

# "Functionalization and Detection of RNA and its Modifications"

Dissertation  
zur Erlangung des Grades

"Doktor der Naturwissenschaften"

im Promotionsfach Pharmazie

am Fachbereich Chemie, Pharmazie und Geowissenschaften  
der Johannes Gutenberg-Universität  
in Mainz

vorgelegt von Stefanie Kellner  
geb. am 06.06.1984 in Schorndorf

Mainz, den 04. Oktober 2012

1. Berichterstatter: [REDACTED]

2. Berichterstatter: [REDACTED]

3. Berichterstatter: [REDACTED]

Datum der mündlichen Prüfung: [REDACTED]



**Danksagung**

[Redacted text block]

[Redacted text block]

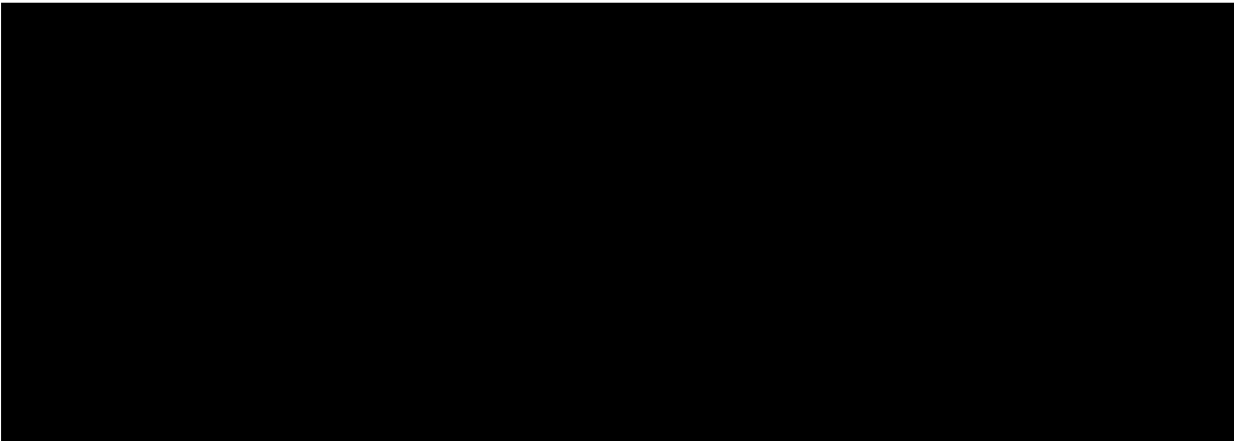
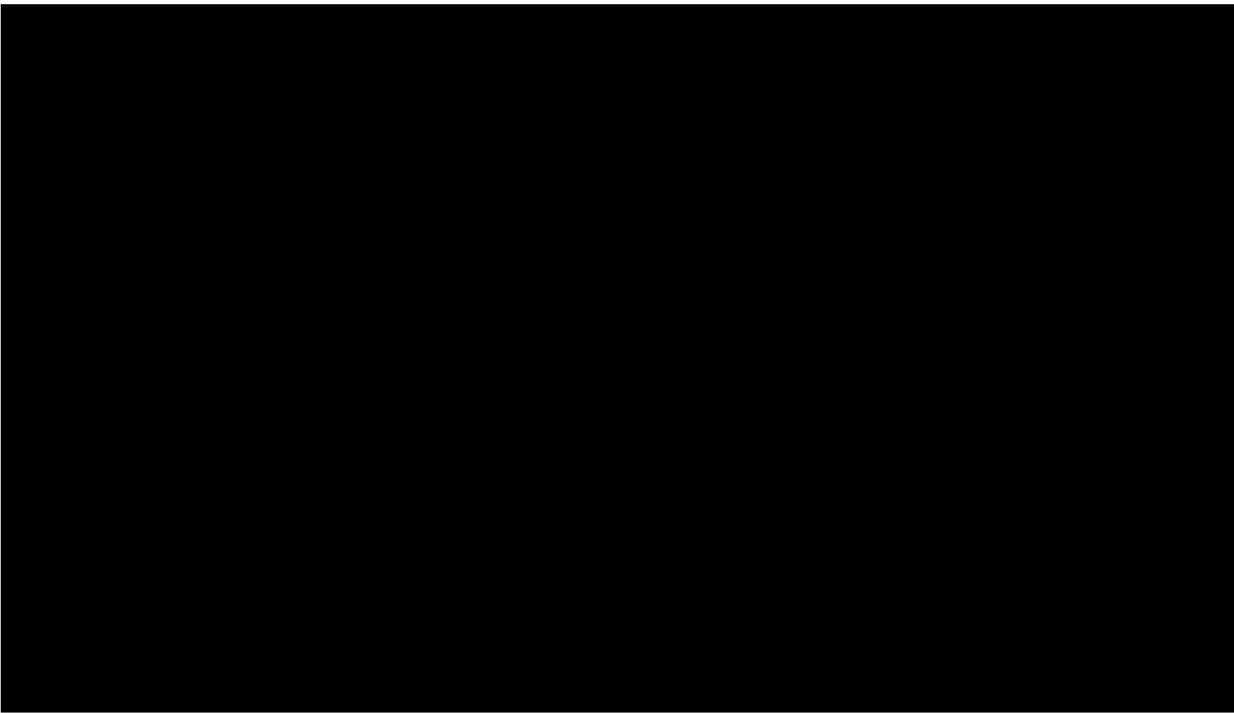
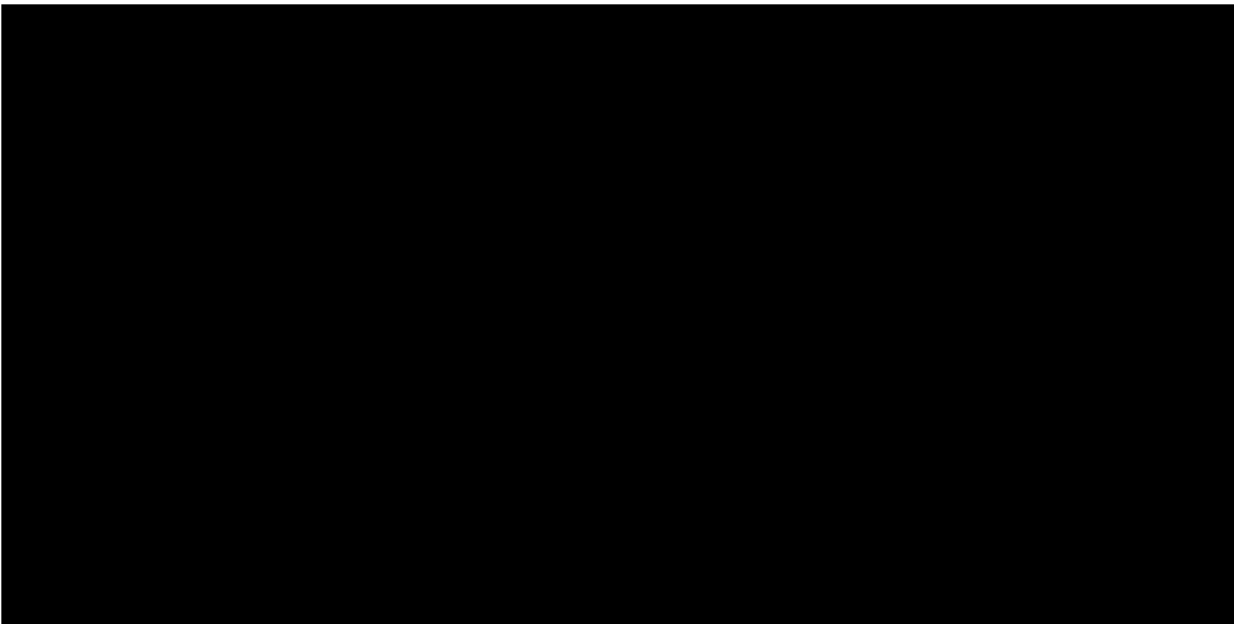
[Redacted text block]

[Redacted text block]

[Redacted text block]

[Redacted text block]







## Abstract

Ribonucleic acids (RNA) are one of the principle biomacromolecules in the life sciences, which are intensively investigated as they represent the central link between two other major biomacromolecules, namely DNA and proteins. To gain more information on RNA, functionalization and labeling techniques for detailed analysis on structure and function are necessary and several techniques have been developed over the past decades.

This work presents a new labeling strategy, based on a small multifunctional chemical reagent, which is specific for uridines in RNAs. This coumarin-based reagent, termed N3BC, has the advantage of: (I) being applicable post-transcriptionally to all kinds of RNAs, (II) displaying fluorescence and (III) having a further functional group for further bioconjugate reactions. The latter comprise, e.g. UV triggered crosslinking experiments with cognate proteins, as well as the bioorthogonal CuAAC reaction with fluorescent alkyne-dyes.

RNAs contain a substantial amount of chemical derivatives of the four major ribonucleosides to facilitate the plethora of RNA functions within the cell. The detection of modified ribonucleosides is challenging and only a few specific reagents that target these modifications have been reported. Therefore, differently substituted reagents, based on the coumarin scaffold, were examined and structure-function relationship studies revealed selectivity for a naturally occurring modified nucleoside, namely 4-thiouridine ( $s^4U$ ). These studies were mainly done by LC-MS/MS analysis of the coumarin products in tRNA samples. During the course of these experiments, another multifunctional coumarin, named PBC, stood out for several reasons. Except for its 2000 fold selectivity for  $s^4U$  over uridine, PBC has an additional terminal alkyne group for subsequent conjugation reactions with azides. It was furthermore used for fluorescent labeling of small interfering RNA, which was successfully transfected and the fluorescence inside the cells was monitored. With PBC, a new chemical reagent for detection of modified nucleosides was added to the available repertoire.

Over 100 other modified nucleosides are described in the literature, most of them being undetectable by small molecules. For reliable detection of these modified nucleosides another principle was applied. The differential physicochemical properties of modified nucleosides were exploited for separation by chromatography in combination with highly sensitive detection by mass spectrometry. In this work, several LC-MS/MS methods for identification and quantification of up to 21 modified ribonucleosides and 5 deoxyribonucleosides in a single run, were developed. In addition, these methods were applied to several studies, mostly on RNA:methyltransferases, where different levels of methylation patterns in the organisms were detected.



## Zusammenfassung

Ribonukleinsäuren (RNA) sind eines der Hauptbiomakromoleküle in den Lebenswissenschaften, die intensiv erforscht werden. Um mehr Informationen über RNA zu erhalten, wurden in den letzten Jahrzehnten Funktionalisierungs- und Labelingtechniken entwickelt, die eine detaillierte Analyse der Struktur und Funktion ermöglichen.

Diese Arbeit zeigt eine neue Labelingstrategie, basierend auf einem kleinen, multifunktionalen chemischen Reagenz, welches spezifisch mit Uridinen in RNA reagiert. Dieses Cumarin-basierte Reagenz, namens N3BC, hat den Vorteil (I) post-transkriptionell gegenüber allen möglichen RNAs einsetzbar zu sein, (II) Fluoreszenz zu zeigen und (III) eine weitere funktionelle Gruppe zu besitzen, die in Biokonjugationsreaktionen einsetzbar ist. Die letzteren umfassen z.B. die durch UV ausgelösten crosslinking Experimente mit verwandten Proteinen, sowie die bioorthogonale CuAAC Reaktion mit fluoreszenten Alkin-Farbstoffen.

RNA enthält eine erhebliche Anzahl von chemischen Abwandlungen der vier Hauptnukleoside, welche die Vielzahl der RNA Funktionen in der Zelle unterstützen. Die Detektion dieser modifizierten Ribonukleoside ist herausfordernd und nur wenige spezifische Reagenzien, die diese Modifikation anzielen, wurden bisher gefunden. Deshalb wurden unterschiedlich substituierte Reagenzien, basierend auf dem Cumarin Körper untersucht und Struktur-Funktions-Beziehungsstudien zeigten eine Selektivität für ein natürlich vorkommendes, modifiziertes Nukleosid, 4-Thiouridine ( $s^4U$ ). Im Verlauf dieser Experimente, fiel ein weiteres multifunktionales Cumarin, namens PBC, aus mehreren Gründen auf. Neben seiner 2000 fachen Selektivität für  $s^4U$  gegenüber Uridin, besitzt PBC ein zusätzliches terminales Alkin für Konjugationsreaktionen mit Aziden. Es wurde zusätzlich zur Fluoreszenzmarkierung von small interfering RNA benutzt, deren Fluoreszenz in Zellen beobachtet werden konnte. Mit PBC kommt ein neues chemisches Reagenz zur Detektion von modifizierten Nukleosiden zum bereits vorhandenen Repertoire hinzu.

Über 100 weitere modifizierte Nukleoside sind in der Literatur beschrieben. Die meisten davon sind allerdings nicht durch kleine Moleküle detektierbar. Für verlässliche Detektion wurden die unterschiedlichen physikochemischen Eigenschaften der modifizierten Nukleoside, zur chromatographischen Trennung in Kombination mit hochsensitiver Detektion durch Massenspektrometrie, ausgenutzt. In dieser Arbeit wurden mehrere LC-MS/MS Methoden, zur Identifizierung und Quantifizierung von bis zu 21 Ribonukleosiden und 5 Deoxyribonukleosiden in einem Einzellauf, entwickelt. Zusätzlich wurden diese Methoden in mehreren Studien, hauptsächlich von Methyltransferasen, welche unterschiedliche Levels in den Methylierungsmustern der Organismen zeigten, angewandt.



# Table of contents

<b>1 INTRODUCTION</b> .....	<b>1</b>
<b>1.1 Functionalization of RNA</b> .....	<b>1</b>
1.1.1 Solid-phase synthesis with functionalized building blocks.....	3
1.1.2 Functionalization by enzymatic techniques .....	4
1.1.3 Chemical functionalization.....	5
1.1.4 Functionality transfer by oligodeoxynucleotides (FTODN) .....	9
<b>1.2 RNA and its modifications</b> .....	<b>11</b>
1.2.1 Nomenclature and structure of RNA modifications.....	11
1.2.2 Function of RNA modifications .....	12
1.2.3 5-methylcytidine (m <sup>5</sup> C).....	13
<b>1.3 DNA and its modifications</b> .....	<b>16</b>
1.3.1 DNA modifications for protection from restriction enzymes .....	16
1.3.2 DNA modifications with epigenetic functions .....	17
<b>1.4 Detection of RNA modifications</b> .....	<b>18</b>
1.4.1 Differential chemical reactivity.....	18
<b>1.5 LC-MS/MS of modified nucleosides</b> .....	<b>22</b>
1.5.1 Basic principles of mass spectrometry of nucleosides .....	22
1.5.2 Basic principles of nucleoside chromatography.....	24
1.5.3 LC-MS(MS) combining methods for RNA and DNA analytics.....	25
<b>2 GOAL OF THE WORK</b> .....	<b>27</b>
<b>3 MATERIALS &amp; METHODS</b> .....	<b>28</b>
<b>3.1 Chemicals and Enzymes</b> .....	<b>28</b>
<b>3.2 Buffers</b> .....	<b>29</b>
<b>3.3 Oligonucleotides</b> .....	<b>30</b>
<b>3.4 Modified nucleosides</b> .....	<b>31</b>
<b>3.5 Methods</b> .....	<b>32</b>
3.5.1 Coumarin derivatization .....	32
3.5.2 Molecular biological methods .....	33
3.5.3 Chromatographic parameters for nucleoside analysis.....	35
3.5.4 MS parameters.....	35
<b>3.6 The RNA methyltransferase Dnmt2 methylates DNA in the structural context of a tRNA</b> .....	<b>36</b>

<b>3.7 rRNA:5-methylcytosine-methyltransferase (rRNA:m<sup>5</sup>C-MTase) activity of yeast proteins Nop2, Ynl022 and human proliferation-associated antigen p120.....</b>	<b>36</b>
3.7.1 Calibration for quantification of m <sup>5</sup> C in rRNA.....	36
3.7.2 Detection and analysis of calibration samples.....	36
3.7.3 Calculations for quantification of m <sup>5</sup> C in rRNA.....	37
3.7.4 Calculations for quantification of injected rRNA.....	37
3.7.5 Calculation for m <sup>5</sup> C quantification.....	37
<b>4 RESULTS.....</b>	<b>38</b>
<b>4.1 General analytics of tRNA <i>E.coli</i> treated with monofunctional coumarins .....</b>	<b>38</b>
4.1.1 Gel assay of conjugated tRNA.....	39
4.1.2 LC-UV and LC-MS/MS analysis.....	40
4.1.3 LC-MS/MS data analysis.....	42
<b>4.2 Structure-function relationship of monofunctional coumarins .....</b>	<b>45</b>
4.2.1 Influence of the substitution pattern on reactivity and selectivity.....	45
4.2.2 Influence of the reaction conditions on reactivity and selectivity.....	45
<b>4.3 A general RNA alkylating reagent with multifunctional properties – N3BC.....</b>	<b>47</b>
4.3.1 Reactivity with nucleosides and homopolymers.....	48
4.3.2 Reactivity with artificial bipartite oligonucleotides.....	49
4.3.3 LC-MS analysis of treated RNA.....	49
4.3.4 Modulation of labeling efficiency by stepwise change of reaction conditions.....	51
4.3.5 Click reaction of successfully labeled RNA.....	52
4.3.6 Crosslinking reaction with cognate proteins.....	52
4.3.7 Reactivity towards DNA.....	53
<b>4.4 PBC – a new multifunctional coumarin selective for 4-thiouridine .....</b>	<b>54</b>
4.4.1 Analysis of treated tRNA.....	55
4.4.2 Comparison with examined monofunctional coumarins.....	56
4.4.3 Click reaction of PBC labeled RNA.....	57
4.4.4 Ring opening of uridine-PBC conjugates upon UV irradiation.....	58
4.4.5 Behavior of monofunctional coumarins upon UV irradiation.....	60
<b>4.5 Application of coumarins .....</b>	<b>62</b>
4.5.1 Preparation of PBC-atto647N labeled siRNA.....	62
4.5.2 Hybridization and FRET experiments of the PBC-Atto647N labeled siRNA.....	63
4.5.3 Cell imaging studies of PBC-atto647N labeled siRNA.....	65
<b>4.6 LC-MS/MS detection of modified nucleosides.....</b>	<b>68</b>
4.6.1 Method development for modified ribonucleosides.....	69
4.6.2 Method development for modified deoxyribonucleosides.....	74
4.6.3 Limit of detection (LOD) and limit of quantification (LOQ) for selected nucleosides	75
4.6.4 General principles of LC-UV-MS/MS quantification of modified nucleosides.....	77
<b>4.7 Investigations on RNA and DNA supported by developed LC-MS methods.....</b>	<b>79</b>
4.7.1 Related haloarchaeal pleomorphic viruses contain different genome types.....	80
4.7.2 Site-specific cytosine-5 methylation by the mouse RNA methyltransferases Dnmt2 and NSun2 promotes tRNA stability and protein synthesis.....	82



4.7.3	The RNA methyltransferase Dnmt2 methylates DNA in the structural context of a tRNA.....	84
4.7.4	rRNA:5-methylcytosine-methyltransferase (rRNA:m <sup>5</sup> C-MTase) activity of yeast proteins Nop2, Ynl022 and human proliferation-associated antigen p120 .....	86
<b>5</b>	<b>SUMMARY &amp; DISCUSSION.....</b>	<b>88</b>
<b>6</b>	<b>CONCLUSION &amp; OUTLOOK.....</b>	<b>91</b>
<b>7</b>	<b>REFERENCES .....</b>	<b>93</b>
<b>8</b>	<b>PUBLISHED ARTICLES &amp; REVIEWS .....</b>	<b>101</b>
<b>9</b>	<b>APPENDIX .....</b>	<b>225</b>
	<b>CURRICULUM VITAE .....</b>	<b>229</b>

## List of figures

Figure 1 Basic principle of the Cu(I) dependent azide-alkyne Huisgen cycloaddition. ....	6
Figure 2 General reaction scheme for the functionality-transfer reaction of an oligodeoxynucleotide (FTODN adjusted from [46]).....	9
Figure 3 Common modifications in RNA named after the established nomenclature. ....	12
Figure 4 Basic principles of bisulfite sequencing. ....	14
Figure 5 Overview and biosynthetic relation of the four epigenetic DNA modifications. ....	17
Figure 6 Overview of reagents for detection of modified nucleotides.....	19
Figure 7 Basic design of MS/MS detection systems demonstrated on an ESI-triple quadrupole MS. ....	23
Figure 8 Overview of the used monofunctional coumarins and the general reaction schemes with nucleosides uridine and 4-thiouridine. ....	38
Figure 9 10% PAGE of total tRNA <i>E.coli</i> treated with 6 monofunctional coumarins under 2 sets of conditions. ....	39
Figure 10 LC-UV-MS analysis of tRNA conjugated with BMB.....	41
Figure 11 General scheme for data acquisition and analysis of coumarin selectivity studies. ..	42
Figure 12 Processed LC-MS/MS data of all six monofunctional coumarins.....	46
Figure 13 Chemical structure of N3BC .....	47
Figure 14 Analysis of N3BC treated bipartite oligomers.....	50
Figure 15 Modulation of labeling efficiency by changing either the DMSO concentration or the N3BC concentration. ....	51
Figure 16 Urea-gels of bipartite oligonucleotides of RNA composition (left) and DNA composition (right). ....	53
Figure 17 Chemical structure of PBC .....	54
Figure 18 Gel analysis of total tRNA <i>E.coli</i> treated with all 8 coumarins, monitoring the coumarin fluorescence under UV illumination at $\lambda=365$ nm .....	55
Figure 19 Diagram of processed data deriving from LC-MS/MS analysis of coumarin treated tRNA digests. ....	56
Figure 20 Urea-PAGE of Atto647N-azide treated oligonucleotides (ON). ....	57
Figure 21 Analysis of UV illuminated PBC conjugated samples. ....	59
Figure 22 LC-MS analysis of coumarin treated tRNA <i>E.coli</i> before and after UV irradiation	61
Figure 23 Fluorescence spectra of commercially and PBC clicked atto647N siRNA hybridized with the complementary alexa555 sense strand done by Markus Hirsch. .	64
Figure 24 Cell imaging studies of PBC-atto647N labeled siRNA uptake performed by Markus Hirsch. ....	67
Figure 25 Chromatogram of 25 ribonucleosides with the gradient nucs 4. ....	69
Figure 26 Fragmentation pattern of ribonucleosides. ....	71
Figure 27 MS/MS spectra of monomethylated cytidines $m^3C$ , $m^5C$ and Cm and their corresponding mass peaks in the chromatogram from the method nucs4_DMRM. ....	73
Figure 28 Chromatogram and exemplaric fragmentation pattern of deoxynucleosides. ....	75
Figure 29 Calibration curves for ribo- and deoxy-5-methylcytidine with indicated LOD and LOQ.....	76
Figure 30 Possibilities for relative quantification of modified nucleosides.....	78
Figure 31 LC-UV/MS analysis of viral DNA .....	81
Figure 32 Analysis of $m^5C$ content in murine tRNA by bisulfite sequencing and LC-MS analysis. ....	83
Figure 33 Methylation efficiency of all-ribo and dC38 tRNA <sup>Asp</sup> analyzed by $^3H$ -SAM incorporation assay and LC-MS/MS analysis. ....	85
Figure 34 Analysis of rRNA yeast by LC-MS/MS analysis. ....	87

## List of tables

Table 1 RNA functionalization techniques..	10
Table 2 Overview of RNA:m <sup>5</sup> C-MTase families and some of their known and putative substrates ..	15
Table 3 Commercial 20 mer oligonucleotides of bipartite composition ..	30
Table 4 Other oligonucleotides used for coumarin derivatization and LC-MS/MS analysis ..	30
Table 5 Exemplary pipetting scheme for coumarin derivatization under reaction conditions 1 ..	32
Table 6 Exemplary pipetting scheme for coumarin derivatization under reaction conditions 2 ..	32
Table 7 Exemplary pipetting scheme for coumarin derivatization under reaction conditions 3 ..	33
Table 8 Exemplary pipetting scheme for the CuAAC (click) reaction ..	33
Table 9 Composition of differently concentrated polyacrylamide gels ..	34
Table 10 Protocol for digestion of RNA to nucleotides ..	34
Table 11 Digestion from nucleotides to nucleosides.....	34
Table 12 Chromatographic parameters used for coumarin and nucleoside analytics ..	35
Table 13 Calculation of adjusted areas for BMB conjugates ..	43
Table 14 Determination of the correction factors for BMB conjugates ..	43
Table 15 Calculation of corrected areas for BMB conjugates ..	43
Table 16 determination of nucleoside composition of total tRNA <i>E. coli</i> ..	44
Table 17 Abundance of coumarin conjugates considering tRNA composition ..	44
Table 18 MS/MS transitions of all 21 analyzed modified nucleosides and retention times ....	72
Table 19 All nucleosides analyzed in method dnucs2 with MS/MS transitions and retention times.....	74
Table 20 LOD and LOQ of several modified nucleosides ..	76
Table 21 Overview of RNA:m <sup>5</sup> C-MTase families and their known and putative substrates ..	90

## Abbreviations

AaRS	Amino acyl tRNA synthetase
CuAAC	Copper catalyzed azide-alkyne cycloaddition
dko	Double knock-out
DMSO	Dimethylsulfoxid
FRET	Fluorescence resonance energy transfer
HPLC	High performance liquid chromatography
i <sup>6</sup> A	N6-isopentenyladenosine
ko	Knock-out
LC-MS/MS	Liquid chromatography coupled with tandem mass spectrometry
LOD	Limit of detection
LOQ	Limit of quantification
m/z	Mass to charge ratio
NTP	Nucleosidetriphosphate
ODN	Oligodeoxyribonucleotide
ON	Oligonucleotide
ORN	Oligoribonucleotide
PAGE	Polyacrylamid gel electrophoresis
Q	queuosine
RNA:m <sup>5</sup> C-MTase	RNA:5-methylcytidine Methyltransferase
S/N	Signal to noise ratio
t <sup>6</sup> A	N6-threonylcarbamoyladenosine
UPLC	Ultra performance liquid chromatography
yW	wybutosine
Ψ	Pseudouridine
Dnmt2	DNA methyltransferase 2





# 1 Introduction

## 1.1 Functionalization of RNA

Ribonucleic acids (RNA) are one of the principle biomacromolecules investigated in the life sciences as they represent the central link to the other major biomacromolecules, namely DNA and proteins. RNA connects the translation of the genetic code stored as DNA into proteins in several ways: messenger RNA (mRNA) acts as the transient information keeper while the genetic code is translated with the help of transfer RNA (tRNA) in the translational machinery of the ribosome with ribosomal RNA (rRNA) guiding both mRNA and tRNA in the requested location of the ribosome. Apart from protein biosynthesis, RNA was discovered to play an important role in cell regulation *e.g.* silencing of genes with microRNA (miRNA). To gain more information on this versatile macromolecule, functionalization and labeling techniques for detailed analysis on structure and function are necessary. The most prominent functionalization approaches go back to the 1970s when Maxam and Gilbert [1] used functionalization of certain DNA and later RNA [2] bases, followed by cleavage at the functionalized site for gel-electrophoretic sequencing. Additionally, this approach requires a radioactive label either at the 3' or the 5' end and samples with homogeneous ends at the labeling position. In contrast, Sanger sequencing used radioactively labeled oligodeoxynucleotides (ODNs) annealed as primers on the target DNA followed by DNA synthesis by DNA polymerase. The DNA synthesis was performed in four different tubes each containing one of the four reaction terminating dideoxynucleotides (ddNTPs). The usage of the ddNTPs generated short fragments of different size, but with a homogeneous 5' end based on the radioactive primer. The 3' ends were determined by the incorporated ddNTP and the sequence could be read by gel-electrophoretic separation of the fragments of all four tubes [3]. Later automated DNA sequencing was achieved by usage of four pre-functionalized ddNTPs with different fluorescent dyes, which allowed sequence reading by colour annotation [4].

Especially the incorporation of fluorescent labels into RNA improved the detection of trace amounts of RNA and facilitated not only sequencing of RNA, but information on structure and structural dynamics or hybridization with other nucleic acids in RNA probing experiments.

Structural probing is used for determination of secondary-structure and uses, like Maxam and Gilbert sequencing, a radioactive end-label for high sensitivity and internal strand functionalization. In addition, chemical functionalization can be combined with enzymatic probing techniques. Structural probing was extensively used *e.g.* for studies on secondary

structure of RNAs [5] and interactions of tRNAs with mRNA [6] or with aminoacyl synthetases [7]. A more detailed overview on these techniques will be given in the following chapters.

In addition, functionalization of RNA is desirable for RNA immobilization experiments, which are performed by conjugation of target RNAs with biotin. Lipid functionalized RNA can be used for cell-specific delivery and uptake of the RNA [8]. Furthermore, the specific labeling of naturally occurring RNA modifications enables a sensitive detection of modified nucleotides and will be discussed in more detail in chapter 1.4.

Depending on the type of RNA study, different functionalization approaches have been developed and can be divided in 4 classes. The first class is solid-phase synthesis, which is used to obtain RNA of precisely defined sequences including functionalized nucleotides at predetermined positions. A second class comprises all techniques that require enzymes for incorporation of functional groups either by labeling of the 3' or 5' end, by coupling of target RNAs with other functionalized RNAs (ligation) or by transcription in the presence of functionalized NTPs. The third class is chemical functionalization which can be subdivided in at least two research fields:

- (a) Post-synthetic or post-transcriptional reactions of incorporated nucleotides with functional groups.
- (b) Functionalization of native RNA with small molecules as applied in structural probing or detection of modified nucleotides.

The last class is a recently developed and already well studied functionalization approach using site-specific group transfer by an oligodeoxynucleotide probe.



### 1.1.1 Solid-phase synthesis with functionalized building blocks

Solid-phase synthesis of RNA, using modified phosphoramidite building blocks, was first established by Gait 1991 [9] by site-specific incorporation of bromodeoxyuridine (BrdU) in an oligoribonucleotide (ORN). This approach has several advantages, for example, site-specificity and relatively easy production of sufficient RNA amounts for further analysis like NMR studies. Limitations, on the other hand, lie in the potential to obtain the modified building block. Fluorophores as building blocks are commercially available and may be incorporated for example in the last step of synthesis at the 5' end [10, 11]. 3' labeling of the desired oligonucleotide can be achieved by usage of already modified solid phases [12]. If a desired functional group is not suitable for incorporation in solid-phase synthesis, either due to its size, or chemical sensitivity, a second option remains. Instead of the fully functionalized nucleoside a building block with a chemical moiety suitable for post-synthetic labeling procedures is incorporated. Such functional groups are mostly thiols, primary amines for application of NHS chemistry and, recently terminal alkynes or azides for the Cu(I) dependent azide-alkyne cycloaddition (CuAAC or click) reaction. Since azides and phosphoramidites undergo a spontaneous chemical reaction, azides are not compatible with the standard phosphoramidite solid-phase chemistry. Therefore, direct incorporation of azides in ORN was only possible at the 5' end [13]. Only recently, the Micura lab presented a solution for this obstacle by synthesizing an azide functionalized phosphodiester building block. They additionally showed successful incorporation of the new building block in standard automated phosphoramidite solid-phase RNA synthesis [14, 15].

A detailed overview on post-synthetic functionalization will be given in chapter 1.1.3.

### 1.1.2 Functionalization by enzymatic techniques

#### a) RNA end-labeling

Radioactive labeling was one of the crucial developments to boost research on RNA in the late 1960s, since it abolished the need for huge amounts of isolated, pure RNA per experiment. T4 polynucleotide kinase (PNK) extracted from *E.coli* was found to attach a radioactive phosphate with  $\gamma$ - $^{32}\text{P}$ -ATP to the free 5'OH of an oligodeoxynucleotide [16]. T4 RNA ligase commonly used for 3' end-labeling does not only accept [ $\alpha$ - $^{32}\text{P}$ ]-ATP as a substrate but additionally biotin- and fluorescein- adenosylpyrophosphates have been tested and successfully incorporated at the 3' end of the RNA molecule [17].

#### b) Ligation of functionalized RNA

Ligation of nucleic acids is the enzymatic conjugation of two nucleic acids. For ligation a free 3'OH is needed on the one strand, while the other nucleic acid must have a phosphorylated 5' end. However, the conjugation reaction is only successful if the free 3'OH and the 5' phosphate of the two RNAs are brought in spatial proximity. Most commonly used is T4 DNA ligase, which facilitates the conjugation by using a DNA splint reverse complementary to the RNA regions that are to be connected, bringing them in the required proximity. In case of tRNA, which forms strong secondary interactions, T4 RNA ligase can be used without a DNA splint. The target tRNA fragments form strong stem-loop-stem interactions, which allows ligation if the cleavage is in the anticodon loop region. While the ligation step itself is not a “functionalization” of the RNA, it is commonly used to link unmodified RNA, like transcripts, with functionalized RNA from solid-phase-synthesis. *E.g.* constructs for FRET (fluorescence resonance energy transfer) experiments are mostly generated by splinted-ligation to reduce costs for the expensive fluorophore-labeled oligonucleotides [18]. Another prominent example for the use of ligation is barcode-labeling with DNA-adapters in new generation sequencing approaches.

#### c) Incorporation of functional groups during transcription

RNA is frequently synthesized by T7 RNA polymerase using NTPs and a DNA template containing the sequence information. It was observed that T7 RNA polymerase accepts modified NTPs with small reactive groups at the 2'OH of the ribose, which allows statistical distribution of modifications in the RNA. Some of these introduced modifications can be used directly like 2'thiols that are used in RNA interference mapping [19] by iodine cleavage. Others like primary, aliphatic amines [20], terminal alkynes or azides need a further post-synthetic derivatization step. It is noteworthy that transcription with “clickable” nucleotides is also successful *in vivo*, as shown for 5-ethynyluridine [21] and 2'-azidouridine [22].

### 1.1.3 Chemical functionalization

Chemical functionalization of RNA is possible with functional groups that occur either naturally in the RNA major nucleosides, their natural derivatives or functionalities previously introduced by synthesis or transcription. Here, post-synthetic functionalization covers all RNAs produced either by solid-phase synthesis or transcription with pre-modified nucleotides. RNA of all other sources is referred to as “native” RNA, since most of the techniques described were used on RNAs extracted and purified from different organisms.

#### *a) Post-synthetic functionalization*

Typical scaffolds used for RNA pre-functionalization by solid-phase synthesis and transcription are primary amines which selectively react with NHS esters. Since a few years, terminal alkynes or azides for the CuAAC reaction have moved into the focus of RNA pre-functionalization techniques.

#### **NHS chemistry:**

NHS chemistry comprises labeling reactions of primary amines with N-hydroxysuccinimide esters and is widely used in protein studies since the natural amino acid lysine contains a primary amine. Primary amines are not found in unmodified RNA and only some modifications like 3-(3-amino-3-carboxypropyl)uridine (acp<sup>3</sup>U) contain a primary amine and can therefore be labeled by NHS reagents [23]. Thus, reactions of synthetically produced RNAs that are pre-modified with an amine are specific for labeling with NHS reagents, while other parts of the RNA stay unaltered. As many NHS reagents are readily available, they have been successfully used for fluorescent [24] and radioactive [25] labeling of RNA.

**CuAAC chemistry:**

The Cu(I)-dependent azide-alkyne Huisgen 1,3-dipolar cycloaddition (CuAAC) reaction is one of the recently developed so called “click reaction” strategies for bioconjugation reactions with high popularity due to its orthogonality. “Bioorthogonal chemical reactions, whose components must react rapidly and selectively with each other under physiological conditions in the presence of the plethora of functionality necessary to sustain life(…)” [26], gave a new boost for research in the life sciences. The chemical principle of the most common triazol ring producing cycloaddition is shown in Figure 1. Click reactions on RNA *in vivo* were first successfully used by Jao and Salic (2008) by transcriptionally introducing 5-ethynyluridine into RNA followed by azide-fluorophore labeling of the newly transcribed RNA [21]. Soon other reports of azide-functionalized RNAs that also react readily with terminal-alkyne fluorophores followed [14, 22, 27]. Additionally, synthetic oligonucleotides with terminal alkynes are commercially available, which would allow easy synthesis of functionalized RNAs for probing experiments. One crucial obstacle, of the otherwise fast and effective click reaction, is the cytotoxicity of the Cu(I) catalyst. It is therefore necessary to reduce the Cu(I) concentration of the reaction, which was recently achieved by using copper-chelating azides that “bring their own copper” to the alkyne reaction partner [28]. Another possibility is the use of strained, cyclic alkynes instead of terminal alkynes that do not require the Cu(I) catalyst [29] which was already applied to oligoribonucleotides [30].

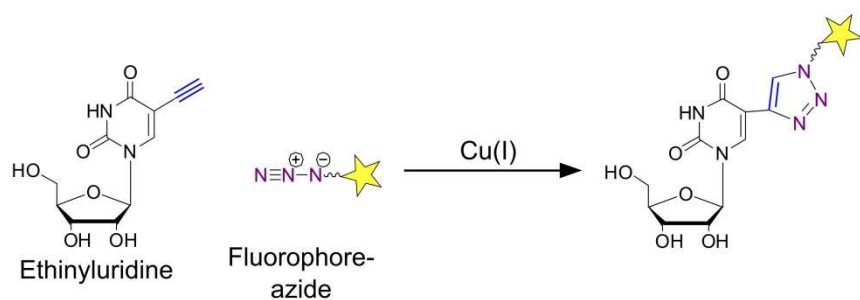


Figure 1 Basic principle of the Cu(I) dependent azide-alkyne Huisgen cycloaddition.

### ***b) Functionalization of native RNA***

In the early years after discovery of nucleic acids, chemical probing was a wide-spread way to gain structural information on the new macromolecules. This resulted in a substantial repertoire of well investigated chemicals and reaction conditions for DNA and RNA. All reagents can be divided in 3 groups:

- 1) Electrophiles and alkylating reagents reacting with the nucleophilic nitrogens of the nucleobases, phosphate groups or ribose-2'-OH.
- 2) Oxidizing reagents that attack nucleophilic sites of the RNA.
- 3) Nucleophiles that attack electron-poor positions in RNA.

The reactions are mostly performed *in vitro* in narrowly defined chemical environments of controlled pH and solvent composition to ascertain target specificity. Some reagents react specifically with uncommon functionalities presented by modified nucleotides which are exploited for detection of modified nucleotides and in structural probing experiments. An extended overview of specific reagents for detection of modified nucleotides will be given in chapter 1.4.1.

Other reagents react with functional groups present in the major nucleotides of RNA, which makes them useful for studies of native RNA and additionally of transcripts containing only the canonical bases. Since nucleotides contain several functional groups, the developed reagents can be grouped by their specific target:

#### **The nucleobase:**

Commonly used reagents, for nucleobase functionalization, are the acylating agent DEPC (diethylpyrocarbonat), which reacts with *N7* of adenosine and the alkylating agent DMS (dimethylsulfate) that preferentially alkylates the *N7* position of guanosine [31]. With these exemplaric reagents it was possible to examine the secondary and tertiary interaction of RNA molecules in structural probing studies, since only sterically accessible nucleotides could react with the probe.

A prominent used nucleophilic reagent that reacts with the C6 of cytosine is sodium bisulfite, which facilitates the conversion of cytosine to uracil (see Figure 4). This reaction does not occur with the C5 methylated nucleoside  $m^5C$ , which can be therefore detected by sequencing approaches.

Fluorescent labeling of the common nucleobases was achieved by alkylation of *N1* of adenosine with 9-bromomethylantracene [32], the guanosine 2-amino group by an epoxy-tetrahydrobenzo[a]pyrene derivative [33], the cytidine 4-amino group after bisulfite conversion

with a fluorescent nitrobenzofurazan [34] and uridine *N3* with 7-methoxy-4-bromomethylcoumarin [35].

### **The phosphate:**

ENU (ethylnitrosourea) was applied for structural probing experiments, as well, since it alkylates free phosphodiester not involved in tertiary interactions [36]. Like ENU the photochemical arenediazonium cations can be used for labeling of phosphodiester in single-stranded loop regions. However, these ethylnitrosourea substitutes are first activated by soft UV-light to produce instable nitrene intermediates that immediately react with nucleophiles in their proximity – the phosphate groups of RNA [37].

### **The ribose:**

Periodate oxidation of the ribose at the 3' end to a dialdehyde allows *e.g.* fluorescent labeling of RNA by subsequent coupling of the oxidized ribose with hydrazine derivatives of fluorophores [38] or even tritium incorporation by reaction with tritium-labeled borohydride [39].

The oxidizing reagent Fe(II)-EDTA reacts with hydrogen peroxide producing hydroxyl radicals, which react with ribose of the RNA backbone used to reveal solvent accessibility [40].

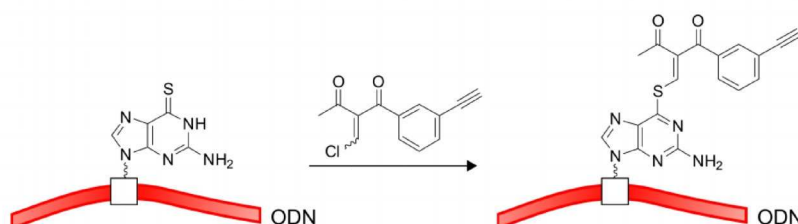
An electrophilic reagent targeting the 2'OH of ribose was discovered in 2004 and used for RNA structure analysis at single nucleotide resolution by selective 2'-hydroxyl acylation and primer extension (SHAPE) [41]. Like bisulfite treatment, SHAPE was successfully combined with modern sequencing technologies and the resulting SHAPE-Seq analysis allows fast and reliable generation of sequence and structural information [42].

Usage of such specific reagents allowed identification of nucleotide sites involved in hydrogen bond formation in tRNA structure and in combination with sequence specific nucleases like RNase T1 an even more reliable prediction of RNA structure is possible [43].

### 1.1.4 Functionality transfer by oligodeoxynucleotides (FTODN)

A site-specific way to introduce small molecules into native RNA is a functionality-transfer reaction of an oligodeoxynucleotide (FTODN) probe with the target RNA [44]. The ODN probe is synthesized with a sequence complementary to the target RNA and with a 6-thio-guanosine opposite to the targeted RNA nucleotide. In a post-synthetic reaction the thio-group of the ODN is functionalized with the group that is supposed to be transferred. The ODN is then hybridized with the target RNA to trigger the selective functionalization of the 4-amino group of the cytosine base at pH 7 or the 2-amino group of the guanine base at pH 9.4 as can be seen in Figure 2 or at pH 7.4 in the presence of  $\text{NiCl}_2$ . The method has been improved over the last few years to allow the transfer of fluorescent groups [45] and even of a terminal alkyne for further functionalization by the aforementioned click reaction [46].

#### 1. Preparation of ODN by solid-phase synthesis and subsequent post-synthetic functionalization



#### 2. Hybridization with target RNA and group transfer

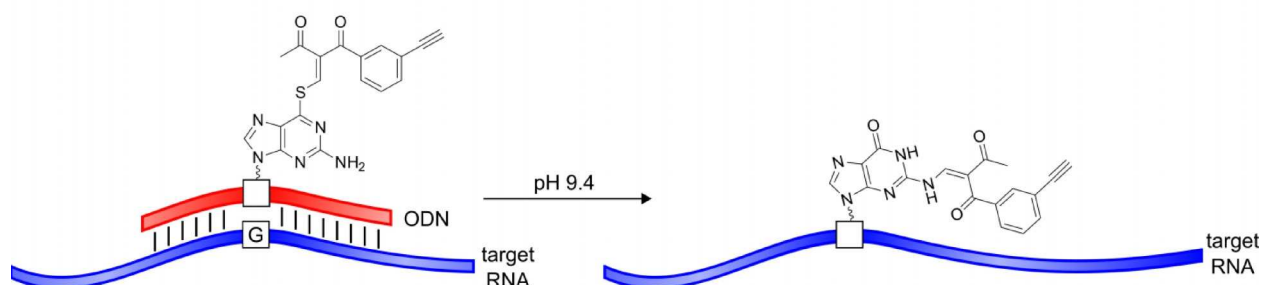


Figure 2 General reaction scheme for the functionality-transfer reaction of an oligodeoxynucleotide (FTODN adjusted from [46])

An overview of all RNA functionalization techniques can be found in Table 1. The column on the right highlights all published studies that combine the aforementioned techniques with the CuAAC reaction.

Table 1 RNA functionalization techniques. Advantages are highlighted in blue, disadvantages in red.

technique	Sequence specific functionalization	Sequence information needed	Applicable to native RNA	Combined with CuAAC reaction
1. Solid phase synthesis	yes	-	no	yes [14]
2. Enzymatic techniques				
a) end-labeling	yes	no	yes	no
b) ligation	yes	no	yes	[13]*
c) transcription	no	no	yes	yes [21]
3. Chemical functionalization				
a) post-synthetic	yes	-	RNA with appropriate modifications	yes [15, 21]
b) small molecules	no	no	yes	no
4. FTODN	yes	yes	yes	yes [46]

\*chemical ligation is possible using click reaction instead of phosphodiester bonds.

As the summarizing table above shows, every technique has its specific advantages (blue) and limitations (red). For example, small molecules on the one hand do not allow a sequence specific introduction of the label (disadvantage), but on the other hand a previous determination of the RNA sequence is not necessary, since the label is distributed statistically in the presence of several targets (advantage). In this case it is noteworthy that small molecule functionalization does not yet allow application of CuAAC reactions.



## 1.2 RNA and its modifications

The monomeric units of RNA are nucleosides, consisting of a ribose sugar connected to one of the four major nucleobases adenine, guanine, cytosine or uracil through a C1'-N1 glycosidic bond in case of pyrimidines, or a C1'-N9 bond in case of purines, respectively. Nucleosides carrying a phosphate are called nucleotides and the formation of 3'-5' phosphodiester bonds connect the nucleotides to long chains of RNA.

In most RNAs modified nucleosides, that are chemical derivatives of the four major nucleosides, can be found. Until today 107 different naturally occurring nucleosides have been published [47] but even more new modified nucleosides are currently identified and will be added to the published list in the near future. However, the distribution of these modifications varies from species to species and from RNA to RNA depending on the complexity of the required RNA function.

tRNA, for example, is an RNA that requires elevated stability and structure to interact with its aminoacyl-tRNA-synthetases (AaRS) and the ribosome. To obtain the needed stability and tertiary structure, tRNA requires an extended chemical diversity for three-dimensional interactions and therefore numerous modified nucleosides can be found in this type of RNA [48]. The following chapter will give a brief overview of the established nomenclature of modifications, the reported function of modifications, and one exemplary biosynthesis of a common RNA modification, 5-methylcytidine.

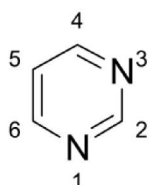
### 1.2.1 Nomenclature and structure of RNA modifications

For nomenclature of RNA modifications the abbreviations C, U, G and A of the four major nucleosides give the basis of the modified nucleosides name. Most modifications are chemical groups, like methyl- or thio-groups attached to the nucleoside and are abbreviated with a one- or two-letter code (like m- for methyl- or s- for thio-). If the modification is placed on the 2'OH of the ribose the letter is placed after calling the base and before the base if the modification is on the nucleobase. The position on the nucleobase is numbered according to the common nomenclature of purines and pyrimidines and indicated by superscript after the modifications abbreviation (see Figure 3A and B). If the modification is on the Hoogsteen face, like the methyl-group of m<sup>5</sup>C its base-pairing properties are not altered compared to modifications at the Watson-Crick site (see m<sup>3</sup>C). More complex modifications like pseudouridine or queuosine are not named after the major nucleoside they biosynthetically originate from, but own abbreviations like Ψ or Q were assigned. Pseudouridine, the isomerization product of uridine, is the most abundant and among all discovered modified nucleosides the only nucleoside that has a carbon-

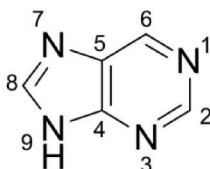
carbon glycosidic bond. For a more detailed overview of RNA modifications the recently published review of Thomas Carell [49] or the complete list at the RNA modification database [47].

### A Numbering of nucleobases

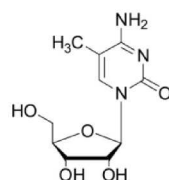
Pyrimidine



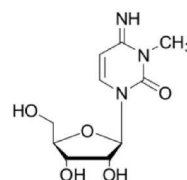
Purine



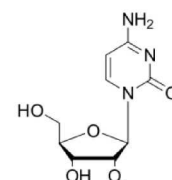
### B Common modified ribonucleosides



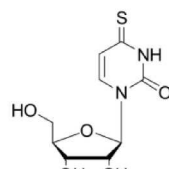
5-methylcytidine  
 $m^5C$



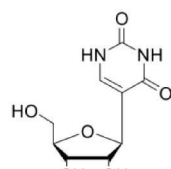
3-methylcytidine  
 $m^3C$



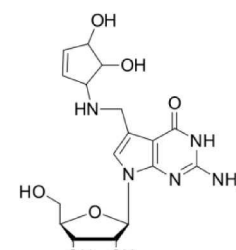
2-O'-methylcytidine  
Cm



4-thiouridine  
 $s^4U$



pseudouridine  
 $\Psi$



queuosine  
Q

Figure 3 Common modifications in RNA named after the established nomenclature.

### 1.2.2 Function of RNA modifications

RNA molecules fulfill a multitude of functions within a cell. This diversity of functions is most likely one of the reasons why so many modified ribonucleosides can be found within a cell. They build a second information layer defined by the position of the modified nucleoside and the chemical structure of the modification.

Most modifications can be found in tRNA (~50% of all modifications) with up to 8 modified nucleosides per molecule. In tRNA, modifications can be divided in three groups corresponding to their function. The first group comprises the mostly complex modified nucleosides like Q or  $t^6A$ ,  $i^6A$  or  $yW$  in the anticodon loop at positions 37 and 34. These modifications are directly involved in the decoding process and influence the conformational flexibility of the anticodon loop [50, 51].

The second group of modifications is involved in the correct folding and stabilization of the complex three-dimensional structure of tRNA. It was shown that Am [52] and  $\Psi$  [53] are commonly found in the helical regions of the tRNA, assumingly stabilizing the structure of the

tRNA. A loss of m<sup>1</sup>A in human mitochondrial tRNA<sup>Lys</sup> was discovered to lead to misfolding [54].

The third group of essential tRNA modifications defines the amino acylation identity of the tRNA. For example the tRNA with the anticodon CAU encodes with an unmodified cytidine at position 34 for methionin, while exchange of position 34 with lysidine (k<sup>2</sup>C) leads to loading of the tRNA with isoleucine. Apparently the unmodified tRNA<sub>CAU</sub> is not (or more slowly) recognized by the Ile-AaRS and isoleucine is only loaded on the lysidine modified tRNA<sup>CAU</sup> [55, 56].

Besides the versatile functions of modifications in tRNA, modified nucleosides play an important role in mRNA and other non-coding RNAs (ncRNAs). In eucaryotic mRNA for example m<sup>7</sup>G plays an important role in the 5' cap structure by defining the mRNA reading frame, influencing the nuclear mRNA export and protecting the mRNA from splicing [57]. In ncRNA 2-O'-methylation alters the RNA structure and therefore protein binding, which leads to a direct regulation of mRNA translation depending on the methylation status [58].

In summary modifications in RNA enable this biomacromolecule to execute all its versatile and important cellular functions by increasing the chemical information content. These functions comprise: Direct involvement in the translational process, stabilization and correct folding of RNA, identification of tRNAs, gene regulation by translational and transcriptional regulation like gene silencing. However, the concrete function of most discovered modifications stays concealed and only some modified nucleosides have elucidated functions and mostly only at certain position of a given RNA. Research of RNA modifications is an ongoing field with many open questions on the positions and functions of modified ribonucleosides.

### 1.2.3 5-methylcytidine (m<sup>5</sup>C)

5-methylcytidine is one of the most prominent and best studied modified nucleosides both in RNA and DNA. In DNA it is known to regulate the heritable gene expression, thereby changing the phenotype of a cell without changing the DNA sequence, which is one important part of epigenetics. In RNA m<sup>5</sup>C is involved in structural folding of RNA, Mg<sup>2+</sup> binding and it was found to affect ribosome precision during translation [59]. A comprehensive review on m<sup>5</sup>C in RNA was published in 2010 and is recommended for a more detailed insight into the topic [60].

### Bisulfite-sequencing

Methylation of cytosine C5, both in DNA and RNA, increases the electron density in the pyrimidine ring, which makes it less prone to reaction with bisulfite. Without the C5 methylation, cytosine readily reacts with bisulfite and is thus deaminated to uracil [61] (Figure 4A). This conversion of all cytosines in a given RNA to uracils results in the introduction of an adenosine in cDNA from the bisulfite treated RNA while the untreated RNA sample has a guanosine incorporated (Figure 4B). By sequence alignment of data derived from treated and untreated RNA, statements about occurrence and positions of  $m^5C$  in the RNA are possible. The technique of bisulfite sequencing was first developed for DNA and adjusted for RNA just a few years ago [62]. The adjustment to RNA facilitated studies to find more RNAs with this modification and locate the position of  $m^5C$  in RNA, which was done transcriptome wide for human RNAs [63]. Additionally straight-forward studies on the enzymes responsible for  $m^5C$  methylation, the RNA:5-methylcytosine-methyltransferases, are now less elaborate.

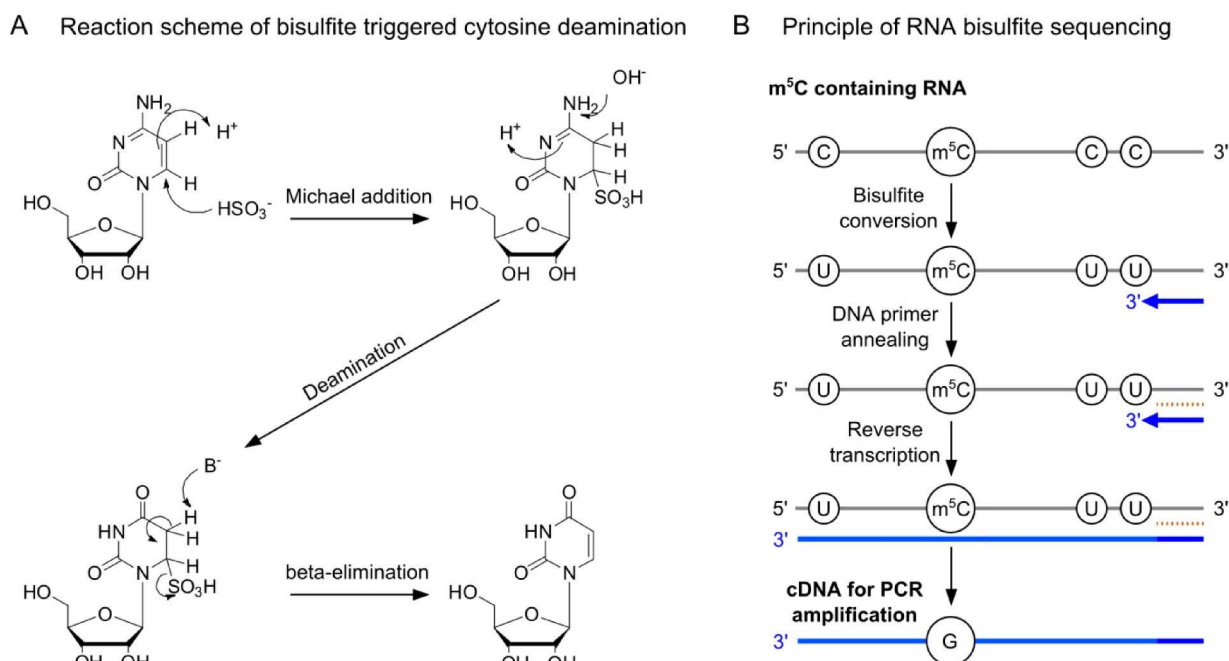


Figure 4 Basic principles of bisulfite sequencing.

A Chemical reactions following addition of bisulfite to the C6 position of cytosine leading to a conversion to uracil.  
 B Only  $m^5C$  is transcribed to G in cDNA generated from bisulfite treated RNA and can be thus detected.

### **RNA:5-methylcytosine-methyltransferases (RNA:m<sup>5</sup>C-MTases)**

m<sup>5</sup>C is a modification that can be found in bacteria, archaea and eucaryotes. All m<sup>5</sup>C-MTases transfer methyl groups from S-adenosyl-methionin (SAM or Adomet) to the C5 of cytosine incorporated in a substrate RNA. Based on sequence similarity RNA:m<sup>5</sup>C-MTases can be divided in four distinct families with examples seen in Table 2 adjusted from a review published in 2010 [60]. Research is now focusing on filling the gaps by assigning substrate RNAs and finding m<sup>5</sup>C positions of putative enzymes by sensitive and robust techniques like bisulfite sequencing and LC-MS/MS analysis.

Table 2 Overview of RNA:m<sup>5</sup>C-MTase families and some of their known and putative substrates (adapted from Motorin *et al.* [60])

Enzyme family	Enzyme name	Other names	Organism	RNA substrate	Modified positions
<b><i>RsmB</i> family</b>					
RsmB/Nol1	Nop2		<i>Saccharomyces cerevisiae</i>	unknown	unknown
YebU	hTrm4	NSun2/Misu	<i>Homo sapiens</i>	Pre-tRNA <sup>Leu</sup>	34
<b><i>Dnmt2</i> family</b>					
Dnmt2	Dnmt2		<i>Homo sapiens</i>	tRNA	38
<b><i>RlmI</i> family</b>					
<b><i>Ynl022</i> family</b>					
Ynl022c	Ynl022c		<i>Saccharomyces cerevisiae</i>	unknown	unknown

Methylation of the non-nucleophilic carbon atom 5 in cytosine residues is a chemical challenge for methyltransferases and requires activation before the methyl group can be transferred. In both DNA:- and RNA:m<sup>5</sup>C-MTases the cytosine is activated by Michael addition of a catalytic cysteine to C6 of the target cytosine in the functional core of the MTase. This leads to an enolate formation which makes the C5 a nucleophile, which is easily methylated by the electrophilic SAM. For enzyme regeneration RNA:m<sup>5</sup>C-MTases use a second cysteine that facilitates the beta-elimination of the C6 bound cysteine and the remaining C5-hydrogen. DNA:m<sup>5</sup>C-MTases on the other hand perform enzyme regeneration without a second cystein [64].

Interestingly Dnmt2 (DNA methyltransferase 2), a highly conserved MTase, has high sequence similarities to DNA:m<sup>5</sup>C-MTases with only one cysteine residue in the catalytic motif but shows nearly no dC methylation but high ribo-C methylation in tRNAs at position 38 [65]. This interesting observation gives room for speculations whether Dnmt2 was a DNA:MTase that switched its substrate specificity to RNA or *vice versa*.

### 1.3 DNA and its modifications

Compared to the amount of nucleoside modifications in RNA, only a small diversity of DNA modifications can be found. This difference of the two nucleic acid species is most likely caused by two reasons: (I) DNA nearly always forms a duplex in B-form, where the bases point into the helix core to form the known base pairs. In this arrangement very little space is left for sophisticated and bulky modifications. (II) In contrast to RNA, DNA has only one known major function – carrier of the genetic code. This important function does not require a high variety of nucleoside modifications.

DNA modifications comprise on the one hand the enzymatically introduced modifications like methylations of cytosine or adenine and on the other hand DNA lesions, which are modifications of the nucleobases due to damaging agents. The most prominent damaging agents are UV irradiation, reactive oxygen species and various molecules from the environment. These modifications in DNA are detrimental by the cells and need to be repaired to keep the genetic information intact. Inefficient repair of these lesions leads to mutagenic and cytotoxic effects [49]. However interesting the formation and effect of DNA lesions and how important a detailed knowledge of the related diseases is, it is not part of this work and only the enzymatically introduced DNA modifications are discussed from here on.

Summarizing DNA modifications found in all trees of life, there are currently 9 known and well described DNA modifications. The nomenclature is similar to the nomenclature of RNA modifications with the difference of a “d” for “deoxy” introduced right before calling the base. In addition, there are several other abbreviations used for DNA modifications like mC, m<sup>5</sup>C and dm<sup>5</sup>C, which normally leave out the “d” or the modification position. In this work the established nomenclature is used. The found modifications can be classified in protecting and epigenetic modifications as presented in a comprehensive review on nucleic acid modifications from 2012 [49].

#### 1.3.1 DNA modifications for protection from restriction enzymes

Bacteria and other non-eucaryotic organisms like trypanosomes mostly use DNA modifications (m<sup>5</sup>dC, m<sup>4</sup>dC, m<sup>6</sup>dA, hm<sup>5</sup>dC and hm<sup>5</sup>dU) for protection of their genome from their own restriction enzymes. These restriction enzymes are used by the organisms to degrade foreign DNA *e.g.* viral DNA that normally does not contain DNA modifications and can be thus distinguished from the hosts DNA. Additionally m<sup>6</sup>dA is involved in regulation of virulence, mismatch repair, timing of DNA replication and control of gene expression [66].

### 1.3.2 DNA modifications with epigenetic functions

Plants, fungi and animals have a broad genome carrying the information for each of the specialized cell types within the organism. Yet, not all genes need to be active in all cell types and regulation of transcriptional activity must be influenced by a second information layer, since the DNA sequence is the same in all cells. This heritable regulation of gene activity can result from histone methylation and methylation of cytosines at crucial positions in the DNA, a part of the research field of epigenetics. The first epigenetic nucleoside found was 5'-methyldeoxycytidine ( $m^5dC$ ), which occurs with a frequency of 4.5 % and is a highly abundant nucleoside in genomic DNA. In general high contents of  $m^5dC$  can be found in promoter regions, where it leads to gene silencing of the following gene. A re-activation of a thus silenced gene, *e.g.* during embryogenesis, would call for a removal of the cytosine methylation. In the past few years, a removal of the methyl group by oxidation was found to be rendered possible by finding three oxidized intermediates of  $m^5dC$ . A first oxidation step of  $m^5dC$  to 5-hydroxymethyldeoxycytidine ( $hm^5dC$ ) was discovered to lead to a reversion of the nucleoside function and a previously silenced gene was shown to become active [67, 68]. In 2011, two more epigenetic nucleosides 5-formyl [69] and 5-carboxydeoxycytidine [70] ( $f^5dC$  and  $ca^5dC$ ) have been discovered. The most reasonable mechanism for replacement of  $m^5dC$  was only recently presented by the Carell group and is shown in Figure 5. After oxidation of  $m^5dC$  to  $ca^5dC$  a decarboxylation reaction leads to an unmodified cytosine residue. This was shown to be chemically possible after saturation of the C5-6 double bond [71]. A second discussed mechanism focuses on the oxidation products as activators of the base-excision repair (BER) system. After excision of the oxidized modified cytosine, a normal, non-methylated, cytosine is introduced in the DNA sequence.

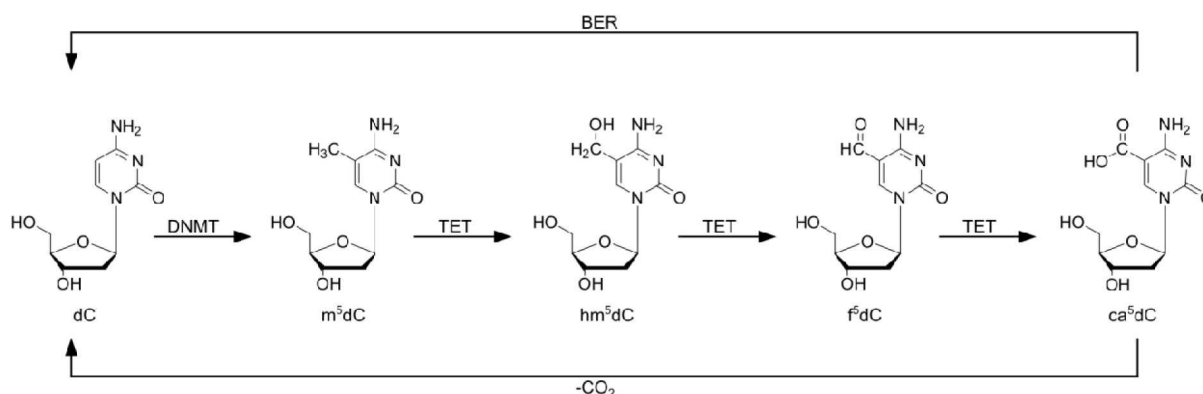


Figure 5 Overview and biosynthetic relation of the four epigenetic DNA modifications. (abbreviations: DNMT – DNA methyl transferase; TET – ten-eleven-translocation enzymes; BER – base excision repair) adjusted from Carell *et al.* [49].

## 1.4 Detection of RNA modifications

As previously described, modifications in RNA play an important role for RNA's versatile functions. To get a better understanding how modifications are synthesized and distributed in the cell, it is necessary to have reliable means of detection. The detection of modified residues in RNA can be divided in three basic principles, which are sometimes combined in several studies to increase the sensitivity of detection: (i) identification according to physicochemical properties, (ii) differential enzymatic turnover, or (iii) differential chemical reactivity using the basics of chemical functionalization (see chapter 1.1.3). We have published a review covering all 3 aspects of RNA modification detection in 2010 for a detailed introduction into the theme [72].

Further Reading:

**Detection of RNA modifications.**

S. Kellner, J. Burhenne & M. Helm

RNA Biol. 2010 Mar-Apr;7(2):237-47.

Epub 2010 Mar 8.

However two principle detection methods are described in more detail in the following chapters, since the differential chemical reactivity and physicochemical properties of RNA modifications were the main focus of this work.

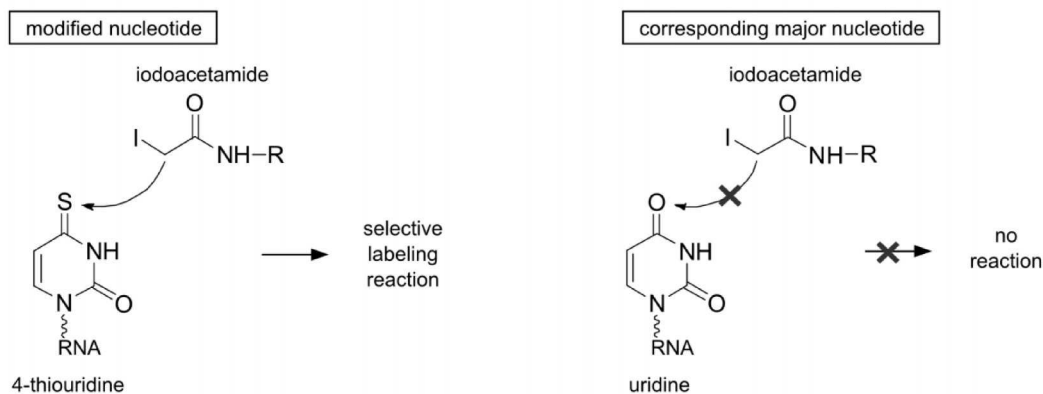
### 1.4.1 Differential chemical reactivity

Even though modified RNA nucleotides are commonly found in nearly all species and RNA types, their detection is cumbersome. One of the main reasons is the low abundance of modified nucleotides compared to their unmodified counterpart, which interfere with the detection of the wanted modified residue.

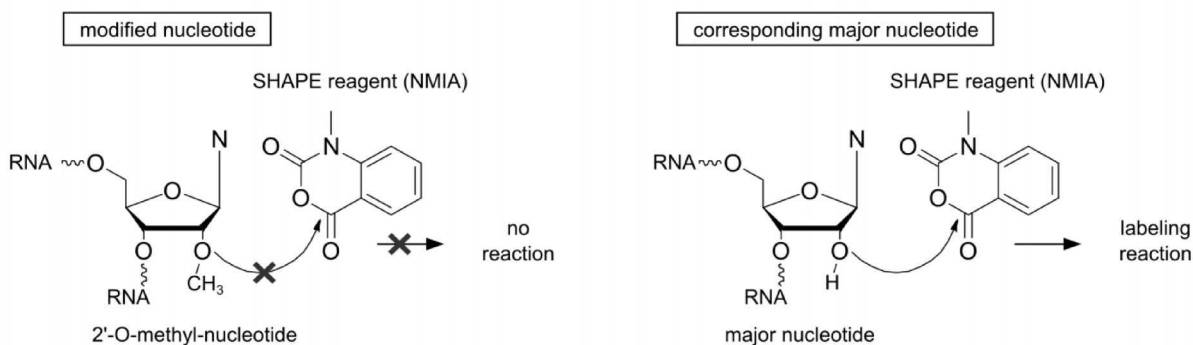
Chemical reagents, used for detection of modified nucleotides, target certain functionalities in the modified residue. This target is in the best case only present in the modified nucleotide but not in the major base, like the sulfur of 4-thiouridine (see Figure 6A), which allows a high selectivity of the reagent. Other reagents target a common functional group of major nucleotides, but their reaction is blocked by a modification (*e.g.* 2'O methylation in Figure 6B), which thus allows the detection of the modification. However, some reagents react with a target that is present in both the modified and the major nucleotide, like the *N3* of pseudouridine and



## A Modifications with targets, not present in major nucleotides



## B Modifications, blocking target sites



## C Modifications detected by differential chemical reactivity of the same target present in the corresponding major nucleotides

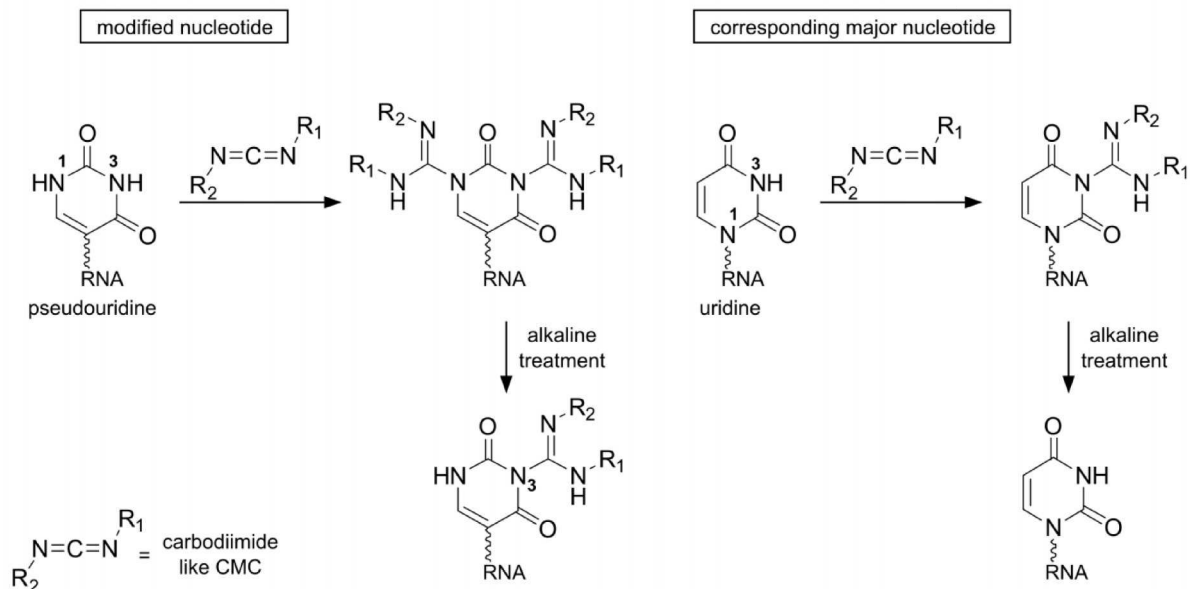


Figure 6 Overview of reagents for detection of modified nucleotides.

A As an example for reagents that target functional groups only present in modified nucleosides, iodoacetamid is presented, a reagent targeting only the sulfur atoms in thiolated nucleotides. B NMIA selectively reacts with non-constrained 2'OH of the ribose. The reaction would most likely not occur in case of 2'OH methylation and can be therefore used for detection of such modified nucleotides. C Detection of pseudouridine with carbodiimide and subsequent alkaline treatment.

uridine (Figure 6C). For acceptable selectivity, the reactivity of the target must be sufficiently altered by the modification to distinguish the modified and major nucleotide. Either way, to achieve highest selectivity for the modified nucleotide, the reaction conditions must be narrowly controlled. Otherwise a lowered selectivity would lead to a side reaction with the more abundant major nucleotide and the modified nucleotide of interest would be masked and would stay undetected.

The second challenge, besides the selective reaction with the modification of interest, is the subsequent detection of the labeled nucleotide. Two different approaches are commonly established, namely detection by reverse transcription (RT) and mass spectrometry (MS). In general, RT profiles are used for bulky labels that stop the reverse transcriptase and lead to significant changes in the RT profile in the presence of a successfully labeled modification [73]. Modifications carrying smaller labels go mostly undetected by the reverse transcriptase but can be detected by MS based techniques, since they increase the mass of the modified nucleotide. A successfully labeled modification, like pseudouridine, can be therefore distinguished from its major nucleoside by its increased mass and a different fragmentation pattern [74]. An advancement of RT for detection of modified nucleotides is the application of new generation sequencing, which can be used for a transcriptome wide search for certain, modified nucleotides. This concept was proven possible by a transcriptome wide mapping of 6-methyladenosine by antibody capturing and subsequent sequencing [75].

As there is a large repertoire of reagents used for detection of modified nucleotides, only some reagents are presented here. For a more detailed overview, the review mentioned before and a more recent review, focusing on exactly this chapters topic Behm-Ansmant *et al.* [76], are recommended.

**Thiols:** Thiols are strong nucleophiles that easily attack all sorts of electrophiles, which is used for selective labeling of tRNA<sup>Phe</sup> s<sup>4</sup>U at position 8 with a fluorescent iodoacetamide dye (Figure 6A) to study the aminoacyl-tRNA.elongation factor Tu.GTP ternary complex [77]. Additionally, s<sup>4</sup>U was reported to react with 7-methoxy-4-bromomethylcoumarin (BMB) in 1973 [78].

**Primary amines:** acp<sup>3</sup>U and queuosine, for example, have a primary amine functionality that selectively reacts with NHS [23] or isothiocyanate reagents [79] and can be therefore selectively labeled.

**2'ribose methylation:** In 2004, a ribose 2'OH specific reagent (N-methylisatoicanhydride, NMIA) was published, which only targets nucleotides not engaged in base-pairing. This feature was exploited for RNA secondary and tertiary structure analysis [41, 80]. As suggested in Figure 6B the 2'OH ribose methylation will most likely not allow a reaction with the used chemicals, which could be applied for detection of ribose methylations by sequencing.

**Pseudouridine ( $\Psi$ ):** The most abundant modified nucleoside, pseudouridine, is the isomerization product of uridine, which results in both having the same Watson-Crick base pairing behavior and the same mass. Detection by enzymatic and mass spectrometric means is therefore problematic, but numerous reagents could be developed that react more or less selective with pseudouridine.

The most selective and therefore most commonly used reagent is CMCT (N-cyclohexyl-N'- $\beta$ -(4-methylmorpholinium)ethylcarbodiimide *p*-tosylate) which acylates G, U, I and  $\Psi$ , but the acyl-moieties can be removed from G, U and I residues by alkaline treatment, leaving only *N*3-acylated pseudouridines. The acylated pseudouridine is relatively bulky, which allows detection not only by mass spectrometry but additionally by RT [73, 81].

Acrylonitrile and methylvinylsulfone are reported to specifically react with pseudouridine, as well, and are mostly applied in mass-spectrometric studies [74, 82] since the reverse transcriptase is not stopped by these labels.

All pseudouridine selective reagents mentioned above are relatively small and do not allow a direct detection of pseudouridine by *e.g.* fluorescence. In 1974, the fluorescent reagent 7-methoxy-4-bromomethylcoumarin (BMB) was reported to react selectively with pseudouridine under specific reaction conditions [83]. Since its characterization it was used in one study on 5S rRNA structure in 1983 [84].

## 1.5 LC-MS/MS of modified nucleosides

LC-MS analysis is the successful combination of two powerful analytical methods, which allows identification and sensitive quantification of compounds. Liquid chromatography separates the analytes according to their physicochemical properties like hydrophilicity, basicity, or size, which can be used for compound identification. Mass spectrometry is a detection system that measures the mass to charge ratio of ionized compounds, which allows identification and sensitive detection of the analyte. The high sensitivity is of great advantage since modified nucleosides are rare compared to their unmodified canonical nucleosides. In this chapter the most important chromatographic methods and the mass spectrometric principles for detection of modified ribonucleosides are presented, followed by examples of successful combination of both methods.

### 1.5.1 Basic principles of mass spectrometry of nucleosides

Mass spectrometry is an analytical method for separation and detection of charged analytes in an electromagnetic field. The basic principle of a 5 module MS/MS device and the typical behavior of a nucleoside analyte are shown in Figure 7.

#### *Ion source:*

The ion source loads the analyte with a charge depending on the ionization mode. In this example, electrospray ionization (ESI) is shown, which adds or removes protons from the analyte in an electric field. Simultaneously, it vaporizes the analyte and transfers it from the normal pressure atmosphere into the high vacuum system by guiding them through an oppositely charged ion capillary. ESI is a relatively gentle ionization method and is frequently used in RNA and DNA analytics. Another commonly used ion source is matrix-assisted-laser-desorption-ionization (MALDI) which is used for gentle ionization of large macromolecules. However, combination of MALDI and liquid chromatography is cumbersome and not commonly done.

#### *Mass analyzer:*

The vaporized and ionized analytes fly through an electromagnetic field, where they are separated by their mass-to-charge ( $m/z$ ) ratio. In Figure 7, a quadrupole MS is shown with 4 parallel rods that produce an oscillating electric field, which allows selective filtering of the desired  $m/z$  charged analyte.

#### *Collision cell:*

The analyte has to pass a highly energetic electrical field, which is filled with an inert gas *e.g.* nitrogen or argon. If the analyte collides with a gas molecule, energy is transferred to the analyte, which causes fragmentation at the chemically weakest bond. Thus, this method is called

collision-induced-dissociation (CID). Since the energetic field is historically based on a non mass-resolving quadrupole, the module is referred to as quadrupole 2. However, modern devices are commonly equipped with a hexapole or an octapole for generation of the energetic field.

### 2<sup>nd</sup> mass analyzer:

In this MS/MS setup a quadrupole (Q3) for mass spectrometric separation of the incoming analyte fragments is used. This allows a highly sensitive detection of the analyte and its fragments. Instead of a quadrupole, a time-of-flight (TOF) analyzer, which allows higher mass precision and resolution, can be used for exact analyte identification.

### Detector:

The detector collects the incoming analyte and multiplies the signal in the shown electron-multiplier horn to produce a signal.

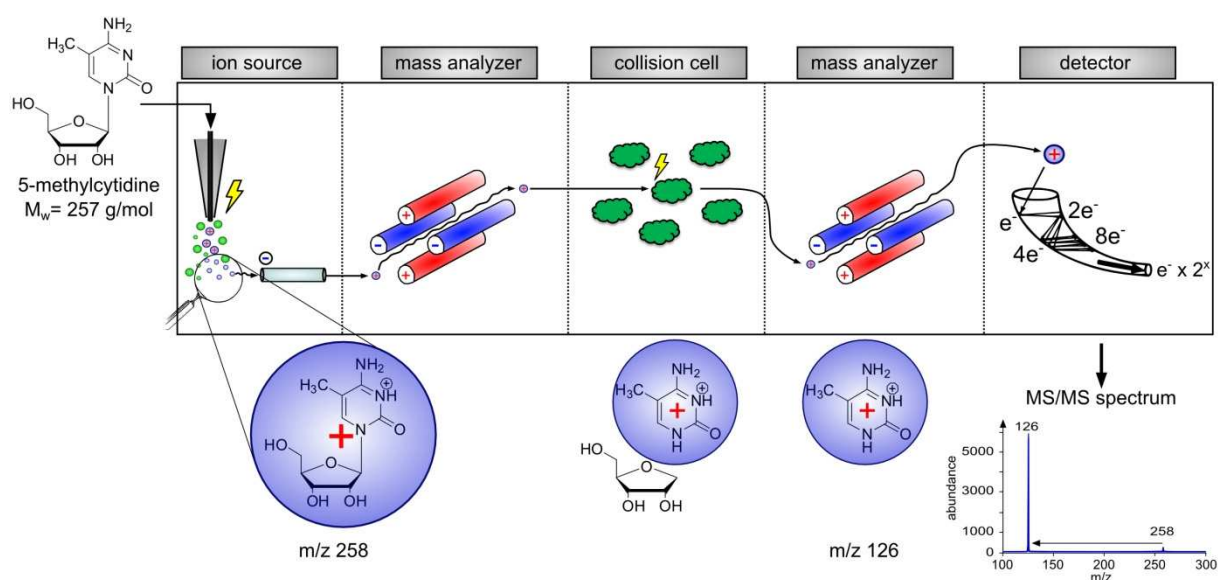


Figure 7 Basic design of MS/MS detection systems demonstrated on an ESI-triple quadrupole MS.

5-methylcytidine enters the mass spectrometer through the electrospray ionization source and is nebulized, evaporated and charged with a proton. The analyte is transferred to the mass analyzer by a dielectric capillary and enters the electromagnetic field of the quadrupole rods. In the collision cell the nucleoside collides with nitrogen atoms, which leads to a cleavage of the glycosidic bond. The charged nucleobase enters the second quadrupole before it is detected in the electronmultiplier, which finally gives a MS/MS signal.

In the second row of Figure 7, the behavior of  $m^5C$  in a triple quadrupole MS is shown. Like  $m^5C$  most nucleosides have a pronounced basic character, which is exploited by ionization with protons in the ESI positive ion mode producing a precursor ion, here  $[M+H]^+$  258. For detection methods with high sensitivity, only this precursor mass is allowed to pass Q1, while other ions collide with the quadrupole rods and are not further analyzed. In this case, only ions of an  $m/z$  of 258 enter the collision cell and the nucleoside collides with nitrogen, yielding two fragments: the uncharged ribose and the charged nucleobase, which is called the product ion  $[B+2H]^+$  126. The uncharged ribose and other possible uncharged fragments are not further transported and only the charged nucleobase enters Q3. For generation of MS/MS spectra, Q3 is set to scan a user defined

mass range for incoming ions (product ion scan mode). In case quantitative MS/MS analysis is required, Q3 is set to filter only the expected product ion mass, which leads to an increase in sensitivity. In this case, the mass transition for the exemplary nucleoside  $m^5C$  would be written like:  $258 \rightarrow 126$  (precursor ion  $\rightarrow$  product ion) for usage in selective-/ or multiple reaction monitoring mode (SRM/MRM).

The high ionization efficiency together with the simple and predictable fragmentation behavior makes nucleosides an extremely well detectable analyte species.

### 1.5.2 Basic principles of nucleoside chromatography

Chromatography for separation of RNA nucleosides is a challenging task, since up till now over 100 modifications are known. Most of these modifications only change the physicochemical properties of the nucleoside to a relatively small extent and in addition the same type of modification can occur at several positions of the nucleoside. As an example, separation of methylated nucleosides is especially challenging, since usually up to 4 isomers of each nucleoside exist (*e.g.*  $m^3C$ ,  $m^4C$ ,  $m^5C$  and  $Cm$ ). Another challenge is the pronounced hydrophilic and basic character of nucleosides, which is caused by numerous hydroxyl-functions of the ribose, amines and amides of the nucleobase. Early chromatographic methods on thin-layer plates used hydrophilic normal-phases (cellulose) and analysis of RNA was done on the nucleotide level to even increase the hydrophilicity. Liquid chromatography of nucleosides is commonly done with reverse-phase C-18 columns and low ionic-strength buffers to keep the nucleosides uncharged.

#### *High-performance-liquid-chromatography (HPLC)*

First reports on RP-18 chromatography of nucleosides combined with UV detection go back to the late 1970s [85] and allowed separation of 9 modified nucleosides. The method was rapidly improved by a detailed analysis of the influence of several chromatographic parameters *e.g.* flow-rate, isocratic vs. gradient elution, buffer pH and column temperature [86]. In 1989, the most sophisticated RP-HPLC method allowed separation of 65 and quantification of up to 31 nucleosides in a single run with an LOD in the single-digit pmol range due to the limited sensitivity of the photodiode array detector (DAD). Variations of this method are still used for simple nucleoside analyses [87, 88]. New column materials that improve retention of the hydrophilic nucleosides by introduction of hydrophilic functional groups to the RP-18 particle surface in the column allow a better separation of similarly modified nucleosides. Even normal-phase HILIC (hydrophilic interaction chromatography) columns are in use today. Additionally,

modern chromatographs allow system pressures of up to 1200 bar, which led in combination with advances in column preparation to the development of ultra-performance-liquid-chromatography (UPLC). UPLC columns are characterized by a combination of smaller length, inner diameter and particle size, which leads to a high back pressure at even low flow-rates below 0.5 mL/min. Their main advantages are:

- Same resolution as the corresponding conventional column,
- massively shortened analysis time and
- lower solvent consumption.

This makes UPLC the most promising recent development for future high-throughput analysis of RNA modifications and is already used in combination with MS based detection methods as will be presented in the third part of this chapter.

### 1.5.3 LC-MS(/MS) combining methods for RNA and DNA analytics

The combination of liquid chromatography and mass spectrometry is troublesome due to the high amount of solvent coming from the LC and the need for vaporized and ionized analyte in the high vacuum MS system. The combination of LC and MS only succeeded with the introduction of atomic pressure ionization methods, like ESI, which allows a high transmission of analyte contained in the LC solvent into the MS [89].

#### *Analysis of nucleosides*

In 1990 a comprehensive LC-MS/MS method for detection of 55 modified nucleosides in RNA digests was presented, which formed the basis for all modern methods [90]. Quantification of modified nucleosides is normally done in percent to the unmodified canonical nucleoside and can be easily done by MS/MS detection methods. The limit of detection and limit of quantification calculated at a signal-to-noise (S/N) of 3, and 10 respectively, are common parameters for the methods quality. Modern methods using tandem MS detection of RNA nucleosides are rare and achieve LODs < 1 pmol [91]. More sophisticated methods with better LODs have been achieved for DNA nucleosides, especially m<sup>5</sup>dC, which is still detectable in single fmol amounts [92-94]. Recently, UPLC-MS/MS detection methods with analysis times below 12 minutes have been reported [95, 96]. New methods for normal-phase nucleoside chromatography with HILIC columns in combination with MS/MS detection are developed for RNA [97] and DNA [98] nucleosides, which further expand the scope of LC-MS based analytics. Reliable quantification is additionally possible with LC-single-MS methods, which require isotope labeled standards of the analyzed modified nucleoside [99, 100]. A drawback of

RNA nucleoside analysis is the complete loss of sequence information by digestion, which can be overcome by LC-MS of short oligonucleotide fragments.

### ***Analysis of short oligonucleotides***

Oligonucleotides are, like other big macromolecules, analyzed by MALDI-MS based methods. In 2009 the group of P. Limbach presented a new approach using specific RNases like RNase T1 to produce RNA species specific signature digestion products (SDPs). These were identified by subsequent MALDI-TOF analysis, which allows mapping of modified nucleosides in the fragment [101]. This approach was further improved by transferring the method to HPLC and UPLC based MS methods to allow identification of organisms by their SDPs [102]. Since recently, this method can be used for quantification of the modified nucleoside pseudouridine with an LOD of  $\sim 5$  pmol [103] requiring only 50 pmol starting material at a 10 % modification rate.

### ***Discovery of novel RNA modifications***

Thanks to more and more sensitive MS based detection methods and RNA isolation techniques it is possible to isolate and identify novel modified nucleosides. The most sophisticated approach was reported 2012, which made use of the high accuracy of QTOF mass spectrometry to identify the novel nucleoside 3-(3-amino-3-carboxypropyl)-5,6-dihydrouridine (acp<sup>3</sup>D) at position 47 of tRNA<sup>Lys</sup><sub>UUU</sub> from *Trypanosoma brucei* [104]. This was possible by combination of the aforementioned nucleoside and oligonucleotide analytics.

### ***Summary of LC-MS(/MS) based methods***

LC-MS/MS analysis of nucleosides is a highly sensitive detection method with LODs in the single digit fmol range, but it does not give any sequence information. MS of oligonucleotides on the other hand is suitable to map modified nucleosides but the sensitivity is 1000 fold worse than nucleoside analytics. In combination both methods can be used for detection of novel modified nucleosides and its exact position in a given sequence.



## 2 Goal of the work

Alkylating agents for nucleic acids have been in the focus of research since the discovery of the structure of DNA and its role in cell heredity in the 1950s. In the course of these examinations several of these agents were found to be specific for certain nitrogens of RNA. Especially the agents selective for particular modified nucleosides contributed to a high extent to studies on RNA structure and function. Up till now 108 modified nucleosides have been identified but only a handful of them can be selectively labeled and much less can be reliably detected due to their very low abundance. This work addresses two possible solutions to overcome this discrepancy.

The first approach is the search for new RNA labeling agents based on a coumarin scaffold. The goal is to examine the potential of differentially substituted coumarins in terms of selectivity for certain functionalities of RNA and its modified nucleosides and thereby to establish a systematic structure-function relationship. This includes examination of different reaction conditions and their influence on the selectivity. In addition LC-MS/MS can be applied for analysis of successfully labeled RNA to identify the nucleoside target of the agent. Using this technique the analytical scope is extended notably which enables a more detailed analysis of the structure-function relationship.

The second approach of this work is to establish a highly accurate and sensitive detection method for naturally modified nucleosides using LC-MS/MS analysis. The developed methods must be applicable to nucleic acid digests and (a) identify as many of the 108 modified ribonucleosides as possible in a single run, or as many of the modified deoxynucleoside modifications, respectively. (b) Allow highly sensitive detection and quantification of certain modified nucleosides, like 5-methylcytidine ( $m^5C$ ), as a particularly timely example. The highly sensitive methods should then be applied to examine RNA or DNA from different origins in order to obtain more information about the nature, function and capabilities of the modifying enzymes with a special focus on methyltransferases.

## 3 Materials & Methods

### 3.1 Chemicals and Enzymes

GelRed	Biotium, Hayward, USA
Acetic acid, glacial for LC-MS	Sigma-Aldrich, Steinheim
Acetone, puris	Sigma-Aldrich, Steinheim
Acetonitril p.a. for HPLC	Sigma-Aldrich, Steinheim
AlexaFluor 594-azide	Invitrogen, Darmstadt
Ammoniaperoxydisulfate (APS)	Roth, Karlsruhe
Ammonium acetate puris, for LC-MS	Sigma-Aldrich, Steinheim
atto488-azide	Atto-Tec, Siegen
atto647 N-azide	Atto-Tec, Siegen
Bromphenolblue	Merck, Darmstadt
Dimethylsulfoxid, pure	Sigma-Aldrich, Steinheim
Ethanol > 99.5 % Ph.Eur.	Roth, Karlsruhe
FastAP Thermosensitive alkaline phosphatase	Fermentas, St.Leon-Roth
Formamide > 99.5 %	Roth, Karlsruhe
Isobutyric acid > 99%	Acros, Geel, Belgium
Lithium perchlorate, battery grade, dry	Sigma-Aldrich, Steinheim
N,N,N',N'-Tetramethylethyldiamin (TEMED)	Roth, Karlsruhe
Nuclease P1	Sigma-Aldrich, Steinheim
Rotiphorese 10 x TBE Buffer	Roth, Karlsruhe
Rotiphorese sequencing gel buffer concentrate	Roth, Karlsruhe
Rotiphorese sequencing gel concentrate	Roth, Karlsruhe
Rotiphorese sequencing gel diluter	Roth, Karlsruhe
Snake venom phosphodiesterase	Worthington, Lakewood, USA
Total tRNA <i>E.coli</i>	Roche, Basel, Swiss
Xylenxyanol	Merck, Darmstadt

All other chemicals not mentioned in the list were provided by the institute's chemical store (1.st floor, Staudingerweg 5, 55128 Mainz).

## 3.2 Buffers

### 5 mM Ammoniumacetate buffer for LC-MS nucleoside analytics

0.385 g ammoniumacetate, LC-MS grade

70  $\mu$ L glacial acetic acid, LC-MS grade

dissolved in 1000 mL MilliQ water in a white-glass Schott bottle

The pH was adjusted to pH 5.3 using glacial acetic acids and 1 mL acetonitril, LC-MS grade, was added for microbial stability .

Storage: Under light protection at 4 °C until use.

On HPLC maximum consuming time: 1 week.

### Buffer Nuclease P1 10 x

90  $\mu$ L Ammoniumacetate 0.25 mM pH 5.0

10  $\mu$ L Zinc chloride 2 mM

### Lithiumperchlorate solution for nucleic acid precipitation

10 g Lithiumperchlorate, battery grade

500 g acetone puris

Storage: in 50 mL Falcon tubes at -20 °C until usage at room temperature

### Loading dye for denaturing PAGE

90 mL formamide

10 mL 10 x TBE buffer

(Blue loading dye: addition of one spatula tip bromphenolblue and xylenxanol)

### TLC running buffer NI

330 mL isobutyric acid

5 mL ammonia 25 %

ad 500 mL with MilliQ water

### TLC running buffer RII

1000 mL phosphate buffer pH 6.8 (0.1 M)

600 g ammonium sulfate

15 ml 1-propanol

### 3.3 Oligonucleotides

Table 3 Commercial 20 mer oligonucleotides of bipartite composition

Database name	Composition	Sequence (5' → 3')	Supplier
MH 572	DNA; A/C	ACCAACAAACCACCACAACC	Biomers, Ulm
MH 573	DNA; C/G	GCGGCCCGCCGGCCGGCGCG	Biomers, Ulm
MH 574	DNA; G/A	GAAGGAAGGGAAAGAGAGGA	Biomers, Ulm
MH 575	DNA; G/T	GGTTGTGGTGGTTTGTGGT	Biomers, Ulm
MH 576	DNA; T/A	TTATATTTATAAATATTTAA	Biomers, Ulm
MH 577	DNA; T/C	TCCTCTCTTCCCTTCCTTC	Biomers, Ulm
MH 527	RNA; A/C	ACCAACAAACCACCACAACC	Sigma-Aldrich, Steinheim
MH 528	RNA; C/G	CCGGCCCGGCGGCCGGCGCG	Sigma-Aldrich, Steinheim
MH 529	RNA; G/A	AAAGGAAGGGAAAGAAAGAA	Sigma-Aldrich, Steinheim
MH 552	RNA; G/U	GGUUGUGGUGGUUUGUUGGU	Biomers, Ulm
MH 553	RNA; U/C	UUCUUUCUUUCCCUUCCUUU	Biomers, Ulm
MH 554	RNA; U/A	UUAUAUUUAUAAAUAUAAA	Biomers, Ulm

Table 4 Other oligonucleotides used for coumarin derivatization and LC-MS/MS analysis

Database name	species	Sequence (5' → 3')	Supplier
MH 449	RNA	GCGGAUUUAGCUCAGUUGGAGAGCGCC	Thermo- Scientific, Dreieich
MH 618	DNA	Alkynyl-CGCGCGAAGCTTAATACGACTCACTATA	IBA, Goettingen
MH 569	RNA	GUCAm <sup>5</sup> dCGCGGGAGACCGGGUUCGAUUCCCCGACG GGGAGCCA	IBA, Goettingen
MH 533	antisense siRNA	GAACUUCAGGGUCAGCUUGCCG	Biomers, Ulm
MH 534	sense siRNA	GCAAGCUGACCCUGAAGUUCAU	Biomers, Ulm
MH 543	antisense siRNA	Atto647N-GAACUUCAGGGUCAGCUUGCCG	IBA, Goettingen
MH 545	sense siRNA	GCAAGCUGACCCUGAAGUUCAU-alexa555	IBA, Goettingen

### 3.4 Modified nucleosides

#### Ribonucleosides

1-methyladenosine	(m <sup>1</sup> A)	Sigma-Aldrich, Steinheim
2'O-methyladenosine	(Am)	Sigma-Aldrich, Steinheim
2'O-methylcytidine	(Cm)	Sigma-Aldrich, Steinheim
2'O-methylguanosine	(Gm)	Berry&Associates, Dexter, USA
2'O-methyluridine	(Um)	Berry&Associates, Dexter, USA
2-methylguanosine	(m <sup>2</sup> G)	Sigma-Aldrich, Steinheim
2-thiocytidine	(s <sup>2</sup> C)	Testmix, Sigma-Aldrich, Steinheim
2-thiouridine	(s <sup>2</sup> U)	Berry&Associates, Dexter, USA
3-methylcytidine	(m <sup>3</sup> C)	Testmix, Sigma-Aldrich, Steinheim
3-methyluridine	(m <sup>3</sup> U)	Sigma-Aldrich, Steinheim
4-thiouridine	(s <sup>4</sup> U)	Sigma-Aldrich, Steinheim
5-methyl-2-thiouridine	(m <sup>5</sup> s <sup>2</sup> U)	Sigma-Aldrich, Steinheim
5-methylcytidine	(m <sup>5</sup> C)	Sigma-Aldrich, Steinheim
5-methyluridine	(m <sup>5</sup> U)	Berry&Associates, Dexter, USA
6-methyladenosine	(m <sup>6</sup> A)	
7-methylguanosine	(m <sup>7</sup> G)	Sigma-Aldrich, Steinheim
Dihydrouridine	(D)	
Inosine	(I)	Sigma-Aldrich, Steinheim
N6,6-dimethyladenosine	(m <sub>2</sub> <sup>6</sup> A)	
Pseudouridine	(Ψ)	Berry&Associates, Dexter, USA
Taurino-5-methyl-2-thiouridine	(τm <sup>5</sup> s <sup>2</sup> U)	

#### Deoxyribonucleosides

N6-methyldeoxyadenosine	(m <sup>6</sup> dA)	
5-carboxydeoxycytidine	(cadC)	Berry&Associates, Dexter, USA
5-formyldeoxycytidine	(fdC)	Berry&Associates, Dexter, USA
5-hydroxymethyldeoxycytidine	(hm <sup>5</sup> dC)	Berry&Associates, Dexter, USA
5-methyldeoxycytidine	(m <sup>5</sup> dC)	Berry&Associates, Dexter, USA

Nucleosides without supplier information were gifts of collaborating laboratories or were of high abundance in tRNA *E.coli* digests.

### 3.5 Methods

#### 3.5.1 Coumarin derivatization

##### Reaction conditions 1 (adjusted from [83])

RNA (0.1-1  $\mu\text{g}/\mu\text{L}$  final concentration) was incubated with 8.8 mM coumarin (20 mM stock solution in pure DMSO), 62.5 mM phosphate buffer pH 6.5 and 75% DMSO at 37 °C for 300 min under light protection. 10 volumes of  $\text{LiClO}_4$ /acetone solution were added, centrifuged at 13 000 rpm for 30 minutes, the pellet washed with pure acetone and redissolved in an appropriate volume of MilliQ water.

Table 5 Exemplary pipetting scheme for coumarin derivatization under reaction conditions 1

	Concentration used	Volume used	final concentration
tRNA <i>E.coli</i>	3 $\mu\text{g}/\mu\text{L}$	1.5 $\mu\text{L}$	0.45 $\mu\text{g}/\mu\text{L}$
coumarin	20 mM	4.4 $\mu\text{L}$	8.8 mM
DMSO		3.1 $\mu\text{L}$	75 %
phosphate buffer pH 6.5	1000 mM	0.625 $\mu\text{L}$	62.5 mM
H <sub>2</sub> O		0.375 $\mu\text{L}$	
total		10 $\mu\text{L}$	

Incubation temperature: 37°C

Incubation time: 300 minutes

Workup: addition of 100  $\mu\text{L}$   $\text{LiClO}_4$ /acetone solution for precipitation

##### Reaction conditions 2 (adjusted from [105])

RNA (0.1-1  $\mu\text{g}/\mu\text{L}$  final concentration) was incubated with 10 mM coumarin (20 mM stock solution in pure DMSO), 100 mM phosphate buffer pH 8.25 and 70% DMSO at 37 °C for 180 min under light protection. 10 volumes of  $\text{LiClO}_4$ /acetone solution were added, centrifuged at 13 000 rpm for 30 minutes, the pellet washed with pure acetone and redissolved in an appropriate volume of MilliQ water.

Table 6 Exemplary pipetting scheme for coumarin derivatization under reaction conditions 2

	Concentration used	Volume used	final concentration
tRNA <i>E.coli</i>	3 $\mu\text{g}/\mu\text{L}$	1.5 $\mu\text{L}$	0.45 $\mu\text{g}/\mu\text{L}$
coumarin	20 mM	5 $\mu\text{L}$	10 mM
DMSO		2.5 $\mu\text{L}$	75 %
phosphate buffer pH 8.25	1000 mM	1 $\mu\text{L}$	100 mM
H <sub>2</sub> O		0 $\mu\text{L}$	
total		10 $\mu\text{L}$	

Incubation temperature: 37°C

Incubation time: 180 minutes

Workup: addition of 100  $\mu\text{L}$   $\text{LiClO}_4$ /acetone solution for precipitation

### Reaction conditions 3 (adjusted from experimental results Katharina Schmid)

RNA (0.1-1  $\mu\text{g}/\mu\text{L}$  final concentration) was incubated with 10 mM coumarin (20 mM stock solution in pure DMSO), 100 mM phosphate buffer pH 8.25 and 50% DMSO at 37 °C for 120 min under light protection. 10 volumes of  $\text{LiClO}_4$ /acetone solution were added, centrifuged at 13000 rpm for 30 minutes, the pellet washed with pure acetone and redissolved in an appropriate volume of MilliQ water.

Table 7 Exemplary pipetting scheme for coumarin derivatization under reaction conditions 3

	Concentration used	Volume used	final concentration
tRNA <i>E.coli</i>	3 $\mu\text{g}/\mu\text{L}$	1.5 $\mu\text{L}$	0.45 $\mu\text{g}/\mu\text{L}$
coumarin	20 mM	5 $\mu\text{L}$	10 mM
DMSO		0 $\mu\text{L}$	50 %
phosphate buffer pH 8.25	1000 mM	1 $\mu\text{L}$	100 mM
H2O		2.5 $\mu\text{L}$	
total		10 $\mu\text{L}$	

Incubation temperature: 37°C

Incubation time: 120 minutes

Workup: short centrifugation (3000 rpm) and removal of supernatant in fresh tube. Addition of 100  $\mu\text{L}$   $\text{LiClO}_4$ /acetone solution for precipitation

### 3.5.2 Molecular biological methods

#### CuAAC (click) reaction

MilliQ water, TPTA, Na-ascorbate and  $\text{CuSO}_4$  are mixed in an eppendorf tube before addition of the fluorescent azide or alkyne dye. Then the RNA is added, mixed and the tube left light-protected for 2 hours at room temperature.

Table 8 Exemplary pipetting scheme for the CuAAC (click) reaction

	Concentration used	Volume used	final concentration
TPTA	50 mM	1 $\mu\text{L}$	2.5 mM
Na-ascorbate	50 mM	2 $\mu\text{L}$	5 mM
$\text{CuSO}_4$	5 mM	2 $\mu\text{L}$	0.5 mM
Fluorescent dye	1 mM	1 $\mu\text{L}$	0.05 mM
H2O		4 $\mu\text{L}$	
RNA	0.1-1 $\mu\text{g}/\mu\text{L}$	10 $\mu\text{L}$	0.01-0.1 $\mu\text{g}/\mu\text{L}$
total		20 $\mu\text{L}$	

Incubation temperature: room temperature

Incubation time: 120 minutes

Workup: addition of 200  $\mu\text{L}$   $\text{LiClO}_4$ /acetone solution for precipitation

### PAGE analysis

50 mL PAGE sufficient for a 20 x 20 cm x 1 mm gel were prepared as follows:

Table 9 Composition of differently concentrated polyacrylamide gels

Concentration	10%	12%	15%	20%	
Gel-concentrate	20	24	30	40	ml
Gel-diluter	25	21	15	5	ml
Gel buffer	5	5	5	5	ml
APS 10 x	200	200	200	200	$\mu$ L
Temed	50	50	50	50	$\mu$ L

### Digestion of nucleic acids for TLC and LC-MS analysis

Standard protocol for 10  $\mu$ g RNA digestion to nucleotides (TLC)

Table 10 Protocol for digestion of RNA to nucleotides

	Concentration used	Volume used	Final
RNA	1 $\mu$ g/ $\mu$ L	10 $\mu$ L	10 $\mu$ g
Nuclease P1	0.3 u/ $\mu$ L	1 $\mu$ L	0.3 units/10 $\mu$ g
Nuclease P1 buffer	10 X	1.1 $\mu$ L	1 x
<i>Incubation:</i>	37 °C	5 hours	
Snake venom phosphodiesterase	0.1 u/ $\mu$ L	1 $\mu$ L	0.1 u/200 $\mu$ g
<i>Incubation:</i>	37 °C	2 hours	
Alkaline Phosphatase	1 u/ $\mu$ L	4 $\mu$ L	4 u/10 $\mu$ g
Buffer alkaline phosphatase	10 x	2 $\mu$ L	1 x
<i>Incubation:</i>	37 °C	1 hour	

For digestion to nucleosides add (LC-MS):

Table 11 Digestion from nucleotides to nucleosides

Alkaline Phosphatase	1 u/ $\mu$ L	4 $\mu$ l	4 u/10 $\mu$ g
Buffer alkaline phosphatase	10 x	2 $\mu$ l	1 x
<i>Incubation:</i>	37 °C	1 hour	

DNA was digested using the same protocol but with 5 minutes denaturing at 100 °C followed by immediate cooling in ice before addition of enzymes.



### 3.5.3 Chromatographic parameters for nucleoside analysis

HPLC: Agilent 1260 equipped with temperature controlled sample tray, autoinjector, column oven, diode array detector and since 07/2012 fluorescence detector

Column: Synergy Fusion RP, 4  $\mu\text{m}$  particle size, 80  $\text{\AA}$  pore size, 250 mm length, 2 mm inner diameter from Phenomenex (Aschaffenburg, Germany)

Buffer: 5 mM Ammoniumacetate pH 5.3

Column Temperature: 35  $^{\circ}\text{C}$

Table 12 Chromatographic parameters used for coumarin and nucleoside analytics

Coumarin analytics		Ribonucleoside analytics „nucs4“		Deoxyribonucleoside analytics „dnucs2“	
Time	% ACN (solvent B)	Time	% ACN (solvent B)	Time	% ACN (solvent B)
0	0	0	0	0	0
2	0	10	8	10	20
10	40	20	40	12	0
12	80	23	0	17	0
15	80	30	0		
18	0				
25	0				
Flow	0.5 mL/min	Flow	0.35 mL/min	Flow	0.5 mL/min

### 3.5.4 MS parameters

Mass spectrometer:

Triple Quadrupol mass spectrometer Agilent 6460 with ESI Jetstream ion source

Delta EMV standard analytics 0 V

Delta EMV high sensitivity +400 V

Fragmentor voltage 80 V

Cell Accelerator Voltage 2 V

Ionisation Electro Spray Ionisation, positive

#### Source parameters

Gas temperature 300  $^{\circ}\text{C}$

Gas flow 5 L/min

Nebulizer 35 psi

Sheath gas temperature 350  $^{\circ}\text{C}$

Sheath gas flow 12 L/min

Capillary 3500 V

Nozzle 500 V

### **3.6 The RNA methyltransferase Dnmt2 methylates DNA in the structural context of a tRNA**

#### HPLC -MS analysis

0.5  $\mu\text{g}$  of either Dnmt2 treated or untreated tRNA was dissolved in 20 mM  $\text{NH}_4\text{OAc}$  pH 5.3 and digested to nucleosides as described above. Additionally the commercial oligomer MH569 containing 5-methyldeoxycytidine was digested and used as a reference sample for MS fragmentation experiments. Commercial 5-methylcytidine was used as a reference for LC-MS parameter optimization.

The digested RNA was analyzed on the Agilent 1260 series equipped with a diode array detector (DAD) and Triple Quadrupol mass spectrometer Agilent 6460 using the method “dnucs2” for analysis of  $\text{m}^5\text{dC}$  and  $\text{m}^5\text{C}$  in MRM mode.

### **3.7 rRNA:5-methylcytosine-methyltransferase (rRNA:m5C-MTase) activity of yeast proteins Nop2, Ynl022 and human proliferation-associated antigen p120**

Isolated rRNA (final concentration 50-100  $\text{ng}/\mu\text{L}$ ) was dissolved in 20 mM  $\text{NH}_4\text{OAc}$  pH 5.3 and digested as described above.

The resulting nucleoside mixture was injected on the Agilent 1260 series equipped with a diode array detector (DAD) and Triple Quadrupol mass spectrometer Agilent 6460 using the method “nucs4” with the MS parameters described above with +400 V EMV for highest sensitivity.

#### **3.7.1 Calibration for quantification of m5C in rRNA**

Adenosine, cytidine and 5-methylcytidine were purchased from Sigma-Aldrich (Munich, Germany) and dissolved in pure water to give 10 mM stock solutions each. All nucleosides were combined in one stock solution of 1 mM adenosine and cytidine and 1  $\mu\text{M}$   $\text{m}^5\text{C}$ .

#### **3.7.2 Detection and analysis of calibration samples**

This stock solution was diluted 1:10 with pure water to give final solutions of 100  $\mu\text{M}$ , 10  $\mu\text{M}$ , 1  $\mu\text{M}$  and 0.1  $\mu\text{M}$  A and C; and 100 nM, 10 nM, 1 nM and 0.1 nM  $\text{m}^5\text{C}$  respectively. 2, 5, 10  $\mu\text{L}$  of these dilutions were injected with the aforementioned LC-MS/MS method starting with the lowest concentrated one. The calibration solutions were measured as triplicates, #1 and #2 before sample measurements and #3 after sample measurements to exclude MS sensitivity loss.

MS-peaks were identified by Agilent Masshunter Qualitative Analysis software using the “find compounds by MRM” tool and MS calibration curves were plotted in Microsoft Excel and a linear regression curve applied. Additionally for A and C the UV peaks at 254 nm were integrated and the areas found, were used for a UV calibration.

### **3.7.3 Calculations for quantification of m<sup>5</sup>C in rRNA**

Digested samples were measured 3 times as mentioned above and analyzed once using the MS calibration data of C, A and m<sup>5</sup>C, and twice using the UV calibration data of C, A and the MS calibration data of m<sup>5</sup>C, whereas each set of sample measurements had its own set of calibration runs.

### **3.7.4 Calculations for quantification of injected rRNA**

The areas found for A and C in each sample, were divided by the respective gradients of the linear regression of the calibration curves to reveal the amount of A and C in fmol per measurement. Note that all measured samples were in the range of the calibration measurements. The amounts of C and A per rRNA molecule and for the DNazyme fragments of 25S rRNA were calculated and used to give the amount of injected molecules in the sample. (25S rRNA: 889 A, 662C; 25S rRNA domain I-IV: 757 A, 545 C; 25S rRNA domain V-VI: 142 A, 117 C)

### **3.7.5 Calculation for m<sup>5</sup>C quantification**

The area found for m<sup>5</sup>C by mass detection in each sample was divided by the respective gradient of the linear regression of the calibration curve to reveal the amount m<sup>5</sup>C in fmol per measurement. The found amount of m<sup>5</sup>C in fmol was then divided by the amount of found RNA in fmol to give the number of m<sup>5</sup>C residues per mol RNA. The LOD of m<sup>5</sup>C was 5 fmol and the LOQ 10 fmol.

## 4 Results

### 4.1 General analytics of tRNA *E.coli* treated with monofunctional coumarins

Labeling of nucleic acids, especially RNA, is of paramount importance in the life sciences to achieve a better understanding on life processes. Such a label should exhibit at least 2 of the 3 following characteristics: (I) it should be specific for a defined target (*e.g.* a modified nucleoside) in the RNA, (II) it should display a useful functionality like fluorescence or an affinity tag and (III) it should be sequence-specific. A broad literature search revealed 4-bromomethyl-7-methoxy-coumarin (BMB) as the only candidate for (I) selectivity for modified nucleosides pseudouridine and 4-thiouridine while (II) displaying fluorescence. Considering these interesting characteristics, it might be possible to further explore the structure-function relationship of similarly substituted coumarins to achieve higher or shifted selectivity for certain modified nucleosides. BMB and 5 other 4-bromomethylcoumarins can be seen in Figure 8A. They are referred to as monofunctional coumarins, since their fluorescence is the only function remaining after the reaction with the RNA as they have no further exploitable substituent for bioconjugate reactions. Figure 8B shows the proposed reaction schemes for bromomethylcoumarins with the nucleobases uridine and 4-thiouridine.

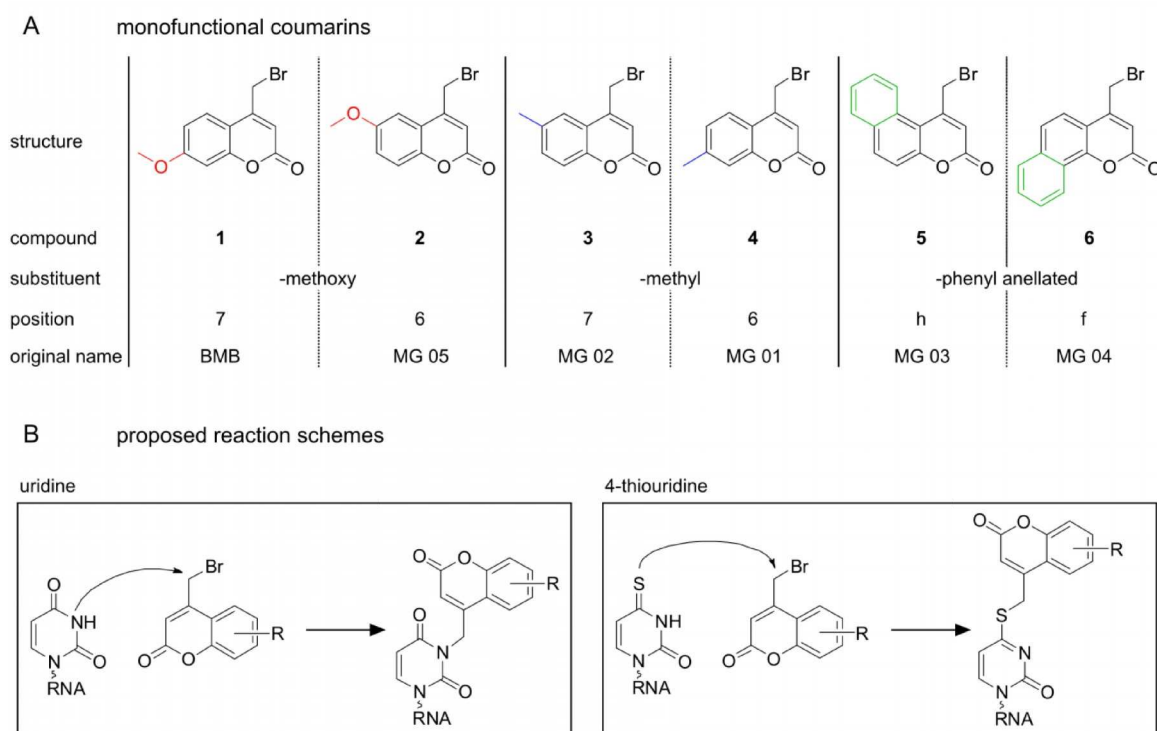


Figure 8 Overview of the used monofunctional coumarins and the general reaction schemes with nucleosides uridine and 4-thiouridine.

#### 4.1.1 Gel assay of conjugated tRNA

The commercially available coumarin 4-bromomethyl-7-methoxycoumarin (**BMB**), described in the literature [78, 83] as selective for modified uridine residues, was reacted with total tRNA *E.coli* under conditions 1 (0.1 M phosphate buffer pH 6.5, 75 % DMSO, 37°C, 5 hours), precipitated to remove unreacted coumarin and analyzed on a polyacrylamide gel. In parallel the 7-methoxy isomer and 4 further monofunctional 4-bromomethylcoumarins were treated accordingly. The results can be seen in the upper panels of Figure 9. The tRNAs are visualized by UV irradiation of the gel leading to a blue fluorescence of the conjugated coumarin-tRNA. For **BMB** and its isomer **2**, similar fluorescence, only slightly reduced for **2**, can be observed. The methyl-substituted labels **3** and **4** show nearly no detectable fluorescence, which could be due to a low reactivity towards tRNA or due to low fluorescence intensity. In comparison the phenyl-anellated compounds **5** and **6** display the highest fluorescence, which is most likely caused by a higher fluorescence of the enlarged aromatic system. A second set of experiments was conducted in a similar way with slightly changed reaction conditions. Reaction conditions 2 comprise a more alkaline pH, slightly reduced DMSO contents and a shorter incubation time. The gel resulting from this experiment can be seen in the lower panels of Figure 9. Compared to conditions 1 the detected fluorescence is increased, which indicates a higher reactivity of the coumarins with the tRNA.

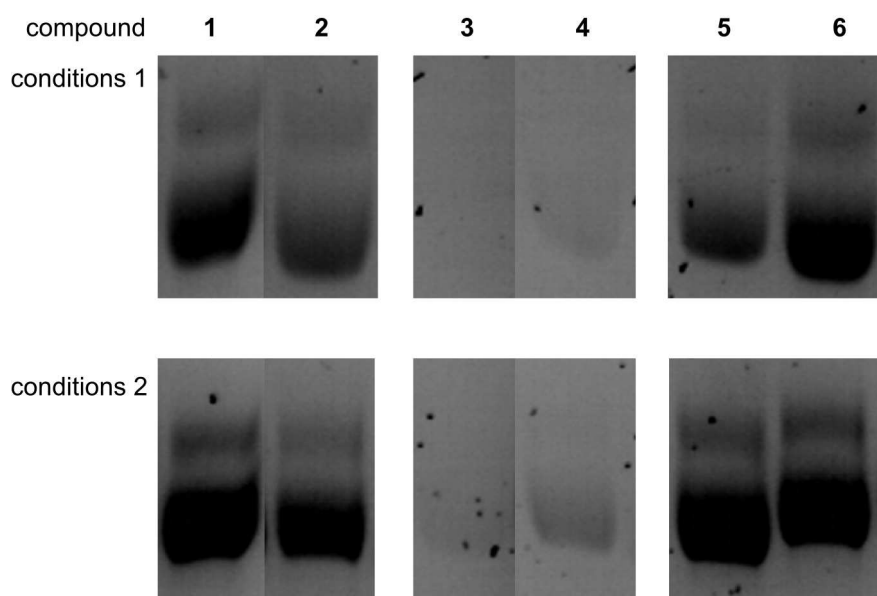


Figure 9 10% PAGE of total tRNA *E.coli* treated with 6 monofunctional coumarins under 2 sets of conditions. Pictures were developed under UV light at 365 nm monitoring the fluorescence of the coumarins.

#### 4.1.2 LC-UV and LC-MS/MS analysis

The gel analysis, revealing a readily detectable conjugation reaction of tRNA with most of the coumarins, gave rise to the question with which ribonucleotides the agents reacted. Each of the coumarin treated samples were digested to nucleosides and injected on LC-MS separately. The chromatography was tuned to clearly separate the hydrophilic nucleosides and the more hydrophobic coumarin-conjugated nucleosides, as can be seen in Figure 10A. Apart from chromatographic separation, coumarin-conjugates can be identified by their typical absorption at 320 nm, whereas unlabeled nucleosides absorb at 254 nm only. In a first run, photometric and mass-spectrometric data was obtained and all peaks displaying a strong absorption at 320 nm in the later part of the chromatogram were chosen for further analysis. This comprised selection of the found masses for filtering in quadrupole 1 followed by collision-induced-dissociation (CID) in a second run. The found MS/MS fragment spectra of these selected masses revealed a fragmentation pattern similar to nucleosides; a loss of 132 Da, which coincides with a cleavage of the glycosidic bond leaving the protonated, coumarin labeled nucleobase. For all six coumarins the reaction products could be identified as uridine, 4-thiouridine, pseudouridine and guanosine conjugates. An overview about the collected spectra for BMB conjugates can be seen in Figure 10B. For each coumarin a multiple-reaction-monitoring (MRM) method was developed using the MS/MS results (see appendix Table A 1).

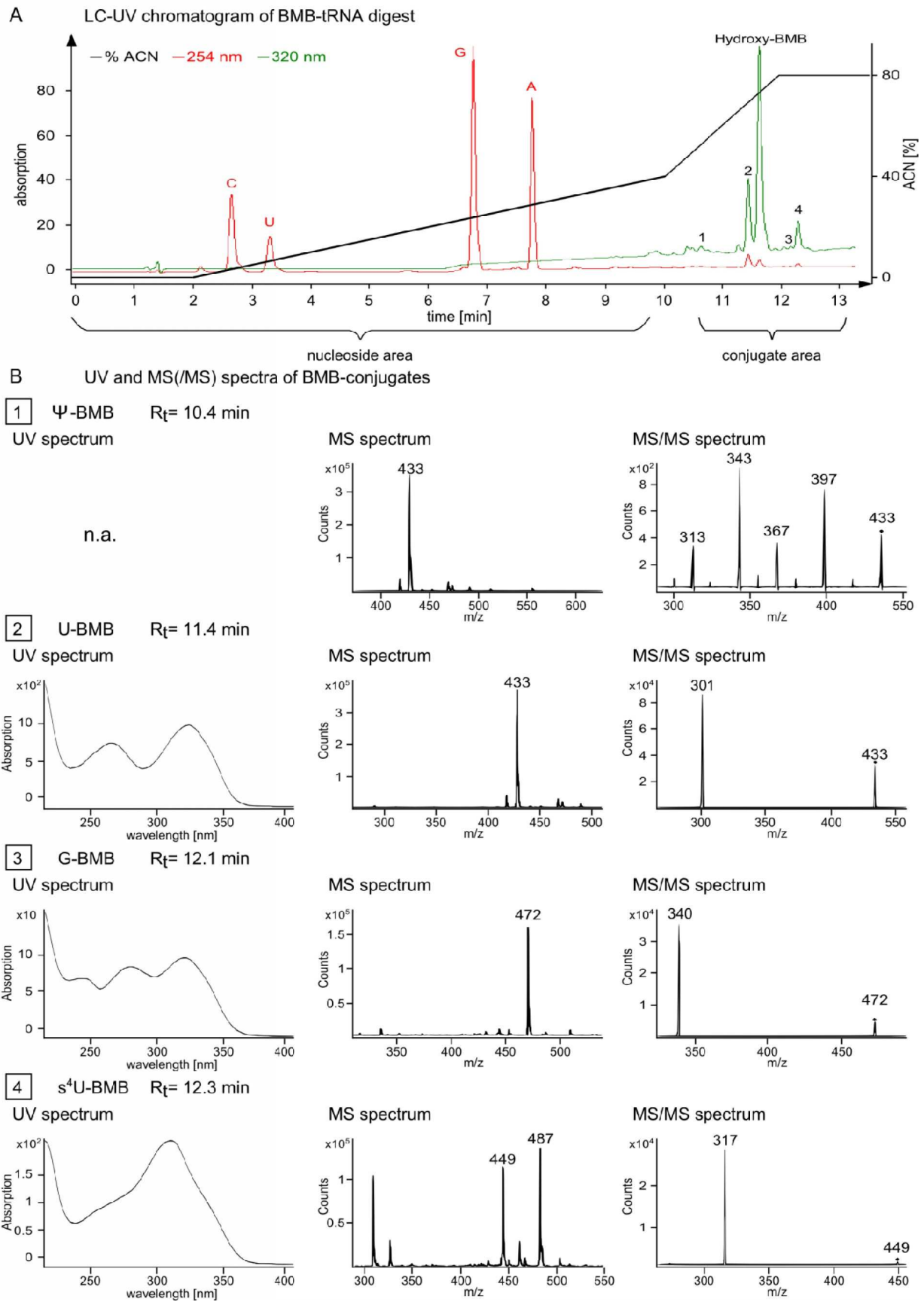


Figure 10 LC-UV-MS analysis of tRNA conjugated with BMB

A Overlay of UV chromatograms at 254 nm monitoring nucleoside absorption (red) and 320 nm monitoring conjugate absorption (green). The black curve indicates the used acetonitrile gradient for separation of hydrophilic nucleosides from the more hydrophobic conjugates. The possible conjugates are numbered from 1-4. B Photometric and mass-spectrometric data of all BMB conjugates 1-4.

### 4.1.3 LC-MS/MS data analysis

Figure 11A shows the first step of coumarin functionalization analysis, starting with tRNA derivatization followed by LC-MS/MS analysis. The measured nucleoside and conjugate peaks of LC-MS/MS analysis are integrated and the resulting areas are referred to as the raw data. Figure 11B shows the second step, which is the workflow of data analysis taking the following parameters into account:

- Amount of injected tRNA per sample for intersample comparability
- Ionization efficiency of the conjugates
- Abundance of the major nucleosides U, G and the modified nucleosides  $\Psi$  and  $s^4U$  in the used tRNA *E.coli* samples

A detailed description of the data analysis workflow to receive final statements on target selectivity is given on the following pages.

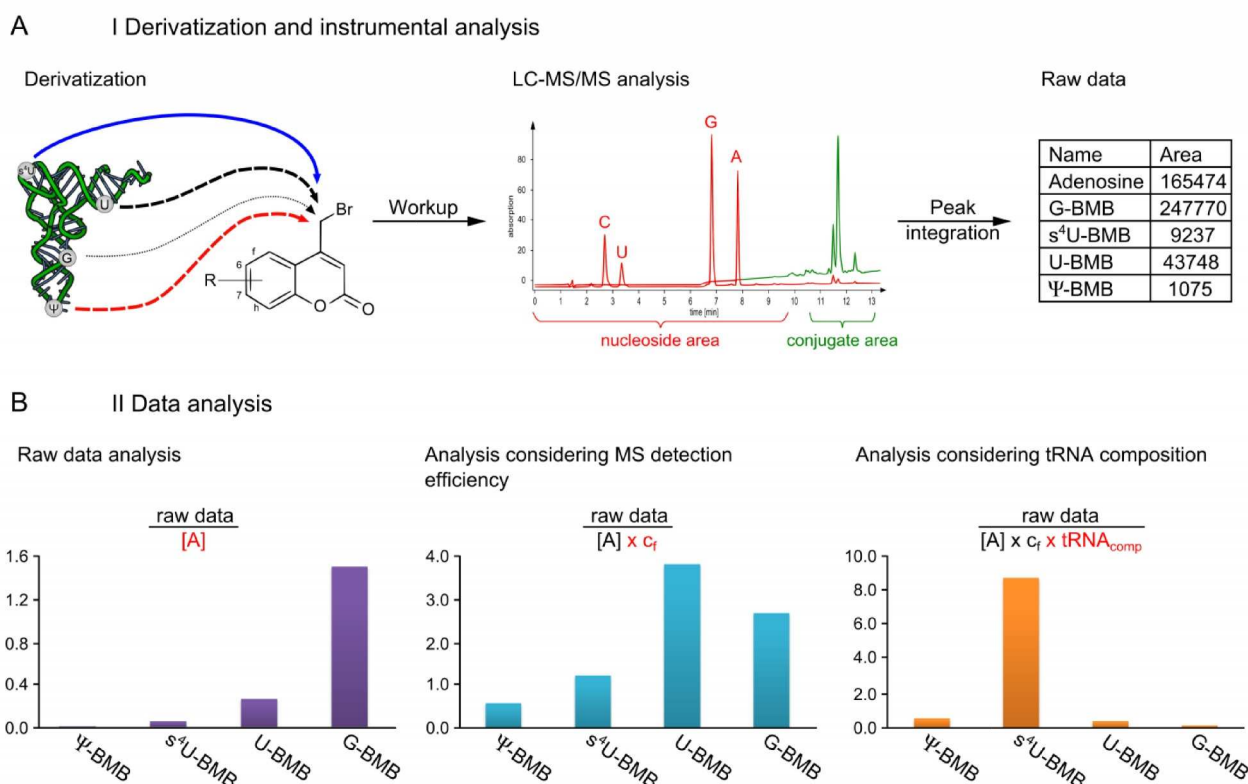


Figure 11 General scheme for data acquisition and analysis of coumarin selectivity studies.

A Derivatization and instrumental analysis of total tRNA *E.coli* treated with BMB. B Workflow of raw data analysis. The diagram on the right shows the raw data normalized to the internal standard adenosine. The diagram in the middle shows the adjusted data considering the different ionization efficiencies of the nucleoside conjugates. The diagram on the right is, additionally, normalized to the nucleoside abundance of the used total tRNA *E.coli*.



For comparability of different samples it is necessary to normalize all measured areas to an internal standard. In this case adenosine was chosen since all tRNA samples contain the same amount of adenosine and it is not influenced by the reaction with the coumarins. The results of the normalization can be seen in Table 13 for BMB conjugates.

IIa

$$\text{Adjusted area} = \frac{\text{area}}{\text{area}[A]}$$

Table 13 Calculation of adjusted areas for BMB conjugates

Name	Area	Adjusted area
Adenosine	165474	1
G-BMB	247770 : 165474	1.497
<sup>4</sup> sU-BMB	9237 : 165474	0.056
U-BMB	43748 : 165474	0.264
Ψ-BMB	1075 : 165474	0.006

The first diagram of Figure 11B shows the raw data normalized to the adenosine abundance. In this case guanosine seems to be the most favored reaction partner of BMB.

Mass spectrometry relies on ionization of the analytes and quantitative measurements are therefore influenced by differences of ionization efficiency among analyte species. In case of nucleosides, which are ionized by nucleobase protonation in the ESI source, the differences are guided by their basicity. For this reason a correction factor ( $c_f$ ) is needed to compare MS data of different nucleosides and nucleoside-coumarin conjugates. This correction factor can be calculated for each analyte by using the absorption at 320 nm of the coumarin chromophore. With the help of this factor the conjugates of a certain coumarin can be compared (Table 15).

IIb

$$c_f = \frac{[\text{MS/MS}]}{[\text{UV}]}$$

Table 14 Determination of the correction factors for BMB conjugates

	MS/MS	UV	$c_f$
G-BMB	134351	2.4	55980
<sup>4</sup> sU-BMB	26017	5.7	4564
U-BMB	3469	0.5	6938
Ψ-BMB	57	0.05	1140

IIc

$$\text{Corr. area} = \frac{\text{adjusted area}}{c_f}$$

Table 15 Calculation of corrected areas for BMB conjugates

Name	Adjusted area / $c_f$	Corrected area
G-BMB	1.497 : 55980	$2.67 \times 10^{-05}$
<sup>4</sup> sU-BMB	0.056 : 4564	$1.22 \times 10^{-05}$
U-BMB	0.264 : 6938 x	$3.81 \times 10^{-05}$
Ψ-BMB	0.006 : 1140	$5.70 \times 10^{-06}$

The second diagram in Figure 11B shows the data of Table 15 after adjustments with the correction factor  $c_f$ . It is now apparent that G-BMB was overrepresented due to a high ionization efficiency of the guanosine base compared to a relatively weak ionization of the more acidic uridine residues. This diagram reveals the absolute abundance of each conjugate in the reacted tRNA and shows uridine as the main product. However to perform statements of selectivity, the analysis must also take into account that uridine and guanosine are more frequent in the substrate tRNA than the modified uridine residues  $s^4U$  and  $\Psi$ . The composition of total tRNA *E.coli* can be accessed by digesting untreated tRNA followed by LC-UV analysis and using the absorption of each nucleoside at its  $\lambda_{max}$  and division by its corresponding molar extinction coefficient known from literature [106](Table 16).

IId

$$\text{tRNA composition} = \frac{\text{mol. frequency (nucleoside)}}{\text{mol. frequency (uridine)}} \times 10$$

Table 16 determination of nucleoside composition of total tRNA *E.coli*

Name	$\lambda_{max}$	$\epsilon_{max}$ ( $\times 10^{-3}$ )	Area ( $\lambda_{max}$ )	mol. frequency Area ( $\lambda_{max}$ ) / $\epsilon_{max}$	Nucleoside abundance
G	253 nm	13.6	284	20912	24.4
$s^4U$	331 nm	21.2	2.6	124	0.14
U	262 nm	10.1	8.7	8616	10
$\Psi$	263 nm	8.1	8.3	1027	1.06

The data of tRNA composition is then used to equalize the adjusted areas to the nucleoside abundance, which finally reveals the favorite reaction partner of the coumarin (Table 17).

Table 17 Abundance of coumarin conjugates considering tRNA composition

Name	Adjusted area ( $\times 10^{-5}$ )	nucleoside abundance	Adjusted area / nucleoside abundance
G-BMB	2.67	24.4	0.11
$s^4U$ -BMB	1.22	0.14	8.74
U-BMB	3.81	10	0.38
$\Psi$ -BMB	0.570	1.06	0.54

The diagram on the right of Figure 11B is a visualization of the results shown in Table 17, illustrating 4-thiouridine as the favorite reaction partner of BMB.

## 4.2 Structure-function relationship of monofunctional coumarins

### 4.2.1 Influence of the substitution pattern on reactivity and selectivity

LC-MS/MS analysis of all six coumarin tRNA conjugates, followed by data processing explained in 4.1.3, revealed a higher reactivity of the C6 substituted coumarins compared to their C7 isomers (Figure 12A). The electron pushing substituents at position 7 lead to an increased electron density in proximity to the 4-bromomethyl function, thereby reducing its electrophilic activity and leading to the observed lower reactivity. This effect is less protuberant for the mildly inductive methyl-substituent of compounds **3** and **4**.

Data processing considering the tRNA composition reveals that all six coumarins favor 4-thiouridine as their main reaction partner and in contrast to reactivity the C7 substituted isomers display the highest selectivity. For s<sup>4</sup>U coumarin **3**, which is the least reactive of the coumarins, is the most selective.

### 4.2.2 Influence of the reaction conditions on reactivity and selectivity

Comparing reaction condition 1 (pH 6.5) with condition 2 (pH 8.25), shows an increased reactivity of uridine most likely caused by a deprotonation of its N3 which leads to a stronger nucleophile and higher reactivity. For the modified uridine residues s<sup>4</sup>U and Ψ the reactivity is unaltered, whereas the reactivity is decreased for guanosine.

Accounting for nucleoside abundance, all coumarins conjugated under reaction conditions 1 show the highest selectivity for s<sup>4</sup>U except for coumarin **3**, which is most selective at conditions 2 (Figure 12B).

Further Reading:

**Structure-function relationship of substituted  
bromomethylcoumarins in nucleoside specificity of RNA  
alkylation**

S. Kellner, L. Kollar, A. Ochel, M. Ghate & M. Helm  
Bioconjugate Chemistry (ACS publications),

manuscript submitted

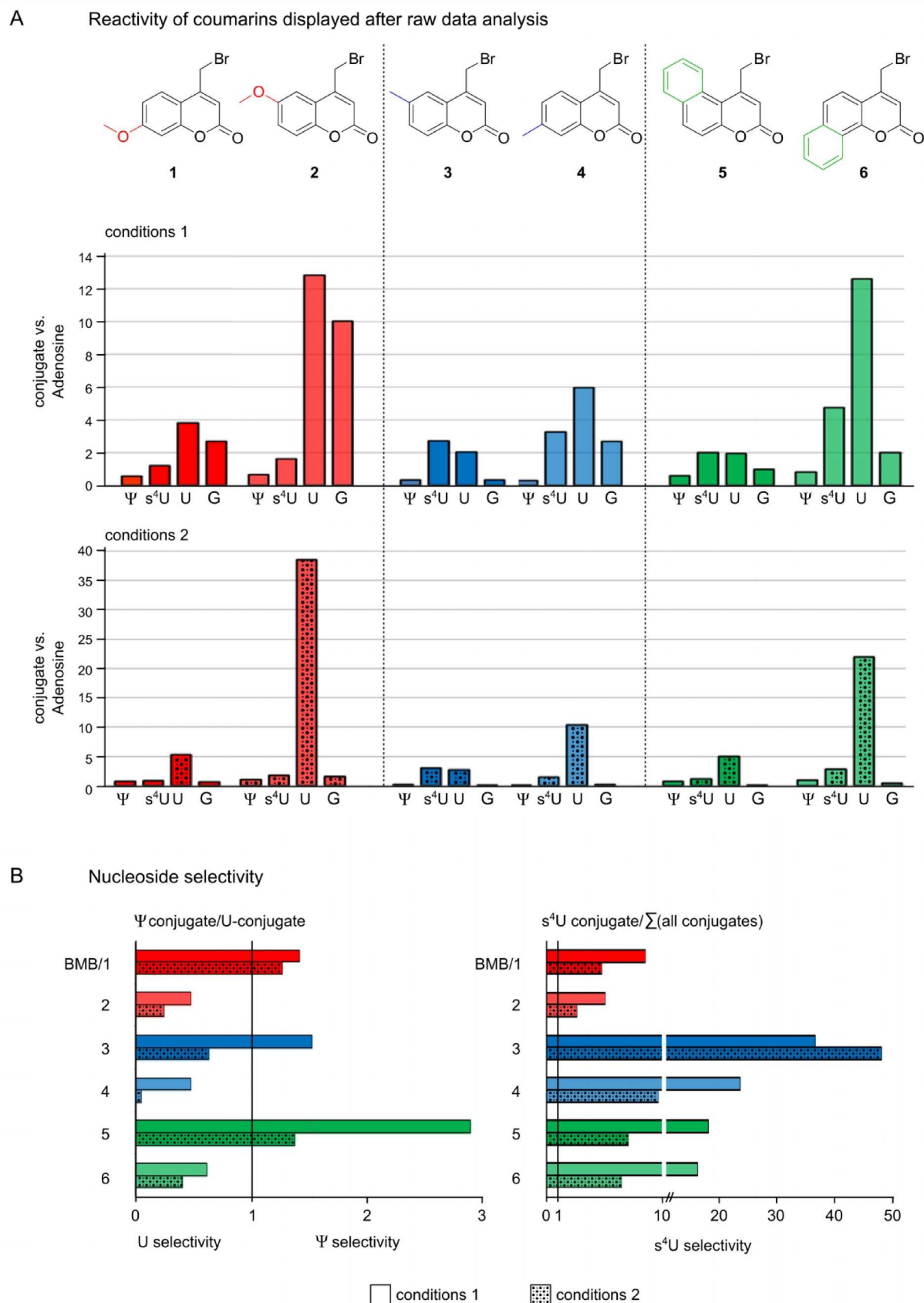


Figure 12 Processed LC-MS/MS data of all six monofunctional coumarins.

A Reactivity of the coumarins towards their nucleophilic targets U, G s<sup>4</sup>U and Ψ under slightly acidic conditions 1 and more alkaline conditions 2. B Overview over nucleoside selectivity of the different coumarins and conditions. Numbers smaller 1 indicate selectivity towards uridine, whereas bigger numbers indicate selectivity for the respective modified nucleoside.

### 4.3 A general RNA alkylating reagent with multifunctional properties – N3BC

The structure-function relationship results of the monofunctional coumarins indicate that the coumarin scaffold is suitable for tuning 4-thiouridine selectivity and furthermore towards selectivity for uridine residues in general. This is of great interest, since post-synthetic labeling of either native or artificial RNA is cumbersome but crucial for studies of these macromolecules. Since not all RNAs contain modified nucleosides, it is desirable to target a reactive group of the RNA that is frequently distributed, for example a specific functionality of a major nucleoside. As it can be seen in Figure 12 compound **2** already displays a high selectivity for uridine compared to the other coumarins. In the light of our previous data on coumarin reactivity, it is likely that a methoxy substitution at position 6 of the coumarin body is advantageous to achieve uridine selectivity. The methoxy substitution at position 6 of the coumarin has an electron donating effect, which increases the electron density at the electrophilic 4-methylbromo function. This insight can be exploited by designing a coumarin that has an electron withdrawing substituent at C7 and is additionally usable as a functional group for later reactions. The azido function of 7-azido-4-bromomethylcoumarin (N3BC) (Figure 13) combines both requirements by (a) being an electron withdrawing substituent and (b) having multifunctional properties like crosslinking under UV illumination or undergoing Cu(I) catalyzed azide-alkyne Huisgen cycloaddition (CuAAC or click) reaction in the presence of a terminal alkyne. The first synthesis of this compound was done by Salifu Seidu-Larry and later syntheses were performed by Maria Adobes-Vidal in our lab.

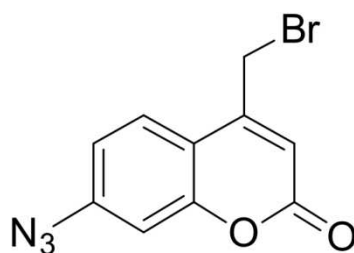


Figure 13 Chemical structure of N3BC

### 4.3.1 Reactivity with nucleosides and homopolymers

N3BC was first tested with nucleosides and the principle reaction conditions were established using cellulose thin layer chromatography (TLC). These conditions comprise a high amount of DMSO (50-70 %) in the final reaction mixture due to the low solubility of the coumarin in aqueous solvents. Other possible hydrophobic solvents like acetonitrile or methanol were excluded at this stage since they did not support the nucleophilic substitution reaction. The same was observed for aqueous buffers with acidic pH (below pH 6.0) and best reactivity was found for buffers in the range of pH 8-9. These preliminary conditions showed a reaction of the coumarin with uridine but not with the other three major nucleosides. To confirm these findings the samples of the conjugated nucleosides were subjected to HPLC analysis with an instrument equipped with a sensitive diode array detector (DAD) and fluorescence detector (FLD). As already described for the monofunctional coumarins, the gradient was chosen to clearly separate the early eluting nucleosides and the later eluting conjugates. A mock incubation revealed several fluorescent N3BC degradation products. The four separately incubated nucleosides showed these degradation products as well, but only uridine containing samples showed an additional fluorescent peak in the later part of the chromatogram. Since this additional peak most likely presents a product peak, it was purified by preparative HPLC and used for later LC-MS/MS analysis.

Homopolymeric RNA was used in a next set of experiments, since their polymeric characteristics allowed precipitation after derivatization to remove free N3BC degradation products. After precipitation and a washing step the RNA was subjected to subsequent digestion to nucleosides and HPLC-FLD analysis. No degradation products of N3BC could be found and only Poly-U containing samples showed a fluorescent peak in the chromatogram.

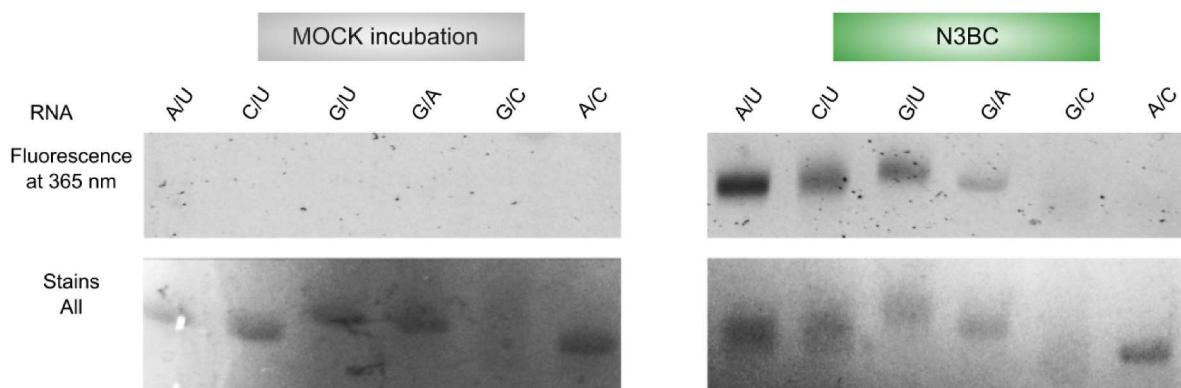
### 4.3.2 Reactivity with artificial bipartite oligonucleotides

Since homopolymers are prone to a large size distribution and RNA unlike behavior on urea-gels, it was necessary to move from these simple substrates to more complex substrates that behave more like natural RNA. Oligoribonucleotides (ORN) of bipartite composition (containing only 2 of the possible 4 major nucleobases) were conjugated with N3BC. After ethanol precipitation and gel filtration for removal of unreacted coumarin, sample aliquots were either analyzed on a urea-gel or digested for further 2d-TLC or LC-MS/MS analysis. Figure 14A shows for all uridine containing ORNs treated with N3BC a strong fluorescence on the gel upon UV irradiation. Additionally, a less intense fluorescence can be observed for the G/A and G/C ORN indicating a side reaction of N3BC with guanosine. A 2D TLC analysis after digestion of these samples shows fluorescence for the uridine containing permutations, only, most probably due to a lower detection limit on the TLC plate (Figure 14B).

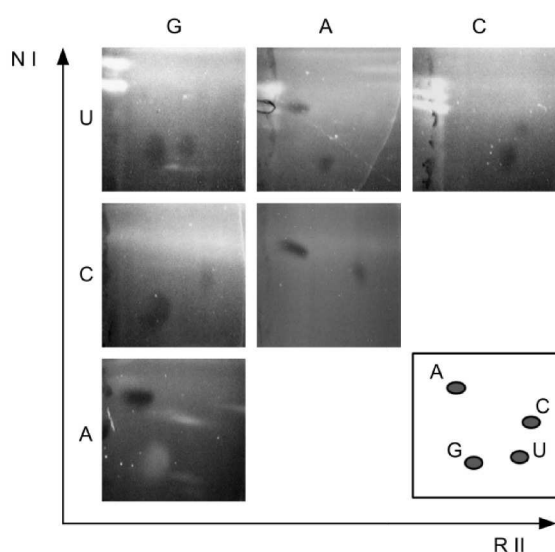
### 4.3.3 LC-MS analysis of treated RNA

In the previous HPLC runs for nucleoside conjugate analysis, the suspected products for uridine and pseudouridine were purified and analyzed by LC-MS/MS. Here, a LC-MS/MS device from Thermo Electron (Dreieich, Germany) was used for product identification by mass transition. MS and MS/MS spectra were recorded for U-N3C and  $\Psi$ -N3C. For the reaction product of guanosine with N3BC, isolation was not possible due to the low reaction yield. The N3BC-treated and digested U/G sample was analyzed on LC-MS and trace amounts of an analyte displaying the mass and fragmentation pattern of a possible G-N3C conjugate were found. It was not possible to calculate a correction factor  $c_f$  for these measurements, since the machine did not provide a functional UV detector. Under assumption of similar ionization efficiency, U-N3C is most likely the main product with a predominance of at least factor 35 (Figure 14C) towards G-N3C.

## A Gel analysis of bipartite ORN



## B 2D-TLC analysis of bipartite ORN



## C LC-MS/MS analysis of bipartite ORN

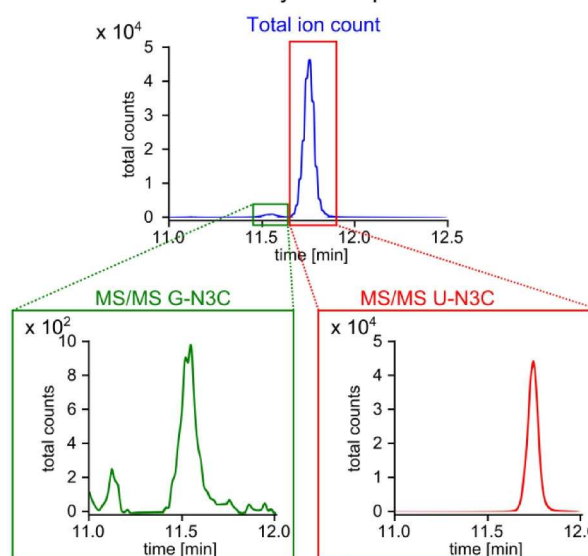


Figure 14 Analysis of N3BC treated bipartite oligomers.

A Urea PAGE of mock and N3BC treated oligoribonucleotides of bipartite composition. The gel is developed under UV illumination at 365 nm monitoring the coumarin fluorescence that can be observed for all 3 uridine containing permutations. Interestingly the G/A and G/C ORN also display minor coumarin fluorescence, indicating a side reaction of N3BC with guanosine. B 2D TLC analysis of digests from N3BC treated bipartite ORN. Only in the first row strong coumarin fluorescence can be observed upon UV irradiation. The plate in the lower right corner shows the spots of main nucleotide migration reported in the literature [107]. C LC-MS/MS analysis of the U/G ORN digest treated with N3BC. In blue the total ion count (TIC) displays two peaks. Below the magnified peaks show that the smaller peak (green) is derived from G-N3C whereas the larger peak is identified as the U-N3C conjugate (red).



#### 4.3.4 Modulation of labeling efficiency by stepwise change of reaction conditions

A crucial point in labeling of macromolecules is the yield. As Figure 15C shows, under the reaction conditions previously used (70% DMSO, 10 mM N3BC, pH 8.5) 50 % of all uridines, in the denatured and therefore completely accessibly RNA, were conjugated. By lowering the DMSO content or the N3BC concentration the alkylation efficiency can be reduced to a more useful yield of 1 to 10 % of all uridines. In the first case, i.e. decreased DMSO concentration, it is most likely that the tRNA secondary structure stays intact and, apart from reduced alkylation efficiency, only uridines not employed in base-pairing have a free target *N3* for nucleophilic attacks (see Figure 15A). In the second case, i.e. decreased concentration of N3BC, the RNA is still denatured due to the high DMSO content and therefore a statistical derivatization of 1-50 % uridines is likely as Figure 15 B and C suggests. These observations suggest that by tuning the reaction conditions of N3BC alkylation either (A) all uridines in a RNA can be statistically labeled to a desired extent or (B) only uridines in loop regions are labeled.

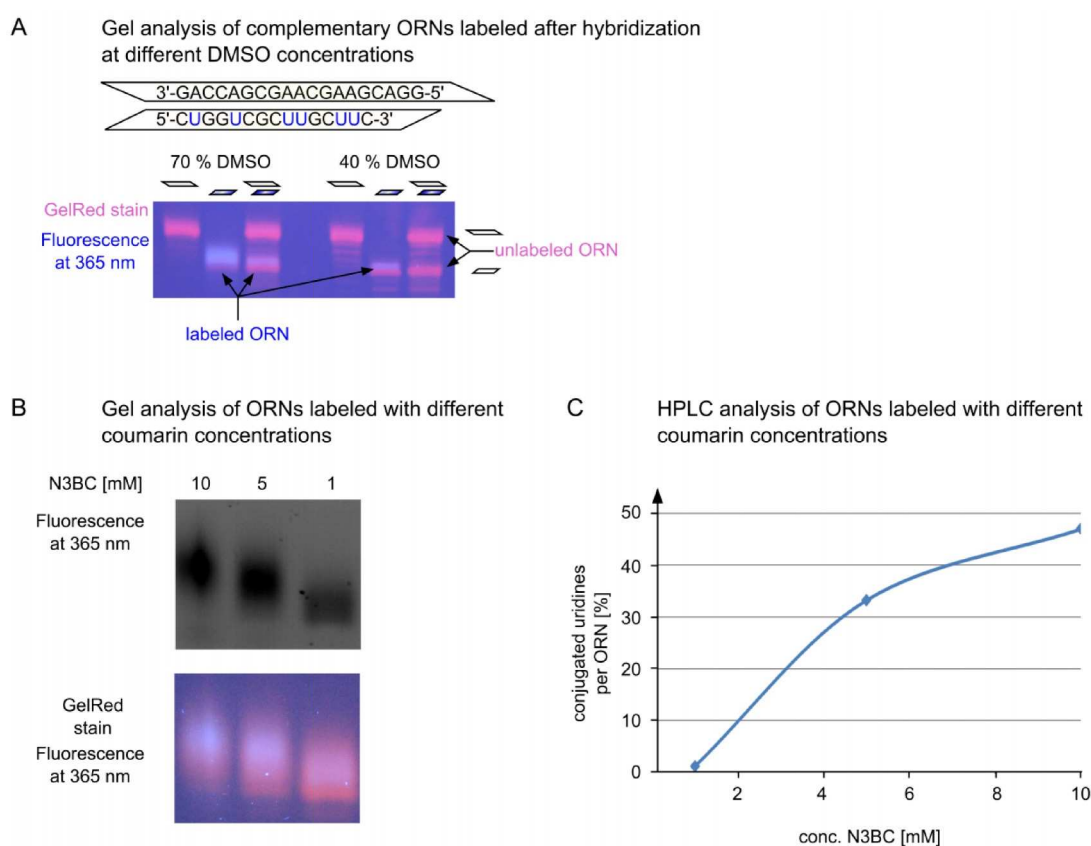


Figure 15 Modulation of labeling efficiency by changing either the DMSO concentration or the N3BC concentration.

A Gel assay displaying the results for the modulation by changing the DMSO content. On the left, labeling with 70 % DMSO in the reaction mix leads to successful alkylation of both free and hybridized RNA (lane 3). On the right, with 40 % DMSO content, no coumarin fluorescence can be detected for the hybridized RNA but for the free RNA. B Gel assay of labeling with decreasing coumarin concentrations reveals a minor fluorescence for RNA alkylated with a low concentration of N3BC. C HPLC analysis of nucleoside digests of RNA shown in B. Here, the ratio of U/A per sample was measured and compared to non alkylated RNA digests to reveal the percentage of coumarin labeled uridines.

#### 4.3.5 Click reaction of successfully labeled RNA

With lower alkylation yields, the coumarin fluorescence is too low for direct detection via UV illumination and high sample amounts would be needed to monitor the labeled RNA. To circumvent this, we took advantage of the azido-function of the coumarin, which was shown to stay intact throughout the labeling process. It is thus available for further functionalization *e.g.* with a fluorescent dye. The N3BC labeled RNA undergoes readily the CuAAC reaction in the presence of Cu(I), THPTA ligand and a terminal alkyne attached to a strong fluorophore like Alexa647 or fluoresceine, respectively. This post-labeling reaction allows an easy and versatile RNA functionalization, which can be used for studies monitoring certain RNAs like aptamer RNAs or siRNAs by fluorescence or for surface immobilization experiments by biotinylation.

#### 4.3.6 Crosslinking reaction with cognate proteins

The azido-function of N3BC is nearly inert under physiological conditions. However, it needs light protection at all steps of derivatization and workup due to a pronounced photolability. Via a nitrene intermediate, the azide is decomposed to an aromatic amine, which cannot be used for the aforementioned click reaction with terminal alkynes. Nitrenes are electron deficient species undergoing a variety of reactions with nucleophiles in their proximity, *e.g.* water, phosphates or aminoacids containing nitrogens or sulfurs, for stabilization. This feature can be exploited for crosslinking of azide-labeled RNA with cognate proteins like RNA modifying enzymes. N3BC treated and control  $^{32}\text{P}$ -tRNA<sup>Phe</sup> from yeast were incubated with different enzymes like RNA:methyltransferases or pseudouridine synthases and illuminated for 30 minutes with UV light at 365 nm. For the N3BC treated tRNA strong crosslink bands were detectable on the SDS gel while no such crosslinks could be observed in the control tRNA. This crosslinking ability of N3BC gives rise to possible applications in protein-RNA binding studies.

Further Reading:

**A multifunctional bioconjugate module for versatile photoaffinity labeling and click chemistry of RNA**

S. Kellner, S. Seidu-Larry, J. Burhenne, Y. Motorin & M.

Helm

Nucleic Acids Res. 2011 Sep 1;39(16):7348-60.

Epub 2011 Jun 6

### 4.3.7 Reactivity towards DNA

N3BC, as an electrophilic, alkylating reagent, can be used for post-transcriptional derivatization of RNA by attack of the nucleophilic *N3* of uridines. The N3BC labeled RNA is then ready for a number of versatile applications. A target nucleophile similar to the *N3* in uridine exists in DNA as well, namely the *N3* of thymidine. Indeed, experiments with bipartite oligodeoxynucleotides (ODN) show comparable results as presented in section 4.3.2 only that thymidine is the main reaction partner instead of uridine. Figure 16 shows the urea-gels of the ribo- and deoxynucleotides treated with N3BC.

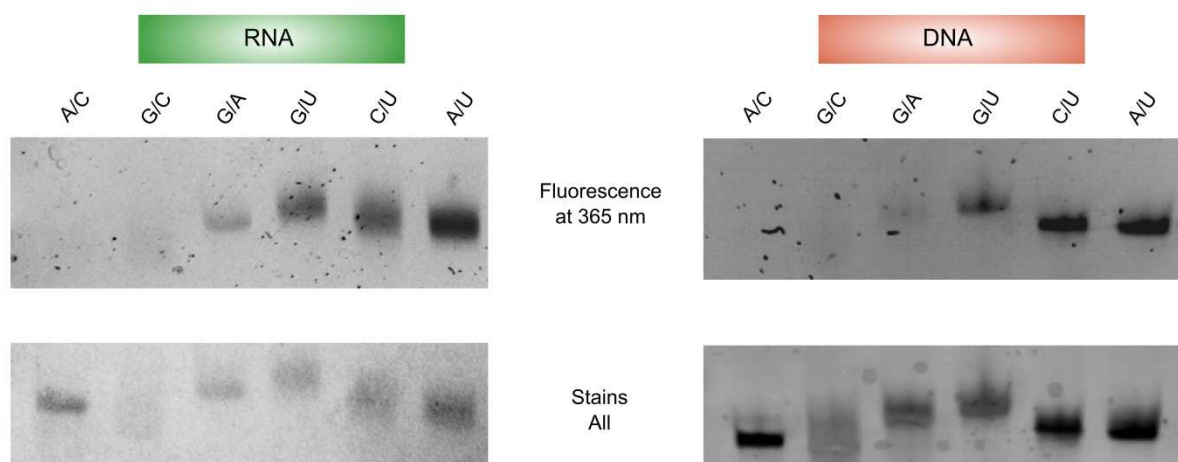


Figure 16 Urea-gels of bipartite oligonucleotides of RNA composition (left) and DNA composition (right). The gels are monitored by illumination at 365 nm and the fluorescence of the coumarin indicates successfully labeled oligonucleotides.

These experiments confirm the hypothesis stated above and expand the possible fields of application from RNA to DNA, allowing the same sort of studies as already described for RNA. N3BC is a multifunctional labeling agent for nucleic acids that contain either uridine or thymidine.

#### 4.4 PBC – a new multifunctional coumarin selective for 4-thiouridine

Guided by the previous structure-function relationship studies in section 4.2 another substituted 4-bromomethylcoumarin with the aim of selective 4-thiouridine conjugation was developed. The substituent must meet two requirements: (I) it must be a substituent with electronic properties similar to compound **3** to achieve the desired selectivity and (II) the substituent must present a further usable functional group. The substitution of compound **3**, a C7 methylation, leads to a more electron-rich bromomethyl group and is the reason for the reduced reactivity of **3**. Thus, a lower reactivity coincides with an increased selectivity for the best nucleophile yet identified in our experiments, namely the sulfur of 4-thiouridine. The substituted coumarin, combining both requirements, can be seen in Figure 17. 4-bromomethyl-7-propargyloxycoumarin, termed PBC, has an electron donating substituent at C7 like compound **3** or BMB and is therefore a good candidate to explore  $s^4U$  selectivity. It was synthesized in our lab by Maria Adobes-Vidal.

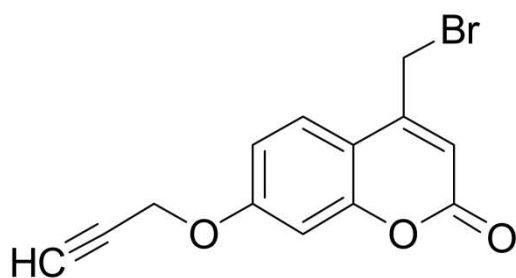


Figure 17 Chemical structure of PBC

#### 4.4.1 Analysis of treated tRNA

First tests of the new coumarin PBC were conducted under the previously described conditions 1 and 2 followed by gel analysis. As can be seen in Figure 18, compared to the other aforementioned coumarins, PBC treated tRNA from *E.coli* shows only a minor fluorescence comparable to compounds **3** and **4**. These differences in fluorescence intensity are most likely due to the different size of the aromatic systems and the resulting differing extinction coefficients. Additionally, a new set of reaction conditions with pH identical to conditions 2 (pH 8.25) but DMSO concentration reduced to 50 % were tested. These reaction conditions 3 were adjusted from conditions for s<sup>4</sup>U selectivity of PBC developed by Katharina Schmid. They are supposed to have a negative influence on the reactivity due to the decreased amount of the aprotic solvent DMSO, which could lead to an advantage in selectivity for the strongest nucleophile in the substrate tRNA, the sulfur of s<sup>4</sup>U. As expected the RNA in the lowest panel in Figure 18 displays the lowest fluorescence of all coumarins, which is caused by decreased reactivity. Further LC-MS/MS analysis of the PBC treated tRNA is necessary to conclude on selectivity in total numbers, as described for the monofunctional coumarins.

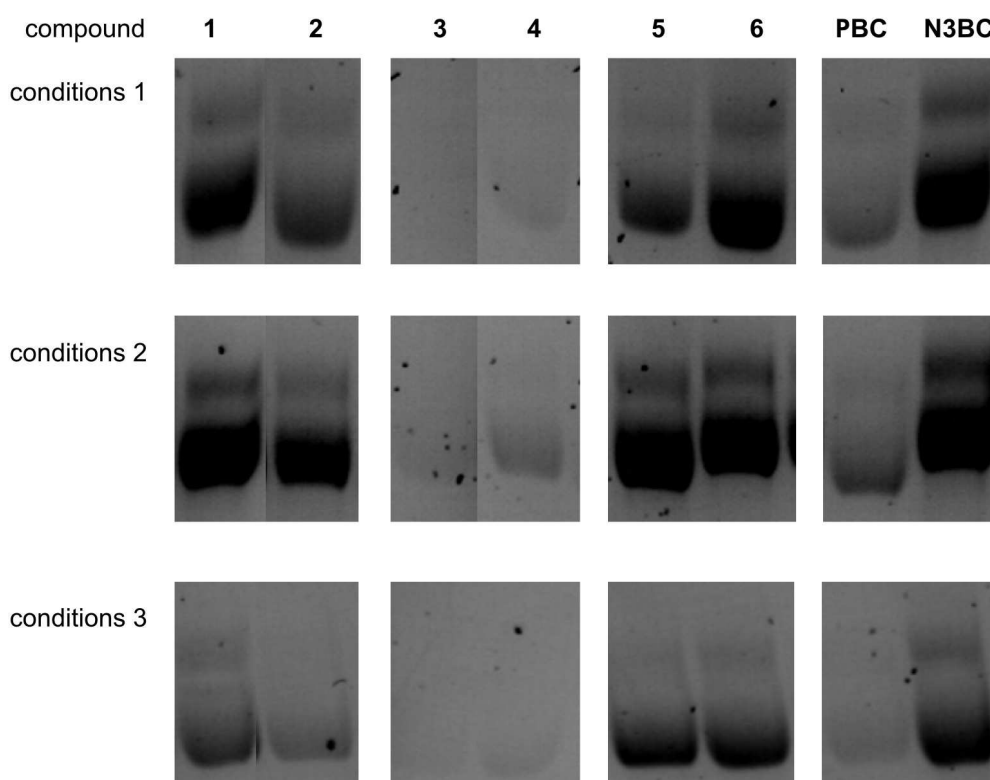


Figure 18 Gel analysis of total tRNA *E.coli* treated with all 8 coumarins, monitoring the coumarin fluorescence under UV illumination at  $\lambda=365$  nm

#### 4.4.2 Comparison with examined monofunctional coumarins

As already described for the monofunctional coumarins, the UV, MS and MS/MS spectra of PBC conjugates were recorded by digestion of PBC treated tRNA *E.coli*. The MS/MS transitions were used for sensitive detection of the found conjugates U-PBC, s<sup>4</sup>U-PBC, G-PBC and Ψ-PBC (see appendix Table A 2). Data of all 7 coumarins under all 3 conditions was processed, as described in section 4.1.3, to reveal the selectivity of the coumarins under consideration of tRNA composition. Figure 19 shows the resulting diagram for all conditions. For conditions 1, the highest selectivity is achieved for the new coumarin PBC (purple bar) that is 3 times higher than the selectivity of compound **3** (solid blue bar). This interesting behavior is not observed at elevated pH in conditions 2, where PBC displays no noticeable selectivity. As assumed for conditions 3 a reduced amount of DMSO in the reaction mixture does improve the selectivity of PBC for s<sup>4</sup>U (factor ~2000) and the monofunctional coumarins especially the C7 substituted compounds. It can be therefore concluded that a decreased reactivity caused by reduced DMSO content leads to a higher selectivity of the strongest nucleophile present in the chosen substrate tRNA, the sulfur atom of 4-thiouridine. Under conditions 3 the highest selectivity is achieved by compound **3** with a selectivity factor of ~2700. These results confirm the coumarin as a possible scaffold for achieving selectivity for modified ribonucleosides, as successfully demonstrated for the monofunctional compounds BMB, **3**, **5** and **6** and the multifunctional coumarin PBC.

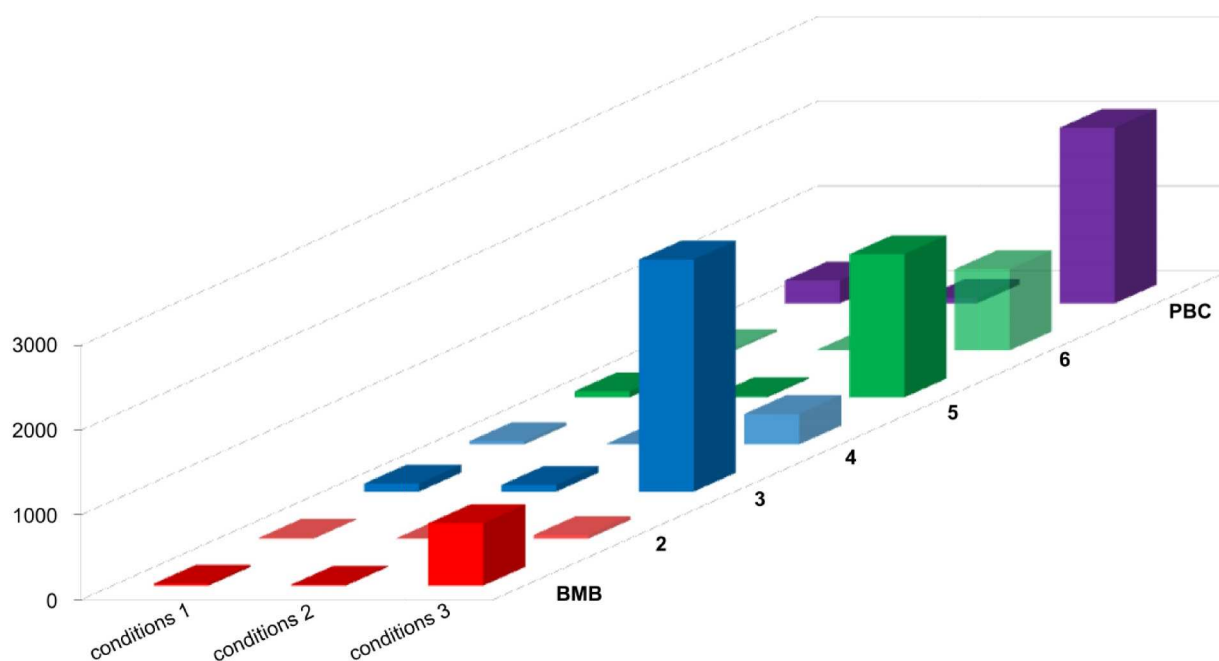


Figure 19 Diagram of processed data deriving from LC-MS/MS analysis of coumarin treated tRNA digests.

#### 4.4.3 Click reaction of PBC labeled RNA

While only slightly less selective than compound **3**, PBC, in addition to displaying a selectivity factor of more than 2000, has the added advantage of a second functional group that can be used for CuAAC type click reactions. This Cu(I) dependent cycloaddition reaction was studied with various fluorescent dyes *e.g.* Atto488, AlexaFluor 594 and Atto647N with azide groups attached to the fluorophores. A synthetic RNA oligomer containing the four canonical bases was used for PBC derivatization under conditions 2 which gives the highest yield for uridines, since no  $s^4U$  is present in the ORN. Additionally, a commercially available ODN with a synthetically introduced alkyne group was used for comparison with the PBC labeled ORN. PBC treated ORN and the commercial ODN were incubated with the Atto647N-azide in the presence and absence of Cu(I) and after acetone/ $LiClO_4$  precipitation analyzed on an urea-PAGE (see Figure 20). The PBC treated ORN and the commercial ODN show a bright blue band on the gel as they are both successfully clicked in the presence of Cu(I). Without Cu(I) the cycloaddition reaction does not take place and no fluorescence is detectable. The GelRed stain of the gel reveals a faint red RNA band below the blue fluorescent PBC-Atto647N ORN band, which is most likely RNA that was not labeled with PBC and subsequently not clicked with the fluorophore.

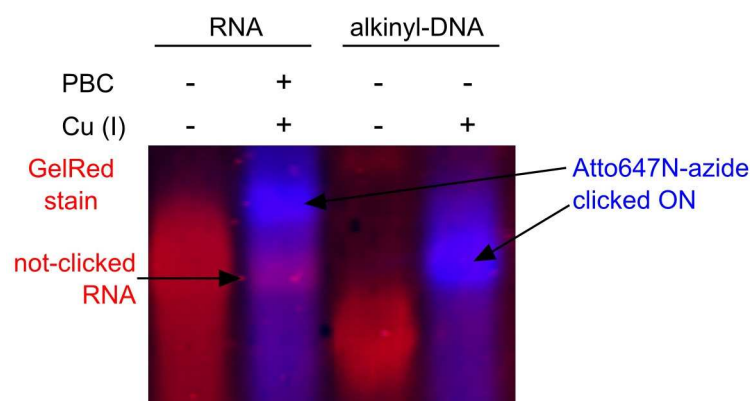


Figure 20 Urea-PAGE of Atto647N-azide treated oligonucleotides (ON).

The blue fluorescence indicates the successfully clicked ON in the presence of Cu(I). Without Cu(I) there is no fluorescence detectable even though the GelRed stain reveals the presence of the ON.

#### 4.4.4 Ring opening of uridine-PBC conjugates upon UV irradiation

First gel analyses of PBC treated tRNA *E.coli* revealed a typical coumarin fluorescence of conjugated samples upon UV irradiation at 365 nm. However, upon continued illumination up to 5-8 minutes, a clear decrease in fluorescence could be observed as is shown in Figure 21A. LC-MS/MS analysis of samples with and without UV illumination showed similar abundance of s<sup>4</sup>U conjugate. G-PBC and Ψ-PBC were not detectable in the UV illuminated samples and the U-PBC abundance was clearly reduced (Figure 21B). These findings might be explained by a UV triggered decomposition of G-PBC, Ψ-PBC and U-PBC. One possible decomposition mechanism might be the loss of the alkynyl-group due to hydrolysis of the ether. In this case it would not be possible to detect the treated samples by using fluorescent azide dyes attached by CuAAC reaction. However, since the s<sup>4</sup>U conjugate seems to stay intact during the UV irradiation an *in-vitro*-transcribed tRNA (IVT) without s<sup>4</sup>U was chosen as an additional control. Figure 21C shows the gel of IVT tRNA and *E.coli* tRNA +/- UV illumination prior to click reaction. It can be observed that the IVT tRNA shows a lower fluorescence than the native tRNA due to the absence of s<sup>4</sup>U, but all tRNAs show a similar click reaction yield independent of UV treatment. This is only possible if the alkyne group is still attached to the tRNA even though the coumarin fluorescence vanishes upon radiation. Another possible mechanism is a hydrolysis of the coumarin lactone ring. This would explain the loss of coumarin fluorescence as a consequence of a reduced π-system of the chromophore. To prove this hypothesis uridine homopolymer RNA was treated with PBC and digested. U-PBC was collected from RP-LC and lyophilized before reconstitution in neutral, acidic or alkaline aqueous buffer. The resulting solutions were illuminated for 10 minutes at 365 nm and injected on the LC-MS in MS scanning mode while monitoring the absorption at 254 nm. A peak, not present in non-UV treated samples, at R<sub>t</sub> = 9.5 minutes with a mass signal of 475 Dalton which is 18 Da (=1 mol H<sub>2</sub>O) more than U-PBC (m/z = 457 and R<sub>t</sub>=12.3 min) was observed. In a second run the corresponding MS/MS spectrum of the found precursor signal m/z=475 was recorded and revealed a fragmentation pattern similar to U-PBC (see Figure 21D). From this data it is suggested that U-PBC is undergoing a UV triggered ring opening reaction by hydrolysis, giving rise to a non-fluorescent coumarin-conjugate with [M+H]<sup>+</sup> 475. The ring-open U-PBC undergoes common mass fragmentation e.g. loss of ribose (475→343), loss of water (475 → 457) or both (475→325) as can be seen from the MS/MS spectrum and the possible reaction scheme Figure 21E. Interestingly, this type of reaction is commonly used in the treatment of psoriasis by the so called PUVA therapy (psoralen plus UV-A). This therapeutic concept from the 1970s makes use of the furanocoumarin psoralen, which is applied to the patients skin followed by UV-A (315-400 nm) irradiation. This to an activation of the drug by ring-opening [108] and the activated



psoralen is crosslinking DNA strands of keratinocytes, inducing death of activated T-cells in the skin.

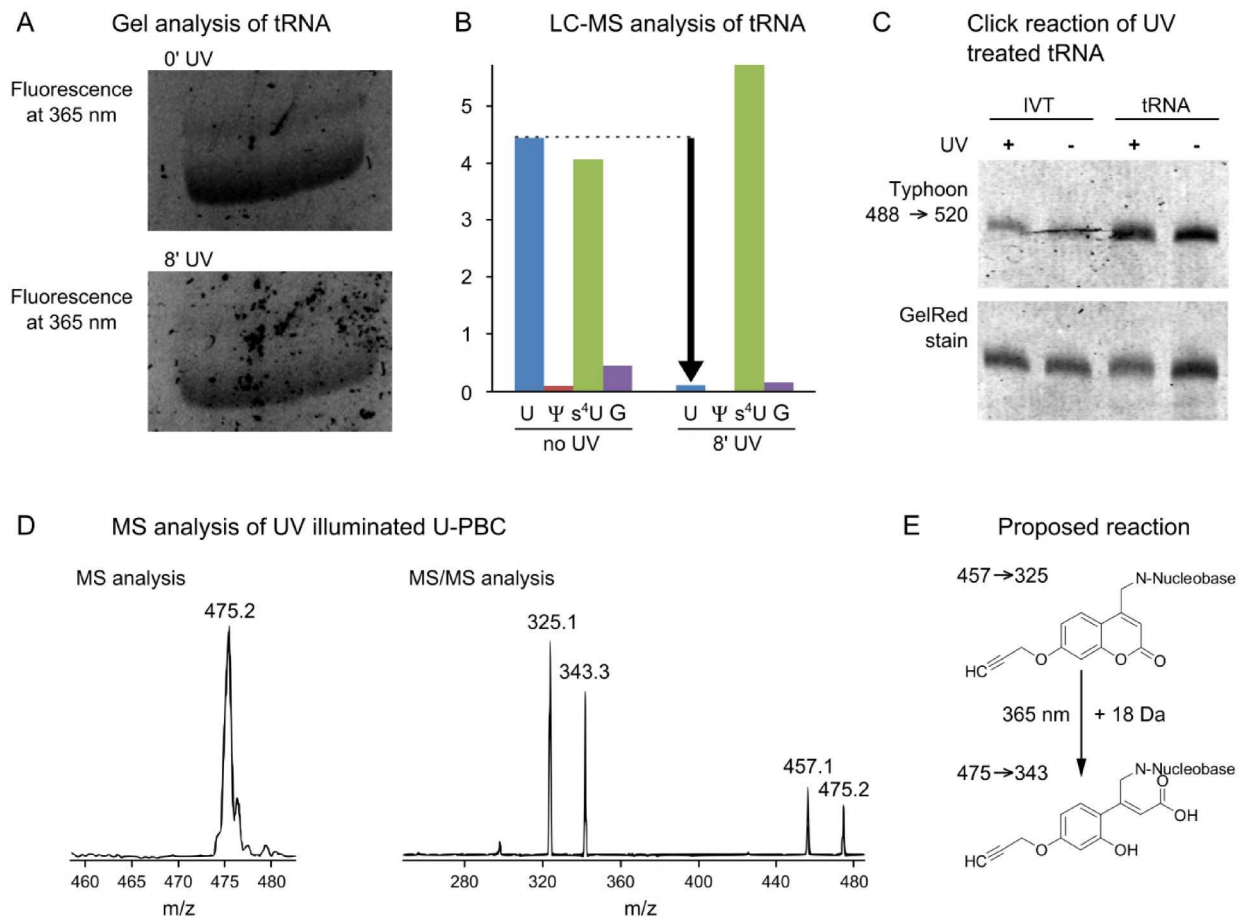


Figure 21 Analysis of UV illuminated PBC conjugated samples.

A tRNA *E.coli* treated with PBC and analyzed on a 10% urea-PAGE monitored at UV light of 365 nm. The upper picture was taken immediately, the lower picture after 8' constant UV irradiation. B LC-MS analysis of treated tRNA *E.coli* digests with and without prior UV treatment. C Gel analysis of IVT (*in-vitro*-transcript) and native tRNA of UV illuminated and mock illuminated samples that were clicked with fluorescein-azide. D LC-MS and LC-MS/MS spectra of isolated U-PBC that was UV illuminated before injection. E Possible ring opening mechanism of U-PBC upon UV illumination, which matches the observed MS/MS transitions shown in D.

#### 4.4.5 Behavior of monofunctional coumarins upon UV irradiation

After observing a UV dependent ring opening of U-PBC but not of  $s^4\text{U}$ -PBC, further analysis in this direction was conducted with samples treated with the monofunctional coumarins. Despite the high sensitivity of LC-MS/MS analysis for detection of minor substance traces, it was not possible to find U-conjugates + 18 Da in coumarin treated, illuminated samples. This might be due to the acidic character of the ring open coumarin that prevents efficient protonation at the eluent pH 5.3 and therefore might show high limits of detection in positive ion mode. Analysis in negative ionization mode was not possible, since it would have called for a complete change in the chromatographic system *e.g.* alkaline buffer pH and column material, which would result in different retention times and different  $m/z$  for all nucleosides and conjugates. However, studies were performed, analyzing the decrease of U-conjugates in the samples after UV irradiation and comparison to non-UV treated samples. The data processing of the LC-MS/MS analysis focuses on the overall yield and reactivity of the coumarins. The results shown in Figure 22 demonstrate that there are only minor changes (maximum factor 5 of total yield) for the  $s^4\text{U}$ -conjugates, which might either stay similar upon UV treatment (compound **2**, **6** and PBC), show increased yields (BMB, **4** and **5**) or decreased  $s^4\text{U}$  conjugate (compound **3**). No pattern in the change of  $s^4\text{U}$ -conjugate yield can be discerned from this dataset and the variability does not allow a conclusion on the influence of UV illumination on the  $s^4\text{U}$ -conjugates. The observed changes are most likely caused by data analysis fluctuations and further experiments are needed to prove this hypothesis.

The diagram on the right of Figure 22 shows the monitored U-conjugates and the dramatic decrease of detectable conjugate upon UV irradiation. U-BMB seems to be most affected with a loss of factor 800 compared to the not irradiated sample. The other conjugate losses range from factor 180 to factor 15. This analysis leads to the assumption that all tested coumarins are prone to ring opening upon UV irradiation if they are covalently attached to uridine (*N3*). On the other hand coumarins conjugated to the sulfur of  $s^4\text{U}$  seem less prone to ring opening reactions most probably due to the sulfurs lower electronegativity compared to the nitrogens. The lower electronegativity leads to a higher electron density in the lacton ring leaving it less attractive for a ring opening hydrolysis. The hypothesis is supported by observations concerning other N-alkylated coumarin conjugates like guanosine or pseudouridine that are found to be less abundant or even absent after UV irradiation of the samples.

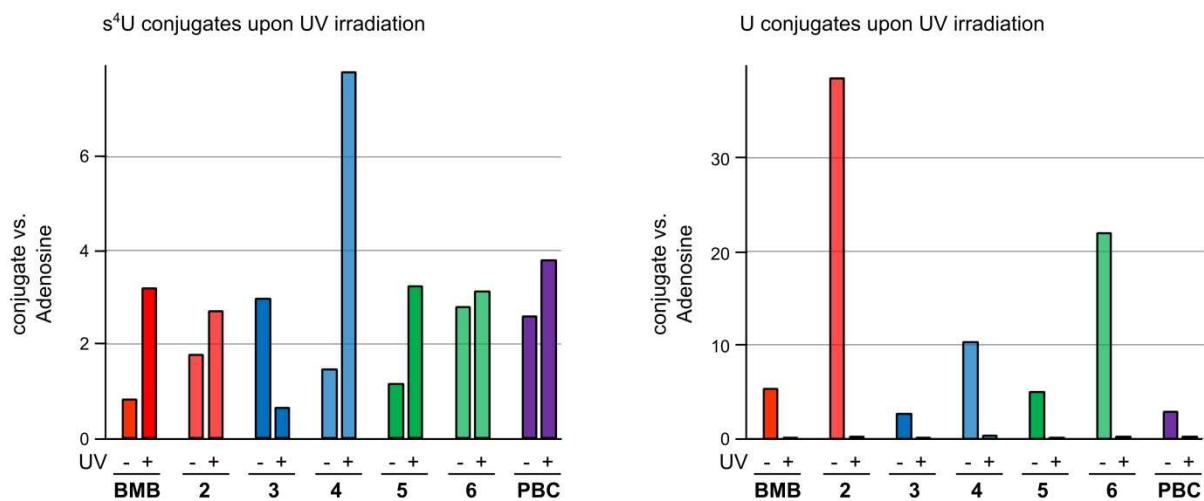


Figure 22 LC-MS analysis of coumarin treated tRNA *E.coli* before and after UV irradiation

## 4.5 Application of coumarins

Coumarins are small molecules used for chemical functionalization of RNA that can be used on all types of RNA and DNA, as presented in the previous chapters. The azide and alkyne functionalized coumarins N3BC and PBC are furthermore used for the CuAAC reaction with *e.g.* fluorescent dyes. This allows an easy way to fluorescently label nucleic acids for all applications where fluorescence is needed but too expensive or troublesome to achieve. As an example, the following chapter presents the successful labeling of siRNA with PBC and subsequent coupling with atto647N-azide. The functionalized siRNA is hybridized with its complementary strand, which is commercially labeled with alexa555 and FRET (fluorescence resonance energy transfer) is observed between both dyes. Additionally, the same construct is successfully delivered into living cells and the fluorescence is monitored by LSCM (laser scanning confocal microscope). This proves that coumarin conjugated and clicked siRNA is taken up by cells and can be traced by its fluorescence.

### 4.5.1 Preparation of PBC-atto647N labeled siRNA

A commercial antisense siRNA to EGFP (enhanced green fluorescent protein) was chosen for the derivatization with the coumarin. EGFP siRNA is commonly chosen, since the efficiency of the siRNA gene knockdown can be measured easily by a reduction of EGFP fluorescence in treated cells. PBC was used in these first tests, since a clickable atto647N-azide dye was available, which allowed the usage of a commercially atto647N labeled siRNA as a control.

The antisense siRNA was conjugated with PBC under reaction conditions 2, which allows a mild uridine alkylation. After LiClO<sub>4</sub> precipitation the siRNA was reacted under CuAAC promoting conditions with atto647N-azide and precipitated. Gel-purification revealed only one fluorescent band, which is interpreted as a singly-labeled siRNA, since multiple labels would lead to several ladder like bands above the respective band due to the increased siRNA size.

#### 4.5.2 Hybridization and FRET experiments of the PBC-Atto647N labeled siRNA

It is emphasized that all ongoing measurements and raw data analyses were performed by my colleague Markus Hirsch.

The commonly used nomenclature of siRNA constructs is as follows:

*e.g.* commercial construct: ds alexa555/atto647N

s: 5'-GCAAGCUGACCCUGAAGUUCAdC-alexa555

as: 3'-GCCGUUCGACUGGGACUUCAAG-atto647N

The coumarin labeled constructs: ds alexa555/PBC-atto647N

The gel-purified PBC-atto647N labeled siRNA was hybridized with its unlabeled or 3'-alexa555 labeled sense strand. As a control, the commercially 5'-atto647N labeled siRNA was hybridized in the same way. For hybridization both strands were diluted in PBS (phosphate buffered saline) and initially denatured for 3 minutes at 90°C followed by one hour annealing at 37°C. The hybridization success was monitored on a non-denaturing polyacrylamide gel (see appendix Figure A 1). Fluorescence measurements of the four resulting constructs upon excitation at 633 nm was done in a cuvette and the results can be seen in Figure 23A. A fluorescent signal with its emission maximum at 670 nm can be observed for all constructs. However, the intensity is higher for the control siRNA compared to the PBC-clicked siRNA probably due to differing sample concentrations.

For FRET studies the four constructs and an alexa555 sense strand hybridized with unlabeled antisense siRNA were used. Here, the alexa555 represents the donor dye which is excited at 543 nm (561 nm corresponding laser line at LSCM) and transfers its energy to the acceptor dye, atto647N, which then emits at 670 nm [109]. The principle of this radiation-free energy transfer can be seen in Figure 23B. The left spectra in Figure 23C show the PBC-atto647N labeled constructs and corresponding controls. For the non-FRET constructs, either lacking the donor or the acceptor dye, no fluorescence signal can be seen at 670 nm. The construct containing both donor and acceptor dye emits at 670 nm with a significant, but low intensity. In the right part of Figure 23C the same constructs with the commercially atto647N labeled siRNA are shown and a more intense fluorescence can be observed for the FRET construct compared to the PBC-atto647N construct. The difference in fluorescence intensity between the two constructs is most likely caused by a higher distance of the 3' positioned donor dye and the statistically distributed acceptor dye in the PBC functionalized siRNA (Figure 23D). According to the graph in Figure 23E the FRET efficiency decreases if the spatial distance is increased above the 1.2 x Förster radius. In case of the commercial donor/acceptor siRNA doublestrand, the dyes are located as close to each other as possible, namely at the 3' of the sense strand and the 5' of the antisense

strand. This minimal spatial distance makes the construct a so called high-FRET construct, whereas the PBC-functionalized FRET pair represents a mixture of different distances, which results in a low-FRET construct [110].

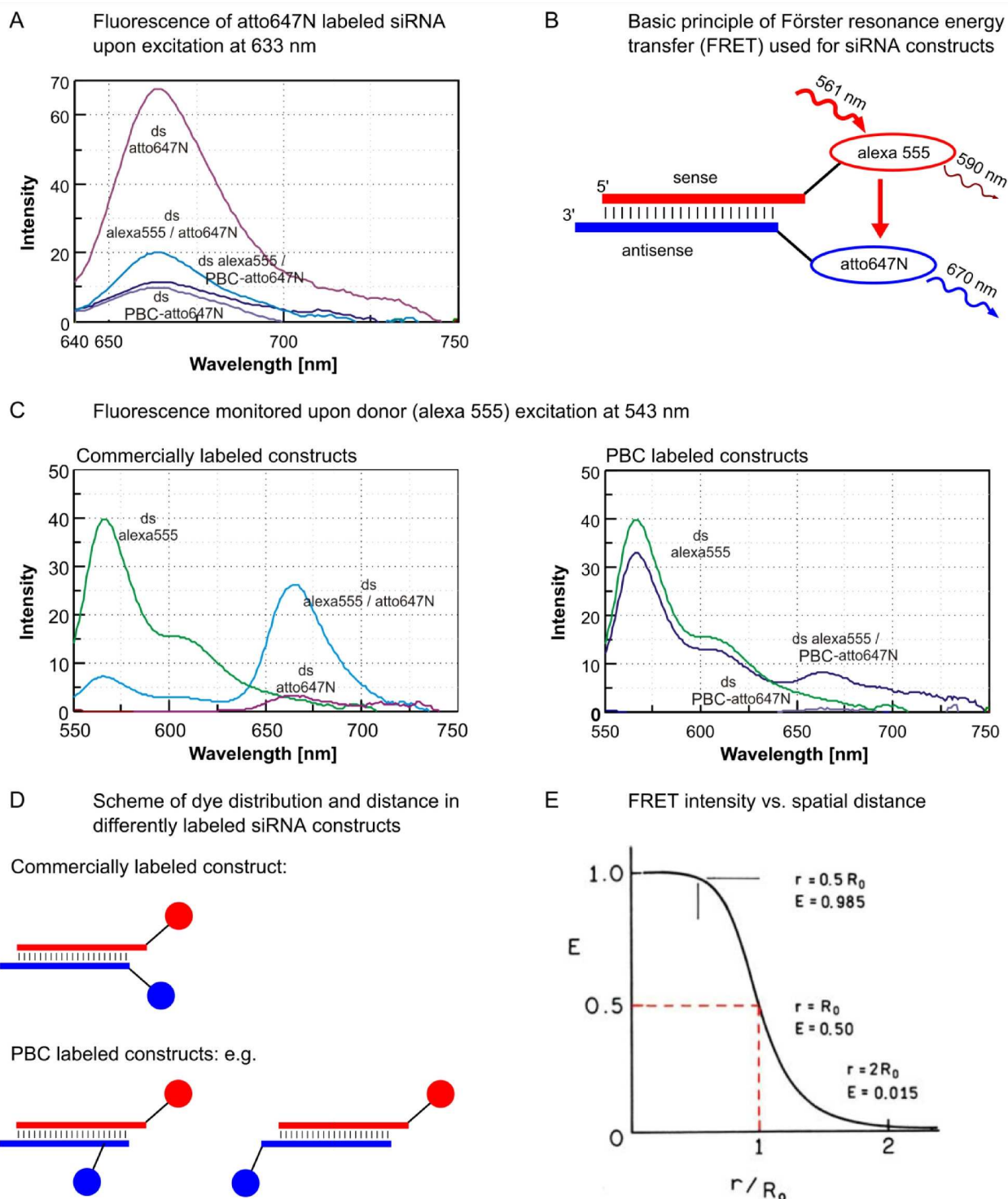


Figure 23 Fluorescence spectra of commercially and PBC clicked atto647N siRNA hybridized with the complementary alexa555 sense strand done by Markus Hirsch.

A Fluorescence of ds siRNA constructs upon excitation of the atto647N acceptor dye. B Principle of Förster resonance energy transfer (FRET). C Fluorescence of ds siRNA constructs upon excitation of the alexa555 donor dye. D Scheme of acceptor dye location in case of PBC and commercially labeled antisense siRNA. E Scheme of FRET efficiency correlation to the Förster radius [111].

### 4.5.3 Cell imaging studies of PBC-atto647N labeled siRNA

#### *Live-cell imaging*

The alexa555/PBC-atto647N doublestrand siRNA was administered to living RBE4 cells and siRNA was transfected using commercially available cationic lipids (i.e. oligofectamine, Invitrogen, Darmstadt) as described [109]. In Figure 24A, the PBC-atto647N fluorescence is monitored by excitation at 633 nm and displayed in red. It is noteworthy that the cells are still alive and dividing even 1.5 hours after administration of the construct (see second row, first column). In ongoing experiments cells were incubated for up to 10 hours without showing any phenotype that suggests cell toxicity. Apparently, the coumarin label has no harmful effect on the cells. The last column of the second row shows a sudden burst of “red” fluorescence in the freshly divided cell which is interpreted as the successful uptake of the coumarin labeled siRNA. The fluorescence spreads further inside the cell and even enters the daughter cell which is most likely still connected to the mother cell after cell division (3<sup>rd</sup> row, 2<sup>nd</sup> column). In parallel, a third, not-freshly divided cell, develops a strong fluorescent signal which makes a total of 2 observed uptake events in the chosen roi (region of interest). From this, it can be concluded, that PBC-atto647N labeled siRNA has no observable cell harming effects and is readily taken up by the cells. It is therefore suitable as an easy labeling technique for siRNA fluorescence studies.

#### *FRET studies in fixed cells*

After the success of observable PBC-atto647N labeled siRNA uptake in living cells, a second transfection approach followed by fixation was conducted. Here, the different constructs, already examined spectrometrically (see chapter 4.5.2) were transfected and the FRET signal analyzed. The first row of Figure 24B shows the signal of the alexa555 donor dye after excitation at 561 nm. The second row shows the fluorescence of the acceptor dye after donor dye excitation, which is FRET. The last row displays the fluorescence emerging from direct acceptor dye excitation at 633 nm.

For untreated cells (first column) no fluorescence can be seen. The next two columns show the cells transfected with control constructs either lacking the acceptor or the donor dye, which both display a faint FRET signal, most probably caused by cross-talk. Column number 4 is the commercially labeled construct control which displays an intense FRET signal, which is in agreement with the previously shown spectra in Figure 23C. The column on the right (ds alexa555/PBC-atto647N) however, shows strong donor and comparatively low acceptor fluorescence. The FRET intensity is as expected lower than the intensity of the commercial construct. In these settings the possible FRET signal of the PBC-atto647N construct is

comparable with the signal intensity of the controls. From these results it is not possible to judge if the observed signal is a real low-FRET signal or simply cross-talk caused by the donor dye.

### ***Summary of coumarin suitability as a cell-compatible labeling agent***

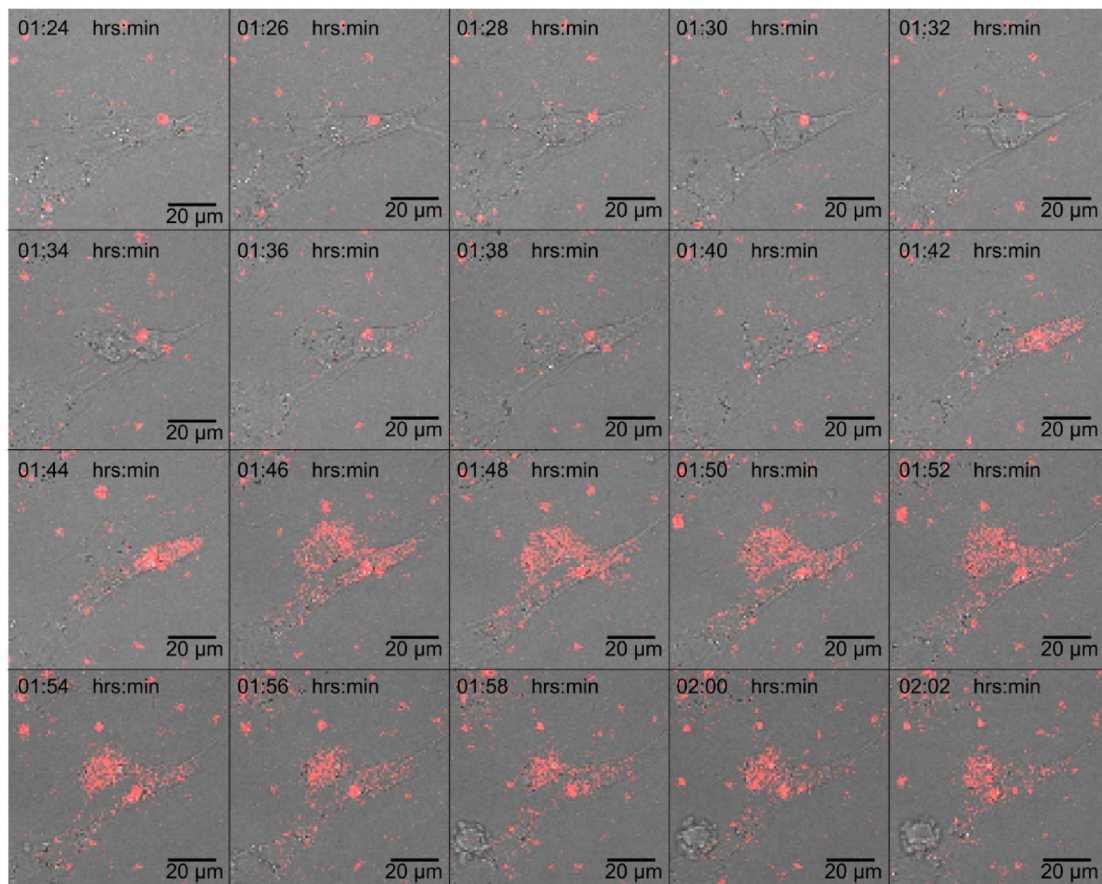
The coumarin PBC was successfully used to alkylate antisense siRNA and further subjected to CuAAC reaction with atto647N-azide. Purified PBC-atto647N antisense and alexa555 sense strand are successfully hybridized, demonstrating that one uridine *N3* alkylation is tolerated by the overall base pairing of the siRNA double strand.

PBC-atto647N labeled and commercial atto647N constructs were fluorometrically analyzed and compared. While the commercially labeled construct shows high-FRET efficiency, the PBC-atto647N labeled displays low-FRET efficiency. This finding is in accordance with the spatial distance of the attached dye molecules, which is larger for the self-labeled construct. With increased spatial distance, the FRET intensity is decreased.

Additionally, it was shown that PBC-atto647N labeled siRNA shows no harmful effects on the observed cells and furthermore that coumarin containing constructs successfully enter cells. These results form a basis for further experiments with self-labeled fluorescent RNA species to monitor their behavior and distribution in cells.



**A** Uptake of PBC-atto647N labeled siRNA monitored by live-cell imaging  
Excitation wavelength = 633 nm



**B** FRET experiments of siRNA in fixed cells

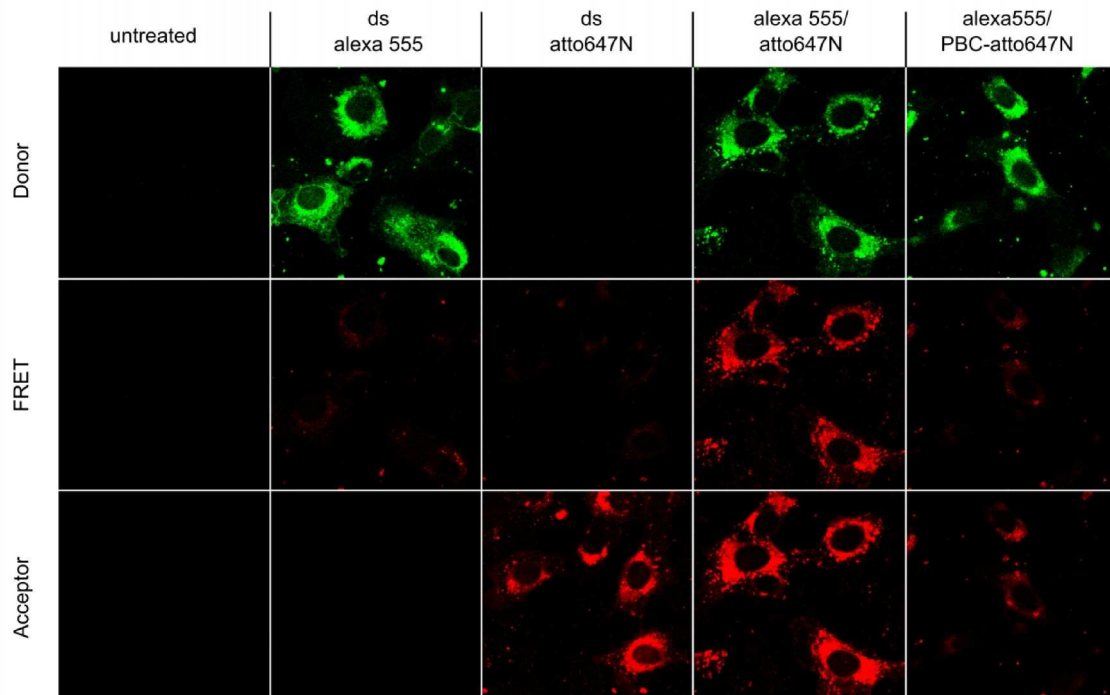


Figure 24 Cell imaging studies of PBC-atto647N labeled siRNA uptake performed by Markus Hirsch. A Time series of live cell imaging of PBC-atto647N transfected siRNA at acceptor excitation. B Fixed cells transfected with donor only, acceptor only, commercial FRET and PBC-FRET construct.

## 4.6 LC-MS/MS detection of modified nucleosides

LC-MS/MS analysis for detection of RNA or DNA nucleosides is confronted with 2 major problems, namely sufficient separation (= chromatographic resolution) and detection of modified nucleosides that are mostly 10.000 times less abundant than their corresponding canonical base. The first problem is characterized by the substantial number and high similarity of modified ribonucleosides. All nucleosides consist of a (deoxy)ribose sugar and either a pyrimidine or a purine nucleobase, which can be modified at one or several positions. In most cases these modifications are methylations or more complex modifications, like additional or enlarged ring systems at the nucleobase or even amino acids. Chromatographic separation is based on the physicochemical properties of the nucleosides, which are altered to some extent *e.g.* by methylation and methylation position. Finding a chromatographic system suitable to separate over 100 modified ribonucleosides is highly cumbersome due to their high physicochemical similarity. The second aforementioned problem arises from the sample composition that normally contains huge amounts of the unmodified nucleosides and only minor amounts of modified nucleosides. This imbalance interferes both with chromatographic separation and detection of the low abundant modified nucleosides. The most powerful approach to overcome this imbalance is the combination of a good resolving chromatography with 2 detection methods *e.g.* UV and mass detection. The dual detection system of UV and MS is the answer to the problem of high abundant main nucleosides, which can be easily detected by absorption. The low abundant, possibly co-eluting modified nucleosides are detectable by highly sensitive mass spectrometry.

Based on these considerations, several chromatographic methods based on an amide-functionalized RP-18 column (Synergy Fusion RP column 4  $\mu\text{m}$  particle size, 80  $\text{\AA}$  pore size, 250 mm length, 2 mm inner diameter from Phenomenex (Aschaffenburg, Germany)) were developed. An advantage of the chosen column is its capability to separate hydrophilic, mildly basic substances, like ribonucleosides and deoxyribonucleosides, while being compatible with MS detection. In addition, optimization of MS parameters for highest sensitive MS/MS detection of nucleosides was performed and implemented in the chromatographic method. A brief overview over the developed methods for detection and quantification of RNA nucleosides followed by DNA nucleosides will be presented in the next two chapters. Details are given in the published papers.

#### 4.6.1 Method development for modified ribonucleosides

##### Chromatography – gradient “nucs4”

As there are more than 100 modified nucleosides, some of them being constitutional isomers, particular attention must be paid to the applied chromatography. The column used was an RP-18 column fused with polar amide groups that allows separation of basic substances at high aqueous conditions. Chromatography was started with 100 % buffer B (5 mM NH<sub>4</sub>OAc pH 5.3) with slowly increasing acetonitrile gradient (< 1% per minute) over time. This allowed sufficient retention of even the most hydrophilic nucleosides dihydrouridine and pseudouridine eluting more than 1 minute after the dead time (Figure 25). Furthermore, the slow gradient allowed sufficient separation of all isomeric nucleosides and of mostly all other nucleosides investigated.

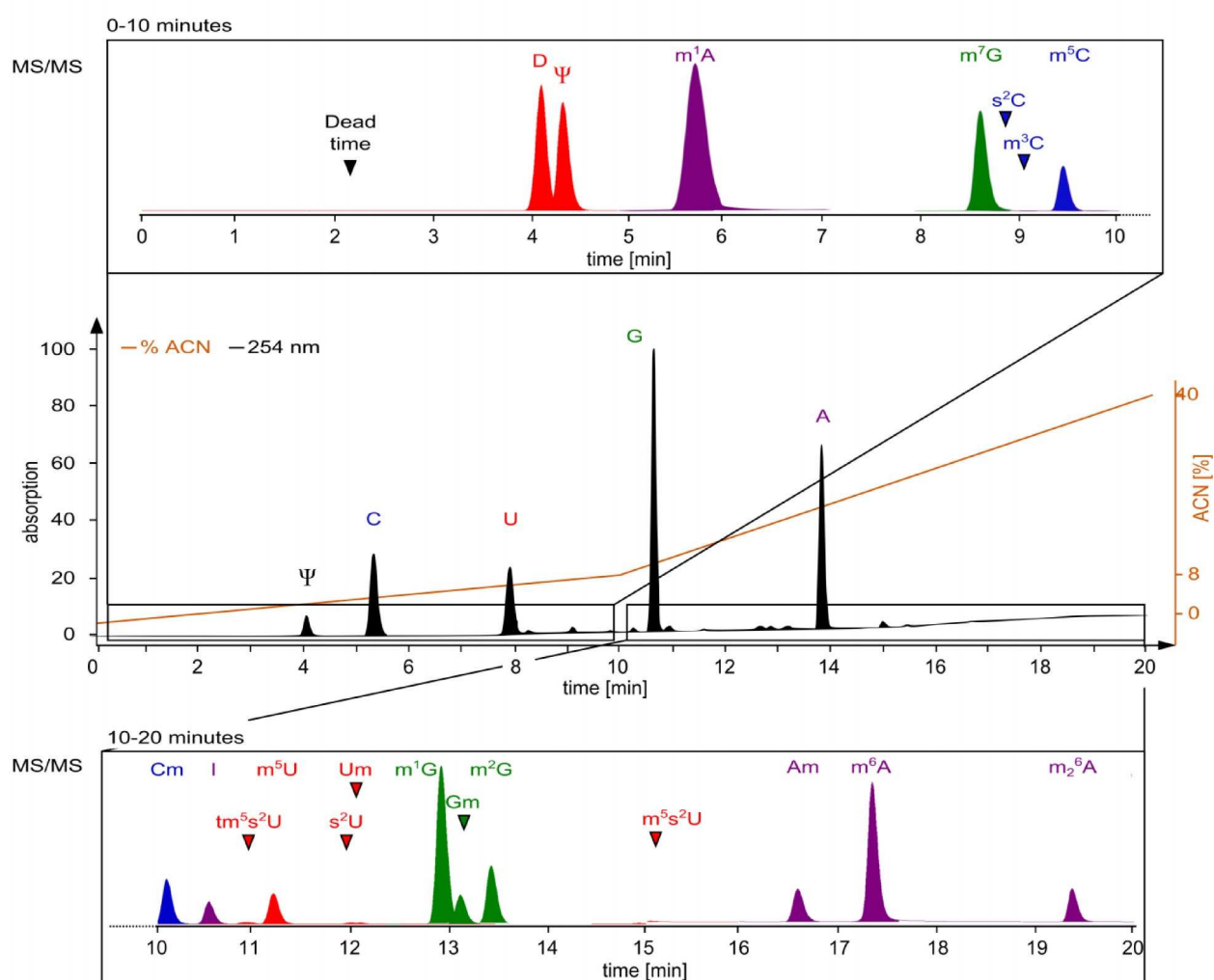


Figure 25 Chromatogram of 25 ribonucleosides with the gradient nucs 4 (indicated with orange line, middle panel). The middle chromatogram shows the main nucleosides and pseudouridine (black) detected by absorption at 254 nm. Above and below are the zoomed MS/MS chromatograms in 10 minute segments. (colour code: blue – C residues; red – U residues; green – G residues; lilac – A residues)

### *Mass spectrometry of ribonucleosides*

RNA nucleosides have a basic character and accept protons easily upon electrospray ionization in positive ion mode. The only exceptions are uridine residues that have lower ionization efficiencies due to their more acidic character. The ionization efficiency is nevertheless yielding enough for detection in positive ion mode by the slightly acidic ammonium acetate HPLC eluent. MS/MS collision induced fragmentation occurs readily at the glycosidic bond of all nucleosides, even at low energy input around 5-15 V (see

Figure 26A) in the collision cell. Pseudouridine, the isomer of uridine, is the only exception of the aforementioned fragmentation pattern. It has a carbon-carbon link between the ribose and the nucleobase that is more stable than the nitrogen-carbon bond. At CID voltages around 10 V the nucleoside undergoes several fragmentation reactions, leading to a large distribution of charged fragments [112] (

Figure 26C). An overview over all MS/MS transitions and retention times is given in Table 18.

The energy needed for cleavage of the nucleosidic bond requires special attention since it is low enough to be reached by the fragmentor voltage, located between the source capillary and quadrupole 1. This has to be taken into account when a MS/MS method (also called MRM method) is created. In case of monitoring the fragmentation of the glycosidic bond, it is necessary to use fragmentor voltages below 100 V to reduce pre-quadrupole fragmentation and to achieve highest sensitivity for nucleoside quantification.

However, as nucleosides are known to undergo reliable fragmentation at the glycosidic bond, a high fragmentor voltage can be advantageous for clear identification of a nucleoside of interest. In this case, the fragmentor voltage is increased to around 200 V, which is in most cases enough to produce the charged nucleobase fragment before entering quadrupole 1. Q1 is then set to let pass the  $m/z$  of the nucleobase, transferring the nucleobase into the collision cell for fragmentation. The fragmentation reaction of the nucleobase can be used for clear identification of *e.g.* isomers, since each modified nucleobase has distinctive fragments caused by its modification type and position. Such an MS<sup>3</sup> like fragmentation for identification of modified nucleosides is well described in the literature [113].

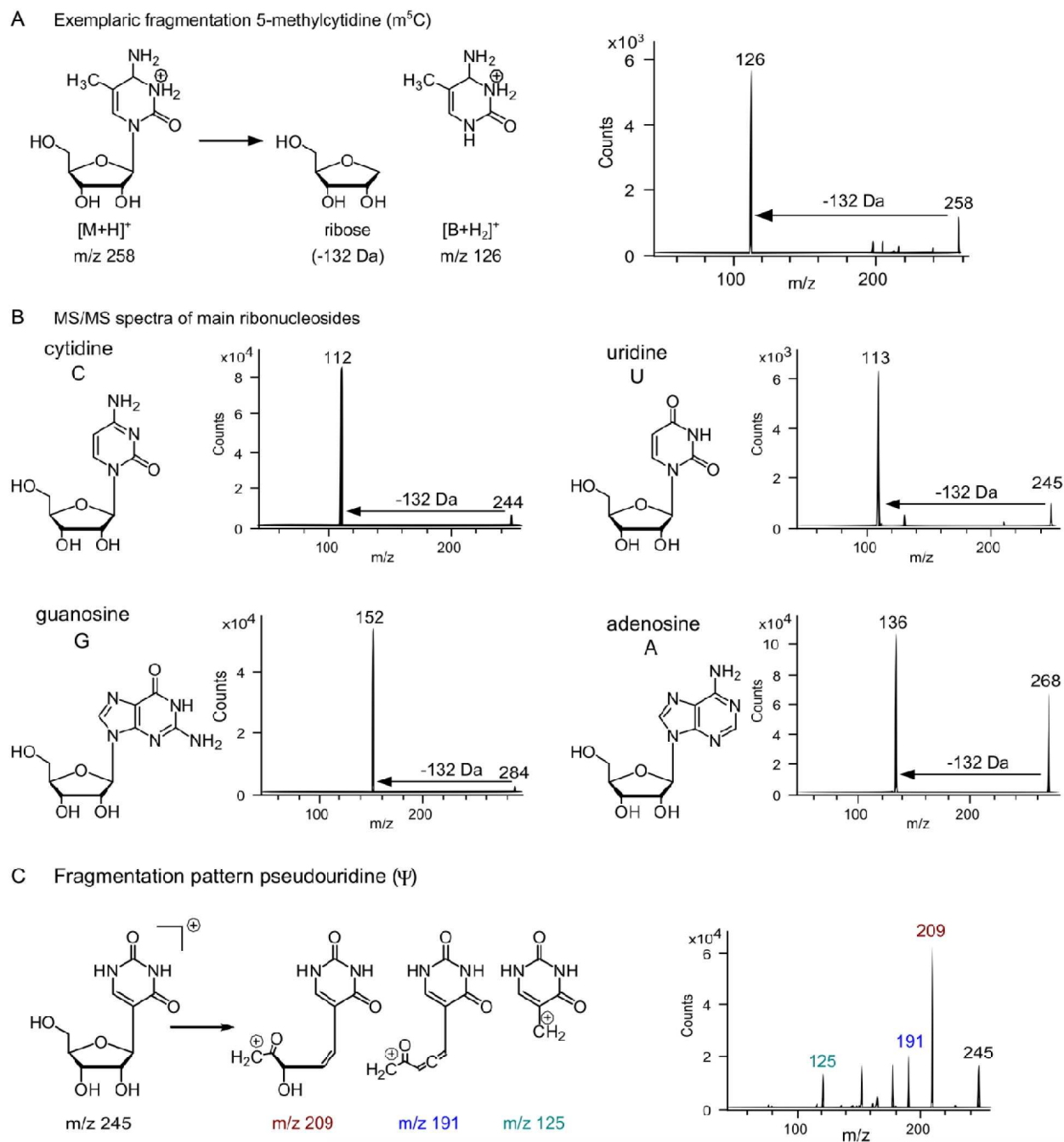


Figure 26 Fragmentation pattern of ribonucleosides.

A Scheme of 5-methylcytidine fragmentation upon CID and the resulting MS/MS spectrum. B MS/MS spectra of all four major nucleosides. C Fragmentation reactions for pseudouridine fragmentation upon CID and the resulting MS/MS spectrum.

Table 18 MS/MS transitions of all 21 analyzed modified nucleosides and retention times (colour code: blue – C residues; red – U residues; green – G residues; lilac – A residues)

Name	Short	Precursor	Product	Fragmentor	CID	R <sub>t</sub> [min]
Dihydrouridine	D	247	115	80	10	4.1
Pseudouridine	Ψ	245	209	81	5	4.3
Cytidine	C	244	112	80	10	5.4
1-methyladenosine	m <sup>1</sup> A	282	150	92	17	5.7
Uridine	U	245	113	80	10	8.0
7-methylguanosine	m <sup>7</sup> G	298	166	82	9	8.6
2-thiocytidine	s <sup>2</sup> C	260	128	40	10	8.9
3-methylcytidine	m <sup>3</sup> C	258	126	40	10	9.0
5-methylcytidine	m <sup>5</sup> C	258	126	40	9	9.4
2'-O-methylcytidine	Cm	258	112	60	9	10.1
Inosine	I	269	137	76	5	10.5
Taurino-5-methyl-2-thiouridine	tm <sup>5</sup> s <sup>2</sup> U	398	266	80	10	10.9
Guanosine	G	284	152	80	10	10.8
5-methyluridine	m <sup>5</sup> U	259	127	76	5	11.2
2-thiouridine	s <sup>2</sup> U	261	129	66	5	12.0
2'-O-methyluridine	Um	259	113	66	5	12.1
1-methylguanosine	m <sup>1</sup> G	298	166	82	9	12.9
4-thiouridine	s <sup>4</sup> U	261	129	66	5	13.1
2'-O-methylguanosine	Gm	298	152	72	5	13.1
2-methylguanosine	m <sup>2</sup> G	298	166	82	9	13.5
Adenosine	A	268	136	80	10	14.9
5-methyl-2-thiouridine	m <sup>5</sup> s <sup>2</sup> U	275	143	66	5	15.0
2'-O-methyladenosine	Am	282	136	92	13	16.3
6-methyladenosine	m <sup>6</sup> A	282	150	92	17	17.0
2,6-dimethyladenosine	m <sup>2</sup> <sub>6</sub> A	296	164	102	17	18.7

### ***LC-MS/MS method “nucs4\_DMRM”***

21 modified nucleosides were available as commercial substances and their retention times and mass transitions examined. A triple quadrupole MS can only focus on one mass transition at a time. If several substances, with different transitions, are to be monitored, the device has to switch from transition to transition. This leads to a timeframe, called dwell time, in which each substance is monitored. Analysis of many substances will lead to dwell times below 50 ms for each transition, which results in lower sensitivity and ultimately even a loss of sharp eluting substance peaks. One possible solution is the dynamic MRM (DMRM) mode provided by Agilent's MassHunter software. The DMRM method is programmed, not only with the mass transitions and mass parameters (like an MRM method), but additionally with the retention times and a time window for each transition. The MS is observing a given mass transition only in the given time frame and therefore the amount of concurring mass transitions is reduced from 21 to a maximum of 8 for the described chromatography “nucs4”. This method allows clear discrimination of isomeric nucleosides, exemplarily shown for monomethylated cytidines m<sup>3</sup>C,



$m^5C$  and Cm. As can be seen in Figure 27, sugar methylated Cm is distinguished from the base methylated isomers  $m^3C$  and  $m^5C$  by mass transition due to the loss of methylated ribose (146 Da) instead of ribose (132 Da) alone.  $m^3C$  and  $m^5C$  have the same fragmentation pattern but differ in their physicochemical properties allowing separation by the used chromatography. Clear identification of the isomers is done by comparison of retention times of commercial standard substances.

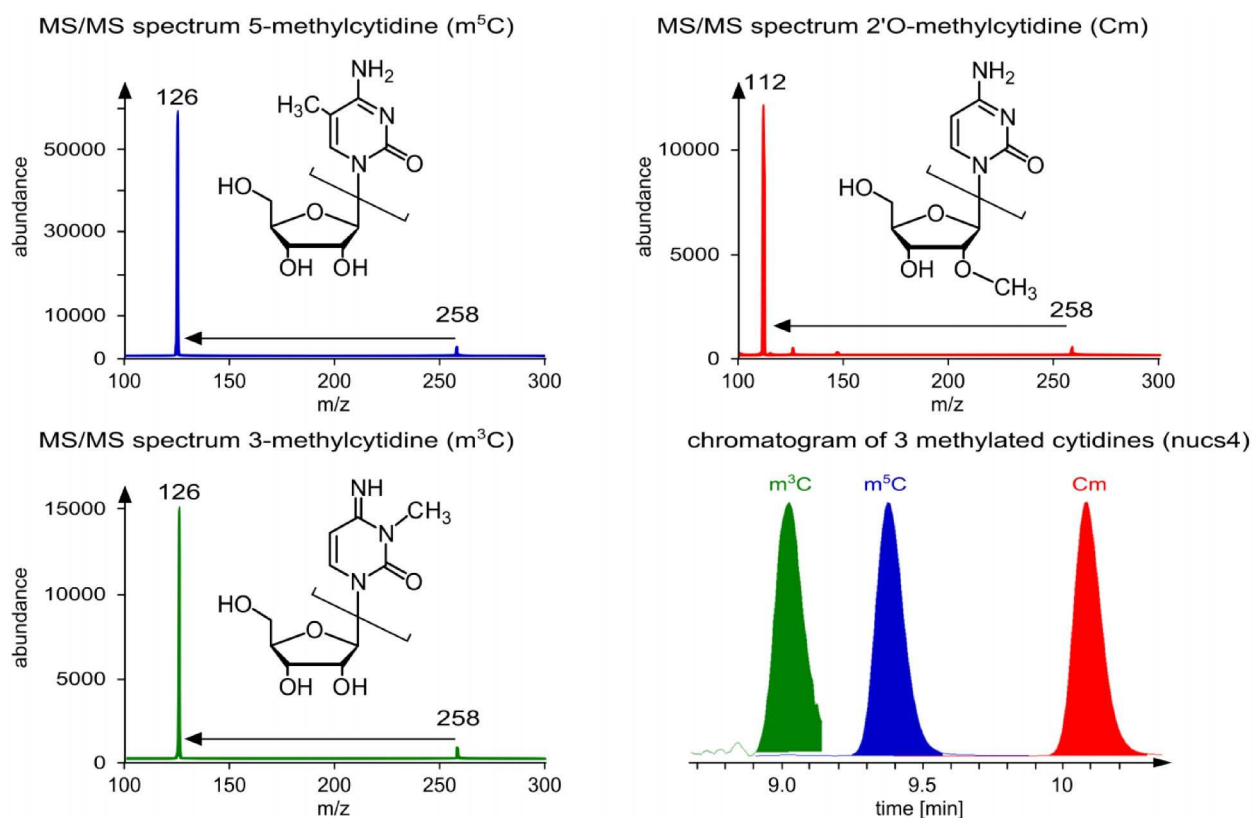


Figure 27 MS/MS spectra of monomethylated cytidines  $m^3C$ ,  $m^5C$  and Cm and their corresponding mass peaks in the chromatogram from the method nucs4\_DMRM.

All cytidines have the same precursor mass 258 Da but only Cm has a product mass of 112 Da (red), while  $m^3C$  and  $m^5C$  both produce 126 Da ions (green and blue). By MS/MS transition these two methylated cytidines cannot be distinguished but chromatographic separation allows sufficient resolution for identification of each isomer.

#### 4.6.2 Method development for modified deoxyribonucleosides

##### *Chromatography – gradient “dnucs2”*

Separation of (modified) deoxyribonucleosides is less challenging, since only half a dozen naturally modified nucleosides are of interest for our research and these modifications differ sufficiently in their physicochemical properties. It was therefore possible to use higher flow-rates and a faster gradient (2% ACN per minute) with the same column and buffer as described for ribonucleosides. The chromatography is separating all deoxynucleosides and overall analysis time is reduced from 30 to 17 minutes (Figure 28A).

##### *Mass spectrometry of deoxyribonucleosides*

All deoxyribonucleosides are easily ionized in positive ion mode and undergo a fragmentation reaction at the glycosidic bond. The resulting product ions are 116 Da lighter than the precursor ion since deoxyribose weighs 116 Da, 16 Da less than ribose. Exemplary MS/MS spectra are shown for m<sup>5</sup>dC and fdC (Figure 28B). An overview of all deoxynucleosides analyzed is given in Table 19.

##### *LC-MS method “dnucs2\_DMRM”*

For DNA nucleosides a DMRM detection method was used, as already described for RNA nucleosides. It has proven useful for studies of cytosine methylating enzymes to additionally monitor the MS/MS transition of 5-methylribocytidine for clear discrimination of the methylated product of deoxycytidine and cytidine.

Table 19 All nucleosides analyzed in method dnucs2 with MS/MS transitions and retention times

<b>Name</b>	<b>Short</b>	<b>Precursor</b>	<b>Product</b>	<b>CID</b>	<b>R<sub>t</sub> [min]</b>
5-carboxydeoxycytidine	cadC	272	156	5	4.6
Deoxycytidine	dC	228	112	10	5.6
5-hydroxydeoxycytidine	hmdC	258	142	5	5.8
5-methyldeoxycytidine	m <sup>5</sup> dC	242	126.1	5	6.8
Deoxyguanosine	dG	268	152	10	7.2
5-formyldeoxycytidine	fdC	256	140	5	7.6
Thymidine	dT	243	127	10	7.8
Deoxyadenosine	dA	252	136	10	8.5
N6-methyldeoxyadenosine	m <sup>6</sup> dA	266	150	10	10.1



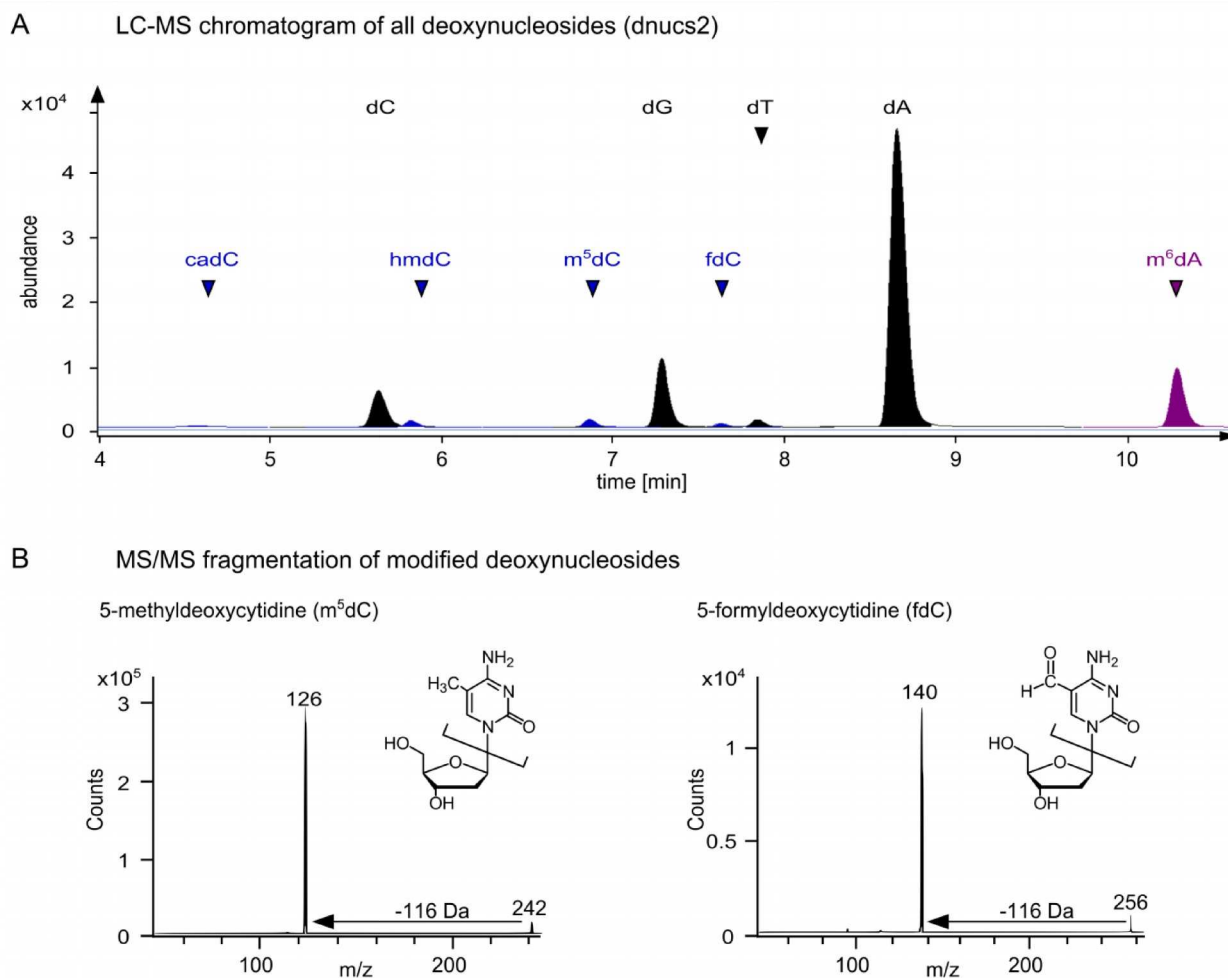


Figure 28 Chromatogram and exemplaric fragmentation pattern of deoxynucleosides.

A Chromatogram of all available natural deoxynucleosides and their modifications with method dnucs2\_DMRM. B MS/MS spectra of exemplary deoxynucleosides  $m^5dC$  and fdC.

#### 4.6.3 Limit of detection (LOD) and limit of quantification (LOQ) for selected nucleosides

The limit of detection (LOD) is the lowest amount of substance still detectable and is defined by a signal-to-noise (S/N) ratio of 3 of a given substance peak. The S/N ratio is specific for each substance peak and can be calculated by the formula:

$$S/N = \frac{2H}{h}$$

2H is the height of the substance peak, while h is the absolute height of the highest peak in a retention time window of 20x width at half the peak height of the substance signal.

The limit of quantification is the lowest amount of substance reliably quantified and it is defined by an S/N of 10. Between the LOD and LOQ the substance can be detected but the substance quantities are not in the linear range. The LOQ marks the beginning of the linear range and for mass spectrometric analysis this range can be up to 5 orders of magnitude before it

becomes nonlinear due to ionization or saturation effects. The calibration curve for determination of LODs and LOQs for ribo- $m^5C$  and deoxy- $m^5C$  are shown in Figure 29. Table 20 lists LODs and LOQs for several analyzed nucleosides. It is noteworthy that 10 amol are sufficient for detection of 7-methylguanosine and only 50 amol are needed for reliable quantification, making  $m^7G$  the best detectable nucleoside in the method setup.

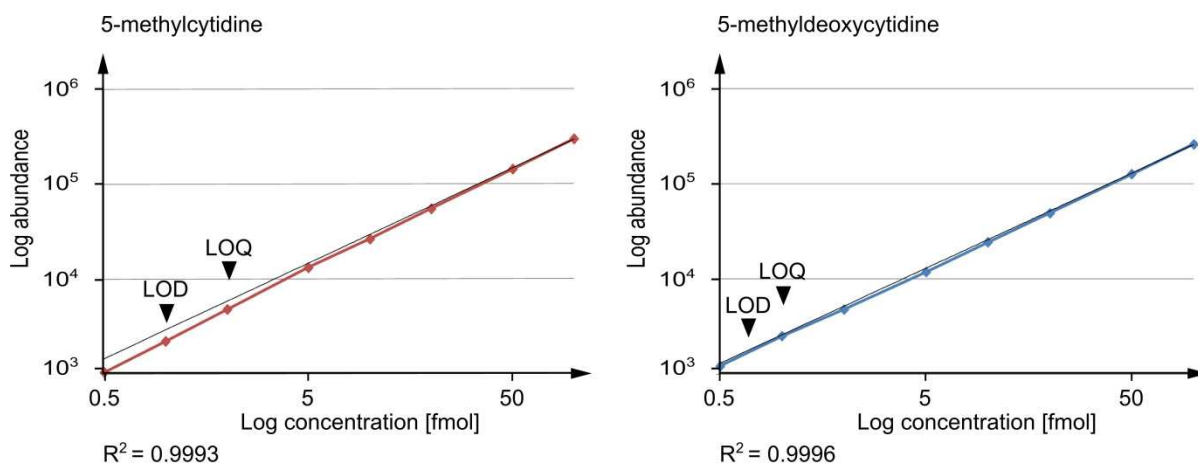


Figure 29 Calibration curves for ribo- and deoxy-5-methylcytidine with indicated LOD and LOQ

Table 20 LOD and LOQ of several modified nucleosides

Name	Short	LOD	LOQ
pseudouridine	$\Psi$	10 fmol	20 fmol
inosine	I	2 fmol	5 fmol
formyldeoxycytidine	fdC	1 fmol	5 fmol
5-methylcytidine	$m^5C$	1 fmol	2 fmol
5-methyldeoxycytidine	$m^5dC$	0.8 fmol	1 fmol
2'-O-methylguanosine	Gm	0.5 fmol	1 fmol
2'-O-methylcytidine	Cm	0.2 fmol	0.5 fmol
1-methyladenosine	$m^1A$	0.1 fmol	0.5 fmol
7-methylguanosine	$m^7G$	0.01 fmol	0.05 fmol

#### 4.6.4 General principles of LC-UV-MS/MS quantification of modified nucleosides

In most cases, quantification of modified nucleosides in RNA or DNA is not done in absolute numbers but in percent relative to the corresponding main nucleoside. This approach must be adapted to several analytical circumstances caused by the sample composition. In case of relatively high amounts of modified nucleosides, *e.g.* around 0.5 % of m<sup>6</sup>dA are present in a given DNA, UV detection of both compounds would be the method of choice. Advantageously, UV detection depends only on substance specific properties like the extinction coefficient and compound concentration. Unlike MS detection, it can be used for absolute quantification without a correction factor with and no day-to-day fluctuations are observed. For quantification of a modified nucleoside the area under the peak can be directly compared to the area under the peak of the corresponding main nucleoside since in most cases the extinction coefficients are similar [106] or are known and can be thus be accounted for. A summarized example and the calculations needed are shown in Figure 30A.

In most cases the amount of modified nucleoside is below 0.1 % and dependable UV detection is not possible due to restricted sample amounts. The method most suitable for this case is MS/MS detection of both the main and the modified nucleoside in the samples. This approach is more cumbersome, since adjustment of the analytes differing ionization efficiencies by a correction factor is necessary. This correction factor is obtained by a calibration analysis of the substances of interest, alongside the sample measurements necessary, due to interday fluctuations. An example with calculations is shown in Figure 30B.

Some modified nucleosides are rare and require a high amount of sample to reach the LOD and even more for quantitative analysis. With high sample amounts a lot of the high abundant major nucleosides are injected as well, which might lead to saturated major nucleoside MS peaks or even altered retention times. Saturated MS peaks are not in the linear range for quantification and the absolute abundance of the major nucleoside cannot be determined by MS. In this case the less sensitive UV detection of the major nucleoside and MS/MS detection of the rare modified nucleoside is a suitable solution. For data analysis a correction factor correlating the UV and MS abundances is needed. Figure 30C shows an overview of this approach.

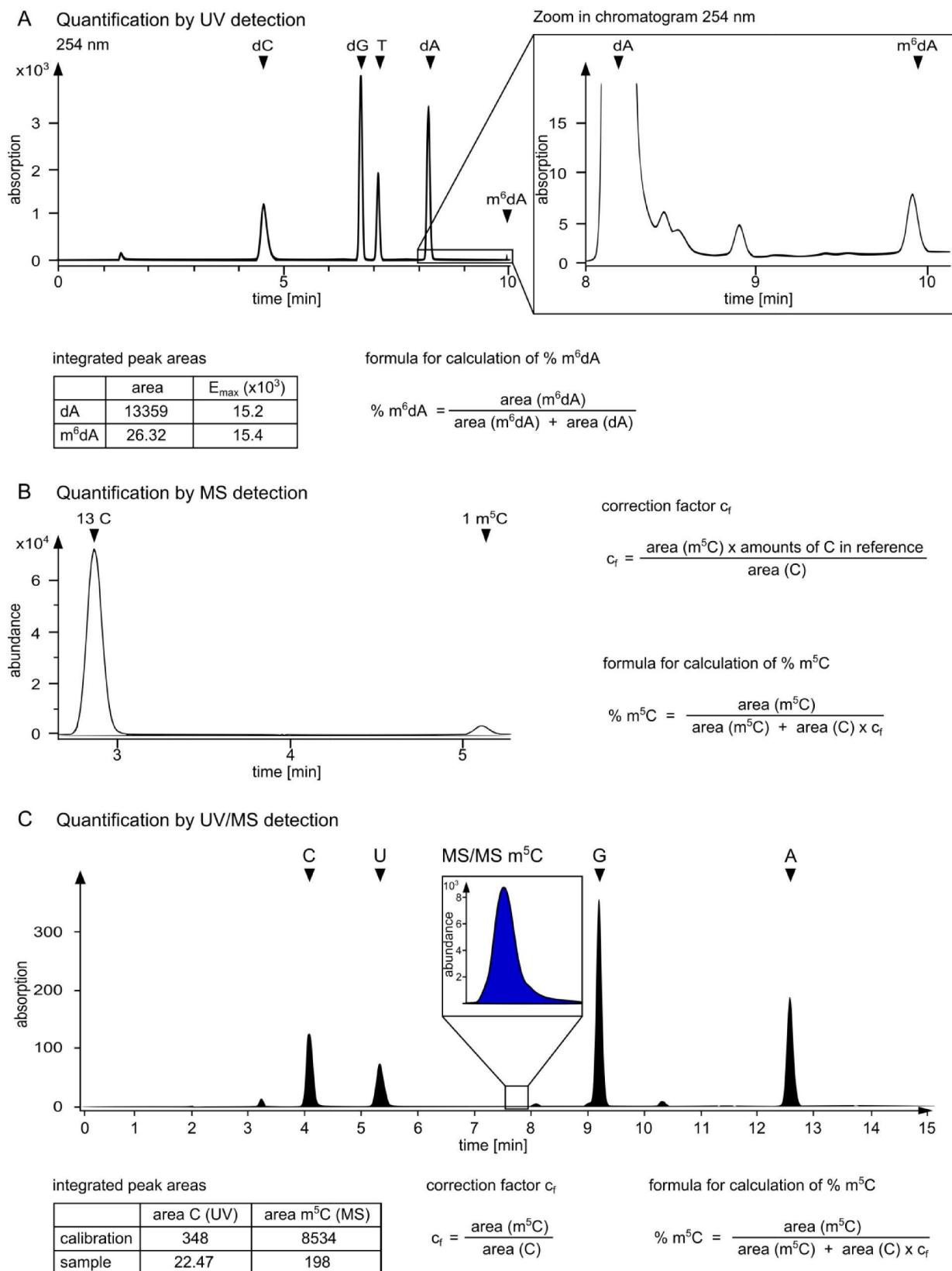


Figure 30 Possibilities for relative quantification of modified nucleosides.

A Relative quantification of the modified nucleoside  $m^6dA$  by UV detection. B Quantification of  $m^5C$  and C by MS detection. C Rare nucleosides with abundances below 0.1 % compared to the major nucleosides are quantified by combined UV and MS detection.

## **4.7 Investigations on RNA and DNA supported by developed LC-MS methods**

The methods described above were applied in several fields of nucleic acid related research. DNA extracted from haloarchaeal pleomorphic viruses was analyzed to reveal the common base modifications of cytosine and adenine. Furthermore, RNA was analyzed to investigate several RNA:m<sup>5</sup>C-methyltransferases like human NSun2 and Dnmt2 or yeast MTases. In all studies LC-MS/MS analysis for detection or quantification was performed by myself (underlined parts), while other parts of the presented studies *e.g.* preparation of knock-out organisms, bisulfite sequencing etc. were performed by the colleagues or collaborators as indicated.

#### 4.7.1 Related haloarchaeal pleomorphic viruses contain different genome types

DNA methylation has a stabilizing effect not only in terms of enzymatic degradation but also on thermal stability. Analysis of DNA extracted from haloarchaeal viruses that have to withstand the host enzymatic machinery and extreme environmental conditions surrounding the host, is of high interest. The group of Elina Roine (Department of Biosciences and Institute of Biotechnology, University of Helsinki, P.O. Box 56, FIN-00014 University of Helsinki, Finland) studied several pleomorphic haloarchaeal viruses. These viruses were sequenced and their protein expression analyzed, which led to three different viral subgroups according to their gene content and organization. One of these subgroups contain the HRPV-3 and HGPV-1 viruses, which showed discontinuities in their genomes as verified by DIG-11-UTP incorporation at the free 3'OH of the short single-stranded DNA by terminal deoxynucleotidyl transferase followed by anti-DIG antibody binding. This interesting finding was the motivation to analyze the DNA of these viruses on its modification pattern, since such single-stranded regions could be caused by disrupting base modifications or not yet reported sugar modifications. DNA from both viruses were extracted and digested for LC-UV-MS analysis. The UV analysis revealed a minor contamination with RNA, which was confirmed by identifying most minor UV peaks as RNA canonical nucleosides or their respective modified nucleosides using mass spectrometry. The main UV peaks found were identified as the main DNA nucleosides C, G, T and A by their respective mass transitions and retention times. Additionally, the presence of the DNA modifications 5-methyldeoxycytidine and N6-methyldeoxyadenosine were observed and quantified as described in Figure 31. HRPV-3 and HGPV-1 DNA contain 0.58 % and 0.66 % m<sup>5</sup>dC respectively and 0.2 % and 1.35 % m<sup>6</sup>dA. No other modified DNA nucleosides could be found in the DNA of the dsDNA haloarchaeal viruses HRPV-3 and HGPV-1. It is therefore concluded that instead of base modifications, the nucleotide sequence must be the reason for the short discontinuities in the genomes. For HRPV-3 it was possible to identify such a sequence of five nucleosides, the GCCCA motif by sequencing and alignments of the end regions. However, it is unclear whether this sequence is recognized by a viral enzyme or a host enzyme and if the short single-stranded stretches have a certain function or not.

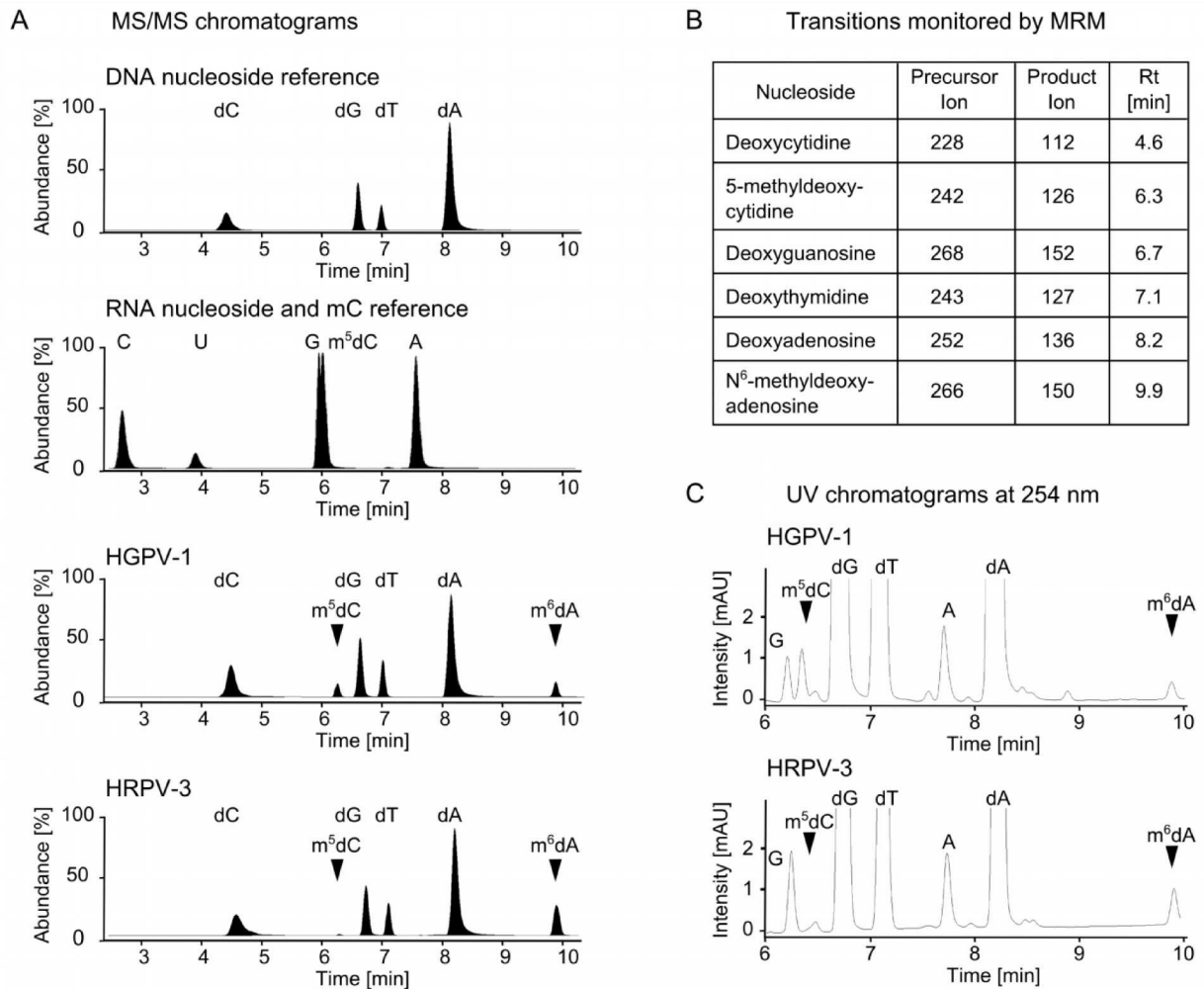


Figure 31 LC-UV/MS analysis of viral DNA

A The upper two MS/MS chromatograms show RNA and DNA digests as references, while the lower chromatograms show the results for the viral DNA digests. B The monitored mass transitions and retention times used for identification are shown here. C Chromatograms of viral DNA digests recorded at 254 nm.

Further Reading:

**Related haloarchaeal pleomorphic viruses contain different genome types**

A. Sencilo, L. Paulin, S. Kellner, M. Helm and E. Roine,

Nucleic Acids Res. 2012 Jul 1;40(12):5523-34.

Epub 2012 Mar 6.

#### 4.7.2 Site-specific cytosine-5 methylation by the mouse RNA methyltransferases Dnmt2 and NSun2 promotes tRNA stability and protein synthesis

Methylation of cytidines in DNA plays a significant role in epigenetic gene regulation and the functional enzymes and effects are intensively studied. In RNA 5-cytosine methylation is observed as a prominent post-transcriptional modification and the responsible enzymes are studied. In mice, two methyltransferases have been identified that promote methylation in cytidines of certain tRNAs, namely Dnmt2 and NSun2. However, it was not known how these enzymes interact and how important their combined functionality is to the organism. Frank Lyko and coworkers (Division of Epigenetics, Deutsches Krebsforschungszentrum, Heidelberg, Germany.) designed and performed this study. Here, single-knock-out mice with either disrupted Dnmt2 or NSun2 gene were generated and bred to obtain double-knock-out (dko) mice. Wildtype, single-ko and dko mice were characterized and comparison revealed a normal phenotype and expected Mendelian frequency for wt and single-ko mice. However, dko mice showed several morphological and phenotypic differences like: reduced size and weight, death within a few days after birth, reduced thickness and organization of cerebral cortex, incomplete skeleton ossification and reduced brown adipose tissue. Molecular analysis of tRNA methylation with bisulfite sequencing revealed a loss of C38 methylation in Dnmt2 single-ko mice, loss of C34, C48-50 methylation in NSun2 mice and a complete loss of all methylated cytidines (C34, C38, C48-50) in dko mice (see Figure 32A). In addition to bisulfite sequencing, LC-MS/MS analysis was performed, for global quantification of m<sup>5</sup>C which is not possible by bisulfite sequencing. In parallel, the effect of the enzyme knockouts on other modified nucleosides was analyzed by LC-MS/MS. tRNA was gel-purified in 3 batches from total RNA extracted from mice testes of wt, single-ko and dko mice, digested to nucleosides and the internal standard theophylline added before performing LC-MS/MS measurements. Analysis of other tRNA modifications (e.g. D, Ψ, Cm, m<sup>3</sup>C see Figure 32B) revealed no differences among the four sample types and m<sup>3</sup>C was chosen as an indicator of this finding. For comparison among sample types and the 3 sample batches, MS peak areas of the internal standard theophylline were used for adjustment. Quantitative analysis of m<sup>3</sup>C and m<sup>5</sup>C levels was performed as described in Figure 30C using the MS/MS peak areas of (modified) cytidine residues in the samples and a reference sample for generation of the correction factor. For Dnmt2 ko samples, a small but significantly reduced level of m<sup>5</sup>C compared to the wt samples was found, which is in consistence with its singular methylation position C38. For NSun2 ko samples, the m<sup>5</sup>C levels are highly reduced, as expected from its higher participation in methylation events at several cytidines within the tRNA sequence. A nearly complete loss of m<sup>5</sup>C is seen for the dko samples



with only a minor residual m<sup>5</sup>C level detectable, which is most probably caused by rRNA contamination (see Figure 32C).

Further analysis of generated immortalized MEF cells revealed a reduced level of substrate tRNAs in the dko cells, but not in the single-ko cell lines, by northern blot analysis. It is therefore concluded that the complete loss of methylation leads to a more rapid degradation of the substrate tRNAs. A loss of several tRNAs might have an effect on protein synthesis, which was observed for dko MEF cells, incubated with radioactive aminoacids, having the lowest incorporation efficiency. Additionally, sucrose gradient centrifugation revealed less polysomes (actively translating ribosomes) in dko cells compared to the wt or single-ko cells. In conclusion, the discovery of RNA methylation as a regulator of protein synthesis and cellular differentiation suggests that RNA modification participates in the epigenetic control of gene expression at the posttranscriptional level.

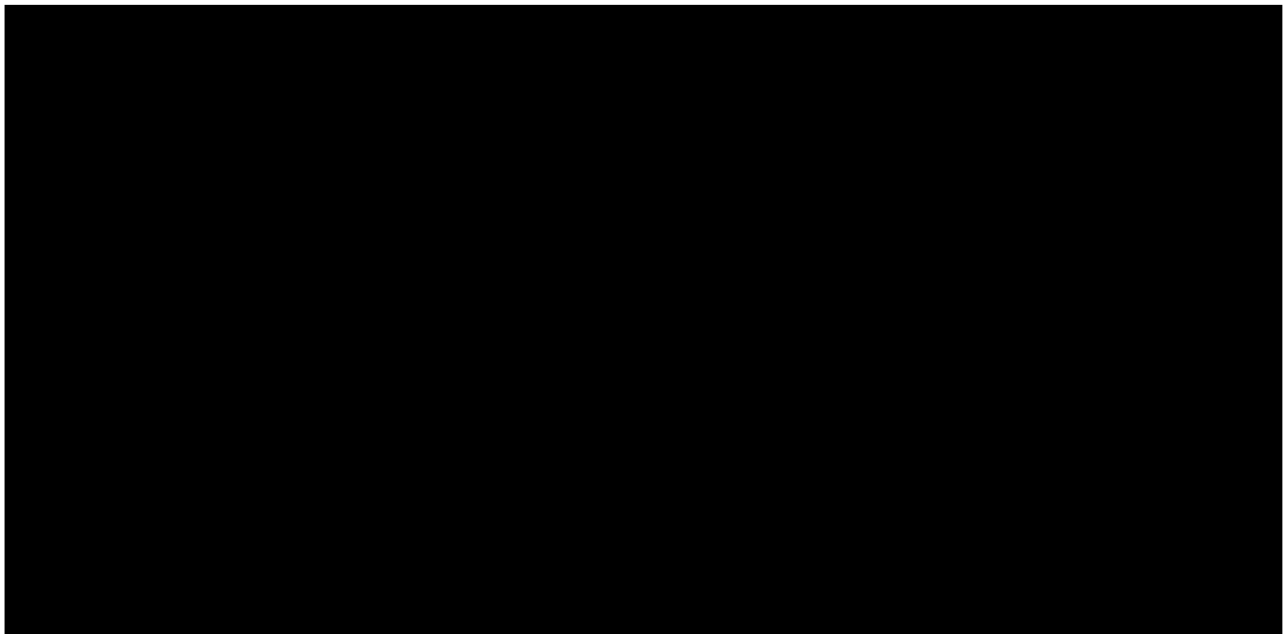


Figure 32 Analysis of m<sup>5</sup>C content in murine tRNA by bisulfite sequencing and LC-MS analysis. A Bisulfite sequencing data of Dnmt2 and NSun2 substrate tRNAs (Matthias Schaefer). B LC-MS/MS chromatogram monitoring cytidine, 5-methylcytidine (m<sup>5</sup>C) and 3-methylcytidine (m<sup>3</sup>C) in wildtype and Dnmt2/NSun2 double ko samples. C Quantitative MS/MS data analysis using MS/MS data.

Further Reading:

**Site-specific cytosine-5 methylation by the mouse RNA methyltransferases Dnmt2 and NSun2 promotes tRNA stability and protein synthesis**

Tuorto F, Liebers R, Musch T, Schaefer M, Hofmann S, Kellner S, Frye M, Helm M, Stoecklin G, Lyko F.  
Nat Struct Mol Biol. 2012 Aug 12.  
doi: 10.1038/nsmb.2357. [Epub ahead of print]

#### 4.7.3 The RNA methyltransferase Dnmt2 methylates DNA in the structural context of a tRNA

Dnmt2 is a well aligned member of the eukaryotic family of DNA 5-methyldeoxycytidine methyltransferases ( $m^5dC$  MTase), but it efficiently catalyzes tRNA methylation at C38 of *e.g.* tRNA<sup>Asp</sup>. Additionally, Dnmt2 shows nearly no activity on DNA, which gives rise to the question if the biochemical properties for DNA methylation are still present in the enzyme. To find an answer to that question several tRNA<sup>Asp</sup> hybrids, containing deoxynucleotides at specific positions of the tRNA, were synthesized in our lab by Steffen Kaiser and their substrate capability tested with a tritium incorporation assay with  $^3H$ -SAM. Figure 33A shows the used standard all-ribo construct, the hybrid dC38 construct where the target cytidine is replaced by a deoxycytidine, and the  $m^5dC38$  construct, where the target cytidine is already methylated as a negative control. The methylation reaction was monitored for 120 minutes and as expected the all-ribo tRNA shows a robust methylation, whereas the  $m^5dC38$  construct shows no methylation. However, the hybrid dC38 tRNA with the single deoxyribose substitution at the target cytidine shows a higher methylation than the natural all-ribo substrate (see Figure 33B). To confirm the structure of the methylated substrates the tRNAs were isolated and digested for LC-MS/MS analysis. In this study the identification of the methylation products was of special importance to prove DNA methylation activity of Dnmt2. Reference substances of 5-methyldeoxycytidine ( $m^5dC$ ) and 5-methylcytidine ( $m^5C$ ) were used and the MS/MS spectra (Figure 33C) and retention times were recorded. Figure 33D shows the chromatograms of the all-ribo tRNA and dC38 tRNA before and after the methylation with Dnmt2. The all-ribo tRNA shows only trace amounts of  $m^5C$  before incubation with Dnmt2 which is most likely caused by the commercial oligomer used for the tRNA synthesis. After reaction with the enzyme the  $m^5C$  peak is highly increased. In case of the dC38 hybrid no  $m^5dC$  was present before the reaction but a clear  $m^5dC$  peak can be seen after the reaction, while the contaminant  $m^5C$  peaks stays unaltered among enzyme reaction. This proves that deoxycytidine is methylated by the tRNA methyltransferase Dnmt2 if it is presented in the structural context of a tRNA. Further hybrid tRNAs with longer deoxynucleotide stretches were synthesized to analyze their substrate quality by tritium incorporation assay. In summary, an exchange of up to 10 nucleotides by deoxynucleotides in the anticodon-loop of tRNA<sup>Asp</sup> is still accepted by Dnmt2 and leads to methylation. However, these findings also show that the proper structural context is required for enzyme activity. This lends strong biochemical support to the notion that the RNA MTase activity of Dnmt2 has evolved from a DNA MTase, rather than vice versa (manuscript in preparation).

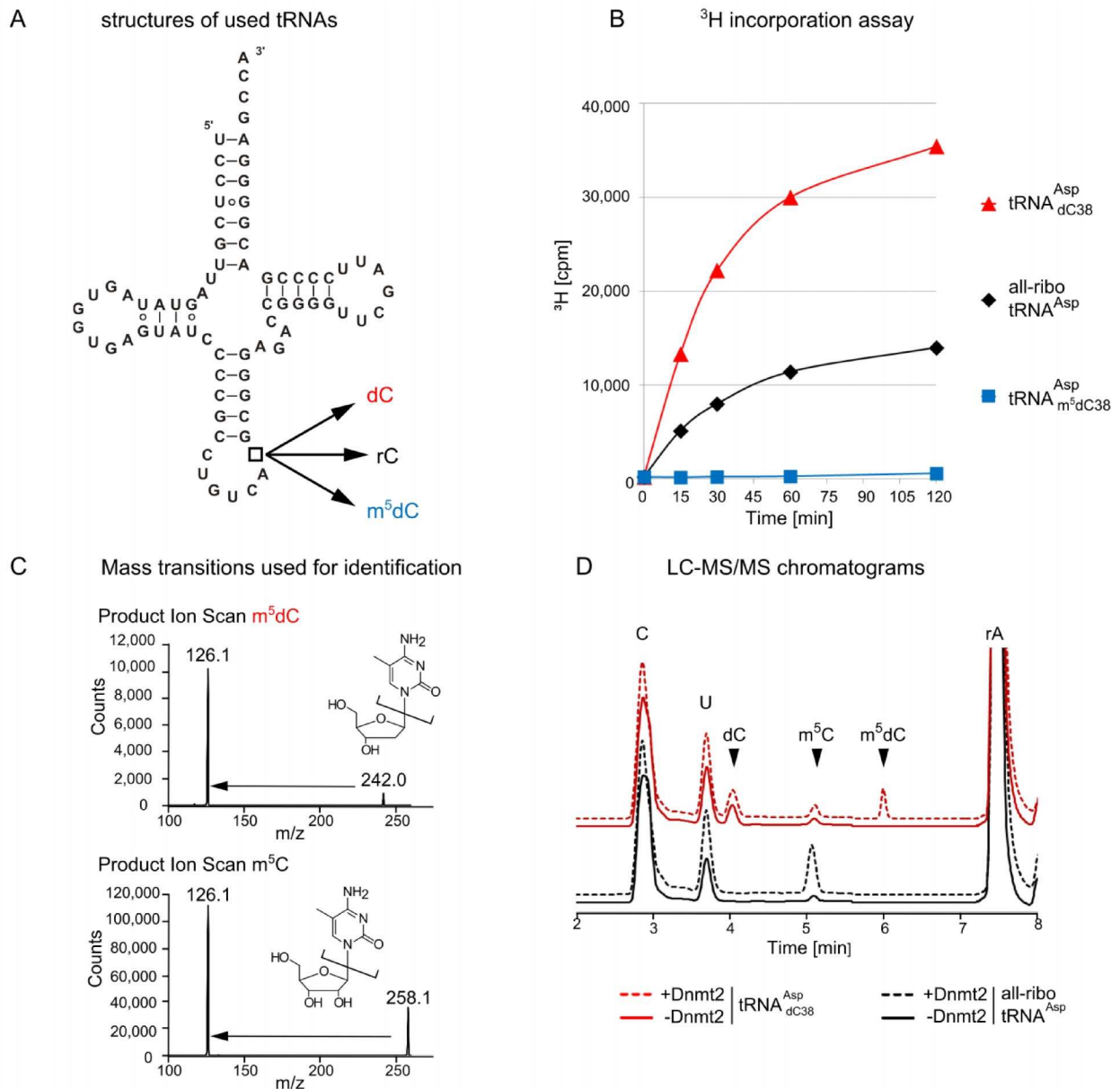
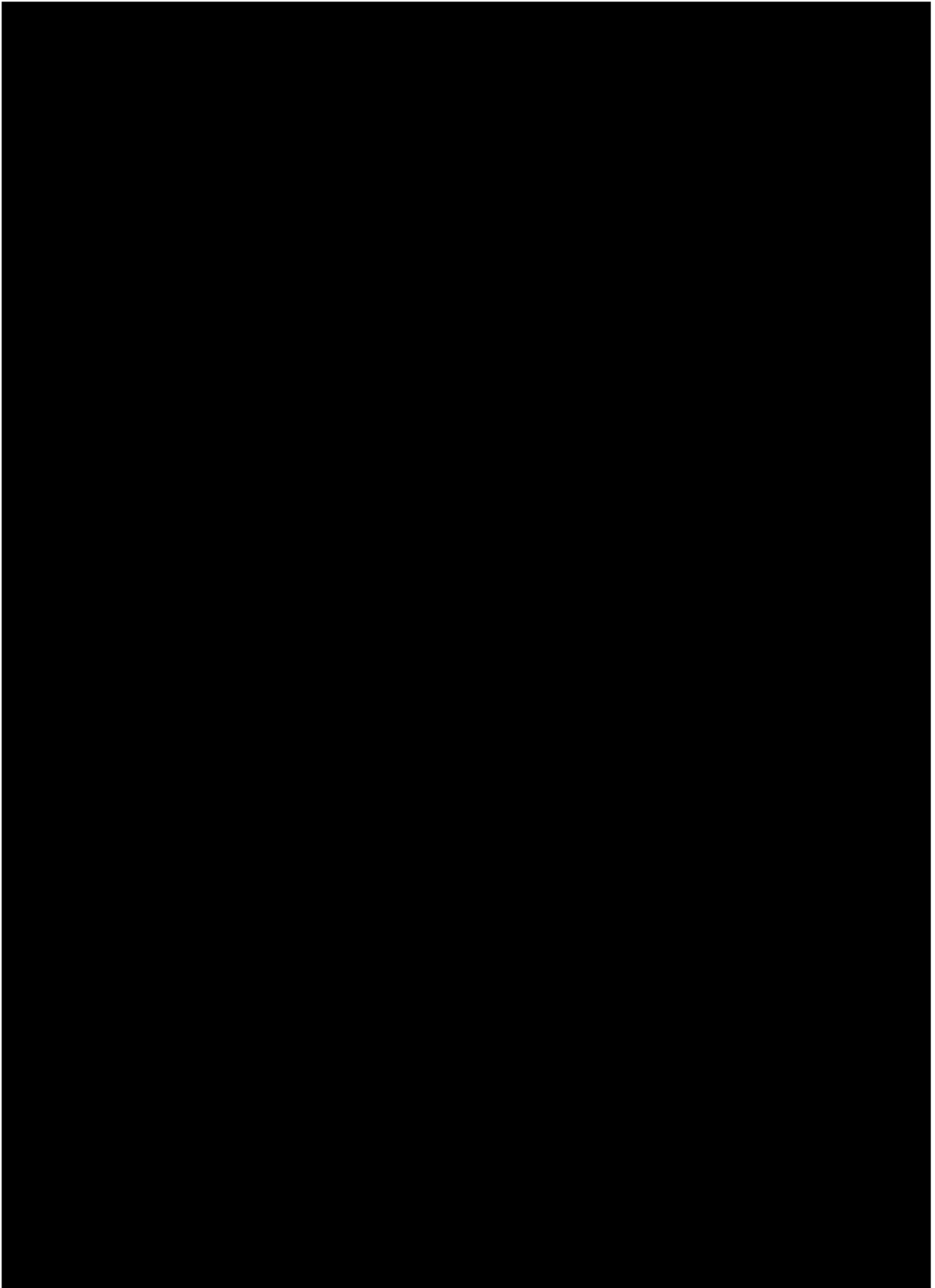
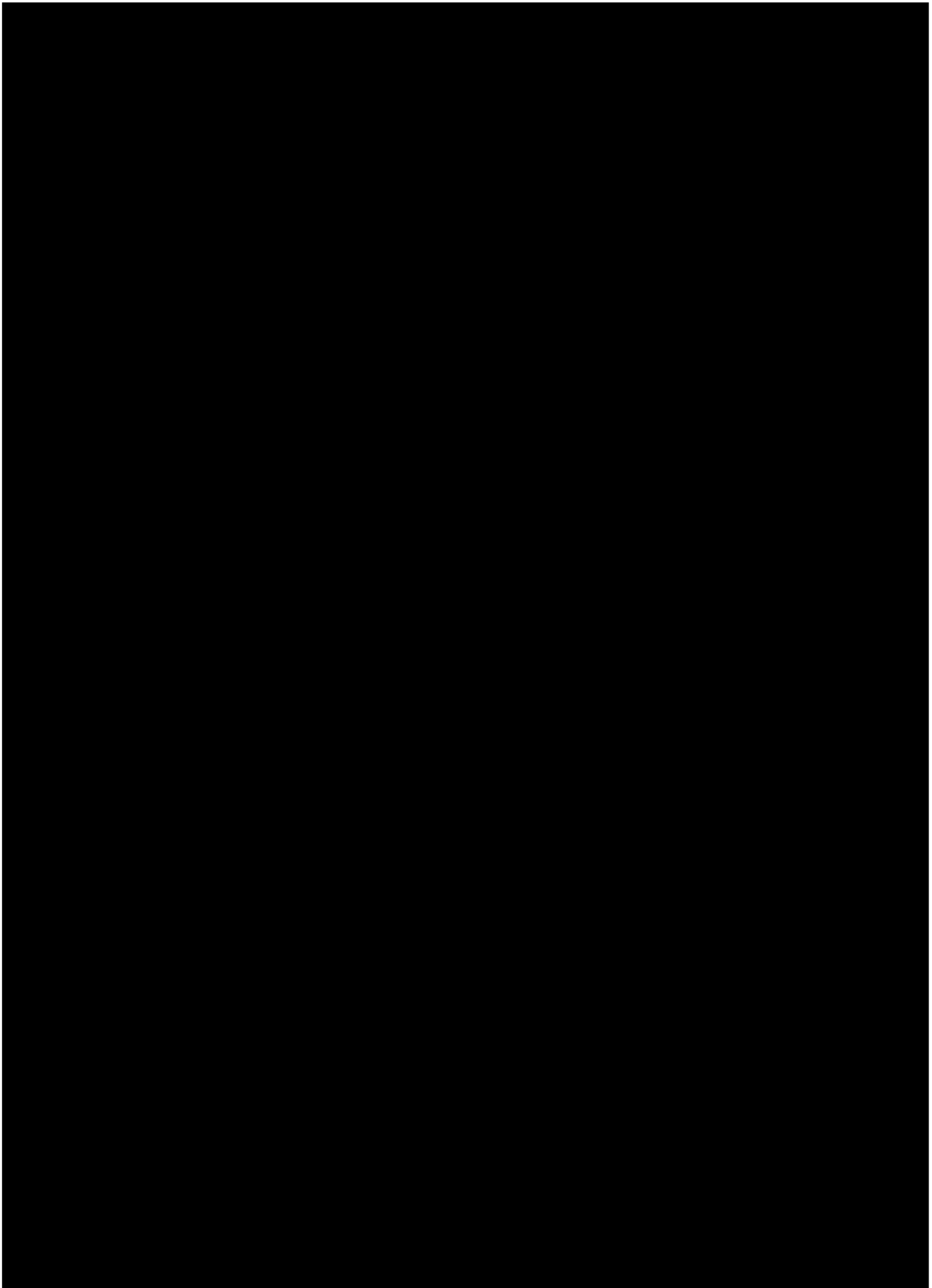


Figure 33 Methylation efficiency of all-ribo and dC38 tRNA<sup>Asp</sup> analyzed by  $^3\text{H}$ -SAM incorporation assay and LC-MS/MS analysis.

A Cloverleaf structure of the substrate tRNA<sup>Asp</sup> with C38 (black), dC38 (red) and m<sup>5</sup>dC38 (blue). B Results of  $^3\text{H}$ -SAM incorporation assay of the 3 constructs shown in A. Construct generation and measurements were done by Steffen Kaiser. C MS/MS transitions for identification of m<sup>5</sup>C and m<sup>5</sup>dC. D MS/MS chromatograms of C38 and dC38 tRNA constructs before and after incubation with Dnmt2.

#### 4.7.4 rRNA:5-methylcytosine-methyltransferase (rRNA:m<sup>5</sup>C-MTase) activity of yeast proteins Nop2, Ynl022 and human proliferation-associated antigen p120





## 5 Summary & Discussion

### *Functionalization and detection of nucleic acids with 4-bromomethylcoumarins*

Functionalization of nucleic acids and detection of modified nucleosides, by chemical reagents are highly important in RNA related research. In this work, the behavior of eight differently substituted 4-bromomethylcoumarins was tested towards nucleic acids under various reaction conditions. The resulting reaction products were identified by LC-MS/MS analysis and quantification of the nucleoside conjugates was used for determination of the coumarin selectivity for target nucleosides. For monofunctional coumarins and the previously described coumarin BMB a structure-function relationship analysis revealed a selectivity for 4-thiouridine. This confirms the findings from 1973 by Yang & Soell [78]. The later reported selectivity for pseudouridine [83] could not be achieved, neither with BMB, nor the other tested coumarins and none of the used conditions. Interestingly, the detailed structure-function relationship analysis indicates that C6 substituted compounds show a higher reactivity than their respective C7 isomers. These findings were used for development of further bromomethylcoumarins, with multifunctional features. This will be the first work publishing results on systematic structure-function relationship analysis for RNA labeling agents (manuscript submitted, see chapter 8 [2]).

Studies with the multifunctional coumarin N3BC support the previously mentioned results. N3BC shows high reactivity towards the canonical nucleoside uridine under the presented reaction conditions but no selectivity for any particular modified uridine residue. This enables the use of N3BC on all RNA and DNA molecules regardless of their modification status and knowledge of the sequence. The N3BC functionalized nucleic acid can be subjected to further applications using the azido function of the coumarin label for *e.g.* bioorthogonal CuAAC reaction with a fluorescent alkyne-dye or photo-crosslinking studies with cognate proteins. Presently N3BC is the only published multifunctional label for non-synthetic, *in vitro* or *in vivo* transcribed RNA and DNA (*e.g.* nucleic acids isolated from organisms). N3BC is a small molecule label, which fills the gap in the repertoire of RNA functionalization techniques, as shown in the summarizing Table 1. Its selectivity for uridine is a unique feature among electrophiles and its application does not require sequence information for the reaction. This is facilitated by its statistical distribution throughout all accessible uridine residues within the target nucleic acid (article published, see chapter 8 [3]).

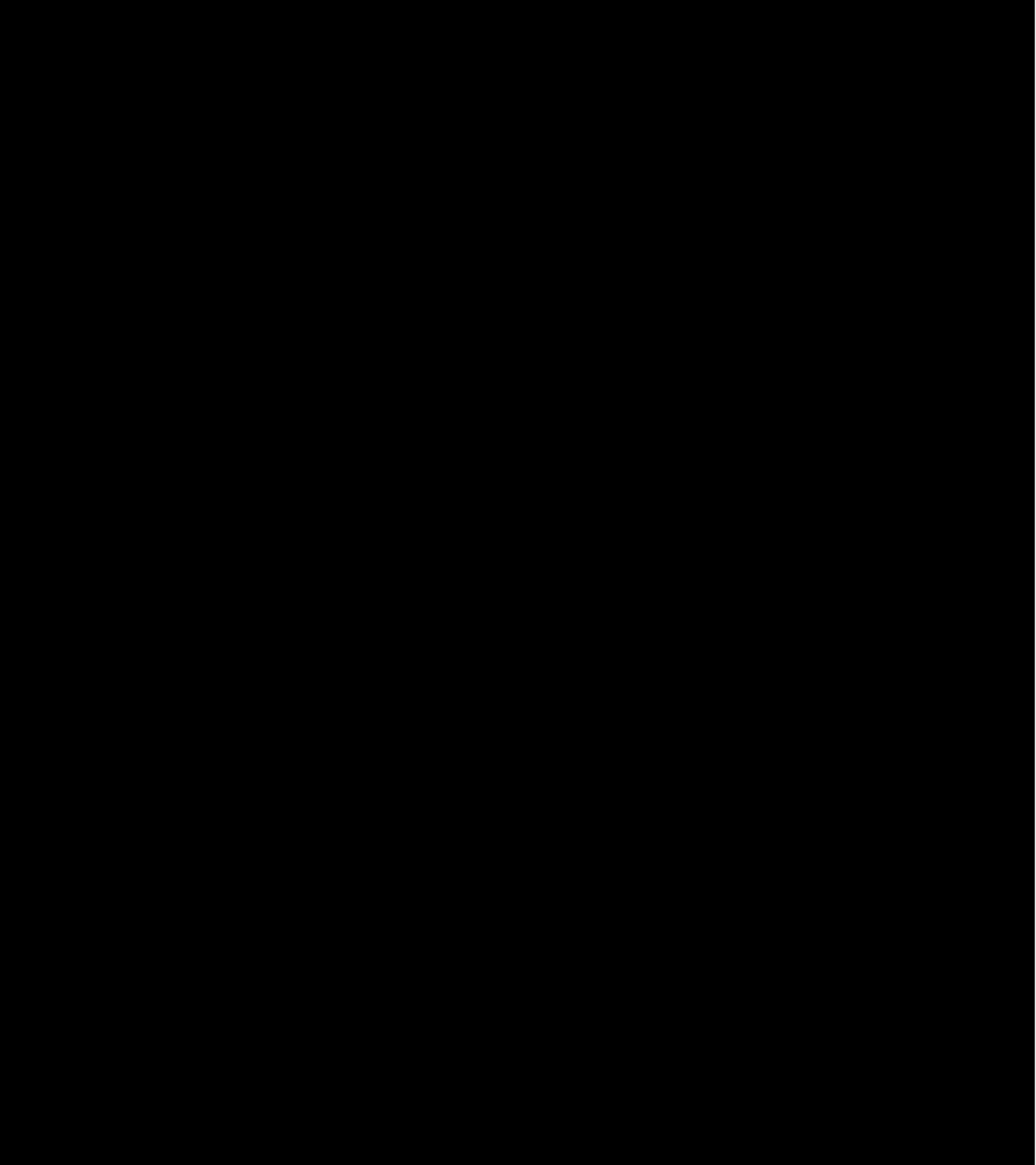
The search for more site-specific labeling agents with a high selectivity for modified nucleosides led to the discovery of the multifunctional coumarin PBC. PBC is 2000 times more selective for  $s^4U$  than for uridine or other nucleosides within total tRNA *E.coli*, when used at low DMSO reaction conditions (50%), developed by Katharina Schmid. A similar or even better selectivity could be observed for the previously described monofunctional coumarins again pointing towards an inverse relation of reactivity and selectivity. However, among all  $s^4U$  selective coumarins, PBC is the only multifunctional compound due to its terminal alkyne group. Like N3BC, this functionality enables a post-labeling functionalization of the RNA with *e.g.* a fluorescent azide-dye (manuscript in preparation). In addition, PBC-atto647N labeled siRNA was successfully transfected into living cells and no cytotoxicity was observed. On-going investigations comprise further improvement of the selectivity on  $s^4U$  especially towards other modified nucleosides like 2-thiouridine ( $s^2U$ ).

#### ***Detection of modified nucleosides with LC-MS/MS analysis***

Several methods for nucleoside analysis with a robust HPLC system equipped with a diode array detector and highly sensitive triple quadrupole detector were established. One of the developed, sophisticated chromatography methods allows separation and quantification of up to 25 (modified) ribonucleosides in a single run and the mass transitions of these nucleosides were optimized. A similar method was developed for 9 naturally occurring deoxyribonucleosides. The limit of detection in MS/MS mode was determined for several modified nucleosides revealing 7-methylguanosine as the best detectable with an LOD of 10 amol and an LOQ of 50 amol. For the epigenetically interesting nucleosides  $m^5C$  and  $m^5dC$  the LOD was determined to be 1 fmol and the LOQ 2-5 fmol. General quantification principles using UV, MS/MS or both detection methods for different sample requirements are presented. All developed methods are of comparable quality and sensitivity to state-of-the-art LC-MS/MS methods used in RNA and DNA related research [92-94]. The methods were applied to support several studies on *e.g.* viral DNA or RNA modifying enzymes.

DNA extracted from haloarchaeal pleomorphic viruses (HRPV-3 and HGPV-1) was subjected to the method dnucs2 and its modification pattern analyzed. HRPV-3 and HGPV-1 DNA contain 0.58 % and 0.66 %  $m^5dC$  respectively and 0.2 % and 1.35 %  $m^6dA$  but no other modified nucleosides could be detected. Single-stranded regions in the DNA of these viruses is most likely caused by specific sequences and not by modified nucleosides, since no anomaly could be found in the analyzed DNA samples. This study was designed and mainly conducted by the lab of Elina Roine (article published, see chapter 8 [4]).

A study on the function of 5-methylcytidine in tRNA made use of knock-out mice lacking either the m<sup>5</sup>C:methyltransferase Dnmt2, NSun2 or both. tRNA from wildtype, single-ko and double-ko mice was isolated and the m<sup>5</sup>C levels quantified by MS/MS detection. In case of the single-ko mice reduced m<sup>5</sup>C levels compared to wildtype were observed while only residual m<sup>5</sup>C was detectable in dko mice tRNA. Bisulfite sequencing additionally revealed 3 more locations of 5-methylated cytosines in tRNA, namely C48-50. Detailed analysis performed by the lab of Frank Lyko revealed a decrease in m<sup>5</sup>C containing tRNA levels and subsequent reduced protein biosynthesis in the dko mice (article published, see chapter 8 [5]).





## 6 Conclusion & Outlook

### *Functionalization and detection of RNA and its modifications*

Coumarins were shown to be interesting scaffolds for derivatization of RNA. The electronic properties of the coumarin are easily modulated by substitution of varying chemical moieties, which alters reactivity and selectivity.

The uridine targeting coumarin N3BC can be used on all RNA species without further knowledge of sequence and modification information. The introduction of an azide function into the RNA is of several advantages and should be further used for standard functionalization of nucleic acids that need *e.g.* fluorescent labels like siRNAs or aptamers. Coumarin labeled siRNA, clicked to a fluorescent dye, is accepted by cells and uptake was observed and described in this work. In the future, this labeling technique must be further investigated and knock-down efficiencies of thus labeled siRNA must be examined. In addition, this technique can be used for generation of triple-FRET siRNA constructs with an internal label introduced by coumarin modification. Such constructs are needed to gain more information on the siRNA integrity after cellular uptake and distribution. Commercially labeled siRNA is commonly labeled at a phosphate group and is therefore prone to enzymatic degradation by phosphatases. This makes studies on siRNA integrity cumbersome as a loss in FRET efficiency might be either caused by siRNA degradation or cleavage of the label, leaving the siRNA otherwise intact. The coumarin label is resistant to enzymatic cleavage and can thus be used for such studies.

With PBC it was shown that selectivity for modified RNA nucleosides can be achieved, while introducing a further functionality, an alkyne group, into the RNA. Future syntheses and analyses of differently modified coumarins should enable selective targeting of other modified nucleosides. These selective labels will expand the possibilities for detection of modified nucleosides and isolation of thus modified RNA.

LC-MS/MS analysis of nucleic acids and its modifications is state-of-the-art for detection of nucleosides. Several methods for identification and quantification of modified nucleosides have been developed and applied to samples including investigations on DNA and RNA. However, with the hitherto developed methods, only 25 of the over 100 known modified nucleosides can be identified and quantified in a single run, due to a lack of available pure nucleosides as references. Furthermore improvement of LODs and LOQs should be attempted by using UPLC methods with smaller column particles and higher system pressures, which are possible up to 600 bar with the used system. UPLC has the advantage of shorter analysis time and sharper eluting

peaks, which leads to a more sensitive high-throughput method with expected 5-10 fold better LODs and LOQs.

A general drawback of the presented nucleoside analytic is the complete loss of sequence information. This could be overcome by avoiding complete digest of the given nucleic acid by using specific endonucleases like RNase T1 or RNase A to yield short oligonucleotides [104]. These could be analyzed by MS compatible ion-pair chromatography and multiple-ionization of the heavy ON which results in lower  $m/z$  values detectable by the used MS/MS system. This approach allows sufficient knowledge on the sequence of the digested nucleic acid and the location of the modification. However, some challenges remain: 1.) Pure sample isolation is necessary, 2.) The multiple-ionization of the digested ON in the MS must be controlled to a defined amount of charges per fragment 3.) The LOD is approximately 1000 fold higher compared to nucleoside analytics.

LC-MS/MS analysis has proven useful in identification of coumarin conjugates and furthermore in quantification of the newly formed conjugates. These methods can be easily adapted to possible new coumarins used for derivatization or even other labeling agents with alternate scaffolds. Additionally, a LC-MS/MS detection method allowing sequence information to exactly locate the labeled nucleoside should be developed using the same principle as mentioned above.

## 7 References

1. Maxam, A.M. and W. Gilbert, *A new method for sequencing DNA*. Proc Natl Acad Sci U S A, 1977. **74**(2): p. 560-4.
2. Peattie, D.A., *Direct chemical method for sequencing RNA*. Proc Natl Acad Sci U S A, 1979. **76**(4): p. 1760-4.
3. Sanger, F., S. Nicklen, and A.R. Coulson, *DNA sequencing with chain-terminating inhibitors*. Proc Natl Acad Sci U S A, 1977. **74**(12): p. 5463-7.
4. Smith, L.M., et al., *Fluorescence detection in automated DNA sequence analysis*. Nature, 1986. **321**(6071): p. 674-9.
5. Felden, B., et al., *Non-canonical substrates of aminoacyl-tRNA synthetases: the tRNA-like structure of brome mosaic virus genomic RNA*. Biochimie, 1993. **75**(12): p. 1143-57.
6. Moras, D., et al., *The structure of yeast tRNA(Asp). A model for tRNA interacting with messenger RNA*. J Biomol Struct Dyn, 1985. **3**(3): p. 479-93.
7. Romby, P., et al., *Yeast tRNA<sup>Asp</sup> tertiary structure in solution and areas of interaction of the tRNA with aspartyl-tRNA synthetase. A comparative study of the yeast phenylalanine system by phosphate alkylation experiments with ethylnitrosourea*. J Mol Biol, 1985. **184**(3): p. 455-71.
8. Godeau, G., C. Staedel, and P. Barthelemy, *Lipid-conjugated oligonucleotides via "click chemistry" efficiently inhibit hepatitis C virus translation*. J Med Chem, 2008. **51**(15): p. 4374-6.
9. Slim, G., et al., *Synthesis of site-specifically modified oligoribonucleotides for studies of the recognition of TAR RNA by HIV-1 tat protein and studies of hammerhead ribozymes*. Nucleic Acids Symp Ser, 1991(24): p. 55-8.
10. Tuschl, T., et al., *A three-dimensional model for the hammerhead ribozyme based on fluorescence measurements*. Science, 1994. **266**(5186): p. 785-9.
11. Qin, P.Z. and A.M. Pyle, *Stopped-flow fluorescence spectroscopy of a group II intron ribozyme reveals that domain I is an independent folding unit with a requirement for specific Mg<sup>2+</sup> ions in the tertiary structure*. Biochemistry, 1997. **36**(16): p. 4718-30.
12. Kopecky, K., et al., *Solid-phase synthesis of azaphthalocyanine-oligonucleotide conjugates and their evaluation as new dark quenchers of fluorescence*. Bioconjug Chem, 2010. **21**(10): p. 1872-9.
13. El-Sagheer, A.H. and T. Brown, *New strategy for the synthesis of chemically modified RNA constructs exemplified by hairpin and hammerhead ribozymes*. Proc Natl Acad Sci U S A, 2010. **107**(35): p. 15329-34.
14. Aigner, M., et al., *Chemical synthesis of site-specifically 2'-azido-modified RNA and potential applications for bioconjugation and RNA interference*. Chembiochem, 2011. **12**(1): p. 47-51.
15. Fauster, K., et al., *2'-Azido RNA, a versatile tool for chemical biology: synthesis, X-ray structure, siRNA applications, click labeling*. ACS Chem Biol, 2012. **7**(3): p. 581-9.
16. Richardson, C.C., *Phosphorylation of nucleic acid by an enzyme from T4 bacteriophage-infected Escherichia coli*. Proc Natl Acad Sci U S A, 1965. **54**(1): p. 158-65.
17. Richardson, R.W. and R.I. Gumport, *Biotin and fluorescent labeling of RNA using T4 RNA ligase*. Nucleic Acids Res, 1983. **11**(18): p. 6167-84.
18. Hengesbach, M., et al., *RNA intramolecular dynamics by single-molecule FRET*. Curr Protoc Nucleic Acid Chem, 2008. **Chapter 11**: p. Unit 11 12.
19. Ryder, S.P. and S.A. Strobel, *Nucleotide analog interference mapping*. Methods, 1999. **18**(1): p. 38-50.

20. Aurup, H., D.M. Williams, and F. Eckstein, *2'-Fluoro- and 2'-amino-2'-deoxynucleoside 5'-triphosphates as substrates for T7 RNA polymerase*. *Biochemistry*, 1992. **31**(40): p. 9636-41.
21. Jao, C.Y. and A. Salic, *Exploring RNA transcription and turnover in vivo by using click chemistry*. *Proc Natl Acad Sci U S A*, 2008. **105**(41): p. 15779-84.
22. Rao, H., et al., *Enzymatic incorporation of an azide-modified UTP analog into oligoribonucleotides for post-transcriptional chemical functionalization*. *Nat Protoc*, 2012. **7**(6): p. 1097-112.
23. Caron, M. and H. Dugas, *Specific spin-labeling of transfer ribonucleic acid molecules*. *Nucleic Acids Res*, 1976. **3**(1): p. 19-34.
24. Silverman, S.K. and T.R. Cech, *RNA tertiary folding monitored by fluorescence of covalently attached pyrene*. *Biochemistry*, 1999. **38**(43): p. 14224-37.
25. Ledoan, T., et al., *High specific radioactivity labeling of oligonucleotides with 3H-succinimidyl propionate*. *Nucleosides Nucleotides*, 1999. **18**(2): p. 277-89.
26. Sletten, E.M. and C.R. Bertozzi, *Bioorthogonal chemistry: fishing for selectivity in a sea of functionality*. *Angew Chem Int Ed Engl*, 2009. **48**(38): p. 6974-98.
27. Winz, M.L., et al., *Site-specific terminal and internal labeling of RNA by poly(A) polymerase tailing and copper-catalyzed or copper-free strain-promoted click chemistry*. *Nucleic Acids Res*, 2012. **40**(10): p. e78.
28. Uttamapinant, C., et al., *Fast, cell-compatible click chemistry with copper-chelating azides for biomolecular labeling*. *Angew Chem Int Ed Engl*, 2012. **51**(24): p. 5852-6.
29. Agard, N.J., J.A. Prescher, and C.R. Bertozzi, *A strain-promoted [3 + 2] azide-alkyne cycloaddition for covalent modification of biomolecules in living systems*. *J Am Chem Soc*, 2004. **126**(46): p. 15046-7.
30. van Delft, P., et al., *Synthesis of oligoribonucleic acid conjugates using a cyclooctyne phosphoramidite*. *Org Lett*, 2010. **12**(23): p. 5486-9.
31. Peattie, D.A. and W. Gilbert, *Chemical probes for higher-order structure in RNA*. *Proc Natl Acad Sci U S A*, 1980. **77**(8): p. 4679-82.
32. Pochon, F. and M. Perrin, *Fluorescent labelling of polynucleotides by 9-bromomethylanthracene*. *Eur J Biochem*, 1974. **43**(1): p. 107-13.
33. Koreeda, M., et al., *Alkylation of polyguanylic acid at the 2-amino group and phosphate by the potent mutagen (+/-)-7beta,8alpha-dihydroxy-9beta,10beta-epoxy-7,8,9,10-tetrahydrobenzo[a]pyrene*. *J Am Chem Soc*, 1976. **98**(21): p. 6720-2.
34. Draper, D.E. and L. Gold, *A method for linking fluorescent labels to polynucleotides: application to studies of ribosome-ribonucleic acid interactions*. *Biochemistry*, 1980. **19**(9): p. 1774-81.
35. Yoshida, S., S. Hirose, and M. Iwamoto, *Use of 4-bromomethyl-7-methoxycoumarin for derivatization of pyrimidine compounds in serum analysed by high-performance liquid chromatography with fluorimetric detection*. *J Chromatogr*, 1986. **383**(1): p. 61-8.
36. Vlassov, V.V., R. Giege, and J.P. Ebel, *The tertiary structure of yeast tRNAPhe in solution studied by phosphodiester bond modification with ethylnitrosourea*. *FEBS Lett*, 1980. **120**(1): p. 12-6.
37. Garcia, A., R. Giege, and J.P. Behr, *New photoactivatable structural and affinity probes of RNAs: specific features and applications for mapping of spermine binding sites in yeast tRNA(Asp) and interaction of this tRNA with yeast aspartyl-tRNA synthetase*. *Nucleic Acids Res*, 1990. **18**(1): p. 89-95.
38. Proudnikov, D. and A. Mirzabekov, *Chemical methods of DNA and RNA fluorescent labeling*. *Nucleic Acids Res*, 1996. **24**(22): p. 4535-42.
39. Randerath, K., *Application of a tritium derivative method to human brain and brain tumor transfer RNA analysis*. *Cancer Res*, 1971. **31**(5): p. 658-61.
40. Latham, J.A. and T.R. Cech, *Defining the inside and outside of a catalytic RNA molecule*. *Science*, 1989. **245**(4915): p. 276-82.

41. Merino, E.J., et al., *RNA structure analysis at single nucleotide resolution by selective 2'-hydroxyl acylation and primer extension (SHAPE)*. J Am Chem Soc, 2005. **127**(12): p. 4223-31.
42. Lucks, J.B., et al., *Multiplexed RNA structure characterization with selective 2'-hydroxyl acylation analyzed by primer extension sequencing (SHAPE-Seq)*. Proc Natl Acad Sci U S A, 2011. **108**(27): p. 11063-8.
43. Giese, R., M. Helm, and C. Florentz, *Classical and Novel Chemical Tools for RNA Structure Probing (Chapter 5)*. 2001. p. 71-89.
44. Onizuka, K., Y. Taniguchi, and S. Sasaki, *Development of novel thioguanosine analogs with the ability to specifically modify cytidine*. Nucleic Acids Symp Ser (Oxf), 2007(51): p. 5-6.
45. Onizuka, K., Y. Taniguchi, and S. Sasaki, *A new usage of functionalized oligodeoxynucleotide probe for site-specific modification of a guanine base within RNA*. Nucleic Acids Res, 2010. **38**(5): p. 1760-6.
46. Sasaki, S., K. Onizuka, and Y. Taniguchi, *Oligodeoxynucleotide containing s-functionalized 2'-deoxy-6-thioguanosine: facile tools for base-selective and site-specific internal modification of RNA*. Curr Protoc Nucleic Acid Chem, 2012. **Chapter 4**: p. Unit 4 49 1-16.
47. Cantara, W.A., et al., *The RNA Modification Database, RNAMDB: 2011 update*. Nucleic Acids Res, 2011. **39**(Database issue): p. D195-201.
48. Kobitski, A.Y., et al., *Single-molecule FRET reveals a cooperative effect of two methyl group modifications in the folding of human mitochondrial tRNA(Lys)*. Chem Biol, 2011. **18**(7): p. 928-36.
49. Carell, T., et al., *Structure and function of noncanonical nucleobases*. Angew Chem Int Ed Engl, 2012. **51**(29): p. 7110-31.
50. Agris, P.F., *Bringing order to translation: the contributions of transfer RNA anticodon-domain modifications*. EMBO Rep, 2008. **9**(7): p. 629-35.
51. Agris, P.F., F.A. Vendeix, and W.D. Graham, *tRNA's wobble decoding of the genome: 40 years of modification*. J Mol Biol, 2007. **366**(1): p. 1-13.
52. Wilkinson, M.L., et al., *The 2'-O-methyltransferase responsible for modification of yeast tRNA at position 4*. RNA, 2007. **13**(3): p. 404-13.
53. Durant, P.C. and D.R. Davis, *Stabilization of the anticodon stem-loop of tRNA<sup>Lys,3</sup> by an A+-C base-pair and by pseudouridine*. J Mol Biol, 1999. **285**(1): p. 115-31.
54. Helm, M., R. Giese, and C. Florentz, *A Watson-Crick base-pair-disrupting methyl group (m1A9) is sufficient for cloverleaf folding of human mitochondrial tRNA<sup>Lys</sup>*. Biochemistry, 1999. **38**(40): p. 13338-46.
55. Muramatsu, T., et al., *Codon and amino-acid specificities of a transfer RNA are both converted by a single post-transcriptional modification*. Nature, 1988. **336**(6195): p. 179-81.
56. Muramatsu, T., et al., *A novel lysine-substituted nucleoside in the first position of the anticodon of minor isoleucine tRNA from Escherichia coli*. J Biol Chem, 1988. **263**(19): p. 9261-7.
57. Cowling, V.H., *Regulation of mRNA cap methylation*. Biochem J, 2010. **425**(2): p. 295-302.
58. Horwich, M.D., et al., *The Drosophila RNA methyltransferase, DmHen1, modifies germline piRNAs and single-stranded siRNAs in RISC*. Curr Biol, 2007. **17**(14): p. 1265-72.
59. Vicens, Q. and E. Westhof, *Crystal structure of paromomycin docked into the eubacterial ribosomal decoding A site*. Structure, 2001. **9**(8): p. 647-58.
60. Motorin, Y., F. Lyko, and M. Helm, *5-methylcytosine in RNA: detection, enzymatic formation and biological functions*. Nucleic Acids Res, 2010. **38**(5): p. 1415-30.

61. Hayatsu, H., *Discovery of bisulfite-mediated cytosine conversion to uracil, the key reaction for DNA methylation analysis--a personal account*. Proc Jpn Acad Ser B Phys Biol Sci, 2008. **84**(8): p. 321-30.
62. Schaefer, M., et al., *RNA cytosine methylation analysis by bisulfite sequencing*. Nucleic Acids Res, 2009. **37**(2): p. e12.
63. Squires, J.E., et al., *Widespread occurrence of 5-methylcytosine in human coding and non-coding RNA*. Nucleic Acids Res, 2012. **40**(11): p. 5023-33.
64. King, M.Y. and K.L. Redman, *RNA methyltransferases utilize two cysteine residues in the formation of 5-methylcytosine*. Biochemistry, 2002. **41**(37): p. 11218-25.
65. Goll, M.G., et al., *Methylation of tRNA<sup>Asp</sup> by the DNA methyltransferase homolog Dnmt2*. Science, 2006. **311**(5759): p. 395-8.
66. Ratel, D., et al., *N6-methyladenine: the other methylated base of DNA*. Bioessays, 2006. **28**(3): p. 309-15.
67. Kriaucionis, S. and N. Heintz, *The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain*. Science, 2009. **324**(5929): p. 929-30.
68. Tahiliani, M., et al., *Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1*. Science, 2009. **324**(5929): p. 930-5.
69. Pfaffeneder, T., et al., *The discovery of 5-formylcytosine in embryonic stem cell DNA*. Angew Chem Int Ed Engl, 2011. **50**(31): p. 7008-12.
70. Ito, S., et al., *Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine*. Science, 2011. **333**(6047): p. 1300-3.
71. Jurkowski, T.P. and A. Jeltsch, *Burning off DNA methylation: new evidence for oxygen-dependent DNA demethylation*. Chembiochem, 2011. **12**(17): p. 2543-5.
72. Kellner, S., J. Burhenne, and M. Helm, *Detection of RNA modifications*. RNA Biol, 2010. **7**(2): p. 237-47.
73. Bakin, A. and J. Ofengand, *Four newly located pseudouridylate residues in Escherichia coli 23S ribosomal RNA are all at the peptidyltransferase center: analysis by the application of a new sequencing technique*. Biochemistry, 1993. **32**(37): p. 9754-62.
74. Emmerechts, G., P. Herdewijn, and J. Rozenski, *Pseudouridine detection improvement by derivatization with methyl vinyl sulfone and capillary HPLC-mass spectrometry*. J Chromatogr B Analyt Technol Biomed Life Sci, 2005. **825**(2): p. 233-8.
75. Dominissini, D., et al., *Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq*. Nature, 2012. **485**(7397): p. 201-6.
76. Behm-Ansmant, I., M. Helm, and Y. Motorin, *Use of specific chemical reagents for detection of modified nucleotides in RNA*. J Nucleic Acids, 2011. **2011**: p. 408053.
77. Watson, B.S., et al., *Macromolecular arrangement in the aminoacyl-tRNA<sub>elongation</sub> factor Tu.GTP ternary complex. A fluorescence energy transfer study*. Biochemistry, 1995. **34**(24): p. 7904-12.
78. Yang, C.H. and D. Soll, *Covalent attachment of a fluorescent group to 4-thiouridine in transfer RNA*. J Biochem, 1973. **73**(6): p. 1243-7.
79. Pingoud, A., R. Kownatzki, and G. Maass, *Fluoresceinylthiocarbamyl-tRNA<sup>Tyr</sup>: a useful derivative of tRNA<sup>Tyr</sup> (E.coli) for physicochemical studies*. Nucleic Acids Res, 1977. **4**(2): p. 327-38.
80. Wilkinson, K.A., E.J. Merino, and K.M. Weeks, *Selective 2'-hydroxyl acylation analyzed by primer extension (SHAPE): quantitative RNA structure analysis at single nucleotide resolution*. Nat Protoc, 2006. **1**(3): p. 1610-6.
81. Durairaj, A. and P.A. Limbach, *Improving CMC-derivatization of pseudouridine in RNA for mass spectrometric detection*. Anal Chim Acta, 2008. **612**(2): p. 173-81.
82. Mengel-Jorgensen, J. and F. Kirpekar, *Detection of pseudouridine and other modifications in tRNA by cyanoethylation and MALDI mass spectrometry*. Nucleic Acids Res, 2002. **30**(23): p. e135.

83. Yang, C. and D. Soll, *Covalent attachment of fluorescent groups to transfer ribonucleic acid. Reactions with 4-bromomethyl-7-methoxy-2-oxo-2H-benzopyran*. *Biochemistry*, 1974. **13**(17): p. 3615-21.
84. Nagamatsu, K. and Y. Miyazawa, *Partial melting of the segment around pseudouridine in yeast 5S RNA*. *Biochem Biophys Res Commun*, 1983. **114**(1): p. 81-7.
85. Davis, G.E., et al., *High-performance liquid chromatographic separation and quantitation of nucleosides in urine and some other biological fluids*. *Clin Chem*, 1977. **23**(8): p. 1427-35.
86. Gehrke, C.W., K.C. Kuo, and R.W. Zumwalt, *Chromatography of nucleosides*. *J Chromatogr*, 1980. **188**(1): p. 129-47.
87. Costa, A., et al., *Determination of queuosine derivatives by reverse-phase liquid chromatography for the hypomodification study of Q-bearing tRNAs from various mammal liver cells*. *J Chromatogr B Analyt Technol Biomed Life Sci*, 2004. **801**(2): p. 237-47.
88. Hou, S. and M. Ding, *Simultaneous separation and determination of eleven nucleosides and bases in beer, herring sperm DNA and RNA soft capsule by high-performance liquid chromatography*. *Anal Sci*, 2010. **26**(10): p. 1111-4.
89. Niessen, W.M., *State-of-the-art in liquid chromatography-mass spectrometry*. *J Chromatogr A*, 1999. **856**(1-2): p. 179-97.
90. Pomerantz, S.C. and J.A. McCloskey, *Analysis of RNA hydrolyzates by liquid chromatography-mass spectrometry*. *Methods Enzymol*, 1990. **193**: p. 796-824.
91. Neubauer, S., et al., *Mass spectrometry based analysis of nucleotides, nucleosides, and nucleobases-application to feed supplements*. *Anal Bioanal Chem*, 2012. **404**(3): p. 799-808.
92. Li, X. and A.A. Franke, *High-throughput and cost-effective global DNA methylation assay by liquid chromatography-mass spectrometry*. *Anal Chim Acta*, 2011. **703**(1): p. 58-63.
93. Zhang, J.J., et al., *Analysis of global DNA methylation by hydrophilic interaction ultra high-pressure liquid chromatography tandem mass spectrometry*. *Anal Biochem*, 2011. **413**(2): p. 164-70.
94. Song, L., et al., *Specific method for the determination of genomic DNA methylation by liquid chromatography-electrospray ionization tandem mass spectrometry*. *Anal Chem*, 2005. **77**(2): p. 504-10.
95. Tao, W.W., et al., *Determination of nucleosides and nucleobases in the pollen of *Typha angustifolia* by UPLC-PDA-MS*. *Phytochem Anal*, 2012. **23**(4): p. 373-8.
96. Rodriguez-Muniz, G.M., et al., *Reactivity of nucleosides with a hydroxyl radical in non-aqueous medium*. *Chemistry*, 2012. **18**(26): p. 8024-7.
97. Zhao, H., et al., *Simultaneous determination nucleosides in marine organisms using ultrasound-assisted extraction followed by hydrophilic interaction liquid chromatography-electrospray ionization time-of-flight mass spectrometry*. *J Sep Sci*, 2011. **34**(19): p. 2594-601.
98. Zhang, L., et al., *Simultaneous determination of global DNA methylation and hydroxymethylation levels by hydrophilic interaction liquid chromatography-tandem mass spectrometry*. *J Biomol Screen*, 2012. **17**(7): p. 877-84.
99. Bruckl, T., et al., *Parallel isotope-based quantification of modified tRNA nucleosides*. *Angew Chem Int Ed Engl*, 2009. **48**(42): p. 7932-4.
100. Globisch, D., et al., *Systems-based analysis of modified tRNA bases*. *Angew Chem Int Ed Engl*, 2011. **50**(41): p. 9739-42.
101. Hossain, M. and P.A. Limbach, *Multiple endonucleases improve MALDI-MS signature digestion product detection of bacterial transfer RNAs*. *Anal Bioanal Chem*, 2009. **394**(4): p. 1125-35.

102. Wetzel, C. and P.A. Limbach, *Global identification of transfer RNAs by liquid chromatography-mass spectrometry (LC-MS)*. J Proteomics, 2012. **75**(12): p. 3450-64.
103. Addepalli, B. and P.A. Limbach, *Mass spectrometry-based quantification of pseudouridine in RNA*. J Am Soc Mass Spectrom, 2011. **22**(8): p. 1363-72.
104. Krog, J.S., et al., *3-(3-amino-3-carboxypropyl)-5,6-dihydrouridine is one of two novel post-transcriptional modifications in tRNALys(UUU) from Trypanosoma brucei*. FEBS J, 2011. **278**(24): p. 4782-96.
105. Kellner, S., et al., *A multifunctional bioconjugate module for versatile photoaffinity labeling and click chemistry of RNA*. Nucleic Acids Res, 2011. **39**(16): p. 7348-60.
106. Hall, D.B.D.a.R.H., *Purines, Pyrimidines, Nucleosides, and Nucleotides in Handbook of Biochemistry and Molecular Biology* May 2010. p. 269-358.
107. Grosjean, H., et al., *Detection of enzymatic activity of transfer RNA modification enzymes using radiolabeled tRNA substrates*. Methods Enzymol, 2007. **425**: p. 55-101.
108. Lebwohl, M., P.T. Ting, and J.Y. Koo, *Psoriasis treatment: traditional therapy*. Ann Rheum Dis, 2005. **64 Suppl 2**: p. ii83-6.
109. Hirsch, M., D. Strand, and M. Helm, *Dye selection for live cell imaging of intact siRNA*. Biol Chem, 2012. **393**(1-2): p. 23-35.
110. Jarve, A., et al., *Surveillance of siRNA integrity by FRET imaging*. Nucleic Acids Res, 2007. **35**(18): p. e124.
111. Lakowicz, J.R., *Principles of Fluorescence Spectroscopy, Third Edition*. Journal of Biomedical Optics, 2008. **13**(2): p. 029901-029901.
112. Dudley, E., et al., *Study of the mass spectrometric fragmentation of pseudouridine: comparison of fragmentation data obtained by matrix-assisted laser desorption/ionisation post-source decay, electrospray ion trap multistage mass spectrometry, and by a method utilising electrospray quadrupole time-of-flight tandem mass spectrometry and in-source fragmentation*. Rapid Commun Mass Spectrom, 2005. **19**(21): p. 3075-85.
113. Crain, P.F., *Detection and Structure Analysis of Modified Nucleosides by Mass Spectrometry*. Modification and editing of RNA, ed. H. Grosjean and R. Benne. 1998, Washington, DC :: ASM Press.







## 8 Published Articles & Reviews

### Review:

1. Kellner, S., J. Burhenne, and M. Helm, Detection of RNA modifications. RNA Biol, 2010. **7**(2): p. 237-47. PMID:20224293

### Articles:

2. S. Kellner, L. Kollar, A. Ochel, M. Ghate & M. Helm, Structure-function relationship of substituted bromomethylcoumarins in nucleoside specificity of RNA alkylation. Submitted to Bioconjugate Chemistry (ACS Publications)
3. Kellner S., Seidu-Larry S., Burhenne J., Motorin Y., Helm M., A multifunctional bioconjugate module for versatile photoaffinity labeling and click chemistry of RNA. Nucleic Acids Res, 2011. **39**(16): p. 7348-60. PMID: 21646334
4. Sencilo A., Paulin L., Kellner S., Helm M., Roine E., Related haloarchaeal pleomorphic viruses contain different genome types. Nucleic Acids Res, 2012. **40**(12): p. 5523-34. PMID: 22396526
5. Tuorto F., Liebers R., Musch T., Schaefer M., Hofmann S., Kellner S., Frye M., Helm M., Stoecklin G., Lyko F., RNA cytosine methylation by Dnmt2 and NSun2 promotes tRNA stability and protein synthesis. Nat Struct Mol Biol, 2012. **19**(9): p. 900-5. PMID: 22885326

### Declaration of contribution of each publication to this work and declaration of my contribution as a co-author:

Publication	contribution of each publication to this work	contribution as a co-author
[1]	5 %	-
[2]	20 %	-
[3]	35 %	-
[4]	20 %	20 %
[5]	20 %	20 %

### Declaration of confirmation:

---

Prof. Mark Helm

---

Stefanie Kellner



**Detection of RNA modifications.**

S. Kellner, J.Burhenne & M. Helm  
RNA Biol. 2010 Mar-Apr;7(2):237-47.  
Epub 2010 Mar 8.



**Structure-function relationship of substituted  
bromomethylcoumarins in nucleoside specificity of RNA  
alkylation**

S. Kellner, L. Kollar, A. Ochel, M. Ghaté & M. Helm

Bioconjugate Chemistry (ACS Publications)

submitted





**A multifunctional bioconjugate module for versatile  
photoaffinity labeling and click chemistry of RNA**

S. Kellner, S. Seidu-Larry, J.Burhenne, Y.Motorin & M. Helm

Nucleic Acids Res. 2011 Sep 1;39(16):7348-60.

Epub 2011 Jun 6



**Related haloarchaeal pleomorphic viruses contain different  
genome types**

A. Sencilo, L. Paulin, S. Kellner, M. Helm and E. Roine,

Nucleic Acids Res. 2012 Jul 1;40(12):5523-34.

Epub 2012 Mar 6.



**Site-specific cytosine-5 methylation by the mouse RNA methyltransferases Dnmt2 and NSun2 promotes tRNA stability and protein synthesis**

Tuorto F, Liebers R, Musch T, Schaefer M, Hofmann S, Kellner S, Frye M, Helm M, Stoecklin G, Lyko F.

Nat Struct Mol Biol. 2012 Aug 12.

doi: 10.1038/nsmb.2357. [Epub ahead of print]



# 9 Appendix

## 1.1 Mass transitions of coumarin conjugates

Table A 1 Mass Transitions, retention times and correction factors of monofunctional coumarin conjugates.

	Parent Ion	Product Ion	Rt [min]	AA [MRM]	AA [320 nm]	Correction factor [MRM]/[320nm]
<b>Compound 1 (BMB)</b>						
G-BMB	472.3	340.2	10.5 12.1	134351	2.4	55980
<sup>4</sup> s U-BMB	449.4	317.1	12.3	26017	5.7	4564
U-BMB	433.1	301.2	11.4	3469	0.5	6938
Ψ-BMB	433.1	343.1	10.4 10.7	57	0.05	1140
<b>compound 2</b>						
G-2	472.2	340.4	9.7, 10.25, 11.15	957934	81	11826
<sup>4</sup> s U-2	449.2	317.2	11.3	139427	121	1152
U-2	433.2	301.4	10.7	639138	928	689
Ψ-2	433.0	343.4	9.7, 10.2	16920	50	338
<b>compound 3</b>						
G-3	456.2	324.4	9.99, 10.6, 11.52	372744	3.5	106498
<sup>4</sup> s U-3	433.2	301.3	11.72	208167	136	1531
U-3	417.2	285.3	11.0	423895	173	2450
Ψ-3	417.2	327.4	10.0	8954	21	426
<b>compound 4</b>						
G-4	456.2	324.4	10.01, 10.65, 11.52	816947	20	40847
<sup>4</sup> s U-4	433.2	301.3	11.72	125518	96	1307
U-4	417.2	285.3	11.1	1111204	602	1846
Ψ-4	417.2	327.4	10.0 10.43	17952	21	855

## Appendix

Table A 1 Continuation: Mass Transitions, retention times and correction factors of monofunctional coumarin conjugates.

	Parent Ion	Product Ion	Rt [min]	AA [MRM]	AA [320 nm]	Correction factor [MRM]/[320nm]
<b>compound 5</b>						
G-5	492.2	360.4	11.9	60218	1.6	37636
<sup>4</sup> s U-5	469.2	337.2	13.3	14586	11,5	1268
U-5	453.2	321.4	12.6	68901	71	970
Ψ-5	453.2	363.4	11.6	7844	22.8	344
<b>compound 6</b>						
G-6	492.2	360.4	12.5	89205	2	44603
<sup>4</sup> s U-6	469.2	337.2	13.5	21464	26	826
U-6	453.2	321.4	12.9	198481	212	936
Ψ-6	453.2	363.4	11.9	5968	15	398

Table A 2 Mass Transitions, retention times and correction factors of multifunctional coumarin conjugates.

	Parent Ion	Product Ion	Rt [min]	AA [MRM]	AA [320 nm]	Correction factor [MRM]/[320nm]
<b>N3BC</b>						
G-N3C	483	351	11.0 11.1 11.6	-	-	-
U-N3C	444	284	11.8	-	-	-
Ψ-N3C	444	354	11.1	-	-	-
<b>PBC</b>						
G-PBC	496.5	364.2	11.6	-	-	-
<sup>4</sup> s U-PBC	473.2	341.2	12.8	105479	71.8	1469
U-PBC	457.2	325.2	12.1	23671	19.7	1197
Ψ-PBC	457	367	11.2	-	-	-



## 1.2 Non denaturing gel of hybridized siRNA

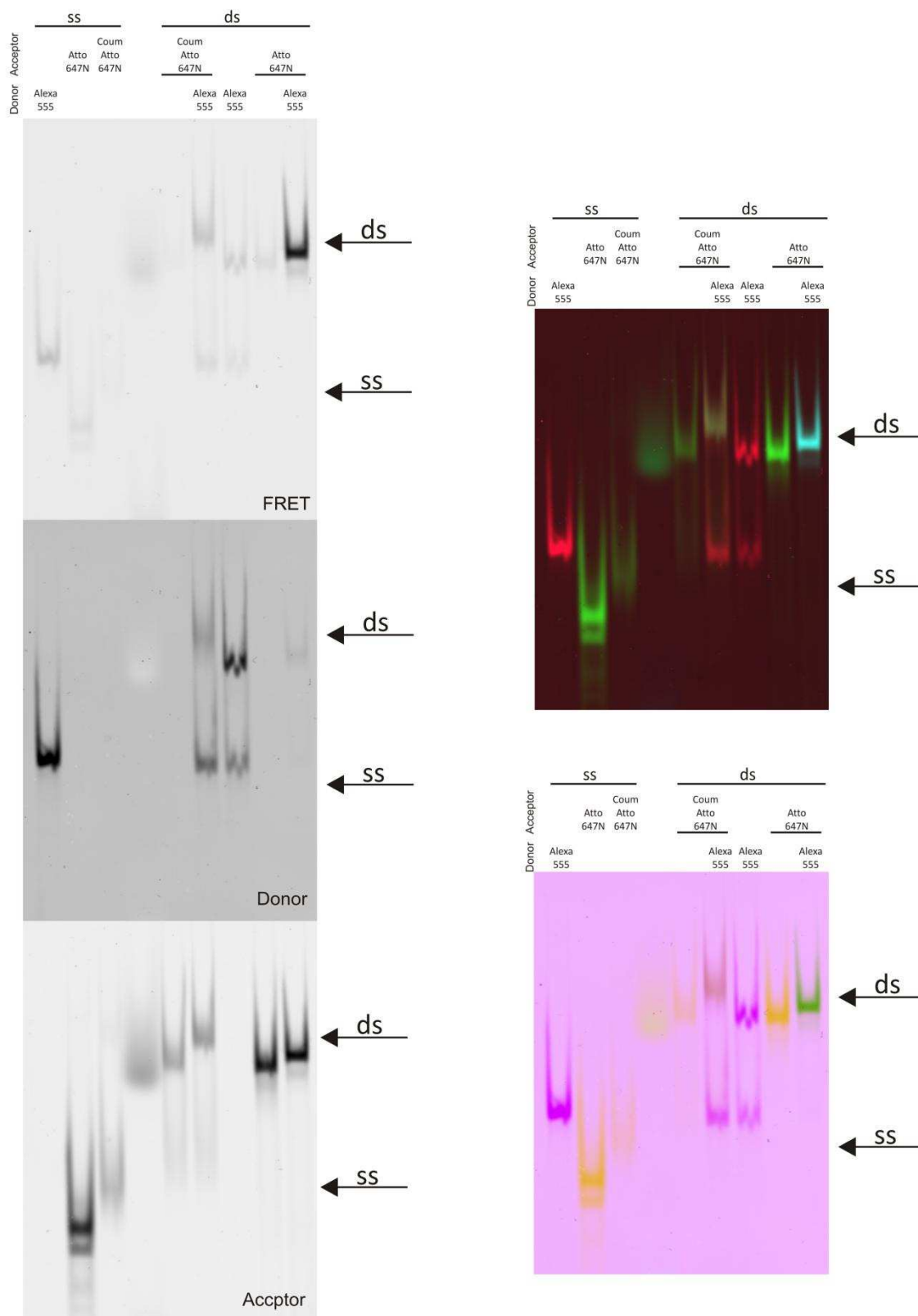
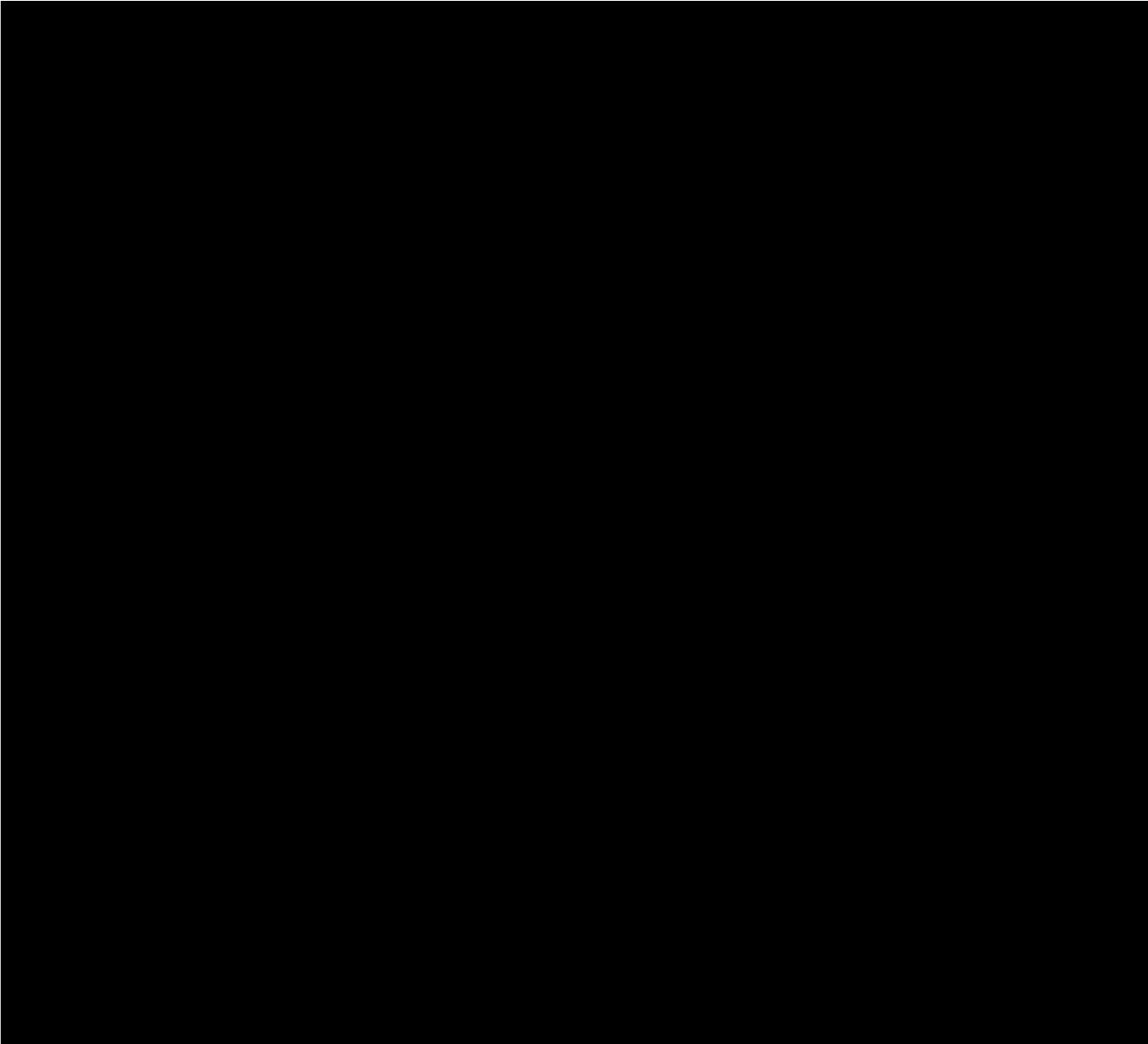
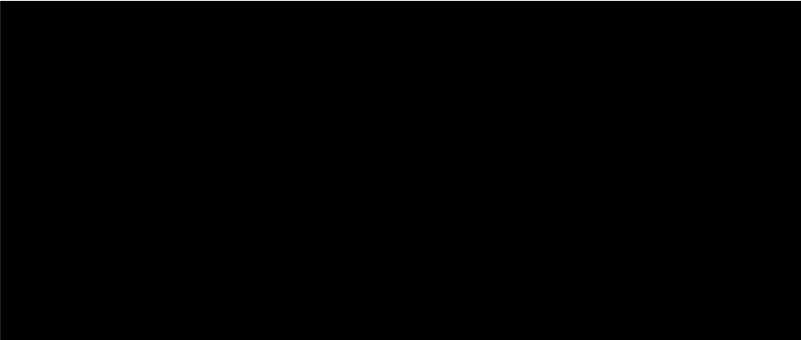


Figure A 1 Non-denaturing gel of hybridized siRNA constructs

The gel scans on the left show FRET, donor and acceptor fluorescence of the constructs. The gel pictures on the right are overlay pictures of the three gel scans. On the upper left gel, is a faint FRET signal for alexa555/PBC-atto647N construct detectable.



# LEBENS LAUF



---

Ort, Datum

---

Stefanie Kellner