of **3´,5´-Di-O-toluoyl-6-ethylthio-2´-deoxyguanosine (24b)**:
2 x -C(O)O-
C-6
C-2
C-4
-C$_{toluoyl}$
C-8
-C$_{toluoyl}$
2 x -C(O)O-**C**-
C-5
C-1´
C-4´
C-3´
C-5´
d$_6$-DMSO
C-2´
-CH$_2$CH$_3$
-CH$_{3(toluoyl)}$
-CH$_{3(toluoyl)}$
-CH$_3$
f1 (ppm)

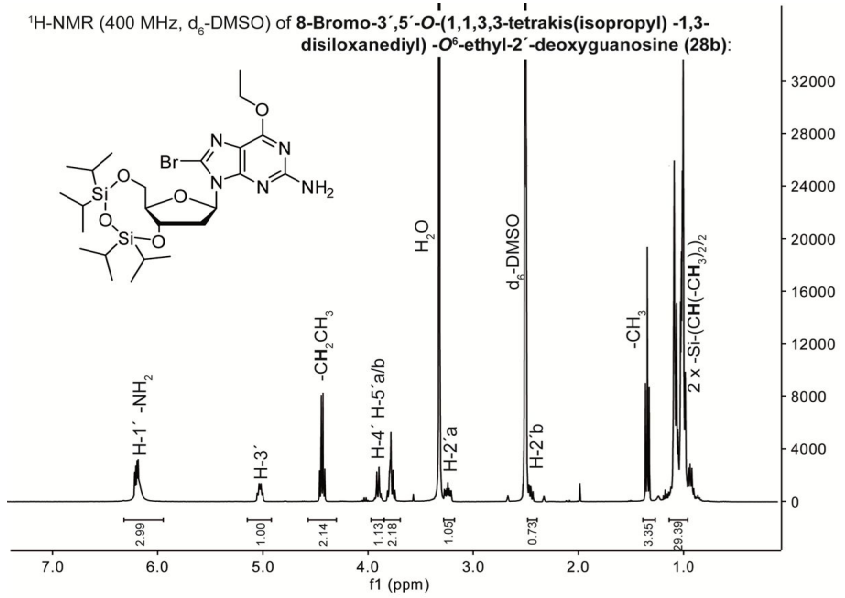

^{1}H-NMR (400 MHz, d$_6$-DMSO) of **8-Bromo-3´,5´-O-(1,1,3,3-tetrakis(isopropyl) -1,3-disiloxanediyl) -O^6-methyl-2´-deoxyguanosine (28a)**:
-NH$_2$
H-1´
H-3´
H-5´a/b
H-4´
-CH$_3$
H$_2$O
H-2´a
d$_6$-DMSO
H-2´b
2 x -Si-(CH(-CH$_3$)$_2$)$_2$
f1 (ppm)

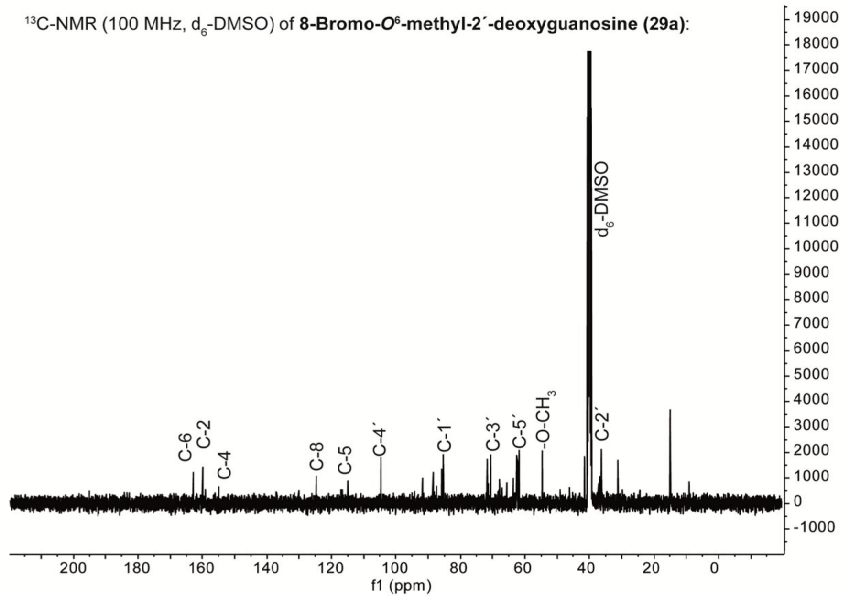

^{1}H-NMR (400 MHz, d$_6$-DMSO) of **8-Bromo-3´,5´-O-(1,1,3,3-tetrakis(isopropyl) -1,3-disiloxanediyl) -2´-deoxyguanosine (27)**:
-NH-
-NH$_2$
H-1´
H-3´
H-4´
H-5´a/b
H-2´a
H$_2$O
d$_6$-DMSO
H-2´b
2 x -Si-(CH(-CH$_3$)$_2$)$_2$
f1 (ppm)

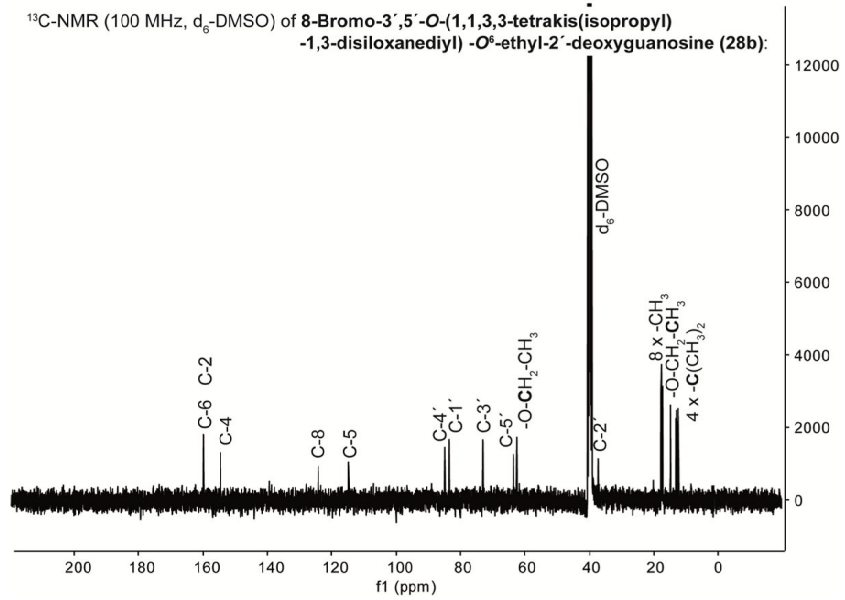

^{13}C-NMR (100 MHz, d$_6$-DMSO) of **8-Bromo-3´,5´-O-(1,1,3,3-tetrakis(isopropyl) -1,3-disiloxanediyl) -O^6-methyl-2´-deoxyguanosine (28a)**:
C-6
C-2
C-4
C-8
C-5
C-4´
C-1´
C-3´
C-5´
-O-CH$_3$
d$_6$-DMSO
C-2´
8 x -CH$_3$
4 x -C(CH$_3$)$_2$
f1 (ppm)

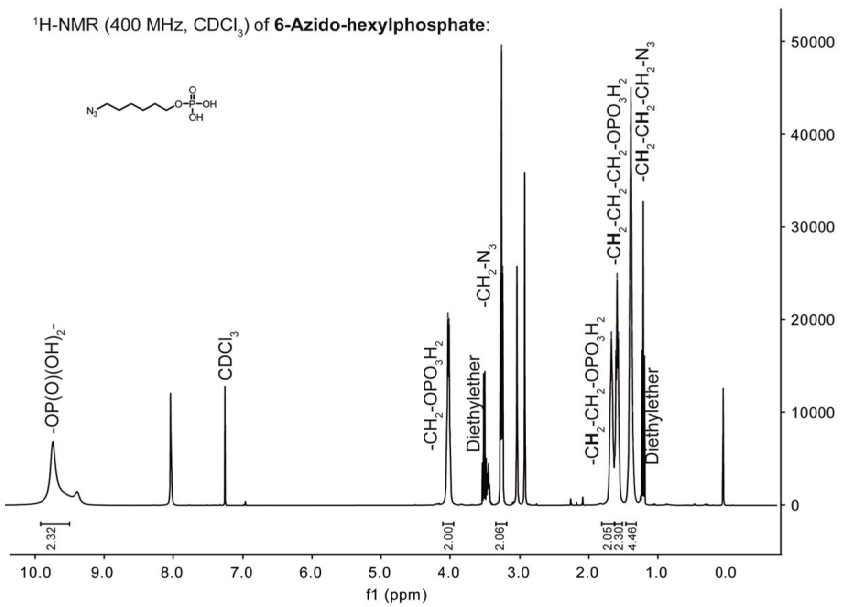

^{13}C-NMR (100 MHz, d$_6$-DMSO) of **8-Bromo-3´,5´-O-(1,1,3,3-tetrakis(isopropyl) -1,3-disiloxanediyl) -2´-deoxyguanosine (27)**:
C-6
C-2
C-4
C-8
C-5
C-4´
C-1´
C-3´
C-5´
d$_6$-DMSO
C-2´
8 x -CH$_3$
4 x -C(CH$_3$)$_2$
f1 (ppm)

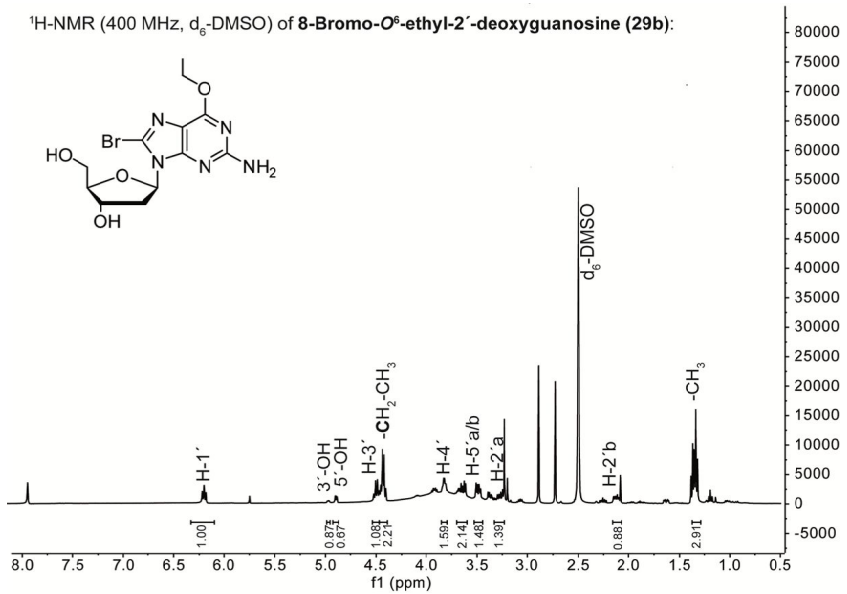

^{1}H-NMR (400 MHz, d_6-DMSO) of **8-Bromo-*O*^{6}-methyl-2´-deoxyguanosine (29a)**:
-NH$_2$
H-1´
3´-OH
5´-OH
H-3´
-CH$_3$
H-4´
H-5´a/b
H$_2$O
H-2´a
d$_6$-DMSO
H-2´b
f1 (ppm)

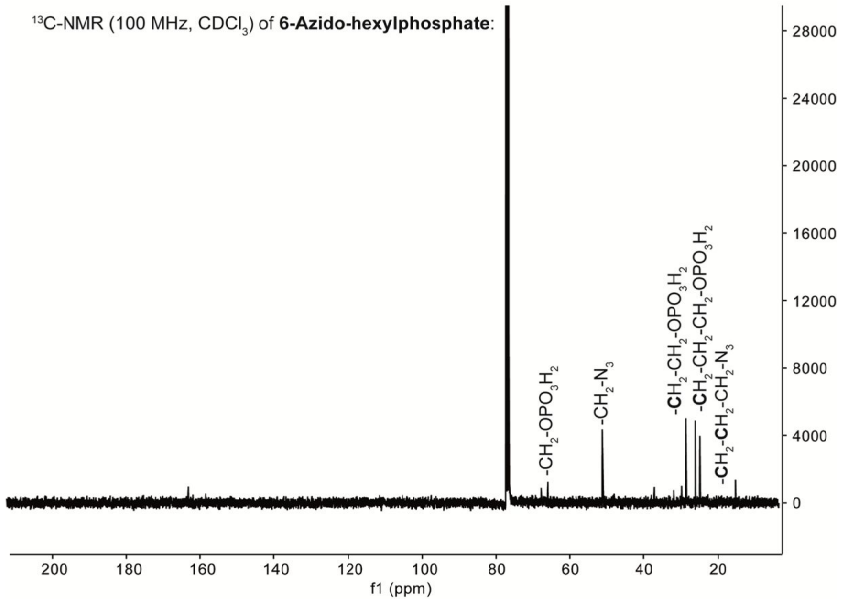

^{1}H-NMR (400 MHz, d_6-DMSO) of **8-Bromo-3´,5´-*O*-(1,1,3,3-tetrakis(isopropyl) -1,3-disiloxanediyl) -*O*^{6}-ethyl-2´-deoxyguanosine (28b)**:
H-1´ -NH$_2$
H-3´
-CH$_2$CH$_3$
H-4´ H-5´a/b
H$_2$O
H-2´a
d$_6$-DMSO
H-2´b
-CH$_3$
2 x -Si-(CH(-CH$_3$)$_2$)$_2$
f1 (ppm)


^{13}C-NMR (100 MHz, d_6-DMSO) of **8-Bromo-*O*^{6}-methyl-2´-deoxyguanosine (29a)**:
C-6
C-2
C-4
C-8
C-5
C-4´
C-1´
C-3´
C-5´
O-CH$_3$
d$_6$-DMSO
C-2´
f1 (ppm)

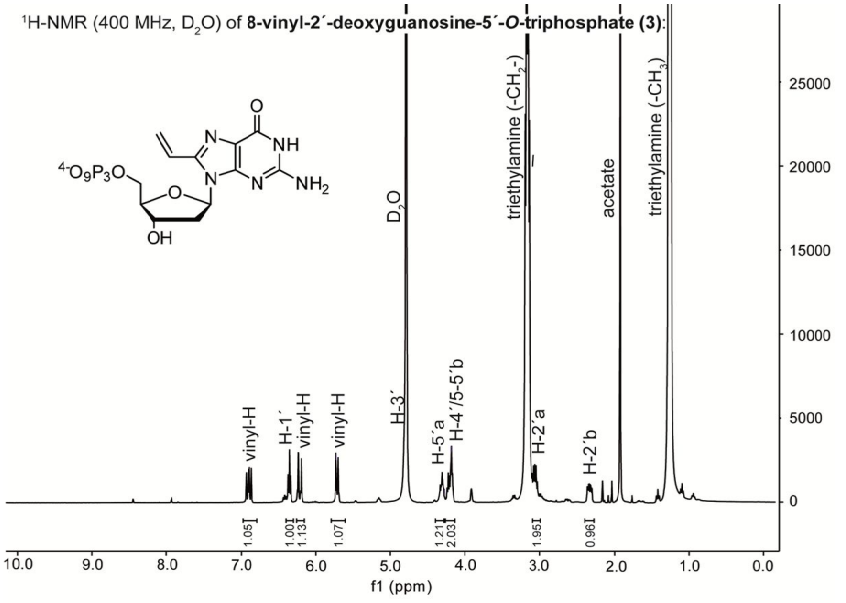

^{13}C-NMR (100 MHz, d_6-DMSO) of **8-Bromo-3´,5´-*O*-(1,1,3,3-tetrakis(isopropyl) -1,3-disiloxanediyl) -*O*^{6}-ethyl-2´-deoxyguanosine (28b)**:
C-6
C-2
C-4
C-8
C-5
C-4´
C-1´
C-3´
C-5´
-O-CH$_2$-CH$_3$
d$_6$-DMSO
C-2´
8 x -CH$_3$
-O-CH$_2$-CH$_3$
4 x -C(CH$_3$)$_2$
f1 (ppm)


^{1}H-NMR (400 MHz, $CDCl_3$) of **6-Azido-hexylphosphate**:
-OP(O)(OH)$_2^-$
$CDCl_3$
-CH$_2$-OPO$_3$H$_2$
Diethylether
-CH$_2$-N$_3$
-CH$_2$-CH$_2$-OPO$_3$H$_2$
-CH$_2$-CH$_2$-CH$_2$-N$_3$
Diethylether
f1 (ppm)

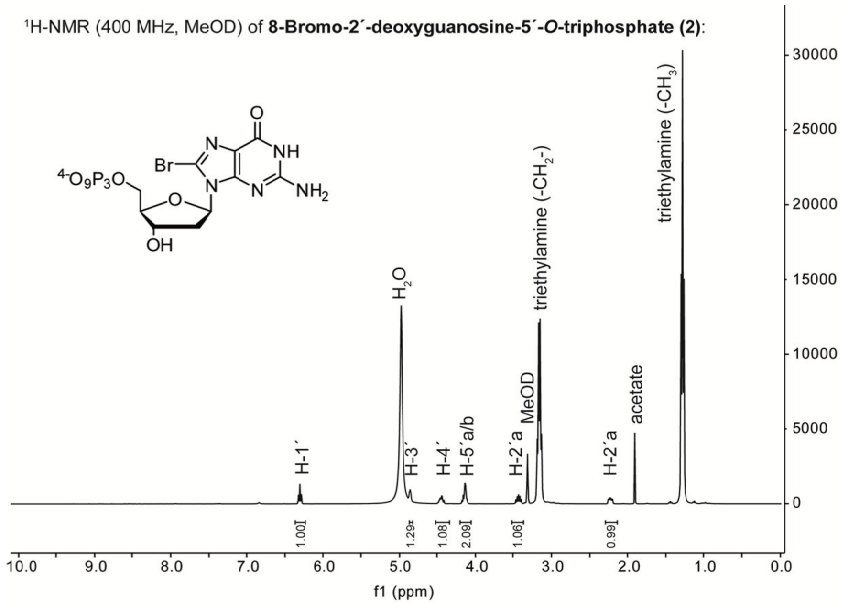

^{1}H-NMR (400 MHz, d_6-DMSO) of **8-Bromo-*O*^{6}-ethyl-2´-deoxyguanosine (29b)**:
H-1´
3´-OH
5´-OH
H-3´
-CH$_2$-CH$_3$
H-4´
H-5´a/b
H-2´a
d$_6$-DMSO
H-2´b
-CH$_3$
f1 (ppm)


^{13}C-NMR (100 MHz, $CDCl_3$) of **6-Azido-hexylphosphate**:
-CH$_2$-OPO$_3$H$_2$
-CH$_2$-N$_3$
-CH$_2$-CH$_2$-OPO$_3$H$_2$
-CH$_2$-CH$_2$-CH$_2$-OPO$_3$H$_2$
-CH$_2$-CH$_2$-CH$_2$-N$_3$
f1 (ppm)

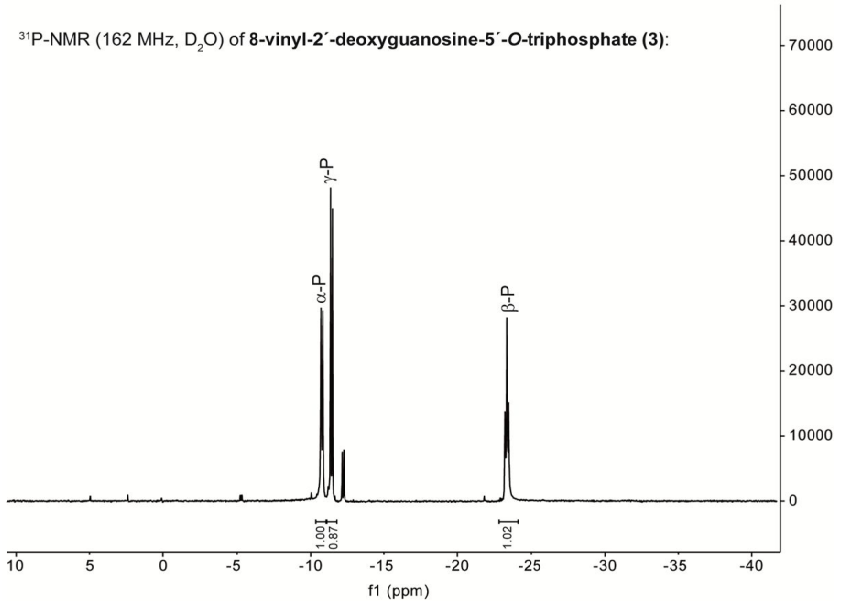
^{1}H-NMR (400 MHz, D_2O) of **8-vinyl-2´-deoxyguanosine-5´-*O*-triphosphate (3)**:



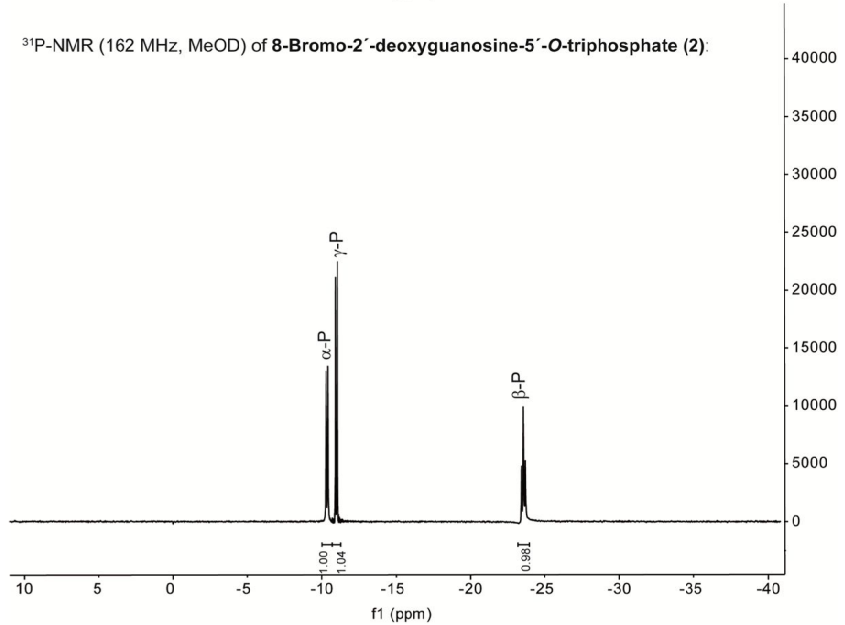
^{1}H-NMR (400 MHz, MeOD) of **8-Bromo-2´-deoxyguanosine-5´-*O*-triphosphate (2)**:



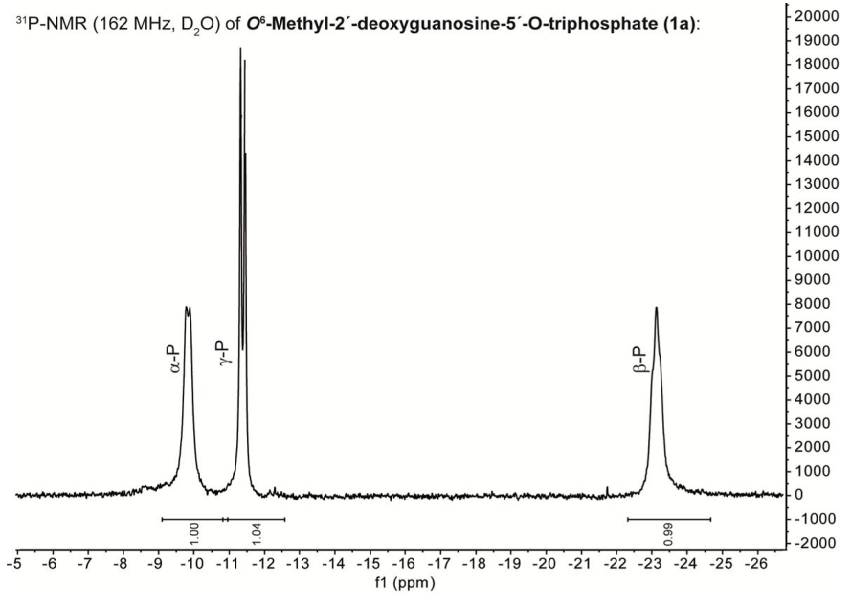
^{31}P-NMR (162 MHz, D_2O) of **8-vinyl-2´-deoxyguanosine-5´-*O*-triphosphate (3)**:



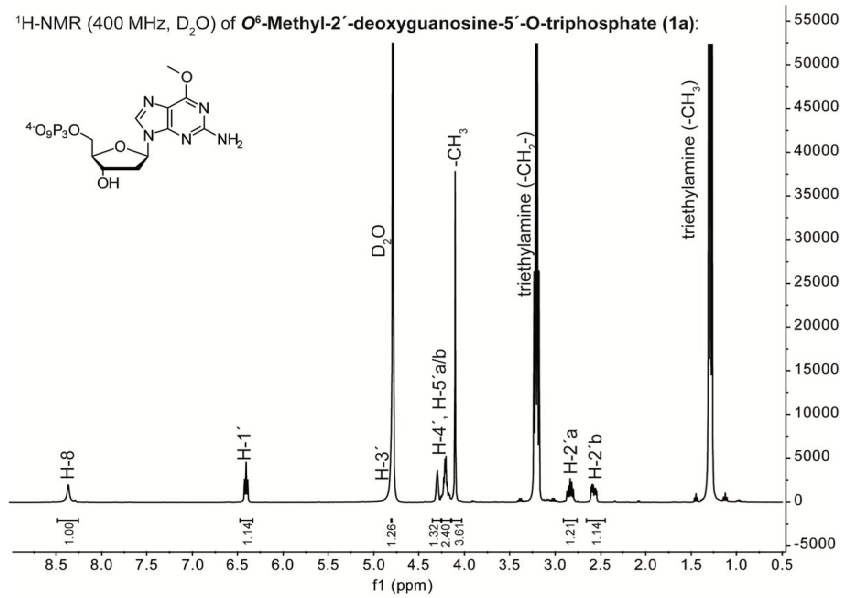
^{31}P-NMR (162 MHz, MeOD) of **8-Bromo-2´-deoxyguanosine-5´-*O*-triphosphate (2)**:



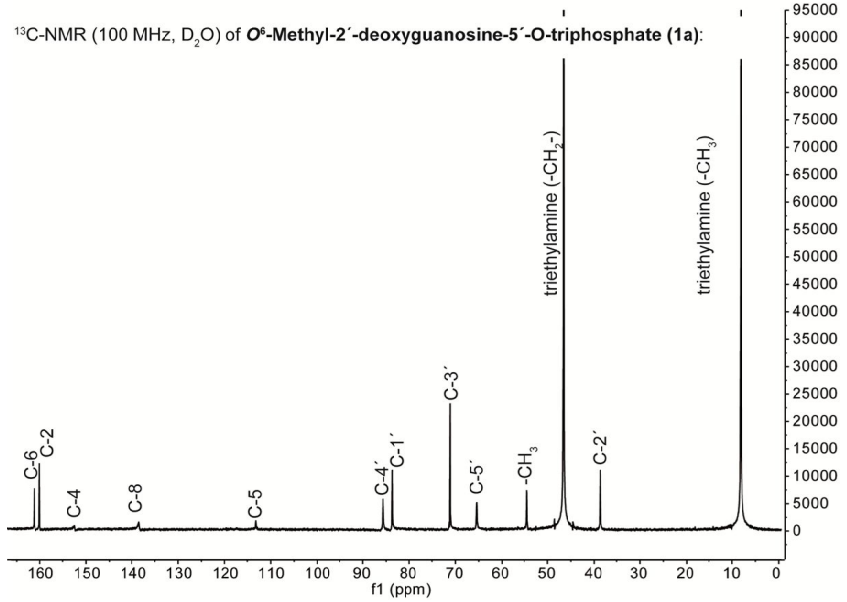
^{31}P-NMR (162 MHz, D_2O) of ***O*^{6}-Methyl-2´-deoxyguanosine-5´-O-triphosphate (1a)**:



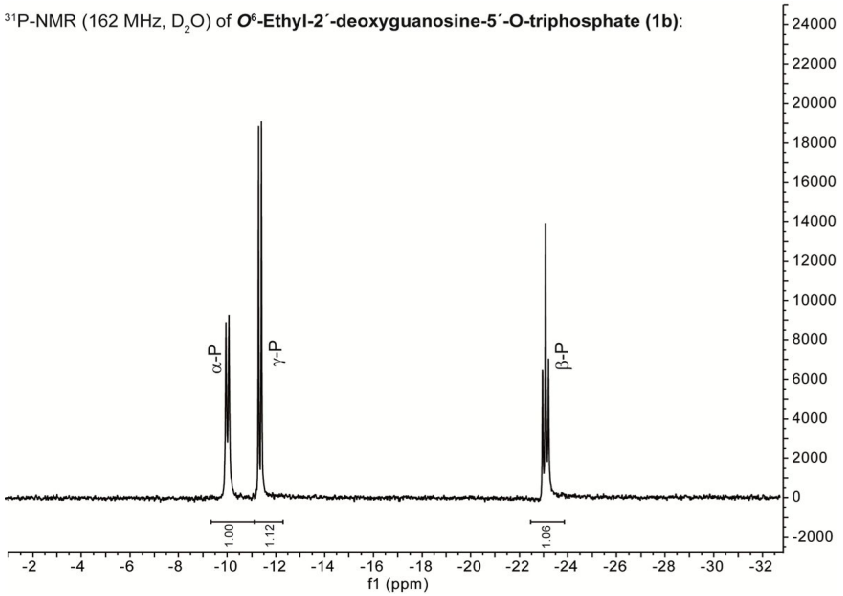
^{1}H-NMR (400 MHz, D_2O) of ***O*^{6}-Methyl-2´-deoxyguanosine-5´-O-triphosphate (1a)**:



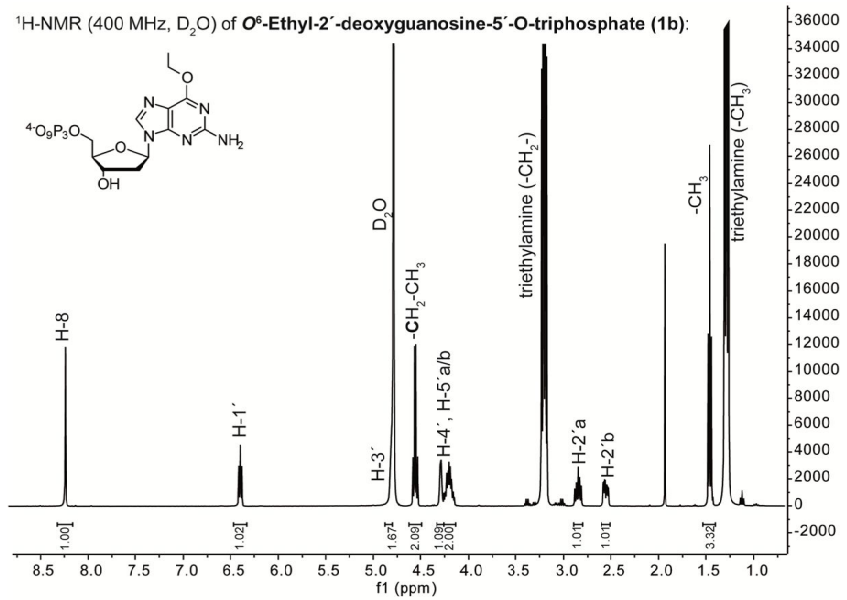
^{13}C-NMR (100 MHz, D_2O) of ***O*^{6}-Methyl-2´-deoxyguanosine-5´-O-triphosphate (1a)**:



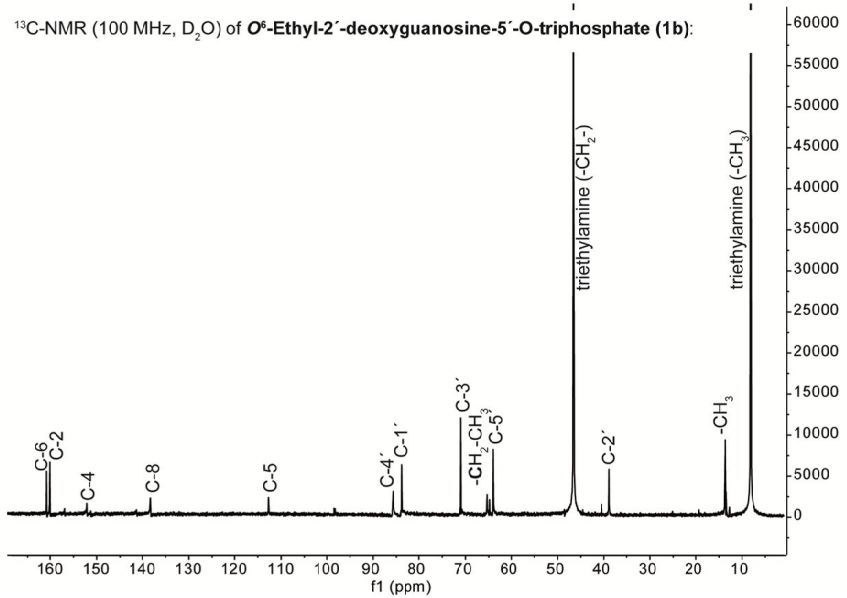
^{31}P-NMR (162 MHz, D_2O) of ***O*^{6}-Ethyl-2´-deoxyguanosine-5´-O-triphosphate (1b)**:



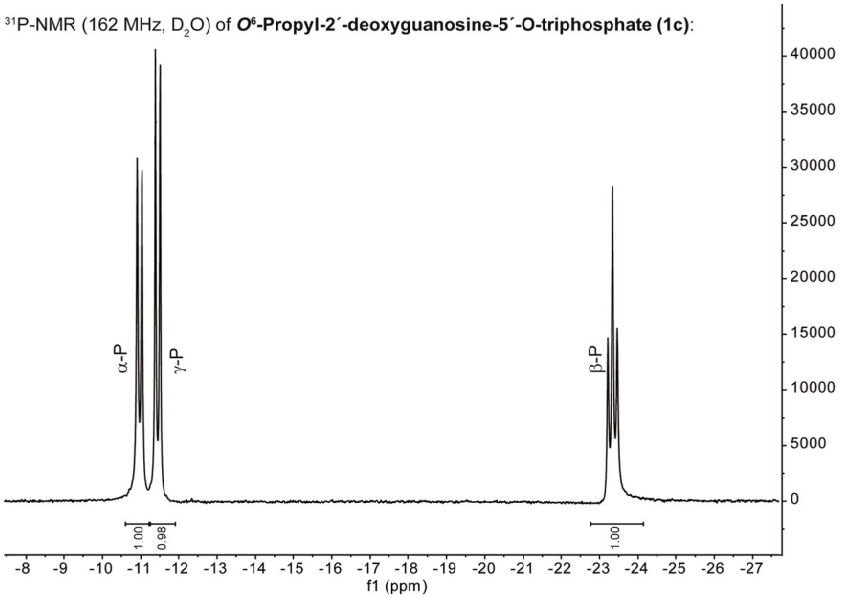
^{1}H-NMR (400 MHz, D_2O) of ***O*^{6}-Ethyl-2´-deoxyguanosine-5´-O-triphosphate (1b)**:



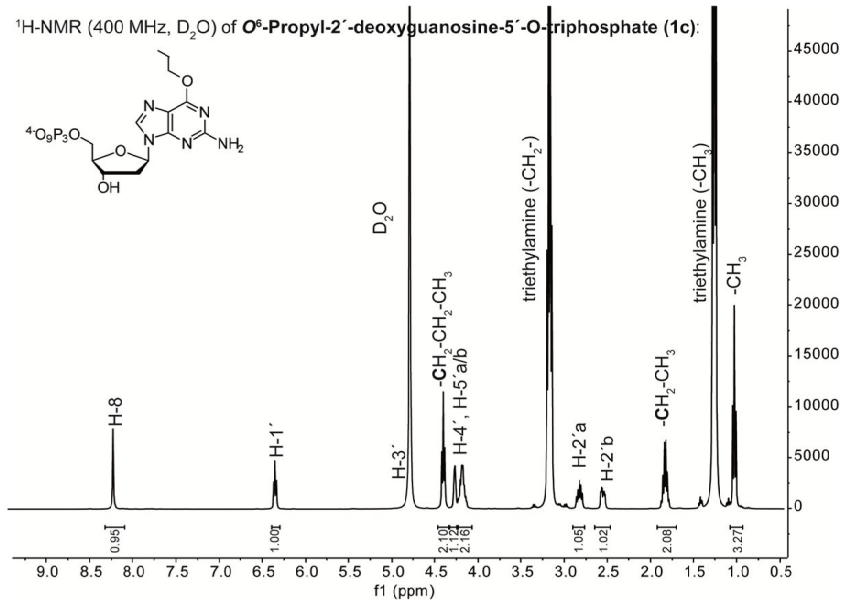
^{13}C-NMR (100 MHz, D_2O) of ***O*^{6}-Ethyl-2´-deoxyguanosine-5´-O-triphosphate (1b)**:



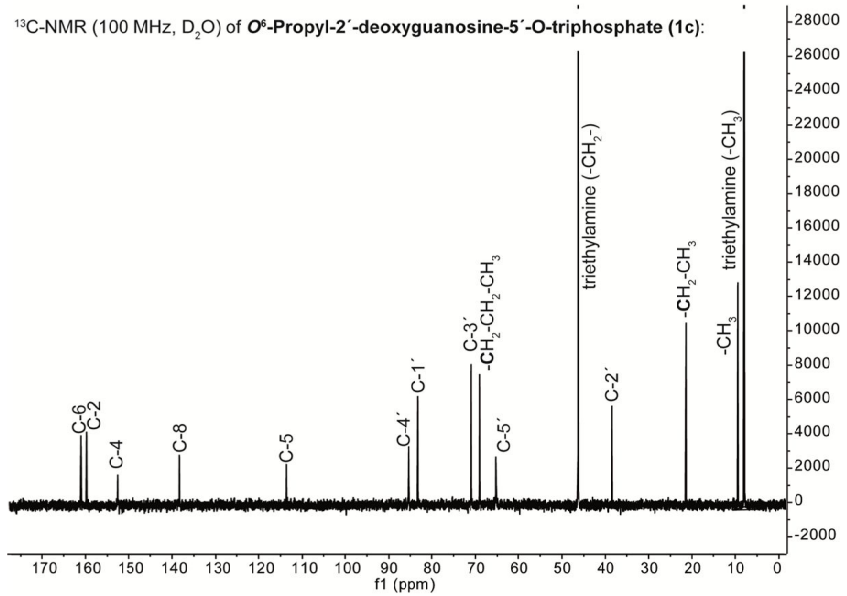
^{31}P-NMR (162 MHz, D_2O) of ***O*^{6}-Propyl-2´-deoxyguanosine-5´-O-triphosphate (1c)**:



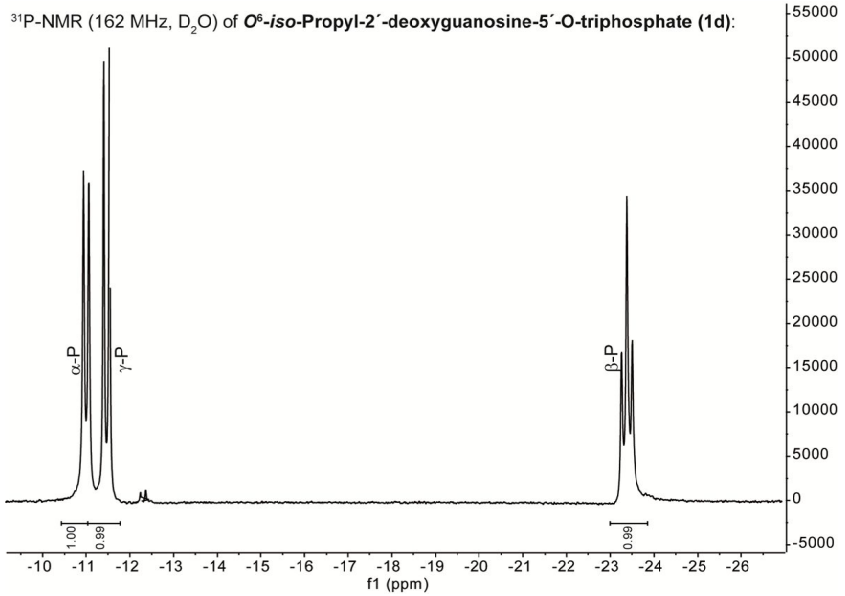
^{1}H-NMR (400 MHz, D_2O) of ***O*^{6}-Propyl-2´-deoxyguanosine-5´-O-triphosphate (1c)**:



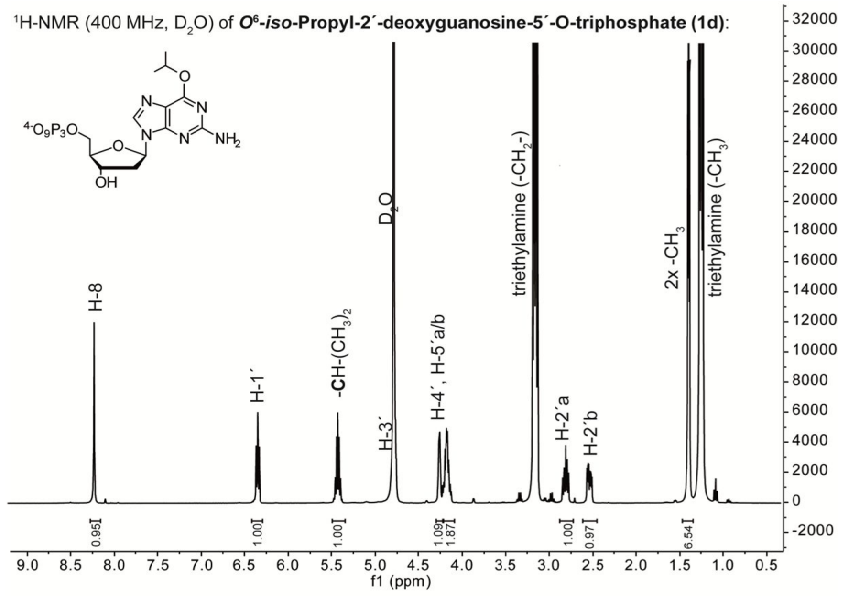
^{13}C-NMR (100 MHz, D_2O) of ***O*^{6}-Propyl-2´-deoxyguanosine-5´-O-triphosphate (1c)**:



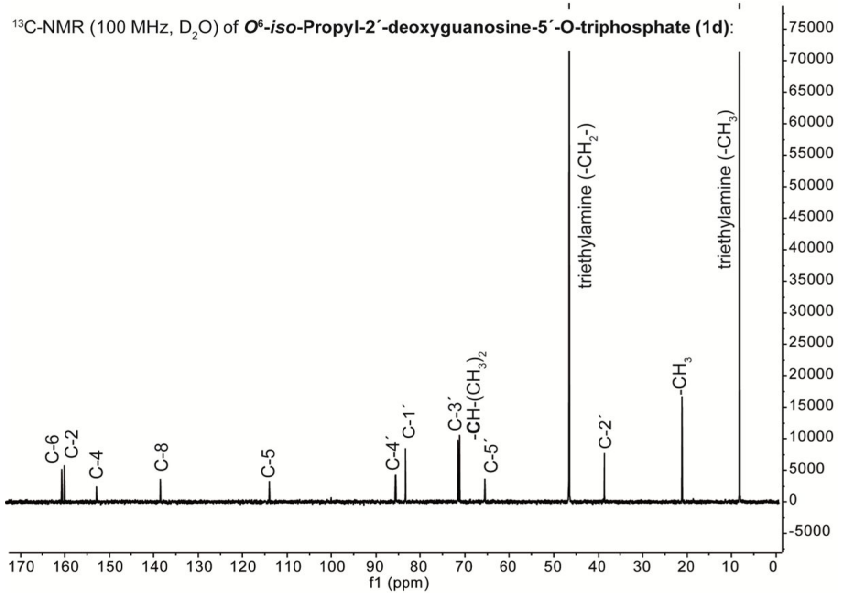
^{31}P-NMR (162 MHz, D_2O) of ***O*^{6}*-iso*-Propyl-2´-deoxyguanosine-5´-O-triphosphate (1d)**:



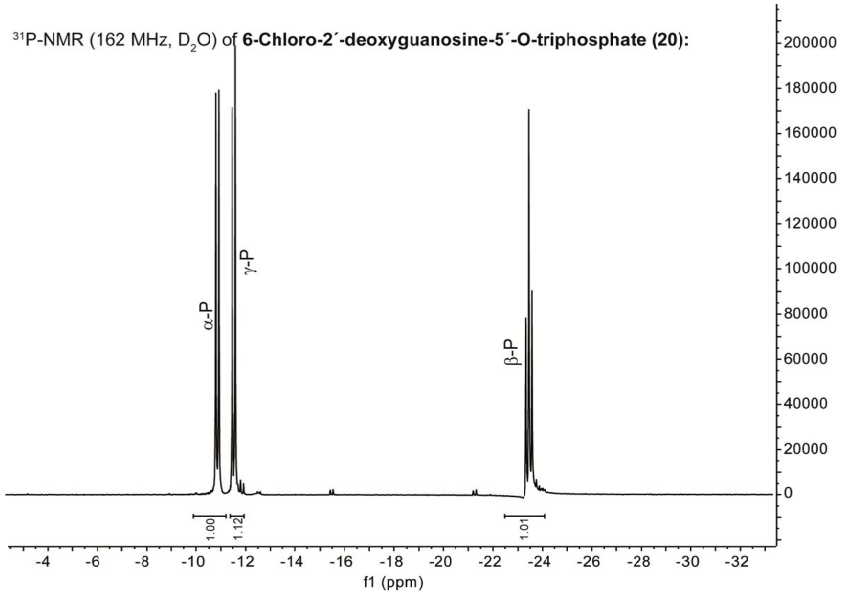
^{1}H-NMR (400 MHz, D_2O) of ***O*^{6}*-iso*-Propyl-2´-deoxyguanosine-5´-O-triphosphate (1d)**:



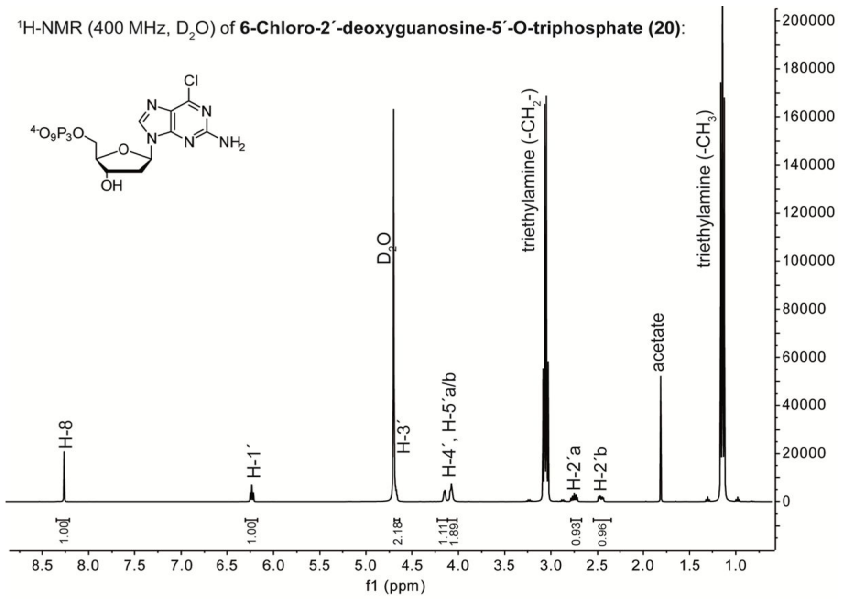
^{13}C-NMR (100 MHz, D_2O) of ***O*^{6}*-iso*-Propyl-2´-deoxyguanosine-5´-O-triphosphate (1d)**:



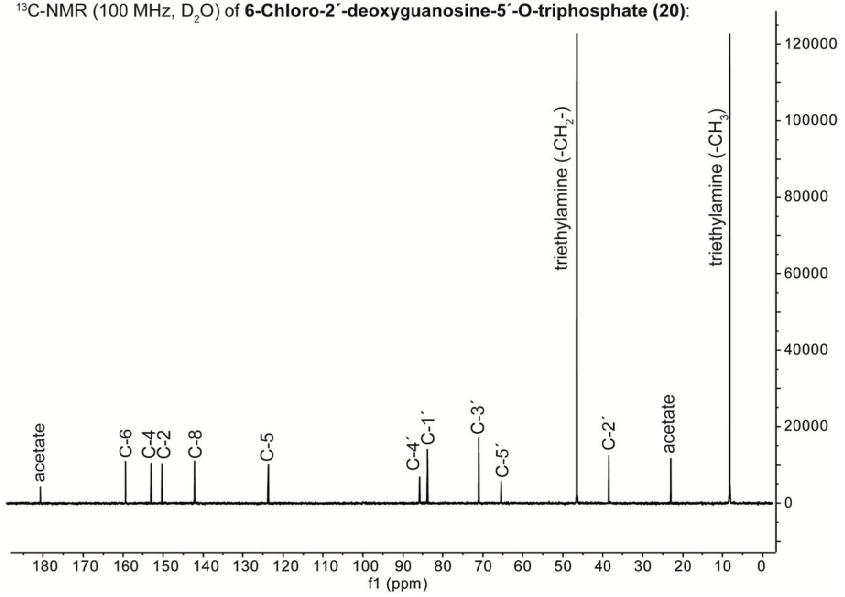
^{31}P-NMR (162 MHz, D_2O) of **6-Chloro-2´-deoxyguanosine-5´-O-triphosphate (20):**



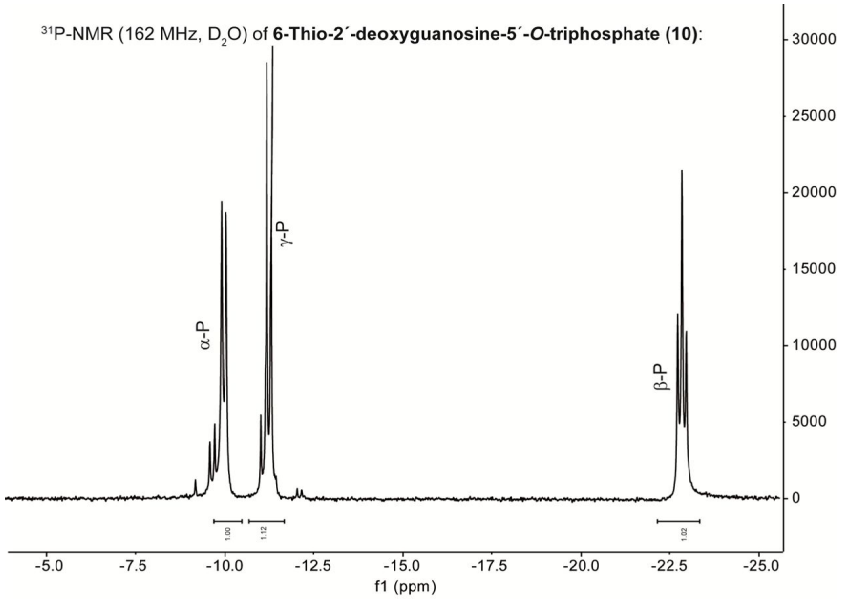
^{1}H-NMR (400 MHz, D_2O) of **6-Chloro-2´-deoxyguanosine-5´-O-triphosphate (20)**:



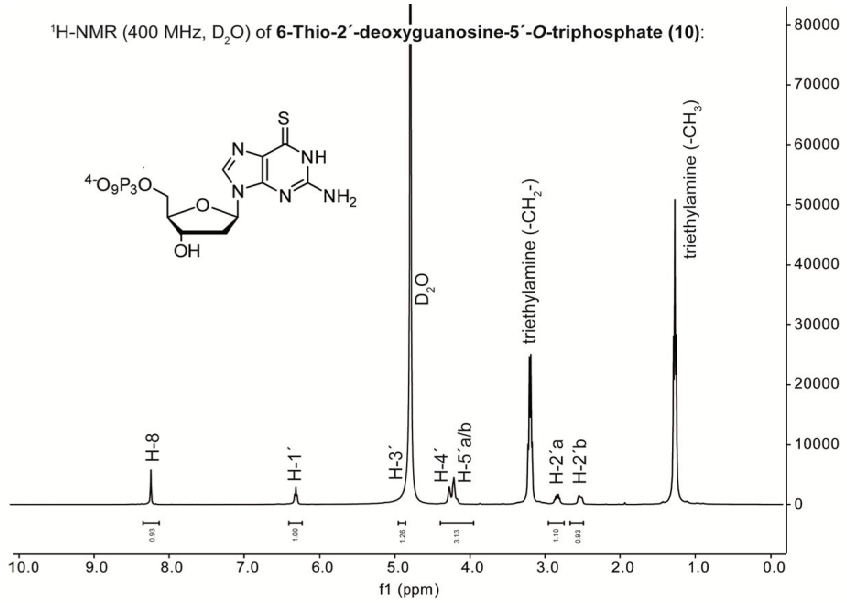
^{13}C-NMR (100 MHz, D_2O) of **6-Chloro-2´-deoxyguanosine-5´-O-triphosphate (20)**:



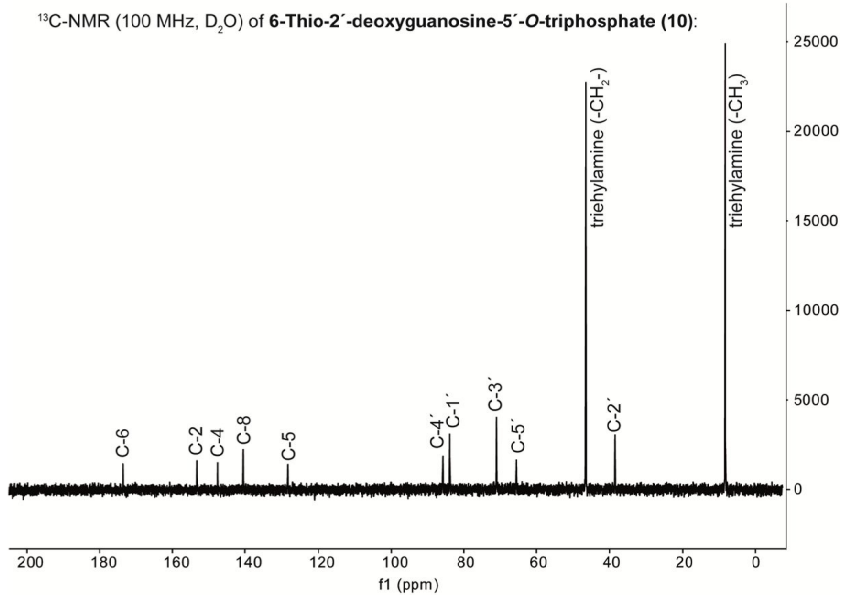
^{31}P-NMR (162 MHz, D_2O) of **6-Thio-2´-deoxyguanosine-5´-*O*-triphosphate (10)**:

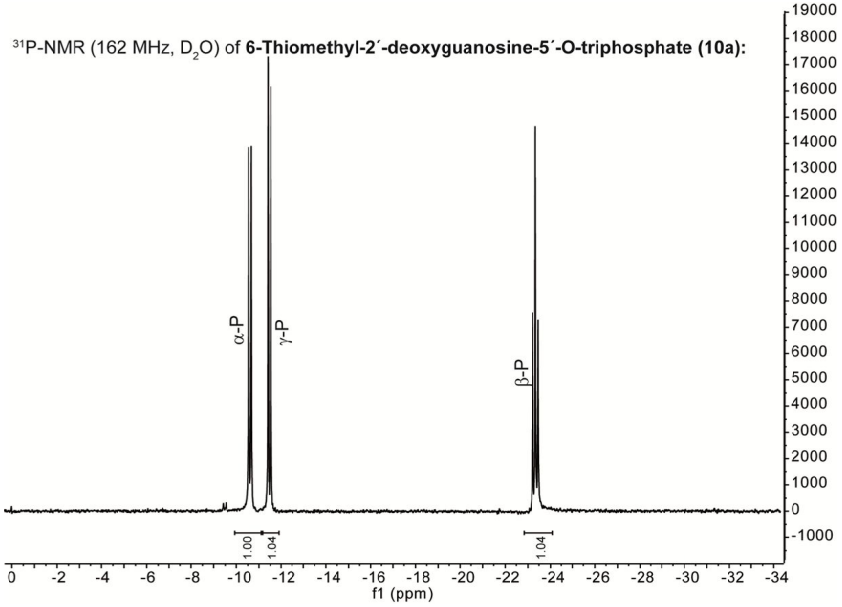


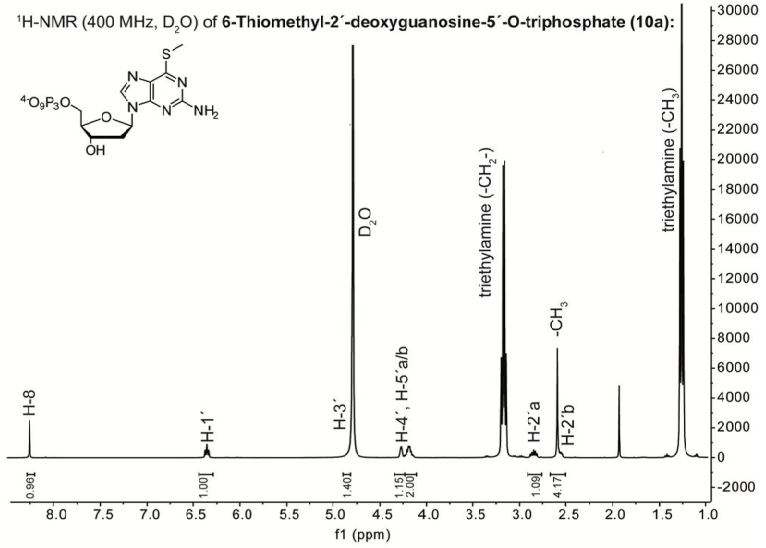
^{1}H-NMR (400 MHz, D_2O) of **6-Thio-2´-deoxyguanosine-5´-*O*-triphosphate (10)**:

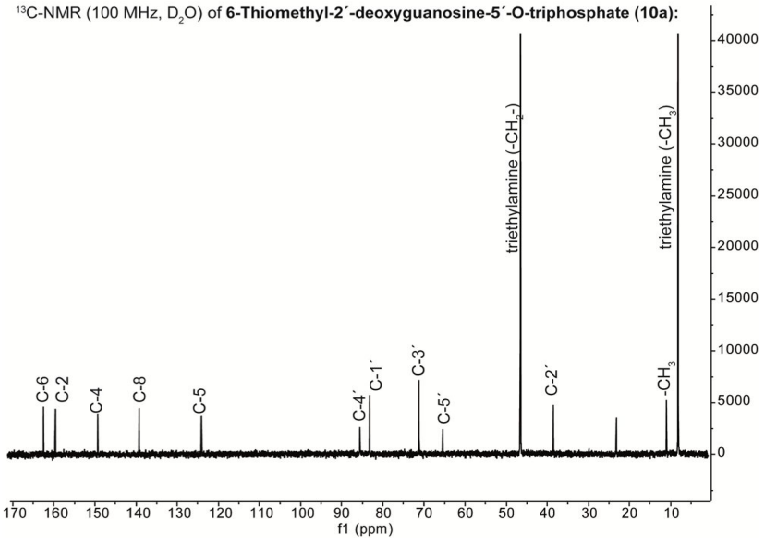


^{13}C-NMR (100 MHz, D_2O) of **6-Thio-2´-deoxyguanosine-5´-*O*-triphosphate (10)**:

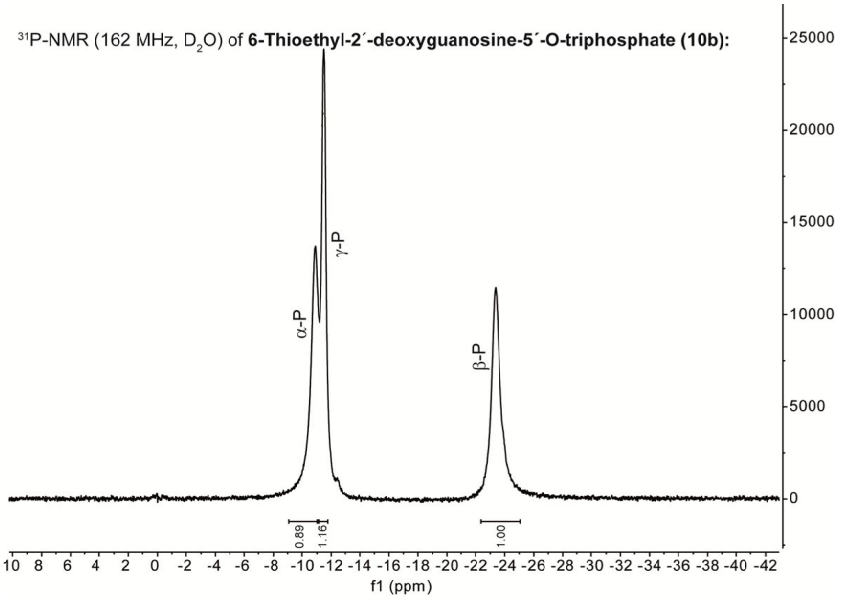



31P-NMR (162 MHz, D2O) of 6-Thiomethyl-2´-deoxyguanosine-5´-O-triphosphate (10a):
α-P
γ-P
β-P
1.00
1.04
1.04
f1 (ppm)

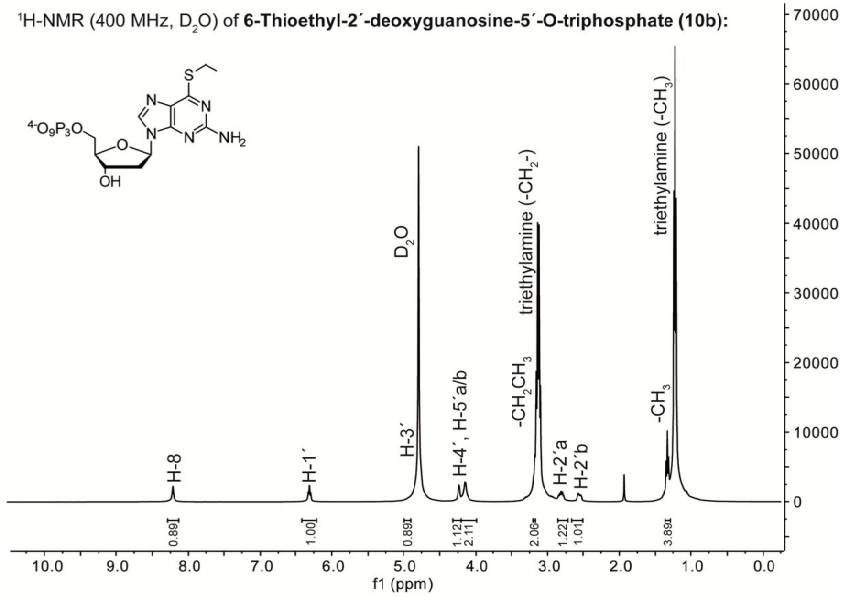

1H-NMR (400 MHz, D2O) of 6-Thiomethyl-2´-deoxyguanosine-5´-O-triphosphate (10a):
4-O9P3O
OH
NH2
H-8
H-1´
H-3´
D2O
H-4´, H-5´a/b
triethylamine (-CH2-)
-CH3
H-2´a
H-2´b
triethylamine (-CH3)
0.96
1.00
1.40
1.15
2.00
1.09
4.17
f1 (ppm)


13C-NMR (100 MHz, D2O) of 6-Thiomethyl-2´-deoxyguanosine-5´-O-triphosphate (10a):
C-6
C-2
C-4
C-8
C-5
C-4´
C-1´
C-3´
C-5´
triethylamine (-CH2-)
C-2´
-CH3
triethylamine (-CH3)
f1 (ppm)

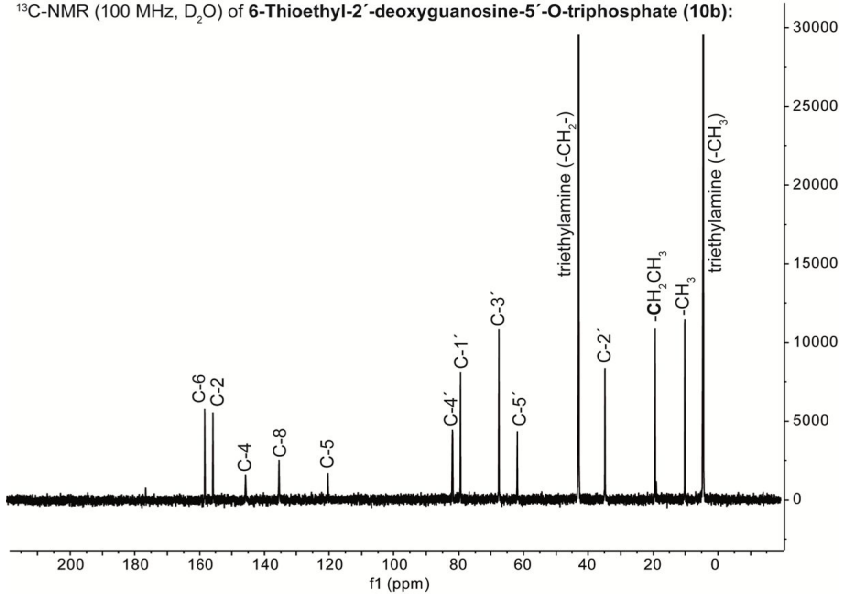
^{31}P-NMR (162 MHz, D_2O) of **6-Thioethyl-2´-deoxyguanosine-5´-*O*-triphosphate (10b):**



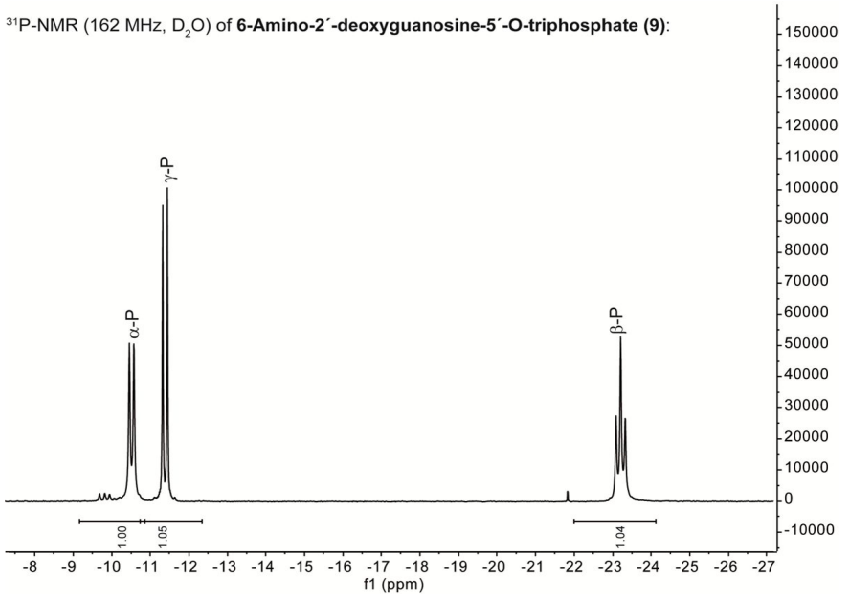
^{1}H-NMR (400 MHz, D_2O) of **6-Thioethyl-2´-deoxyguanosine-5´-*O*-triphosphate (10b):**



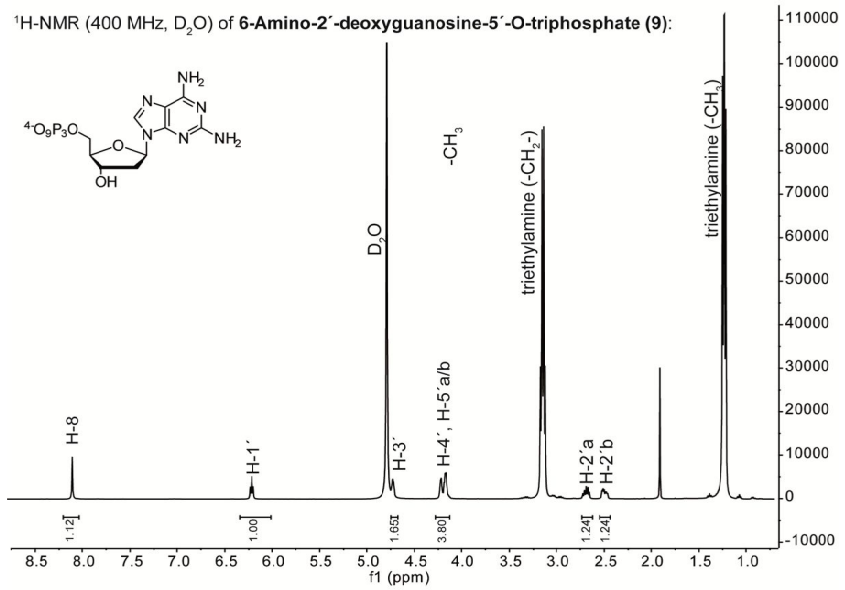
^{13}C-NMR (100 MHz, D_2O) of **6-Thioethyl-2´-deoxyguanosine-5´-*O*-triphosphate (10b):**



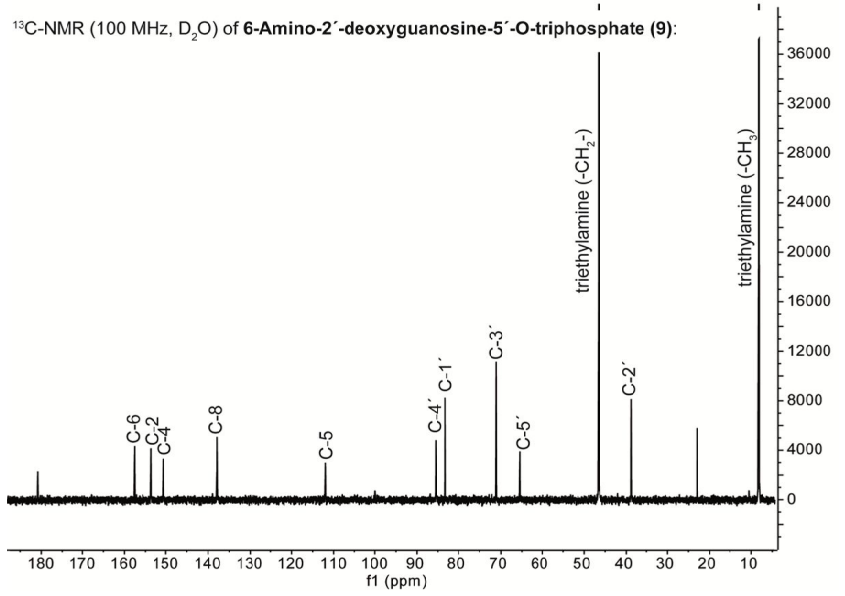
^{31}P-NMR (162 MHz, D_2O) of **6-Amino-2´-deoxyguanosine-5´-O-triphosphate (9)**:



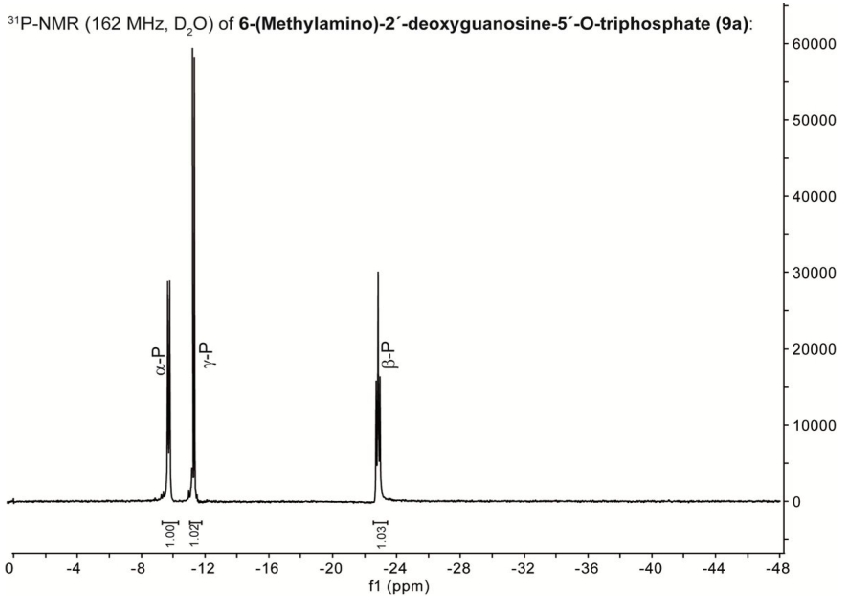
^{1}H-NMR (400 MHz, D_2O) of **6-Amino-2´-deoxyguanosine-5´-O-triphosphate (9)**:



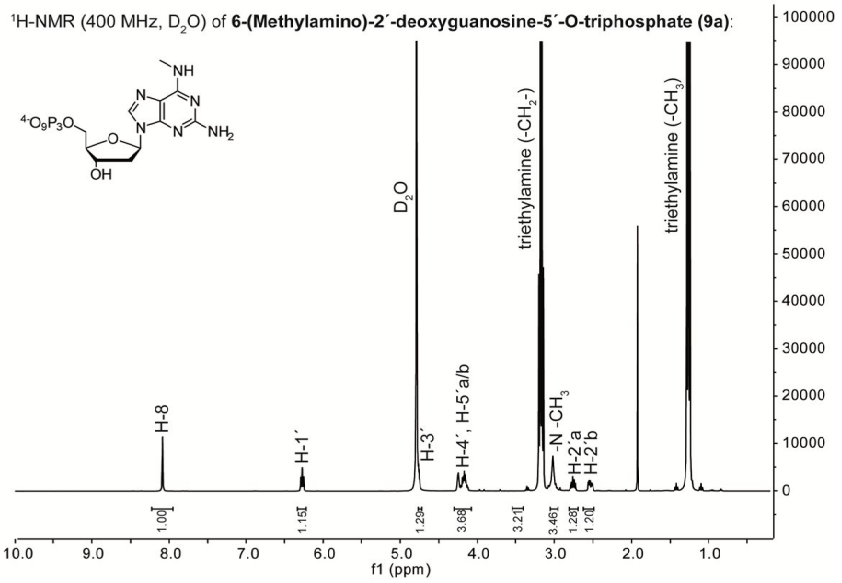
^{13}C-NMR (100 MHz, D_2O) of **6-Amino-2´-deoxyguanosine-5´-O-triphosphate (9)**:



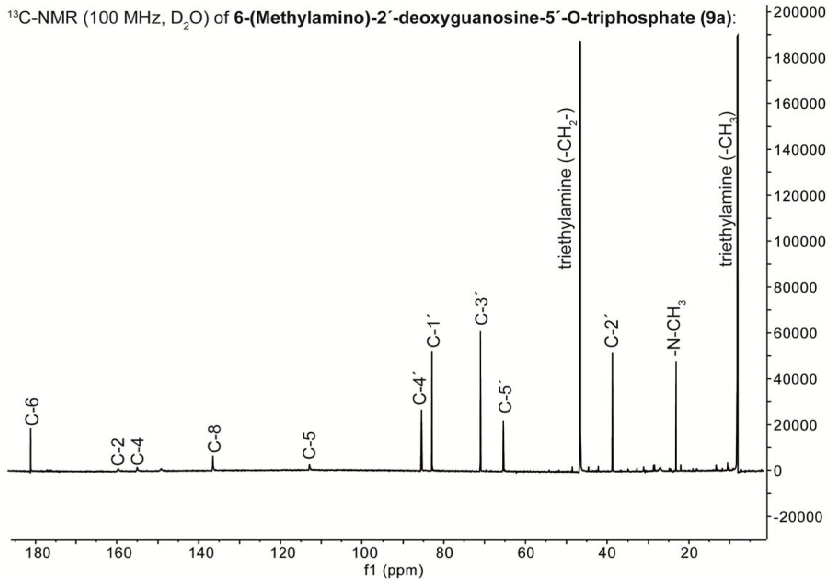
^{31}P-NMR (162 MHz, D_2O) of **6-(Methylamino)-2´-deoxyguanosine-5´-O-triphosphate (9a)**:

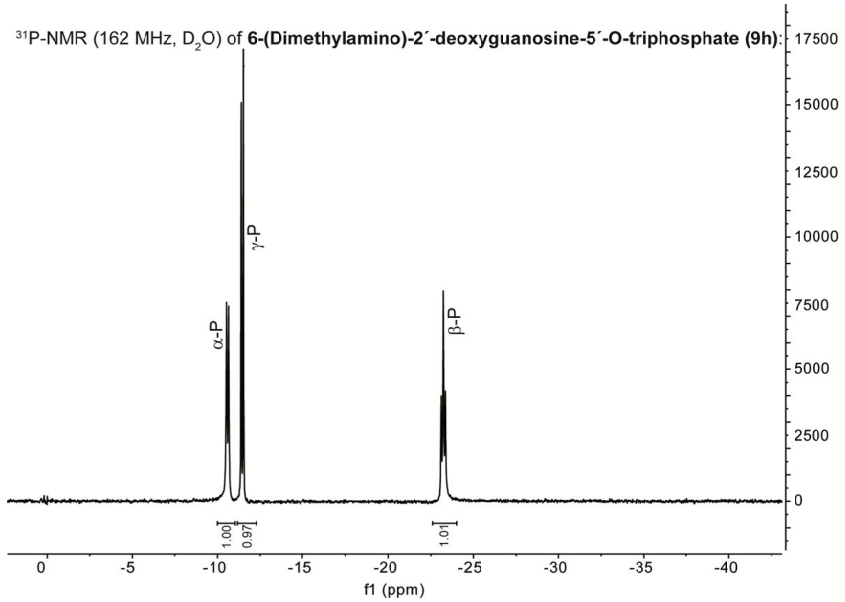


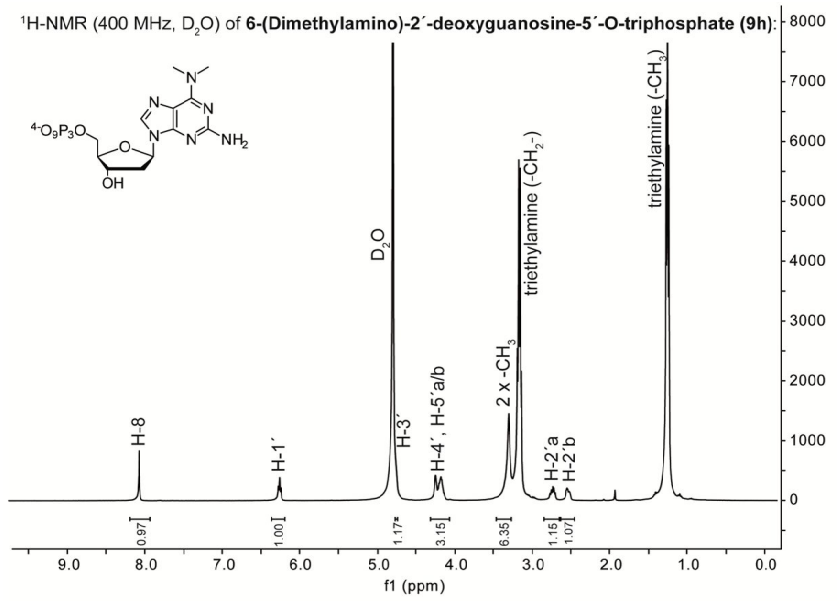
^{1}H-NMR (400 MHz, D_2O) of **6-(Methylamino)-2´-deoxyguanosine-5´-O-triphosphate (9a)**:

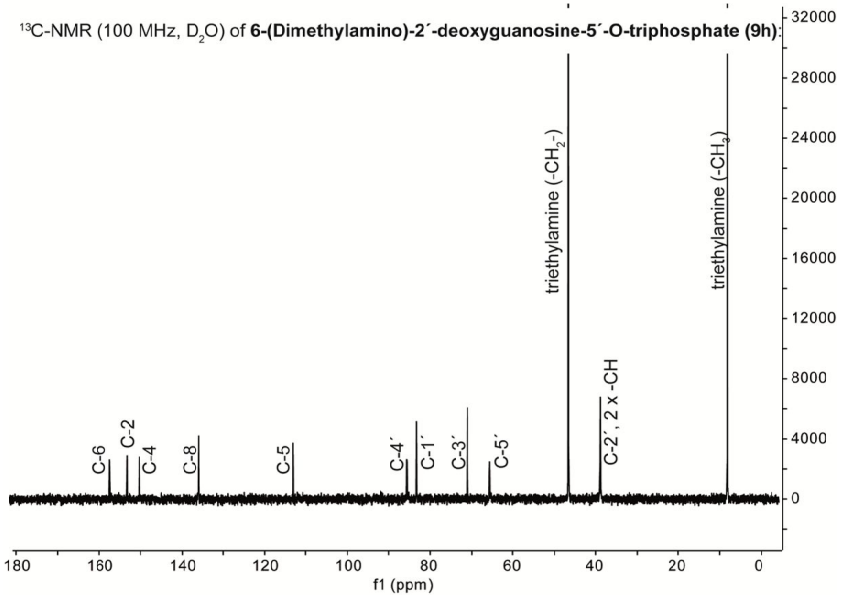


^{13}C-NMR (100 MHz, D_2O) of **6-(Methylamino)-2´-deoxyguanosine-5´-O-triphosphate (9a)**:

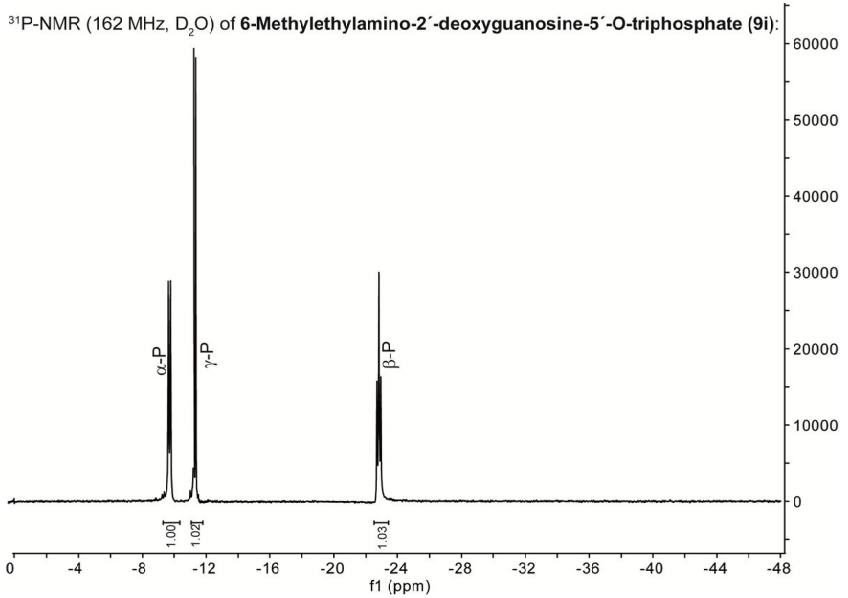



31P-NMR (162 MHz, D2O) of 6-(Dimethylamino)-2´-deoxyguanosine-5´-O-triphosphate (9h):
α-P
γ-P
β-P
1.00
0.97
1.01
f1 (ppm)

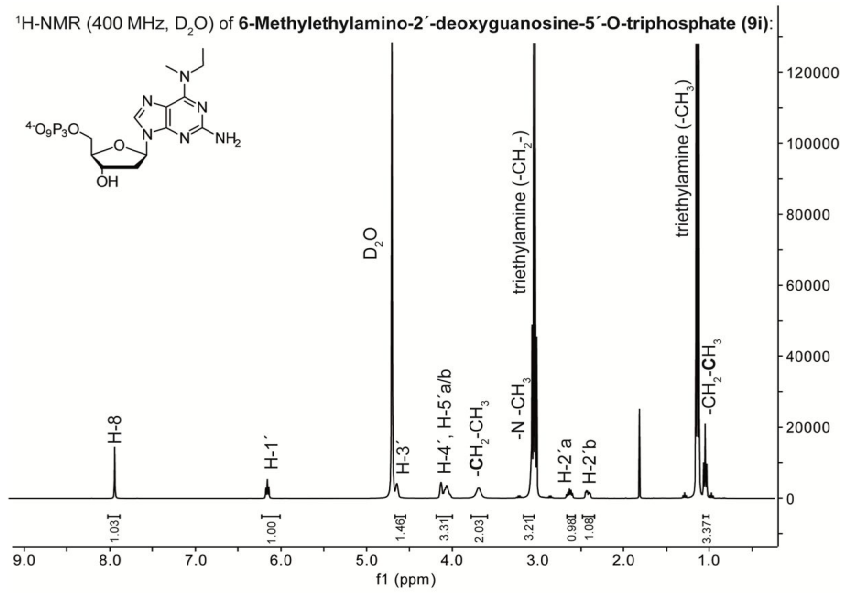

1H-NMR (400 MHz, D2O) of 6-(Dimethylamino)-2´-deoxyguanosine-5´-O-triphosphate (9h):
4-O9P3O
OH
NH2
H-8
H-1´
D2O
H-3´
H-4´, H-5´a/b
2 x -CH3
triethylamine (-CH2-)
H-2´a
H-2´b
triethylamine (-CH3)
0.97
1.00
1.17*
3.15
6.35
1.15
1.07
f1 (ppm)


13C-NMR (100 MHz, D2O) of 6-(Dimethylamino)-2´-deoxyguanosine-5´-O-triphosphate (9h):
C-6
C-2
C-4
C-8
C-5
C-4´
C-1´
C-3´
C-5´
triethylamine (-CH2-)
C-2´, 2 x -CH
triethylamine (-CH3)
f1 (ppm)

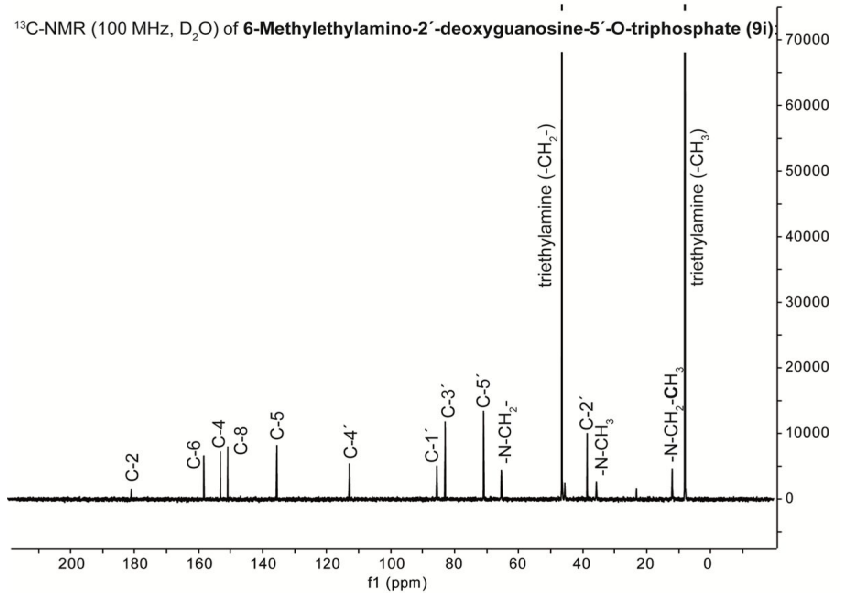
^{31}P-NMR (162 MHz, D_2O) of **6-Methylethylamino-2´-deoxyguanosine-5´-O-triphosphate (9i)**:



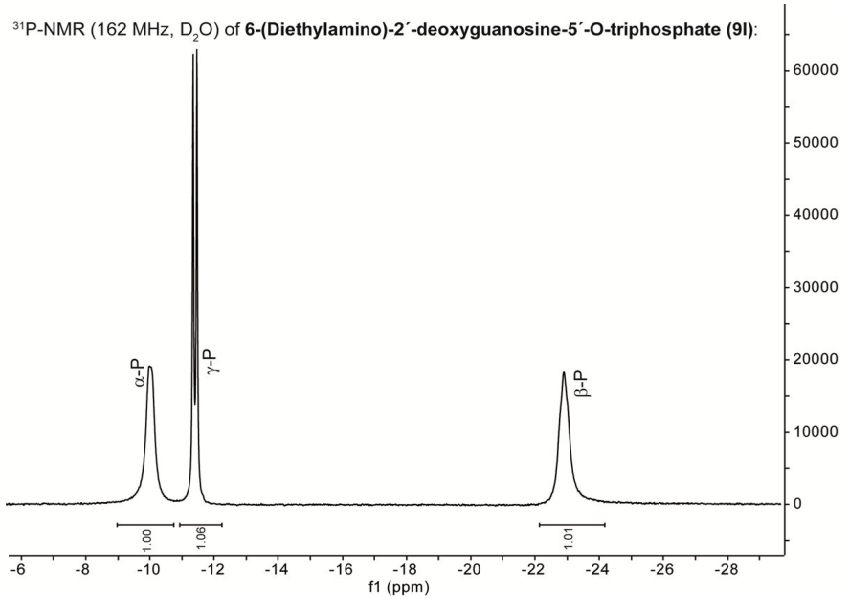
^{1}H-NMR (400 MHz, D_2O) of **6-Methylethylamino-2´-deoxyguanosine-5´-O-triphosphate (9i)**:



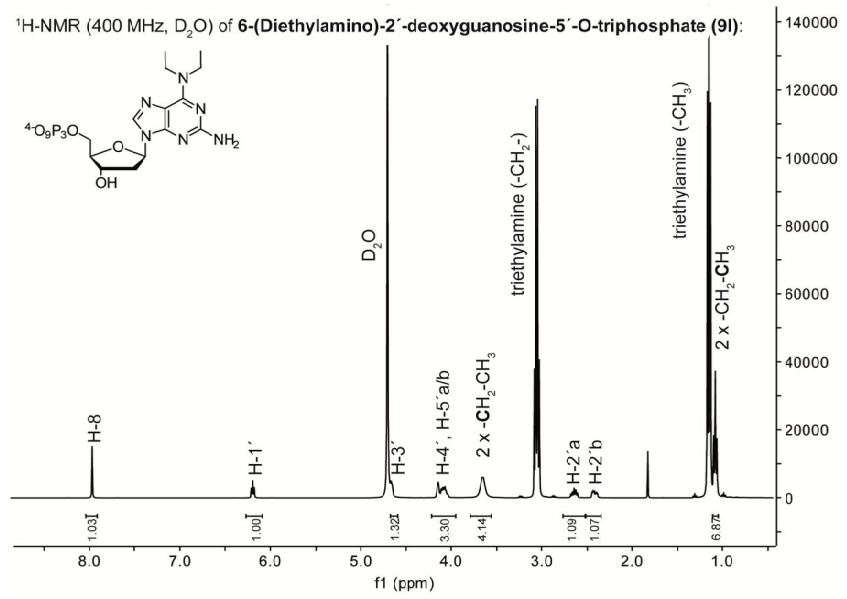
^{13}C-NMR (100 MHz, D_2O) of **6-Methylethylamino-2´-deoxyguanosine-5´-O-triphosphate (9i)**:



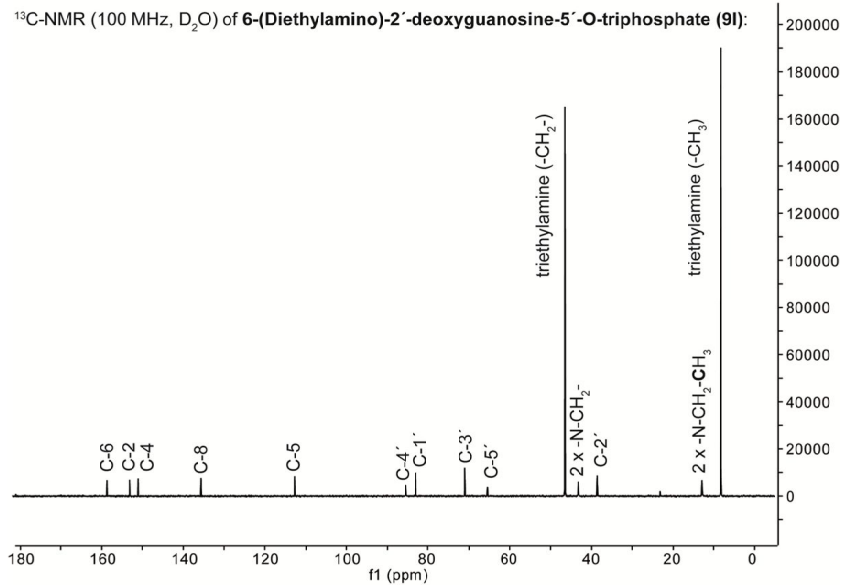
^{31}P-NMR (162 MHz, D_2O) of **6-(Diethylamino)-2´-deoxyguanosine-5´-O-triphosphate (9l)**:

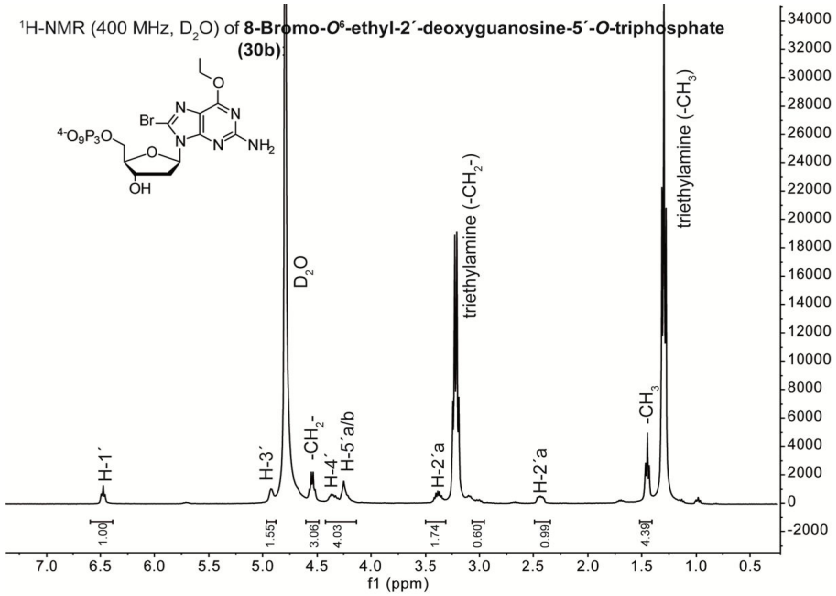


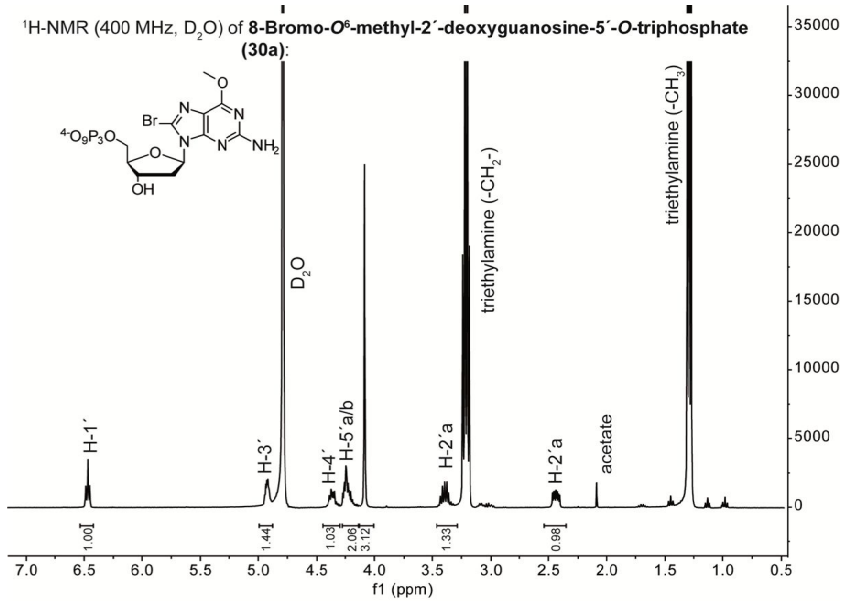
^{1}H-NMR (400 MHz, D_2O) of **6-(Diethylamino)-2´-deoxyguanosine-5´-O-triphosphate (9l)**:

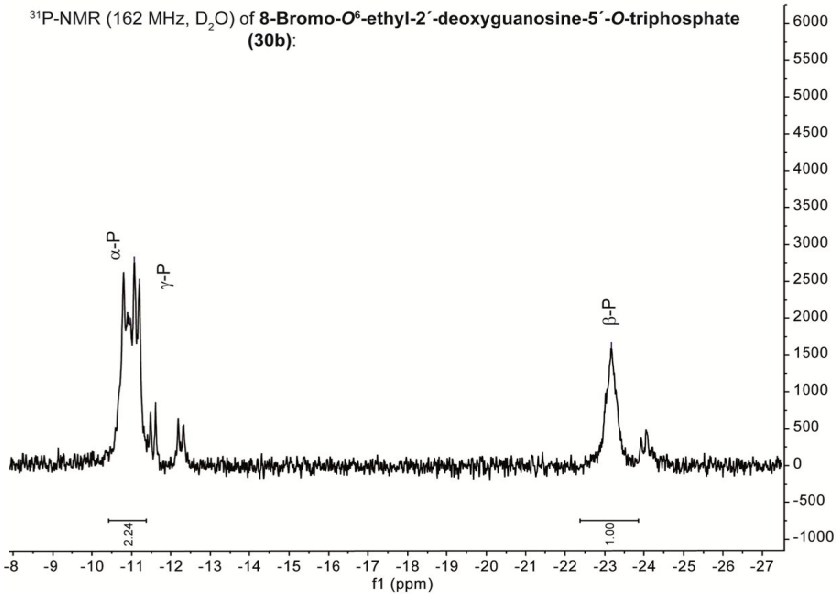


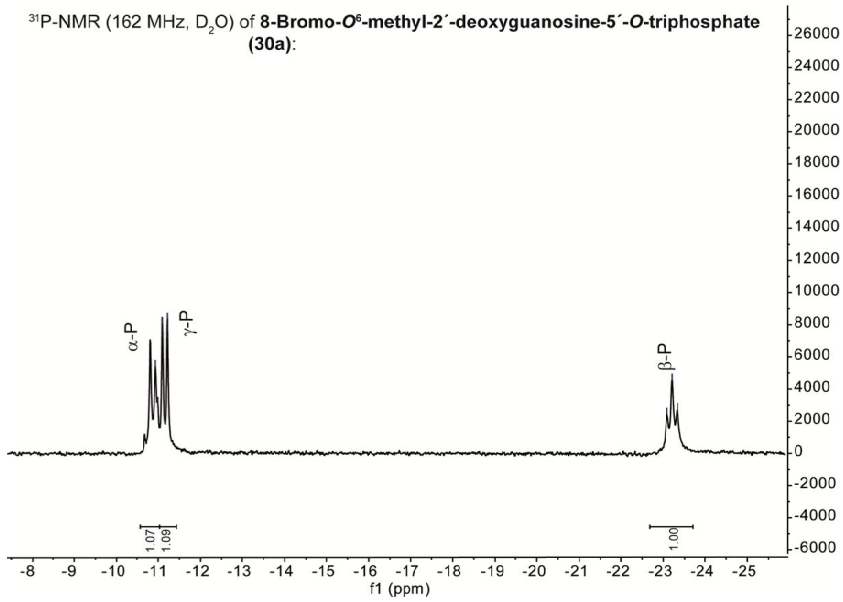
^{13}C-NMR (100 MHz, D_2O) of **6-(Diethylamino)-2´-deoxyguanosine-5´-O-triphosphate (9l)**:

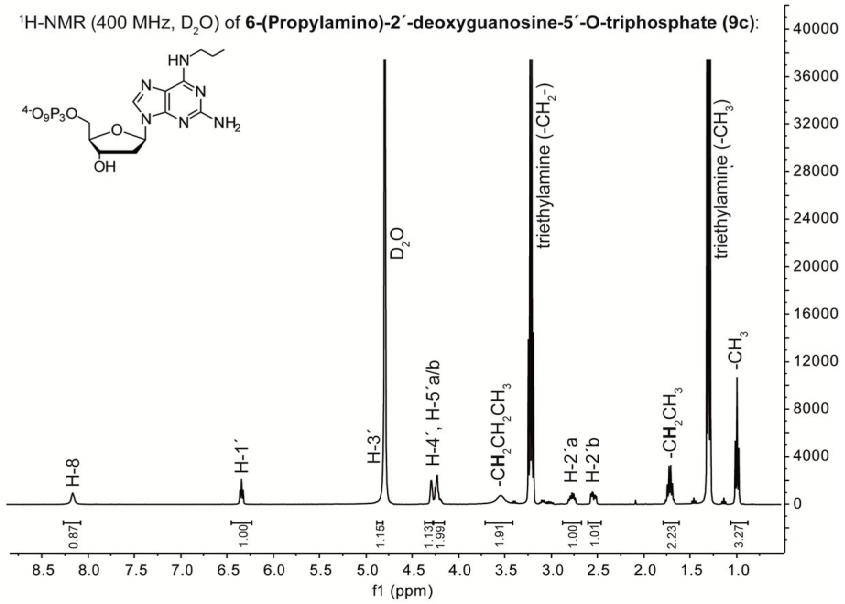


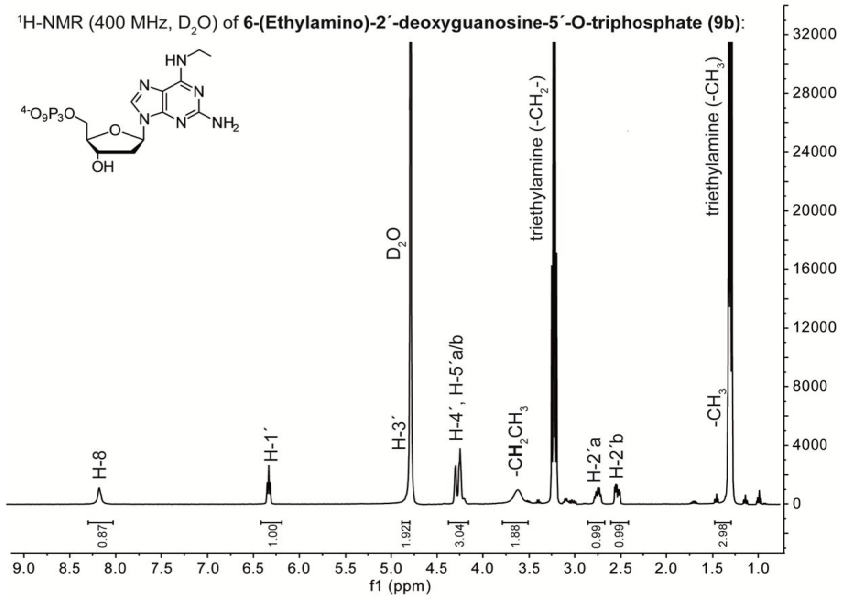

^{1}H-NMR (400 MHz, D_2O) of **8-Bromo-*O*^{6}-ethyl-2´-deoxyguanosine-5´-*O*-triphosphate (30b)**:
H-1´
H-3´
D$_2$O
-CH$_2$-
H-4´
H-5´a/b
H-2´a
triethylamine (-CH$_2$-)
H-2´a
-CH$_3$
triethylamine (-CH$_3$)
f1 (ppm)

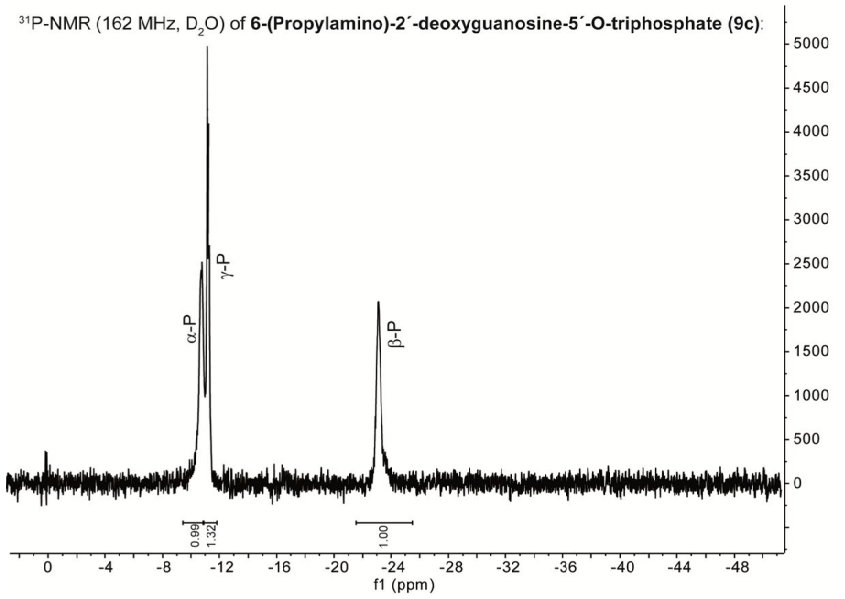

^{1}H-NMR (400 MHz, D_2O) of **8-Bromo-*O*^{6}-methyl-2´-deoxyguanosine-5´-*O*-triphosphate (30a)**:
H-1´
H-3´
D$_2$O
H-4´
H-5´a/b
H-2´a
triethylamine (-CH$_2$-)
H-2´a
acetate
triethylamine (-CH$_3$)
f1 (ppm)

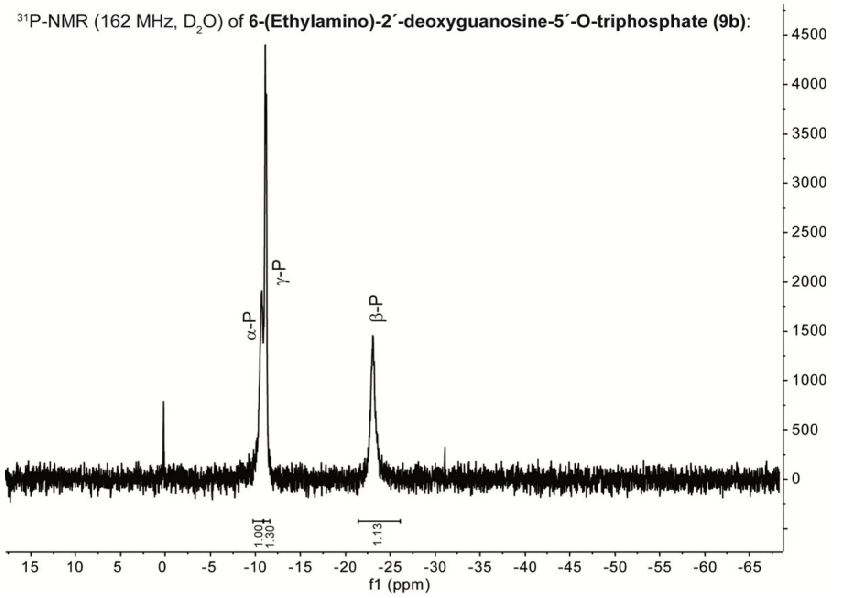

^{31}P-NMR (162 MHz, D_2O) of **8-Bromo-*O*^{6}-ethyl-2´-deoxyguanosine-5´-*O*-triphosphate (30b)**:
α-P
γ-P
β-P
f1 (ppm)

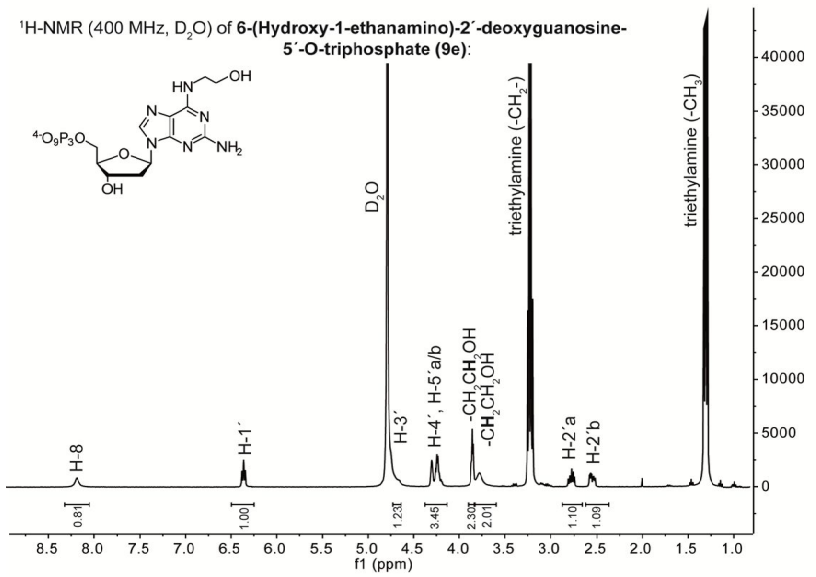

^{31}P-NMR (162 MHz, D_2O) of **8-Bromo-*O*^{6}-methyl-2´-deoxyguanosine-5´-*O*-triphosphate (30a)**:
α-P
γ-P
β-P
f1 (ppm)

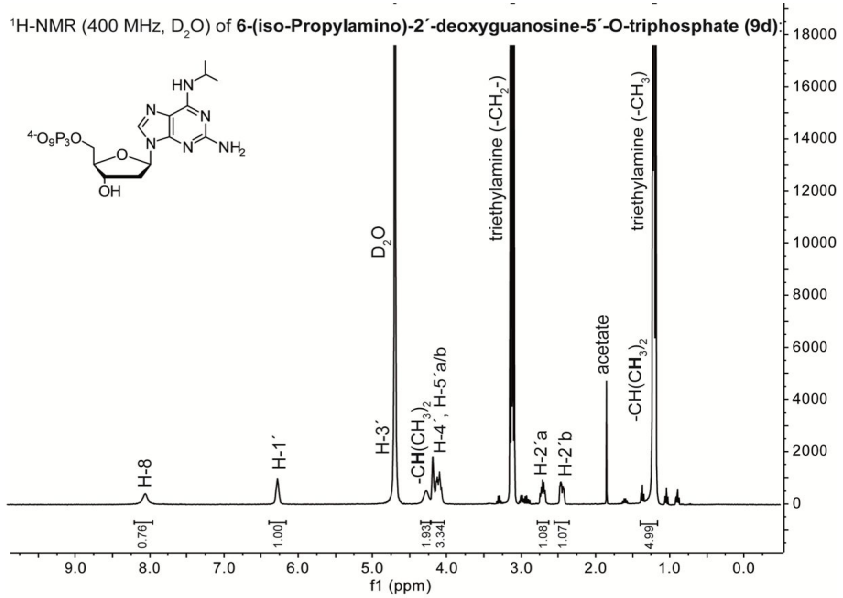

^{1}H-NMR (400 MHz, D_2O) of **6-(Propylamino)-2´-deoxyguanosine-5´-O-triphosphate (9c)**:
H-8
H-1´
H-3´
D_2O
H-4´, H-5´a/b
$-CH_2CH_2CH_3$
triethylamine ($-CH_2-$)
H-2´a
H-2´b
$-CH_2CH_3$
triethylamine ($-CH_3$)
$-CH_3$
f1 (ppm)

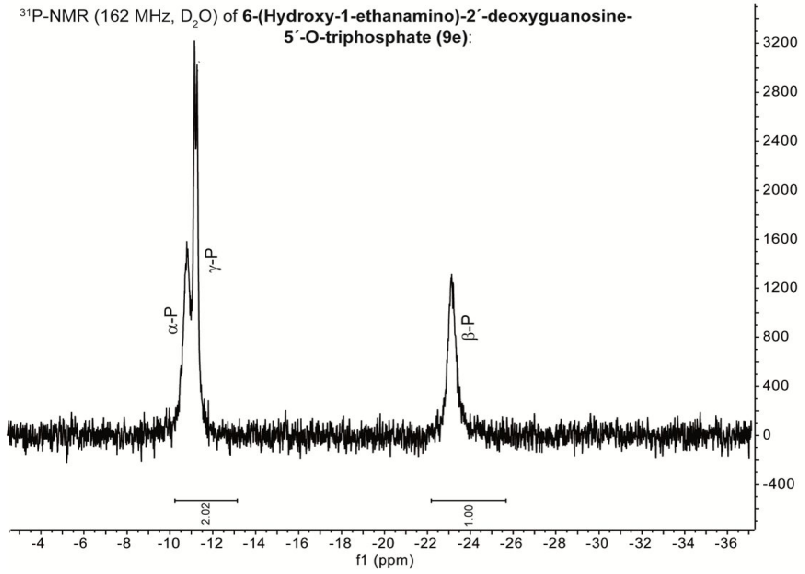

^{1}H-NMR (400 MHz, D_2O) of **6-(Ethylamino)-2´-deoxyguanosine-5´-O-triphosphate (9b)**:
H-8
H-1´
H-3´
D_2O
H-4´, H-5´a/b
$-CH_2CH_3$
triethylamine ($-CH_2-$)
H-2´a
H-2´b
triethylamine ($-CH_3$)
$-CH_3$
f1 (ppm)

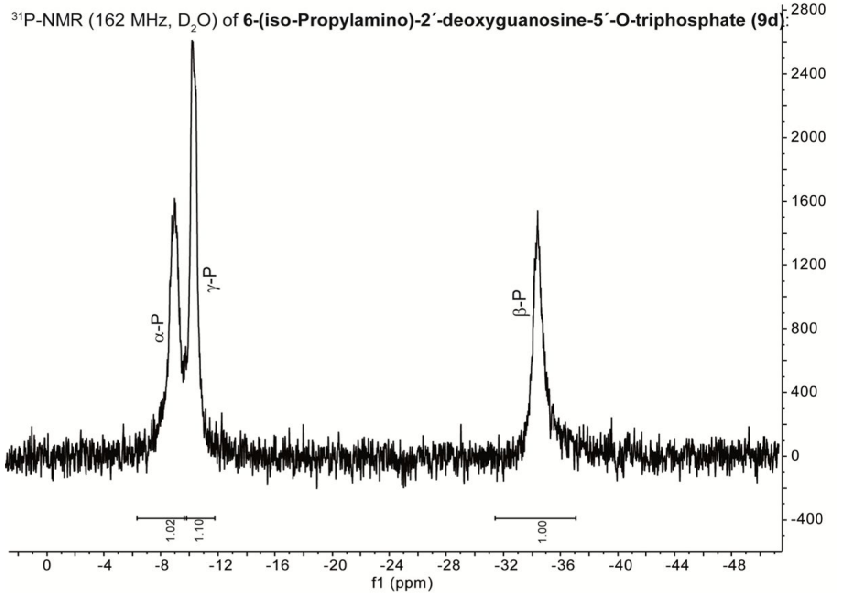

^{31}P-NMR (162 MHz, D_2O) of **6-(Propylamino)-2´-deoxyguanosine-5´-O-triphosphate (9c)**:
α-P
γ-P
β-P
f1 (ppm)

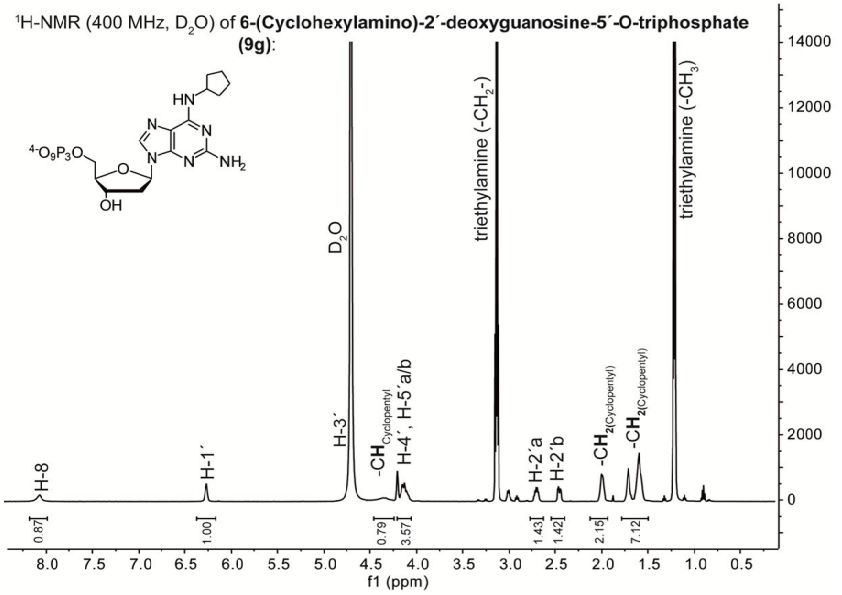

^{31}P-NMR (162 MHz, D_2O) of **6-(Ethylamino)-2´-deoxyguanosine-5´-O-triphosphate (9b)**:
α-P
γ-P
β-P
f1 (ppm)

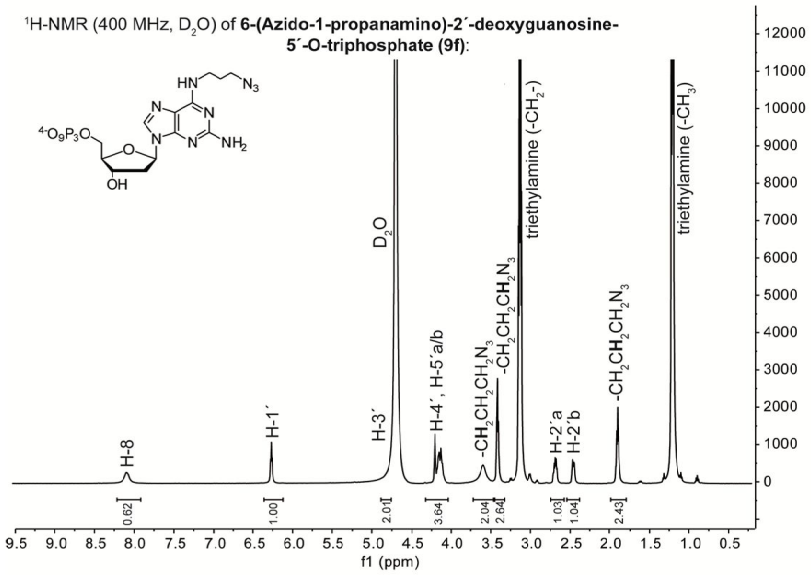

1H-NMR (400 MHz, D2O) of 6-(Hydroxy-1-ethanamino)-2´-deoxyguanosine-5´-O-triphosphate (9e):
H-8
H-1´
D2O
H-3´
H-4´, H-5´a/b
-CH2CH2OH
-CH2CH2OH
triethylamine (-CH2-)
H-2´a
H-2´b
triethylamine (-CH3)
f1 (ppm)

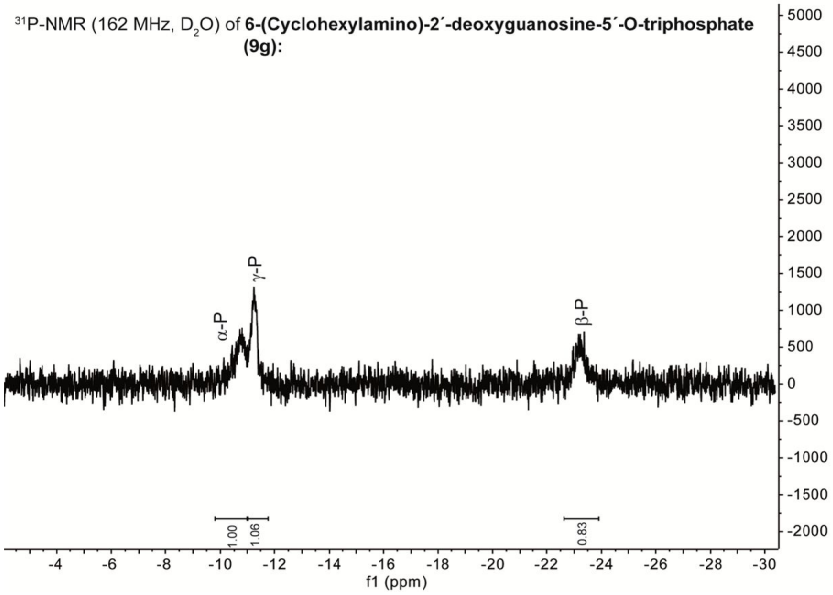

1H-NMR (400 MHz, D2O) of 6-(iso-Propylamino)-2´-deoxyguanosine-5´-O-triphosphate (9d):
H-8
H-1´
D2O
H-3´
-CH(CH3)2
H-4´, H-5´a/b
triethylamine (-CH2-)
H-2´a
H-2´b
acetate
-CH(CH3)2
triethylamine (-CH3)
f1 (ppm)

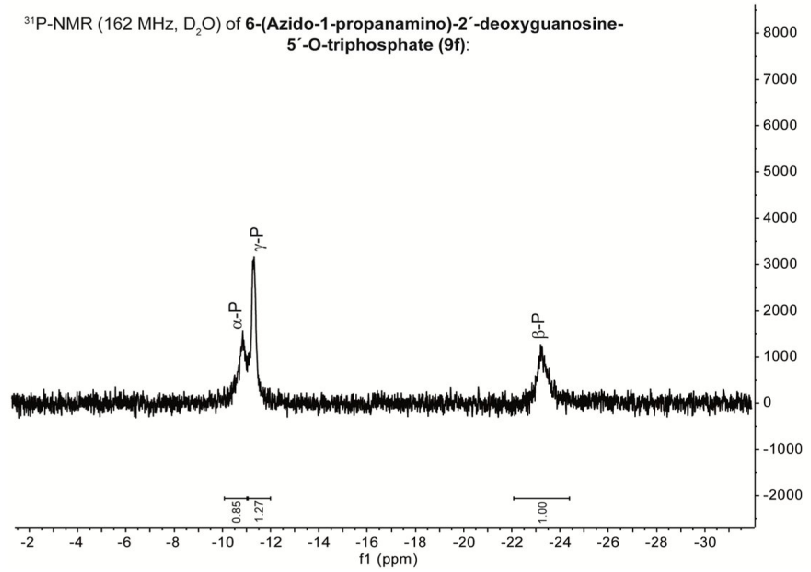

31P-NMR (162 MHz, D2O) of 6-(Hydroxy-1-ethanamino)-2´-deoxyguanosine-5´-O-triphosphate (9e):
α-P
γ-P
β-P
f1 (ppm)

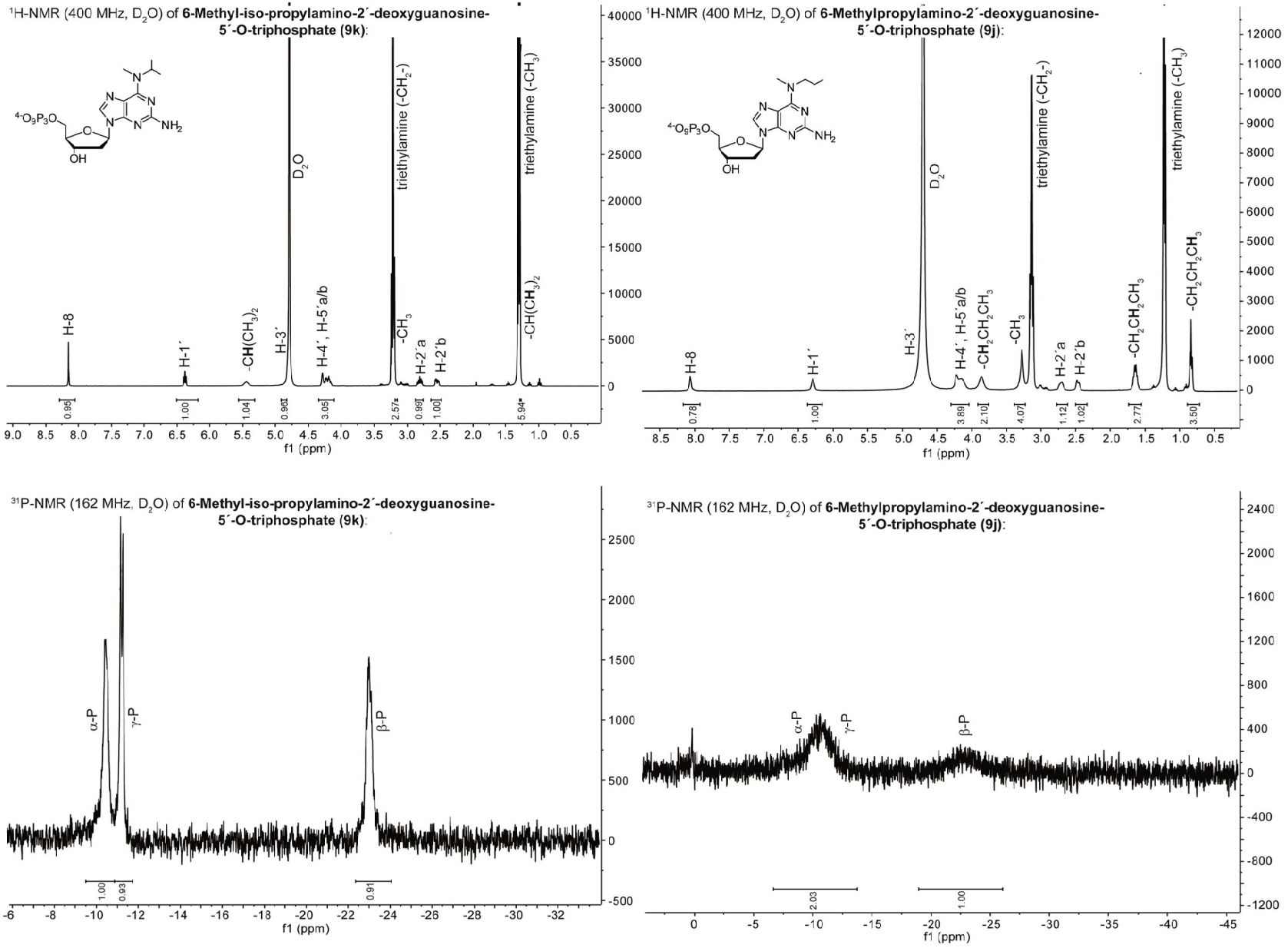

31P-NMR (162 MHz, D2O) of 6-(iso-Propylamino)-2´-deoxyguanosine-5´-O-triphosphate (9d):
α-P
γ-P
β-P
f1 (ppm)

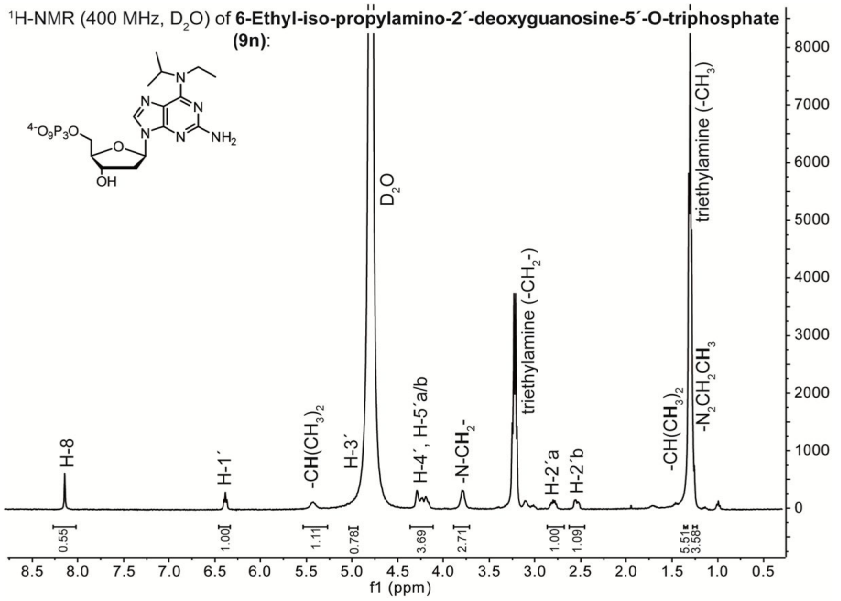

1H-NMR (400 MHz, D2O) of 6-(Cyclohexylamino)-2´-deoxyguanosine-5´-O-triphosphate (9g):
H-8
H-1´
H-3´
D2O
-CHCyclopentyl
H-4´, H-5´a/b
triethylamine (-CH2-)
H-2´a
H-2´b
-CH2(Cyclopentyl)
-CH2(Cyclopentyl)
triethylamine (-CH3)
f1 (ppm)

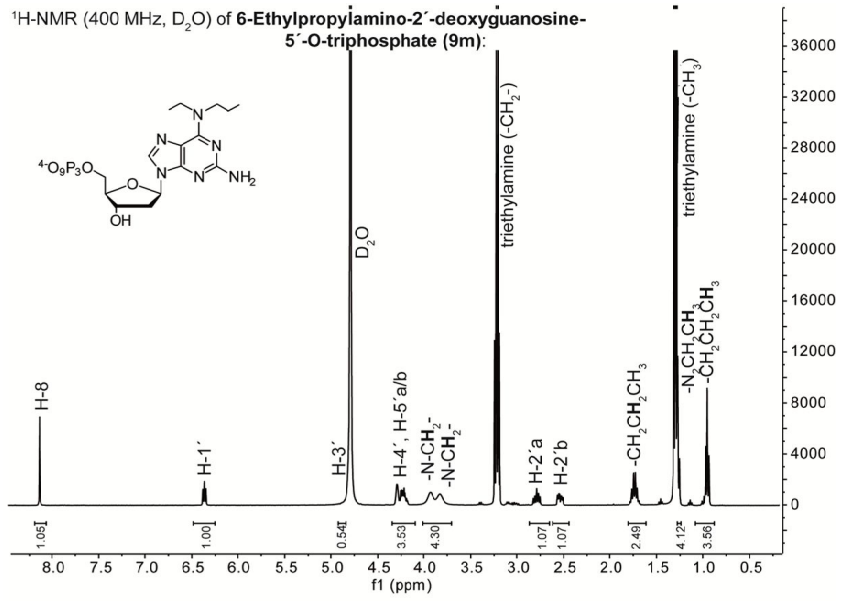

1H-NMR (400 MHz, D2O) of 6-(Azido-1-propanamino)-2´-deoxyguanosine-5´-O-triphosphate (9f):
H-8
H-1´
H-3´
D2O
H-4´, H-5´a/b
-CH2CH2CH2N3
-CH2CH2CH2N3
triethylamine (-CH2-)
H-2´a
H-2´b
-CH2CH2CH2N3
triethylamine (-CH3)
f1 (ppm)

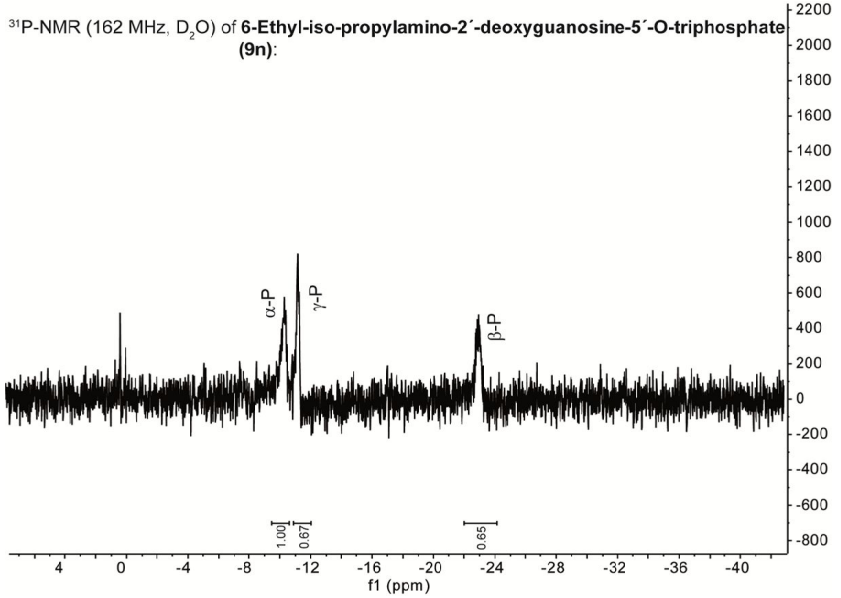

31P-NMR (162 MHz, D2O) of 6-(Cyclohexylamino)-2´-deoxyguanosine-5´-O-triphosphate (9g):
α-P
γ-P
β-P
f1 (ppm)

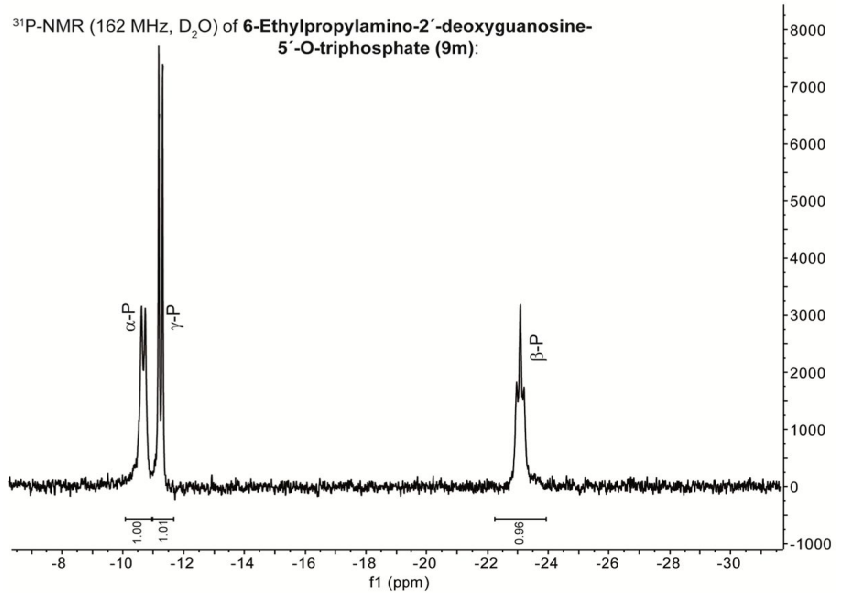

31P-NMR (162 MHz, D2O) of 6-(Azido-1-propanamino)-2´-deoxyguanosine-5´-O-triphosphate (9f):
α-P
γ-P
β-P
f1 (ppm)

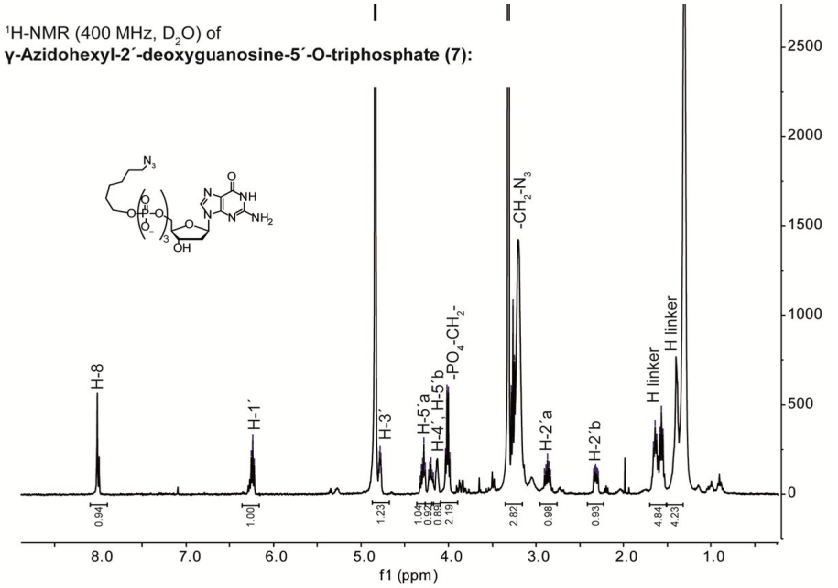

¹H-NMR (400 MHz, D2O) of 6-Methyl-iso-propylamino-2´-deoxyguanosine-5´-O-triphosphate (9k):
H-8
H-1´
-CH(CH3)/2
H-3´
H-4´, H-5´a/b
D2O
triethylamine (-CH2-)
-CH3
H-2´a
H-2´b
triethylamine (-CH3)
-CH(CH3)/2
f1 (ppm)
¹H-NMR (400 MHz, D2O) of 6-Methylpropylamino-2´-deoxyguanosine-5´-O-triphosphate (9j):
H-8
H-1´
H-3´
D2O
H-4´, H-5´a/b
-CH2CH2CH3
-CH3
triethylamine (-CH2-)
H-2´a
H-2´b
-CH2CH2CH3
triethylamine (-CH3)
-CH2CH2CH3
f1 (ppm)
31P-NMR (162 MHz, D2O) of 6-Methyl-iso-propylamino-2´-deoxyguanosine-5´-O-triphosphate (9k):
α-P
γ-P
β-P
f1 (ppm)
31P-NMR (162 MHz, D2O) of 6-Methylpropylamino-2´-deoxyguanosine-5´-O-triphosphate (9j):
α-P
γ-P
β-P
f1 (ppm)

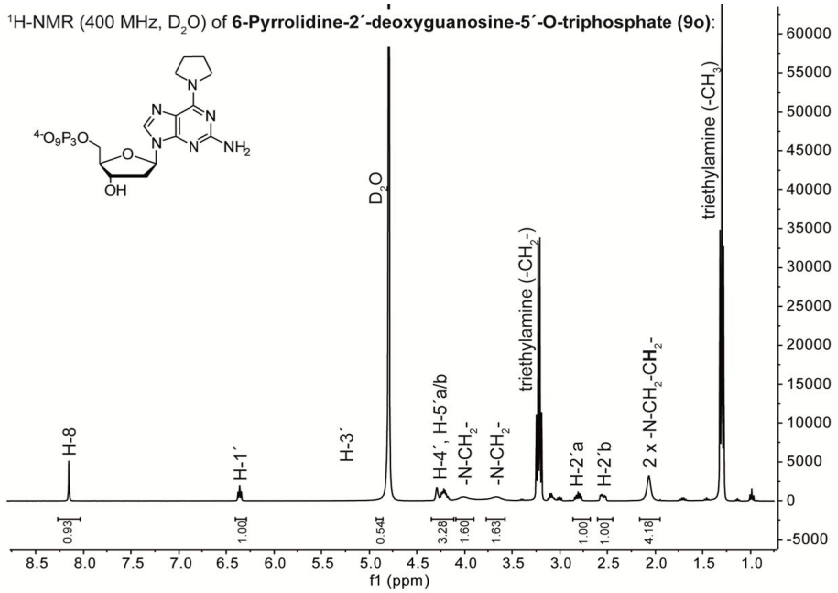

^{1}H-NMR (400 MHz, D_2O) of **6-Ethyl-iso-propylamino-2´-deoxyguanosine-5´-O-triphosphate (9n)**:
H-8
H-1´
-CH(CH3)2
H-3´
D2O
H-4´, H-5´a/b
-N-CH2-
triethylamine (-CH2-)
H-2´a
H-2´b
-CH(CH3)2
-N2CH2CH3
triethylamine (-CH3)
f1 (ppm)

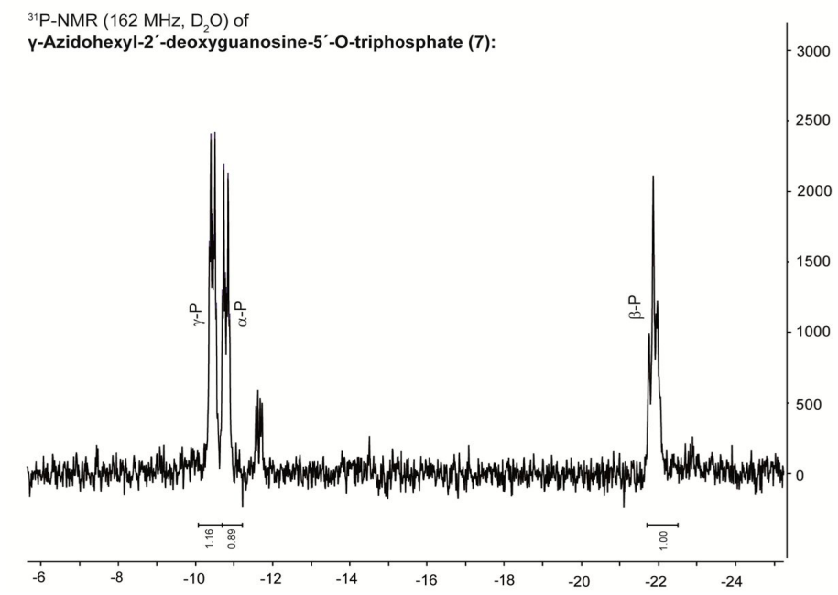

^{1}H-NMR (400 MHz, D_2O) of **6-Ethylpropylamino-2´-deoxyguanosine-5´-O-triphosphate (9m)**:
H-8
H-1´
H-3´
D2O
H-4´, H-5´a/b
-N-CH2-
-N-CH2-
triethylamine (-CH2-)
H-2´a
H-2´b
-CH2CH2CH3
-N2CH2CH3
-CH2CH2CH3
triethylamine (-CH3)
f1 (ppm)

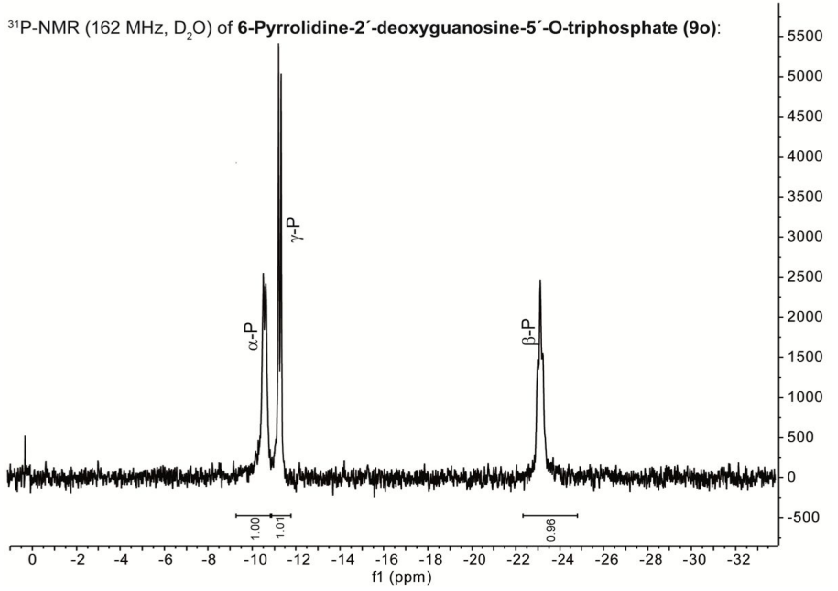

^{31}P-NMR (162 MHz, D_2O) of **6-Ethyl-iso-propylamino-2´-deoxyguanosine-5´-O-triphosphate (9n)**:
α-P
γ-P
β-P
f1 (ppm)


^{31}P-NMR (162 MHz, D_2O) of **6-Ethylpropylamino-2´-deoxyguanosine-5´-O-triphosphate (9m)**:
α-P
γ-P
β-P
f1 (ppm)

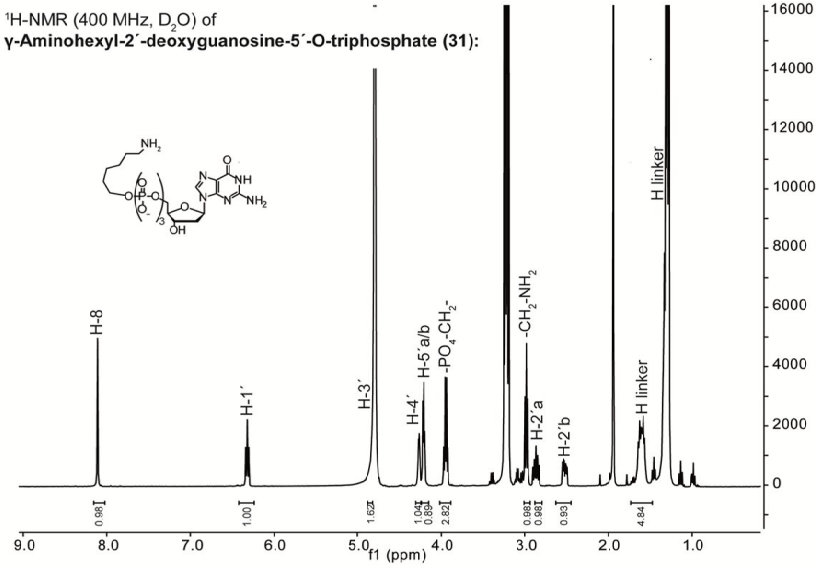

^{1}H-NMR (400 MHz, D_2O) of
γ-Azidohexyl-2´-deoxyguanosine-5´-O-triphosphate (7):
H-8
H-1´
H-3´
H-5´a
H-4´, H-5´b
-PO$_4$-CH$_2$-
-CH$_2$-N$_3$
H-2´a
H-2´b
H linker
H linker
f1 (ppm)


^{1}H-NMR (400 MHz, D_2O) of **6-Pyrrolidine-2´-deoxyguanosine-5´-O-triphosphate (9o)**:
H-8
H-1´
H-3´
D$_2$O
H-4´, H-5´a/b
-N-CH$_2$-
-N-CH$_2$-
triethylamine (-CH$_2$-)
H-2´a
H-2´b
2 x -N-CH$_2$-CH$_2$-
triethylamine (-CH$_3$)
f1 (ppm)

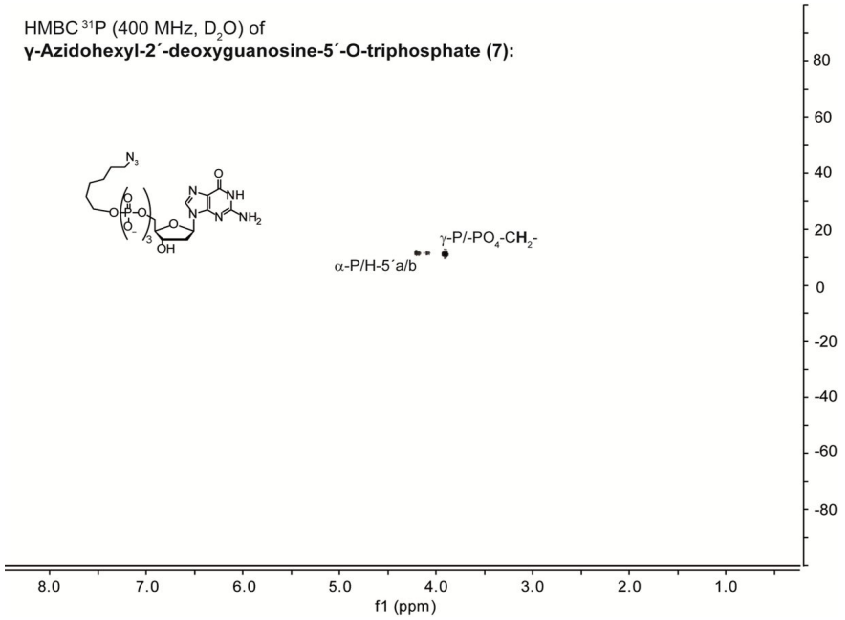

^{31}P-NMR (162 MHz, D_2O) of
γ-Azidohexyl-2´-deoxyguanosine-5´-O-triphosphate (7):
γ-P
α-P
β-P


^{31}P-NMR (162 MHz, D_2O) of **6-Pyrrolidine-2´-deoxyguanosine-5´-O-triphosphate (9o)**:
α-P
γ-P
β-P
f1 (ppm)

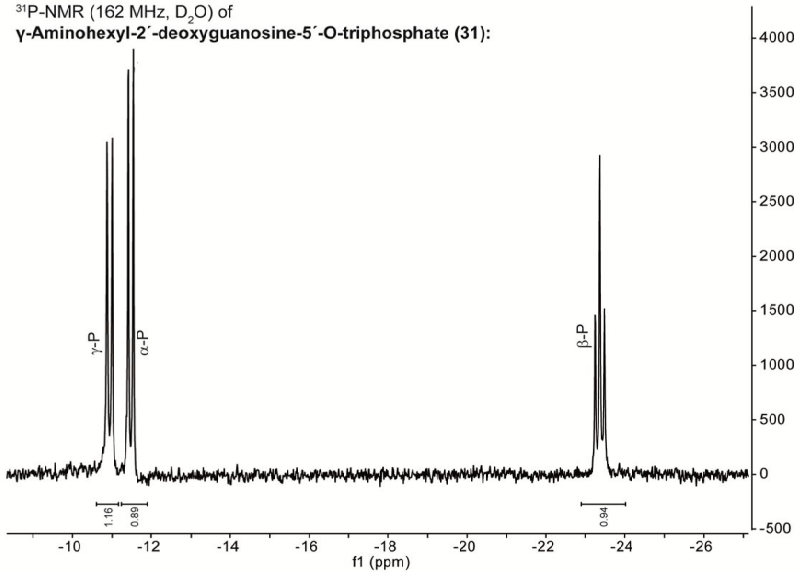
^{1}H-NMR (400 MHz, D_2O) of
γ-Aminohexyl-2´-deoxyguanosine-5´-O-triphosphate (31):



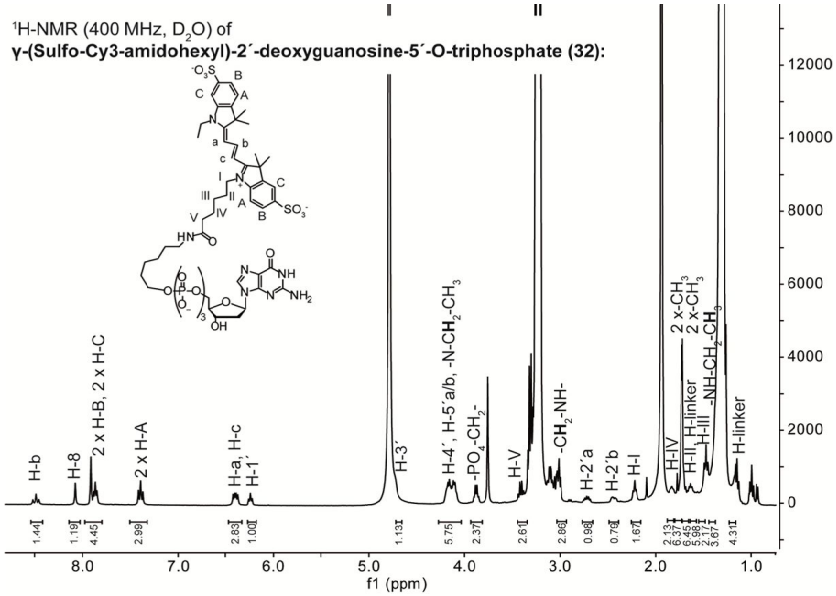
HMBC ^{31}P (400 MHz, D_2O) of
γ-Azidohexyl-2´-deoxyguanosine-5´-O-triphosphate (7):



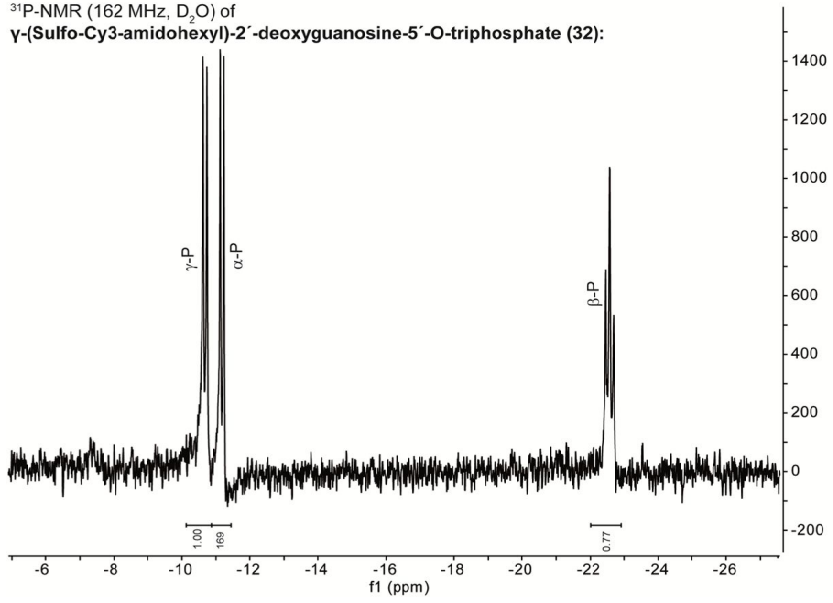
^{31}P-NMR (162 MHz, D_2O) of
γ-Aminohexyl-2´-deoxyguanosine-5´-O-triphosphate (31):



^{1}H-NMR (400 MHz, D_2O) of
γ-(Sulfo-Cy3-amidohexyl)-2´-deoxyguanosine-5´-O-triphosphate (32):

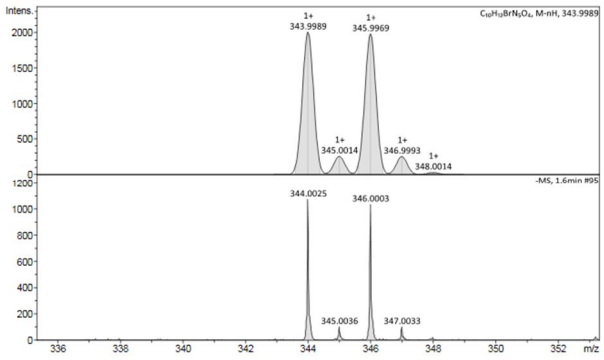


^{31}P-NMR (162 MHz, D_2O) of
γ-(Sulfo-Cy3-amidohexyl)-2´-deoxyguanosine-5´-O-triphosphate (32):

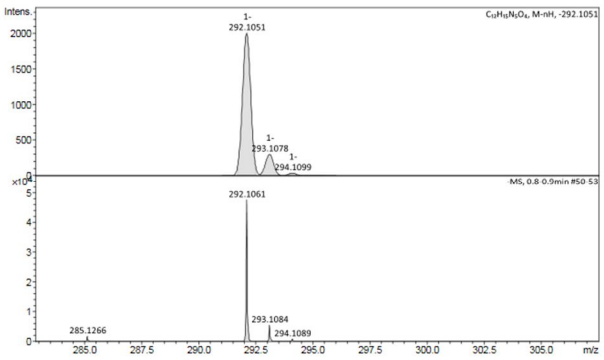


8.2 HR-MS

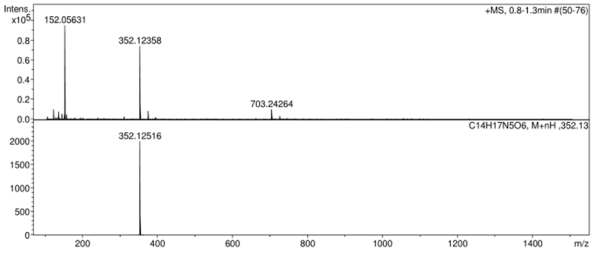
8-Bromo-2´-deoxyguanosine 14



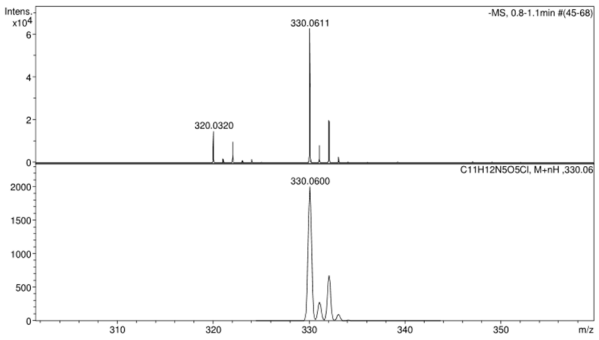
8-Vinyl-2´-deoxyguanosine 15



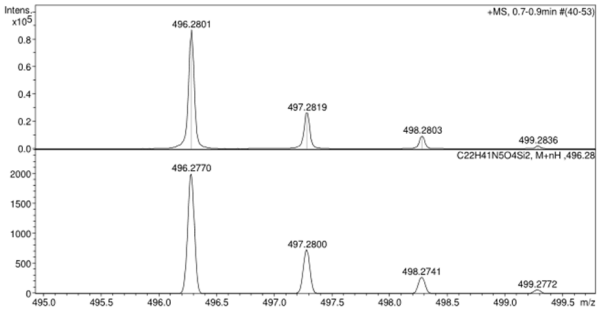
3´, 5´-Di-*O*-acetyl-2´-deoxyguanosine 11



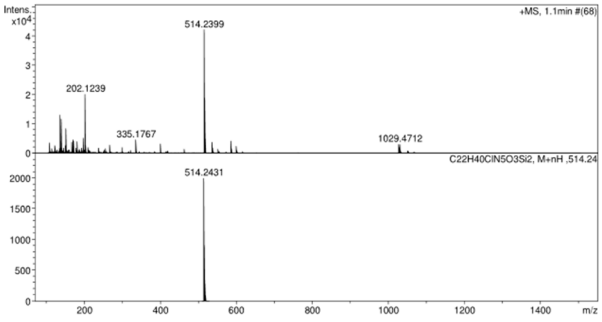
3´, 5´-Di-*O*-acetyl-6-deoxy-6-chloro-2´-deoxyguanosine 12



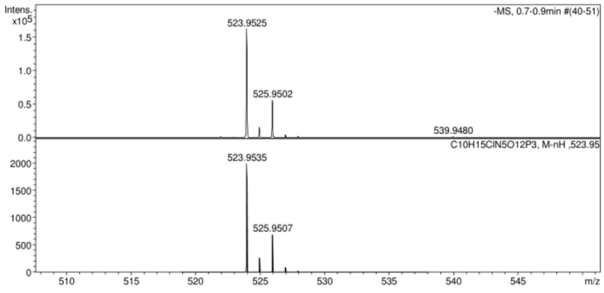
3´,5´-Di-*O*-*tert*-butyldimethylsilyl-2´-deoxyguanosine 17



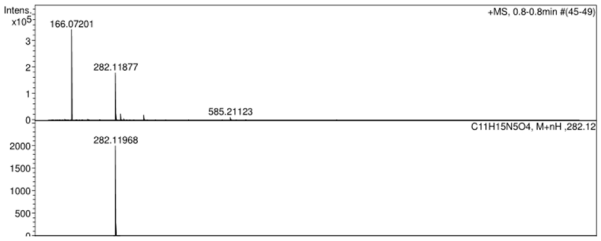
3´,5´-Di-*O*-*tert*-butyldimethylsilyl-6-chloro-2´-deoxyguanosine 18



6-Chloro-2´-deoxyguanosine 19

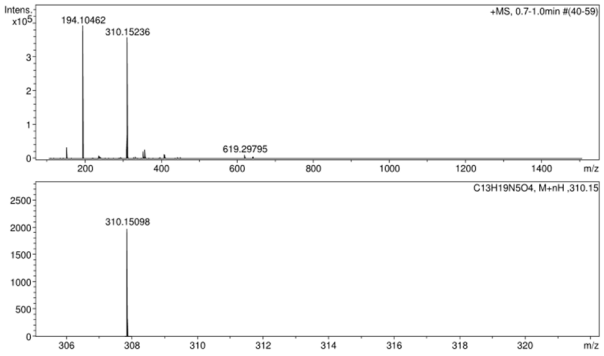


O^6-Methyl-2´-deoxyguanosine 13a

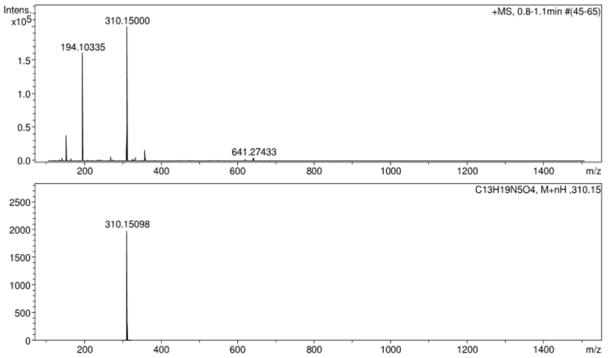


O6-Ethyl-2´-deoxyguanosine 13b

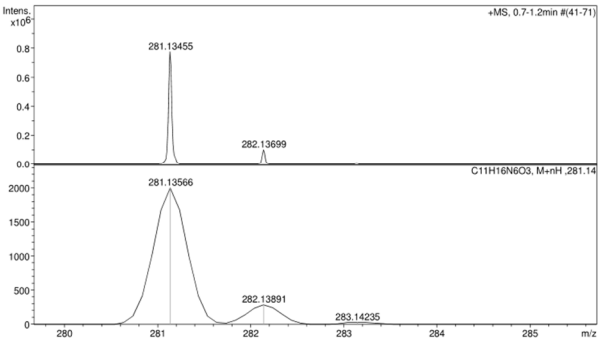
O^6-Propyl-2´-deoxyguanosine 13c



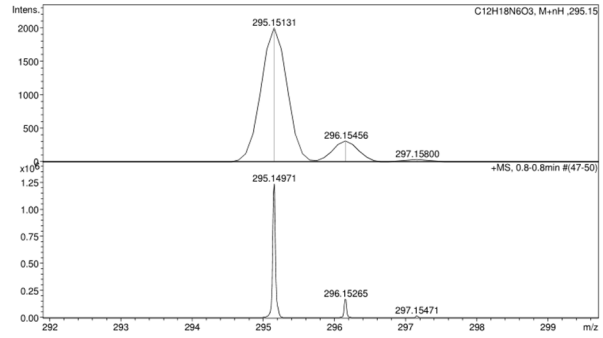
O^6-*Iso*-propyl-2´-deoxyguanosine 13d



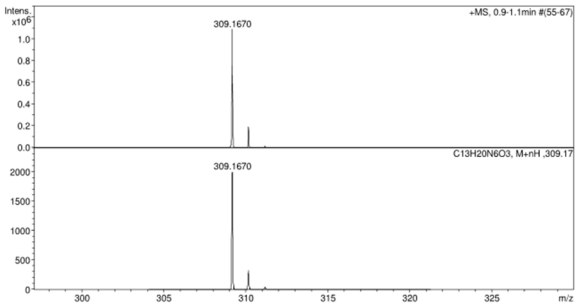
6-Methylamino-2´deoxyguanosine 16a



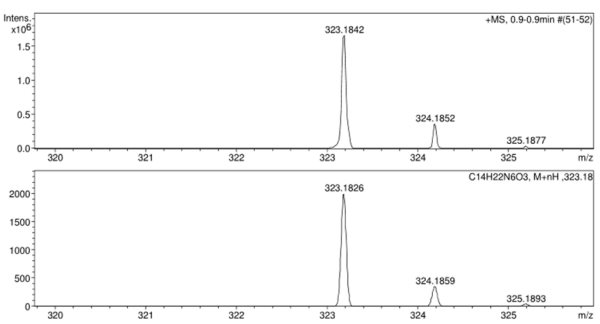
6-Dimethylamino-2´-deoxyguanosine 16h



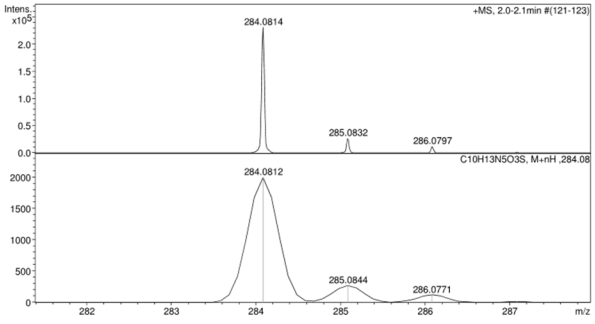
6-Methylethylamino-2´-deoxyguanosine 16i



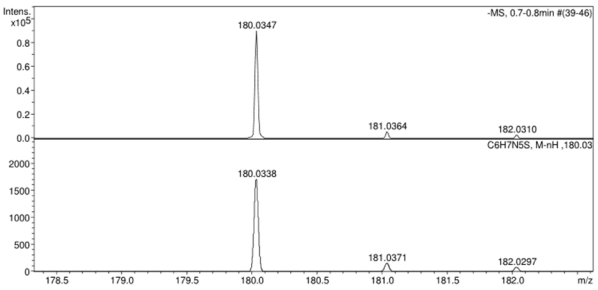
6-Diethylamino-2´-deoxyguanosine 16l



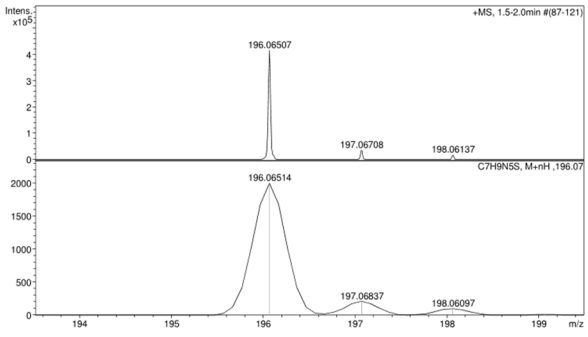
6-Thio-2´-deoxyguanosine



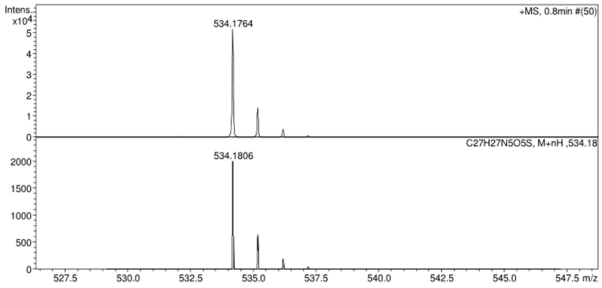
2-Amino-6-methylthiopurin 23a



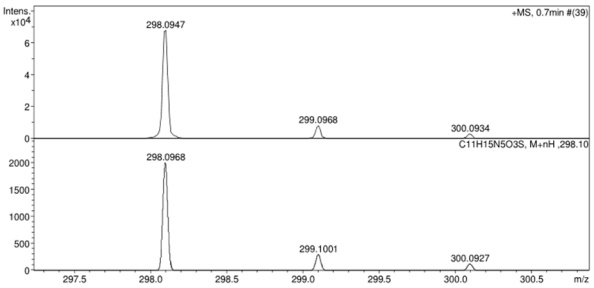
2-Amino-6-ethylthiopurin 23b



3´,5´-Di-*O*-toluoyl-6-methylthio-2´-deoxyguanosine 24a

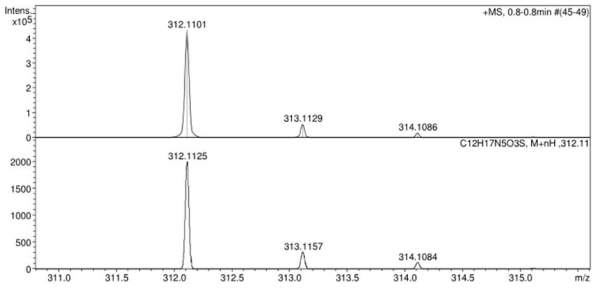


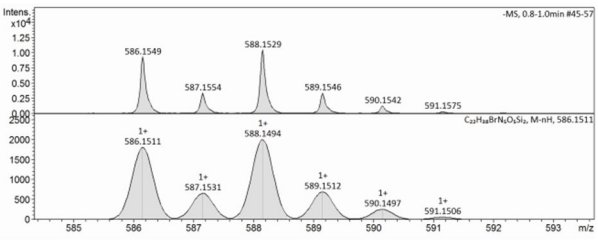
6-Methylthio-2´-deoxyguanosine 21a



6-Ethylthio-2´-deoxyguanosine 21b

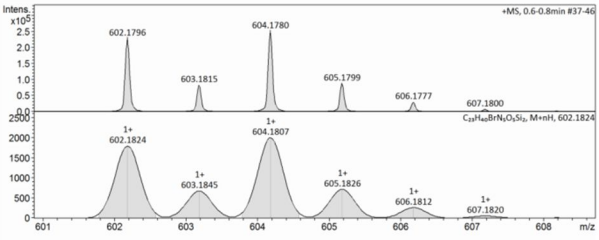
8-Bromo-3´,5´-O-(1,1,3,3-tetrakis(-*iso*-propyl)-1,3-disiloxanediyl)-2´-deoxyguanosine 27

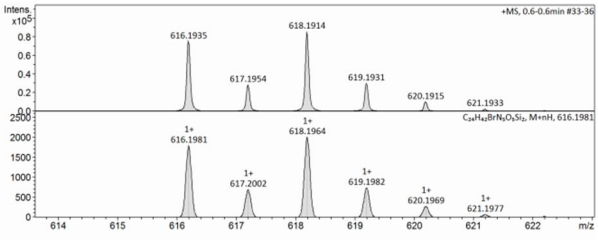




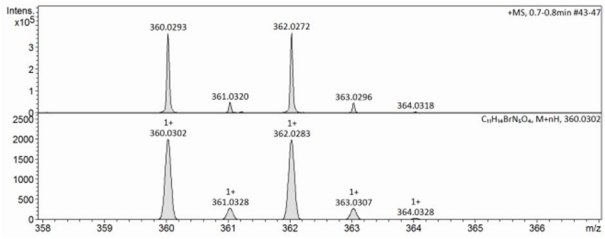
8-Bromo-3´,5´-*O*-(1,1,3,3-tetrakis(*iso*-propyl)-1,3-disiloxanediyl)-O^6-methyl-2´-deoxyguanosine 28a

8-Bromo-3´,5´-*O*-(1,1,3,3-tetrakis(*iso*-propyl)-1,3-disiloxanediyl)-O^6-ethyl-2´-deoxyguanosine 28b

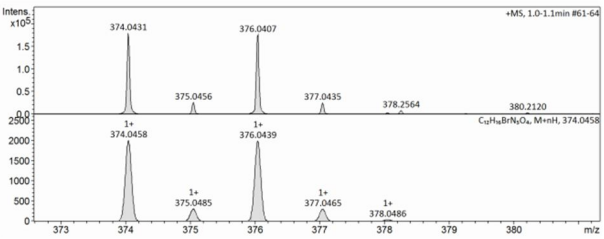




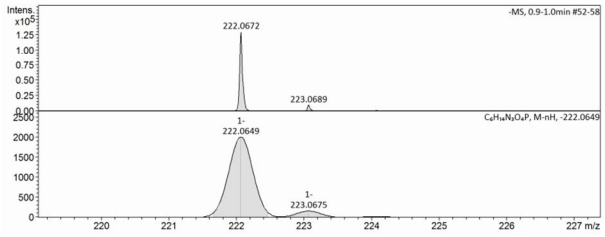
8-Bromo-O^6-methyl-2´-deoxyguanosine 29a



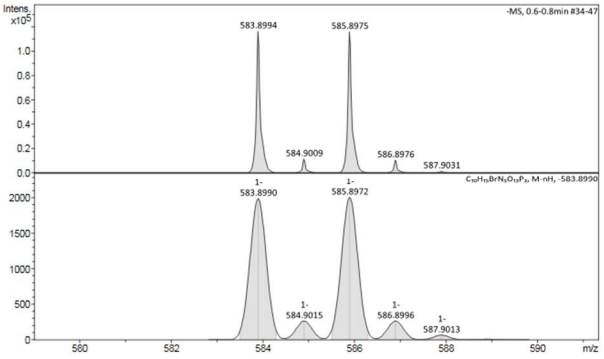
8-Bromo-O^6-ethyl-2´-deoxyguanosine 29b



6-azido-hexylphosphate

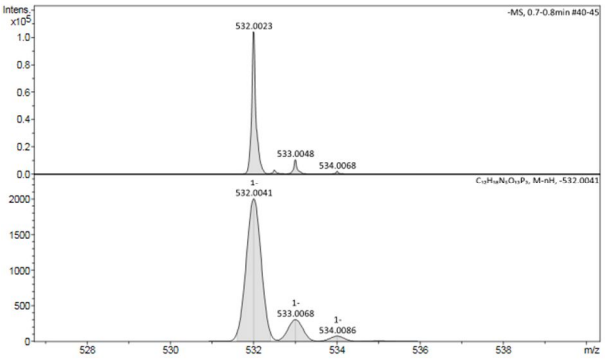


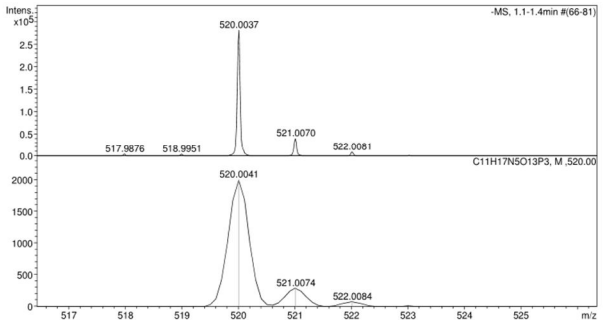
8-Bromo-2´-deoxyguanosine-5´-*O*-triphosphate 2



8-Vinyl-2´-deoxyguanosine-5´-*O*-triphosphate 3

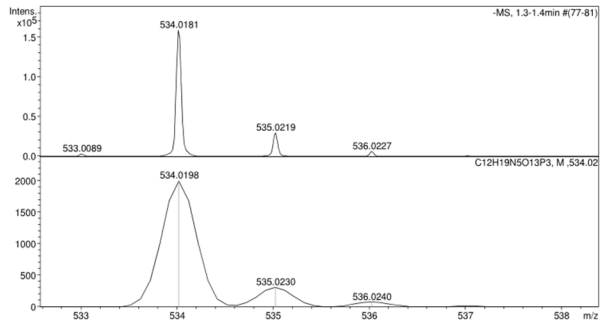
O^6-Methyl-2´-deoxyguanosine-5´-*O*-triphosphate 1a

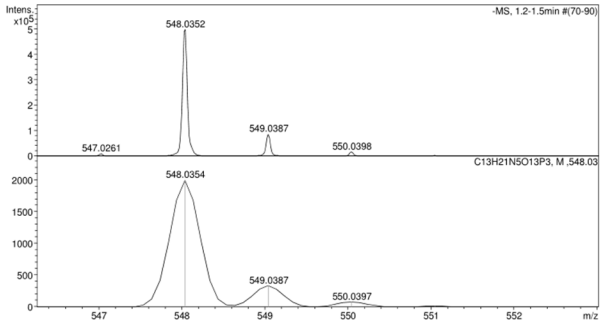




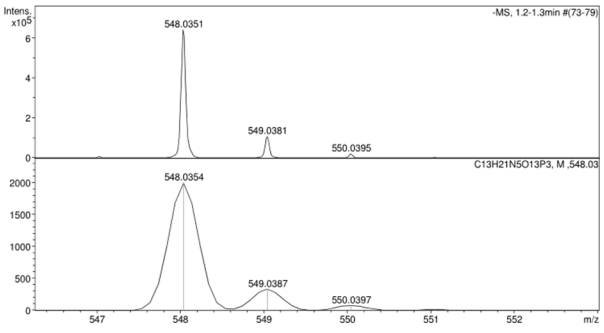
O^6-Ethyl-2´-deoxyguanosine-5´-*O*-triphosphate 1b

O^6-Propyl-2´deoxyguanosine-5´-*O*-triphosphate 1c

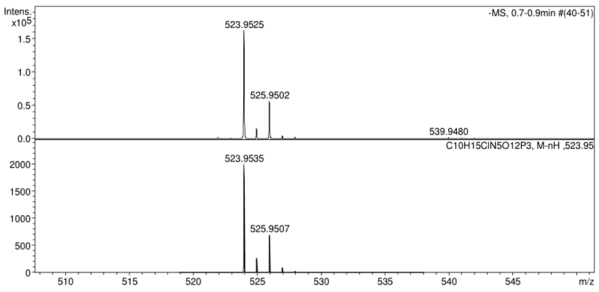




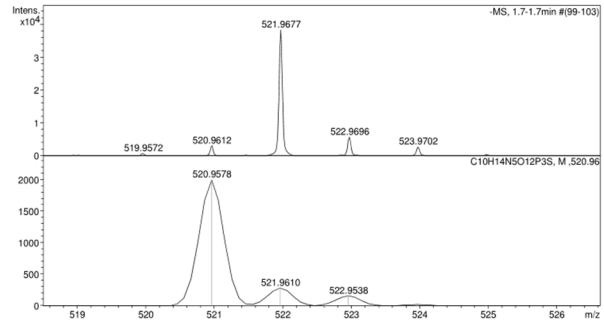
O^6-*Iso*-propyl-2´-deoxyguanosine-5´-*O*-triphosphate 1d



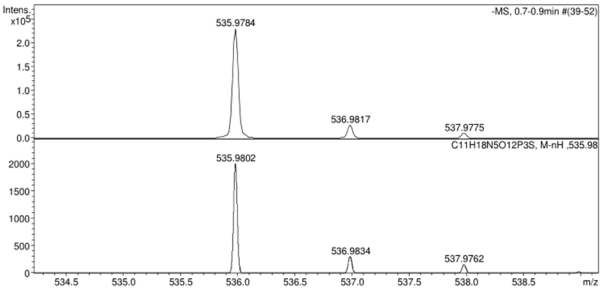
6-Chloro-2´-deoxyguanosine-5´-*O*-triphosphate 20



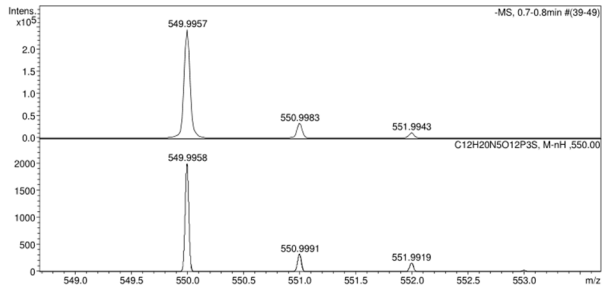
6-Thio-2´-deoxyguanosine-5´-*O*-triphosphate 10



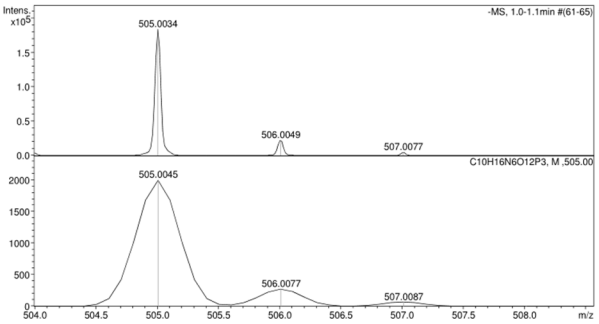
6-Thiomethyl-2´-deoxyguanosine-5´-*O*-triphosphate 10a



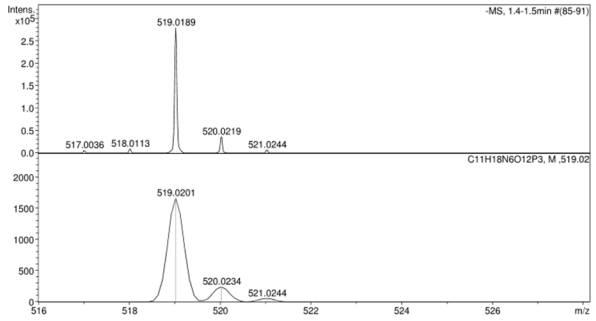
6-Thioethyl-2´-deoxyguanosine-5´-*O*-triphosphate 10b



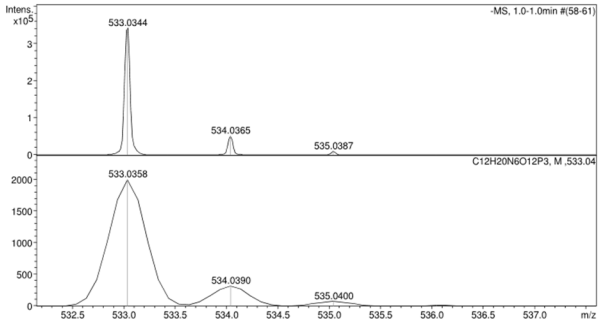
6-Amino-2´-deoxyguanosine-5´-*O*-triphosphate 9



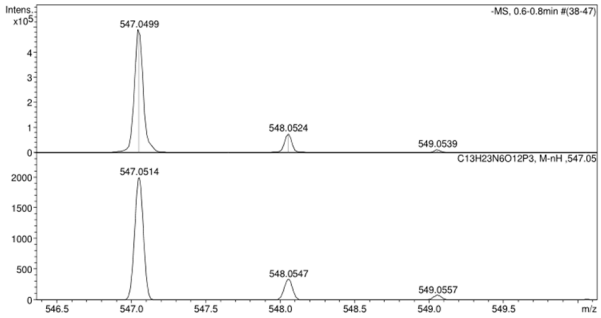
6-Methylamino-2´-deoxyguanosine-5´-*O*-triphosphate 9a



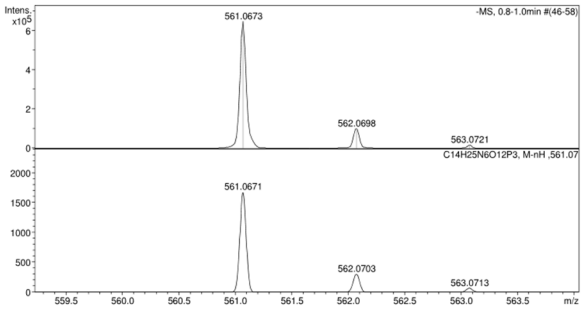
6-Dimethylamino-2´-deoxyguanosine-5´-*O*-triphosphate 9h



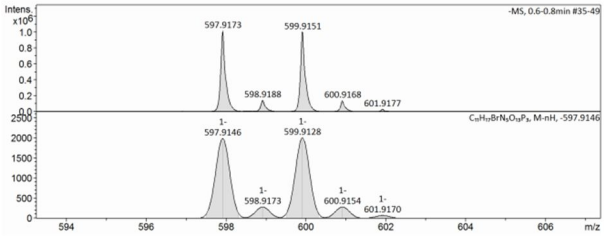
6-Methylethylamino-2´-deoxyguanosine-5´-*O*-triphosphate 9i



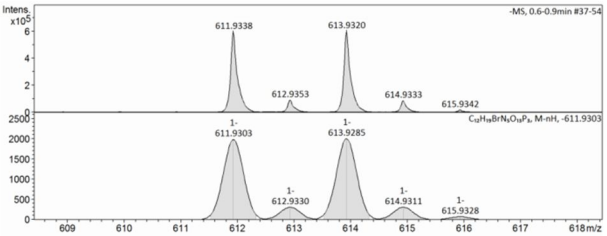
6-Diethylamino-2´-deoxyguanosine-5´-*O*-triphosphate 9l



8-Bromo-*O*^{6}-methyl-2´-deoxyguanosine-5´-*O*-triphosphate 30a

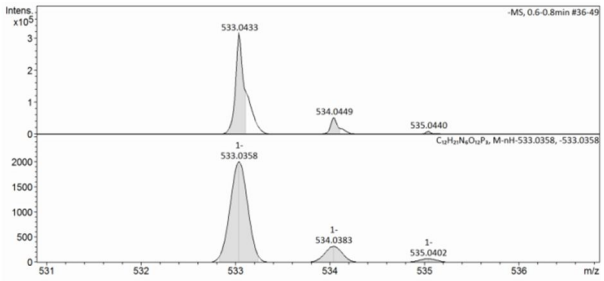


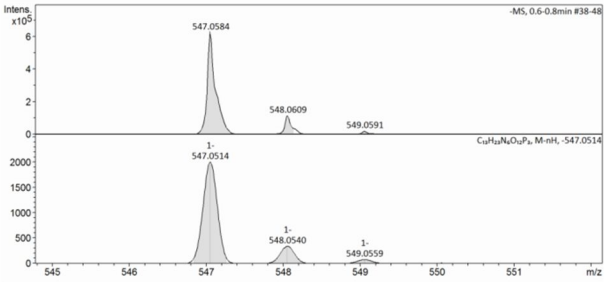
8-Bromo-*O*^{6}-ethyl-2´-deoxyguanosine-5´-*O*-triphosphate 30b



6-Ethylamino-2´-deoxyguanosine-5´-*O*-triphosphate 9b

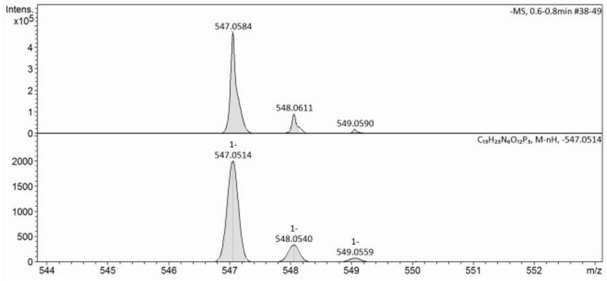
6-Propylamino-2´-deoxyguanosine-5´-*O*-triphosphate 9c

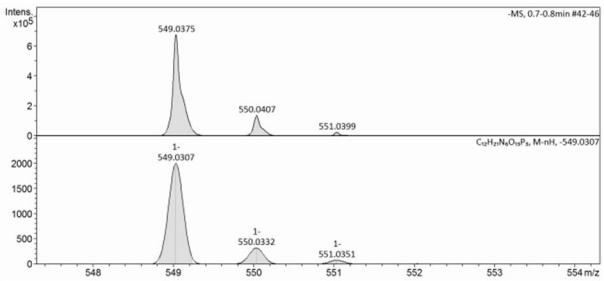




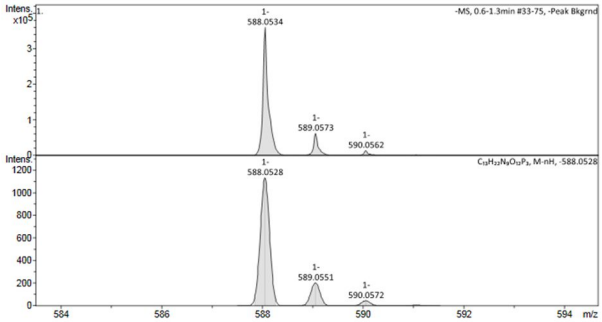
6-*Iso*-propylamino-2´-deoxyguanosine-5´-*O*-triphosphate 9d

6-Hydroxy-1-ethanamino-2´-deoxyguanosine-5´-*O*-triphosphate 9e

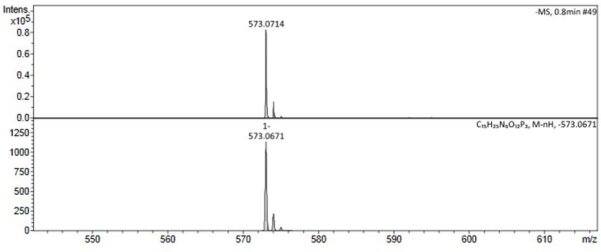




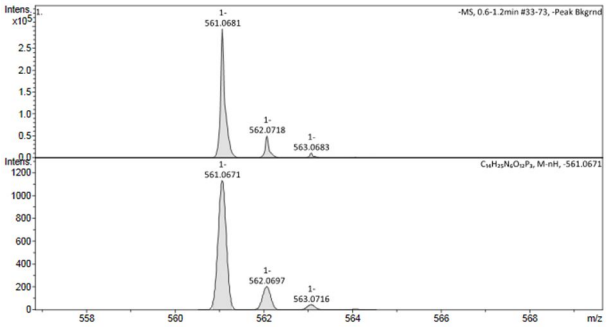
6-Azido-1-propanamino-2´-deoxy-guanosine-5´-*O*-triphosphate 9f



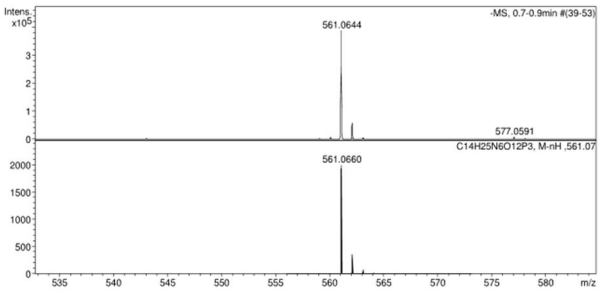
6-Cyclohexylamino-2´-deoxy-guanosine-5´-*O*-triphosphate 9g



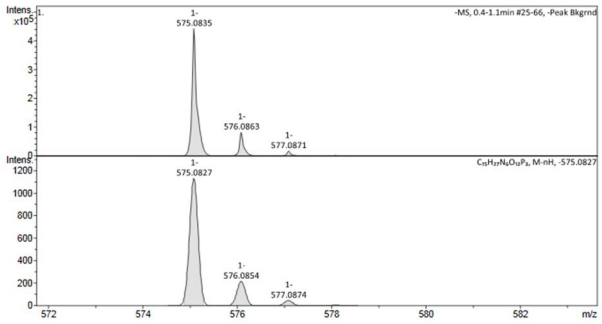
6-Methylpropylamino-2´-deoxy-guanosine-5´-*O*-triphosphate 9j



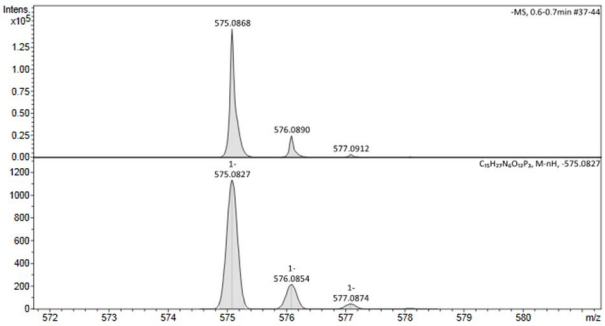
6-Methyl-*iso*-propylamino-2´-deoxy-guanosine-5´-*O*-triphosphate 9k



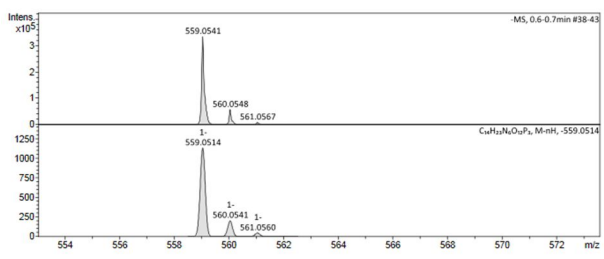
6-Ethylpropylamino-2´-deoxy-guanosine-5´-*O*-triphosphate 9m



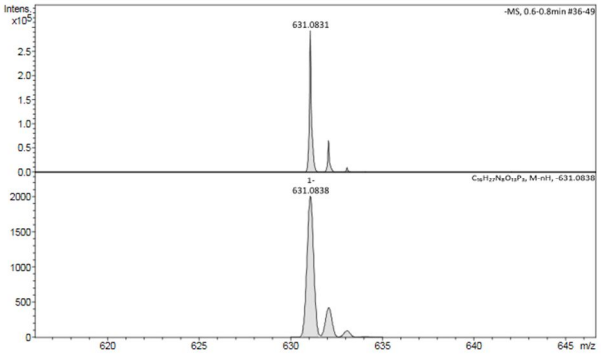
6-Ethyl-*iso*-propylamino-2´-deoxy-guanosine-5´-*O*-triphosphate 9n



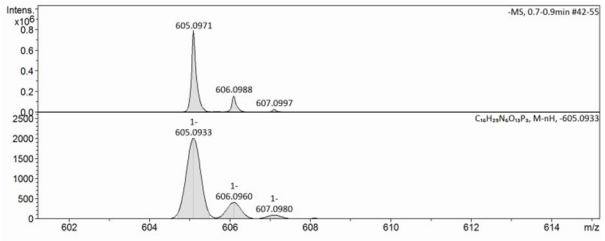
6-Pyrrolidine-2´-deoxyguanosine-5´-*O*-triphosphate 9o



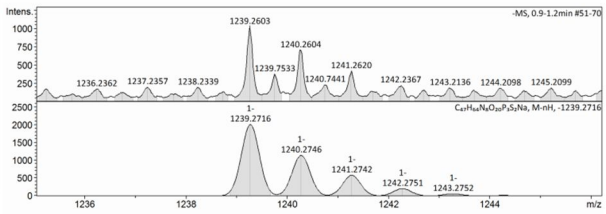
γ-Azidohexyl-2´-deoxyguanosine-5´-*O*-triphosphate 7



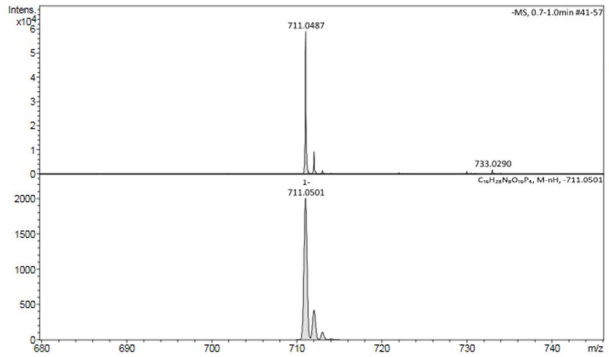
γ-Aminohexyl-2´-deoxyguanosine-5´-*O*-triphosphate 31



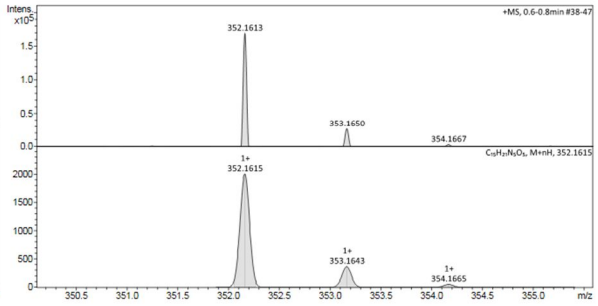
γ-(Sulfo-Cy3-amidohexyl)-2´-deoxy-guanosine-5´-*O*-triphosphate 32



δ-Azidohexyl-2´-deoxyguanosine-5´-*O*-triphosphate 8

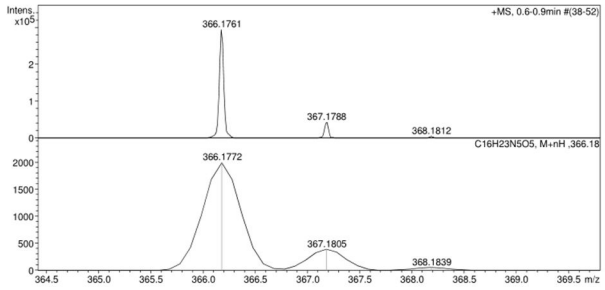


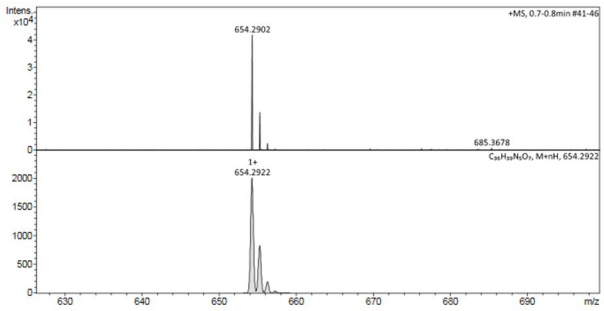
***N*^{2}-*iso*-butyryl-*O*^{6}-methyl-2´deoxy-guanosine 25a**



***N*^{2}-*iso*-butyryl-*O*^{6}-ethyl-2´deoxy-guanosine 25b**

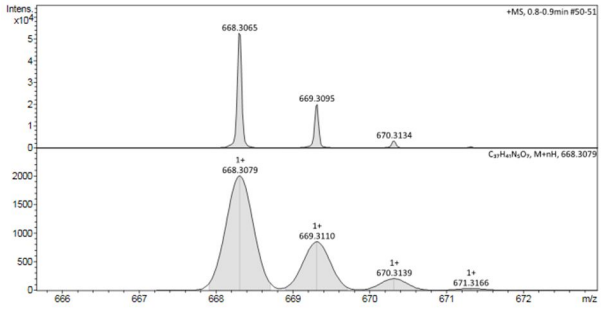
5´-*O*-(4,4´,-dimethoxytrityl)-*N*^{2}-*iso*-butyryl- *O*^{6}-methyl-2´deoxyguanosine 26a

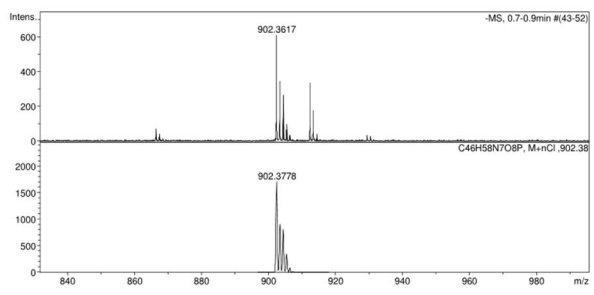




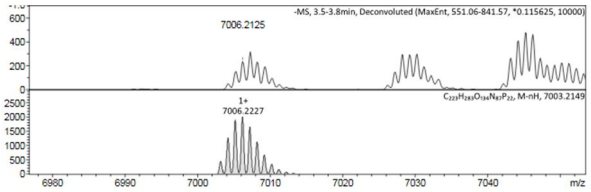
5´-*O*-(4,4´,-dimethoxytrityl)-*N*^{2}-*iso*-butyryl-*O*^{6}-ethyl-2´deoxyguanosine 26b

5´-*O*-(4,4´,-dimethoxytrityl)-*N*^{2}-*iso*-butyryl- *O*^{6}-ethyl-2´deoxyguanosine - 3´- (2-cyanoethyl-di-iso-propyl-phosphoramidite

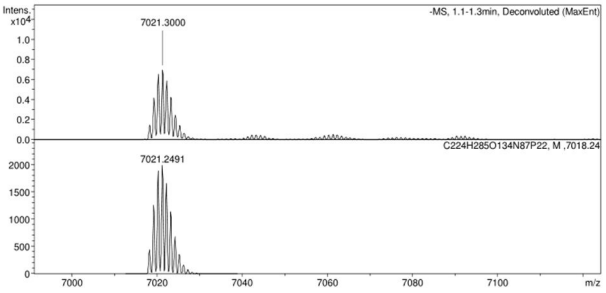




Sept9 - Primer O^6-methyl-G

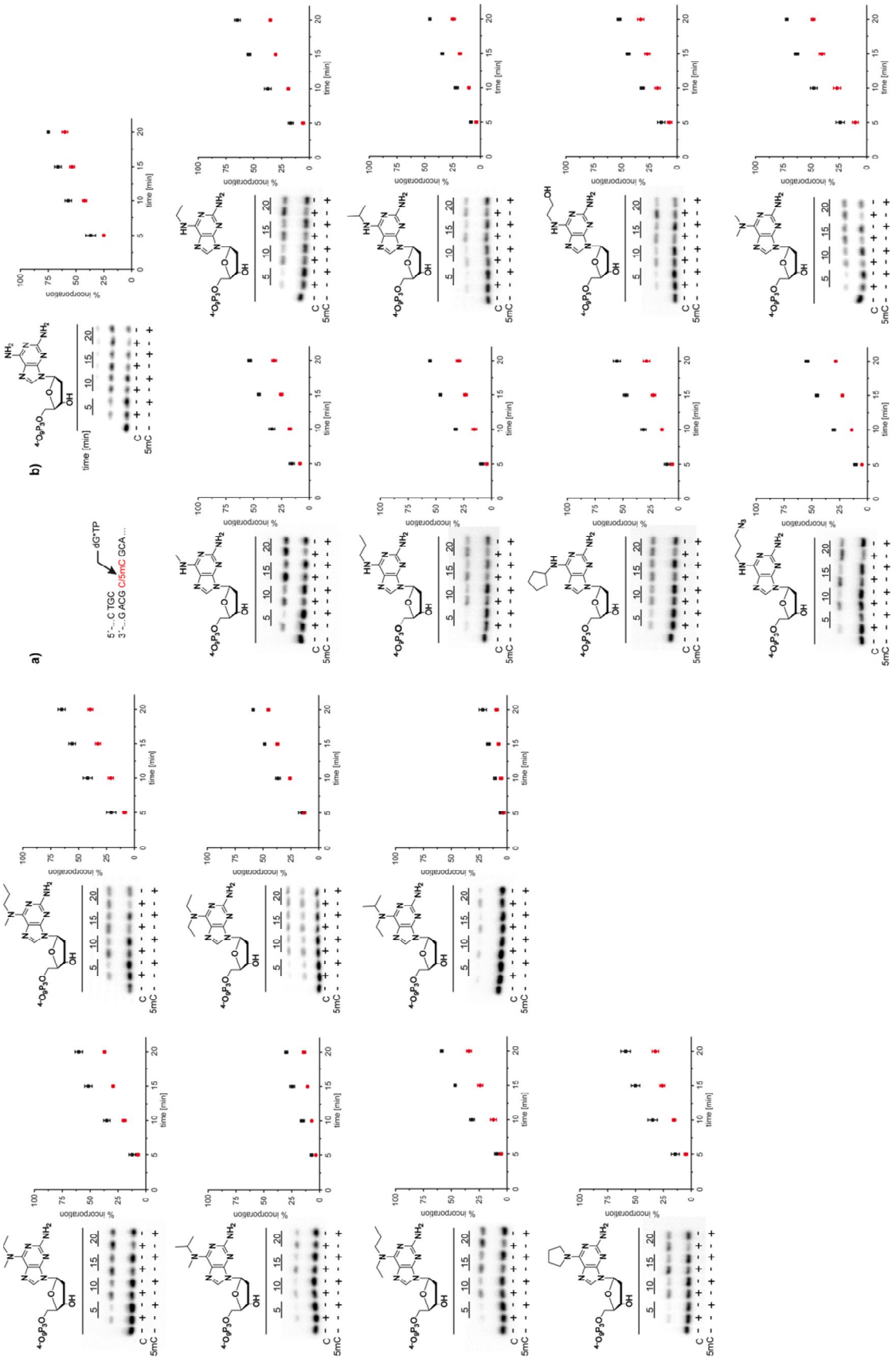


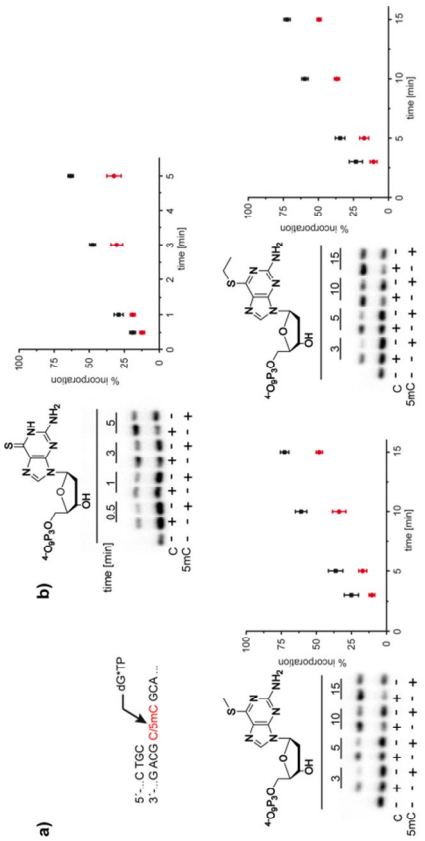
Sept9 - Primer O^6-ethyl-G



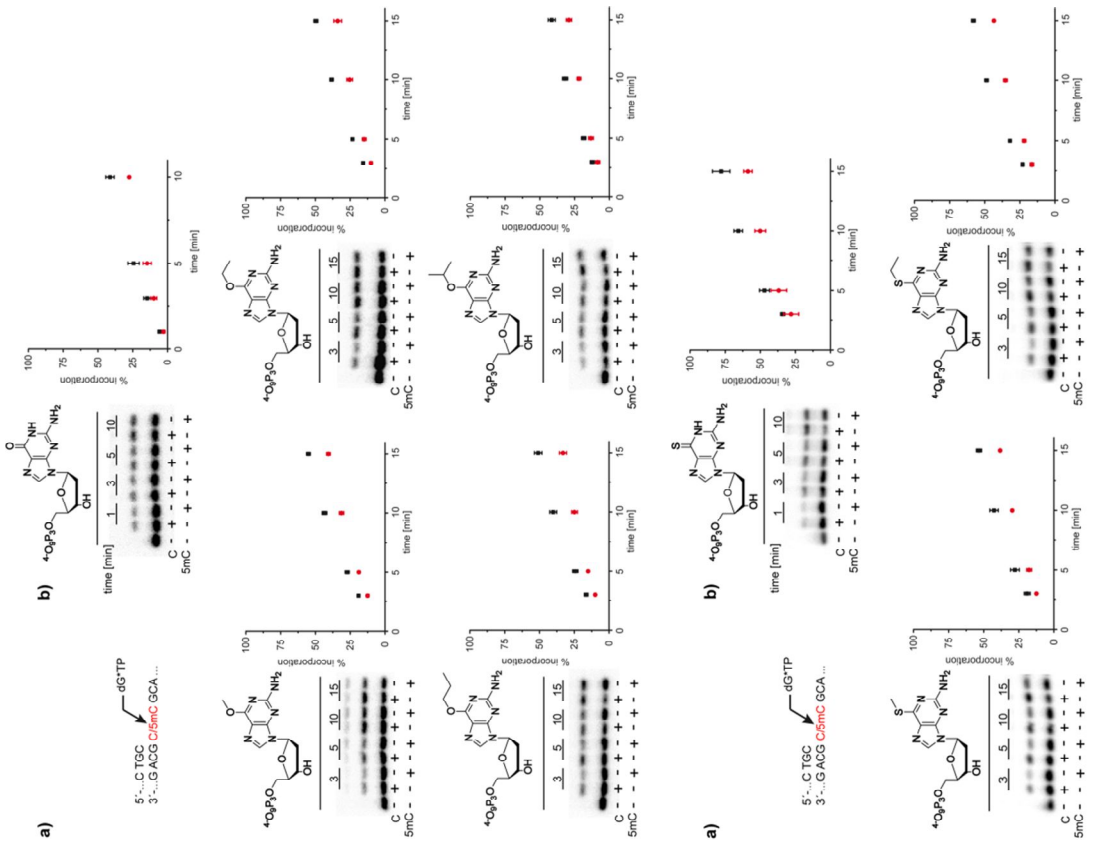
8.3 Primer Extension Experiments:

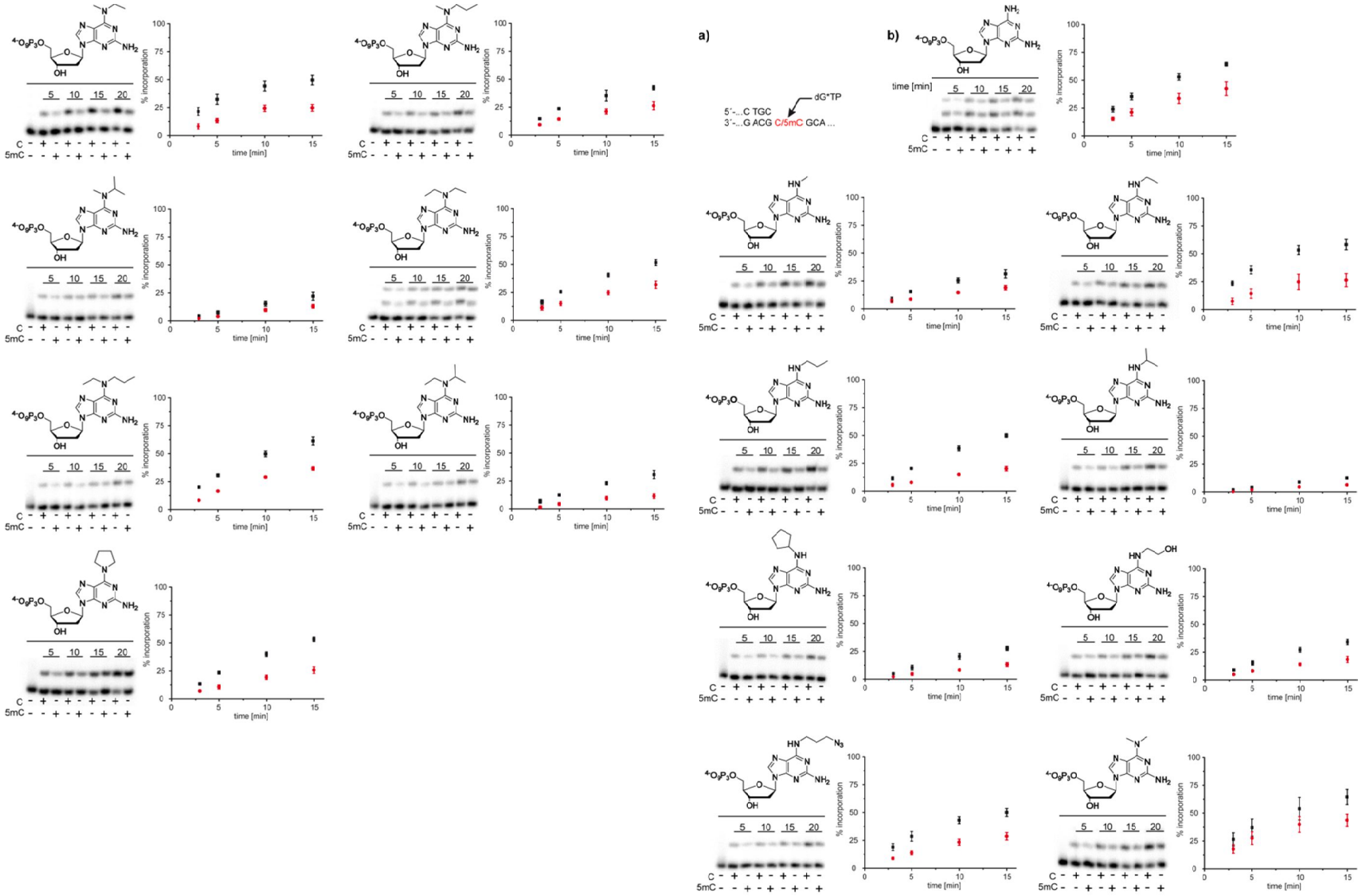
Screening of modified nucleotides using KOD exo$^-$:

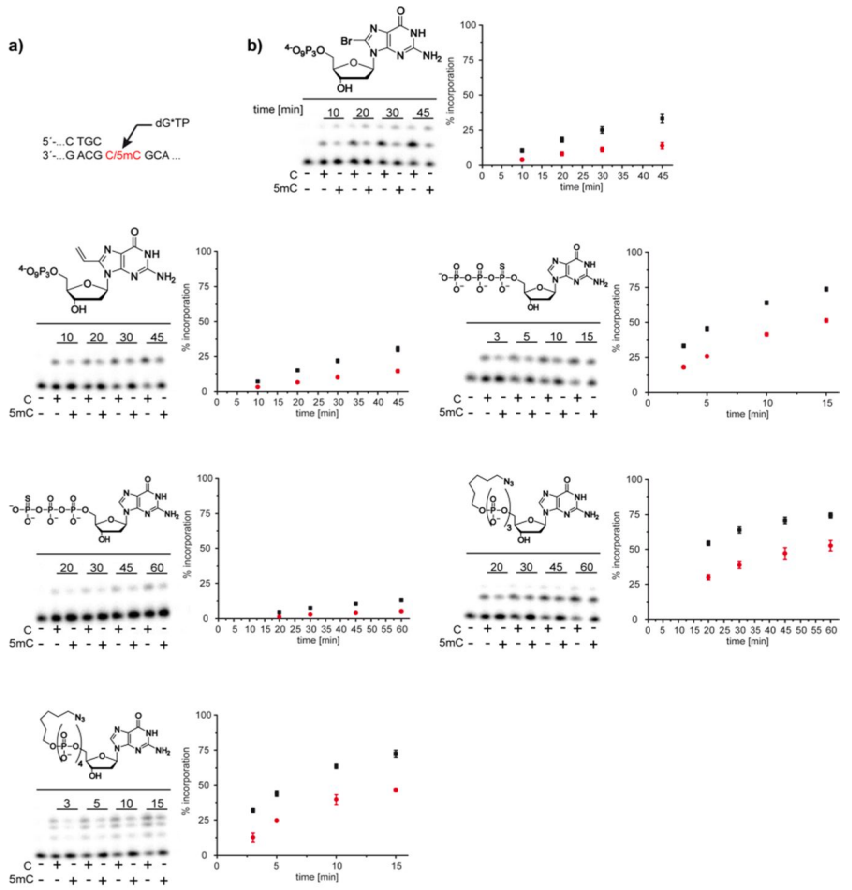




Screening of modified nucleotides using 9° North exo⁻ :

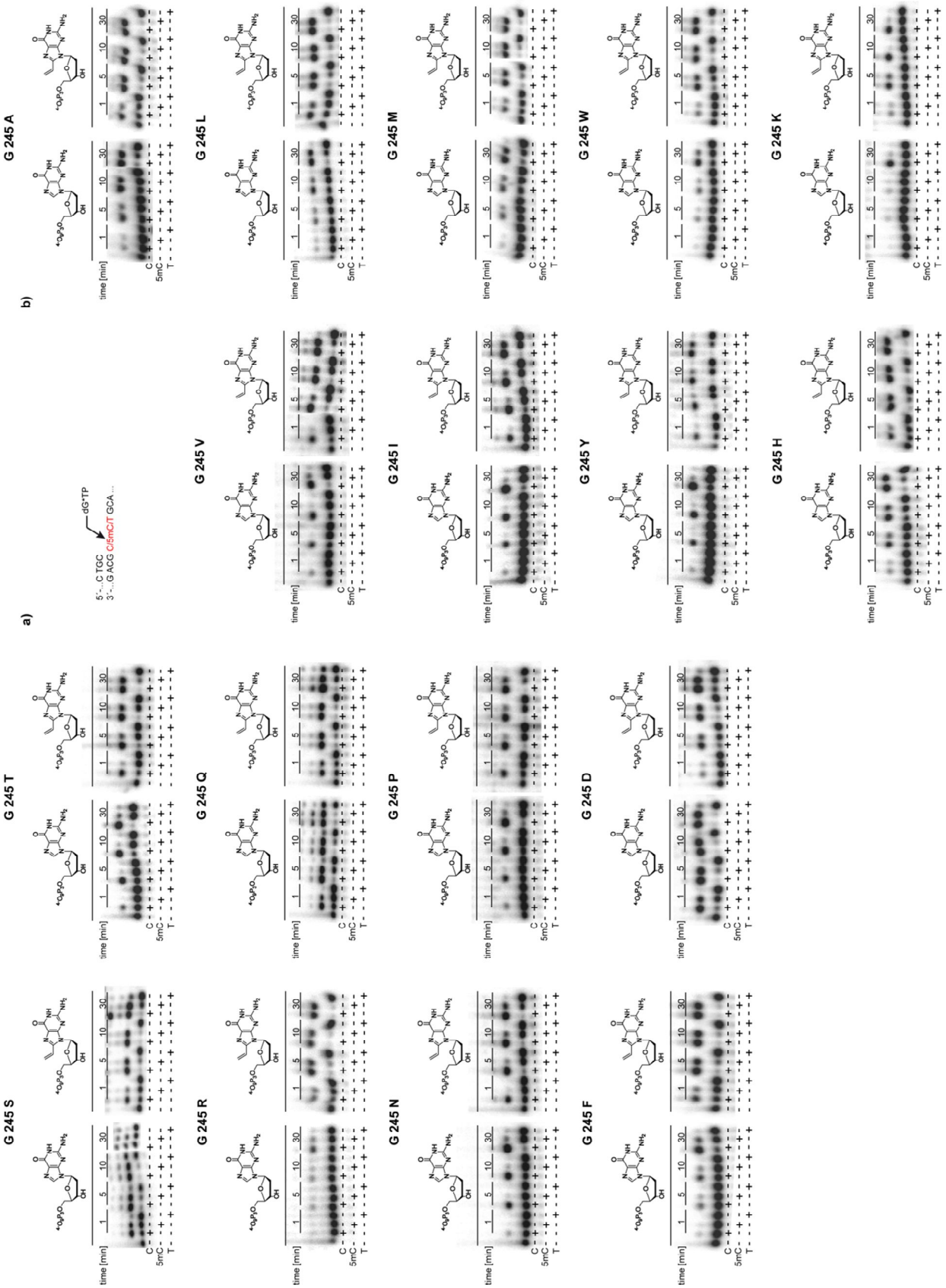


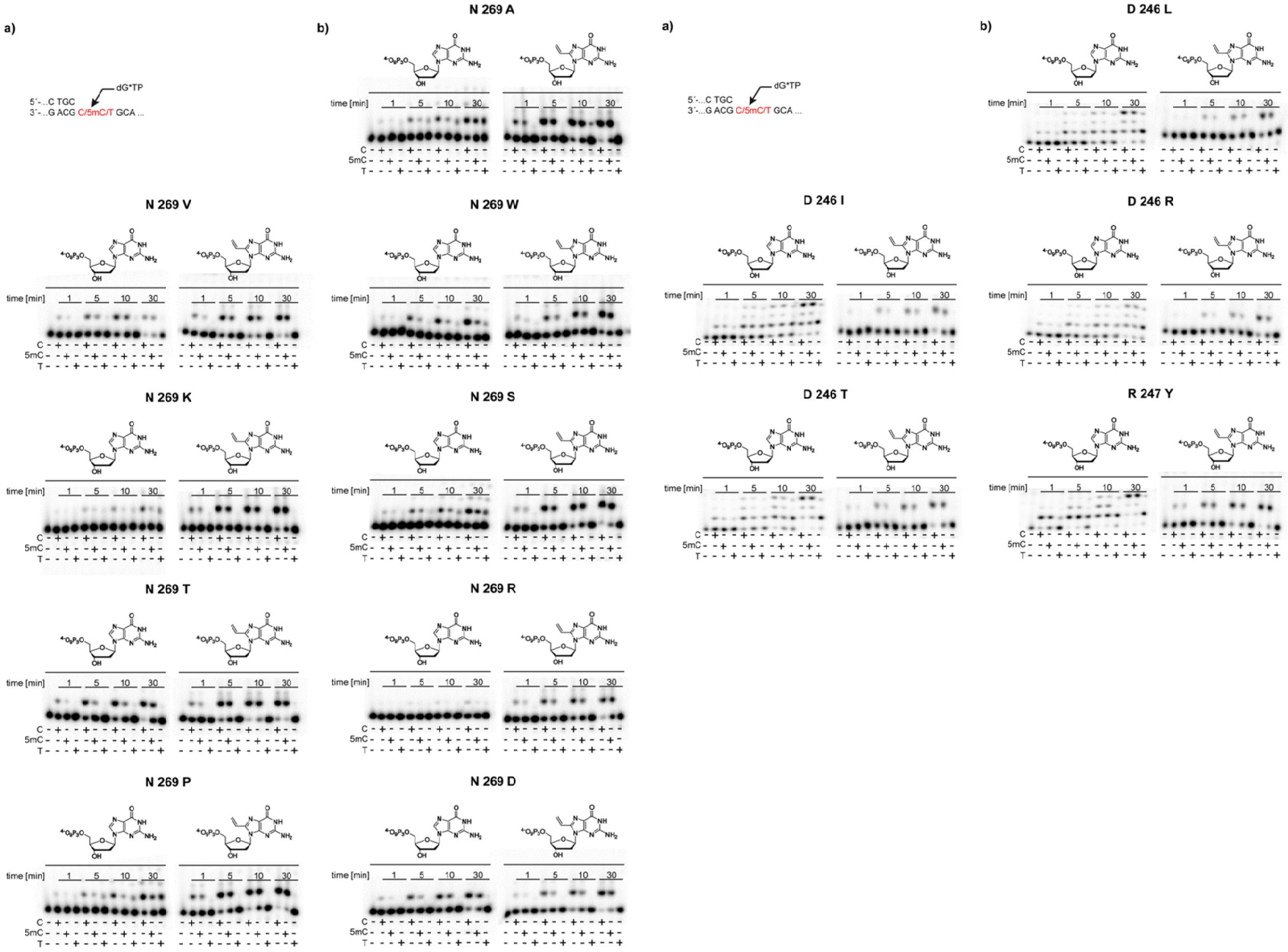

a)
b)
dG*TP
5´-...C TGC
3´-...G ACG C/5mC GCA ...
time [min]
5 10 15 20
C - + - + - + - +
5mC - - + - + - + -
% incorporation
0 25 50 75 100
0 5 10 15
time [min]

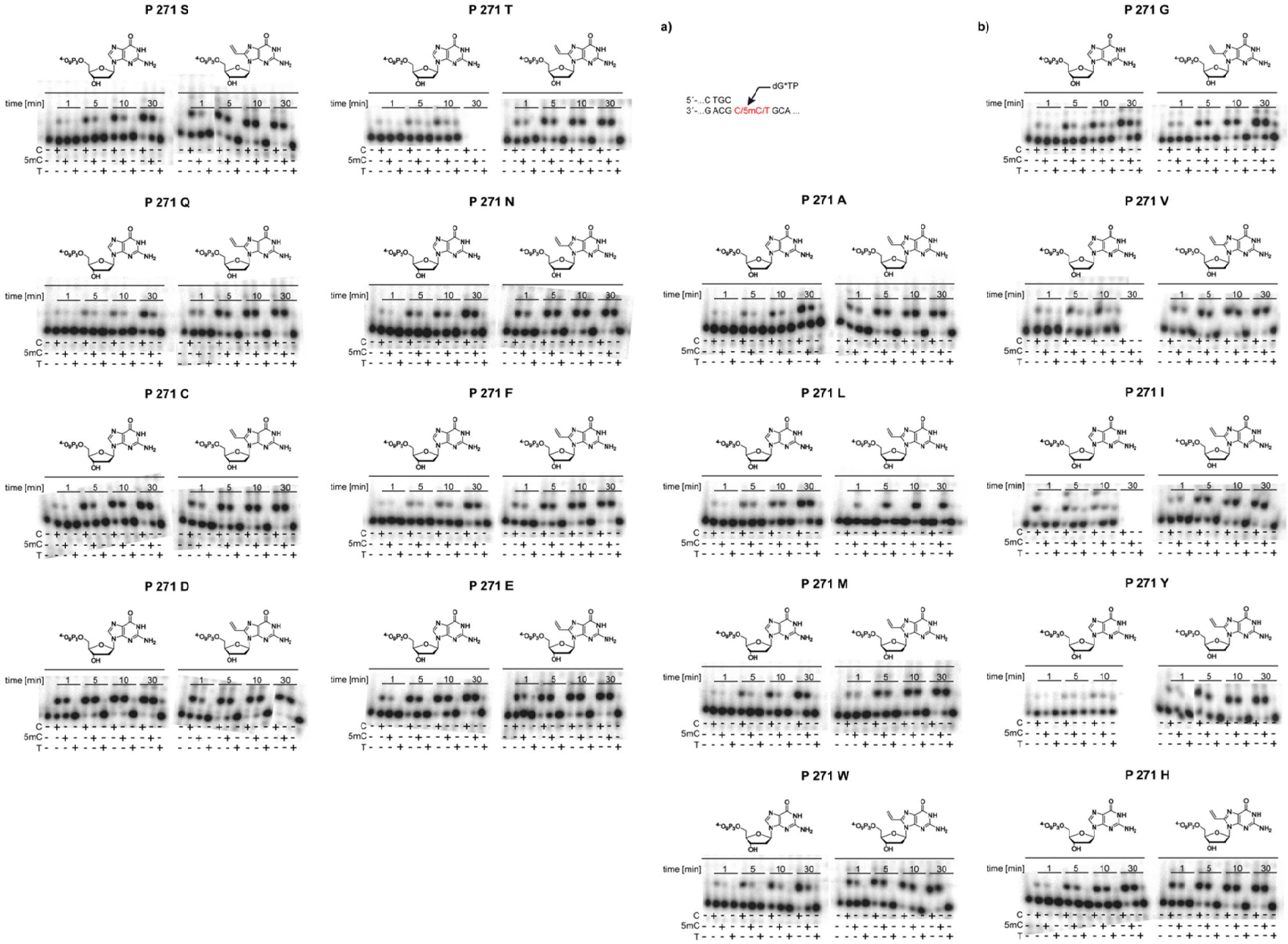

a)
dG*TP
5'-...C TGC
3'-...G ACG C/5mC GCA ...
b)
time [min]
10 20 30 45
C
5mC
% incorporation
time [min]
3 5 10 15
20 30 45 60

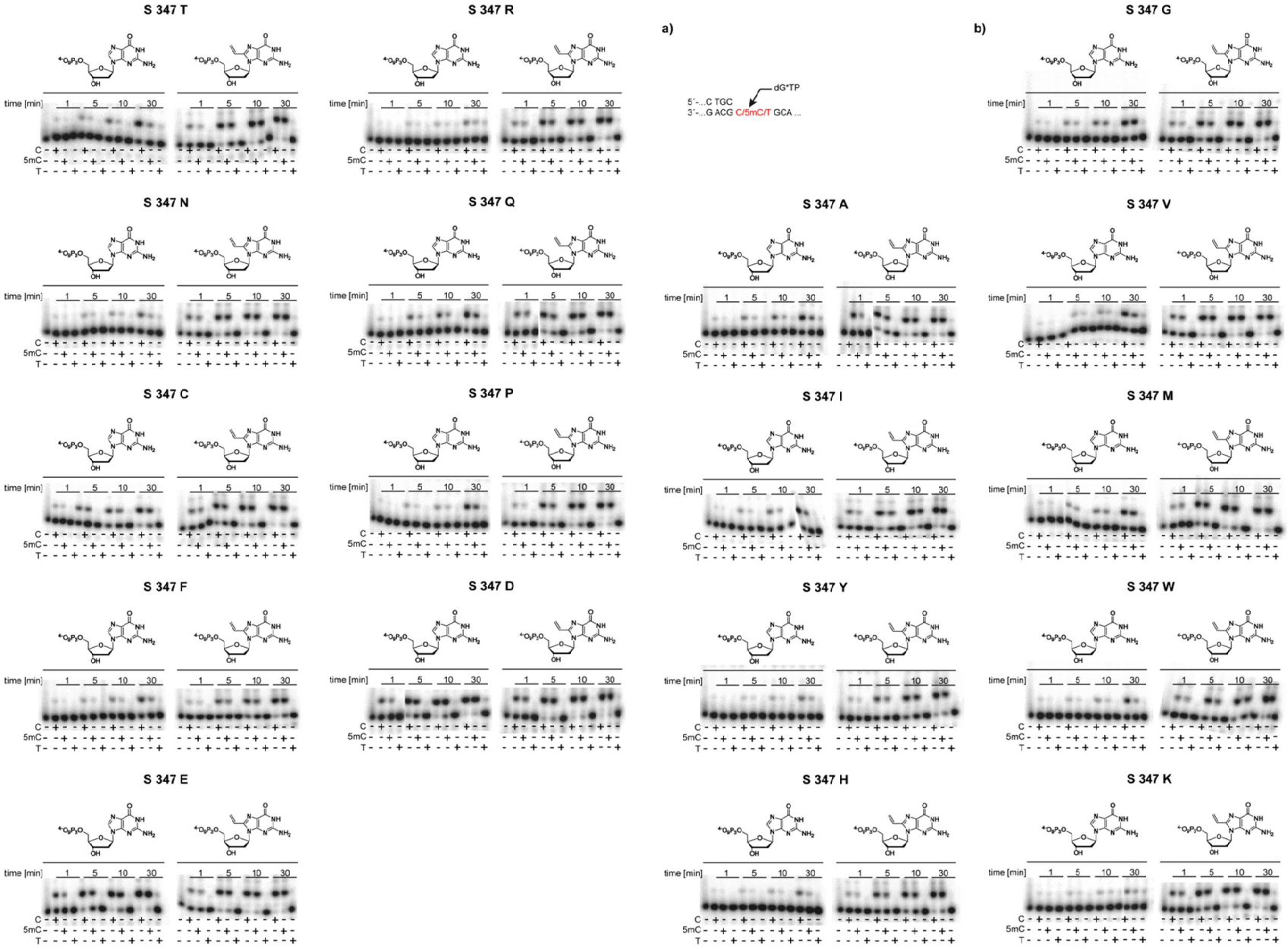
Screening KOD exo$^-$ Lysates

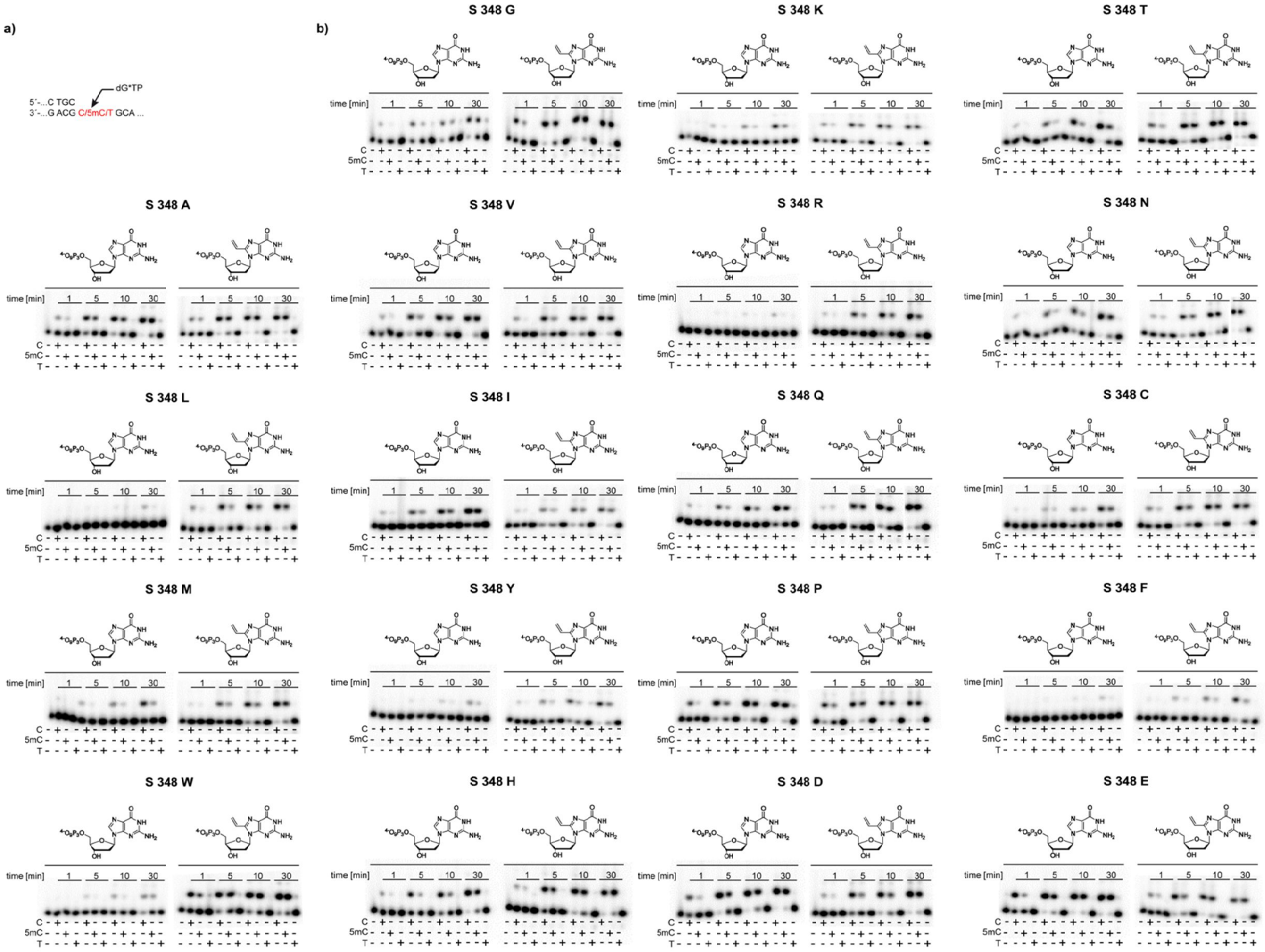
exo$^-$ domain:

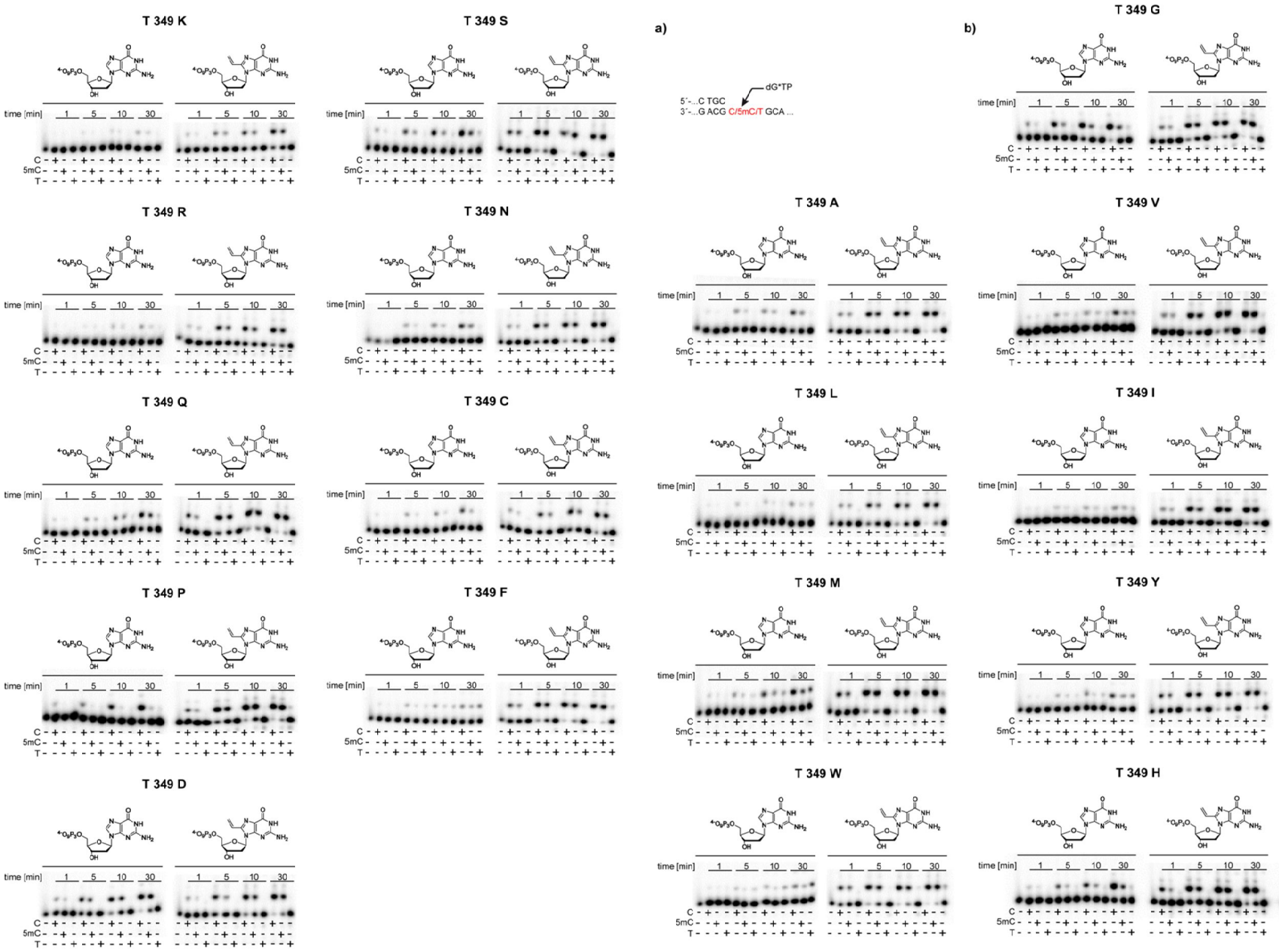


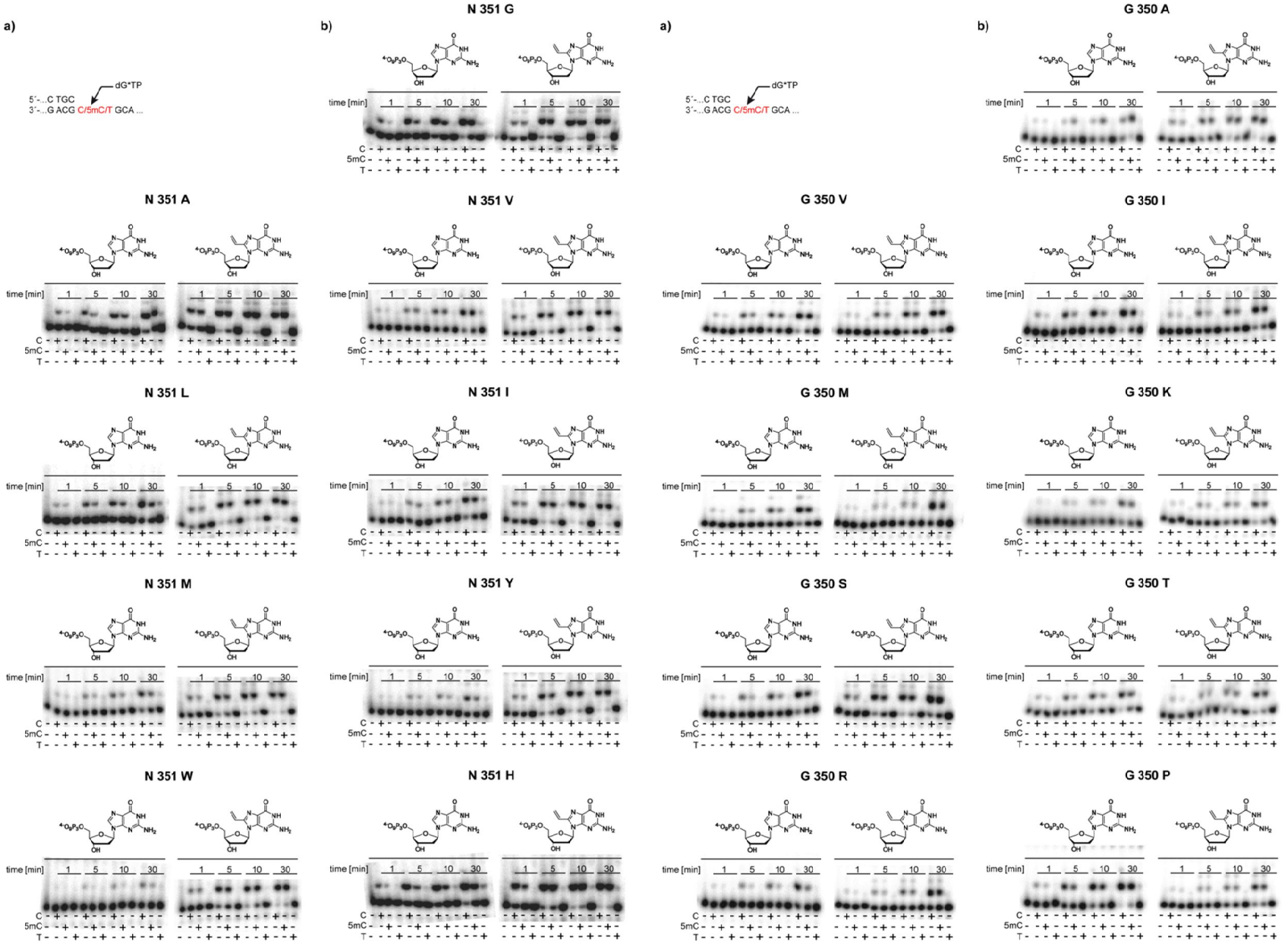

a)
dG*TP
5´-...C TGC
3´-...G ACG C/5mC/T GCA ...
b)
N 269 A
N 269 V
N 269 W
N 269 K
N 269 S
N 269 T
N 269 R
N 269 P
N 269 D
D 246 L
D 246 I
D 246 R
D 246 T
R 247 Y
time [min] 1 5 10 30
C
5mC
T

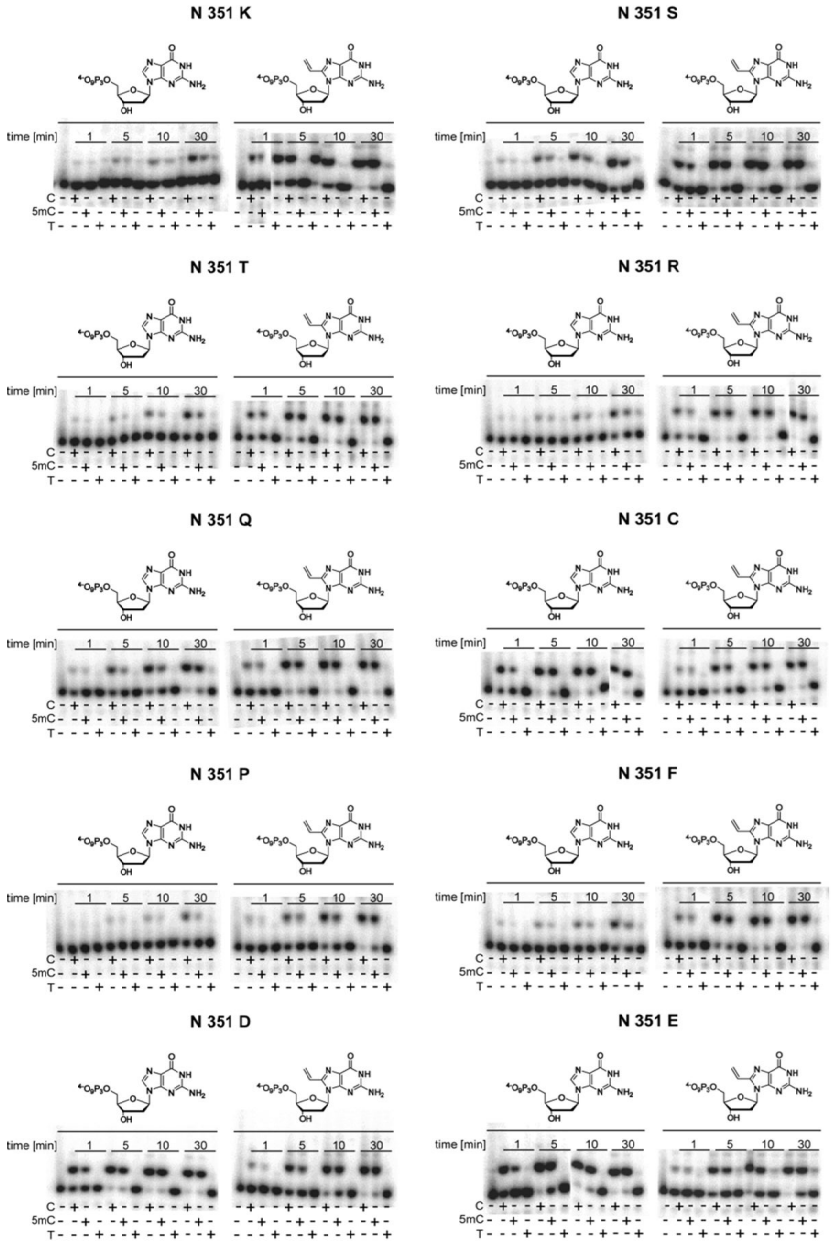

a)
b)
5′-…C TGC
3′-…G ACG C/5mC/T GCA …
dG*TP
P 271 S
P 271 T
P 271 G
P 271 Q
P 271 N
P 271 A
P 271 V
P 271 C
P 271 F
P 271 L
P 271 I
P 271 D
P 271 E
P 271 M
P 271 Y
P 271 W
P 271 H
time [min]
1
5
10
30
C
5mC
T


S 347 T
S 347 R
a)
b)
S 347 G
S 347 N
S 347 Q
S 347 A
S 347 V
S 347 C
S 347 P
S 347 I
S 347 M
S 347 F
S 347 D
S 347 Y
S 347 W
S 347 E
S 347 H
S 347 K
dG*TP
5´-...C TGC
3´-...G ACG C/5mC/T GCA ...
time [min] 1 5 10 30
C
5mC
T
N-terminal domain:

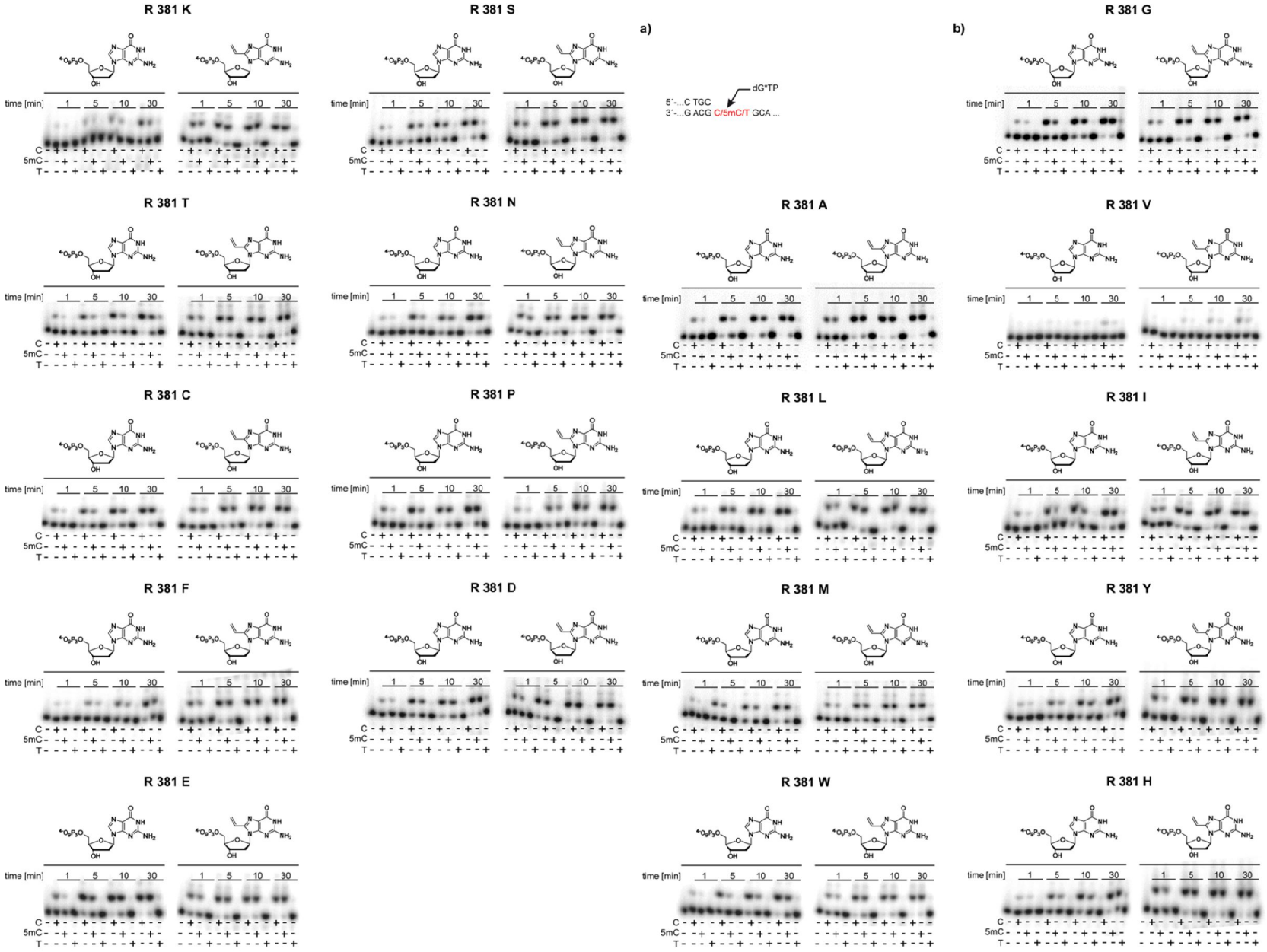

a)
dG*TP
5'-...C TGC
3'-...G ACG C/5mC/T GCA ...
b)
S 348 G
S 348 K
S 348 T
S 348 A
S 348 V
S 348 R
S 348 N
S 348 L
S 348 I
S 348 Q
S 348 C
S 348 M
S 348 Y
S 348 P
S 348 F
S 348 W
S 348 H
S 348 D
S 348 E
time [min] 1 5 10 30
C
5mC
T

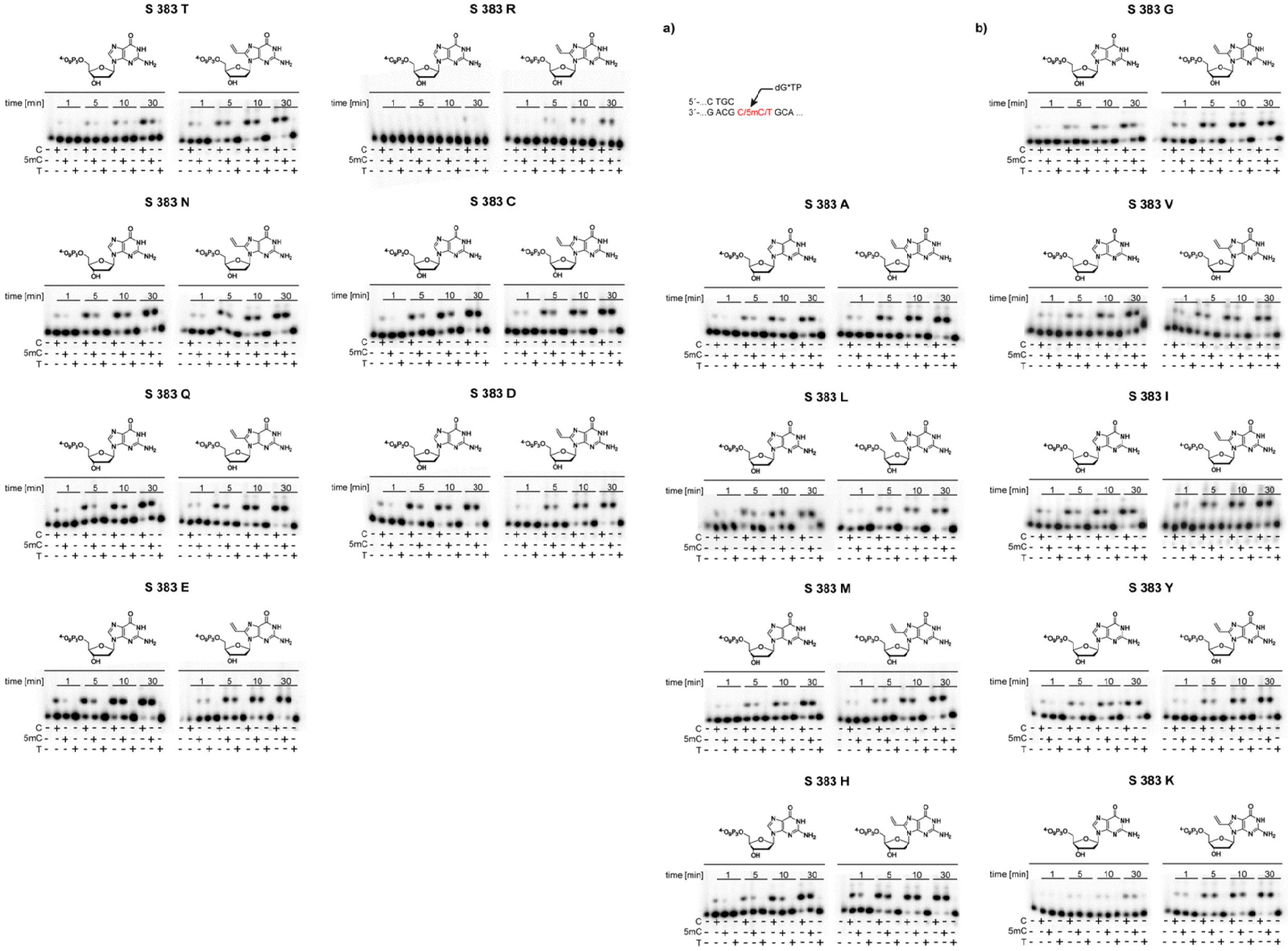

a)
b)
dG*TP
5´-...C TGC
3´-...G ACG C/5mC/T GCA ...
T 349 K
T 349 S
T 349 G
T 349 R
T 349 N
T 349 A
T 349 V
T 349 Q
T 349 C
T 349 L
T 349 I
T 349 P
T 349 F
T 349 M
T 349 Y
T 349 D
T 349 W
T 349 H
time [min] 1 5 10 30
C
5mC
T

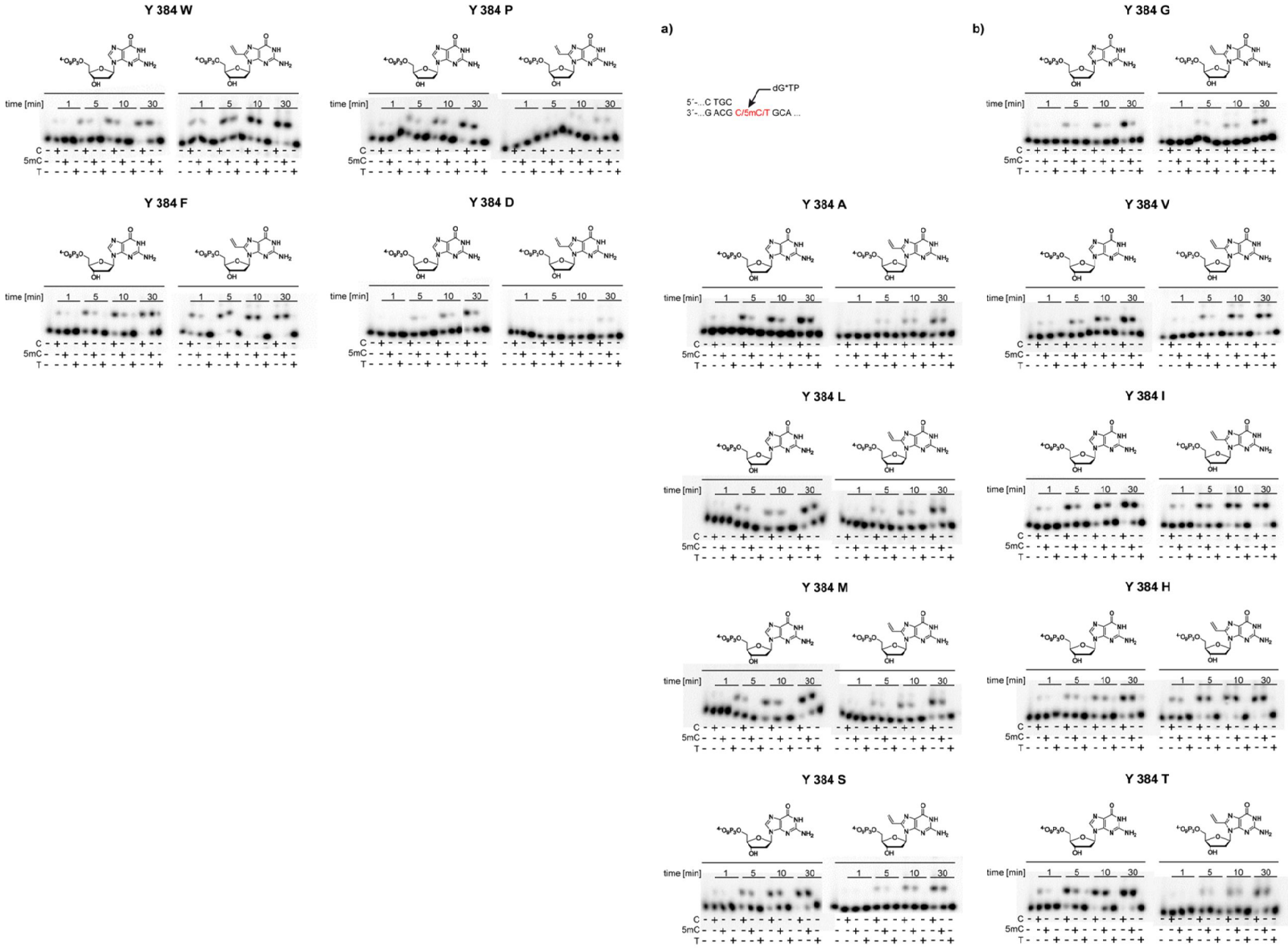

a)
dG*TP
5′-…C TGC
3′-…G ACG C/5mC/T GCA ...
b)
N 351 G
N 351 A
N 351 V
N 351 L
N 351 I
N 351 M
N 351 Y
N 351 W
N 351 H
G 350 A
G 350 V
G 350 I
G 350 M
G 350 K
G 350 S
G 350 T
G 350 R
G 350 P
time [min] 1 5 10 30
C
5mC
T

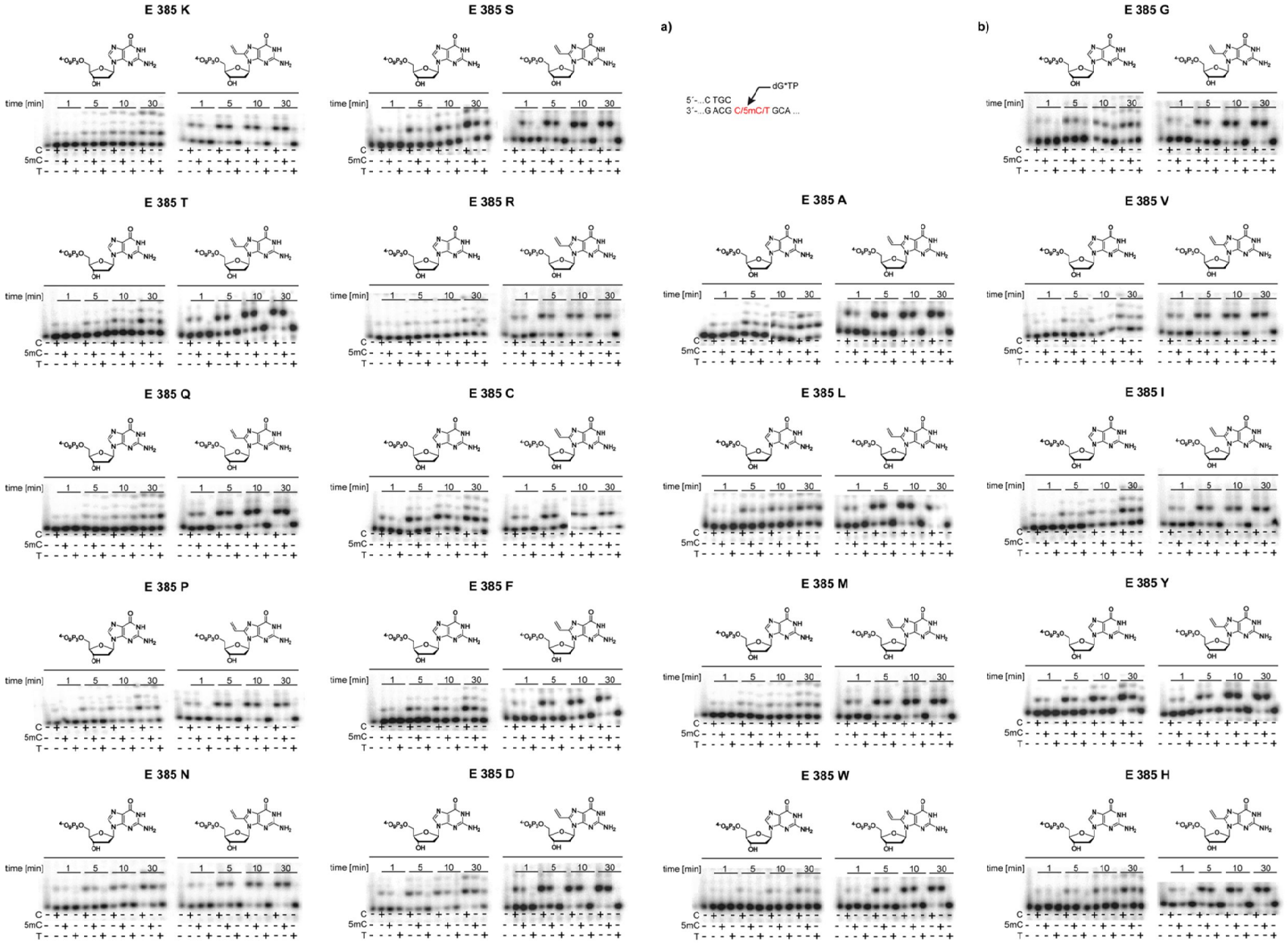

N 351 K
N 351 S
N 351 T
N 351 R
N 351 Q
N 351 C
N 351 P
N 351 F
N 351 D
N 351 E
time [min] 1 5 10 30
C
5mC
T

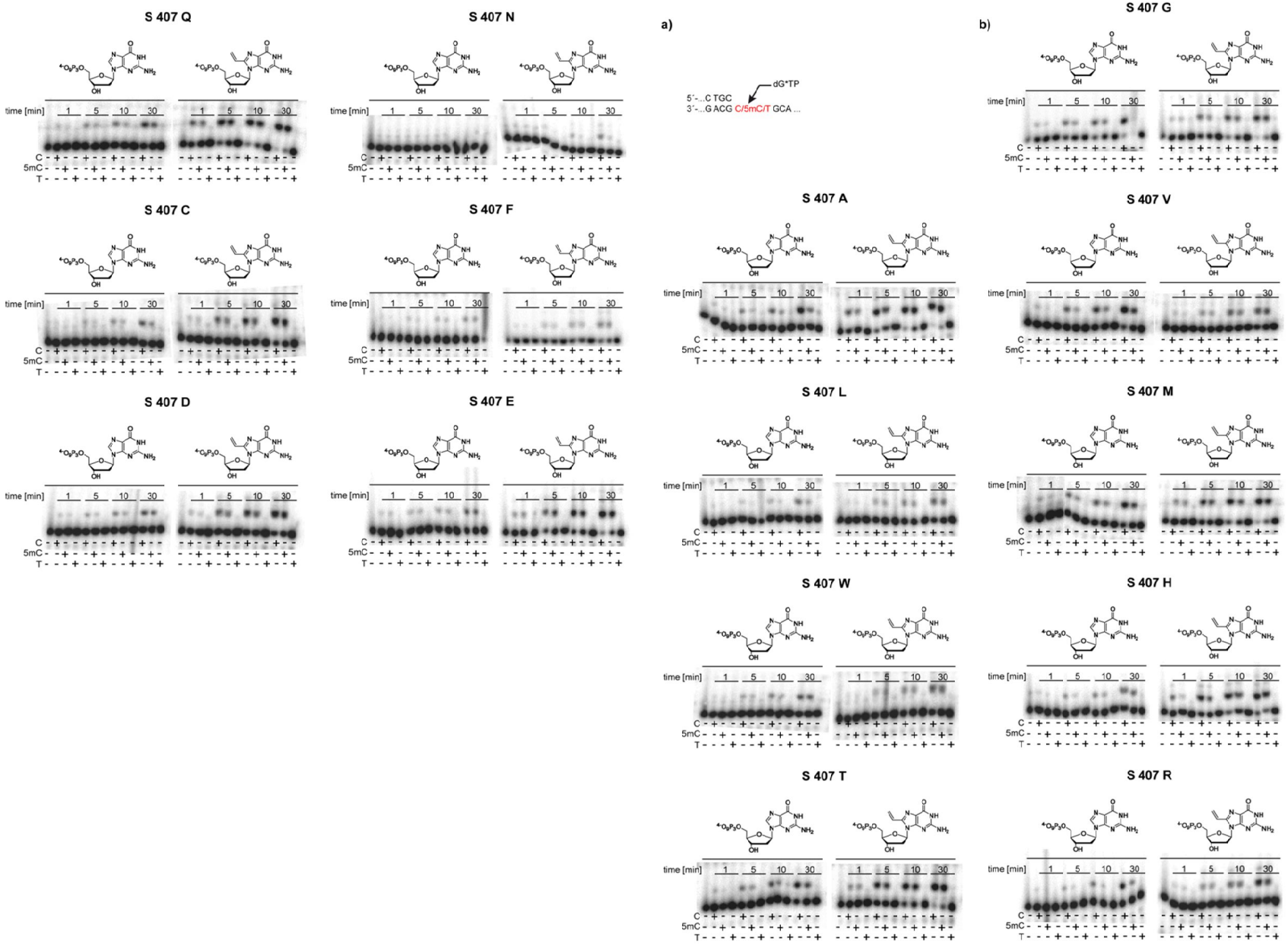
palm domain:

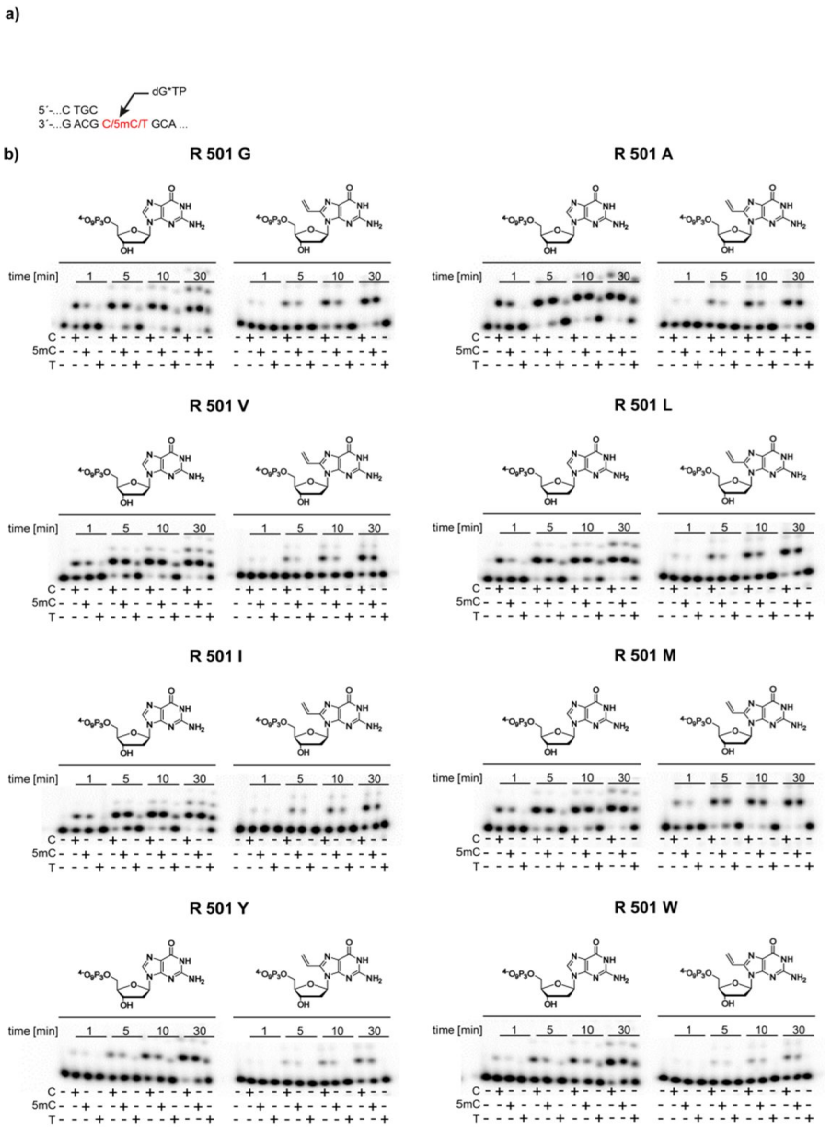

R 381 K
R 381 S
a)
5´-...C TGC
3´-...G ACG C/5mC/T GCA ...
dG*TP
b)
R 381 G
R 381 T
R 381 N
R 381 A
R 381 V
R 381 C
R 381 P
R 381 L
R 381 I
R 381 F
R 381 D
R 381 M
R 381 Y
R 381 E
R 381 W
R 381 H
time [min] 1 5 10 30
C
5mC
T

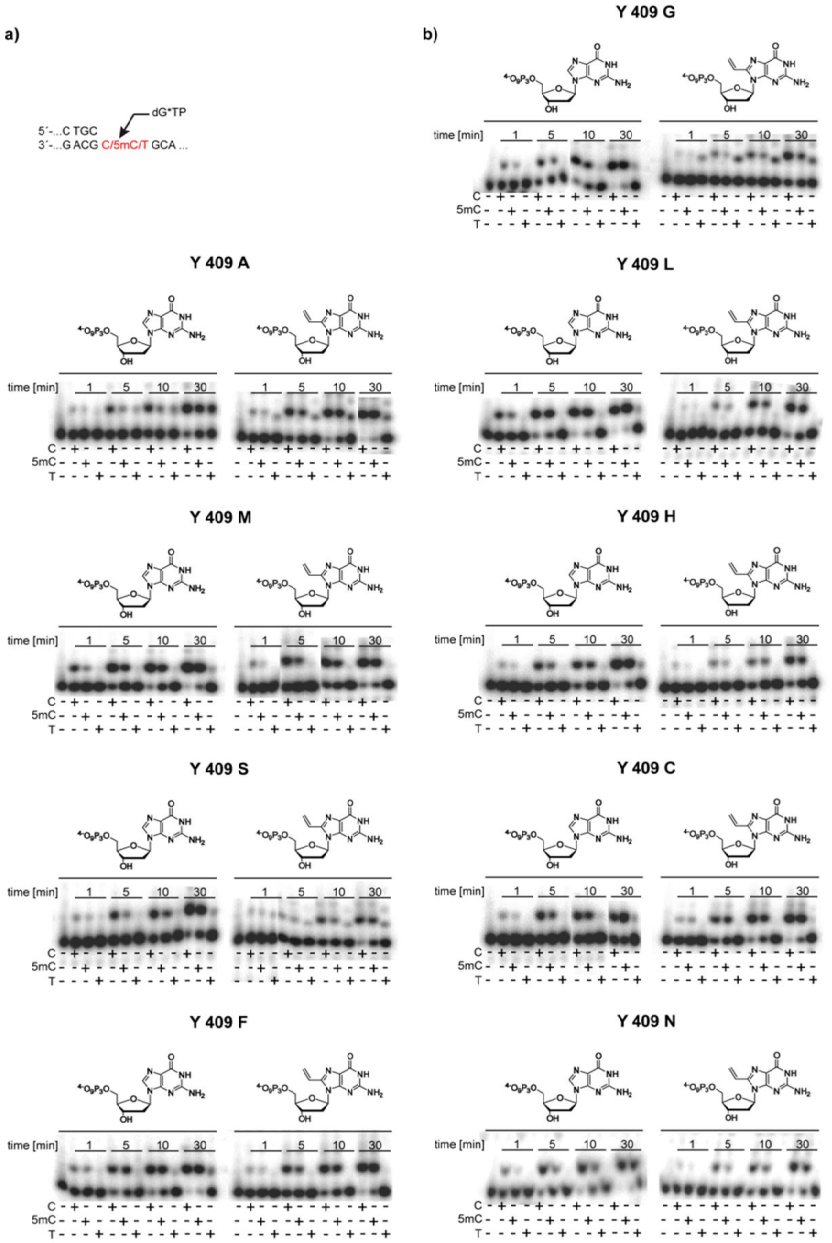

S 383 T
S 383 R
a)
5′-...C TGC
3′-...G ACG C/5mC/T GCA ...
dG*TP
b)
S 383 G
S 383 N
S 383 C
S 383 A
S 383 V
S 383 Q
S 383 D
S 383 L
S 383 I
S 383 E
S 383 M
S 383 Y
S 383 H
S 383 K
time [min] 1 5 10 30
C
5mC
T

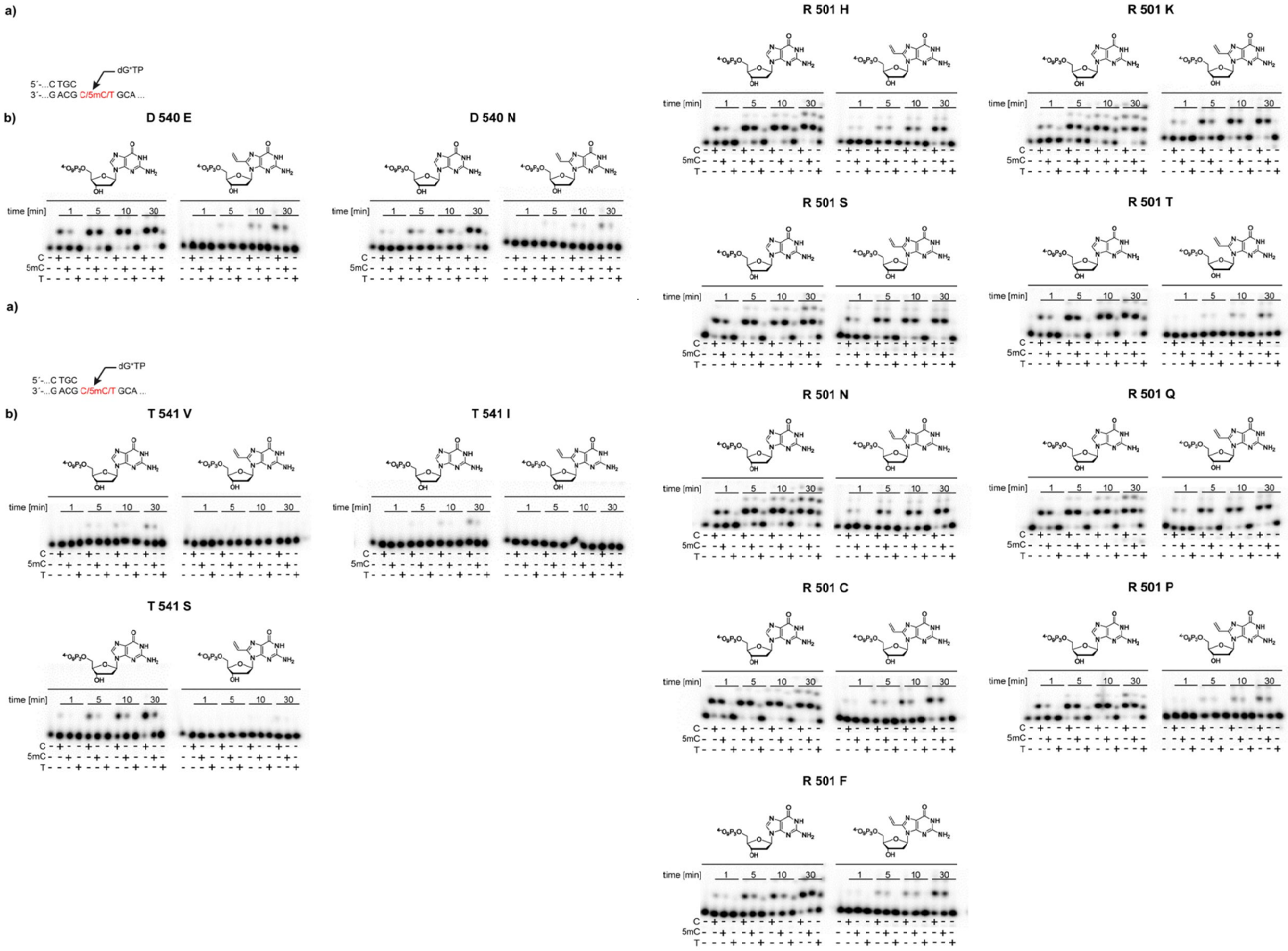

Y 384 W
Y 384 P
a)
b)
Y 384 G
5′-...C TGC
3′-...G ACG C/5mC/T GCA ...
dG*TP
Y 384 F
Y 384 D
Y 384 A
Y 384 V
Y 384 L
Y 384 I
Y 384 M
Y 384 H
Y 384 S
Y 384 T
time [min] 1 5 10 30
C
5mC
T


E 385 K
E 385 S
a)
b)
E 385 G
E 385 T
E 385 R
E 385 A
E 385 V
E 385 Q
E 385 C
E 385 L
E 385 I
E 385 P
E 385 F
E 385 M
E 385 Y
E 385 N
E 385 D
E 385 W
E 385 H
dG*TP
5´-…C TGC
3´-…G ACG C/5mC/T GCA …
time [min]
1
5
10
30
C
5mC
T

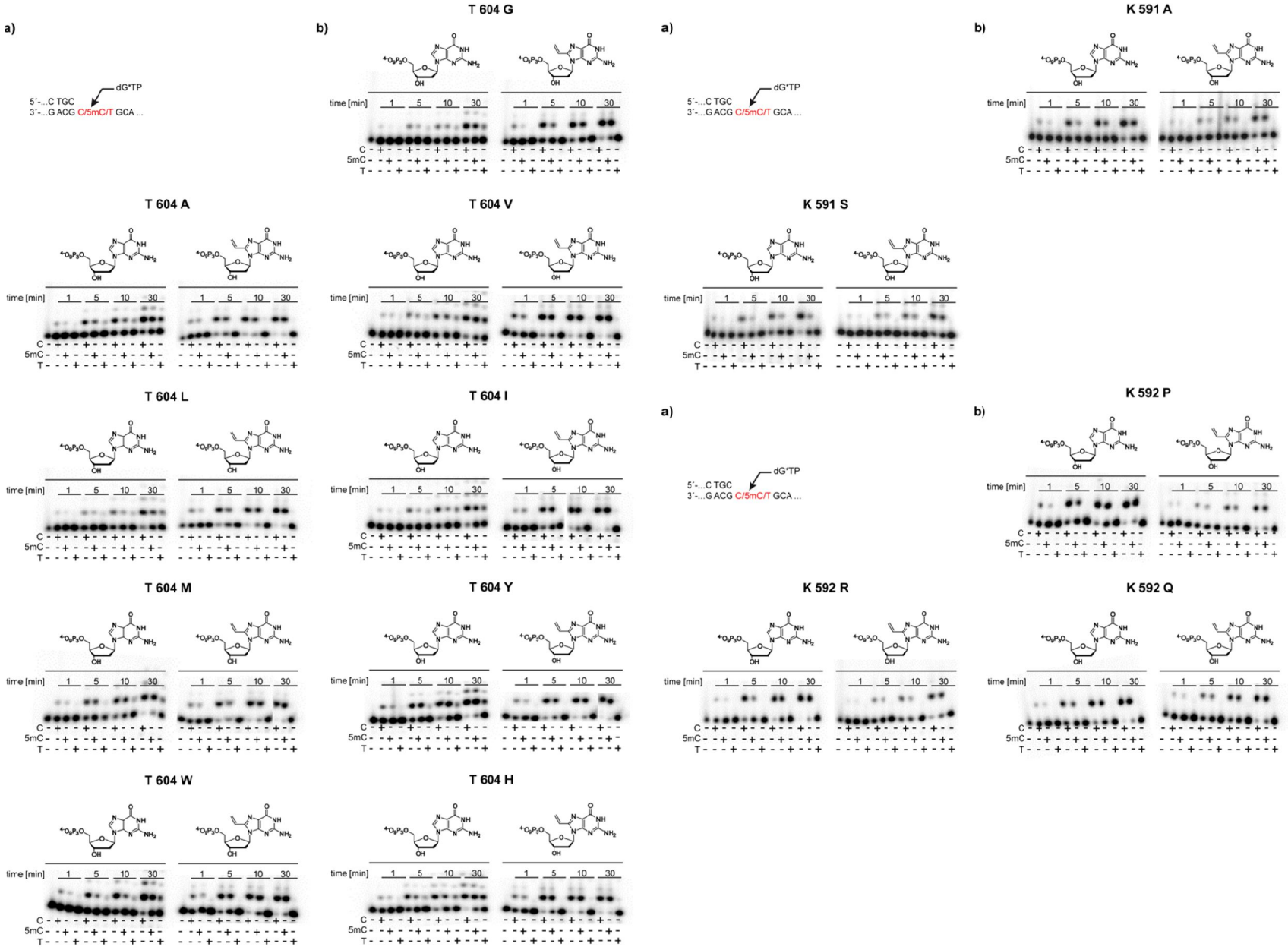

S 407 Q
S 407 N
S 407 C
S 407 F
S 407 D
S 407 E
a)
dG*TP
5´-…C TGC
3´-…G ACG C/5mC/T GCA …
S 407 A
S 407 L
S 407 W
S 407 T
b)
S 407 G
S 407 V
S 407 M
S 407 H
S 407 R
time [min] 1 5 10 30
C
5mC
T

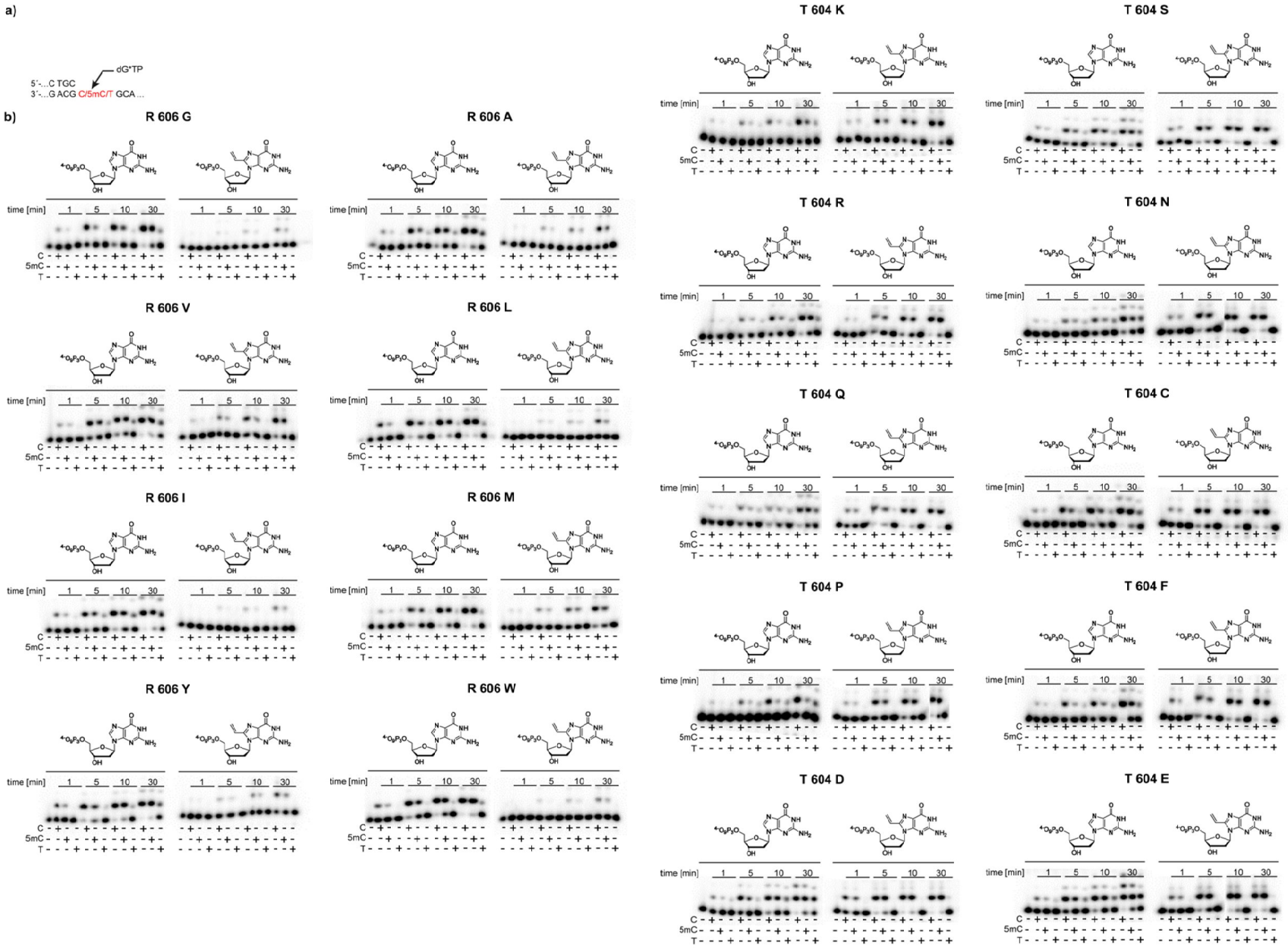

a)
dG*TP
5´-...C TGC
3´-...G ACG C/5mC/T GCA ...
b)
R 501 G
R 501 A
R 501 V
R 501 L
R 501 I
R 501 M
R 501 Y
R 501 W
time [min] 1 5 10 30
C
5mC
T

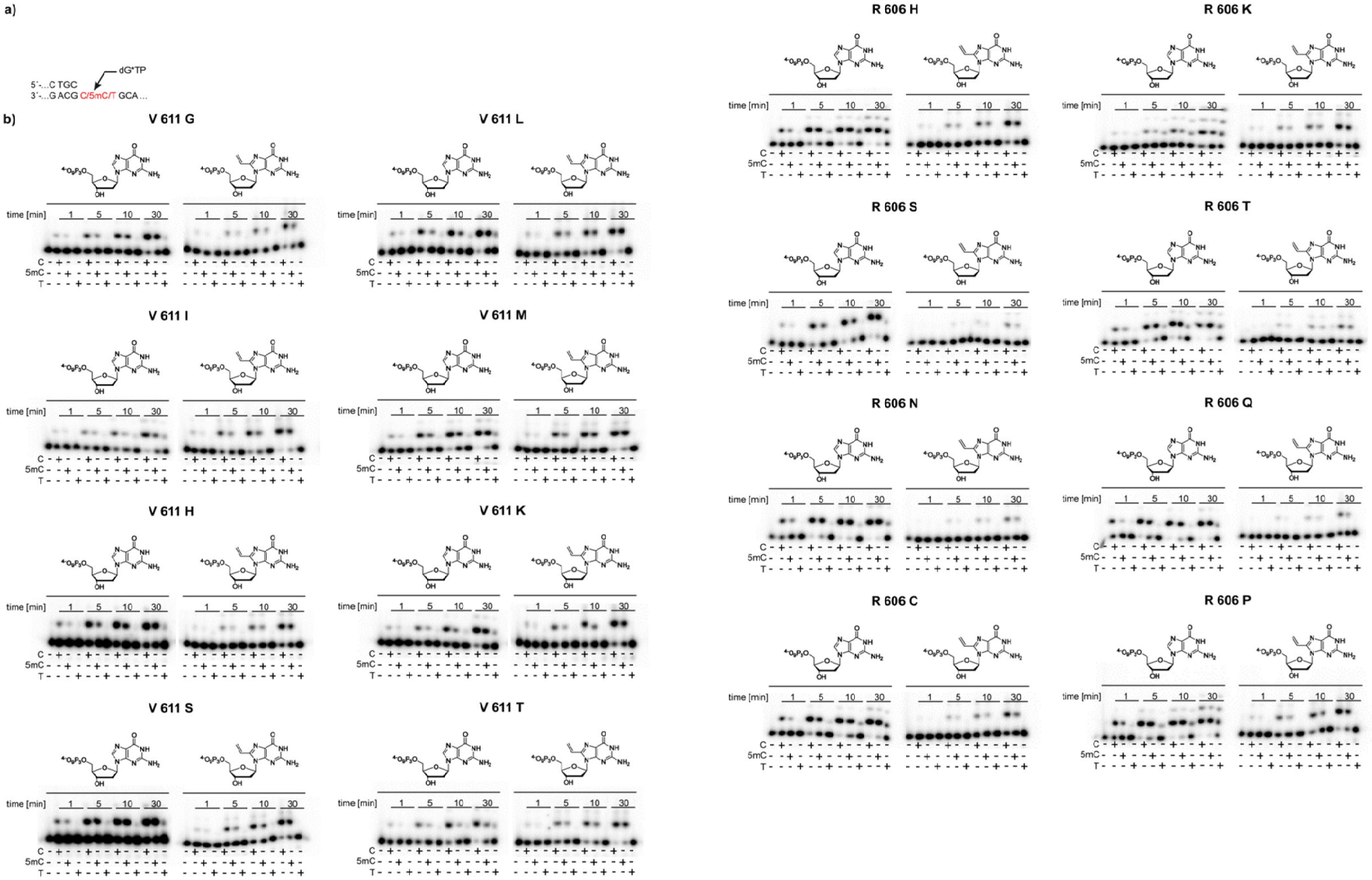

a)
dG*TP
5´-...C TGC
3´-...G ACG C/5mC/T GCA ...
b)
Y 409 G
Y 409 A
Y 409 L
Y 409 M
Y 409 H
Y 409 S
Y 409 C
Y 409 F
Y 409 N
time [min] 1 5 10 30
C
5mC
T

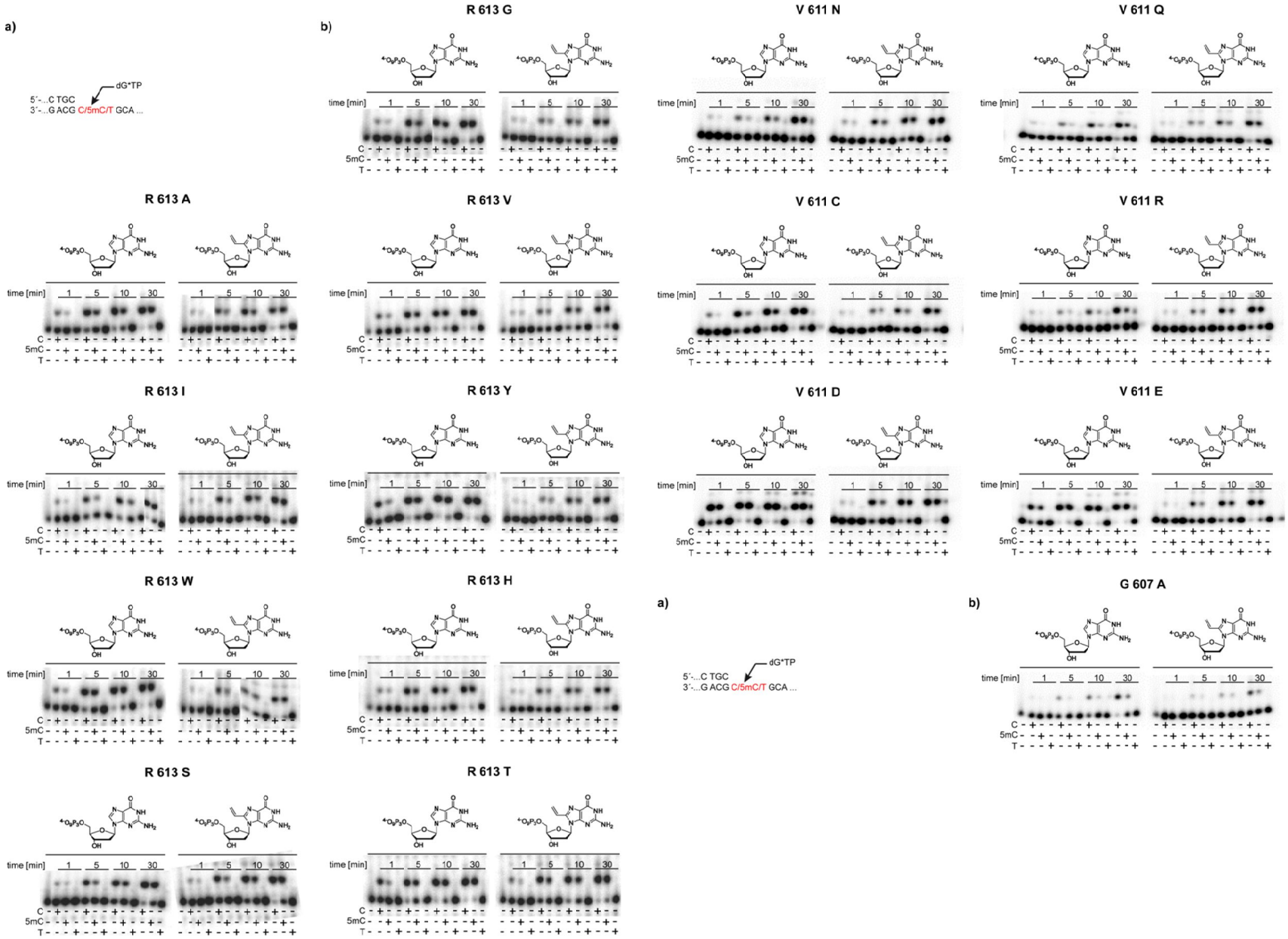

a)
dG*TP
5´-...C TGC
3´-...G ACG C/5mC/T GCA ...
b)
D 540 E
D 540 N
T 541 V
T 541 I
T 541 S
R 501 H
R 501 K
R 501 S
R 501 T
R 501 N
R 501 Q
R 501 C
R 501 P
R 501 F
time [min] 1 5 10 30
C
5mC
T

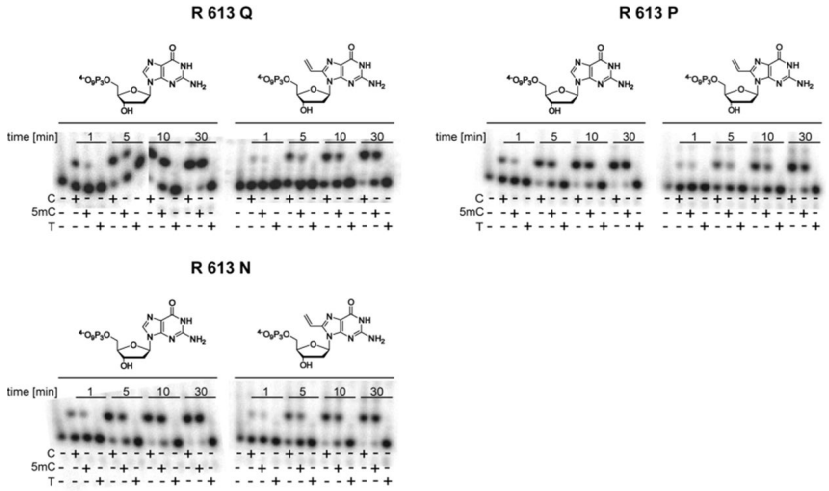
thumb domain:



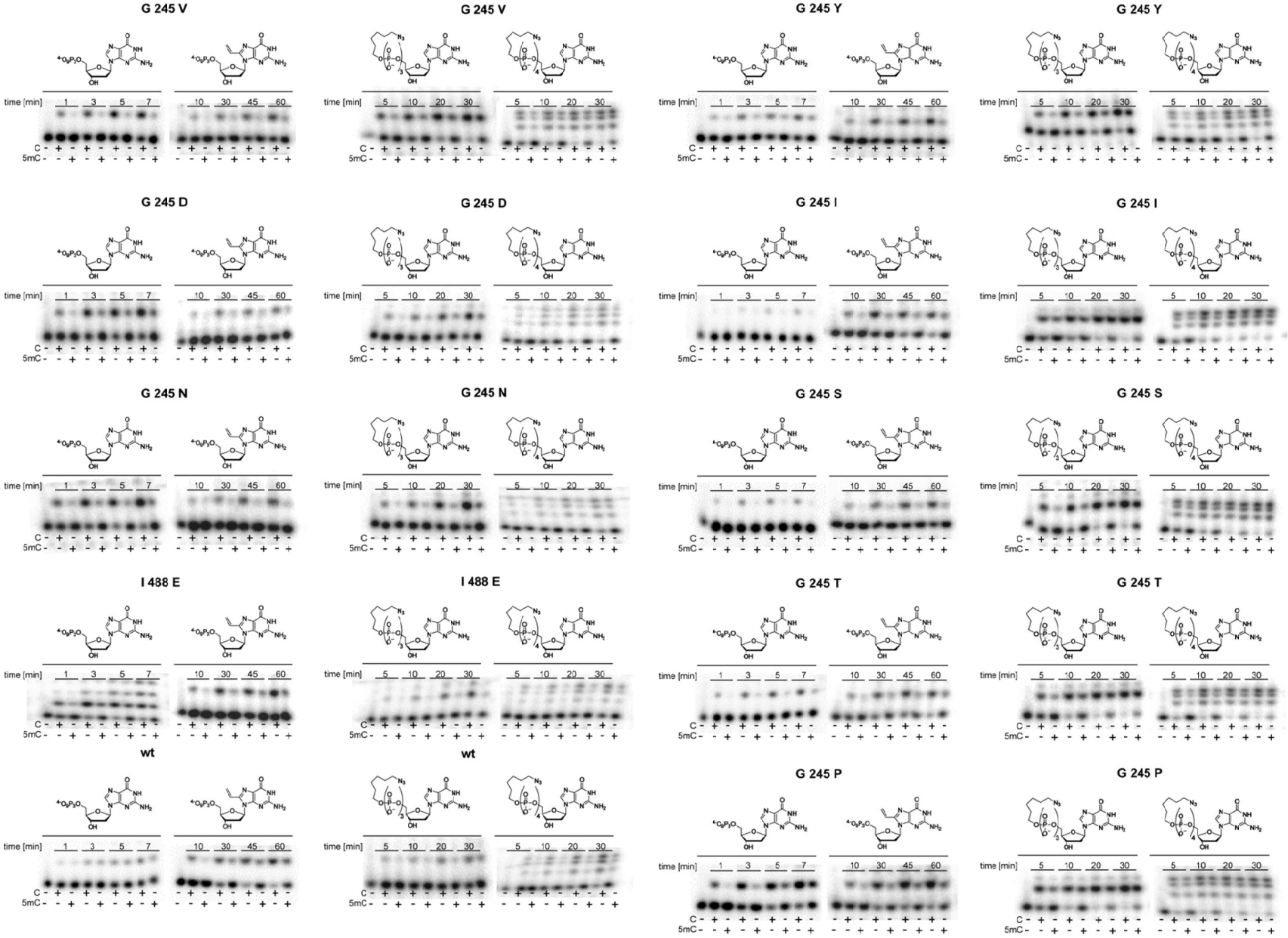

a)
dG*TP
5´-...C TGC
3´-...G ACG C/5mC/T GCA ...
b)
R 606 G
R 606 A
R 606 V
R 606 L
R 606 I
R 606 M
R 606 Y
R 606 W
T 604 K
T 604 S
T 604 R
T 604 N
T 604 Q
T 604 C
T 604 P
T 604 F
T 604 D
T 604 E
time [min] 1 5 10 30
C
5mC
T


a)
dG*TP
5'-…C TGC
3'-…G ACG C/5mC/T GCA…
b)
V 611 G
V 611 L
V 611 I
V 611 M
V 611 H
V 611 K
V 611 S
V 611 T
R 606 H
R 606 K
R 606 S
R 606 T
R 606 N
R 606 Q
R 606 C
R 606 P
time [min] 1 5 10 30
C
5mC
T

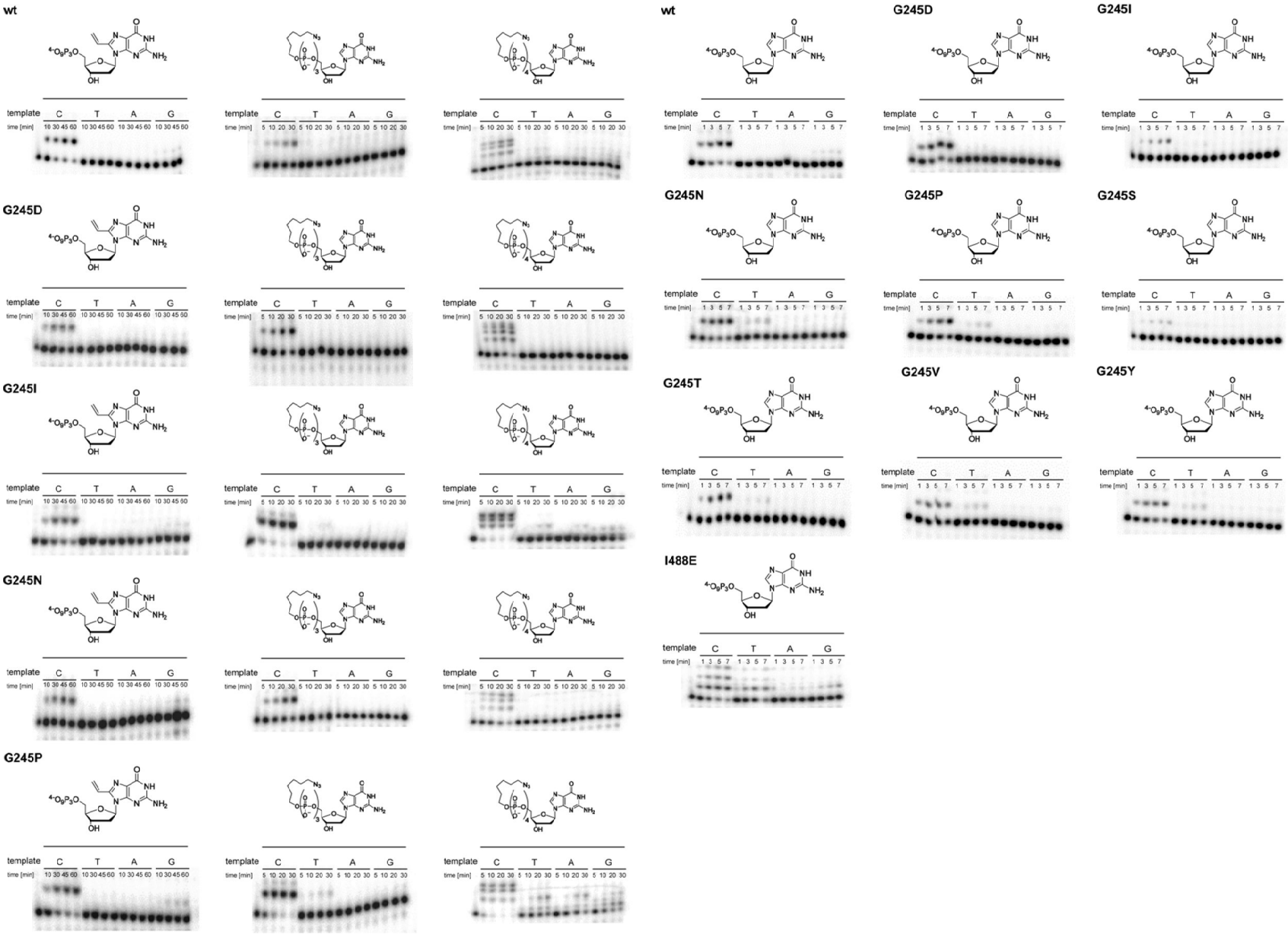

a)
dG*TP
5´-...C TGC
3´-...G ACG C/5mC/T GCA ...
b)
R 613 G
R 613 A
R 613 V
R 613 I
R 613 Y
R 613 W
R 613 H
R 613 S
R 613 T
V 611 N
V 611 Q
V 611 C
V 611 R
V 611 D
V 611 E
G 607 A
time [min] 1 5 10 30
C
5mC
T

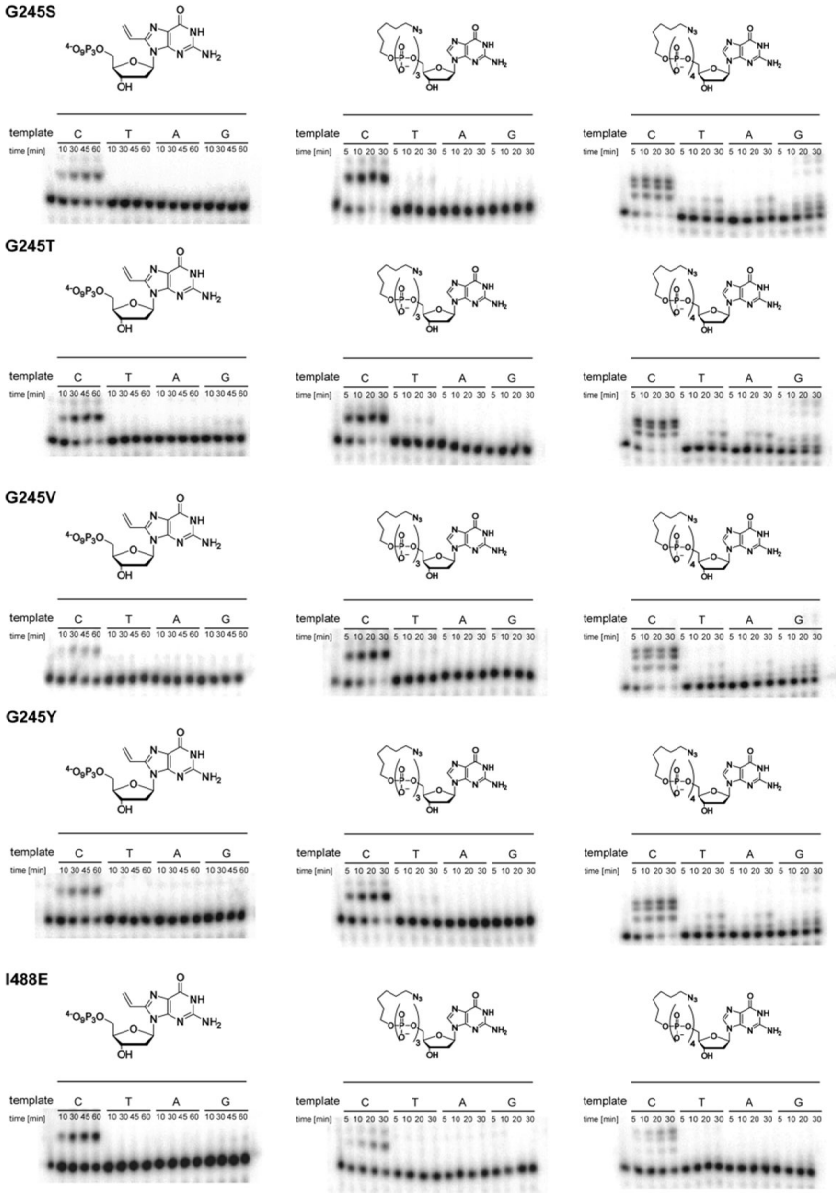

R 613 Q
R 613 P
R 613 N
time [min]
1
5
10
30
C
5mC
T

Screening of purified KOD exo⁻ mutants:



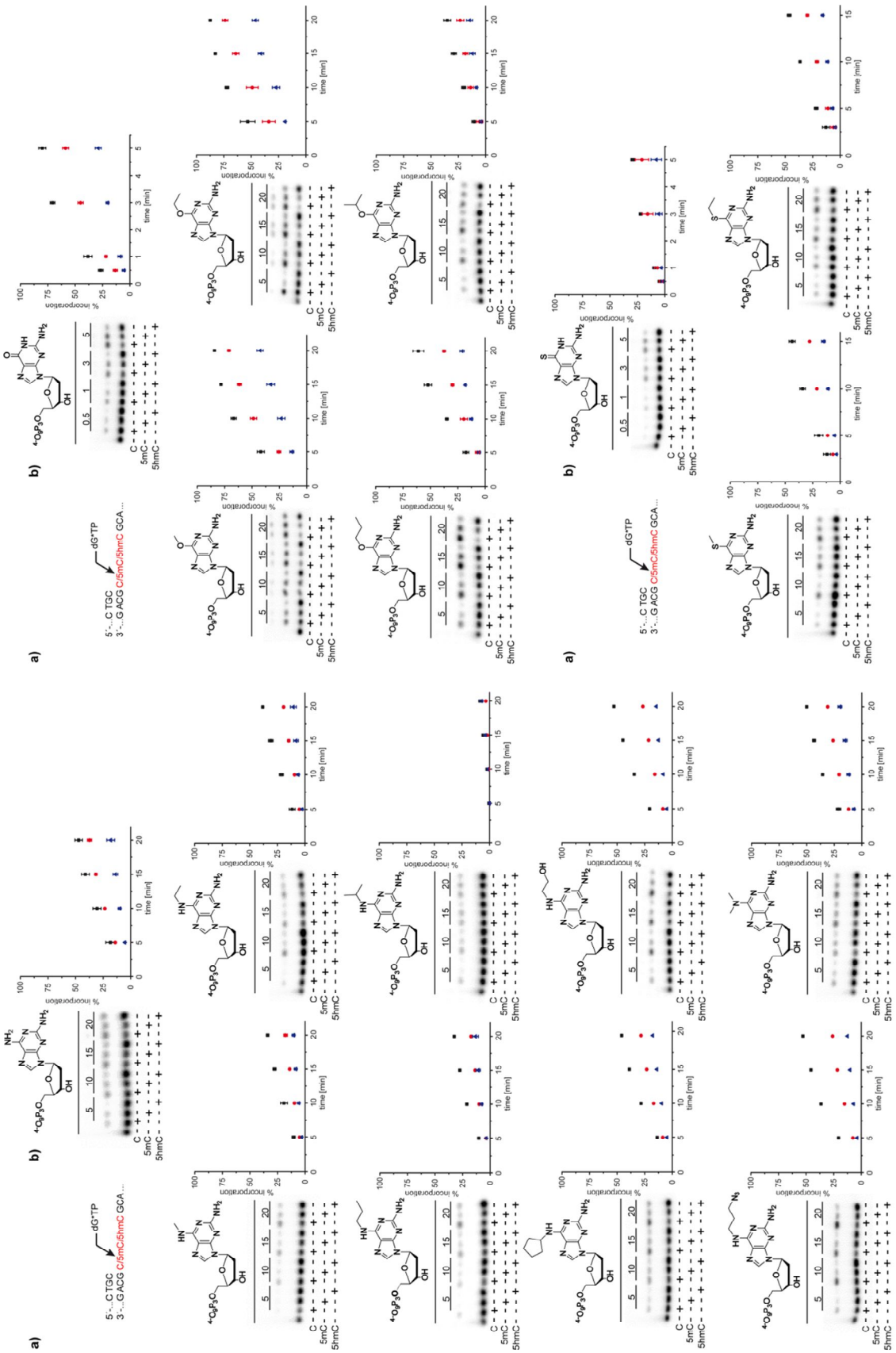
Selectivity studies of purified KOD exo⁻ mutants:

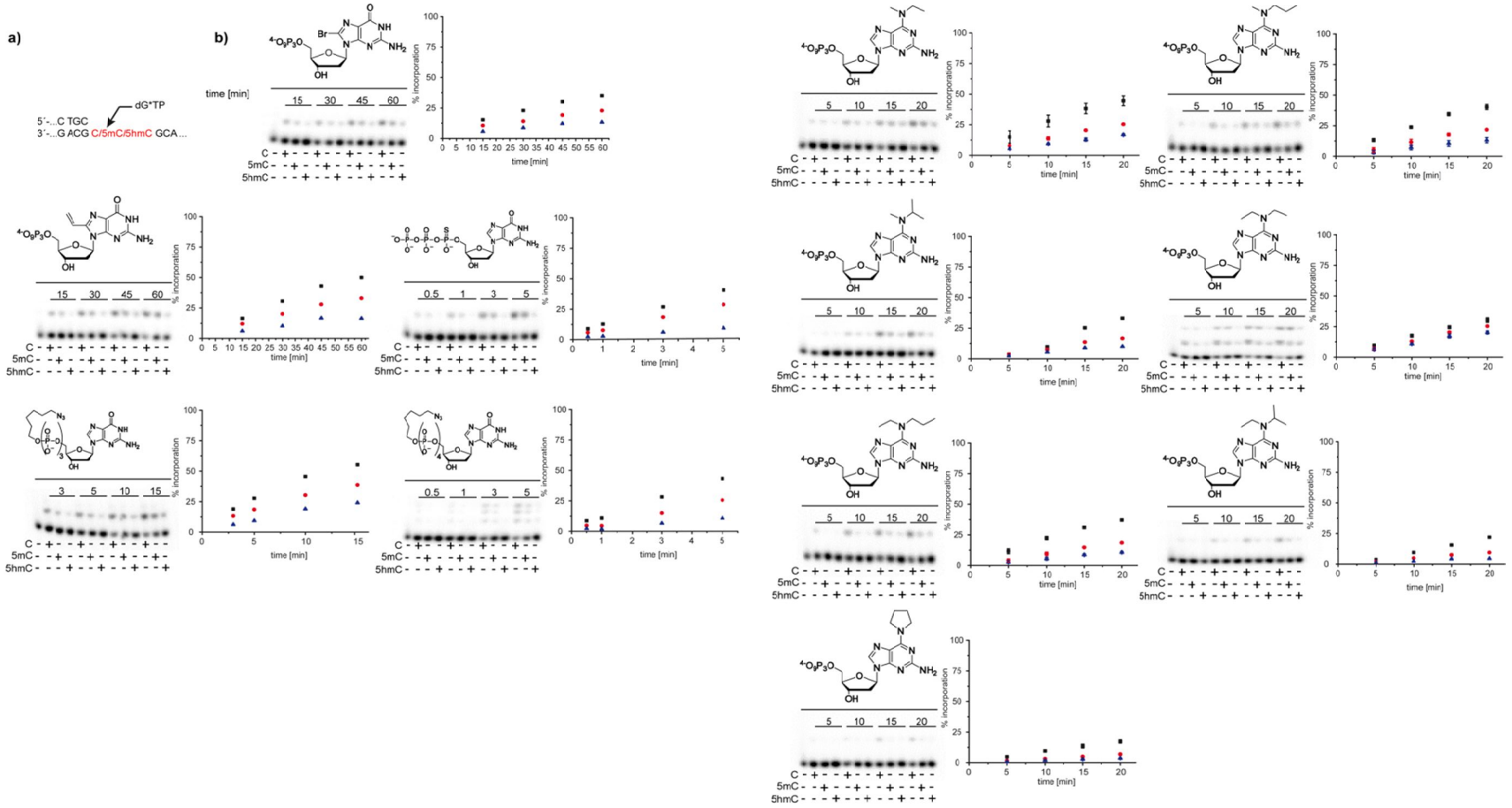



G245S
G245T
G245V
G245Y
I488E
template
C T A G
time [min]
10 30 45 60
5 10 20 30

Screening of modified nucleotides for 5hmC detection:

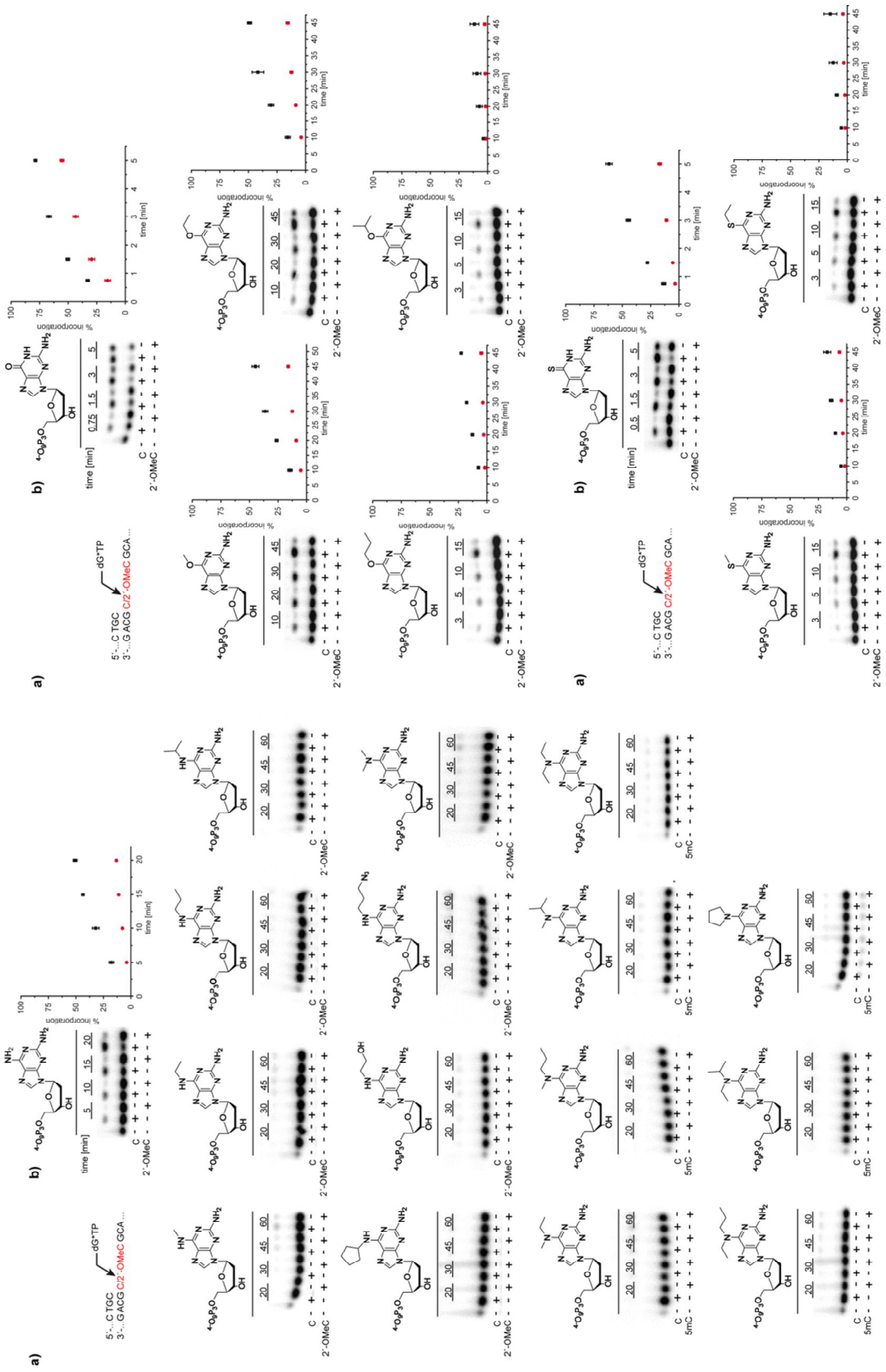
KOD exo$^-$:

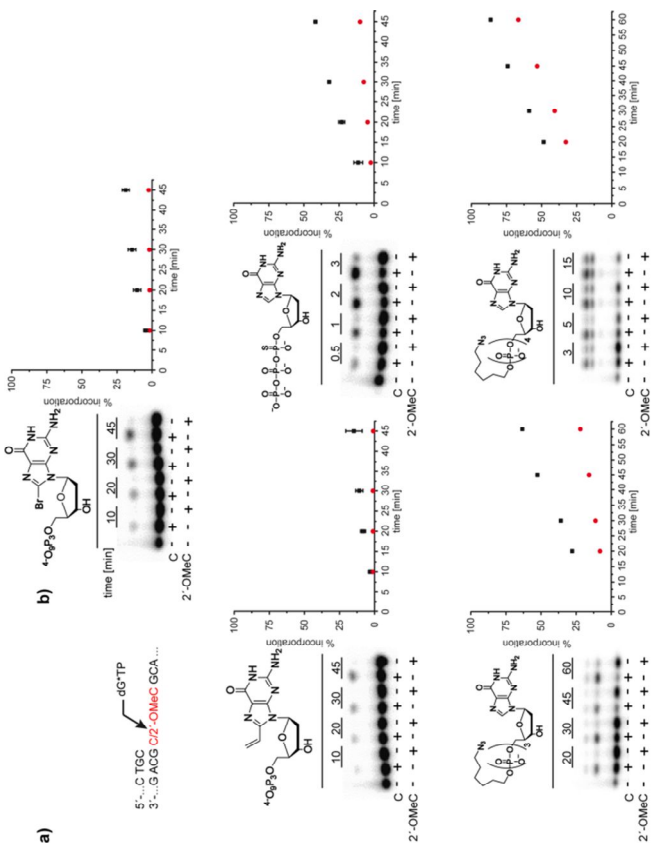



a)
dG*TP
5´-...C TGC
3´-...G ACG C/5mC/5hmC GCA...
b)
time [min]
% incorporation
C
5mC
5hmC

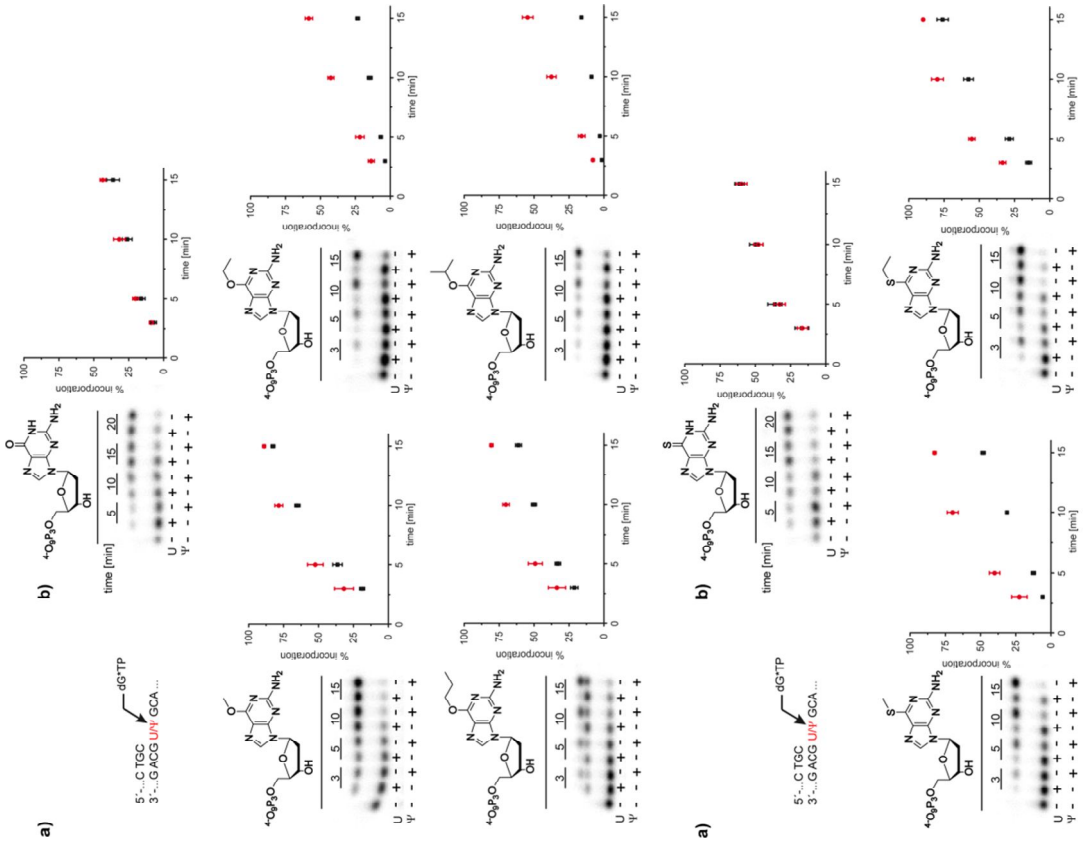
Screening of modified nucleotides for detection RNA modifications:

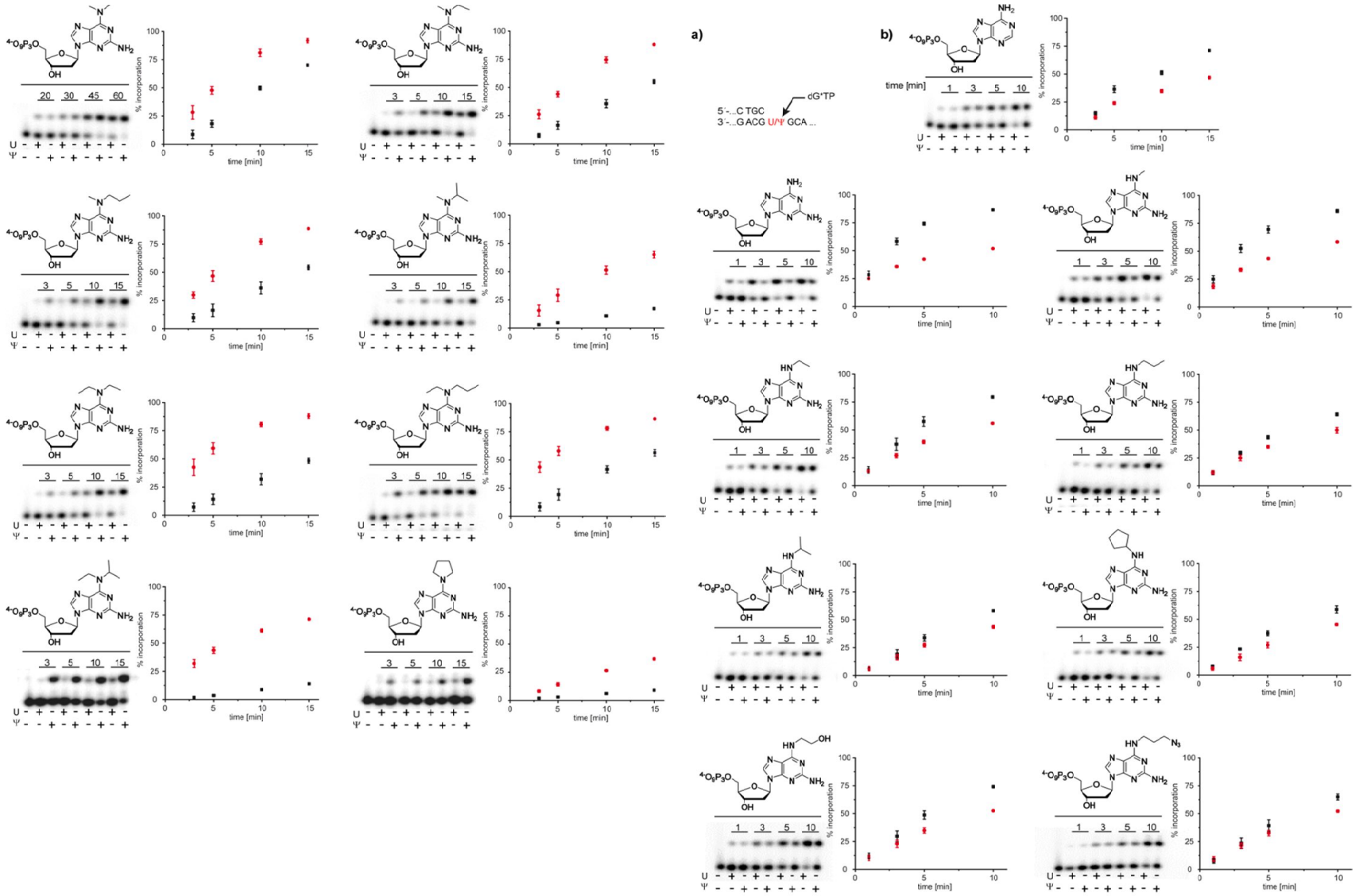
2´-*O*Methyl-C :





Pseudouridine:




a)
b)
dG*TP
5´-...C TGC
3´-...G ACG U/Ψ GCA ...
time [min]
% incorporation
U
Ψ

8.4. Sequences

5mC/5hmC:

template:

5'- CCG CTG CCC ACC AGC CAT CAT GTC GGA CCC CGC GGT CAA CG **C/5mC/5hmC** GCA GCT GGA TGG GAT CAT TTC GGA CT -3'

primer:

5'- CGA AAT GAT CCC ATC CAG CTG C -3'

CE:

primer:

5´-CGA AAT GAT CCC ATC CAG CTG C -3'

32 mer:

5´-A AAA AAA AAA CGA AAT GAT CCC ATC CAG CTG C -3'

31 mer:

5´- AAA AAA AAA CGA AAT GAT CCC ATC CAG CTG C -3'

39 mer:

5´-AA AAA AAA AAA AAA AAA CGA AAT GAT CCC ATC CAG CTG C -3'

47 mer:

5´-A AAA AAA AAA AAA AAA AAA AAA AAA CGA AAT GAT CCC ATC CAG CTG C -3'

55 mer:

5´-AA AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA CGA AAT GAT CCC ATC CAG CTG C -3'

RNA:

template:

5'- AUA GGG GAA UGG GCC GUU CAU CUG CUA AAA GG **U/C/Ψ/2´-OMe-C** CUG CUU UUG GGG CUU GUA GU -3'

primer:

5'- ACT ACA AGC CCC AAA AGC AG -3'

Primer-reverse:

5'- AUA GGG GAA UGG GCC GUU CAU -3'

primer (RNA extract):

primer1 (unmodified U):

5'- GCA GAC GTT CGA ATG GGT CGT CGC -3'

primer1 (unmodified U) - rev:

5´-CGC CGG CGG CTT TGG TGA CTC-3´

Primer2 (unmodified U):

5'- AGT TCG ACC GTC TTC TCA GCG CTC CG -3'

Primer2 (unmodified U) - rev:

5´- CCG TCG CTA CTA CCG ATT GGA TGG TTT AG -3´

Primer3 (Ψ 966):

5'- GTC CGT CTT GCG CCG GTC CAA G -3'

Primer3 (Ψ 966) - rev:

5´- AAG AGG GAC GGC CGG GGG CA -3´

Primer3 (Ψ 1081):

5'- CGC ATC GCC GGT CGG CAT CGT TT -3'

Primer3 (Ψ 1081) - rev:

5´- GGT TCG AAG ACG ATC AGA TAC CGU CGT AG -3´