

Structural dynamics of m⁶A formation in mRNA

Dissertation

zur Erlangung des Doktorgrades
der Naturwissenschaften

vorgelegt beim Fachbereich Biochemie, Chemie und Pharmazie
der Johann Wolfgang Goethe–Universität
in Frankfurt am Main

von

Nathalie Meiser

aus Frankfurt am Main

Frankfurt 2023

(D 30)

Vom Fachbereich Biochemie, Chemie und Pharmazie der
Johann Wolfgang Goethe-Universität als Dissertation angenommen.

Dekan: Prof. Dr. Clemens Glaubitz

Gutachter: Dr. Martin Hengesbach
Prof. Dr. Harald Schwalbe

Datum der Disputation: 29.06.2023

Table of contents

Summary	1
1 Introduction	9
1.1 RNA modifications.....	9
1.1.1 N6-methyladenosine.....	11
1.1.2 The role of N6-methyladenosine in mRNA	12
1.1.3 Identification of N6-methyladenosine sites.....	15
1.2 Human N6-methyladenosine writer complex.....	18
1.2.1 Catalytic core METTL3/METTL14	18
1.2.2 Accessory proteins	21
1.3 Principles of single molecule FRET spectroscopy.....	23
1.3.1 Förster resonance energy transfer.....	23
1.3.2 Single molecule FRET spectroscopy	25
1.3.3 Total internal reflection fluorescence microscopy	26
1.3.4 TIRF sample preparation	29
2 Motivation	35
3 Results and discussion – Specificity of the METTL3/METTL14 complex	39
3.1 Structural specificity of the full-length METTL3/METTL14 complex.....	39
3.1.1 Preparation of the full-length METTL3/METTL14 complex	40
3.1.2 Target RNA constructs.....	42
3.1.3 Site-specific RNA <i>in vitro</i> methylation assay	43
3.1.4 Methylation activity determination	45
3.1.5 Conclusion and outlook.....	47
3.2 Influence of protein domains on substrate specificity and catalytic activity	50
3.2.1 Preparation of shortened METTL3/METTL14 complexes	50
3.2.2 General methyltransferase activity.....	55
3.2.3 RNA binding properties.....	56
3.2.4 Secondary structure specificity	59
3.2.5 Conclusion and outlook.....	62
3.3 Influence of WTAP.....	64
3.3.1 Generation of expression constructs (Gibson assembly).....	65
3.3.2 Protein preparation.....	67
3.3.3 Secondary structure specificity	71
3.3.4 Conclusion and outlook.....	73

Table of contents

4	Results and discussion - smFRET studies on the METTL3/METTL14 complex	77
4.1	smFRET setup construction.....	77
4.1.1	Excitation.....	78
4.1.2	Microscope.....	80
4.1.3	Creating of total internal reflection	82
4.1.4	Detection.....	84
4.1.5	Remote control software	85
4.1.6	Software application	87
4.2	smFRET measurements.....	89
4.2.1	smFRET construct design and preparation	89
4.2.2	smFRET state analysis	92
4.2.3	Analysis of dynamic behavior	99
4.2.4	Conclusion and outlook	118
5	Materials and methods.....	123
5.1	Buffers and Media.....	123
5.1.1	Media	123
5.1.2	Buffers.....	123
5.2	General biochemical methods	125
5.2.1	Concentration determination of biomolecules	125
5.2.2	Precipitation of nucleic acids	125
5.2.3	Electrophoretic methods	126
5.2.4	Polymerase chain reaction.....	129
5.2.5	<i>In vitro</i> transcription	129
5.2.6	Fluorescent labeling of nucleic acids	131
5.3	Nucleic acid constructs	132
5.4	RNA melting point analysis	133
5.5	Protein constructs.....	134
5.5.1	Amino acid sequences	134
5.5.2	Expression constructs	138
5.5.3	DNA plasmid preparation.....	140
5.6	Protein expression in <i>Spodoptera frugiperda</i> 9 (Sf9) cells.....	144
5.6.1	Handling of Sf9 cells.....	144
5.6.2	Recombinant baculovirus generation via the flashBAC system.....	144
5.6.3	Recombinant baculovirus generation via the Bac-to-Bac system	144

5.6.4	Virus amplification	145
5.6.5	Test expressions	145
5.6.6	Large scale expressions	146
5.7	Protein purification	146
5.7.1	Ni-NTA affinity purification	146
5.7.2	Size exclusion chromatography	147
5.8	Methyltransferase activity determination	147
5.8.1	Site-specific ³² P-labeling of substrate RNAs	147
5.8.2	Methyltransferase activity assay	149
5.8.3	Modification analysis	150
5.9	Microscale thermophoresis	151
5.10	smFRET spectroscopy	151
5.10.1	Materials for construction of smFRET microscope setup	151
5.10.2	Sample slide preparation	153
5.10.3	Sample preparation	154
5.10.4	smFRET data acquisition	155
5.10.5	smFRET data analysis	155
6	Appendix	163
6.1	Protein coding gene sequences	163
6.1.1	METTL3	163
6.1.2	METTL14	164
6.1.3	WTAP	165
6.1.4	RBM15	165
6.1.5	VIRMA	166
6.1.6	ZC1H13	167
6.2	RNA melting point analysis	169
6.3	Optimization of assay conditions	169
6.4	Protein purification of shortened METTL3/METTL14 complexes	170
6.4.1	M3ΔZF1/METTL14	170
6.4.2	M3MTD/METTL14	171
6.4.3	METTL3/M14ΔRGG	172
6.4.4	METTL3/M14ΔNTD	173
6.4.5	METTL3/M14MTD	174
6.4.6	M3MTD/M14MTD	175

Table of contents

6.5	Protein purification of the METTL3/METTL14/(short)WTAP complex	176
6.6	Custom made components for the smFRET setup	177
6.7	Matlab scripts for smFRET data analysis.....	180
6.7.1	Analysis of histogram data.....	180
6.7.2	Selection of movie traces.....	181
6.7.3	Export movie traces for HaMMY.....	184
6.7.4	Analysis of HaMMY identified states	185
6.7.5	Analysis of HaMMY dwell times.....	188
6.7.6	Population analysis of dynamic molecules	189
6.8	smFRET data.....	191
6.8.1	EMSA of the smFRET construct.....	191
6.8.2	Transition density plots.....	192
6.8.3	Movie analysis by individual transitions	194
6.8.4	Exponential decay fits for dwell time distribution.....	201
6.8.5	Transition rates from exponential decay fit.....	204
6.8.6	Transition rates from average dwell times	205
	References.....	206
	List of abbreviations.....	216
	Deutsche Zusammenfassung	218

Summary

N6-methyladenosine (m^6A) is the most abundant and well understood modification in eukaryotic mRNA and was first identified in polyadenylated parts of the mRNA.^{1,2} The distinct distribution of m^6A in the transcriptome with special enrichment in long internal exons, 3'UTRs and around stop codons was uncovered by early biochemical work and later on antibody based sequencing techniques.³⁻⁷ The so called m^6A writer, reader and eraser machinery is responsible for the dynamic and with that regulatory nature of the m^6A modification.⁸ As m^6A writer, the human N6-methyltransferase complex (MTC) cotranscriptionally methylates the central adenine within a RRACH (preferably GGACU) sequence context to form m^6A in the nascent RNA chain.⁹⁻¹⁵ The catalytic core of the complex is formed by the two proteins METTL3 and METTL14, with the active site located in the methyltransferase domain (MTD) of METTL3.¹⁶⁻¹⁸ The DPPW motif near the methyl donor S-adenosylmethionine (SAM) binding site in this MTD was postulated to bind the target adenine during catalysis. Moreover, a positively charged groove in the METTL3-METTL14 interface, the C-terminal RGG domain in METTL14 and the zinc finger motifs in METTL3 were identified as important domains for RNA binding.^{19,20} However, to date there are no full-length or substrate-RNA-bound structures of the catalytic METTL3-METTL14 complex.

In addition, a set of accessory proteins assembles to the METTL3-METTL14 heterodimer to form the full MTC, mediated by WTAP that firmly binds to the N-terminal leader helix in METTL3.²⁰ WTAP was shown to locate the whole complex to the nuclear speckles and can modulate m^6A deposition to specific sites in the RNA.^{13,21,22} Moreover, WTAP acts as binding platform for other accessory proteins including VIRMA, RBM15, ZC3H13 and HAKAI that are mostly identified to mediate position specific methylation.²³⁻²⁶ For example, RBM15 was shown to mediate region-

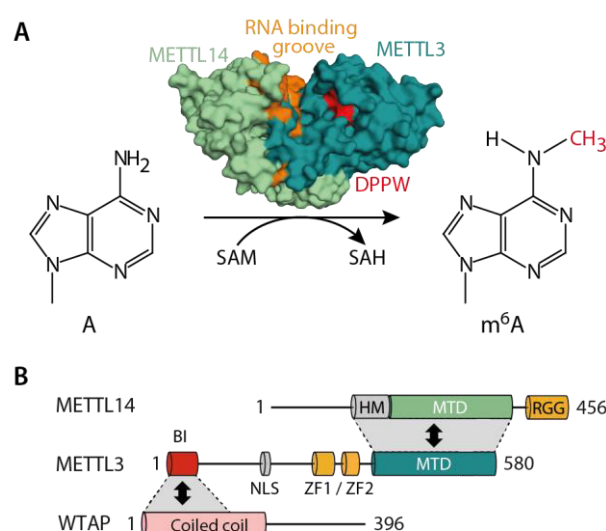


Figure 1: A m^6A formation by the METTL3/METTL14 complex (crystal structure of the MTDs, PDB 5IL1). **B** Schematic representation of the METTL3/METTL14/WTAP domains and interactions.

Summary

selective methylation in a WTAP dependent manner, directing specificity towards U-rich sequences.²⁴

The observed specificity of the methyltransferase complex to methylate only site specific DRACH sequenced is still poorly understood. Some possible modulators like the role of the accessory proteins are under investigation, however, the structural context of the RNA methylation sites or a structural preference of the complex have been mainly neglected so far. Moreover, the structural dynamics of this methylation process still remain elusive. This thesis contributes to the afore-mentioned aspects by analysis of the methylation process regarding RNA structure sensitivity with enzymatic activity assays and its dynamic nature by implementing a smFRET approach.

We hypothesized the target RNA secondary structure to be an additional important modulator of methylation efficiency, based on the RNA binding elements of the complex (positively charged binding groove, zinc finger domain, RGG domain) and the supposed target adenine binding in the active site. Here, we postulated the possibility for a flipped-out adenine to be of special relevance, which is closely related to the local stability of the target adenine containing structure. Moreover, efficient binding of the protein complex to the RNA should require the ability to anchor the RNA on both sides of the target sequence.

To investigate this hypothesis, five GGACU containing RNAs with different secondary structure were designed and used as substrates for enzymatic activity assays with the full-length METTL3/METTL14 complex. The resulting methylation yields showed significant differences between the respective secondary structures with the most efficient m⁶A formation in the unstructured RNA. For a duplex with rather loose base pairing in the target region, the methylation yield was significantly decreased to 75 % with regard to the unstructured RNA. In contrast, no methylation activity was detected on a target adenosine bound by tight Watson-Crick (WC) base pairing inside the center of a duplex with full complementarity. This proved the local stability of the target adenine containing RNA structure being anticorrelated to efficient catalysis (high stability = low methylation yield). Further comparison of the loose duplex to a hairpin structure with similar

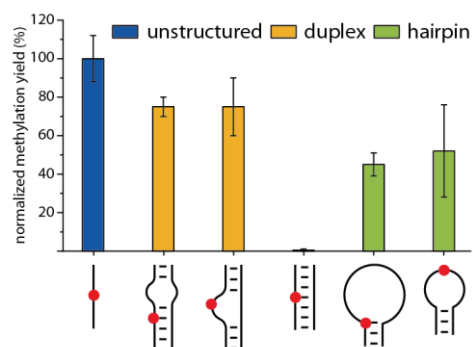


Figure 2: Activity assay results for differently structured RNAs. Local secondary structures are indicated (not the full-length RNA constructs).

local target stability, resulted in significantly lower methylation yields for the hairpin. This decreased methylation yield for a rather bended structure presumably confirms the requirement for effective protein binding on both sides of the target RNA sequence for efficient methylation.

Furthermore, we tested unstructured and hairpin RNAs with different RRACH sequence contexts. Changes to the predominant GGACU sequence resulted in reduced methylation yields, with the central GAC being an important factor for productive catalysis. For the investigated RNA constructs, the influence of the RRACH sequence context was observed to be more pronounced for target nucleotides in a hairpin structure. Together this demonstrates that methylation efficiency of the METTL3/METTL14 complex is highly dependent on the combination of both, target RNA secondary structure and methylation sequence.

As a next step, the role of several protein domains on RNA binding and catalysis was investigated, focusing on the RNA binding domains. Therefore, six truncated protein complexes were successfully expressed and purified. In METTL3, the complex was subsequently shortened from the N-terminal domain (NTD) to leave both zinc fingers (M3 Δ NTD), one zinc finger (M3 Δ ZF1) or only the MTD (M3MTD). The complexes shortened in METTL14 are deficient in the RGG domain (M14 Δ RGG), the NTD (M14 Δ NTD) or both (M14MTD). A minimal complex consisting of only the two MTDs was used as negative control (M3MTD/M14MTD). After purification, the enzymatic activity as well as the RNA binding properties of all truncated complexes were tested on an unstructured RNA (GGACU). The results show a slightly increased methylation efficiency for the M3 Δ NTD complex in comparison to the full-length complex. For all complexes deficient in any RNA binding domain a significantly decreased methylation yield was observed. This indicates that all of the RNA binding domains are necessary for efficient catalysis, which is mostly in line with the respective RNA binding properties determined. Moreover, depletion of the structurally important α -helical motif adjacent to the MTD in METTL14 leads to a decreased methylation activity together with a reduced RNA binding. Interestingly, the lowered methylation yield for the complex with only one of the two zinc

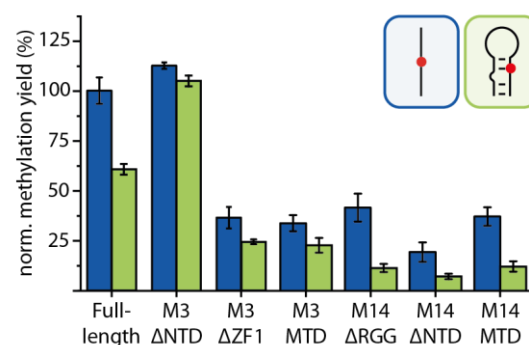


Figure 3: Activity assay results for the truncated protein complexes on two different RNAs (color coded). Local secondary structures of the RNAs are indicated (not full construct structure).




Summary

fingers left in METTL3 was observed together with a good RNA binding ability. Further RNA binding experiments including an off-target construct lead to the assumption of an unproductive RNA binding of this complex. Here, RNA binding might occur in the wrong orientation for catalysis. Moreover, binding to the off-target RNA was also shown for the full-length complex and complex M3ΔNTD, which hints towards a general RNA binding ability of the METTL3/METTL14 complex with mainly the active site exhibiting target sequence specificity.

All truncated protein complexes were additionally tested for their sensitivity towards RNA secondary structure with a hairpin structured target RNA. Most of the complexes exhibited the expected reduced methylation activity on the hairpin RNA. The relative decrease in methylation yield was more pronounced for the complexes lacking the RGG domain than for the ZFD deficient complexes. This hints towards a preferred RNA orientation in favor of the RGG domain during catalysis on this RNA construct. Again, complex M3ΔNTD exhibited a unique behavior with no significant reduction in methylation for the hairpin RNA. To further validate this observation, enzymatic activity assays with two additional structured RNAs (hairpin, loose duplex) were performed for this complex. The results confirmed the already observed decreased sensitivity towards RNA secondary structure when depleting the NTD in METTL3. This protein domain harbors the binding interface for WTAP that itself acts as binding platform for other accessory proteins.

In the following experiments, the sensitivity towards target RNA secondary structure was therefore investigated for the METTL3/METTL14/WTAP complex. First, this complex was expressed and purified with a different approach where only METTL14 carries a His-tag for purification. After successful purification of the complex, enzymatic activity assays were performed with differently structured RNAs. The methylation results on the unstructured RNA were similar with regard to the METTL3/METTL14 complex. For the structured RNAs, a more pronounced decrease in methylation efficiency was observed for the WTAP containing complex. This demonstrates an increased structural sensitivity for this complex in comparison to the METTL3/METTL14 core complex. The obtained results were referenced to the previous data for a better comparison with the M3ΔNTD complex. Comparison of the three complexes reinforced the special role of

Table 1: Methylation yields for different protein complexes on diverse RNA targets. (*referenced yields)

normalized methylation yield			
M3ΔNTD / METTL14	112 %	118 %	109 %
METTL3 / METTL14	100 %	75 %	45 %
METTL3 / METTL14 / WTAP	102 %	57 %*	28 %*

the NTD in METTL3 for the sensitivity towards target RNA secondary structure. Depletion of the NTD vastly diminished this sensitivity, opposed to providing the NTD binding partner WTAP where an increased sensitivity was observed. This might be related to differences in the ability to locally destabilize the target structure, the general RNA binding mode, the dynamics in the complex or the accessibility of the RNA binding domains.

In the second part of this project, the structural dynamics of the METTL3/METTL14 complex were investigated in different catalytic stages via a smFRET approach. At first, a total internal reflection fluorescence (TIRF) based microscope was built to perform the smFRET measurements. The excitation beam was constructed with three different laser sources (637 nm, 532 nm, 488 nm) that were parallelized to one single beam path. For every laser, a beam shutter was set directly after the laser source to provide fast on and off switching of the excitation light on the sample slide. The laser beam was brought onto an inverted light microscope and focused onto a prism to create total internal reflection on the underlying sample slide. From there, the water immersion objective collects the sample fluorescence and passes it to the detection path. Here, the emitted light is filtered to block excitation light and separated into two channels (wavelength above/below 642 nm). The channels are then focused on separate areas of the EMCCD camera that is used for detection. The whole setup can be controlled remotely by individual software tools for the three main components (laser, shutter, camera). To gain a significantly faster and improved control over these components, a custom remote-control software was written. With this, all required settings can be applied in one software. Most importantly, excitation and acquisition can be coupled if desired, which opens up a great variety of applications, including donor-only filtering.

For smFRET measurements, the target RNA was fluorescently labeled and immobilized by hybridization with DNA handles. Here, three RNA constructs were prepared that differ only in the center of the target sequence: one with the substrate nucleotide (GGACU), one with the product nucleotide (GGm⁶ACU) and one with an abasic site (GGXCU). Upon protein binding and catalysis, the distance deviations of the dyes attached to the RNA are captured with changes in FRET efficiency. Catalytic activity of the protein complexes with and without WTAP was proven on the substrate RNA. First, the smFRET data was acquired, processed into histograms and the results for the different ribonucleoproteins (RNPs) consisting of RNA (substrate/product/abasic),

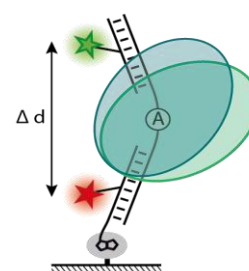


Figure 4: Schematic representation of the smFRET construct.

Summary

protein complex (METTL3/METTL14 w/o WTAP) and cofactor (SAM/SAH) were compared. The center of the protein bound RNA state for the catalytically competent RNPs (substrate RNA, protein complex and SAM) was shifted towards higher FRET efficiencies compared to the post-catalytic RNPs (product RNA, protein complex and SAH). This indicates an enhanced contraction of the RNA conformation for the catalytically competent RNPs. Further on, these differences were investigated with the analysis of time traces for their dynamic behavior. The single molecules were individually fitted by hidden Markov modelling (resulting in 4 states at $E_{\text{FRET}}=0.38, 0.48, 0.56$ and 0.66) and subject to an extensive summarized analysis. The dynamic molecules were identified and the population distribution as well as the transitions between the identified states were determined. Moreover, transition rates were calculated. This analysis revealed robust differences between the catalytically competent, post-catalytic and catalytically incompetent (abasic) RNPs. The population of the highest FRET state ($E_{\text{FRET}}=0.66$) was significantly increased for the catalytically competent RNPs in relation to the post-catalytic RNPs. This renders this state as relevant for catalysis and indicates a contraction of the RNA during methylation. Comparison to the abasic RNPs uncovered this difference originating from the presence of substrate or product nucleotide rather than SAM or SAH. These differences were also mirrored in the population of the corresponding transitions. The same behavior was observed for the respective WTAP containing complexes, however, with a general increase of the overall dynamics and the population of the FRET state at $E_{\text{FRET}}=0.66$.

Based on the results presented in this thesis, a hypothesis for the catalytic cycle was postulated. The observed catalytically relevant FRET state ($E_{\text{FRET}}=0.66$) is associated with a contraction of the RNA conformation with respect to the free RNA. RNA binding in the METTL3/METTL14 complex was shown to be dependent on the presence of the RGG and zinc finger domains, adjacent to the positively charged binding groove.^{17–20} Moreover, the here reported sensitivity of the protein complex towards RNA secondary structure showed the importance for effective protein anchoring on both sides of the target RNA. This is most likely related to an RNA binding mode

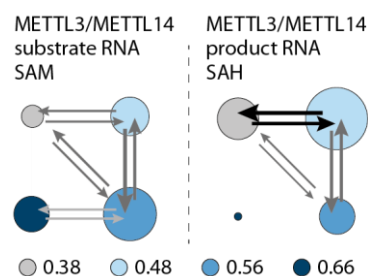


Figure 5: Populations of FRET state (circle sizes) and transitions for the dynamic molecules.

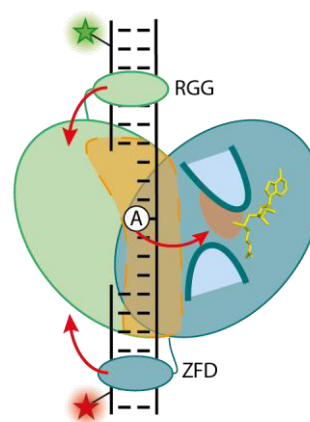


Figure 6: Hypothesized model for the RNA motion during catalysis.

through the RGG on one side and the ZFD on the other side of the target sequence. Taken together, this hints towards a possible "squeezing" motion of the two RNA binding protein domains to contract the bound RNA during the methylation process. Upon contraction, the substrate adenosine might be flipped out of the RNA structure towards the active site and is presumably bound by the DPPW motif and the two active site loops¹⁶⁻¹⁸ for catalysis. After methylation, the formed m⁶A most likely exhibits a lower affinity towards the active site and is therefore flipped back inside the RNA structure. Rebinding of the product nucleotide in the active site is probably sterically hindered.

Taken together, this thesis gives new insights into RNA methylation by the METTL3/METTL14 complex. The methylation process was uncovered to be sensitive towards the secondary structure of the target RNAs with special importance of local structural stability and protein binding on both sides of the target adenosine. This sensitivity was shown to be closely related to the presence of the NTD in METTL3 and to be more pronounced for the WTAP containing complex. Moreover, all RNA binding domains of the complex were shown to be indispensable for efficient methylation. The results indicated general RNA binding by the RGG and zinc finger domains and the sequence specificity presumably originating from the active site. SmFRET measurements with the self-built TIRF microscope hinted towards a contraction of the bound RNA in catalytically active states. Together with the previous results, this allowed for setting up a hypothesis for methylation catalysis. This hypothesis includes a possible squeezing mechanism of the bound RGG and zinc finger domains to contract the RNA and with this presumably flipping out the target adenosine into the active site for catalysis.

1 Introduction

1.1 RNA modifications

Secondary structure and function of all RNA species is modulated by posttranscriptional modifications where to date over 150 naturally occurring modified nucleosides are known.^{27–31} The types of modification span a large variety of chemical changes, including addition of small or larger chemical groups to ribose or base (methyl groups, sugars, amino acids, ...), base isomerizations or substitutions of oxygen with heavy atoms (Figure 7). This extensive toolbox gets even larger through modification cascades where combinations of different modifications are formed, leading to sometimes highly complex modification patterns. The implications and biological functions of the majority of these modifications is, however, still unknown.

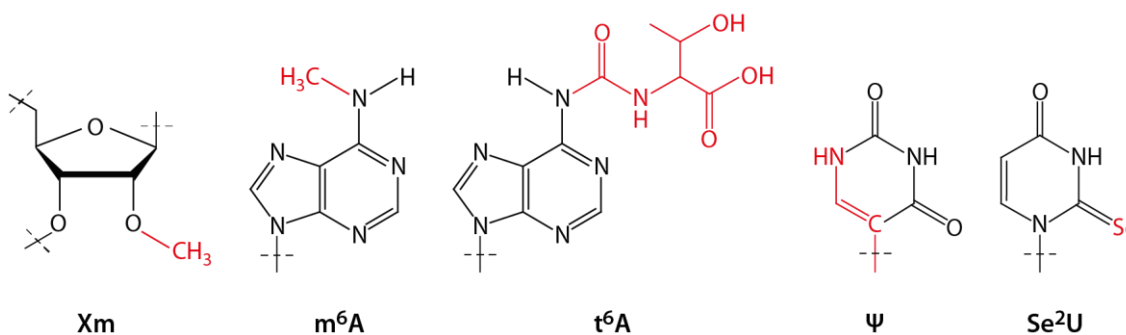


Figure 7: Representative natural RNA modifications for ribose methylation (Xm), base methylation (N6-methyladenosine (m^6A)), addition of amino acids (N6-threonylcarbamoyladenine (t^6A)), isomerizations (pseudouridine (Ψ)) and insertion of heavy atoms (2-selenouridine (Se^2U)). Modifications are colored in red.

Some of the modifications are specific to certain species while others are prominent in all kingdoms of life: in eukarya, bacteria and archaea (e.g. N6-methyladenosine or pseudouridine).³² The first modification that was discovered in RNA is also the most prevalent one: pseudouridine.³³ Pseudouridine is an isomer of uridine, where the base is flipped 180° forming a C-C glycosidic bond and an additional NH group in the ring. In general, modifications can influence RNA secondary structure stability through weakening or strengthening of specific interactions. Pseudouridine for example has stabilizing effects on RNA structure as the newly formed NH group acts as additional hydrogen donor on the Hoogsteen (HG) base pairing site (Figure 8).^{33,34} In contrast, destabilizing effects on Watson-Crick (WC) base pairing can be observed upon formation of 1-methylguanosine. Insertion of the methyl group into the WC site

Introduction

of the base sterically disrupts base pair formation.³⁵ These different effects of modifications on base interactions are particularly important for highly structured types of RNA, the most prominent example being tRNAs. tRNAs are the most often and diversely modified RNAs and play a key role in protein biosynthesis. Their “L-shaped” 3D structure is highly sensitive towards small changes in secondary structure interactions and the different modification sites are needed to finetune and ensure the proper folding.^{32,36} Along the same line, the “tip” of the structure, responsible for codon-anticodon interactions, is oftentimes modified at position 34 and/or 37 to prevent frame shifting.^{36–38}

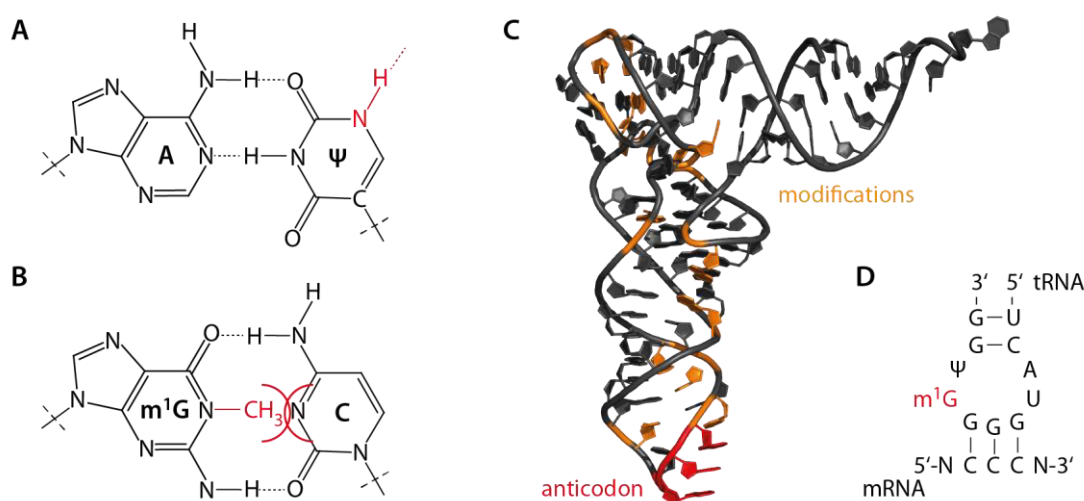


Figure 8: **A** Pseudouridine has an additional H-donor in the HG site and stabilizes RNA. **B** 1-methylguanosine sterically blocks WC base pairing. **C** The “L-shaped” 3D structure of tRNAs (PDB 1EHZ) is ensured by multiple modifications. **D** Codon-anticodon recognition is prevented from frame shifting by 1-methylguanosine formation at position 37 in the tRNA.

In mRNAs, the base pairing partner of the anticodon loop in tRNAs, the most abundant modification is N6-methyladenosine (m^6A). It is formed through methyl transfer from the methyl donor S-adenosylmethionine (SAM) to the N6-position in the adenosine, mediated by N6-methyltransferases (Figure 9). Details to this specific modification are described in the following chapters.

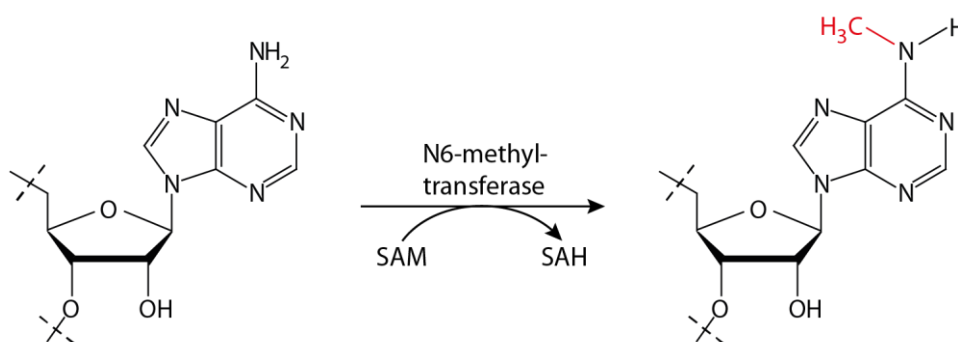


Figure 9: Formation of N6-methyladenosine by methyl transfer from S-adenosylmethionine to adenosine, enzymatically catalyzed by an N6-methyltransferase.

1.1.1 N6-methyladenosine

The first identification of m^6A was achieved in polyadenylated parts of the mRNA in 1974.^{1,2,39} However, further investigation of this modification remained inaccessible until the first m^6A mapping techniques were developed in 2012.^{3,5} Since then, m^6A was proven to be present and to be relevant in RNAs in a variety of organisms, including viruses, bacteria, yeast, archaea, plants and vertebrates.⁸ In eukaryotic RNA, m^6A is most abundant in mRNA but also plays a role in a variety of other RNAs like long non-coding RNAs (lncRNAs).

Some examples of m^6A functions in bacteria and viruses have uncovered its importance for related human diseases. The first example becomes evident in the antibiotic resistance of bacteria.⁴⁰ Proteins of the erythromycin ribosome methylase (ERM) family are responsible for the single or double methylation at the N6 position of adenosine A2058 in *E.coli*. Being located in the 23S rRNA of the ribosome, this (double) methylated adenosine sterically blocks binding of macrolide, lincosamide and streptogramin B (MLS_B) antibiotics which confers resistance to these antibiotics. Moreover, recent studies report on the influence of m^6A formation in the SARS-CoV-2 genome on viral replication.⁴¹⁻⁴³ Methylation by host enzymes in different parts of the SARS-CoV-2 RNA was shown to help the virus to bypass the immune system by blocking RIG-I binding and thereby the activation of downstream immune pathways for inflammatory gene expression.^{42,43} Moreover, N6-methylation at position A74 in the 5'UTR of SARS-CoV-2 was shown to increase translation initiation presumably through secondary structure destabilizing effects and with that enhanced accessibility for the ribosomal pre-initiation complex.⁴¹

1.1.2 The role of N6-methyladenosine in mRNA

The most abundant modification in eucaryotic mRNA is m⁶A. Its widespread distribution was revealed to show modification hotspots in long internal exons, 3'UTRs and around stop codons.^{3,5} The first successful steps towards specification of the m⁶A forming methyltransferase complex (MTC) was achieved by the identification and purification of METTL3, one of the two catalytic core proteins.^{9,10} Since then, only the discovery of m⁶A demethylating enzymes (FTO⁴⁴, ALKBH5⁴⁵) reinforced the interest in this modification, almost 40 years after its discovery in 1974. The proven reversibility and with that dynamic nature of m⁶A, together with its specific distribution, gave a strong indication on regulatory effects of m⁶A and paved the way for numerous studies on all aspects of the m⁶A cycle. The methylation cycle for mRNAs (Figure 10) is commonly described with the separation in:

1. The methylation process enzymatically catalyzed by the “m⁶A writer” protein complex transferring a methyl group from the cofactor S-adenosylmethionine to the N6 position of the adenosine
2. Functions of m⁶A by altering RNA secondary structure or mediated by “m⁶A reader” proteins that recognize and bind the methylated site
3. The reformation of adenosine by demethylation through “m⁶A erasers”

Since the discovery of METTL3, more proteins of the MTC were identified. Besides the two catalytic core proteins METTL3 and METTL14¹³, WTAP²¹ and several assembling proteins (VIRMA, ZC3H13, RBM15, HAKAI)^{22,23,25,26} were shown to be part of the complex (for details see chapter 1.2). This MTC was observed to target mainly DRACH motif sequences as substrate for methylation with the central adenine being modified.^{3,5} However, as only a subset of all DRACH sequences is methylated there have to be additional layers of target control. These elements can be intrinsic (complex specific) influences like the preference for DRACH sequences as well as the surrounding local structure.^{22,46} Extrinsic factors can be other proteins guiding the methylation complex to a specific RNA site, as it was shown for some transcription factors, solely the influence of the transcription rate.^{15,47-49} However, the main modulator of m⁶A deposition was just recently identified to be the physical exclusion from specific sites through the exon junction complex.⁵⁰

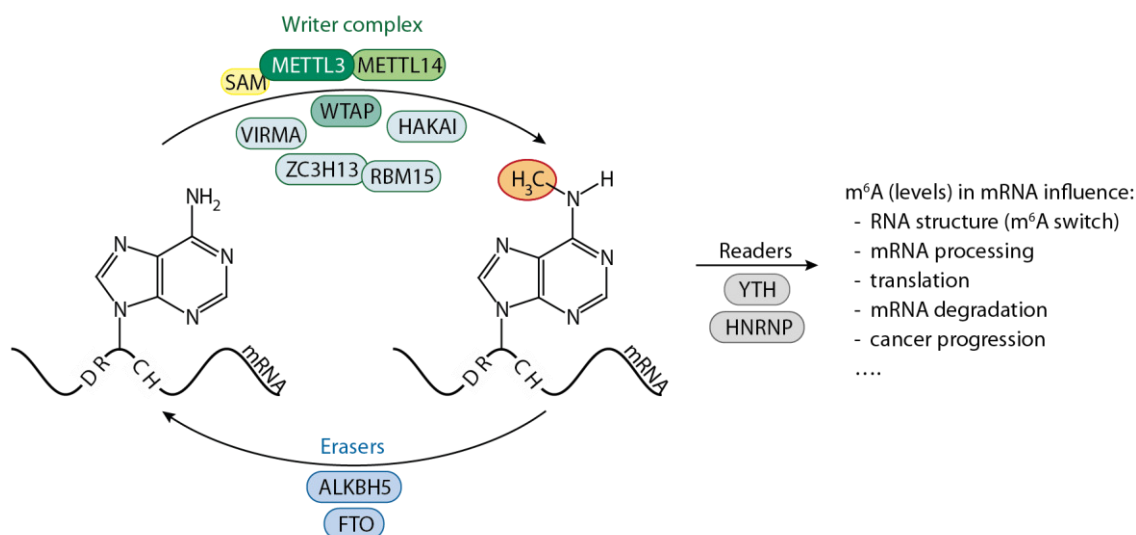


Figure 10: The methylation cycle for mRNAs is induced by the m⁶A writer complex transferring a methyl group from S-adenosylmethionine to the RNA. m⁶A based functions are mediated by reader proteins. Demethylation catalyzed by m⁶A erasers restores the unmethylated adenosine.^{based on 8}

The biological functions of the m⁶A modification to selectively regulate specific transcripts can be based on modification induced RNA secondary structure changes or mediated by m⁶A reader proteins. m⁶A interplays with many processes in the cell depending on its location in the transcript and the cellular context, influencing the whole life cycle of an mRNA.⁵¹ Deposition of m⁶A on the lncRNA XIST even influences gene silencing on the X chromosome and thereby helps to prevent production of various transcripts.²⁴ Most of the reader proteins known to date are from the YTH and HNRNP protein families. Binding of the reader protein YTHDC1 affects pre-mRNA processing⁵² and mRNA nuclear export.⁵³ In the cytosol, YTHDF1, YTHDF3⁵⁴ and IGFBP1, IGFBP2 and IGFBP3⁵⁵ can enhance translation by binding different parts of the transcriptome. While IGFBP1 to 3 were also shown to stabilize mRNA, binding of YTHDF2 is responsible for mRNA degradation.⁵⁶

Besides the reader protein mediated functions of m⁶A, there are also functions related to changes in RNA secondary structure induced by the modification. Box C/D small nucleolar ribonucleoproteins (snoRNPs) are responsible for the site-specific 2'-O methylation at the ribose ring in several rRNAs and tRNAs in eukaryotes and archaea.⁵⁷⁻⁵⁹ For this, part of the snoRNA is forming a kink turn motif upon binding of the 15.5k protein, which is required as a first step prior to full assembly of the ribonucleoprotein (RNP) complex (Figure 11A). This kink turn folding involves formation of a guanosine-adenosine trans sugar HG base pair above a bulge. For snoRNAs with a GAC sequence context, the adenine above the bulge can be methylated and

thereby prevents the HG base pair formation. This blocks the required kink turn formation and with that the binding of the 15.5k protein to form a functional Box C/D snoRNP complex. Another example for m⁶A related structural changes was termed as “m⁶A switch” and modulates the accessibility of single stranded RNA binding motifs.⁶⁰ Insertion of the methyl group into the mRNA weakens its secondary structure by destabilizing duplexes and therefore enhances binding of RNA binding proteins like HNRNPC (Figure 11B). Binding of HNRNPC controls abundance and alternative splicing of mRNAs, indicating the importance of these m⁶A regulated mechanisms for gene expression and mRNA processing.

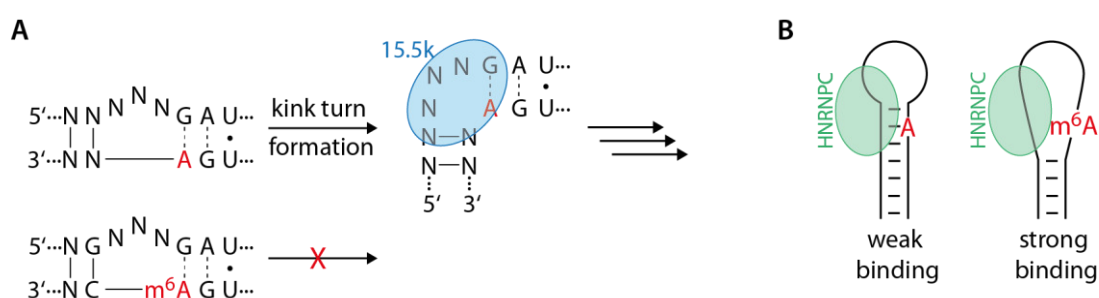


Figure 11: Examples for m⁶A functions related to changes in RNA secondary structure. **A** In box C/D snoRNPs, binding of the 15.5k protein induces kink turn folding with formation of a G-A Hoogsteen base pair. Methylation of the involved adenosine blocks this structure folding and with that RNP formation. **B** The m⁶A switch modulates the secondary structure of the mRNA by destabilizing the duplex and releasing the single stranded binding platform for RNA binding proteins like HNRNPC for enhanced binding.

With all the various functions of m⁶A in several biological processes known to date, it is not surprising that m⁶A is highly relevant in human diseases. It was shown to influence tumor progression in different types of cancer like hepatocellular carcinoma,⁶¹ leukemia⁶² and breast cancer⁶³ to name only a few. This raises great interest in m⁶A and especially its depositing enzymes as a potential therapeutic target. In November 2022 the first inhibitor of RNA methyltransferases in humans entered clinical trials (STORM therapeutics).⁶⁴ The small-molecule inhibitor STC-15, developed for treatment of acute myeloid leukemia, blocks the active site of METTL3, responsible for catalytic activity of the MTC.⁶⁵

1.1.3 Identification of N6-methyladenosine sites

The first identification of m⁶A in mammalian RNAs was based on incorporation of radioactive [methyl-³H]-methionine into the investigated cells.^{1,39} Upon cell growth, [methyl-³H]-methionine is metabolized to SAM and with that the radioactive methyl group is transferred to the RNA in all SAM related RNA methylations. After purification of polyadenylated RNA and digestion to nucleosides, the samples were divided into 2'-O-methylnucleosides and base-methylnucleosides via DEAE-cellulose (borate) chromatography. The separated fractions were then analyzed with high-speed liquid chromatography (Aminex A-5), where m⁶A was shown to be the most prominent modification in the base-methylnucleoside fraction.¹ Although identifying m⁶A to be present in these RNAs, location and distribution of the modification remained inaccessible with this technique.

A site-specific approach to identify m⁶A modification in RNA is based on site-specific cleavage and radioactive-labelling followed by ligation-assisted extraction and thin-layer chromatography (SCARLET).⁶⁶ The major advantage of this technique is the site-specific RNA analysis, the high confidence in stoichiometry as well as its applicability to other RNA modifications. However, SCARLET is not suitable for a transcriptome wide analysis.⁶⁷

State of the art, transcriptome wide analysis techniques are based on the development of m⁶A specific antibodies.⁶⁸ The first approaches built on methylated RNA immunoprecipitation followed by RNA sequencing (MeRIP-Seq) (Figure 12A).^{3,5} For this, the RNA is fragmented (approx. size: 100 nucleotides) and immunoprecipitated with m⁶A specific antibodies. Then, the isolated RNA fragments are subject to RNA sequencing and the results are mapped onto the genome. Comparison with sequencing results from the initial pool of RNA fragments yields “m⁶A peaks” around m⁶A sites. To increase robustness, pulldown with at least two different antibodies is performed and only m⁶A peaks present in both evaluations are further analyzed.⁶⁹ However, identification of the specific m⁶A site in high resolution is not possible since the m⁶A peaks are 200 nucleotides wide and rarely symmetric, due to biased RNA fragmentation and m⁶A sites in closer proximity.⁶⁷

An improved transcriptome wide RNA analysis with individual nucleotide-resolution was achieved with the miCLIP (m⁶A individual-nucleotide-resolution cross-linking and immunoprecipitation) method (Figure 12B).⁷⁰ The first steps are similar to meRIP-Seq, namely fragmentation and immunoprecipitation of the RNA with m⁶A specific antibodies. Then, the antibody-bound RNA is analyzed with the iCLIP protocol⁷¹: RNA and antibody are UV crosslinked in the first step, treated with proteinase K and then subject to reverse transcription. During

reverse transcription, antibody-recognized m⁶A sites lead to truncations in the produced cDNA. In the last step, a cDNA library is prepared and sequenced. Analysis of the sequencing results leads to a precise identification of the m⁶A site with individual nucleotide resolution. A recent improvement for the experimental protocol (miCLIP2) includes the reduction of input material amounts required and a developed machine-learning based analysis tool (m6Aboost).⁷²

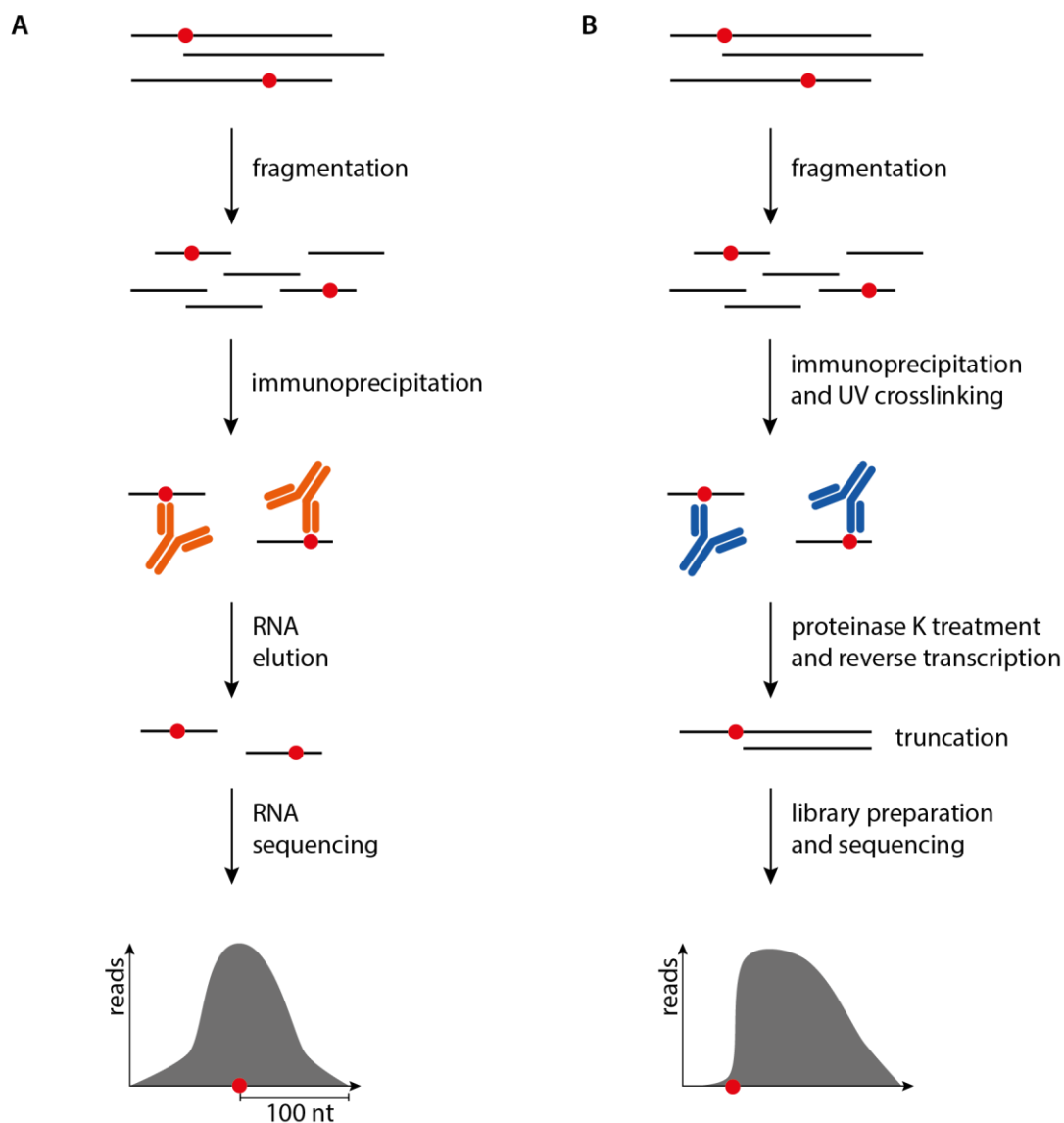


Figure 12: **A** Schematic representation of the MeRIP-Seq protocol. After RNA digestion to fragments of approx. 100 nt, immunoprecipitation with m⁶A specific antibodies and RNA elution, sequencing of RNAs results in a broad m⁶A peak. **B** Schematic representation of the miCLIP method. The RNA is fragmented, immunoprecipitated with m⁶A specific antibodies and UV crosslinked to the antibody. After proteinase K treatment and reverse transcription, a cDNA library is prepared and sequenced, yielding m⁶A sites with individual-nucleotide-resolution.

Taken together, transcriptome wide mapping techniques revealed the non-uniform distribution of m⁶A within the transcriptome and between individual cell types. In general, m⁶A levels are enriched in liver, kidney and brain cells. In the transcriptome, methylation hotspots were shown to be in the 3'UTR, around stop codons and in long internal exons.^{3,5,70} The DRACH sequence motif was uncovered to be the most commonly methylated sequence with GGACU as the prime methylation target, independently supported by biochemical studies on the MTC.^{11,13}

1.2 Human N6-methyladenosine writer complex

The human N6-methyladenosine writer complex is a multicomponent methyltransferase that methylates adenosines in a DRACH sequence context to form m⁶A on the central adenosine.^{11,13} This process occurs in the nucleus, most likely cotranscriptionally.^{12,14,15,73} The catalytic core of the multicomponent protein complex is comprised of METTL3 and METTL14, which form a tight heterodimeric complex.¹³ These two proteins are sufficient for efficient methylation *in vitro*, while other proteins of the complex are required additionally for productive methylation *in vivo*.^{13,21,24,25} The individual roles of the accessory proteins are described in chapter 1.2.2. In summary, the METTL3/METTL14 heterodimer is bound by WTAP to be localized to the nuclear speckles.^{20,21} ZC3H13 mediates nuclear retention and bridges the interaction to RBM15, which guides towards U rich RNA sites.^{23,24} VIRMA provides methylation specificity in the 3'UTR and near stop codons²⁵, while the role of HAKAI in the methylation process is still unknown.²⁶

1.2.1 Catalytic core METTL3/METTL14

Already the first study on the human methyltransferase complex identified METTL3 as the catalytically active protein with the ability to bind the methyl donor SAM.⁹ The first crystallization studies on the METTL3/METTL14 complex further shed light on its more detailed structure.¹⁶⁻¹⁸ The overall architecture of the two proteins comprises a methyltransferase domain (MTD) in each of them, a tandem CCCH zinc finger domain (ZFD) in METTL3 and an RGG domain in METTL14. Significant catalytic activity was reported to require both proteins hinting towards a synergistic effect. However, all of the crystallized complexes show no catalytic activity covering only the methyltransferase domains of the two proteins¹⁶⁻¹⁸ or the isolated ZFD of METTL3.¹⁹

The two MTDs form a compact heterodimer through an extensive network of hydrogen bonds and hydrophobic contacts on the interface between the two proteins.¹⁶⁻¹⁸ Cocrystallization with SAM revealed the catalytically active site being located in METTL3, whereas METTL14 contains a degenerate active site. In METTL14, the potential active site is sterically hindered to harbor methyl donor or RNA target. Mutation experiments on the relevant amino acids revealed complete loss of enzymatic activity when targeting the METTL3 active site, while mutation on the potential active site in METTL14 showed no significant activity changes.¹⁶⁻¹⁸

The synergistic effect of the two proteins becomes evident when taking a closer look at the heterodimeric interface. A positively charged groove comprised of 4 to 6 positively charged

residues of each protein (METTL3: Arg465, Arg468, His474, His478 and METTL14: Arg245, Arg249, Arg254 and Arg255) is located alongside the interaction surface near the SAM binding site (Figure 13A).^{16–18} This evolutionary conserved groove is assigned to bind the negatively charged RNA backbone and with that locate the target nucleotide near the active site and explains the requirement of METTL14 for target RNA binding and catalytic activity.^{16,18} This is supported by mutational analysis where replacement of positively charged residues of METTL14 within this groove (Arg254A, Arg255A) diminished the enzymatic activity of the complex to background levels.¹⁶ Moreover, the overall electrostatic potential on the “front side” with the active site in METTL3 is rather positively charged whereas the “back side” with the possible METTL14 active site shows negative electrostatic potential.^{16,18} This adds to the positively charged groove, guiding the target RNA to the active site in METTL3.

The active site in METTL3 binds the methyl donor SAM with several hydrogen bonds and the adenine moiety located in a hydrophobic pocket.^{16–18} The SAM molecule is surrounded by two loops (active site loop1 (ASL1) and 2(ASL2)) with the reactive methyl group located directly adjacent to a DPPW motif (METTL3 Asp395–Trp398) in ASL1 (Figure 13C). This DPPW motif was shown to be characteristic for N6-methyltransferases.⁷⁴ It is thought to coordinate the target adenosine for correct positioning and indicating a nucleophilic attack based mechanism towards the target nitrogen. This also holds true for the bacterial DNA methyltransferases M.TaqI and EcoP15I with the latter showing similar structures as observed for METTL3/METTL14 heterodimer formation.^{16,18,75,76} A structure overlay of substrate DNA bound EcoP15I with METTL3 shows high similarity between the active sites. The overlay places the target adenosine inside the catalytic pocket, hinting towards a possible structure of the target adenosine bound state (Figure 13D).^{16,75} Similar positioning of the target adenosine between the two active site loops suggests hydrogen bonds to Asp395, Tyr406 and Pro396 and a rearrangement of ASL1 to form π - π -stacking interactions with Trp598 to bind the target adenosine in the active site.¹⁶

Upon comparison of the SAM bound, SAH bound and the cofactor free crystal structures, Wang *et al.* discovered a significant change in ASL1 and ASL2 (Figure 13D).¹⁸ For ASL1, opening of the binding pocket is visible only for the SAM bound state. For ASL2, closing of the binding pocket is triggered upon SAM or SAH binding. Structural changes in the loops around the active site were also reported for the DNA methyltransferase M.TaqI, suggesting an important role in target binding and catalysis.^{18,76}

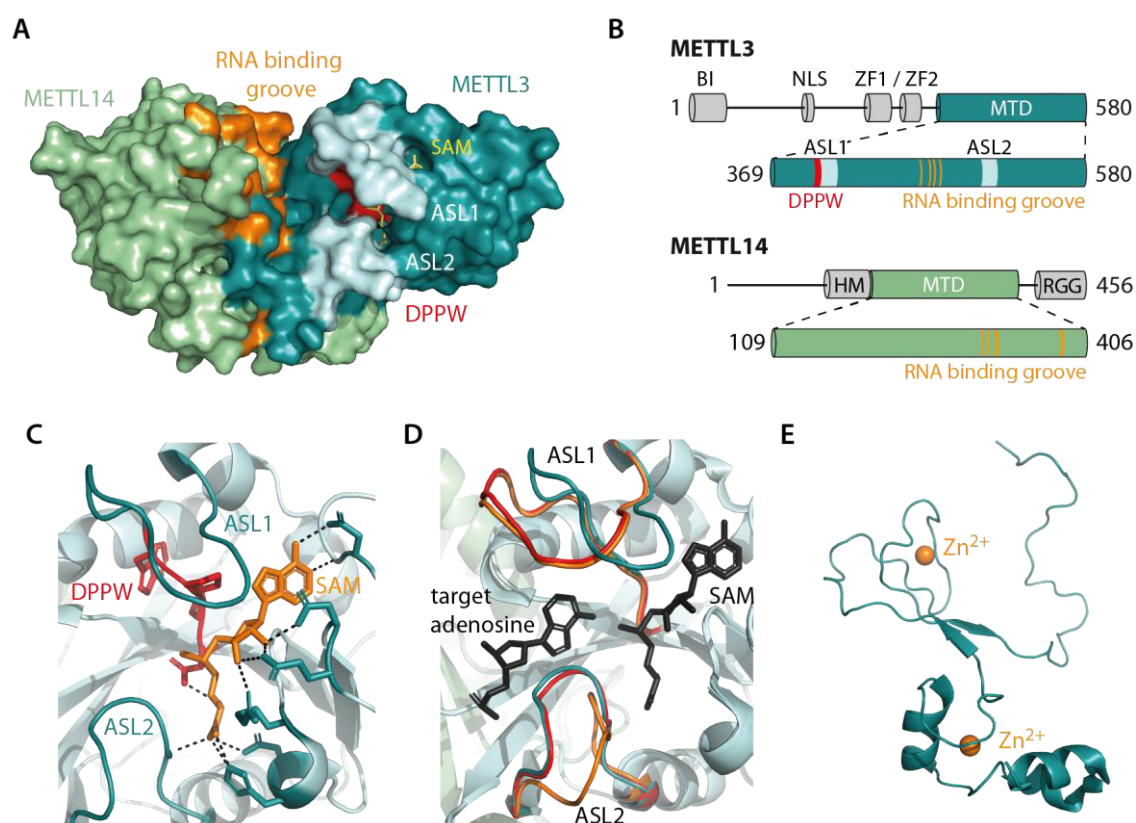


Figure 13: **A** Crystal structure of the heterodimeric complex consisting of METTL3 MTD (369-580) and METTL14 MTD (109-406) with the methyl donor SAM (PDB 5IL1). Functional parts of the complex are color coded. **B** Schematic representation of METTL3 and METTL14 with indicated functional parts. The color code corresponds to A. **C** Details of the hydrogen bonded SAM in the active site of METTL3. The donor methyl group is located next to the conserved DPPW motif in ASL1. **D** The overlay of METTL3 in the crystal structure with the active site of substrate bound EcoP15I (PDB 4ZCF) uncovers a possible target adenosine positioning in the active site. Comparison of the SAM bound (dark cyan, PDB 5IL1), SAH bound (red, PDB 5IL2) and cofactor free (orange, PDB 5IL0) crystal structures uncovers structural changes in ASL1 and ASL2. **E** Crystal structure of the isolated zinc finger domain in METTL3 (PDB 5YZ9).

Besides the positively charged binding groove on the heterodimeric interface, the zinc finger motif in METTL3 as well as the C-terminal RGG domain of METTL14 are important for target RNA binding and methyltransferase activity.^{16,17,19,20} Schöller *et al.* reported on the reduced RNA binding abilities of the RGG domain depleted complex together with low enzymatic activity.²⁰ Another group, however, showed the same activity levels for the full length complex compared to a complex with METTL14 only present as MTD without the RGG domain.¹⁹ While the RGG domain is thought to contribute to unspecific RNA binding, the ZFD was reported to introduce sequence specific binding abilities.¹⁹ However, this data was obtained with ITC measurements on the isolated ZFD showing RNA binding towards a GGACU containing RNA and no detectable

binding on a polyadenosine RNA construct. Although an enzymatically active form of the complex showed similar binding affinity towards the GGACU containing RNA, the measurement on the polyadenosine RNA was not repeated with this complex.

The cotranscriptional methylation process requires the methyltransferase complex to be imported from the cytosol into the nucleus. Schöller *et al.* identified a nuclear localization signal (NLS) in METTL3 and WTAP but not in METTL14, indicating that METTL3 and METTL14 are imported as a complex.²⁰ Mutation of the NLS in METTL3 retains METTL3 in the cytosol, and together with the unaffected ability of the mutant to bind to WTAP, this also indicates a WTAP independent nuclear import. WTAP has its own NLS and is imported into the nucleus independently from the heterodimeric complex. The contact between METTL3 and WTAP takes place after nuclear import and the interaction surface was mapped to be located in the N-terminal domain of both proteins.²⁰ A leader helix structure in the N-terminal domain of METTL3 acts as interaction platform to contact the coiled coil structure of WTAP (Figure 14A). However, the interaction between WTAP and METTL3 was shown to be weaker than between METTL3 and METTL14.²¹

1.2.2 Accessory proteins

WTAP was the first protein that was identified to be an additional part of the human N6-methyltransferase writer complex.²¹ It localizes the whole complex to the nuclear speckles as proven by immunofluorescence staining. Depletion of WTAP did not only significantly reduce the overall m⁶A levels, but also diminished the accumulation of the METTL3/METTL14 complex in the nuclear speckles. Knockdown of METTL3 or METTL14 in contrast did not alter the localization of WTAP to the nuclear speckles. Moreover, depletion of WTAP leads to a reduced binding of the core complex to specific RNAs.²¹ These findings are in line with a study reporting on WTAP dependent and independent m⁶A sites with the WTAP independent sites being the RNA cap structures.²²

Besides the interaction with the METTL3/METTL14 core complex, WTAP was shown to bind to a variety of other proteins where some of them were identified to be part of the methyltransferase complex.⁷⁷ One of these proteins is ZC3H13, a zinc finger domain containing protein.^{23,78} In fractionation assays, knockdown of ZC3H13 decreased the amount of all identified methyltransferase proteins in the nucleus, indicating the role on methyltransferase complex nuclear retention for ZC3H13.⁷⁸ Moreover, ZC3H13 bridges the interaction between

Introduction

WTAP and RBM15 in an RNA independent manner.²³ In *Drosophila*, the homologs of RBM15 and ZC3H13 were both uncovered to modulate alternative splicing of Sxl in an m⁶A dependent process.^{23,79} RBM15 was shown to bind the lncRNA XIST to mediate m⁶A dependent gene silencing on the X-chromosome.²⁴ In mRNA, RBM15 recruits the methyltransferase complex to U rich sequences near DRACH sites and knockdown of RBM15 results in decreased m⁶A levels.²⁴ Another accessory protein is VIRMA.²² VIRMA directs the methyltransferase complex to specific RNA sites in the 3'UTR and near stop codons.²⁵ VIRMA also associates with proteins involved in alternative polyadenylation in the 3'UTR.²⁵ The entanglement of this process with methylation becomes evident, when depleting METTL3 or VIRMA, where in both cases lengthening of the 3'UTR was observed. Interaction studies revealed the binding of VIRMA towards WTAP and a stabilizing effect on the WTAP/ZC3H13/HAKAI complex. HAKAI is an E3 ubiquitin-protein ligase that was also shown to bind to WTAP, however, its role in the methyltransferase complex is still unknown.²⁶ Parts of the structure of the WTAP/ZC3H13/VIRMA complex were recently solved by Su *et al.* with cryo electron microscopy (Figure 14B,C).⁸⁰ WTAP is present as dimer representing a saddle like shape on top of VIRMA which adopts a "horse shaped" conformation. The visible part of ZC3H13 is described as "barding" of the "horse", comparing the whole structure to a "war horse".⁸⁰ Attempts to solve a structure with METTL3/METTL14 present, gave hints to a possible anchoring point of METTL3 on WTAP.

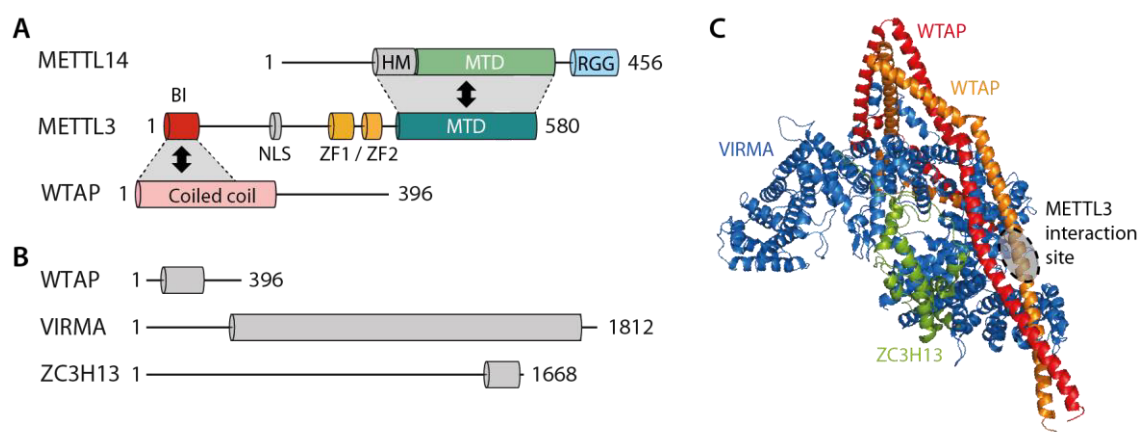


Figure 14: **A** Schematic representation of the interaction surfaces of the METTL3/METTL14/WTAP trimer. **B** Schematic representation of the protein domains from WTAP, VIRMA and ZC3H13 represented in the structure shown in C. **C** Cryo electron microscopy structure of the WTAP/VIRMA/ZC3H13 complex with indicated METTL3 interaction site (PDB 7VF2).

1.3 Principles of single molecule FRET spectroscopy

1.3.1 Förster resonance energy transfer

The first description of fluorescence was reported 1845 by Sir John Frederick William Herschel, who characterized the fluorescence of a quinine sulphate solution as "an extremely vivid and beautiful celestial blue color".⁸¹ Meanwhile, the phenomenon of fluorescence is defined as the emission of light from any substance that occurs from electronically excited singlet states.^{82,83}

The physical description of fluorescence is illustrated by the Jablonski diagram (Figure 15A). Every molecule has an electronic energy structure consisting of the electronic ground state (S_0) and several excited states (S_1, S_2, \dots). The electronic states are further divided into several vibrational levels (v). Without external energy sources except the thermal energy at room temperature, most molecules are mainly populated in the electronic and vibrational ground state ($S_0, v=0$). Excitation of an electron to higher electronic and vibrational states can occur by means of energy absorption. In the case of fluorescent dyes, this energy is usually provided by photons. The energy of the absorbed photon defines the energy level to which the electron is excited during this process. In the depicted Jablonski diagram the transition from the vibrational ground state of S_0 to the second excited vibrational level of S_1 is shown.^{82,83}

After excitation to a higher electronic and vibrational state, fast energy loss through internal conversion brings the electron to the next vibrational ground state (here $S_1, v=0$). From there,

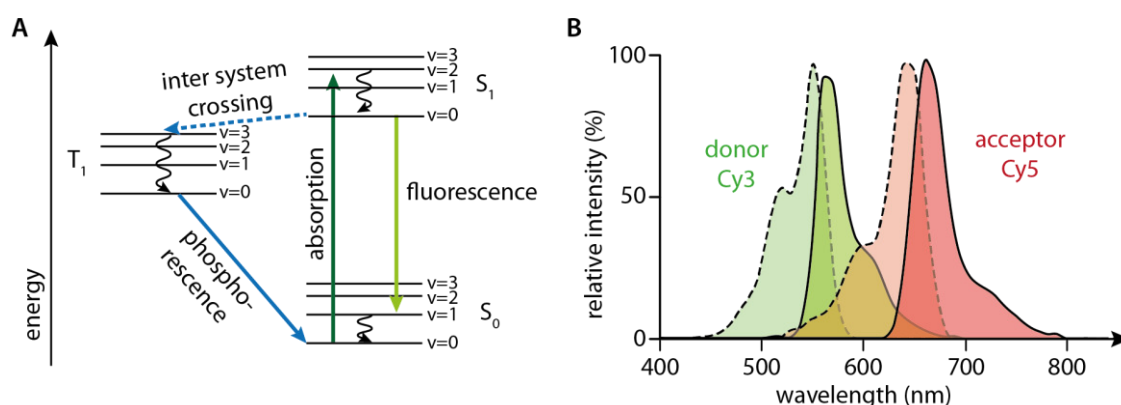


Figure 15: **A** Schematic Jablonski diagram of a fluorophore. The molecule absorbs a photon and transitions from the electronic ground state (S_0) to the excited state (S_1). After internal conversion to the vibrational ground state, the molecule relaxes back to S_0 through fluorescence. Alternatively, the molecule can relax back to S_0 over inter system crossing and phosphorescence. **B** Absorption (dashed lines) and emission (solid lines) spectra of the Cy3 (green) and Cy5 (red) fluorophore with spectral overlap (yellow).

the electron relaxes back to higher vibrational states of S_0 under emission of a photon, defined as fluorescence. Fluorescence occurs from the vibrational ground state as internal conversion takes place on the timescale of femtoseconds while fluorescence life times are usually several nanoseconds. This results in a redshift of the emitted light in relation to the absorbed light (Stokes shift). Besides internal conversion, additional effects (e.g., solvent effects) can contribute to the Stokes shift. The summarized spectral behavior of fluorophores is depicted in their absorption and emission spectra (Figure 15B).^{82,83}

Alternatively to fluorescence, relaxation back to the ground state can occur over inter system crossing (ISC) and phosphorescence (Figure 15A). In this case, the excited molecule undergoes spin conversion from the electronically excited singlet state S_1 to the triplet state T_1 with subsequent internal conversion to the vibrational ground state. Relaxation back to S_0 under emission of a photon is called phosphorescence and results in an additional spin conversion. As spin conversion is a forbidden transition, this pathway is usually very sparsely populated with phosphorescence lifetimes of at least milliseconds.^{82,83}

When bringing two fluorophores into close proximity, an additional way of relaxation for the excited molecule becomes evident. The Förster resonance energy transfer (FRET) describes the radiationless energy transfer from a donor molecule to an acceptor molecule (Figure 16A). In this process, the energy is transferred over dipole-dipole interaction and is dependent on the dye-to-dye distance. Typically, FRET based methods can span dye-to-dye distances that are between 2 and 10 nm.^{82,83}

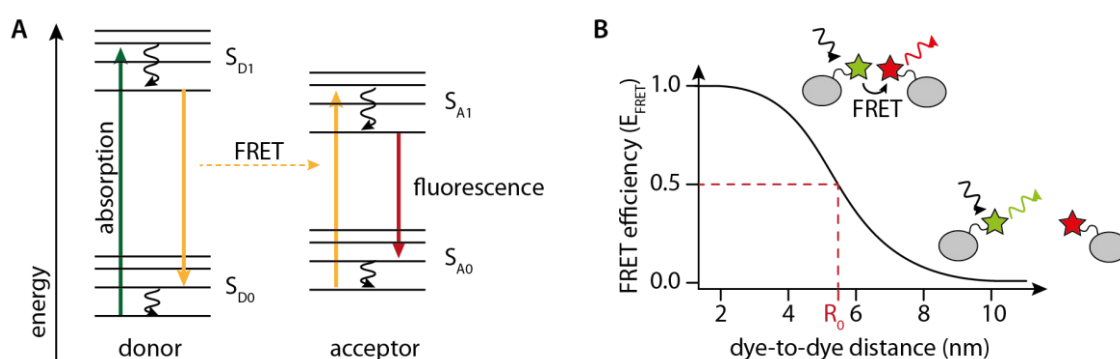


Figure 16: **A** Schematic Jablonski diagram for a typical donor acceptor FRET pair. After excitation of the donor molecule, the energy is transferred by Förster resonance energy transfer (FRET) to a suitable acceptor fluorophore nearby. The excited acceptor transitions back to the electronic ground state by fluorescence. **B** FRET efficiency is correlated to the dye-to-dye distance. The Förster radius (R_0) is defined as the dye-to-dye distance resulting in a FRET efficiency of 0.5, individually for each FRET pair.

The efficiency (E) of the energy transfer can be described as function of the dye-to-dye distance (r), dependent on the Förster radius (R₀) (Figure 16B):

$$E(r) = \frac{R_0^6}{R_0^6 + r^6} \quad (1)$$

R₀ defines the dye-to-dye distance for a specific donor-acceptor fluorophore pair, where the FRET efficiency is 0.5. It is dependent on the spectral overlap between the emission spectrum of the donor fluorophore and the absorption spectrum of the acceptor fluorophore, the relative dye-to-dye orientation, the donor quantum yield and the surrounding medium.^{82,83}

In FRET microscopy applications, the FRET efficiency (E) is usually calculated by comparing the relative intensities of the donor (I_D) and acceptor (I_A) fluorophores:

$$E = \frac{I_A}{I_D + I_A} \quad (2)$$

Due to its strong dependence on the dye-to-dye distance, FRET can be used as 'spectroscopic ruler'. While exact distance measurements can be quite challenging, relative distance changes of the two fluorophores are utilized in a variety of applications.⁸⁴ FRET microscopy is used especially for biomolecular applications as the covered distances are in the range of the size of biological macromolecules, like proteins and nucleic acids.^{82,83}

1.3.2 Single molecule FRET spectroscopy

One of the great advantages of fluorescence spectroscopy in contrast to many other spectroscopic methods is the opportunity for single molecule approaches. Conventional spectroscopic methods mainly yield the average of the investigated ensemble which oftentimes does not represent the complexity of the system. Single molecule approaches are able to capture the structural and dynamical heterogeneity of a system and with that uncover minor populated substates. The analysis of individual molecules also allows to follow complex processes like kinetic pathways with high sensitivity and under equilibrium conditions.⁸⁴⁻⁸⁶

The first single molecule FRET (smFRET) experiment was implemented by Ha *et al.* in 1996 investigating a short DNA molecule.⁸⁷ Since then, smFRET approaches were applied in various studies spanning a wide array of biological systems. Besides studies on functional RNAs like riboswitches^{88,89} and individual proteins like protein folding kinetics⁹⁰⁻⁹² (reviewed by Schuler

and Eaton)⁹³, the big and complex field of RNPs highly benefited from investigations based on smFRET approaches (reviewed in Meiser *et al.*)⁹⁴. The huge variety of examined RNPs spans from DNA related RNPs like the telomerase^{95–98}, transcription related proteins like polymerases^{99,100}, RNA processing RNPs like the spliceosome^{101–104} up to translation related machineries like the ribosome^{105–107}.

There are two main methodical ways to study FRET samples on a single molecule level, where the molecules to be analyzed are either diffusing freely or are immobilized on a surface.^{84–86} Samples with freely diffusing molecules have the advantage to display near physiological conditions and the possibility to determine diffusion constants. Illumination of these samples is done on a confocal based setup, where the excitation volume is kept to one femtoliter in order to illuminate only one molecule at a time. This has the advantage of working on a single molecule level, however, also brings the major drawback of this method: single molecules can only be followed for a very short time period closely related to the diffusion behavior of the molecule. Associated to this, molecule dynamics are hard to study when being slower than the average diffusion time. Immobilizing molecules on a surface overcomes this problem. It opens up the possibility to follow individual molecules for a longer period of time, which is mainly limited by the lifetime of the fluorophores (up to hours). Surface immobilized molecules are illuminated either by scanning confocal microscopy (SCM) or by total internal reflection fluorescence (TIRF) microscopy. SCM is similar to the above-mentioned confocal microscopy, however, uses raster scanning of the surface to assess multiple immobilized molecules. This technique needs longer measurement time to observe a significant number of molecules, yet usually achieves higher time resolution than TIRF microscopy. The latter illuminates a broader near surface area through creation of an evanescent field which allows to observe many single molecules simultaneously. In the following chapter this method is described in more detail.

1.3.3 Total internal reflection fluorescence microscopy

Total internal reflection fluorescence (TIRF) microscopy is applied for investigation of surface immobilized molecules. The excitation of the fluorophores through an evanescent field on the sample surface has several advantages. In contrast to confocal microscopy, only molecules close to the surface are excited. This reduces background levels significantly, which makes this technique the weapon of choice for most smFRET approaches.^{84,108}

The evanescent wave on the sample surface of a TIRF microscope is implemented by total internal reflection of the laser beam on the interface between the sample and the sample slide.

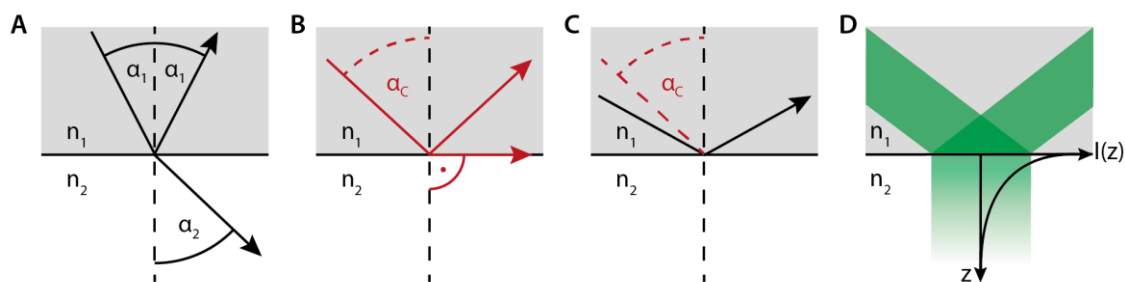


Figure 17: **A** Reflection and transmission of light on the interface between two media with different optical densities (n_1 , n_2). The angle of incidence (α_1) equals the angle of reflection (α_1), while the angle of transmission (α_2) is following Snell's law of refraction. **B** The critical angle for total reflection (α_c) is given when the angle of transmission (α_2) is equal to 90° . This is only possible for $n_1 > n_2$. **C** Total reflection is reached when the angle of incidence is higher than the critical angle for total reflection (α_c). **D** During total reflection, an evanescent field is intruding the optically thinner medium at the area of reflection. The intensity (I) is exponentially decaying with rising penetration depth (z).

When light passes the interface between two different optical media, the light beam is partially reflected and transmitted at the same time. In general, the angle of reflection is equal to the angle of incidence (α_1) with respect to the surface normal, independent of the two media.¹⁰⁹ However, the angle of transmission (α_2) is dependent on the optical density (n) of the two media on the interface (Figure 17A). This can be described with Snell's law of refraction:

$$n_1 \cdot \sin(\alpha_1) = n_2 \cdot \sin(\alpha_2) \quad (3)$$

Under specific conditions, the transmission of light can be eliminated which results in total reflection of the light beam. The boundary condition for total reflection is given with (Figure 17B):

$$\alpha_2 = 90^\circ \leftrightarrow \sin(\alpha_2) = 1 \quad (4)$$

With that, the critical angle for total reflection (α_c) results in:

$$\alpha_c = \sin^{-1} \left(\frac{n_2}{n_1} \right) \quad (5)$$

To reach rational results for equation (5), n_1 has to be bigger than n_2 , which results in the transition from a medium with higher to a medium with lower optical density (e.g. glass to water). Whenever the angle of incidence exceeds the critical angle α_c in the given conditions, total reflection occurs on this interface (Figure 17C). However, at the area of reflection the

reflected light enters the optically thinner medium on a microscopic scale ($\sim 100\text{-}200\text{ nm}$). The intensity of this evanescent field is exponentially decaying with growing penetration depth (z) (Figure 17D).¹⁰⁹

Experimentally, the creation of total internal reflection-based sample illumination can be achieved with two different setup constructions: the prism-type and the objective-type setup (Figure 18). Both setup types are built on the basis of an inverted microscope where the fluorescence is collected through the objective directly below the sample slide (Figure 18A).^{84,85,108} However, the evanescent field is formed in two different ways. For the objective-type setup, the laser beam is focused off axis on the back focal plane of the objective (Figure 18C). The objective hereby needs at least a numerical aperture (NA) of 1.45 to create total internal reflection on the sample slide, which is only applicable for oil immersion objectives. In the case of prism-type setups, a prism is placed on top of the sample slide and the excitation laser beam is focused on the sample slide through this prism. Here, typically water immersion objectives with long range focus are used.^{84,85,108}

When comparing the two setup types there are several advantages and disadvantages to consider.^{84,85,108} The different excitation path and the use of different objective types uncovers some of the major contrasts. The completely separated excitation and detection path for prism-

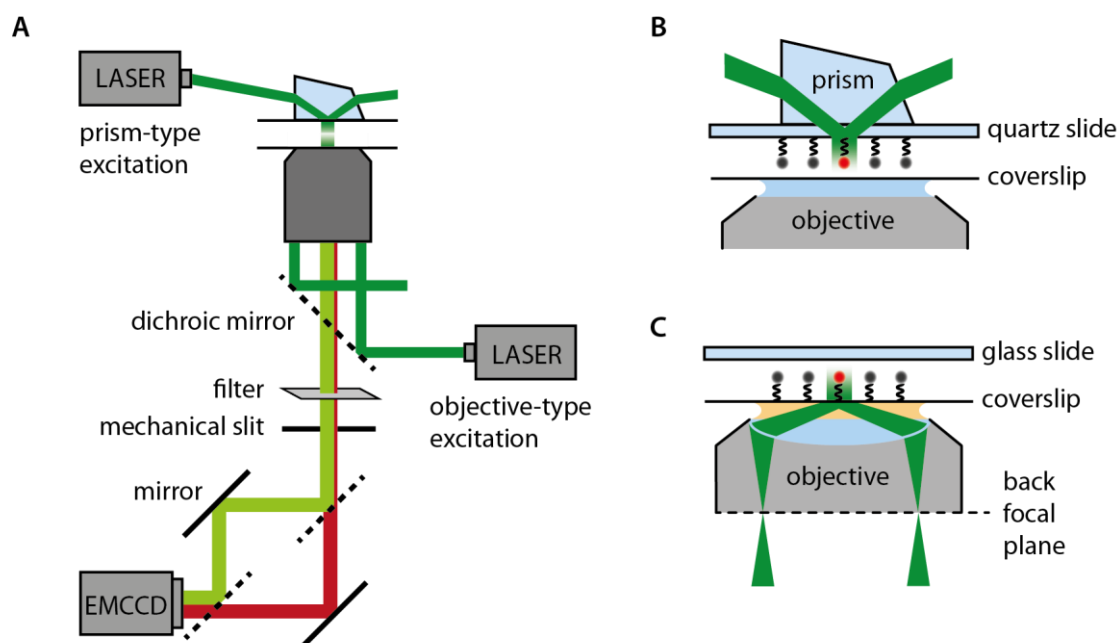


Figure 18: **A** Prism-type and objective-type TIRF setup with detection. **B** Details for prism-type TIRF: The laser is focused onto the slide through a prism. **C** Details for objective-type TIRF: The laser is focused off axis on the back focal plane of the objective.

type TIRF has the advantage of very low background noise through straylight from the excitation. Additionally, the used water immersion objectives are usually cheaper than high numerical aperture oil immersion objectives needed for objective-type TIRF. However, these oil immersion objectives have a higher photon collection efficiency and, although being quite expensive, newer objectives result in comparable background levels as in prism-type TIRF. An additional point to consider is the modular setup of prism-type TIRF. It has the major drawback of being harder to set up and align and small readjustments need to be done after every sample slide change. Nevertheless, the modular setup opens up a great flexibility in extending the setup by e.g. more excitation lasers and makes a wider range of incidence angles possible. This also includes a more flexible illumination by varying the laser spot size. Moreover, fluid injection experiments on immobilized molecules are only possible with a prism-type setup.⁸⁴

Detection of the resulting fluorescence is the same for both TIRF types.^{84–86,108,110} First, a suitable filter (long pass or band pass) is set to filter out remaining stray light from the excitation. After this, a mechanical slit is positioned to create sharp boundaries on both sides of the image. In the next step, donor and acceptor fluorescence are separated by a dichroic mirror and then focused onto different areas of an electron-multiplying charge-coupled device (EMCCD) camera chip. EMCCD cameras have a high quantum efficiency (85-95 %) between 450 and 700 nm. Moreover, the low readout and multiplying noise, high frame rates and readout speeds assign this camera type ideal for smFRET approaches. The more recently developed complementary metal-oxide-semiconductor (CMOS) based cameras are able to work at a higher frame rate, however, the EMCCD cameras are still superior in signal to noise ratio on a single molecule level.^{111,112}

1.3.4 TIRF sample preparation

The samples used for FRET applications in TIRF microscopy have two main requirements: a good FRET pair attached to reasonable sites on the biomolecule and immobilization of the latter. There are several dye properties that are needed for a good FRET pair.^{84,85} First of all, the dyes have to be photostable under the desired (especially buffer) conditions. Moreover, the brightness is of importance which means the extinction coefficient should be above $50000 \text{ M}^{-1}\text{cm}^{-1}$ and the quantum yield greater than 0.1. Additionally, a good FRET pair needs a high spectral overlap between donor emission and acceptor excitation and the quantum yield and detection efficiency on the camera is ideally the same for both fluorophores. However, donor

Introduction

and acceptor emission spectra have to be clearly separable for efficient detection. The fluorophores for biomolecular applications additionally require water solubility, a small size and bio-conjugation chemistry options. There are three main types of fluorophores: fluorescent proteins, quantum dots and organic dyes. Small organic dyes are the most useful type of fluorophores for biomolecular FRET applications due to their small size (smaller than 1 nm) and excellent photo physics. The most widely used small organic dye FRET pairs are from the cyanine, Alexa Fluor and ATTO dye families. To prevent the dyes from fast bleaching an oxygen scavenger system is implemented in the imaging buffer where two main systems are applicable. The older system relies on the enzymatic oxidation of glucose through glucose oxidase yielding hydrogen peroxide as side product. This toxic side product is then removed by reduction with the enzyme catalase to water and half of the starting oxygen concentration (Figure 19A).¹¹³ The newer and more stable oxygen scavenger system uses protocatechuate-3,4-dioxygenase (PCD) as enzyme to oxidize protocatechuic acid (PCA), thereby removing all the reacting oxygen (Figure 19B).^{114,115} Additionally, 6-hydroxy-2,5,7,8-tetramethylchromane-2-carboxylic acid

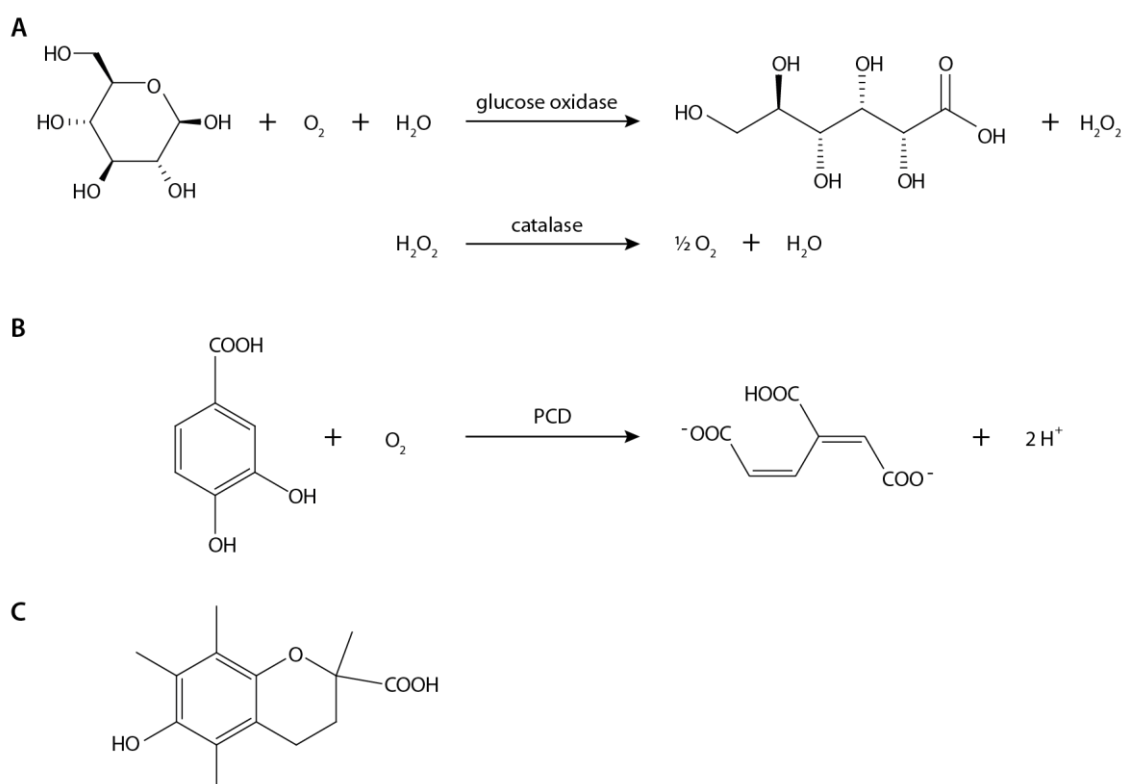


Figure 19: **A** Oxidation of glucose through glucose oxidase. Resulting hydrogen peroxide is enzymatically reduced to molecular oxygen and water. **B** Enzymatic oxidation of protocatechuic acid (PCA) through protocatechuate-3,4-dioxygenase (PCD). **C** Structural formula of 6-hydroxy-2,5,7,8-tetramethylchromane-2-carboxylic acid (trolox).

(trolox) is added to the imaging buffer to prevent the dyes from blinking due to triplet states (Figure 19C).¹¹⁶

The attachment sites of the FRET pair should be chosen to pass the Förster radius upon structural changes of the biomolecule for best sensitivity.¹¹⁷ However, attachment of the dyes itself must not yield structural or functional changes of the biomolecule. This is not only important for the product of the dye attachment but also for the whole attachment process (e.g. buffer/reaction conditions especially for proteins).¹¹⁷⁻¹¹⁹ Moreover, the photo physics of the dyes may not be altered which also includes the free rotation of the attached dye. Bio-conjugation of the dyes in general can be divided into two steps: first, introduction of biorthogonal functional groups into the biomolecule and second coupling of the desired dye. Looking at RNPs, both, the RNA and the protein can be subject to desired fluorescent labeling.¹²⁰ For RNA there are solid phase synthesis or enzymatic (cotranscriptional as well as posttranscriptional) options to insert the functional groups for coupling (reviewed in Hanspach *et al.*).^{118,121} The variety of inserted functional groups also offer a diversity of subsequent coupling reactions. One prime example is the coupling of NHS ester dyes to inserted aliphatic amine groups of the RNA (Figure 20A and B). Insertion of two different fluorophores into one RNA molecule needs either insertion of two functionally orthogonal groups or ligation of two individual RNA strands to the full-length RNA. Another possibility to (transiently) label RNA with one or more dyes is by hybridization with fluorescently labelled DNA¹²², PNA¹²³ or molecular beacons¹²⁴. The toolbox for dye attachment on proteins was historically very limited. Cysteine residues were and are still used to couple thiol reactive dyes.¹²⁵ However, when there is more than one cysteine residue inside the protein this has the major drawback of either unspecific labeling or the need of mutating all undesired cysteines. Moreover, position specific labeling of one or more interesting sites is mostly not applicable using only existing cysteines. Newer techniques rely on the incorporation of a set of unnatural amino acids (reviewed in Müller *et al.*).^{99,126} This enables the insertion of two or more functionally orthogonal groups that can be easily coupled to fluorophores under protein friendly conditions.

Immobilization of the dye coupled biomolecule is essential for TIRF microscopy. For this, additional functional groups for immobilization need to be attached to the biomolecule without disturbing structure or function. One of the most applied immobilization techniques relies on the tight interaction between biotin and streptavidin.^{84,127} Here, biotin is attached to the biomolecule and the slide surface is biotinylated as well. The addition of streptavidin to the surface enables tight binding of the biomolecule to the surface. Biotinylation of the glass surface

Introduction

can be done by adsorption of BSA/BSA-biotin. For protein samples, the passivation of the glass surface with aminosilane and subsequent coupling of NHS ester functionalized PEG/PEG-biotin is used (Figure 20C).^{84,128}

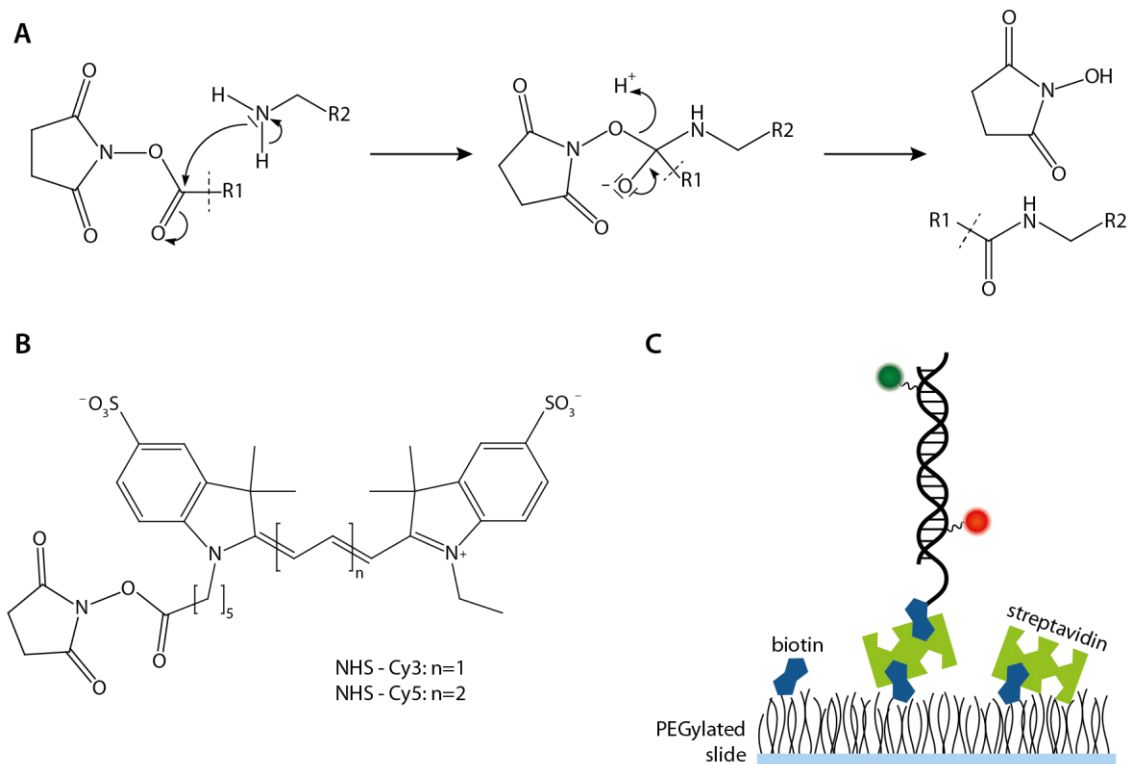


Figure 20: **A** Reaction mechanism of amine groups with NHS esters. Reaction products are a newly formed amide connecting the desired molecules R1 and R2 and NHS as side product. In typical applications R1 is a fluorophore (see B) and R2 is a biomolecule (e.g. RNA). **B** Structural formula of Cy3 and Cy5 NHS esters. **C** Biotin-streptavidin immobilization of FRET constructs on a PEGylated quartz slide.

2 Motivation

The human N6-methyltransferase complex and the corresponding formation of m⁶A has gained more and more interest. After the first identification of the complex in 1994,¹⁰ several aspects of the METTL3/METTL14 complex were uncovered (see chapter 1.2) including the positively charged binding groove on the METTL3/METTL14 interface, the RGG domain in METTL14 and the zinc finger domain in METTL3 being involved in target RNA binding.^{16–20} Most information is derived from crystal structures of the truncated complex, mutational analysis or biochemical work. However, there are no full-length or substrate RNA bound structures available until now. This leaves the RNA binding mode of the complex undefined. The structural organization of the RNA binding elements of the complex as well as the specificity received from the DRACH recognition mechanism need further investigation. Additionally, the modulation of the precise distribution of m⁶A sites in the transcript remained elusive at the beginning of this work. Since not all DRACH motifs are methylated, the specificity of the complex could be dependent on additional characteristics of the target RNA. Some of the m⁶A sites were shown to be dependent on mediation by some of the accessory proteins to target specific sites in the transcript.^{13,21,24,25} Besides the possibility of the presence of other, unidentified accessory proteins, this still does not sufficiently explain the m⁶A distribution. Current analysis utilizes antibody based sequencing approaches to determine m⁶A sites. This largely displays the overall distribution of m⁶A sites *in vivo*, however, is also vulnerable for antibody related errors. This could include artificial over- or underrepresentation of specific sites or contexts (sequence, structure, location, ...). Moreover, the level of methylation for any given specific sites is not resolved with this technique.

The first goal of this thesis was to shed further light on the specificity of the methyltransferase complex. The intrinsic specificity of the core complex consisting of METTL3 and METTL14 was chosen to be the best starting point for investigation. Besides RNA sequence, RNA secondary structure is apparently the second main characteristic of RNAs in general. Accordingly, RNA secondary structure should also influence methylation by the METTL3/METTL14 complex. The influence of the RNA binding domains was shown for the general methylation ability of the complex,^{16–20} however, this also raises the question of the influence of these domains on the specificity of the complex. First, the specificity of the full-length complex was investigated. For this, a set of different RNAs with various secondary structures needed to be designed and generated. Moreover, the full-length complex had to be expressed and purified. For efficient analysis of the methylation behaviour, site specific enzymatic activity assays were to be

performed on the different RNA structures. Based on the results from these first experiments, a subset of these RNA constructs was chosen for further experiments with truncated protein complexes. The truncated protein complexes needed to be designed, expressed and purified. The design of these proteins intended to be focused on truncation of RNA binding domains to investigate their role on RNA specificity and catalysis. Possible differences in RNA methylation were attempted to be linked to the RNA binding properties of the respective complexes. Further on, the influence of the accessory proteins on target RNA specificity were added to the analysis, starting with WTAP as binding platform. This included expression and purification of several complex compositions and repetition of the enzymatic activity assays with these complexes.

Together with the target RNA specificity of the MTC, the mechanism of the catalytic methyl transfer still remains uncovered. The active site of the catalytic core was postulated to be adjacent to the SAM binding site in METTL3, with a DPPW motif responsible for substrate adenine binding.¹⁶⁻¹⁸ The two flanking active site loops were shown to govern different positions around the active site for the cofactor-free, SAM- or SAH-bound states. This gives a slight hint towards possible dynamics of the complex during methylation. However, additional aspects of the dynamics of the methylation process are not known to date. Here, not only the dynamics during the actual methyl transfer but also the RNA binding and correct positioning in the active site are highly interesting.

The second part of this thesis was focused to gain insights into the dynamics of the complex during RNA binding and catalysis. Again, the first experiments focused on the full-length METTL3/METTL14 complex to capture its intrinsic behaviour. Prior to that, a suitable smFRET microscope needed to be set up. In the given case, this setup required TIRF sample illumination with a prism-based setup for highest flexibility. Fluorophores for the smFRET construct were attached to the RNA for minimal effort together with a great flexibility in measurement conditions, especially regarding the protein complex composition. If this construct design turns out to be suitable for conclusive measurements, RNAs with different target nucleotides including the substrate, target and an inhibitory nucleotide can be set up. Together with the protein complex, supplemented with SAM or SAH this can then be used to form catalytically competent, post-catalytic and catalytically incompetent complexes. Analysis of the measured data can be performed as bulk analysis to obtain histograms and gain information on the overall FRET state as well as on an individual scale for analysis of the dynamic behaviour.

3 Results and discussion – Specificity of the METTL3/METTL14 complex

3.1 Structural specificity of the full-length METTL3/METTL14 complex

The mode of target RNA recognition of the METTL3/METTL14 complex is mostly unknown. For the METTL3/METTL14 complex, some RNA interaction domains have been identified including the ZFD in METTL3, the RGG domain in METTL14 and the positively charged groove on the METTL3/METTL14 interface.^{16–20} Since not all of the DRACH sequences are methylated *in vivo*,⁶⁷ there presumably have to be additional layers for target RNA recognition. It is known that some of the assembling proteins guide methylation towards specific RNA sites,^{21,23–25} however, the structural context of the methylation process was mostly neglected to date. The structural context of the m⁶A sites in the transcript as well as the structural preference of the MTC are largely unknown.

For target RNA recognition and especially target specificity, the mode of target adenosine binding in the active site is of special relevance. Since there are only substrate RNA-free crystal structures of the shortened METTL3/METTL14 complex available, the RNA binding mode especially in the active site remains elusive. The comparison to structurally similar DNA methyltransferases, where substrate bound structures exist, provides an indication of a possible target adenosine location in the active site (see Figure 13D).¹⁶ This specific binding mode locates the RNA backbone alongside the positively charged groove with the adenosine in the active site adjacent to the DPPW motif and the donor methyl group, which structurally excludes a base paired target adenosine. For hydrogen bonded target adenosines this would require that the adenosine is flipped out of the existing structure as it has been reported for DNA methyltransferase.¹⁶ The reported possible mode of RNA binding leads to several assumptions for the RNA specificity of the METTL3/METTL14 complex. First of all, an unpaired adenosine is most likely required to be located in the active site for catalysis. This would render the stability of the target sequence containing secondary structure an important modulator for methylation efficiency, with tightly WC base paired adenosines being methylated least efficiently. Moreover, the overall secondary structure of the RNA should influence RNA binding for correct target positioning in the first place. The RNA-binding elements mentioned above (RGG domain, ZF domain, positively charged groove) suggest that anchoring of the RNA on both sides of the target nucleotide leads to an increased m⁶A formation. Taken together, RNAs with a tightly bound target adenosine such as complementary duplexes or bended RNA structures like

hairpins should decrease methylation yield, whereas an unstructured RNA presumably represents a hypothetically ideal target RNA for methylation with the METTL3/METTL14 complex. In 2014, Liu *et al.* assigned the complex to have no apparent preference towards a specific structure for methylation (e.g. only hairpins being methylated). However, they reported on different methylation yields for structured targets, indicating our assumption to be valid.¹³

To test the described hypothesis and look deeper into the structural specificity of the methyltransferase complex, the methylation efficiency of the full-length METTL3/METTL14 complex on target RNAs with different structural contexts was determined. The details of these experiments are presented in the following chapters.

3.1.1 Preparation of the full-length METTL3/METTL14 complex

In the first step, the full-length METTL3/METTL14 complex was expressed and purified for the following experiments. The basis for protein expression of the full-length METTL3/METTL14 complex from the pTriEx vector system was already established in previous work. Based on this work, METTL3 and METTL14 were coexpressed in line with the already set up protocol (see chapter 5.6 and 5.7).¹²⁹ After lysis of the expression pellet, the METTL3/METTL14 complex was isolated from the cell lysate via Ni-NTA affinity chromatography. Although METTL3 and METTL14 form a stable heterodimeric complex, both proteins carry a His¹⁰-tag at their N-terminal end. This leads to a mixture of heterodimeric complex and individual monomeric proteins that are

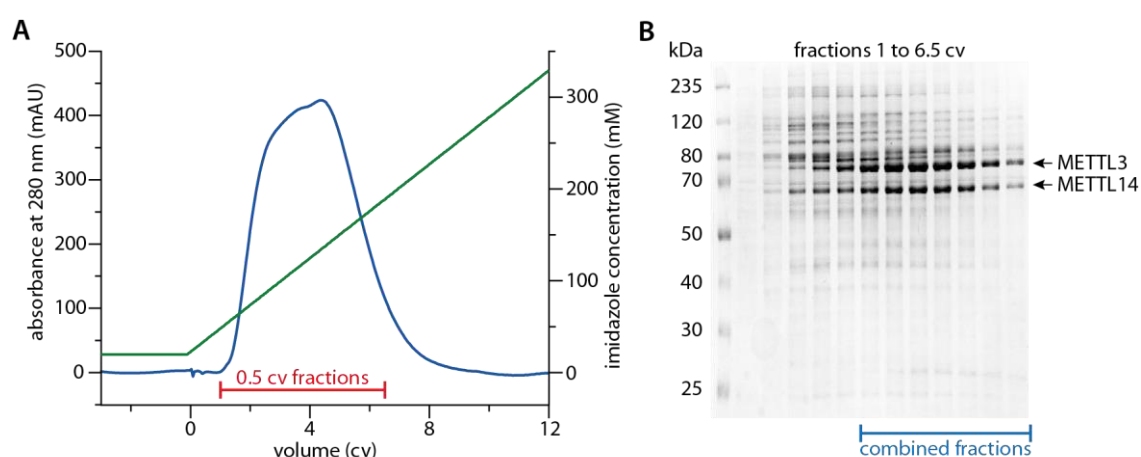


Figure 21: HisTrap purification of the METTL3/METTL14 complex expressed with the baculovirus generated via the flashBAC system. **A** Baseline corrected chromatogram with absorption at 280 nm (blue line), imidazole gradient (green line) and indicated analyzed fractions. **B** SDS-PAGE of the analyzed elution fractions. Combined fractions for further purification are indicated.

bound to the Ni-NTA column. After a washing step to eliminate unbound proteins, the desired proteins were eluted by an imidazole gradient. An example for the resulting baseline corrected chromatogram is shown in Figure 21A. Fractions with significant absorption were analyzed on an SDS-PAGE (Figure 21B). In the given example, protein bands for METTL3 and METTL14 were present in all analyzed fractions. However, protein bands for METTL3 are more pronounced indicating higher expression levels for this protein. Moreover, impurities of mainly higher molecular weight were present especially in the first fractions. Fractions containing primarily the desired proteins (3.5 to 6.5 cv) were combined and concentrated to 500 μ L for further purification.

To separate the intact heterodimeric METTL3/METTL14 complex from the undesired impurities and especially the copurified monomers, size exclusion chromatography was performed with a Superdex 200 Increase 10/300 GL column. The resulting chromatogram (Figure 22A) shows two main peaks that are partially overlapping. After analysis with SDS-PAGE (Figure 22B), the first main peak was identified as originating from the METTL3/METTL14 complex and the second peak from predominantly METTL3 monomers. Impurities of higher molecular weight could not be separated completely, however reduced to a very low proportion in relation to the pure complex. Fractions containing the purified METTL3/METTL14 complex (9.5 to 10.5 mL) were combined and concentrated to a protein concentration of 10 μ M. The 260 nm/280 nm absorption ratio of all samples was determined to be around 0.6 which attests to a nucleic acid free protein sample.¹³⁰

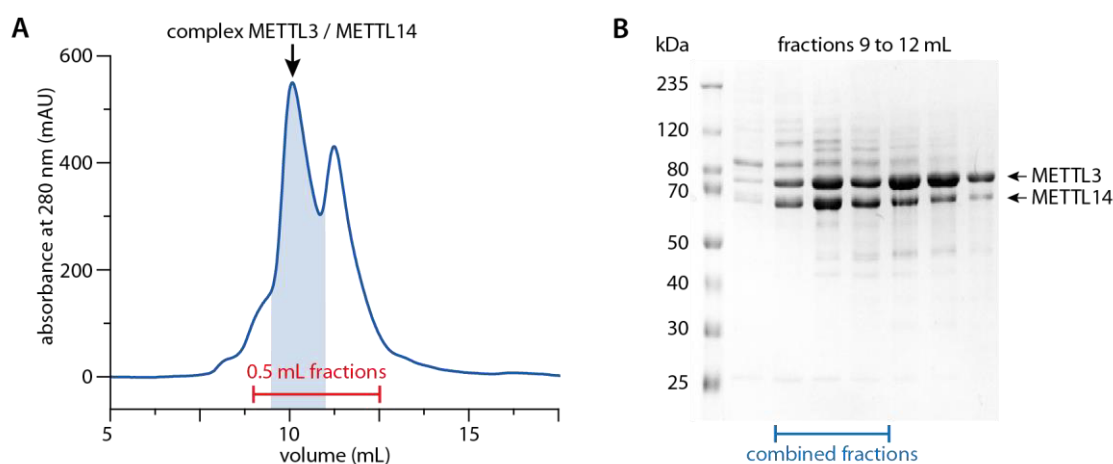


Figure 22: Size exclusion chromatography of the METTL3/METTL14 complex. **A** Chromatogram with absorption at 280 nm (blue line) and indicated analyzed (red) and combined (blue) fractions. **B** SDS-PAGE of the analyzed fractions. The combined fractions that were concentrated for use in further experiments are indicated.

3.1.2 Target RNA constructs

To test our hypothesis regarding the influence of RNA secondary structure on the methyltransferase activity of the METTL3/METTL14 complex, a set of RNA constructs with different secondary structures was designed (Figure 23). The RNA constructs were derived from known methylation sites of the suppressor of cytokine signaling 2 (SOCS2) transcript,¹³¹ which was shown to play a role in tumor progression in hepatocellular carcinoma.⁶¹ The m⁶A levels of this mRNA influence the YTHDF2 dependent mRNA degradation and with this the expression of the tumor suppressor SOCS2. We chose six different natural targets from this mRNA with different structure and DRACH sequence contexts. Three of them contain the predominant GGACU sequence motif bound in different structural contexts. The first one is loosely base paired in a duplex, adjacent to a bulge (D1). The second construct places the GGACU sequence inside a bulge within a duplex (D2), while the last one is bound adjacent to a hairpin loop (H1). The three additional constructs contain different DRACH sequence motifs that are located inside or adjacent to a hairpin loop (H3-H5). To complement the set of RNA structures, three additional RNA targets were constructed with a GGACU sequence context. One RNA was designed to be unstructured (U1), one construct has the target adenosine bound in a tight duplex (D3) and in the last RNA the GGACU sequence is located in the center of a hairpin loop (H2). With this set of natural m⁶A targets, complemented by artificial RNA constructs, different overall secondary structures (unstructured/duplex/hairpin) are covered together with a variety of local structures (unpaired/loosely bound/tightly bound) effecting the target sequence.

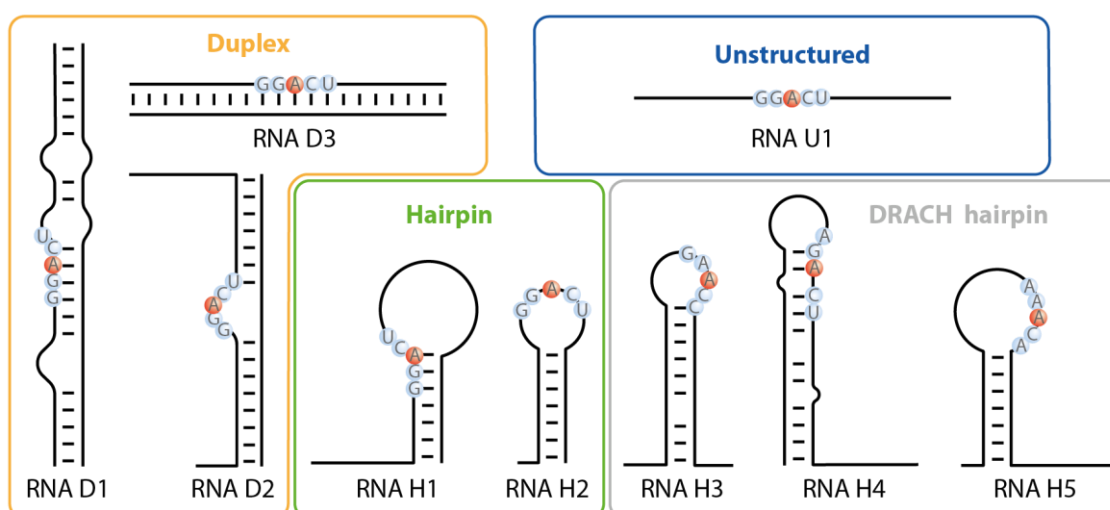


Figure 23: RNA constructs with different secondary structures and DRACH sequence contexts used for methylation activity experiments.

To ensure proper folding of the designed constructs under assay conditions, CD-spectroscopy melting curve analysis was performed. The RNA constructs (D1, D2, H1-H5) were first transcribed, purified and dissolved in activity assay buffer. PAGE analysis (Figure 24A) of the resulting RNA samples shows the expected RNA bands for all constructs. Two RNA bands are visible for the RNA duplexes as desired. Minor impurities from RNAs of lower (RNA D2, H2) or higher (RNA H1) molecular weight are visible for three constructs. The purified RNA constructs were then analyzed with CD-spectroscopy, yielding melting points between 46 °C (H1) and 67 °C (H2) (exemplarily shown in Figure 24B, remaining data see Figure 74 in appendix, page 169). For the RNAs with minor impurities visible in the PAGE analysis, no unusual melting curves or additional melting points were observed, indicating no disturbing effects. The given melting point analysis proves that all RNA constructs are folded under assay conditions at 37 °C.

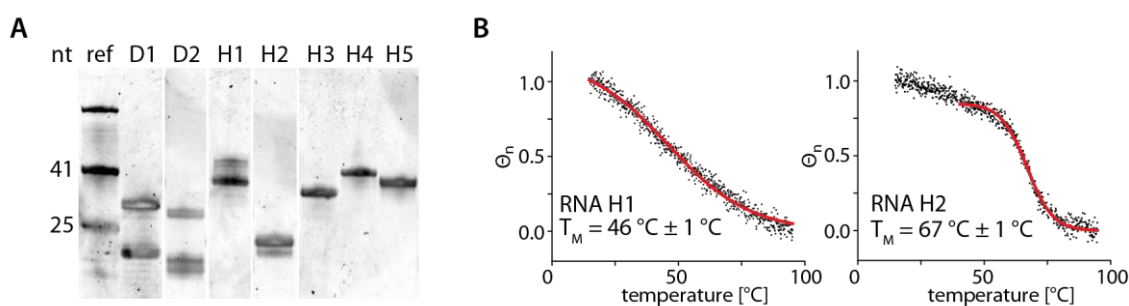


Figure 24: **A** PAGE analysis of purified RNAs for melting point estimation. **B** Graphical determination of melting points for the RNA construct with the lowest (H1) and highest (H2) melting point, based on CD spectroscopy measurements.

3.1.3 Site-specific RNA *in vitro* methylation assay

The common techniques described in the literature to study N6-methyltransferase activity are mostly based on either m⁶A specific antibodies (mostly for *in vivo* experiments) or rely on the use of radioactive SAM (mostly for *in vitro* experiments). For the latter, [³H-methyl]-SAM is used as methyl donor and the change in radioactivity of the investigated RNAs is detected as measure for methylation status. The advantage of detecting the sum of any formed m⁶A in the targeted RNA, however, also comes with the major disadvantage to be unable to site-specifically locate the methylated adenosines. To precisely investigate the influence of RNA secondary structure on the methylation activity of the METTL3/METTL14 complex, a site-specific approach is indispensable. To gain a site-specific and precisely quantifiable approach, we adapted an *in vitro*

Structural specificity of the full-length METTL3/METTL14 complex

assay based on ^{32}P -labeled RNA substrates (Figure 25).¹³² The ^{32}P -labeled RNA is synthesized from 2 fragments, where the 3'-fragment carries the target adenosine at its 5'-end. This 3'-fragment is then phosphorylated and with that partially ^{32}P labeled. This is achieved by using a mixture of ^{32}P - and ^{31}P - γ -ATP for phosphorylation. Then, the phosphorylated 3'-fragment is joined to the 5'-fragment via DNA splinted ligation, forming the final site-specifically ^{32}P -labeled target RNA. After methylation with the METTL3/METTL14 complex, the RNA is purified and completely digested to nucleoside monophosphates (NMPs). The sample is then analyzed via thin layer chromatography (TLC) where under the given conditions adenosine monophosphate has a different retention factor than N6-methyladenosine monophosphate.^{133,134}

Substrate RNA ligation



Methyltransferase activity assay

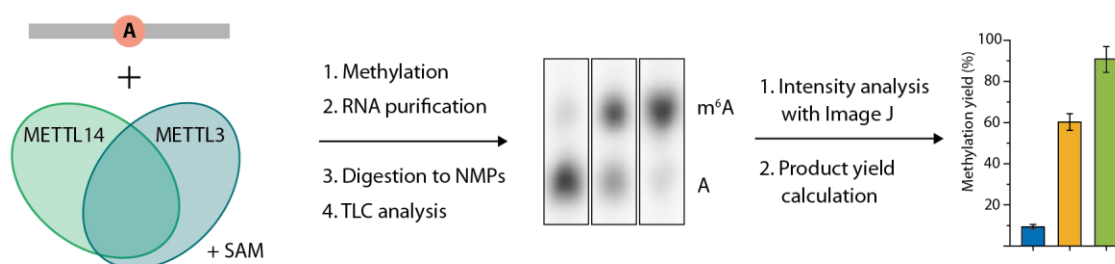


Figure 25: Site-specific methylation assay based on radioactive labeling of the target adenosine with ^{32}P . After substrate RNA production and methylation with the METTL3/METTL14 complex, sample analysis via thin layer chromatography allows quantification of the methylation efficiency.

As only radioactive NMPs are detected on the TLC, the NMPs originating from the ^{32}P -labeled target adenosine (methylated and unmethylated) are visible. This renders this approach not only site-specific but also allows for direct quantification of methylated against unmethylated target giving the absolute methylation yield, which is not always the case for the approaches mentioned above. Moreover, the use of ^{32}P as radioactive label has the advantage of a very sensitive detection option in comparison to ^3H .¹³⁵ Together with the variable radioactive proportion, this opens up the approach for a wide variety of sample concentrations and enzymatic activities. Taking into account the protein to RNA ratio during methylation, very low methylation yields can also be detected. We therefore chose a single turnover approach with excess protein over RNA to detect even small amounts of m^6A formation.

3.1.4 Methylation activity determination

After preparation of the METTL3/METTL14 complex, target RNA design and analysis of its folding, single turnover methylation assays with all target RNAs shown in Figure 23 were performed. The site-specifically ^{32}P -labeled RNAs were produced as described in chapter 5.8.1. To begin with, methylation activities of the full-length METTL3/METTL14 complex on all RNA constructs with the preferred GGACU sequence were analyzed (Figure 26). The results show the highest methylation yield for the unstructured RNA U1 with an absolute methylation yield of 70 % for this experiment. To have a better comparison between the different experiments, all other results were normalized to the methylation yield for RNA U1 (set to 100 %). For RNA D1 and D2 this gives a normalized methylation yield of 75 %, indicating a significantly lower methylation efficiency on these structures. D1 and D2 represent duplexes with a rather loose base pairing of the target adenosine. For the RNA structure with the GGACU sequence bound in a tight duplex (D3), no m^6A formation was detectable. By comparing these different RNAs, the biggest difference is the strength of WC binding towards the target adenosine as both the RNAs' length and possibly also their shape are similar for all of them. With increasing binding strength, the methylation yield is decreased. These results are perfectly in line with the hypothesis that stability of the target sequence containing secondary structure influences methylation efficiency in an anticorrelated manner.

Moving on to the hairpin structures H1 and H2, the determined methylation yields were reduced to approximately half (45 % and 52 %, respectively) compared to the unstructured RNA U1.

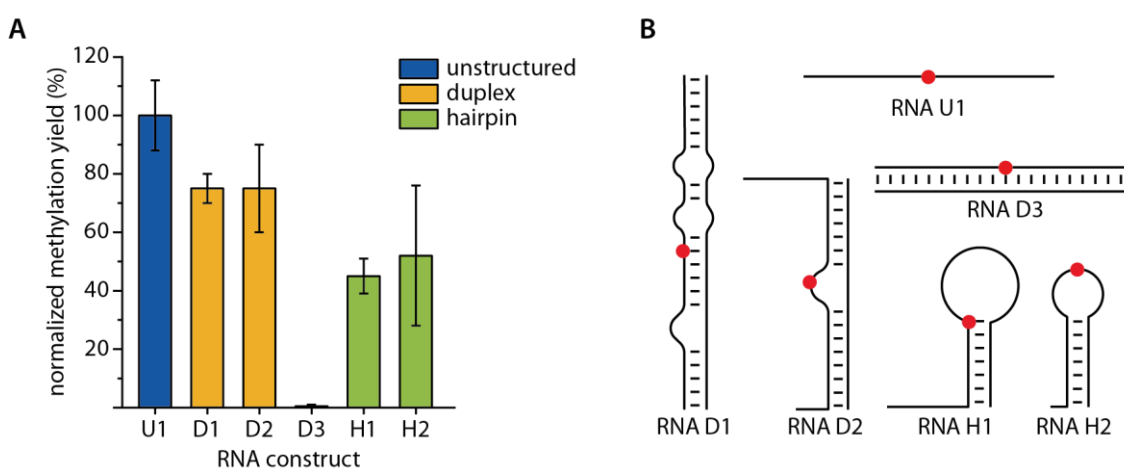


Figure 26: **A** Methyltransferase activity assay results on RNAs with different secondary structure, methylated with the full-length METTL3/METTL14 complex. The methylation results were normalized to RNA U1 (absolute methylation yield 70 %). **B** RNA structures of the tested RNAs with indicated target adenosine position (red dot, all inside GGACU sequence).

Structural specificity of the full-length METTL3/METTL14 complex

These hairpin structures provide a structural context for the target adenosine similar to RNA D1 and D2. However, the overall RNA secondary structure changes from a more stretched form (duplex) to a bent structure (hairpin). This presumably excludes protein binding on both sides of the target nucleotide sequence for the latter without breaking up the hairpin fold. In line with the previously described hypothesis, this renders the possibility for anchoring the RNA on both sides important for efficient methylation.

Next, the three remaining natural target RNAs with different DRACH sequence contexts inside a hairpin were investigated (Figure 27A). For all of them, even lower methylation yields than for the GGACU hairpin constructs were detected (7–19 % versus 45–52 %, respectively). Harper *et al.* already showed in 1990 that changes in the preferred GGACU sequence lower m⁶A formation using HeLa cell extracts as methyltransferase source.¹¹ Further on, we dissected whether target RNA structure or changes towards the consensus GGACU sequence within a DRACH motif predominantly effect methylation efficiency. To have a direct comparison, already examined RNA structures were mutated to other DRACH sequences largely maintaining the same secondary structure. Since the vast majority of the reported methylated DRACH sequences is actually within a RRACH context, we focused on mutating only towards the latter. RNA U1 was mutated to the different RRACH sequences, focusing on the two nucleotides preceding the target adenosine (RNA U1.1: AAACU, RNA U1.2: AGACU, RNA U1.3: GAACU), while RNA H4 was mutated from AGACU to the preferred GGACU sequence (RNA H4.1). For RNA 4.1 the secondary

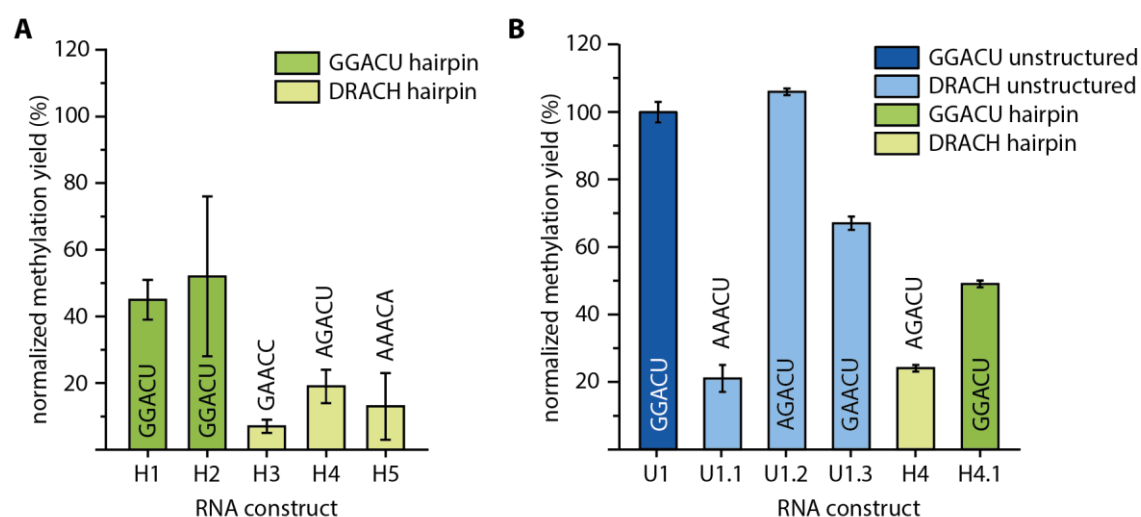


Figure 27: A Activity assays results for methylation with the full-length METTL3/METTL14 complex and hairpin RNAs of different DRACH sequences as target. The methylation results were normalized to RNA U1 (absolute methylation yield 70 %). **B** Activity assay results for the RRACH mutated versions of RNA U1 and H4. The methylation results were normalized to RNA U1 (absolute methylation yield 85 %).

structure largely remains the same as for RNA H4 and mutation of RNA U1 still keeps the unstructured form(predicted with m-fold).¹³⁶ The methylation results of the mutated unstructured RNAs (U1.1-1.3) showed the largest decrease in methylation yield for the AAACU sequence motif (normalized methylation yield 21 %). Mutation to GAACU also significantly decreased methylation yields (67 %), whereas methylation efficiency is unaltered, if not slightly increased for the AGACU sequence context (106 %). These results render the central GAC motif as most important element of the RRACH sequence context for efficient methylation in unstructured RNAs, also supported by literature.¹¹ For the hairpin RNA, the methylation activity was doubled upon mutation of the AGACU sequence (24 %) towards GGACU (49 %). Forming the preferred sequence context in this RNA raised the methylation yield to the same level as observed for the other GGACU hairpin constructs (H1, H2), while methylation levels of the unstructured RNA (even with a GAACU target sequence) were out of reach. This again proves the importance of the RNA secondary structure for methylation efficiency.

3.1.5 Conclusion and outlook

In this chapter, the influence of target RNA secondary structure and target sequence were investigated for the full-length METTL3/METTL14 complex. The sensitivity towards the local target RNA structure was investigated via enzymatic activity assays on substrate RNAs with different structural contexts. The highest methylation efficiency was observed for an unstructured RNA. Significantly lower product formation was detected for duplexes with a loosely base paired target region (GGACU) whereas no m⁶A formation was detected for duplexes in full complementarity. RNAs with the GGACU target sequence in or around the loop of a hairpin structure were shown to be methylated in all cases, but with the lowest methylation efficiencies. The results render the target sequence containing secondary structure stability and the possibility for anchoring on both sides of the target nucleotide as especially relevant for efficient catalysis. These results are in line with other *in vitro* methylation studies.¹³ A recent study on the methylation activity of the METTL3/METTL14 complex on DNA constructs similarly showed that double stranded DNA and DNA/RNA hybrids are not methylated by the complex.¹³⁷

Alterations of the GGACU sequence within the RRACH sequence context decreased methylation yields, where the central GAC sequence is of special relevance. We observed a stronger influence of the RRACH sequence context on the methylation efficiency in a hairpin structured RNA than for the unstructured RNA. Taken together, the results show an influence of RNA secondary structure and target sequence context on the methylation efficiency of the

METTL3/METTL14 complex. Both factors have similar significance for the methylation yield *in vitro*.

The local RNA secondary structure as well as the RRACH sequence context were both shown to be important modulators for methylation efficiency *in vitro*. Mapping techniques for the *in vivo* distribution of m⁶A sites in the transcript are mostly based on m⁶A specific antibodies. These studies uncovered the common DRACH sequence context of the identified m⁶A sites.^{3,5,70} However, the structural context of these non-randomly distributed sites was not investigated and still remains elusive. This leaves the question whether the here observed structural sensitivity towards the RNA targets is mirrored in the *in vivo* distribution of m⁶A sites. Closely related to this are several considerations regarding the specificity of the m⁶A antibodies.¹³⁸ One of the fundamental concerns are false positive reads on the one hand and low methylated sites being buried in the background noise on the other hand. Moreover, the possible structural sensitivity of the m⁶A antibodies themselves is reasonable to be taken into consideration. There might be a different affinity of the antibodies towards m⁶A sites in varying structural contexts leading to artificial up- or downscaling of identified m⁶A sites in a specific RNA secondary structure. In future work, these questions and considerations would be interesting targets to investigate. Some of these questions are subject to collaborative work with the groups of Kathi Zarnack and Julian König, however, this project did not yield robust results yet.

Methylation of mRNAs by the MTC can occur cotranscriptionally *in vivo* with the nascent RNA chain being directly methylated.^{12,14,15} The relevance of RNA secondary structures in this case is highly dependent on the length of the nascent RNA and relatedly the kinetics with the possibility to form (locally) stable secondary structures that can impact the methylation process. This also includes the influence of other (cotranscriptional) proteins on the methylation process. The given results indicate that the core complex itself has no apparent RNA unwinding activity. However, this does not rule out the presence of other protein cofactors such as helicases that would diminish the influence of RNA secondary structure on methylation efficiency. Moreover, additional proteins of the MTC such as WTAP themselves might influence the RNA structure dependent behavior. This also raises the question on the role of (RNA binding) protein domains in target recognition and secondary structure sensitivity of the METTL3/METTL14 complex *in vitro*, which is addressed in the next chapter.

Very recently, new insights on the distinct distribution of m⁶A sites in the transcript were published by Uzonyi *et al.*⁵⁰ They showed this distinct distribution being mostly dependent on the accessibility of the DRACH sites, which in case of (short) exons are occupied by the exon

junction complex. This physical exclusion of a defined set of DRACH sites was moreover shown to be largely predictable and in line with the measured m⁶A profile. Although this study provides evidence on the general m⁶A profile in the transcript, the level of methylation for individual sites is not fully covered by this. This includes that not every DRACH motif in the parts physically accessible for methylation is (measurably) methylated. Moreover, the level of methylation for the individual m⁶A sites is not captured in the antibody-based mapping approaches, which also refers back to the antibody-related concerns discussed above. The secondary structure of the specific DRACH sites might still be a relevant factor especially for the methylation level of individual sites within the limits discussed above.

3.2 Influence of protein domains on substrate specificity and catalytic activity

Based on the results described in the previous chapter, the role of the individual protein domains in RNA binding and catalysis, focusing on RNA secondary structure sensitivity was of special interest. To assess the contributions of the individual protein domains, a set of subsequently shortened protein constructs was designed (Figure 28). For this, we focused on the RNA binding domains of the complex. In METTL3, first the N-terminal domain (NTD) was removed (M3 Δ NTD) leaving only the protein domains relevant for *in vitro* catalytic activity, as described previously.^{16,17,19} Further on, the first CCCH motif of the ZFD was eliminated (M3 Δ ZF1), followed by the whole ZFD, remaining only with the MTD (M3MTD). METTL14 was shortened from both sides, first eliminating the RNA binding RGG domain (M14 Δ RGG). To finalize the set of truncated METTL14 constructs, the NTD (M14 Δ NTD) as well as both, NTD and RGG domain, were cut off again leaving only the MTD (M14MTD).

To get an approximation of the impact of the individual protein domains on RNA catalysis, the methylation efficiency was first tested on the unstructured RNA U1 and based on these results additional RNAs of different secondary structure were tested. Moreover, the RNA binding abilities of the complexes were investigated. The respective experiments and results are depicted in detail in the following chapters.

3.2.1 Preparation of shortened METTL3/METTL14 complexes

First all truncated protein complexes were cloned, expressed and purified. The gene coding plasmid for M3MTD was derived from the METTL3 full-length plasmid through site directed mutagenesis (see chapter 5.5.3.1). All other shortened genes were purchased from *Genscript* in the same vector backbone. After virus generation via the flashBAC system, small scale test expressions were carried out for the individual expression of all shortened proteins. The western blot analysis (anti-His) showed the desired protein bands in the correct size range, indicating suitable protein expression for all constructs (Figure 28).

In the next step, all shortened proteins were coexpressed together with the full-length complex partner in larger scales. Moreover, a minimal complex consisting of only the two MTDs (M3MTD/M14MTD) was expressed.¹⁸ After having the cell pellets for all different complex

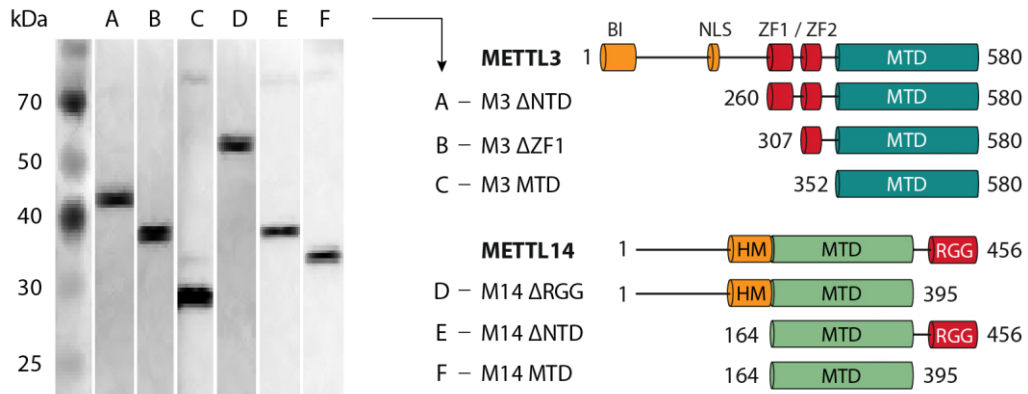


Figure 28: Test expressions of all shortened METTL3 and METTL14 constructs analyzed with a western blot (anti-His). Shortened constructs are shown schematically with RNA binding elements colored in red and additional interesting domains colored in orange.

compositions in hand, the complexes were purified according to the already established protocol for the full-length complex (see chapter 5.7). Initial purifications and optimizations were performed as part of an internship with Nicole Mench. In the following section, the purification steps for the truncated complexes are shown on the example of complex M3ΔNTD/METTL14. The purification chromatograms and analytical SDS-PAGEs for all other shortened complexes can be found in chapter 6.2.

In the first step of the purification process after cell lysis, the His tagged proteins were bound to a Ni-NTA column and separated from unspecifically bound proteins by imidazole gradient

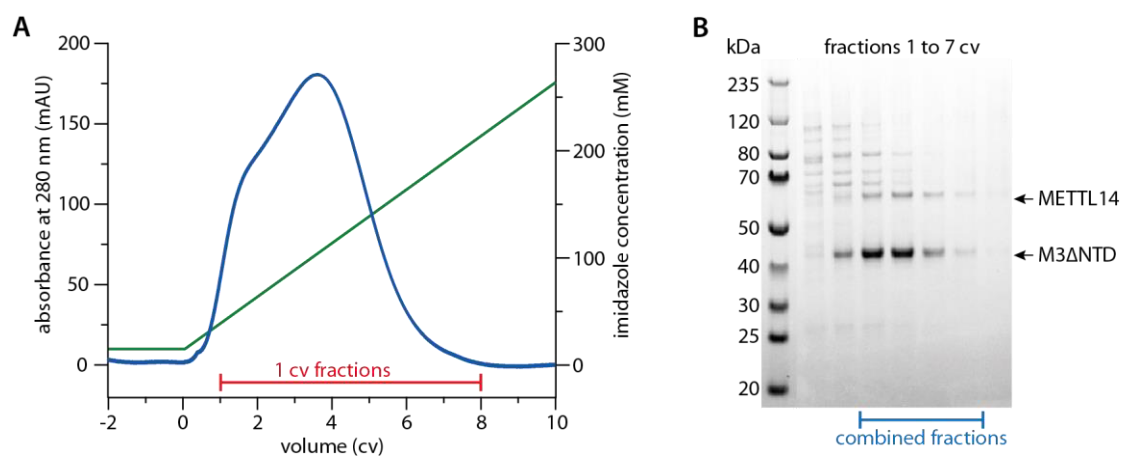


Figure 29: HisTrap purification of the M3ΔNTD/METTL14 complex. **A** Baseline corrected chromatogram with absorption at 280 nm (blue line), imidazole gradient (green line) and indicated analyzed fractions. **B** SDS-PAGE of the analyzed elution fractions. Fractions combined for further purification are indicated.

elution. Protein containing fractions were identified by significant absorption at 280 nm in the resulting chromatogram (Figure 29A) and analyzed via SDS-PAGE (Figure 29B). It is visible that both complex partners are present in several fractions, however, the protein band for M3 Δ NTD is significantly more intense. The fractions containing the desired proteins were combined and concentrated to 500 μ L for further purification by size exclusion chromatography.

The separation of the proteins regarding their size in the next step is important to compensate for imbalance in expression levels, in this case to remove excess M3 Δ NTD monomer. At the same time, minor high molecular weight impurities from the preceding step are eliminated. The chromatogram of the size exclusion chromatography (Figure 30A) shows two clearly separated major peaks. The analysis via SDS-PAGE (Figure 30B) proves the first major peak to originate from the desired M3 Δ NTD/METTL14 complex with very good separation from the M3 Δ NTD monomers in the second peak. Fractions with the pure complex (11 to 13 mL) were combined, checked for an OD 260/280 ratio of approximately 0.6 for a nucleic acid free sample and concentrated to a protein concentration of 6 μ M.

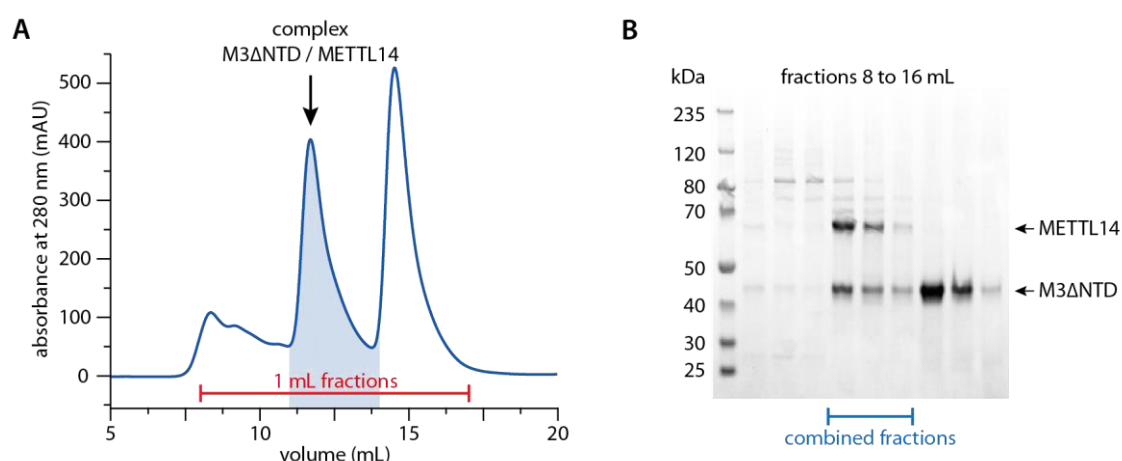


Figure 30: Size exclusion chromatography of the M3 Δ NTD/METTL14 complex. **A** Chromatogram with absorption at 280 nm (blue line) and indicated analyzed (red) and combined (blue) fractions. **B** SDS-PAGE of the analyzed fractions. The combined fractions that were concentrated for use in further experiments are indicated.

This procedure was repeated for all shortened complexes. Figure 31 shows an analytical SDS-PAGE of all the purified METTL3/METTL14 complexes. For all of them, the major protein bands match the expected protein sizes. Minor impurities are still visible; however, the overall sample purity is sufficient.

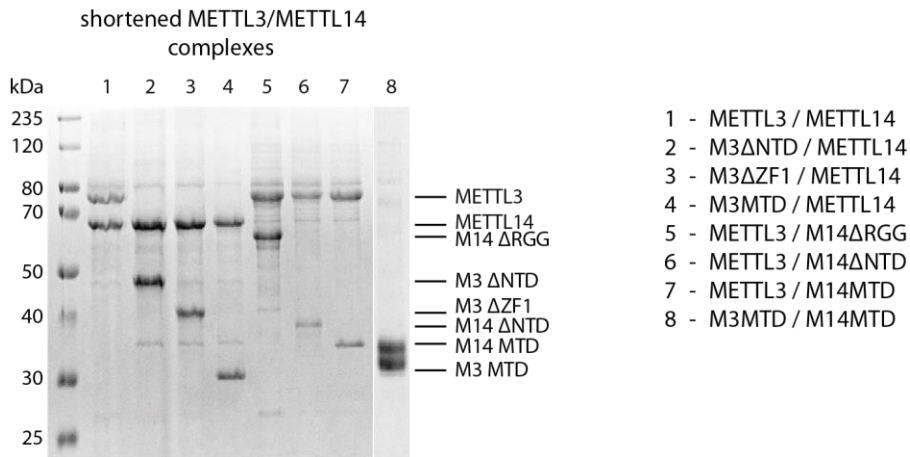


Figure 31: SDS PAGE of all purified truncated constructs.

After purification, the size exclusion chromatograms for the different complexes were compared and illustrated with indicated retention volumes (Figure 32). For a better overview, retention volumes together with the molecular weights of the complexes are listed in Table 2. The complexes shortened in METTL14 (Figure 32A) show the expected behavior in comparison

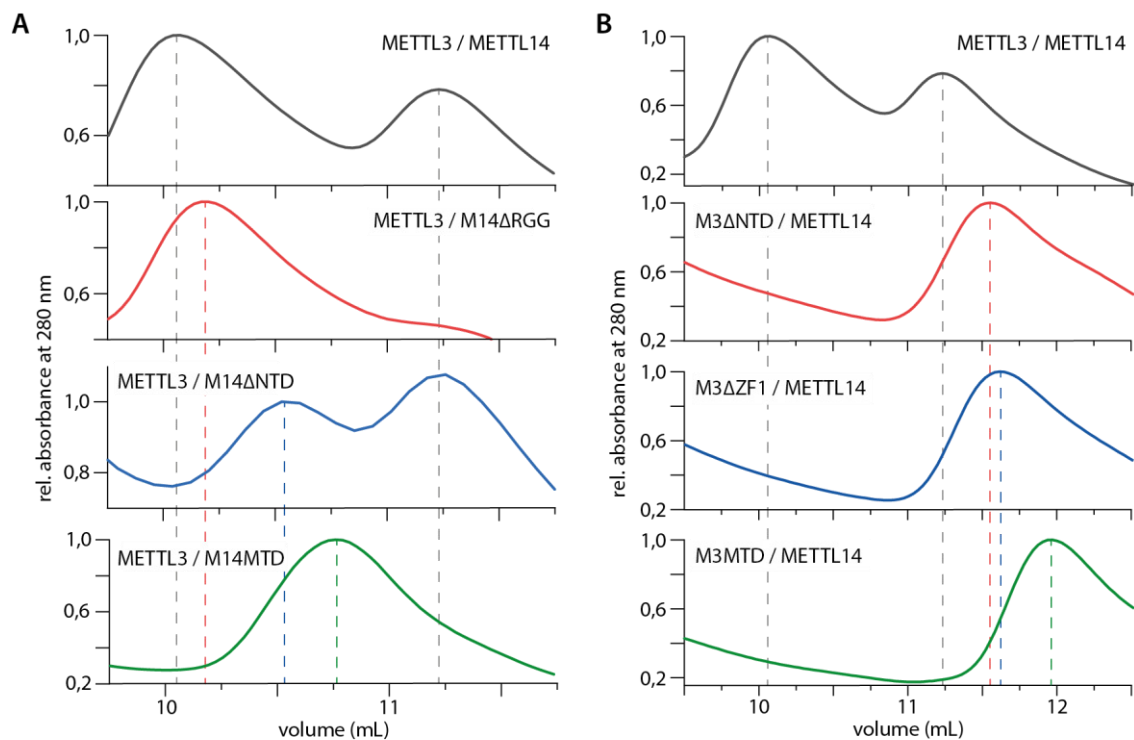


Figure 32: Comparison of the retention volumes (dashed lines) during size exclusion chromatography for the METTL3/METTL14 complexes shortened in METTL14 (A) and METTL3 (B). Sections of the main peaks are shown for comparison to the full-length complex (black chromatogram).

to the full-length complex: with decreasing molecular weight of the complex, the retention volume is increasing. Moreover, in comparison with the retention volume of the METTL3 monomer all METTL14 shortened complexes show a lower retention volume due to their higher molecular weight. Interestingly, this is not the case for the METTL3 shortened complexes (Figure 32B). Although the smaller complexes also show an increased retention volume, all complexes have higher retention volumes than the METTL3 monomer. This was unexpected as all of the METTL3 shortened complexes are of higher molecular weight than the METTL3 monomer.

Table 2: Summary of the retention volumes for all shortened complexes during size exclusion purification. The compared retention volumes were derived with the same column, purification system and flowrate. Molecular weights are given with expression tags.

Complex	Retention volume (mL)	Molecular weight (kDa)
METTL3/METTL14	10.05	127.56
METTL3/M14ΔRGG	10.18	121.30
METTL3/M14ΔNTD	10.55	108.31
METTL3/M14MTD	10.75	102.05
METTL3 monomer	11.23	69.94
M3ΔNTD/METTL14	11.55	99.74
M3ΔZF1/METTL14	11.60	94.09
M3MTD/METTL14	11.95	89.06
M3MTD/M14MTD	13.75	63.55

Although this was unexpected, it is in line with a recent study solving cryo electron microscopy structures of the accessory protein complex.⁸⁰ Attempts to solve structures together with the full-length METTL3/METTL14 largely failed due to the possibly high flexibility and stretched form of the domain between the MTD and the WTAP binding interface in METTL3. Moreover, previous crystal structures of the METTL3/METTL14 complex only cover the two MTDs.^{16–18} This also hints towards flexible regions in the remaining domains, which could explain the larger differences for retention volumes upon shortening of METTL3 in the NTD. The two MTDs themselves form a tight heterodimer with an almost „ball-like” structure. Adding the flexible N-terminal regions in METTL3 presumably extends the “ball” with a “tail” and influences the behavior on the size exclusion column. The “tail” increases the hydrodynamic radius more than a “ball” of the same weight would do, which could explain the observed retention volumes.

3.2.2 General methyltransferase activity

Before preparing the activity assays with the truncated protein complexes, the assay conditions were further optimized regarding the protein to RNA ratio (appendix page 169, Figure 75). For the following experiments, the protein to RNA ratio was lowered from 20:1 to 5:1, which is less protein consuming but leads to similar methylation yields. In the following section, the truncated protein complexes (Figure 28) are named after the shortened protein without mentioning the full-length complex partner (M3 Δ NTD = complex M3 Δ NTD/METTL14), unless stated otherwise. In all experiments, the full-length complex was added as a positive control while the inactive M3MTD/M14MTD complex was used as negative control. To begin with, all newly expressed and purified complexes were tested for methyltransferase activity on the unstructured RNA U1, representing the most efficient methylation target for the full-length complex (see chapter 3.1.4). All methylation yields were normalized to the results for the full-length complex (absolute methylation yield 81 %).

An increased methylation yield was observed for complex M3 Δ NTD (112 %) (Figure 33). For all other constructs, a significant decrease in methylation efficiency was detected (19 to 42 %). These results support that all the RNA binding elements are required for efficient methylation *in vitro*. In METTL3, one zinc finger is not sufficient for productive methylation, which is in line with results published by Huang *et al.*¹⁹ Depletion of the RNA binding RGG domain in METTL14 reduced the methyltransferase activity to 42% compared to the full-length complex. This result is in between the findings of Schöller *et al.* and Huang *et al.* / Slédz *et al.*^{16,19,20} Schöller *et al.* reported on a reduction to almost background levels for a RGG domain deleted complex,²⁰ while

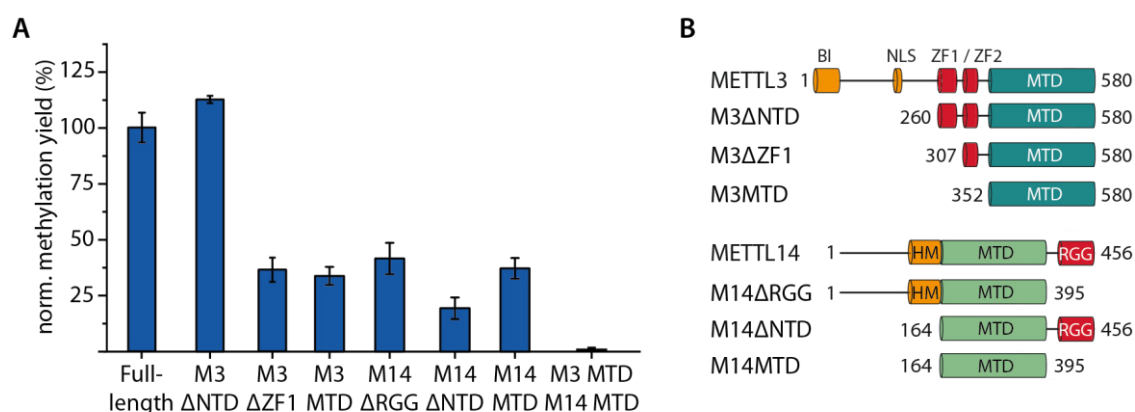


Figure 33: A Activity assay results for all truncated protein complexes with the unstructured RNA U1 as target. All methylation yields are normalized to the full-length complex (absolute methylation yield 81 %). **B** Truncated proteins used as complexes with the full-length complex partner.

the others did not detect any significant activity changes for a complex with METTL14 present only as MTD.^{16,19} In our case, the decrease in methylation yield upon deletion of the RGG domain supports the findings of Schöller *et al.* identifying this RNA binding domain as important for efficient catalysis. Aside from this, the findings of Huang *et al.* and Slédz *et al.* are in contrast to our methylation data for complexes M14 Δ NTD and M14MTD as we observed significant reduction in methylation yield. However, the N-terminal boundaries of the MTD in METTL14 were chosen differently. Huang *et al.* and Slédz *et al.* included the adjacent N-terminal α -helical motif (HM, aa117-136)^{16,19} while our constructs span solely the “pure” MTDs as identified by Wang *et al.*¹⁸ This suggests an important role for the HM in METTL14 for catalytic activity presumably related to its structural impact. It was shown that this domain is an essential part of the heterodimeric interaction framework, adjusting the relative positions of the two MTDs and protecting hydrophobic interaction regions from solvent exposure.^{16–18}

3.2.3 RNA binding properties

Investigation of the methylation activities of the truncated complexes showed that all RNA binding domains are indispensable for efficient methylation *in vitro*. Next, we examined if this is indeed a result of the diminished RNA binding ability of the complexes.

To begin with, we performed microscale thermophoresis (MST) measurements to exactly determine the dissociation constant (K_D) of the complexes towards the unstructured RNA U1. These experiments were carried out as part of an internship with Claudia Catapano. As a first

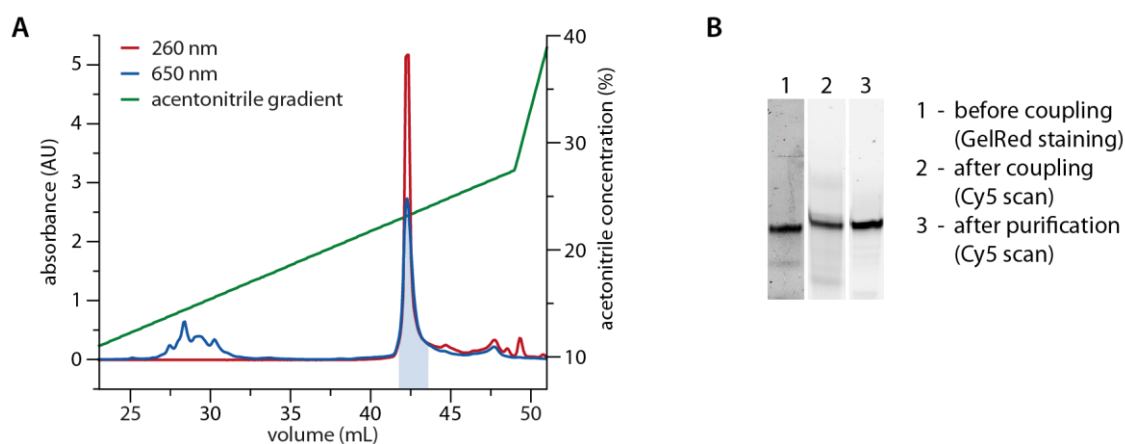


Figure 34: **A** Chromatogram of the purification of RNA U1-Cy5 over a Kromasil C8 HPLC column after labeling. **B** Analytical PAGE of the sample before fluorescent labeling and after Cy5 coupling prior and after purification.

step, the RNA U1 was fluorescently labelled with Cy5. The purification chromatogram and the analytical gel are shown in Figure 34, indicating a quantitatively labelled and pure RNA sample.

The first MST measurements were performed with the full-length complex with and without SAM or SAH (Figure 35). This resulted in a K_D of approximately 4 μM for all combinations with minor differences (3.5 to 4.7 μM). The derived K_D is lower than the one reported by Huang et al. (21 μM)¹⁹ but approximately in the same range.

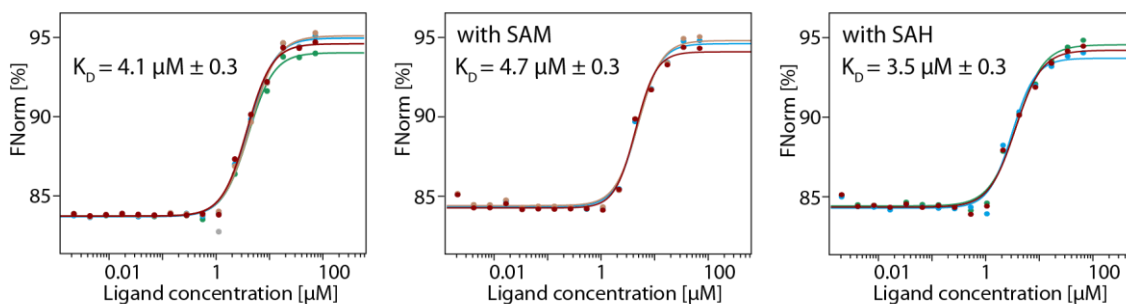


Figure 35: Results of the K_D determination derived with microscale thermophoresis (MST) measurements for the full-length METTL3/METTL14 complex and the fluorescently labelled RNA U1 (U1-Cy5). The subsequent measurements are color coded.

MST is able to precisely determine the K_D , however, in this case the measurements were restricted by a protein concentration of 72 μM . In the given MST data for the full-length complex, this limits the upper part of the titration which did not completely reach saturation in all cases. As decreased RNA binding abilities were expected for the truncated complexes, this would require approximate protein concentrations in the upper micromolar range. These high concentrations were not applicable in our case, which is why we decided to shift towards a different technique, and selected electrophoretic mobility shift assays (EMSAs) as method of choice. Here, precise K_D determination is quite challenging, however, qualitative changes in RNA binding can be observed already with small sample amounts. Moreover, we again were able to use fluorescently labelled RNAs as target to enhance detection limits and further lower the required protein concentrations. This resulted in an about 10-fold reduction of the required protein concentration in comparison to the MST samples.

The EMSAs with the different truncated complexes were first done with the unstructured RNA U1-Cy5 as target. The resulting gel (Figure 36A) shows no significant differences in RNA binding for complexes M3 Δ NTD and M3 Δ ZF1 in comparison to the full-length complex. For all other truncated complexes, the RNA binding strength was significantly decreased. These results are

Influence of protein domains on substrate specificity and catalytic activity

mainly in line with the observed methyltransferase activities of the complexes, linking low methylation yields to a loss of RNA binding strength. However, for the complex with M3 Δ ZF1 a reduced methylation yield is not associated with a decreased RNA binding ability, which suggests an unspecific RNA binding by the remaining zinc finger. Besides, the truncated ZFD might be partially altered in protein folding, which can also not be completely ruled out for the other truncated complexes. To test the assumption of an unspecific RNA binding, EMSAs were repeated with a partial RNA duplex (D4) and an unstructured off-target RNA (OT) (no DRACH sequence) for complexes M3 Δ NTD, M3 Δ ZF1 and the full-length complex as a reference. Surprisingly, the full-length complex showed the same RNA binding strength towards the two RNAs as observed for the unstructured RNA U1. Interestingly, complex M3 Δ ZF1 and M3 Δ NTD displayed even an increased RNA binding behavior towards the additional RNAs. This supports

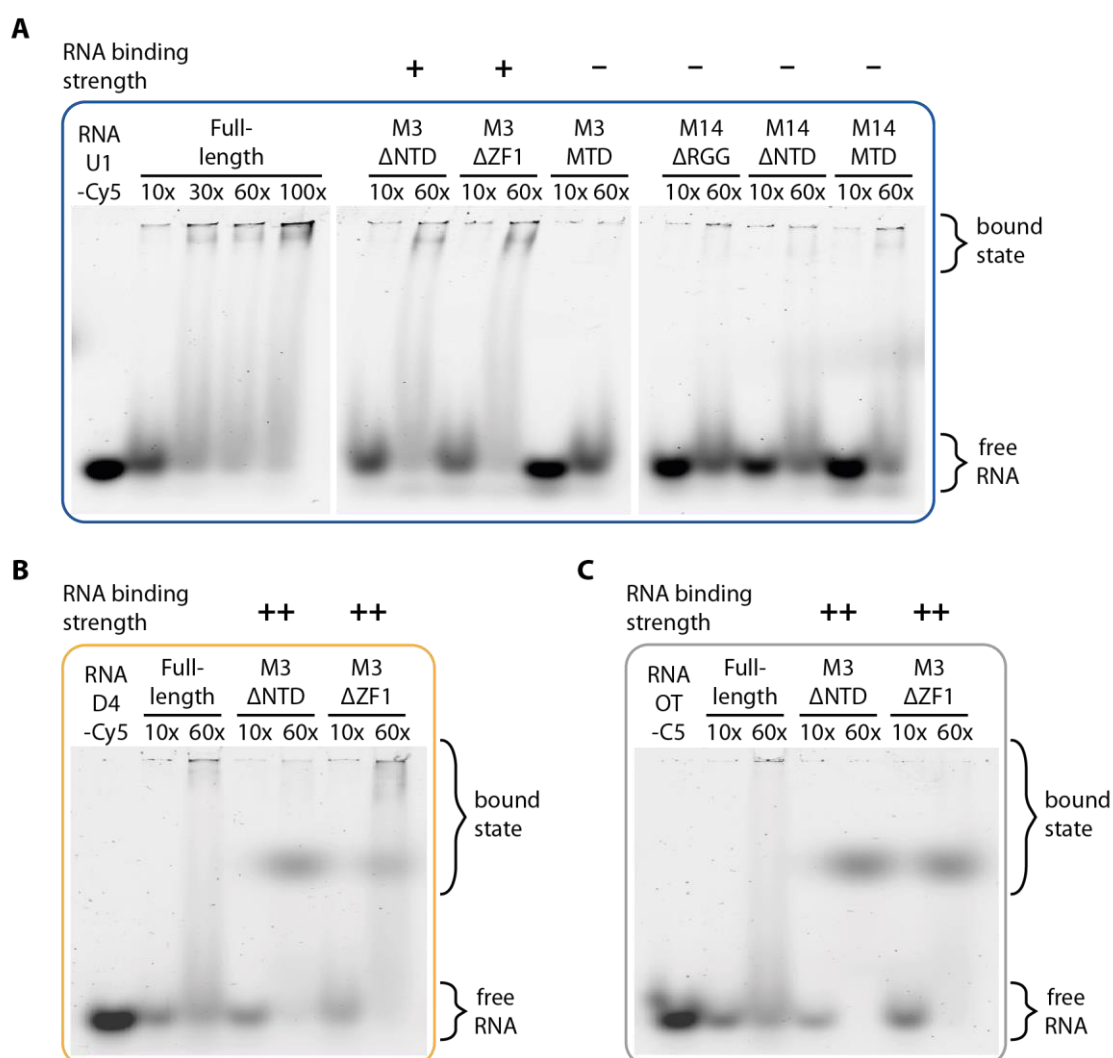


Figure 36: EMSAs of the shortened complexes to assess differences in RNA binding towards the unstructured RNA **(A)** an RNA duplex **(B)** and an unstructured RNA without the consensus DRACH sequence **(C)**.

the assumption of an unspecific RNA binding through the remaining zinc finger in complex M3 Δ ZF1. It is possible, that the RNA is bound, however, not always in the right orientation for catalysis. The increased binding affinity for M3 Δ NTD likely does not originate from the same mechanism since both zinc fingers are present and a very efficient methylation was observed for this complex. This may hint towards a better accessibility of the RNA binding domains in the NTD depleted complex with respect to the full-length complex.

3.2.4 Secondary structure specificity

After assessing the efficiency of the truncated protein complexes to methylate the unstructured RNA U1, the role of the different protein domains on RNA secondary structure specificity was investigated further. The activity assay experiments were repeated with RNA H4.1 representing a hairpin structure. All methylation yields were normalized to the yield for the full-length complex with RNA U1 as target (absolute methylation yield 81 %). The results (Figure 37) show a significantly lower methylation efficiency on the structured RNA for all complexes, except for complex M3 Δ NTD. This complex exhibits a unique behavior where only a minor decrease in methylation efficiency was observed. In contrast, the two complexes that also lack the NTD in METTL3 (M3 Δ ZF1, M3MTD) show approximately the same decrease in methylation efficiency on the structured RNA as observed for the full-length complex. This becomes apparent when directly comparing the methylation efficiencies of the same complex on the two different RNAs. For the structured RNA H4.1, this gives a reduction to 62 % (full-length), 68 % (M3 Δ ZF1) and

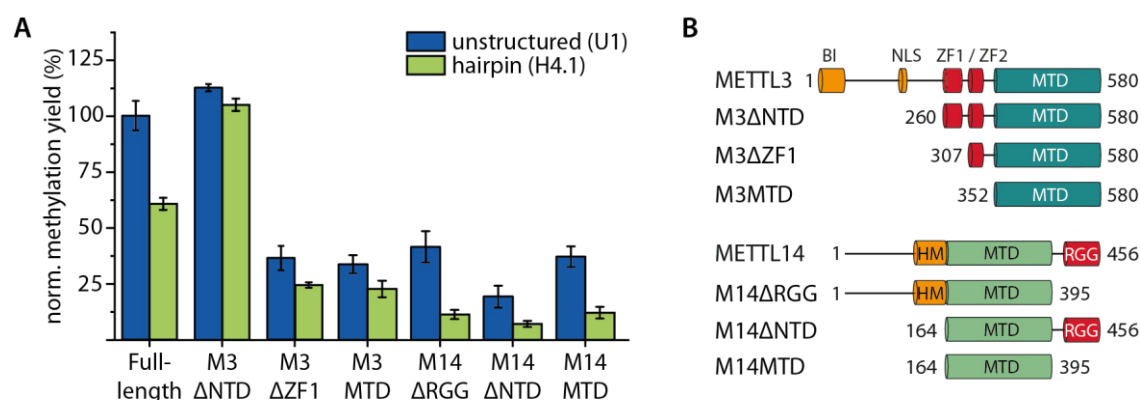


Figure 37: A Activity assay results for the truncated protein complexes with a structured RNA (RNA H4.1) as target in comparison to RNA U1. All methylation yields are normalized to the full-length complex with RNA U1 as target (absolute methylation yield 81 %). **B** Truncated proteins used as complexes with the full-length complex partner.

66 % (M3MTD) of the methylation yield on the unstructured RNA U1, respectively. However, it is questionable if the comparison of the two complexes M3 Δ ZF1 and M3MTD towards M3 Δ NTD is valid, since these complexes already show an altered methylation behavior on the unstructured RNA. Nevertheless, we can state that this behavior is consistent regardless of RNA structure.

The complexes shortened in METTL14 all showed very low methylation yields for the structured RNA (8 to 12 % absolute methylation yield). Here the largest difference between structured and unstructured RNA was observed for complex M14 Δ RGG (reduction to 27 % of unstructured). This further supports the importance of the RGG domain as an RNA binding domain. Moreover, this strengthens the previous finding that the possibility for anchoring the target RNA on both sides of the complex is important for catalytic activity. For the unstructured RNA, the ZFD in METTL3 probably compensates the missing RGG domain in METTL14 partly. When changing from the rather stretched form of the unstructured RNA to the bended hairpin structure, this compensation is most likely less effective. This again raises the question if the target RNA is bound with a preferred or random orientation inside the RNA binding groove and whether this plays a role in RNA binding through the ZFD and RGG domain. Binding of the target sequence in and near the active site occurs most likely always in the same orientation as only RRACH sequences are methylated and not their inversed counterparts (5' to 3' only GGACU and not UCAGG). Moreover, this is presumably supported with the comparison of the structural sensitivity of the ZFD deleted complexes (M3 Δ ZF1, M3MTD) to the RGG deleted complexes (M14 Δ RGG, M14MTD). Here, the RGG depleted complexes show a more pronounced decrease in methylation yield for the hairpin RNA. This hints either towards a stronger compensation ability of the RGG domain or a preferred orientation of the hairpin in favor of the RGG domain for this construct. In support of the latter, the ZFD deleted complexes show a sensitivity towards the substrate RNA structure similar to the full-length complex.

The most interesting behavior however was observed for complex M3 Δ NTD which still harbors all identified RNA binding domains. To get a more refined understanding of the decreased structural preference of this truncated complex, additional structured RNAs were tested as methylation targets: a loosely base paired duplex (RNA D1) and another hairpin RNA with a different target nucleotide location (RNA H1). All of these tested RNAs harbor the preferred methylation sequence GGACU. For better comparison, the experiments were also carried out with the full-length complex and complex M3 Δ ZF1. Methylation yields were normalized to the yield of the full-length complex with RNA U1 as target (absolute methylation yield 81 %). The

derived methylation efficiencies (Figure 38) supported the observed behavior for complex M3 Δ NTD. Overall, the derived methylation yields complex M3 Δ NTD were significantly higher than for the full-length complex. There was no significant decrease in activity of complex M3 Δ NTD on all structured RNAs with reference to the unstructured RNA. In comparison, the expected differences in methylation yield were robustly observed for the full-length complex and complex M3 Δ ZF1. Here, methylation yields were significantly lower for the structured RNAs than for the unstructured RNA as target.

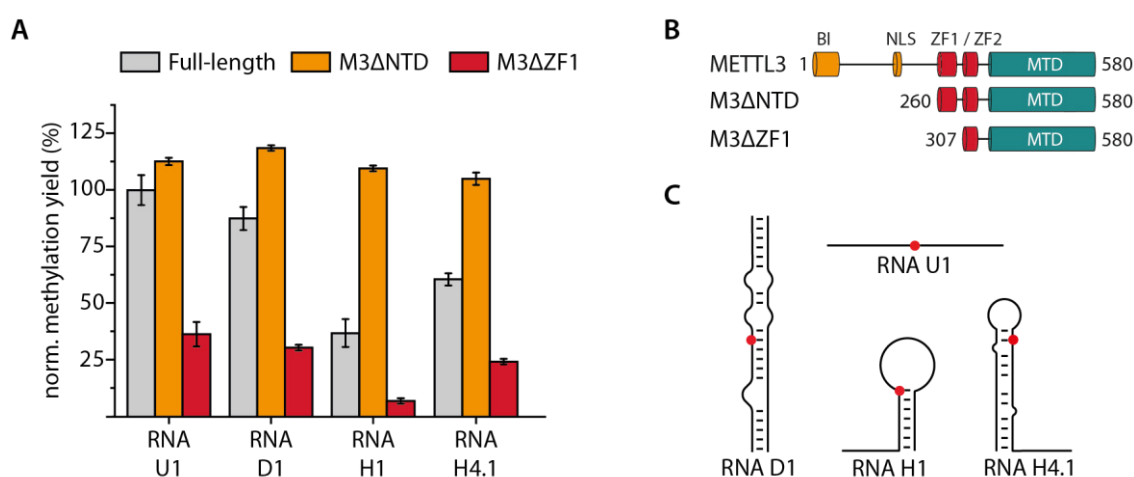


Figure 38: **A** Activity assay results on RNAs with different secondary structure with three different protein complexes. All methylation yields are normalized to the full-length complex with RNA U1 as target (absolute methylation yield 81 %). **B** (Truncated) METTL3 proteins used in complex with full-length METTL14 for activity assays. **C** Schematic representation of the differently structured target RNAs with indicated target adenosine location.

Together these results indicate that the structural specificity of the METTL3/METTL14 core complex is governed to some extent by the NTD in METTL3. One possible reason for this is a locally destabilizing effect on the structure surrounding the target adenosine upon deletion of the NTD. Moreover, the structural RNA binding mode and the overall complex dynamics may be altered in absence of the NTD. It is imaginable, that the ZFD as well as the RNA binding groove near the active site are more accessible and with that may enhance catalytic activity of the complex. This is also supported by the increased binding ability of the NTD shortened complexes (M3 Δ NTD, M3 Δ ZF1) towards the off-target RNA and the loose RNA duplex.

The results presented in chapter 3.1 and 3.2 were published in: Meiser *et al.* (2020): RNA secondary structure dependence in METTL3-METTL14 mRNA methylation is modulated by the

N-terminal domain of METTL3. *Biological Chemistry* 402 (1), S. 89-98. DOI: 10.1515/hsz-2020-0265.

3.2.5 Conclusion and outlook

To obtain insights on the impact of the individual METTL3/METTL14 protein domains on RNA catalysis, a set of six different truncated protein complexes was expressed and purified. Truncation of the proteins was mainly focused on the (postulated) RNA binding domains of the core complex. The methylation activity of the complexes was tested first on the unstructured RNA U1 and as a second step on the hairpin structured RNA H4.1, both with the optimal target sequence GGACU. Overall, deletion of the RNA binding domains resulted in a decreased methylation yield. EMSA experiments confirmed the RNA binding properties of the respective complexes being mostly in line with their enzymatic activity. For METTL14, the RNA binding RGG domain as well as the structurally important α -helical motif were identified as indispensable for efficient catalysis. For METTL3, the intact ZFD consisting of two individual zinc fingers was observed to be relevant for catalysis, however, a single zinc finger is sufficient for RNA binding. This suggests an unspecific RNA binding mode of the remaining zinc finger, supported by additional EMSA experiments. When bound to only the remaining zinc finger, the RNA is presumably not always oriented in the right direction for catalysis, which leads to a decreased methylation yield. Moreover, the ability of the full-length complex to efficiently bind an off-target RNA construct, hints towards a general RNA binding property of the complex, not only restricted to target sequences. It might as well be, that the RNA binding domains bind the RNA rather unspecifically and target specificity arises mainly from the structural organization of the active site. This interpretation is however not in line with the study of Huang *et al.*, where a sequence specificity of the isolated zinc finger domain was claimed.¹⁹

Comparison of the methylation yields on the unstructured RNA with those on the hairpin structured RNA showed the expected decrease in methylation efficiency for most truncated complexes. The RGG depleted complexes exhibited a higher sensitivity towards the structure of the tested RNAs than the ZFD depleted complexes. This again raised the question of a preferred RNA orientation for catalysis, in the case of the hairpin RNA H4.1 in favour of an RNA binding via the RGG domain. The specific methylation of RRACH sequences (and not their inversed counterparts) indicates a fixed RNA orientation in the active site of the methyltransferase complex during catalysis, which is closely related to the question on RNA orientation in the RNA binding domains. To investigate this further, the enzymatic activity assays and EMSAs of the

truncated protein complexes may be repeated with additional hairpin constructs, differing only in target sequence position. However, smFRET experiments with suitable labelling schemes would presumably give a better insight on this. Here, labelling of the RNA with one fluorophore of the FRET pair and labelling of the protein complex with the corresponding fluorophore would give insights into their relative orientation. Attachment sites on the RNA can be chosen at the 5'- or 3'-end of the target RNA. On the side of the protein complex, multiple attachment sites including the RGG and zinc finger domains would be interesting to label. Moreover, fluorophore attachment to the upper or lower side of the RNA binding groove could give an estimation of the preferred RNA orientation in the active site. Comparison between all labelling schemes then possibly yields an estimation on the relative molecular organisation of the RNA binding domains. However, the enzymatic activity of the differently labelled complexes would need to be verified prior to all smFRET measurements to rule out disturbing effects of the attached fluorophores.

The most unique behaviour was observed for the methyltransferase complex shortened at the NTD in METTL3. Here, a significant increase in methylation efficiency was observed on all tested RNAs including the unstructured RNA, a loose duplex and two hairpin structures. With this, no sensitivity towards the structural context of the target RNA could be observed in contrast to the full-length complex. This renders the impact of the target RNA structure on enzymatic activity of the METTL3/METTL14 complex to originate at least in part from the NTD in METTL3. This might be related to an increased ability to locally destabilize the structure surrounding the target adenine or a generally altered RNA binding mode. Moreover, the lack of the NTD in METTL3 presumably has impact on the dynamics of the protein complex as well as the accessibility of the RNA binding domains (especially ZFD and positively charged binding groove). To distinguish between a lacking sensitivity of the structural RNA context in general and the possibility of a faster methylation process exhibited by the NTD depleted complex, the performed activity assays would need to be adapted. It might be that the structural sensitivity of the M3 Δ NTD complex is just not resolved under the given conditions. A higher time resolution or the use of a multiple turnover approach would presumably give further insight into the observed differences.

The NTD in METTL3 harbours the binding interface for WTAP²⁰, which itself acts as binding platform for other accessory proteins.²³⁻²⁵ This raises the question how providing the binding partner WTAP, in contrast to depletion of the NTD, influences the structural sensitivity of the complex, which was investigated in in the next step.

3.3 Influence of WTAP

Based on the results described in the preceding chapters, the role of the NTD domain in METTL3 was an interesting target to investigate further. Therefore, a set of new, extended protein complexes was designed where the core complex (METTL3/METTL14) should be expressed together with the accessory proteins (VIRMA, ZC3H13/ RBM15) with and without WTAP as mediator. Moreover, a second isoform of WTAP (shortWTAP)¹³⁹ resulting from alternative mRNA splicing was added to the constructs. The planned set of complex compositions can be found in Table 3.

Table 3: Complex compositions of the planned complexes with accessory proteins.

Complex	METTL3 / METTL14	WTAP	ZC3H13 / RBM15	VIRMA
A1	+	-	-	-
A2	+	+	-	-
A3	+	isoform 2	-	-
AB-11	+	-	+	-
AB-21	+	+	+	-
AB-12	+	-	-	+
AB-22	+	+	-	+
AB-13	+	-	+	+
AB-23	+	+	+	+

In the previous work, the proteins METTL3 and METTL14 were coexpressed, however, from two individual plasmids. With the new set of protein complexes, a different expression and purification strategy was applied to overcome the copurification of different monomers. The protein coding genes of the desired complex were first cloned into one single vector backbone (see chapter 3.3.1). In these constructs, only METTL14 was designed with a His-tag for purification to copurify complex bound proteins only. All other proteins were extended with a FLAG- or Strep-tag to have orthogonal purification and detection options.

3.3.1 Generation of expression constructs (Gibson assembly)

Synthesis of the new expression constructs with different complex compositions (Figure 39A) was performed by applying the biGBac method described by Weissmann *et al.*¹⁴⁰ This method is based on a set of vectors and primers optimized for rapid construct generation by Gibson assembly. To start with, all individual protein coding genes were amplified by PCR with the specified primers required for the desired gene assemblies (details see Table 19 in chapter 5.5.3.2). Successful amplification was achieved for all genes of interest except for VIRMA as shown in the analytical agarose gel (Figure 39B). The main visible DNA bands were matching the expected PCR product lengths, except for VIRMA where a strong smaller band (approx. 1 kbp) was visible. Although annealing temperature and elongation time were varied several times, the

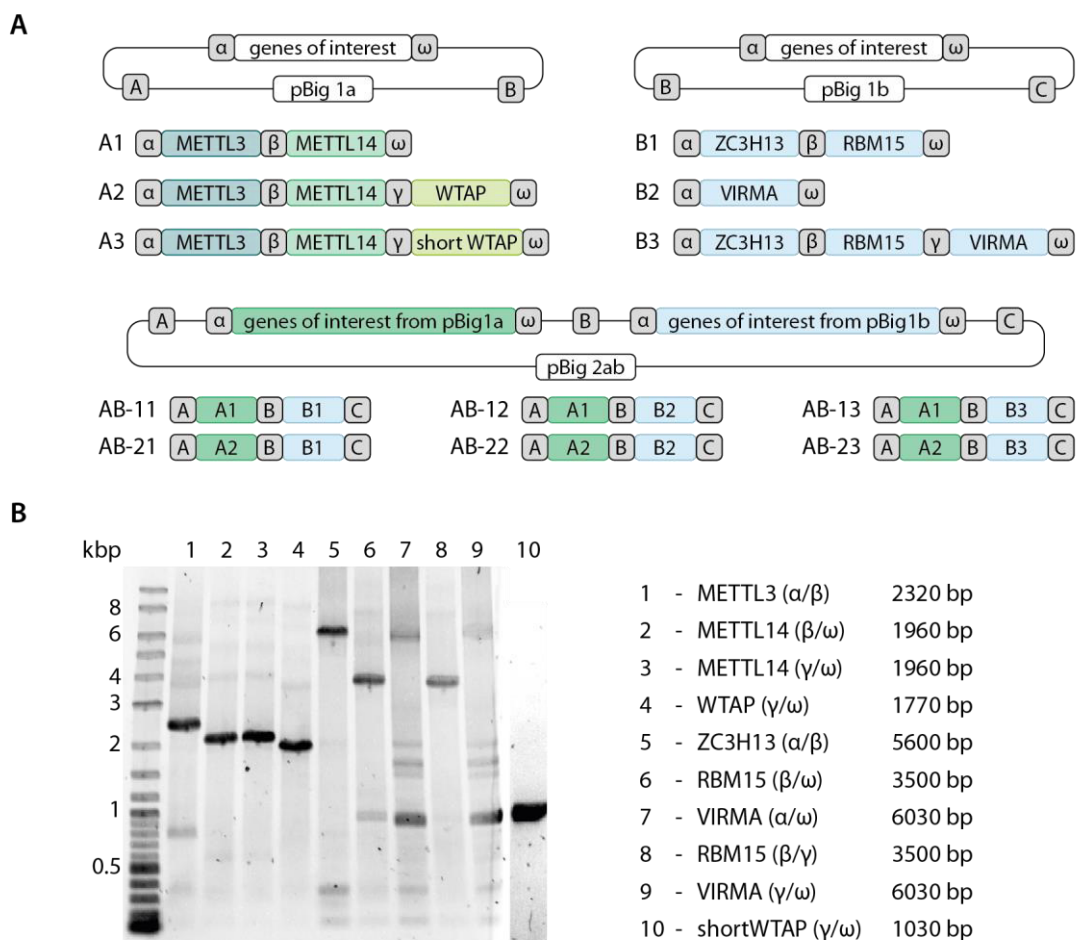


Figure 39: A Gene assemblies in pBig vector backbones that were planned for the formation of several expression constructs. Finally, only gene assemblies A1, A2, A3, B1, AB-11 and AB-21 were successfully formed. **B** Agarose gel (0.8 %) of the PCR reaction for all genes preceding the Gibson assembly reaction.

PCR reaction for VIRMA continued to fail (data not shown). For the following steps, I decided to continue the synthesis of all gene assemblies without VIRMA and express this accessory protein separately.

In the next step, Gibson assemblies were performed to combine the genes that were successfully amplified by PCR into one vector. First, the proteins were divided into two groups and assembled into two different vector backbones (pBig1a/pBig1b, Figure 39A). After isolation from the *E. coli* cells, all plasmids were analyzed by analytical digestion (Figure 40) and verified by Sanger sequencing (*Eurofins Genomics*). Here, plasmids for combination A1, A2, A3 and B1 were successfully obtained. For the larger plasmid formation with up to five individual genes, the gene cassettes from construct A1, A2 and B1 were released from the vector by digestion (PmeI) and inserted in both combinations (AB-11 and AB-21) into the pBig2ab vector backbone (Figure 39A). Again, both plasmids were analyzed by analytical digestion and verified by Sanger sequencing. The analytical digestion of all successfully cloned and purified expression constructs is shown in Figure 40. During analysis, digestion with the Swal enzyme releases the individual genes from the vector whereas digestion with PmeI cuts out the whole gene cassette. The analysis for the presence of the individual genes showed the desired genes for all constructs. However, the gene band for METTL3 in A3 was not clearly visible. Moreover, the ZC3H13 fragment has almost the same size as the vector backbone, making it difficult to prove its presence via analytical digestion with the Swal enzyme. To overcome this problem, the size of the whole gene cassette was tested in addition. As Gibson assembly into pBig2ab destroys the PmeI restriction site, this analysis was only done for constructs with the pBig1 vector backbone. For B1 a DNA band in the correct size range (9.5 kbp) was visible on the agarose gel in addition to the vector backbone (5.5 kbp). As the size of the gene cassettes for A1 (4.7 kbp), A2 (6.5 kbp) and A3 (5.8 kbp) are very close to the size of the vector backbone it is hard to tell where the

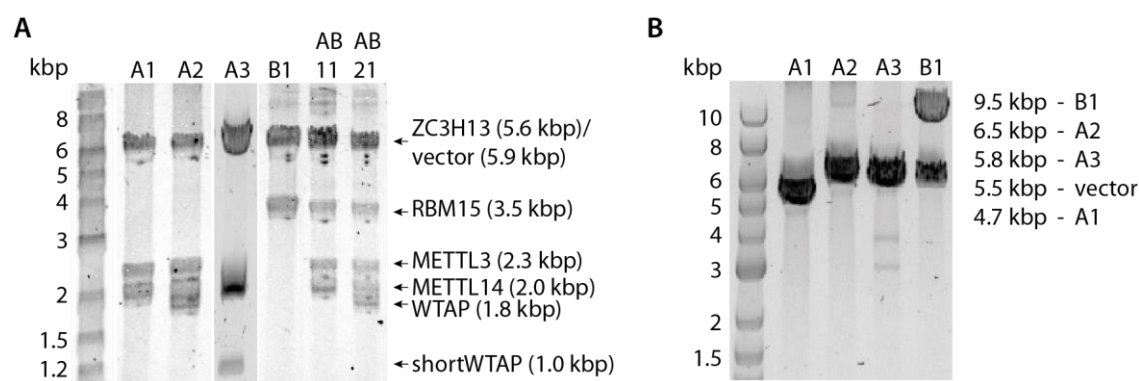


Figure 40: Analytical digestion of the Gibson assembly products with Swal (**A**) and PmeI (**B**).

DNA band originates from. However, no additional DNA band was visible and small size differences were observed. Together with the correct sequencing analysis, this ensured the right size and the correct ratio of the inserted genes for all constructs. With that, the gene assemblies A1, A2, A3, AB-11, AB-21 and additionally the original VIRMA plasmid were used for subsequent expression and purification.

3.3.2 Protein preparation

For expression in Sf9 cells, a recombinant baculovirus was produced and amplified from each plasmid of the accomplished set of expression constructs, including VIRMA. With these constructs, test expressions were performed with the generated P₂ baculovirus respectively. A successful protein expression was achieved for gene assemblies A1, A2 and A3 where the desired protein bands for METTL3, METTL14, WTAP and shortWTAP could be observed in the western blot analysis (Figure 41). All proteins run at a higher molecular weight reference which was observed also for previous expressions. An additional band at a reference size of approximately 50 kDa was visible in all three FLAG-tag-stained western blots, indicating either slight METTL3 degradation or undesired METTL3 termination fragments. Unfortunately, no protein expression was observable for constructs AB-11, AB-21 and VIRMA (data not shown).

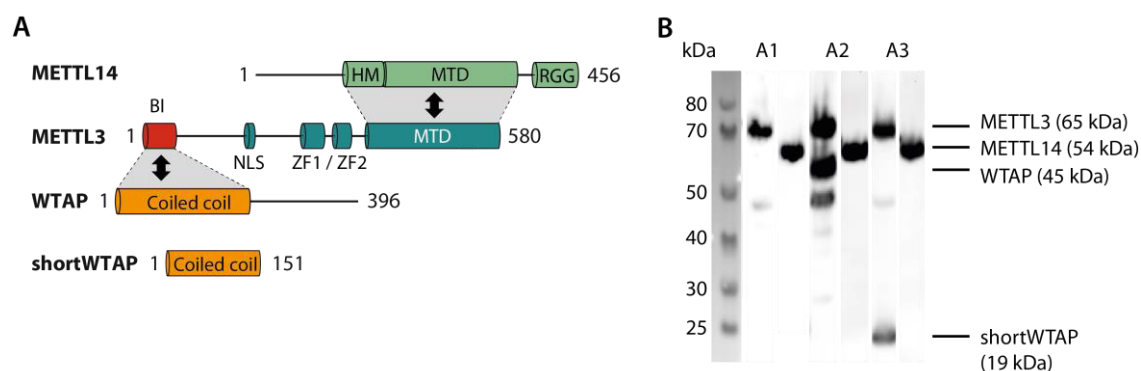


Figure 41: **A** Schematic representation of the expressed proteins. **B** Western blot stained with antibodies against FLAG-tag (1st, 3rd and 5th lane) and His-tag (2nd, 4th and 6th lane) from test expressions of all three expression constructs (A1: METTL3/METTL14, A2: METTL3/METTL14/WTAP, A3: METTL3/METTL14/shortWTAP).

After the large-scale expression for the METTL3, METTL14, WTAP and shortWTAP containing constructs, purification of the different complex compositions was performed with the established protocol. In contrast to the proteins expressed from the pTriEx vector system, only

METTL14 is tagged with a His(8) purification tag whereas the other proteins are expressed with a FLAG-tag. Therefore, only proteins that firmly bind to METTL14 or one of its binding partners are copurified during the protein purification described in the following sections.

3.3.2.1 METTL3/METTL14 complex

First, the METTL3/METTL14 complex was newly expressed and purified to have a direct comparison in the following experiments. With Ni-NTA affinity chromatography (Figure 42A), the proteins were separated from the cell lysate. The analytical SDS-PAGE (Figure 42B) shows the presence of both complex partners in the analyzed fractions. Excess METTL14 was observed, which is however in line with METTL14 being the only complex partner that carries a His-tag. The protein containing fractions (2 to 10 cv) were combined and concentrated for further purification.

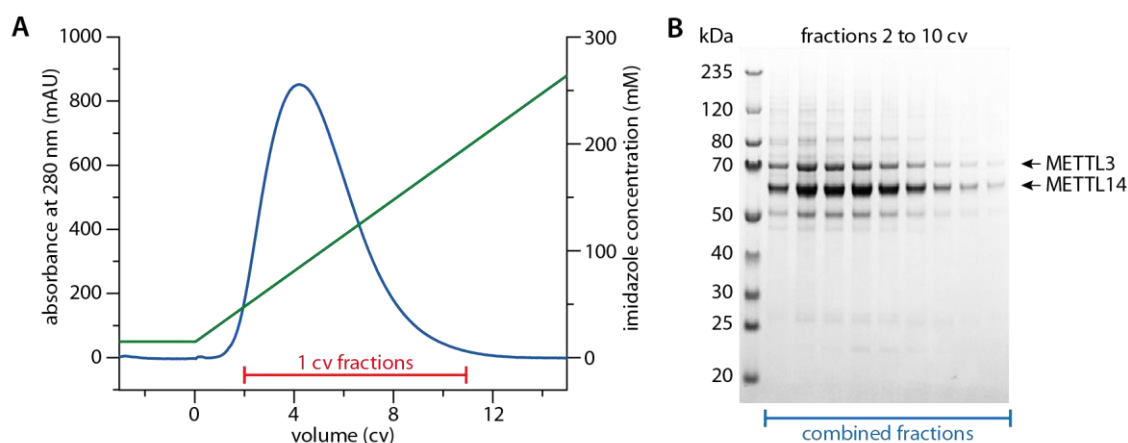


Figure 42: HisTrap purification of the METTL3/METTL14 complex from the pBig vector system. Here, only METTL14 carries a His-tag for purification. **A** Baseline corrected chromatogram with absorption at 280 nm (blue line), imidazole gradient (green line) and indicated analyzed fractions. **B** SDS-PAGE of the analyzed elution fractions. Combined fractions for further purification are indicated.

To get the pure complex especially without the excess METTL14 monomers, a size exclusion chromatography was performed. The resulting chromatogram shows a clear main peak that was analyzed by SDS-PAGE (Figure 43). The gel shows the presence of both proteins and a good separation from the METTL14 monomers. However, the protein band at a reference size of 50 kDa is still visible and is coeluting with the complex. This was already observed in the test expressions and presumably originates from either slight METTL3 degradation or METTL3 termination fragments. In favor of the latter, METTL3 degradation was not observed for the

expression and purification from the pTriEx vector system (see chapter 3.1.1). The resulting sample purity of the full-length METTL3/METTL14 complex is still suitable for further experiments. Fractions at 10 and 11 mL retention volume were combined, concentrated to a protein concentration of 27 to 84 μ M and flash frozen in liquid nitrogen. The 260 nm/280 nm absorption ratio of all samples was determined to be around 0.6 for a nucleic acid free protein sample.¹³⁰

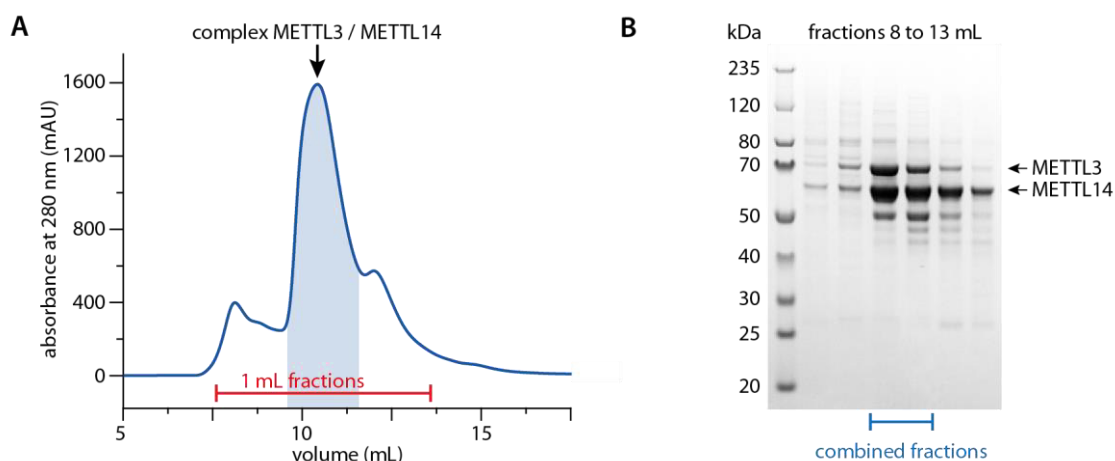


Figure 43: Size exclusion chromatography of the METTL3/METTL14 complex. **A** Chromatogram with absorption at 280 nm (blue line) and indicated analyzed (red) and combined (blue) fractions. **B** SDS-PAGE of the analyzed fractions. The combined fractions that were concentrated for use in further experiments are indicated.

3.3.2.2 METTL3/METTL14/(short)WTAP complex

The METTL3/METTL14 complexes expressed together with WTAP or shortWTAP were purified following the same protocol. After Ni-NTA affinity chromatography (the chromatograms can be found in Figure 82 and Figure 83 on page 176), the proteins coeluting with METTL14 were analyzed via western blot (anti-FLAG staining) to ensure the presence of all desired proteins after separation from the cell lysate. The western blot shows that only the full-length WTAP binds firmly enough to the complex to be copurified (Figure 44). Although both isoforms are clearly present in the cell lysate, the shorter WTAP construct was only detected in the flowthrough and not in the elution as desired. This indicates that this isoform is not able to firmly bind to the METTL3/METTL14 core complex, which is in line with Schöller *et al.*²⁰ This in the best case weak binding of the short isoform renders the purification as complex challenging and would also raise the question whether the trimeric complex would form in further experiments.

Attempts to purify the complex via the FLAG-tags in METTL3 and the shorter WTAP failed (data not shown). We therefore decided to continue with only the full-length WTAP as part of the extended complex.

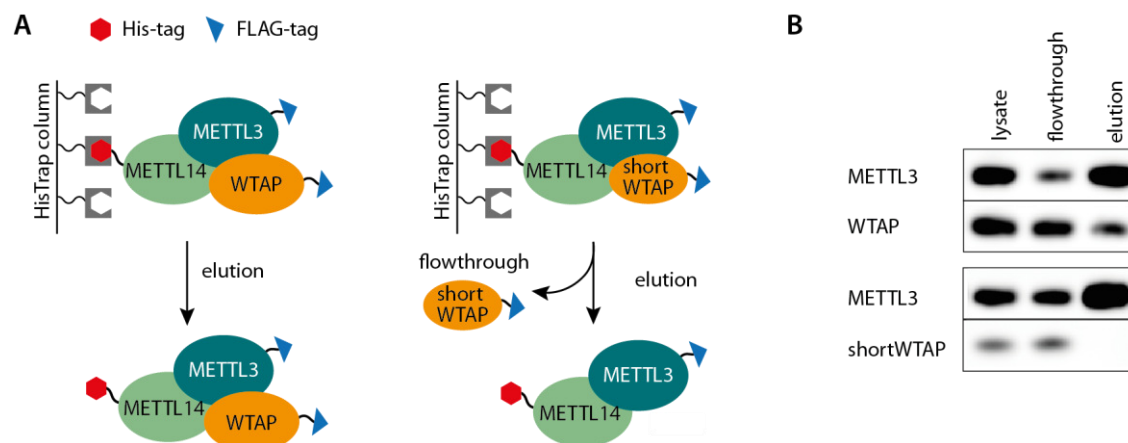


Figure 44: **A** Schematic representation of the Ni-NTA purification as ternary complex, with only METTL14 carrying a His-tag. **B** Western blot (anti-FLAG staining) of the different purification steps for the METTL3/METTL14 complex with WTAP or shortWTAP as complex partner.

To further purify the METTL3/METTL14/WTAP complex, size exclusion chromatography was performed. The protein containing fractions were analyzed via SDS-PAGE (Figure 45A,B) and the fractions at 10 and 11 mL were combined and concentrated to 62 μ M. Again, the 260 nm/280 nm absorption ratio of the samples was determined to be around 0.6. Purified METTL3/METTL14 samples with and without WTAP were analyzed via western blot (Figure 45C). The western blot shows that for both samples all the desired proteins are present. For the METTL3/METTL14/WTAP complex, the METTL3 band in the FLAG-stained western blot is significantly more intense than the WTAP band. This indicates that the complexes with and without WTAP both are present in the sample and could not be separated on the size exclusion column. However, this still results in one sample containing the WTAP assembled complex, although WTAP dependent effects may be not as pronounced in the following experiments as for a sample that exclusively contained the pure METTL3/METTL14/WTAP complex.

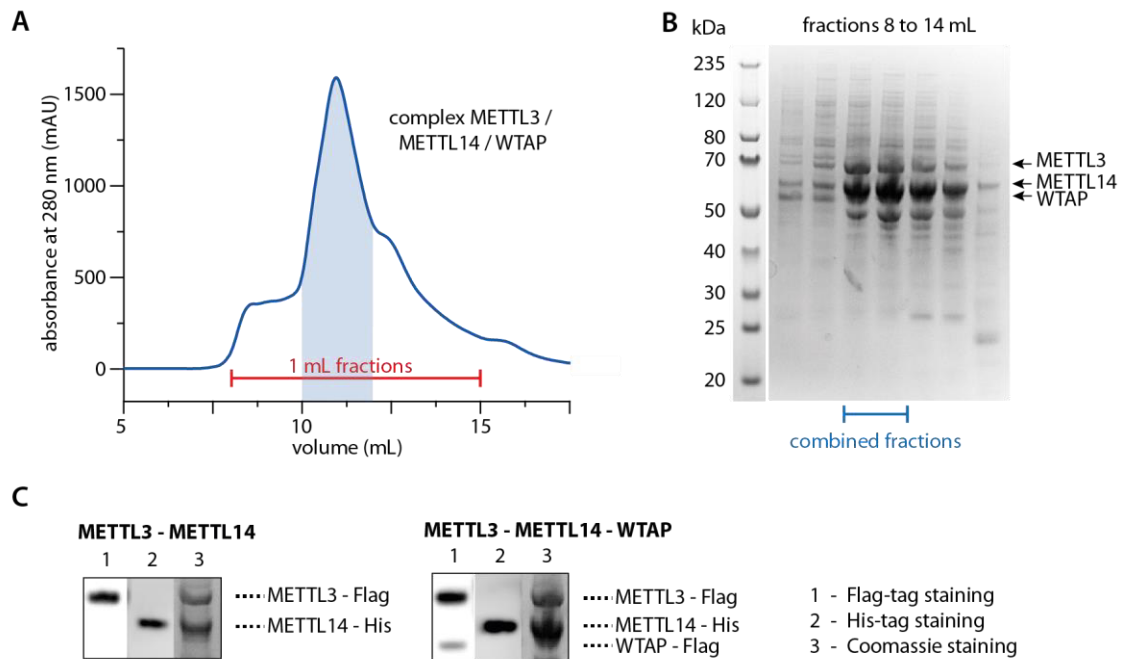


Figure 45: Size exclusion chromatography of the METTL3/METTL14/WTAP complex. **A** Chromatogram with absorption at 280 nm (blue line) and indicated analyzed (red) and combined (blue) fractions. **B** SDS-PAGE of the analyzed fractions. The combined fractions that were concentrated for use in further experiments are indicated. **C** Western blot analysis of the purified METTL3/METTL14 complex with and without WTAP in comparison to SDS-PAGE with Coomassie staining.

3.3.3 Secondary structure specificity

Before moving on to the methylation experiments with WTAP, the methylation behavior of the METTL3/METTL14 complex from the previous and the new expression system were compared (pTriEx vs pBig vector system). Here, also the complex from the new expression system showed robust methyltransferase activity and even a slightly increased methylation efficiency was noted (90 % absolute methylation yield). Following the described role of the NTD in METTL3, increased sensitivity towards the RNA structure of the WTAP complex was expected. Therefore, the assay conditions were further optimized by lowering the SAM concentration (see Figure 75 in appendix (page 169), refer to chapter 3.2.2) to shorten reaction times and still be able to capture small activity changes.

To assess the enzymatic methyltransferase activity of the METTL3/METTL14/WTAP complex, the activity assay was repeated for RNA U1, D1, D3 and H1. For better comparison, the activity assays were not only performed with the METTL3/METTL14/WTAP complex but also with the

METTL3/METTL14 complex to have a direct reference point under the same conditions (reaction conditions and expression/purification system). The resulting methylation yields (Figure 46) for the structured RNA targets show a significant difference for the presence of WTAP in the complex. For the unstructured RNA U1, no difference between the two complexes was detectable. This was confirmed with experiments where the reaction time was reduced to one third. For the loosely base paired duplex D1, a significant decrease in methylation yield was observed for the METTL3/METTL14/WTAP complex in comparison to the METTL3/METTL14 complex. Strikingly, the expected lower methylation for the METTL3/METTL14 complex on RNA D1 compared to RNA U1 was not detected. Most likely, the optimized reaction conditions are not suitable to resolve these differences. This is supported by the results on the hairpin RNA H1, where in previous assays a maximum normalized methylation yield of 45 % was detected (81 % in this case). However, the data shows that the conditions are suitable to resolve the differences between the complexes with and without WTAP. The observed decrease in methylation efficiency for the WTAP containing complex with RNA D1 as target was also detected for RNA H1 with a slightly more pronounced relative decrease for RNA H1. Methylation experiments on RNA D3, representing a tightly duplexed RNA, yielded no detectable m⁶A formation for both complexes, as expected. This data indicates an even more pronounced sensitivity of the METTL3/METTL14/WTAP complex towards structured RNAs with the same trend as already observed for the core METTL3/METTL14 complex.

To have a better comparability with methylation results of the previous experiments, the observed methylation yields were calculated back to the results presented in chapter 3.1.4. Transferring the relative decrease down to 76 % (RNA D1) and 63 % (RNA H1) of the methylation

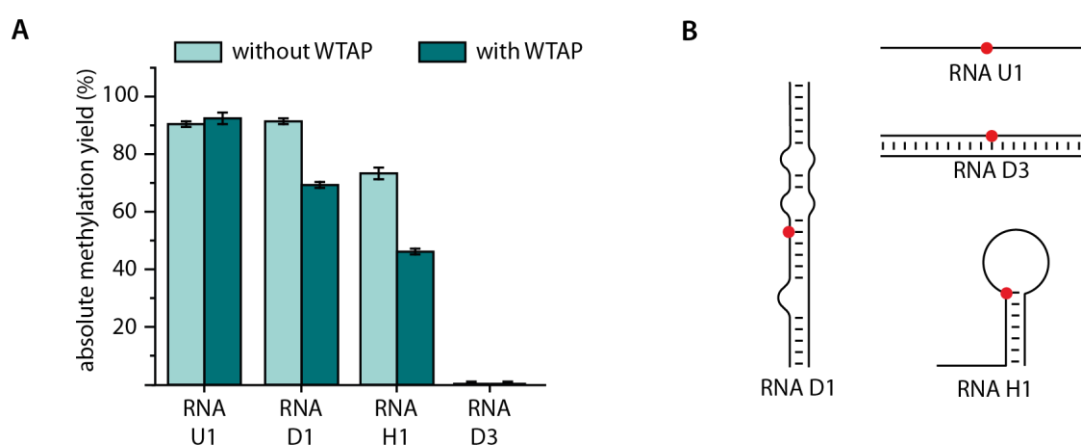


Figure 46: **A** Activity assay results on RNAs with different secondary structure methylated with the METTL3/METTL14 complex with and without WTAP. Absolute methylation yields are given. **B** Schematic representation of the structured target RNAs with indicated adenosine location.

yield from complex METTL3/METTL14, this gives an estimated normalized methylation yield of 57 % for RNA D1 (0.76x75 %) and 28 % for RNA H1 (0.63x45 %) for methylation with the WTAP containing complex. Although these normalized methylation yields for the METTL3/METTL14/WTAP complex are a very rough estimation, they undoubtedly represent the robustly observed trend, summarized in Table 4. The core complex itself shows a clear secondary structure sensitivity towards its target RNA during methylation. This effect is diminished upon depletion of the NTD domain in METTL3, which harbors the binding interface for WTAP. In contrast, “extending” the NTD domain by addition of the binding partner WTAP increases the sensitivity of the complex towards target RNA structure.

Table 4: Summarized methylation yields for different complex compositions with diverse RNA targets. (*estimated normalized methylation yields based on results shown in Figure 26 and Figure 46)

Normalized methylation yields for	RNA U1	RNA D1	RNA H1	RNA D3
M3ΔNTD / METTL14	112 %	118 %	109 %	n.d.
METTL3 / METTL14	100 %	75 %	45 %	0 %
METTL3 / METTL14 / WTAP	102 %	57 %*	28 %*	0 %

As already stated for the complex that lacks the NTD in METTL3, this can be related to several possible changes in the complex. First, there might be an altered ability to destabilize the local target adenine containing structure. In this case, the complex without the NTD would exhibit an increased ability for destabilization while in the case of WTAP binding this ability would be decreased. The described differences could also be based on changes in the RNA binding mode or the dynamics in the protein complex. This both is most likely related to the structural organization of the complex and with that also to the accessibility of the RNA binding domains. All of this might also be influenced by other proteins of the complex that were shown to effect m⁶A formation *in vivo* (VIRMA, ZC3H13, RBM15, HAKAI).^{22,23,25,26}

3.3.4 Conclusion and outlook

In this chapter, the role of additional proteins (in particular WTAP) on the sensitivity of the METTL3/METTL14 core complex towards the target RNA secondary structure was investigated. To achieve this, DNA plasmids for different complex compositions were prepared via Gibson

assembly followed by small scale test expressions. Successful expression was achieved for complexes consisting of the METTL3/METTL14 core with and without WTAP in two different isoforms. The new expression and purification technique was designed to purify the proteins directly as a complex, with only METTL14 carrying a His-tag for purification. The shorter WTAP isoform was readily coexpressed but did not bind firmly enough to the METTL3/METTL14 heterodimer to be successfully copurified. In contrast, purification of the enzymatically active METTL3/METTL14 and METTL3/METTL14/WTAP (full-length) complexes was achieved with this method.

After purification of the two complexes, enzymatic activity assays were repeated for the unstructured RNA U1, the loose duplex D1, the duplex with full complementarity D3 and the hairpin structured RNA H1. Similar methylation results were observed for the complexes with and without WTAP on the unstructured RNA with nearly quantitative methylation and for the tight duplex D3 with no detectable m⁶A formation. The methylation of the two remaining RNA constructs, however, uncovered the increased sensitivity of the WTAP containing complex towards RNA secondary structure in comparison to the METTL3/METTL14 core complex. With reference to the diminished structural sensitivity of the M3ΔNTD complex, this assigns the NTD in METTL3 an important role in secondary structure dependent methylation of target RNAs. As already discussed in the preceding chapters, this might be related to changes in the general RNA binding mode, the ability to locally destabilize the target structure, the dynamics in the complex or the structural organization and with that the accessibility of the RNA binding domains. Additional EMSA experiments especially with the METTL3/METTL14/WTAP complex and RNAs of different secondary structure might give further insight into the RNA binding properties.

The observed strong RNA secondary structure dependence of the methylation process of the methyltransferase complex *in vitro* is presumably modulated by several factors *in vivo*. Besides the region-selective methylation conducted by additional accessory proteins of the complex, these proteins might also alter the complex characteristics. This is presumably related to the structural sensitivity towards target RNAs. This extends also towards the cotranscriptionality of the methylation and proteins that were shown to recruit the complex to specific cell processes such as transcription or polyadenylation.^{15,25,47–49} Moreover, the influences already discussed in chapter 3.1.5 are likewise valid for the WTAP containing complex. It is highly interesting if the here reported structure dependent methylation efficiency is mirrored in the transcriptome wide distribution of m⁶A sites.

4 Results and discussion - smFRET studies on the METTL3/METTL14 complex

4.1 smFRET setup construction

The used smFRET setup (Figure 47) was constructed by Julius Blechar, Martin Hengesbach and myself taking into account protocols and manuals from literature.^{84,85,108,141} It can be divided into four main sections: excitation, microscope, detection and remote control. The whole setup was built on an optical breadboard on a table frame with passive isolation mounts to prevent vibrations on the setup. First, the excitation beam path was built bringing all three lasers into a

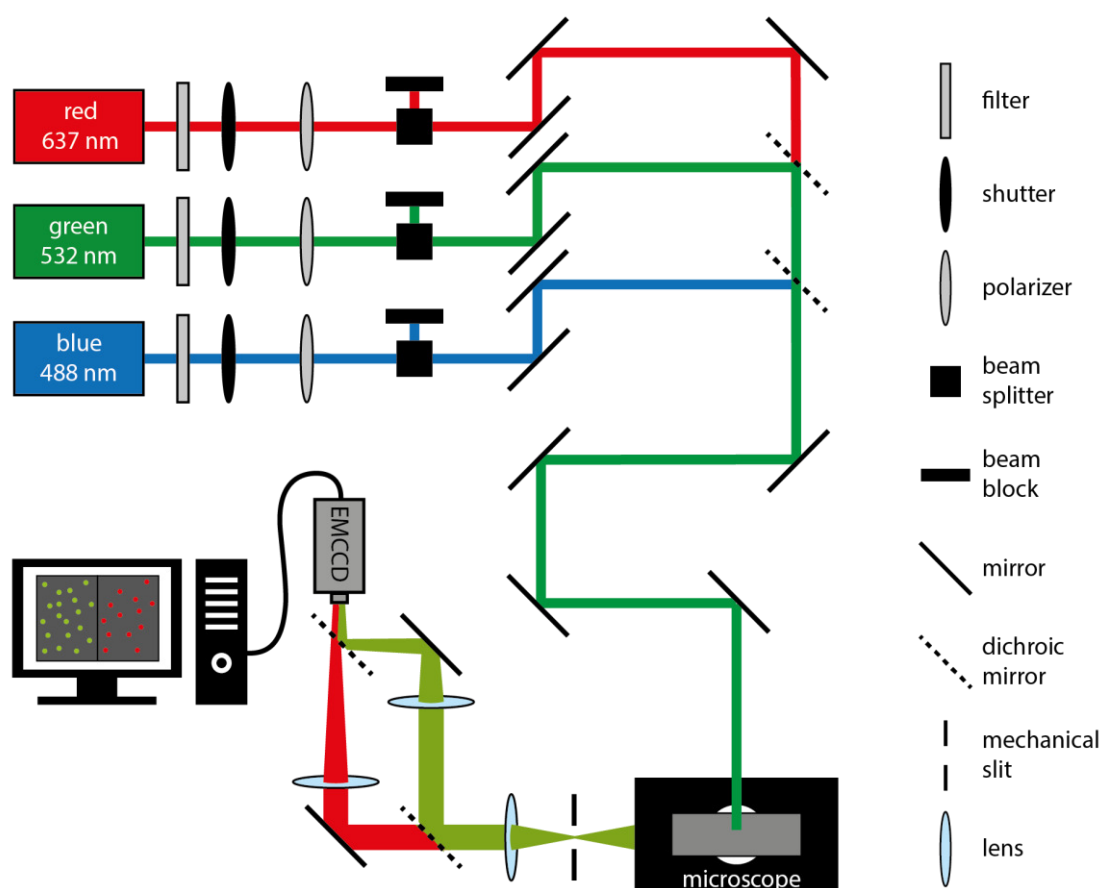


Figure 47: Schematic representation of the complete smFRET setup. The running lasers can be controlled by opening and closing the shutters. The laser intensity range is controlled with the filter, polarizer, and beam splitter. All three lasers are brought in a parallel beam path onto the slide on the microscope (here only shown in green for better overview). For detection, the light from the sample slide is filtered inside the microscope to block excitation light (not shown). Emitted wavelengths were separated with a dichroic mirror and focused onto different areas of the EMCCD camera.

parallel beam path and onto the microscope. The inverted light microscope was adapted to realize a prism based TIRF approach. Here, the parallel laser light was focused onto a prism to get total internal reflection on the sample slide. To analyze the sample, the emitted light was separated into two channels and focused onto an EMCCD camera. The detected image is then shown on the attached computer system and subject to further data analysis.

The main construction sections are described in more detail in the following chapters. For details on the technical equipment used see chapter 5.10.1. During the whole setup construction, appropriate laser safety goggles were worn.

Additionally, several safety precautions were installed to rule out undesired leakage of laser light and with that dangerous eye injuries. First, the room was completely darkened by foiling all windows with lightproof, matt foil. This is also necessary to shield the setup from stray light. Moreover, each section was individually covered in black hardboard, with only the necessary holes for the laser beam to pass onto the next section. Furthermore, all shutters were set to closed state by default when switching on the laser setup.

4.1.1 Excitation

The first section to assemble was the excitation beam path. The lasers were chosen according to the requirements for the most widely used FRET fluorophore pairs. Therefore, one red (637 nm) and one green (532 nm) laser were selected together with a blue (488 nm) laser for further applications. All three lasers can be remotely controlled with the computer software supplied by the manufacturer. However, the lasers need some time to assume a constant, stable laser emission after being switched on. For fast on- and off-switching of the laser light in the beam path, a (remotely) fast switchable shutter was placed at the start of each laser beam path. This enables fast (scale of 10 ms) switching between the different wavelengths focused on the sample slide.

To control the intensity of the laser light on the sample, a filter wheel together with a half-wave plate and a beam splitter was mounted into the laser beam path each. First, the laser intensity range is determined by the filter wheel, where 5 different neutral density filters can be chosen besides a blank spot. The laser light intensity is weakened depending on the density of the used filter. In the next step, the polarization plane of the linearly polarized light is rotated according to the adjustable half-wave plate (Figure 48). The light then passes the beam splitter, where only its p-polarized component is transmitted into the downstream beam path while the s-

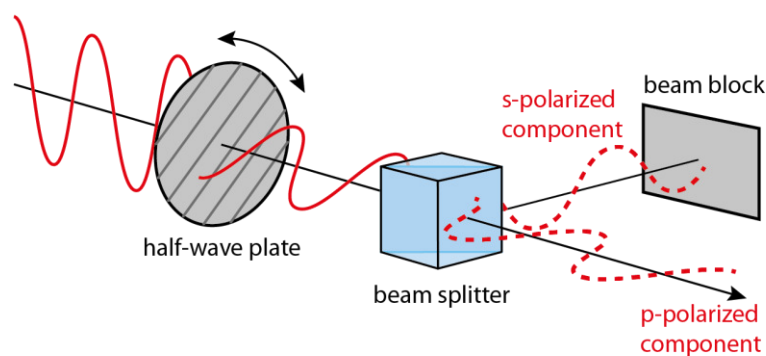


Figure 48: Schematic representation of the setup for the manual laser intensity adjustment. The polarization plane of the linearly polarized laser light is rotated through the adjustable half-wave plate. Afterwards, the beam is split into its s- and p-polarized components, resulting in a p-polarized beam for the following excitation path. The laser intensity is regulated by adjusting the half-wave plate.

polarized component is reflected into the nearby beam block. Depending on the adjusted polarization plane, the proportion of the s- and p-polarized components vary. With that, the laser intensity can be precisely adjusted, resulting in a linear p-polarized laser beam individually for each laser wavelength. Moreover, all three lasers can be modified in intensity using the corresponding computer software for remote control. As the effective laser intensity is important for the measurement conditions, a digital power and energy console (silicon sensor) was mounted onto a sliding construction and placed in the beam path in front of the microscope desk. In this way, the effective laser intensity can be easily measured for each wavelength.

In the next step, the three lasers were brought into a parallel beam path (Figure 49) which is necessary to focus all onto the same spot on the sample slide. First, the red laser beam was adjusted to a horizontal line, passing a pinhole in short-distance and falling onto the center of a long-distance target (approx. 5 m distance from the pinhole). Afterwards, a dichroic mirror was mounted into the red laser beam for each additional wavelength. The mirrors were positioned to transmit the red laser beam in the middle of the mirror plate but also to reflect the additional laser beam at the same point with about a 45° angle. It is important to keep the laser order sorted from higher to lower wavelength, as the dichroic mirrors only transmit light lower than the given cut off. To align the additional lasers onto the red beam, the reflection points on the dichroic mirrors were adjusted to match the transmission point of the red laser beam. This was achieved by fine tuning the preceding beam path of the green and blue laser with two additional mirrors each. To have a completely parallel beam path, the dichroic mirrors were tilted to vary the vertical and horizontal reflection angle while keeping the reflection point the same. With

that, all three laser points were adjusted to pass the pinhole and strike exactly the same point on the long-distance target. The resulting parallel beam path was then brought into the correct height and direction to be reflected on the mirror on the microscope desk.

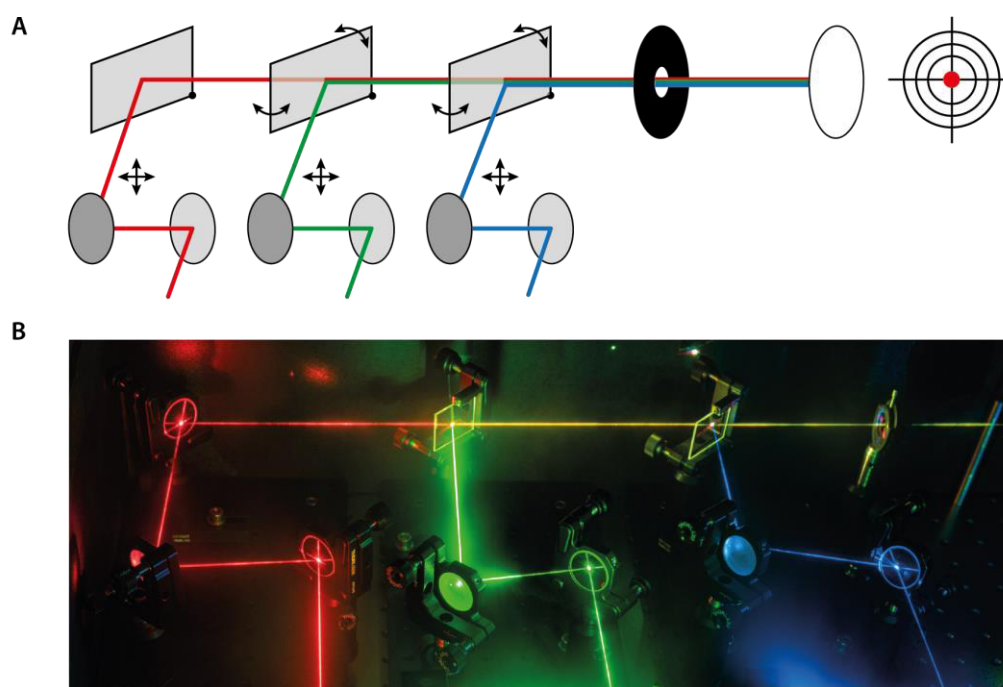


Figure 49: **A** Alignment of all three laser beams to a parallel beam path. Reflection points of the green and blue laser were adjusted with the two preceding mirrors to match the transmission point of the red laser on the dichroic mirrors. The dichroic mirrors were tilted to fine tune the horizontal and vertical reflection angle to build a parallel beam path. A pinhole in short-distance and a target in long-distance were used to verify the beam parallelism. **B** Picture of the setup with three lasers aligned to a parallel beam path.

4.1.2 Microscope

The basis of the microscope is formed by an inverted light microscope (Olympus IX73, Olympus) with a water immersion objective (UPLSAPO 60XW, NA 1.2, Olympus). The sample slide is fixed above the objective on the microscope desk. As this setup was built for a prism based TIRF approach, the sample slide is always mounted upside down. To create total internal reflection on the sample slide, the prism was mounted right on top of the sample slide. In this setup, the prism has to be in optimal contact with the sample slide which requires a fine adjustment of the height and vertical position of the prism. However, after being perfectly adjusted to the

individual sample slide, the prism position has to be fixed. This is necessary to allow independent horizontal movement of the sample slide under the prism without altering the evanescent field. Considering these requirements, a bridge was designed to locate the prism in the correct position (Figure 50). The middle of the bridge was constructed to be above the center of the objective. However, anchoring of the bridge was planned to leave room for fine adjustments of 9 mm in the horizontal direction. The prism itself was mounted into a prism holder that is clamped with a screw into the prism bridge. Loosening the screw enables optimal contact between prism and sample slide. After adjusting its position and fit, the prism position is fixed by tightening the screw. For frictionless sliding of the sample slide under the fixed prism, immersion oil was used between the two materials to rule out damage of the prism surface, but most importantly to prevent refraction on the interface.

After correct positioning of the prism over the center of the objective with optimal contact to the sample slide, the laser beam was adjusted to the prism to obtain total internal reflection on the sample slide.

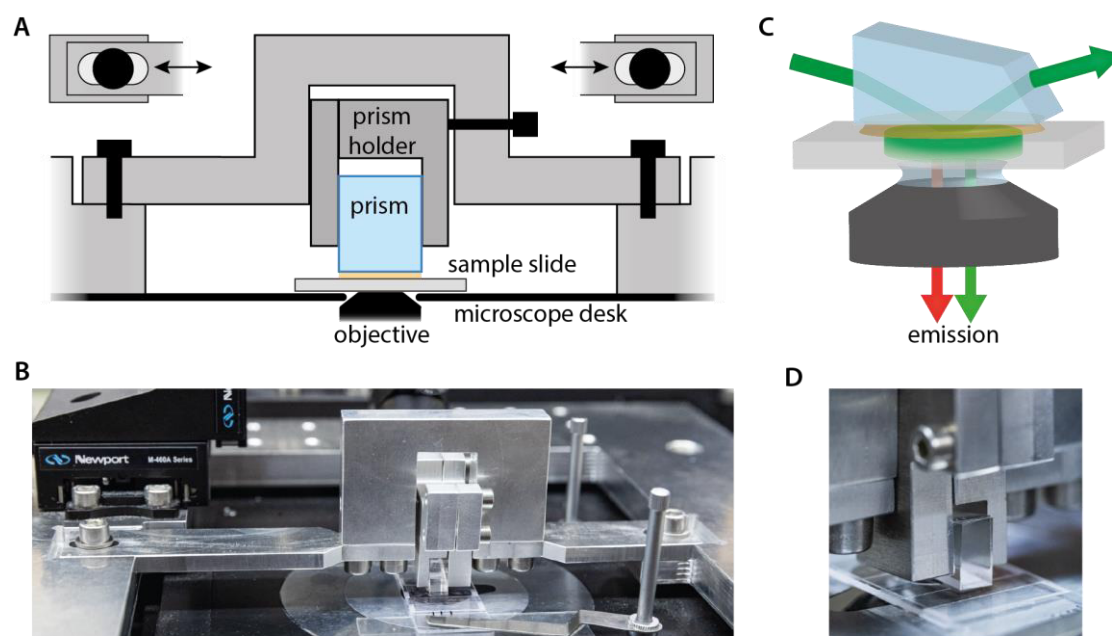


Figure 50: Schematic representation (A) and picture (B) of the prism bridge with adjustable positioning. The prism is clamped in a prism holder into the middle of the prism bridge, located above the center of the objective. C The prism is in optimal contact with the sample slide to obtain a stable evanescent field. Immersion oil between the materials enables frictionless adjustment of the sample slide. The water immersion objective passes the emission fluorescence into the detection section. D Detailed picture of the adjusted prism.

4.1.3 Creating of total internal reflection

The essential part of a TIRF microscope is to create an evanescent field on the sample slide. To obtain total internal reflection, the conditions described in chapter 1.3.3 have to be fulfilled.¹⁰⁹

In the given case, total internal reflection is desired on the interface between the quartz slide (UVFS glass) and the sample, mainly consisting of water. With that, the critical angle for total internal reflection (α_c) is given with:

$$\alpha_c = \sin^{-1} \left(\frac{n_{\text{water}}}{n_{\text{UVFS}}} \right) \quad (6)$$

On the microscope, the sample slide is covered by immersion oil with the same refractive index and a prism on top. The sample slide is made from the same material as the prism that is why refraction on the prism/immersion oil interface and the immersion oil/sample slide interface is negligible.

However, before being reflected on the slide/water interface, the laser beam has to pass the air/prism interface to enter the prism. Here, the laser light is refracted following Snell's law of refraction¹⁰⁹ (Figure 51A):

$$\frac{\sin(\gamma)}{\sin(\beta)} = \frac{n_{\text{UVFS}}}{n_{\text{air}}} \quad (7)$$

$$\beta = 90^\circ - \alpha_c \quad (8)$$

With these equations, the maximum angle of incidence for total internal reflection (γ) was calculated from the critical angle (α_c). The refractive index of water (n_{water}) equals to 1.333¹⁴² and the refractive index of air (n_{air}) is 1.000,¹⁴³ while the refractive index of UVFS glass (n_{UVFS})¹⁴⁴ is dependent on the wavelength λ of the refracted light. Calculation results are shown in Table 5.

Table 5: Calculated wavelength (λ) dependent critical angles for total internal reflection (α_c) based on the refractive index of UVFS glass (n_{UVFS}). To reach α_c , the light has to enter the prism within the angle γ and is refracted to angle β .

λ [nm]	n_{UVFS}^{144}	α_c	β	γ
637	1.457	66.229°	23.771°	35.954°
532	1.461	65.893°	24.107°	36.623°
488	1.463	65.692°	24.308°	37.026°

Corresponding to these results, the angle of incidence has to be smaller than 35.954° to obtain total internal reflection for all three wavelengths. An angle of 30° (this is $\alpha \approx 70^\circ$) was chosen to be set up, resulting in only $36 \mu\text{m}$ distance deviation between the center of the red and blue laser spot on the sample slide.

To form a 30° angle for the light to enter the prism, a right-angled triangle between the last mirror of the beam path and the objective was constructed (Figure 51B). The height and distance between the two points can be chosen within the limits of the following equation:

$$\tan(30^\circ) = \frac{\text{height} - 0.5 \cdot \text{height of prism}}{\text{distance} - 0.5 \cdot \text{width of prism}} \quad (9)$$

For the given setup, the distance was limited to 170 mm, the height of the prism to 11 mm and the width of the prism to 20 mm. That gives a height of 97 mm for the mirror on the microscope desk to form a 30° angle. After positioning, the preceding mirror was brought to the same height to create a horizontal beam path. The mirror on the microscope desk was then tilted to reflect the laser beam onto the objective.

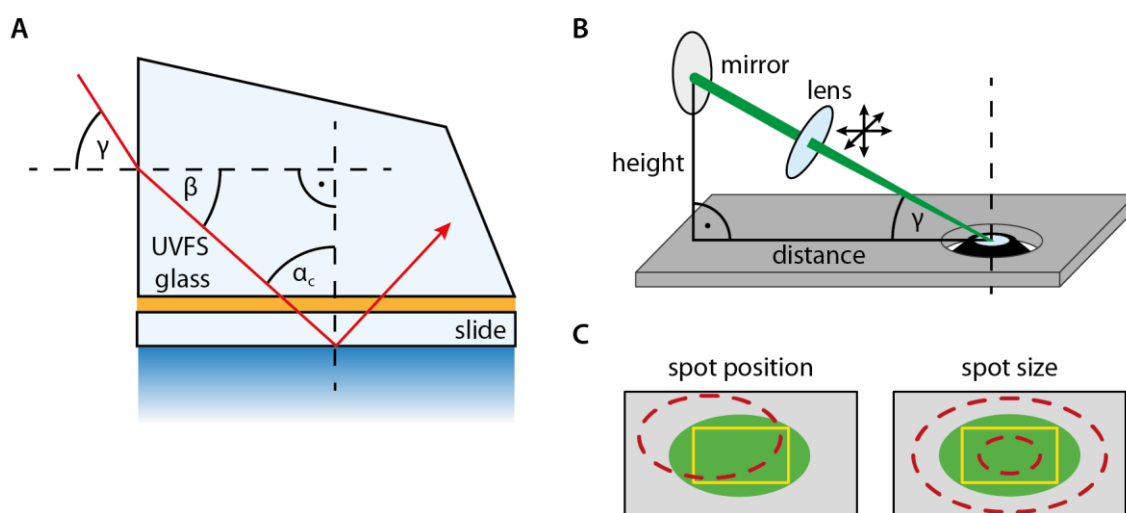


Figure 51: **A** Schematic representation of the prism with angles used for calculation of the angle of incidence (γ) for the laser beam to enter the prism. **B** Schematic representation of the beam path calculated to form the desired angle of incidence (γ) with the laser beam passing the focusing lens on a XYZ stage for fine adjustment. **C** Schematic representation of the spot position and size optimization to evenly illuminate the region of interest (yellow rectangle, laser spot is shown in green). The dashed red line indicates an example for wrong spot size or position.

To obtain a high energy density for the evanescent field, the laser beam was focused with a lens ($f=75$ mm) onto the objective. The lens was positioned into the laser beam with the laser beam passing the center of the lens. To optimize the size and position of the laser spot on the sample slide, the lens was mounted onto a XYZ linear stage with micrometer drives for each direction. Using a surface adherent fluorophore on the sample slide, the size and position of the laser spot were preadjusted to evenly illuminate the region of interest (Figure 51C) with all three lasers. Final adjustments were made for each individual sample slide before measurement.

4.1.4 Detection

After excitation of the fluorophores on the sample slide, the resulting emission fluorescence is collected through the water immersion objective in the microscope. Depending on the excitation wavelength used, different filters can be chosen to block excitation stray light between objective and detection beam path (Figure 52). Here, one long pass filter for blue or green excitation and one triple band pass filter for red excitation were equipped. The filtered emission light is then entering the detection beam path through the side of the microscope. At the intermediate image plane of the microscope a mechanical slit was mounted to define sharp boundaries of the region of interest. The following lens ($f=100$ mm) parallelizes the beam path before the green and red emission fluorescence are separated through a dichroic mirror. In the

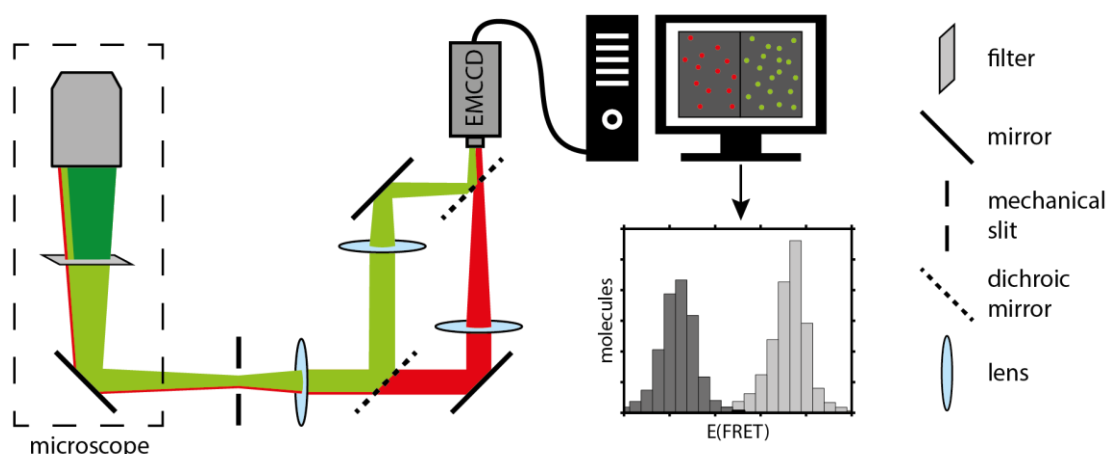


Figure 52: Detection beam path. The emitted light is collected through an objective and filtered inside the microscope before the region of interest is defined by a mechanical slit. The parallelized light is split into a red and green channel and focused on different areas of the EMCCD camera. The picture is transferred to the computer, saved, and analyzed.

following beam path both light beams are focused individually through a lens ($f=150$ mm) onto the EMCCD camera. The distance between the last lens and the camera is identical in both beam paths to ensure the same image sizes. However, the red and green channel are focused on different areas of the detector in the camera to be able to detect them simultaneously. The picture from the EMCCD camera is then transferred onto the computer monitor, saved and in a second step subject to data analysis (see chapter 5.10.5).

4.1.5 Remote control software

In contrast to individual control over the lasers and shutters, integrated control for these instruments is more convenient. Lasers and shutters can be controlled remotely together with the EMCCD camera by individual softwares, supplied by the manufacturers. However, adding up to five individual software instances with at least 7 control windows in total, this is quite complex to control and with that inefficient in usage. Moreover, in this way it is only hardly possible to consistently couple excitation schemes with data acquisition.

In order to obtain a more simple and efficient control over the remotely accessible instruments, I combined all parts in a single user interface (Figure 53) with LabVIEW (version 2021). Here, I

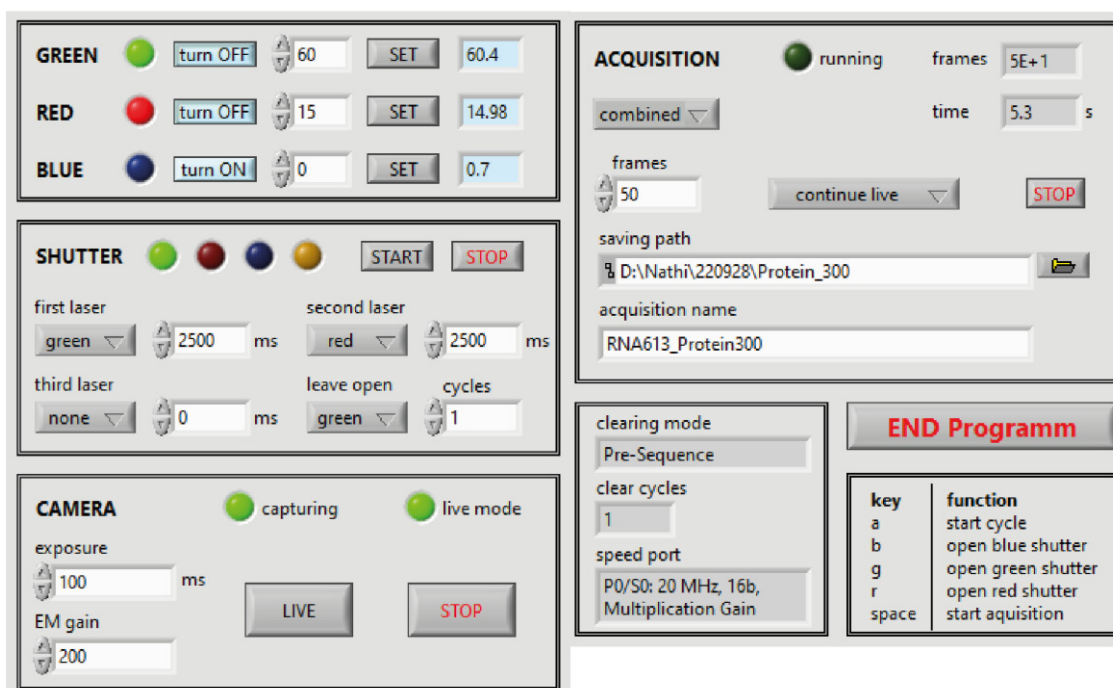


Figure 53: User interface of the smFRET setup control software with panels controlling the green, red and blue laser, the corresponding shutters as well as camera and acquisition parameters.

focused on the most important settings, reasonable flexibility and added a keyboard control tailored to our needs. Visually, the program is divided into five different control panels: laser, shutter, camera, acquisition and additional information. The detailed control functions are described in the following sections.

Laser

The lasers can be switched on and off by mouse click and the laser intensity can be set (white fields). Active lasers are indicated by an illuminated virtual lamp in the corresponding color and the actual laser intensity is shown (blue fields).

Shutter

Opening and closing of the individual shutters was implemented with keyboard control to gain a comfortable and fast way for switching lasers on the sample slide (“g” for the green shutter, “r” for the red shutter and “b” for the blue shutter). Again, open shutters are indicated with illuminated virtual lamps in the corresponding color (yellow for a running shutter cycle). Furthermore, a shutter cycle of maximal 3 phases with specified laser type (green, red, blue or none) and duration can be set up and executed repeatedly (number of cycles). After initializing the cycle by mouse click on “Start” or with pressing the “a” key, the cycle can be aborted with the “stop” button or will finish automatically and continue with opening the specified shutter after the cycle (“leave open”: green, red, blue or none).

Camera

A second window with the camera image will open as soon as the program is initialized. The camera can be set on live view (“Live” button) with the specified camera settings for exposure time and electron multiplying gain. Virtual lamps indicate if the camera is capturing and if it is in live mode. The “Stop” button will set the camera back to idle.

Acquisition

The acquisition panel together with the camera panel controls the settings for the image data. By pressing the “space” key, acquisition is triggered and the given number of frames is captured. A running acquisition is indicated by a virtual lamp and acquisition time and captured frames are shown in real time. Afterwards, the image data is saved as TIFF stack under the given saving path and acquisition name, however, the image name is formed by the acquisition name, extended by successive numeration, binning and frame values. The camera behavior after the acquisition (continue in live mode, continue closed) can be set to ensure effortless and fast repeat measurements. Most importantly, the acquisition can be combined (select “combined”, not “individual”) with the specified shutter cycle. With this, acquisition and shutter cycle will

start correlated when acquisition is initialized. Acquisition can be stopped prematurely by pressing the “Stop” button, so only the frames already captured are saved. In the combined mode the shutter cycle will also stop upon stopping the acquisition.

Preadjusted settings

Important and preadjusted settings are indicated in the lower left edge of the user interface. The clearing mode is set to “pre-sequence” (1 cycle) and the 20 MHz speed port is used.

End program

The program will automatically close all shutters and turn off the lasers as well as the camera when shutting down, which is important especially for safety reasons.

The newly created user interface is not only noticeably more convenient to use, but also enables significantly faster measurement times. Especially with the introduction of keyboard controls, the clearly structured interface and the introduction of “convenience” settings the acquisition time is reduced to a minimum. However, the most important feature is the possibility to couple acquisition with shutter cycles. This opens up a great variety of applications that require precise laser switching in an automated fashion with up to three different wavelengths.

4.1.6 Software application

The probably simplest experiment is used to filter out donor only and acceptor bleached molecules from individual measurements. Adapted from alternating laser excitation (ALEX) approaches¹⁴⁵, this requires starting the measurement with the donor excitation laser for an exactly defined time to capture the experimentally desired molecule data. This is then followed by fast switching to the acceptor excitation laser to trace only molecules attached to an intact acceptor fluorophore. Comparison of both excitation schemes will yield only molecules with intact donor and acceptor molecule attached over the whole measurement (Figure 54). In principle this can also be done using individual programs and switching the lasers by hand, however, an automated and directly acquisition coupled program has some major benefits. The most important point is the laser switching at the exact same time point for every single measurement, ideally within one frame (normally 100 ms). This is only reliably possible in an automated setup, as it is impossible to rely on the human speed of reaction (about 300 ms for reaction on visual stimuli).¹⁴⁶ This software automatically sets the green laser excitation, starts acquisition, switches to the red laser after an exactly defined number of frames and switches back to the green laser after ending the acquisition. This ensures the same excitation scheme

for every measurement and with that makes a uniform data processing possible. Moreover, programmed acquisition setups keep measurement times as short as possible which is important for high quality data. For the given example, automation significantly reduces the effective excitation and simplifies the user action from 8 mouse clicks in 3 different windows to 1 single key press. This yields a reduction of measurement time by at least 50 % for histogram data. All these benefits are even more pronounced with rising complexity of the excitation schemes.

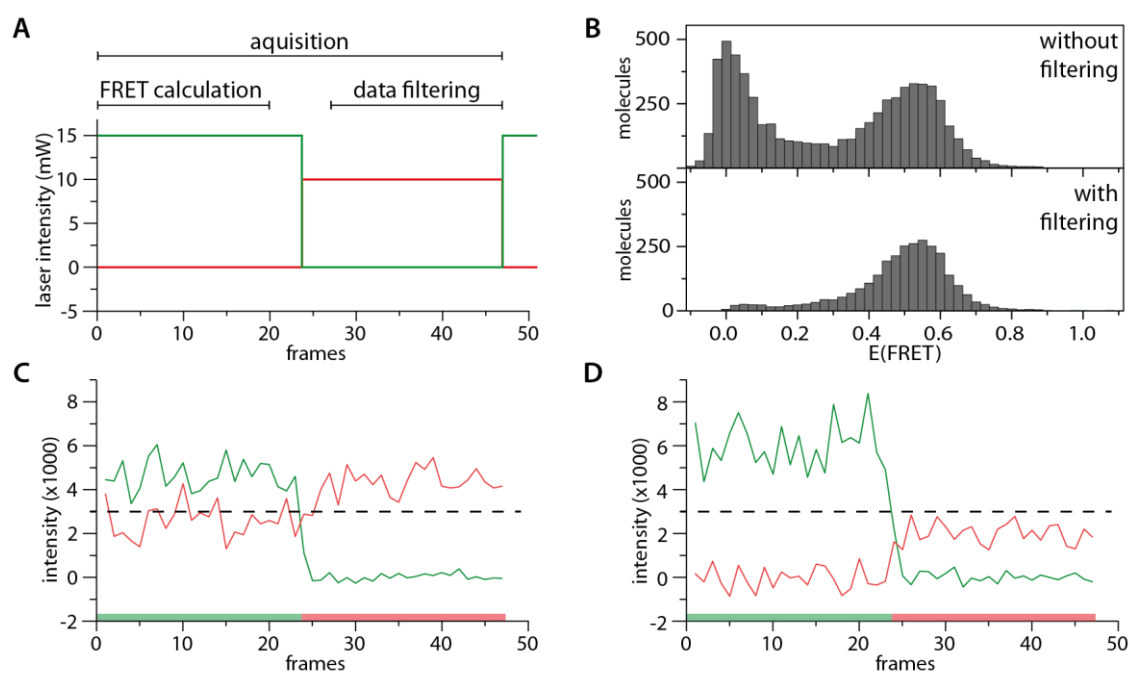


Figure 54: **A** Automated excitation scheme with indicated regions for FRET calculation and data filtering. **B** Exemplary histograms with and without threshold filtering. **C/D** Exemplary molecule trace for a molecule that is considered valid for analysis(**C**) and for a donor-only molecule that is filtered out (**D**), depending on the average acceptor intensity in last 20 frames (above or below the set threshold (dotted line)). Laser excitation is indicated above the x-axis.

4.2 smFRET measurements

In the second part of this thesis, the methyltransferase complex was investigated regarding its dynamic behavior during the catalytic cycle. The applied smFRET approach is focused on differences in the target RNA conformation. This gave insights on the differences between catalytically competent, post-catalytic and catalytically incompetent complexes by analysis of FRET state histograms and the dynamics of individual molecules.

4.2.1 smFRET construct design and preparation

As a first step, a smFRET construct was designed that allows for a variety of measurement conditions and is conveniently accessible. We selected the RNA as target for fluorescent labelling as many labelling techniques for this are commercially available and easier to apply than suitable protein labelling techniques. Based on the previously described methyltransferase activity assay results (see chapter 3.1), a linear RNA with the optimal target sequence GGACU was designed. For highest flexibility in target RNA choice, the fluorescent dyes were not directly attached to the RNA but rather through hybridization with DNA handles above and below the target nucleotide (Figure 55A). This allows for easy exchange of the target RNA without the need to repeat fluorescent labelling and ligation to form a new RNA construct. To create an unpaired target region, a 9 nucleotide long single stranded stretch was left between the two hybridization regions which is the same size as the GGACU containing loop of RNA H2 from the activity assays. A PEG₄-biotin linker was attached to the acceptor labelled DNA handles for biotin-streptavidin immobilization of the whole construct. The advantage of immobilization through the acceptor labelled DNA handle is to minimize background noise from unpaired DNA handles, as the acceptor is not directly excited during the measurement. This is of special relevance for the longer traces (movie data) as these are not automatically filtered for intact FRET pairs in the analysis in contrast to the histogram data (see Figure 54, page 88). The sequence of the second DNA handle was chosen being inverted from the first and the sequence of the RNAs 3'- and 5'-end were adapted accordingly. For optimal sensitivity, the attachment sites of the fluorophores were chosen to span approximately the Förster radius of the dye pair (5.3 nm for Cy3/Cy5).¹⁴⁷ This is required to capture also small distance changes upon protein binding or during the methylation cycle with differences in FRET efficiency (Figure 55B). The distance was calculated to 21 base pairs, assuming the whole construct in A-form helix shape (0.26 nm height per base pair). As the required amino modifier (C6 linker) is best available as modified dT base, the

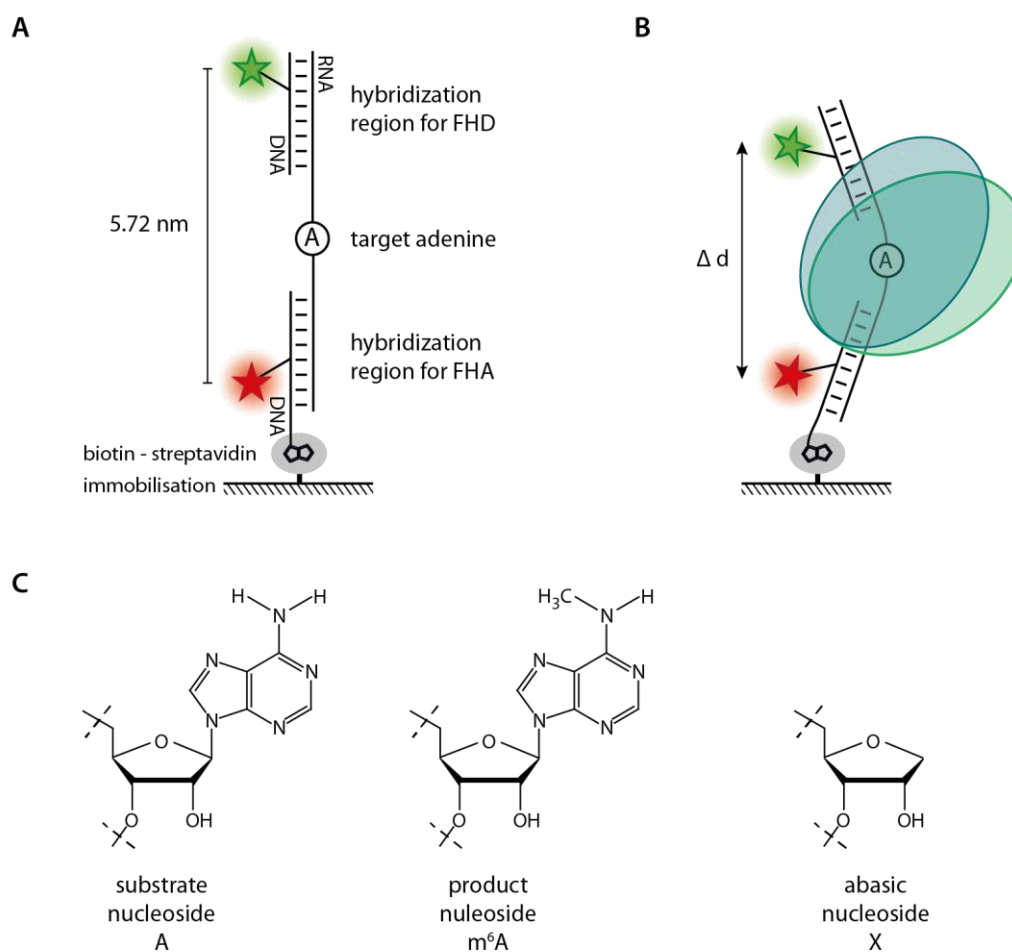


Figure 55: **A** The fluorophores for the smFRET RNA construct are attached by hybridization of labeled DNA handles (FHD, FHA). After hybridization, the fluorophores are approximately 5.72 nm away from each other which is close to the Förster radius of Cy3 and Cy5 (5.3 nm). **B** Distance changes of the two fluorophores upon protein binding to the RNA or during the methylation cycle will be detectable through FRET efficiency changes. **C** Nucleosides in the center of the GGACU target sequence of the different RNA constructs used in this work.

attachment sites were chosen based on the DNA handle sequence and finally are 22 nucleotides apart.

To gain detailed insight into the methylation catalysis, three RNAs with the same sequence but different target nucleotides inside the GGACU motif were designed (Figure 55C):

1. an RNA construct with the regular adenosine (GGACU) to represent the substrate for catalysis
2. an RNA construct with the methylated adenosine (GGm⁶ACU) to mimic the post-catalytic state, and

- an RNA construct with the adenosine exchanged with an abasic site (GGXCU) to represent a catalytically incompetent RNA.

The product RNA as well as the abasic RNA were purchased from Dharmacon, while the substrate RNA was prepared via in vitro transcription. Figure 56A depicts an analytical PAGE of all three constructs where the purity of the samples is shown.

The DNA handles were purchased with an amino modified C6 linker at the respective positions. After coupling with the NHS-ester dyes (FHD: Cy3 / FHA: Cy5), reversed phase liquid chromatography with an acetonitrile gradient elution was applied to purify the samples. Fractionation was performed by eye waiting for the characteristic colour of the dyes (Cy3 red, Cy5 blue). Analysis of the fractions via PAGE with different visualization techniques showed a pure sample in both cases (Figure 56B) and UV-vis analysis confirmed a nearly quantitative labelling of the DNA handles. To ensure binding of the DNA handles to the RNA constructs, an EMSA (see appendix page 191, Figure 87) was performed that displayed good binding of the smFRET construct elements to each other.

In the next step, the activity of the two full-length protein complexes METTL3/METTL14 and METTL3/METTL14/WTAP on the smFRET substrate RNA was assessed (Figure 56C). The methylation assay conditions were chosen to have the same protein to SAM ratio (1:50), temperature (18 °C) and buffer conditions (without the oxygen scavenger system) as in the respective smFRET experiments. The resulting TLC analysis confirmed m⁶A formation for both protein complexes with higher methylation results for the complex without WTAP. This renders both complexes very likely to be catalytically competent under the conditions during smFRET measurements.

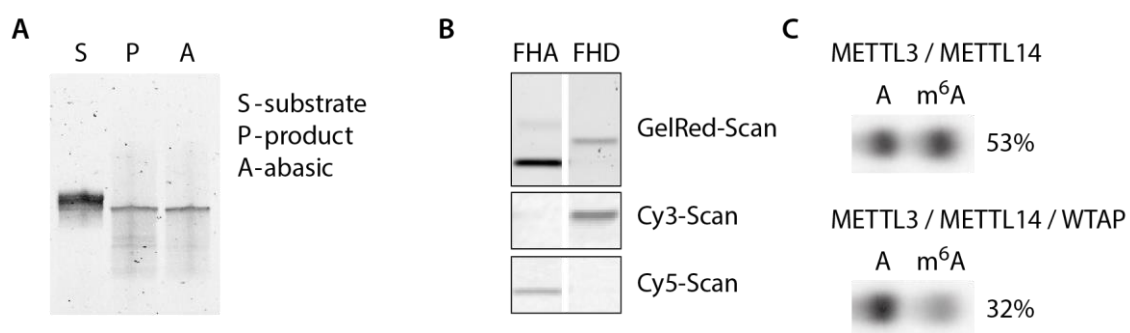


Figure 56: **A** RNA page of the three different RNA constructs used for smFRET measurements. **B** RNA PAGE of the labeled and purified DNA handles. **C** TLC of the enzymatic activity assay of the METTL3/METTL14 complex with and without WTAP on the smFRET substrate RNA with attached fluorophores (DNA handles)

4.2.2 smFRET state analysis

4.2.2.1 RNA constructs

The first smFRET measurements were performed on the individual RNAs without protein addition to validate the designed constructs especially regarding suitable FRET efficiency and comparability between them. The resulting smFRET histograms (Figure 57) show one sharp main peak for all three constructs. The center of the peak is located between a FRET efficiency of 0.37 (product and abasic RNA) and 0.38 (substrate RNA), which indicates very similar conformational states. The resulting FRET efficiency is in the expected range and with that the chosen attachment sites assumed suitable to sensitively display small distance changes with differences in FRET efficiency. In the best case, addition of protein contracts the RNA and shifts the FRET efficiency towards higher values for best sensitivity. If widening of the RNA is observed upon protein binding with shifts towards lower FRET efficiencies, depending on the efficiency change the attachment sites need to be reevaluated to confirm suitable sensitivity.

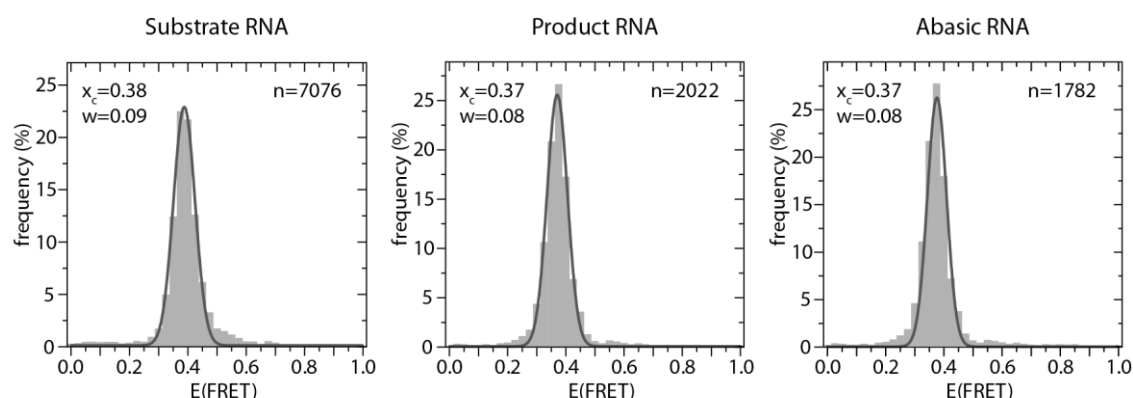


Figure 57: smFRET histograms of the three FRET RNA constructs with main peaks at E_{FRET} of 0.37 (product and abasic RNA) and 0.38 (substrate RNA). The peak center (x_c), peak width at half maximum (w) and analyzed number of molecules (n) are given.

4.2.2.2 Protein titration

After some initial experiments to optimize the conditions (especially buffer) for RNA immobilization and subsequent protein addition (data not shown), the METTL3/METTL14 complex was titrated to the RNA. All these measurements were performed on the substrate nucleotide containing RNA and without SAM or SAH addition to the protein buffer. The resulting smFRET histograms (Figure 58) show a significant shift in FRET efficiency towards higher values and a widening of the peak upon protein addition. A second, wider peak at $E_{\text{FRET}}=0.58$ appears and gets more pronounced for increasing protein concentrations. To get an estimation of the

free RNA ($E_{\text{FRET}}=0.38$) to protein bound RNA ratio, the two peaks were individually fitted with a Gaussian function and the area under the curves were compared. For the given concentrations, the population of the protein bound state changes from approximately 33 % for 300 nM protein complex to approximately 80 % for a concentration of 2700 nM protein complex in the imaging buffer. Repeated measurements directly (approx. 5 min) after protein addition reliably yielded similar results. The system can be assumed to be in nearly equilibrium conditions for the time of measurement. Supporting this, without SAM addition as methyl group donor no catalytic activity may occur that could alter binding and dissociation rates. The RNA-protein binding itself is very stable on the sample slide as buffer flush without protein does not flush away the protein, leaving an unaltered histogram afterwards (data not shown). However, measurements to a later time point (approx. 1.5 h) show a backshift of the populations towards the RNA only state (0.38) (data not shown). This can be caused by starting degradation or adsorption of the protein on the sample slide. All measurements were therefore performed directly (approx. 5 min) after protein addition to the RNA on the sample slide to capture near equilibrium conditions.

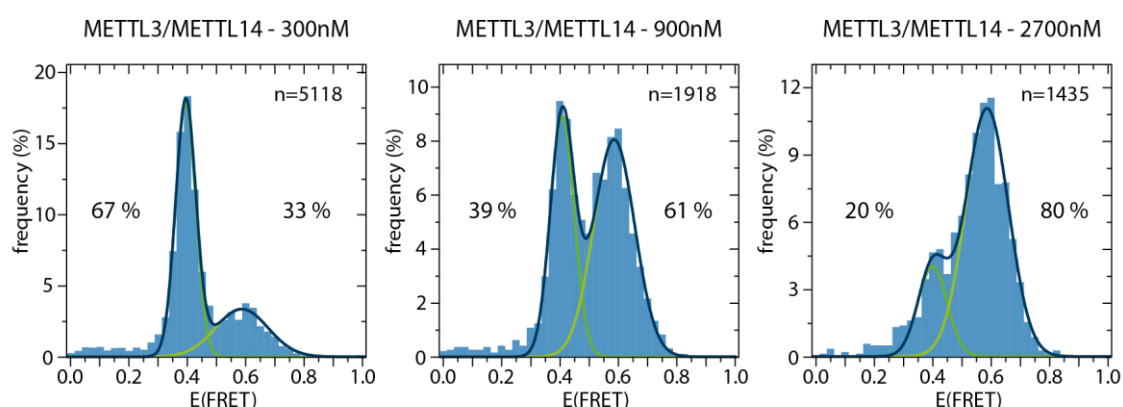


Figure 58: smFRET histograms of the titration of complex METTL3/METTL14 on the substrate adenosine containing RNA. RNA ($E_{\text{FRET}}=0.38$) and RNP ($E_{\text{FRET}}=0.58$) peak were fitted to determine the approximate the proportion of protein bound RNA.

The protein concentration of the complex for all following experiments was chosen to be 2 μM . With this concentration and in equilibrium conditions, all states including the RNA only state are accessible, however, the major population is on the protein bound state. This is preferable as the further analysis focuses on small differences and dynamics in the protein bound state(s).

4.2.2.3 SAM/SAH free measurements

In the next step, all three RNA constructs were measured with the METTL3/METTL14 complex added to the imaging buffer without SAM or SAH. Figure 59 shows the smFRET histograms with very similar results. For all histograms a small population of RNA only molecules ($E_{\text{FRET}}=0.38$) is visible, which shows that the protein concentration is suitable for a more detailed analysis in later stages. The main RNP peak is located at a FRET efficiency of 0.52. In comparison to the previous titration experiments, this represents a slight shift towards lower values ($\Delta=0.06$) for the protein bound substrate RNA. These small differences were observable in repeated measurements with the peak center varying in between a FRET efficiency of 0.52 and 0.58, also for the protein bound product RNA. However, the majority of the measurements resulted in a protein bound state at a FRET efficiency of approximately 0.54 for both RNAs. As the protein bound states of substrate RNA, product RNA, and abasic RNA are very similar, this indicates no major conformational differences in the respective protein bound RNA that are captured with this histogram analysis.

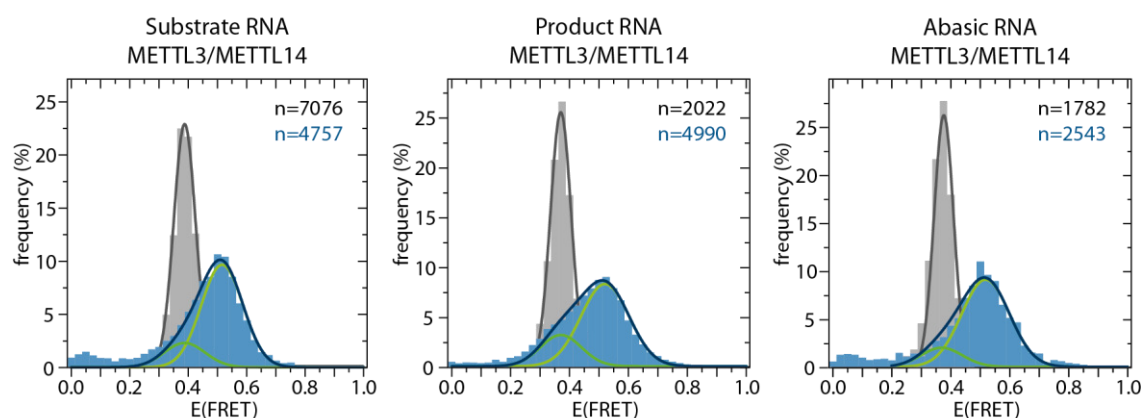


Figure 59: smFRET histograms of the METTL3/METTL14 protein complex (2 μM , blue histogram) with different RNA targets (substrate/product/abasic target nucleotide) with reference to the RNA (grey histogram). Histograms were fitted with one (RNA at $E_{\text{FRET}}=0.38$) or two states (RNA at $E_{\text{FRET}}=0.38$, protein bound at $E_{\text{FRET}}=0.52$). The analyzed number of molecules is given with n .

The same experiments were repeated with the METTL3/METTL14/WTAP complex (Figure 60). Here, the resulting smFRET histograms show a small difference of the protein bound peak of the abasic RNA with a FRET efficiency of 0.52 in comparison to the other protein bound RNAs with a FRET efficiency of 0.56. As reported for the protein complex without WTAP, the peak center of the protein bound states for the substrate and product RNA differed slightly between a FRET efficiency of 0.54 and 0.58. However, for the WTAP containing complex, the majority of the

measurements resulted in higher FRET efficiencies than observed for the METTL3/METTL14 complex alone. On average, this indicates a minor contraction of the substrate and product RNA construct upon WTAP addition. This difference was not observed for the abasic RNA in the analyzed histogram data.

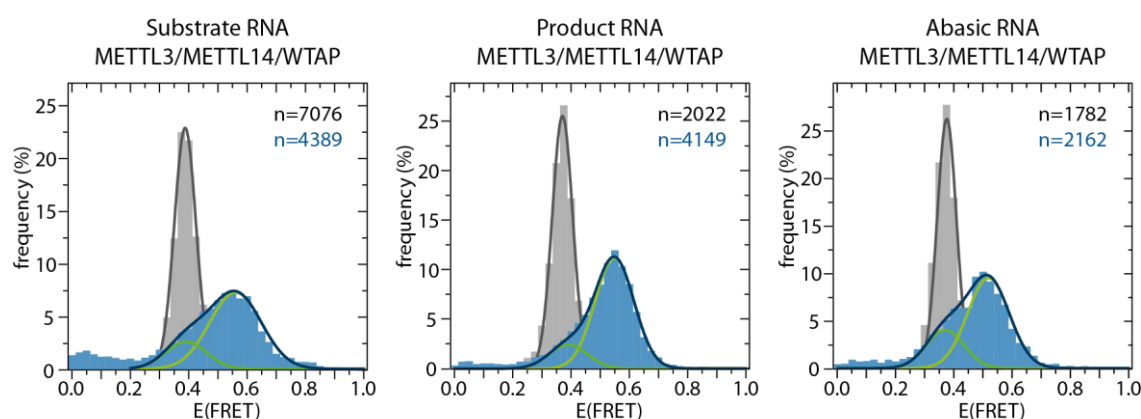


Figure 60: smFRET histograms of the METTL3/METTL14/WTAP protein complex (2 μ M, blue histogram) with different RNA targets (substrate/product/abasic target nucleotide) with reference to the RNA (grey histogram). Histograms were fitted with one (RNA at $E_{\text{FRET}}=0.38$) or two states (RNA at $E_{\text{FRET}}=0.38$, protein bound at $E_{\text{FRET}}=0.56$ or $E_{\text{FRET}}=0.52$ (abasic RNA)). The analysed number of molecules is given with n .

The obtained smFRET data of the METTL3/METTL14 complexes with and without WTAP give a hint towards an RNA contracting role of WTAP. These measurements were performed without the addition of SAM or SAH as cofactor, which renders this data a valuable reference point. However, the most likely directly catalytically relevant RNP compositions require the addition of the corresponding cofactors.

4.2.2.4 Influence of SAM/ SAH addition

It is unknown whether there is an order in substrate and cofactor binding to the protein complex. It remains elusive, if catalytic binding of the substrate RNA requires preceding SAM insertion into the SAM binding pocket or if the RNA is bound before the cofactor. The same question can be asked for the dissociation of the product RNA in the post-catalytic SAH containing complex. Here, the dissociation of the product RNA can occur prior or after release of SAH. For both however, pre- and post-catalytic complex formation, there is also the possibility for a mostly random form of RNP association and dissociation without the need for a specific

binding order. Besides these considerations, in between the RNA and cofactor association and dissociation events it should be undoubtedly required to have an assembled pre-catalytic state that involves the bound substrate RNA and SAM in the same complex. This complex composition is then followed by a post-catalytic RNP, consisting of the methyltransferase bound to the product RNA with the formed m⁶A and the after methyl group transfer remaining SAH. Both of these catalytically relevant RNP compositions were investigated in the following sections for the METTL3/METTL14 and the METTL3/METTL14/WTAP complex.

As shown before (Figure 56), both protein complexes are catalytically competent upon the addition of SAM at conditions similar to the smFRET measurements. This is desired to capture a functioning complex during the measurements. However, this adds another layer of complexity for the measurements on the substrate RNA with SAM addition to the protein complex. The previous equilibrium conditions are altered by the possible formation of product RNA and SAH, which most likely have different binding constants to the complex. To also capture the influence of the resulting additional catalytically incompetent RNP compositions (SAM/product RNA, SAH/substrate RNA), those were individually added to the measurements.

The abasic RNA construct represents a catalytically incompetent target RNA and was measured in combination with both, SAM and SAH. This allows to compare possible pre- and post-catalytic cofactor dependent effects on the RNA without the influence of the methylation state of the RNA. Most importantly, the addition of SAM was included to mimic a pre-catalytic RNP complex without the ability for catalytic turnover as close as possible. To complement the set of conditions, the measurements with SAH in the imaging buffer on the abasic RNA were added. Another possibility would be the use of a METTL3 inhibitor like STM2457⁶⁵. To efficiently inhibit METTL3, this small molecule binds in the SAM binding pocket with high affinity (1.4 nM). However, it additionally binds in the DPPW region and extends towards the possible adenine binding site (Figure 61). STM2457 therefore most likely alters the adenine binding of the substrate RNA towards the methyltransferase complex active site, which renders this the less suitable option. Moreover, the probably altered RNA binding may influence the protein complex dynamics which were analyzed in a later stage.

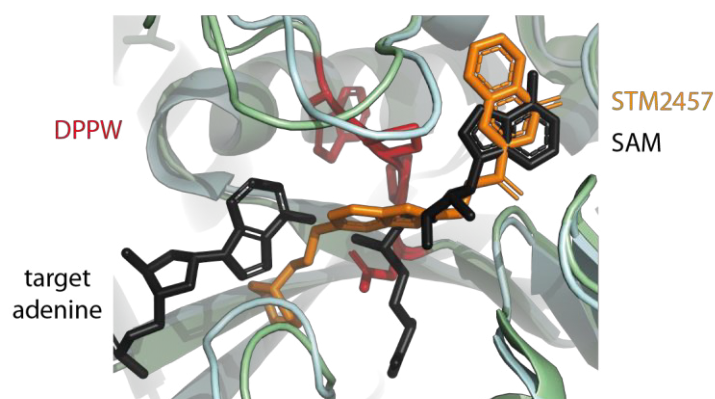


Figure 61: Overlay of the crystal structures of the SAM and STM2457 bound METTL3/METTL14 complexes (PDB 5IL1, 7O2I) with the active site of the substrate bound EcoP15I (PDB 4ZCF). The STM2457 molecule interferes with the possible target adenine positioning in the active site.

The different RNP compositions were first measured for the METTL3/METTL14 complex without WTAP. The resulting smFRET histograms are shown in Figure 62. The comparison between the two catalytically relevant states shows a minor shift in FRET efficiency for the main peak. The substrate RNA bound state for the SAM supplemented complex displays the main population at a FRET efficiency of 0.55 while the product RNA bound state for the SAM supplemented complex is shifted to a FRET efficiency of 0.50. This indicates a more compact conformation in the SAM bound state. The same tendency was observed for the WTAP containing protein complex, however with a small FRET efficiency increase for both conditions to 0.58 and 0.55 respectively. This again confirms the already observed shift to higher FRET efficiencies for the addition of WTAP to the protein complex.

The catalytically irrelevant RNP compositions with the product RNA bound to a SAM containing complex and the substrate RNA bound to a SAM supplemented complex were investigated in the next step. Here, a very similar pattern was observed for both complexes in both conditions. However, an increased RNA only peak at a FRET efficiency of 0.38 was observed for complex METTL3/METTL14 with SAM on the product RNA. This is most likely originating from a lower binding affinity of the product RNA to the SAM bound protein complex. This is reasonable as the respective methyl groups of the product RNA and the SAM molecule might structurally interfere in the active site. Nevertheless, as this is neither observed in all the measurements nor the case for the complex with WTAP, the protein concentration might also slightly vary on the sample slide.

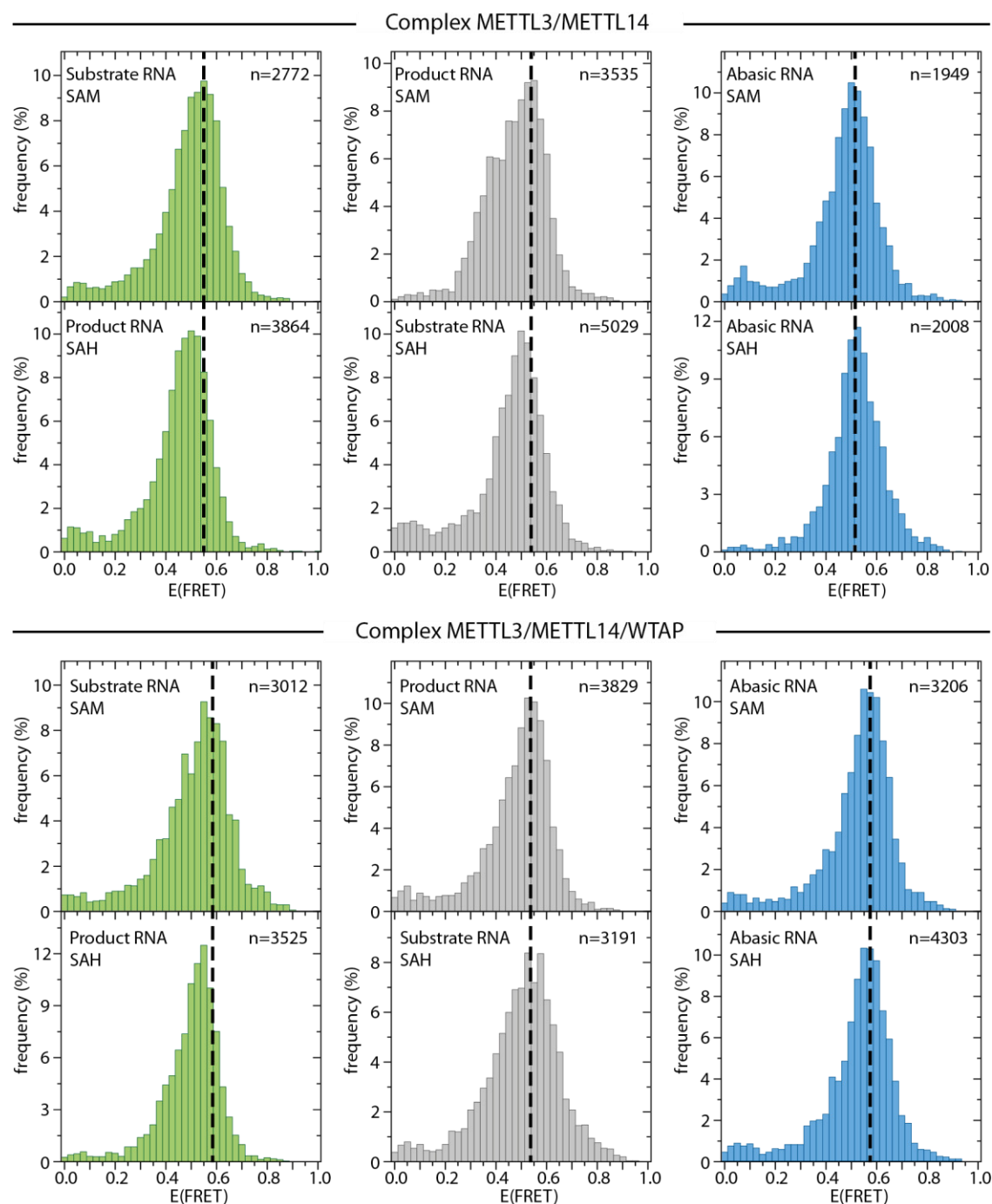


Figure 62: smFRET histograms of the catalytically relevant RNP compositions (green), other possibly occurring combinations (grey) and catalytically incompetent RNPs with the abasic RNA (blue). The number of molecules is given with n.

The smFRET histograms for the measurements on the abasic RNA are very similar for the addition of SAM and SAH. The comparison of the two protein complexes again shows a FRET efficiency increase from 0.52 to 0.57 upon WTAP addition.

The described results indicated two main findings: the pre- and post-catalytic states of the complex display a slightly different population distribution. The SAM supplemented complexes shift the substrate RNA bound state towards higher FRET efficiencies when compared to the product RNA bound to the SAH supplemented complexes. Moreover, addition of WTAP changes the respective populations to a higher FRET efficiency. In both cases, this indicates a slightly more compact conformation of the RNA in the respective conditions. The observed differences are however of a rather minor nature (around $\Delta E_{\text{FRET}} = 0.05$) and not completely reproducible in all cases. They are nevertheless a valid tendency for the behavior of the complex during catalysis.

4.2.3 Analysis of dynamic behavior

To gain a more detailed insight into the dynamic nature of the methyltransferase complex, the analysis was complemented by the analysis of longer molecule traces (60 s long movies). The measurement data of the different RNP complexes formed by the RNAs and proteins specified in Table 6 were included in the analysis. Moreover, the substrate RNA without protein addition was analyzed as a reference.

Table 6: Protein complex and RNA combinations measured for analysis of movie data.

Protein complex with cofactor	Substrate RNA	Product RNA	Abasic RNA
METTL3/METTL14 with SAM	X		X
METTL3/METTL14 with SAH		X	X
METTL3/METTL14/WTAP with SAM	X		X
METTL3/METTL14/WTAP with SAH		X	X

First, the acquired molecule traces were sorted by eye into a group of molecules with a typical FRET pair behavior and a group of donor-only molecules or other undesired behavior (two-step acceptor or donor bleaching, blinking, no bleaching ...). A typical FRET pair behavior requires a reasonable FRET efficiency (between 0 and 1 in average) and following first a one-step bleaching of the of the acceptor molecule with anticorrelated donor behavior and subsequent donor bleaching (Figure 63A). Moreover, molecules with only bleaching of the acceptor molecule with anticorrelated donor behavior (no donor bleaching) or simultaneous donor and acceptor

bleaching were accepted for analysis. In all cases, the first bleaching event was required to be after approximately 10 s earliest.

The traces accepted for analysis (around 200 molecules per measurement condition with protein complex, 85 for the substrate RNA only measurement) were trimmed to the non-bleached regions and subject to hidden Markov modelling (HMM) with HaMMY¹⁴⁸ (Figure 63B). The first experiments were fitted with 3, 4 and 5 states each, yielding the most reasonable results for a 3-state analysis for RNA only data and for a 4-state analysis for protein related measurements. The HaMMY analysis was performed in two different ways: for all traces on an individual scale and in a “stitched” format. For the individual analysis, all molecule traces were independently fitted by HaMMY as exported beforehand. In the stitched analysis, all molecule traces for one measurement condition were stitched together to form one very long trace before the HaMMY analysis so all traces are fitted together in one fit.

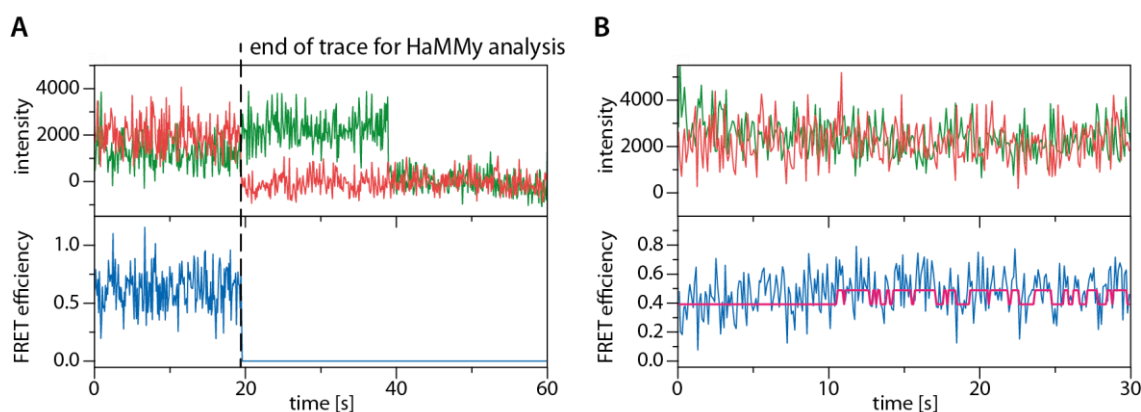


Figure 63: Exemplary molecule traces for donor fluorescence (green), acceptor fluorescence (red) and resulting FRET efficiency (blue). **A** The full-length trace for molecules with intact FRET pair (acceptor and donor bleaching visible) are exported for HaMMY analysis (shortened to the non-bleached regions). **B** The exported traces are then fitted by HaMMY (purple line) for further analysis.

After HaMMY analysis, three output files per input file are generated. The “report” file contains information on the fitted states and the probability and total number of the transitions between those states. The “path” file contains the calculated FRET efficiency per frame and the state which this frame was fitted to. The last output file (“dwell”) reports on the dwell time for which the molecule stays in the given state before transitioning to the next (e.g. 3 s state A, followed by 2 s state B, followed by 5 s state A, ...). These output files are concise to analyze for the stitched format as it provides a good overview with only 3 files generated for the combined

accepted traces. However, when moving forward to the individual analysis this becomes quite complex, as three files are generated per molecule (approx. 600 files per condition). One way to analyze this data is the generation of a transition density plot (TDP).

4.2.3.1 Transition density plots

TDPs are generated from one or more “report” files, representing a heat map of the identified transitions (optionally weighted by number of transitions) (Figure 64). Here, a TDP for the stitched analysis as well as a combined TDP for all individual reports per measurement condition was created (see also chapter 6.8.2, Figure 88 and Figure 89). Differences between the stitched and individual analysis are already showing for the substrate RNA only data and are even more pronounced upon protein addition (Figure 64). In the first case, less diverse transitions are indicated in the individual analysis while the variety of transitions in the protein containing condition increases. A more diverse transition distribution for the individual analysis is reasonable as the individual fitting results in a distribution of similar fitted states, differing only slightly. However, the less diverse transitions for the individual RNA only trace analysis uncovered a general bias of the stitched analysis form. During HaMMMy analysis, the stitched molecule traces are treated as one long trace originating from one single molecule. The molecules themselves show, even assuming only static molecules, not completely identical FRET efficiencies. This is perfectly visible in the population distribution of the respective histogram (Figure 57). Stitching together the traces slightly differing in overall FRET efficiency creates an artificial transition between two states for every molecule-to-molecule crossover in the combined long trace. This bias is closely related to not only the amount of stitched molecule traces, but also to the (random) stitching order and most importantly the complexity of the system.

To overcome this bias, the following analysis of the movie traces was performed from the individual trace fitting by HaMMMy. However, the analysis of this data via TDPs in the given form faces several problems. The TDP allows a rough estimation which transitions are higher populated than others. The more detailed information on the exact number of transitions is, however, missing. For the stitched analysis, this information is easily accessible within the “report” file. In contrast to this, for the individual analysis there is no combined report file of all molecules. This leaves the question on how many transitions (in total and per transition between two distinct states) were identified for the analyzed molecules. Moreover, in both cases one of the most significant information is lacking in the analysis of the HaMMMy fit: How

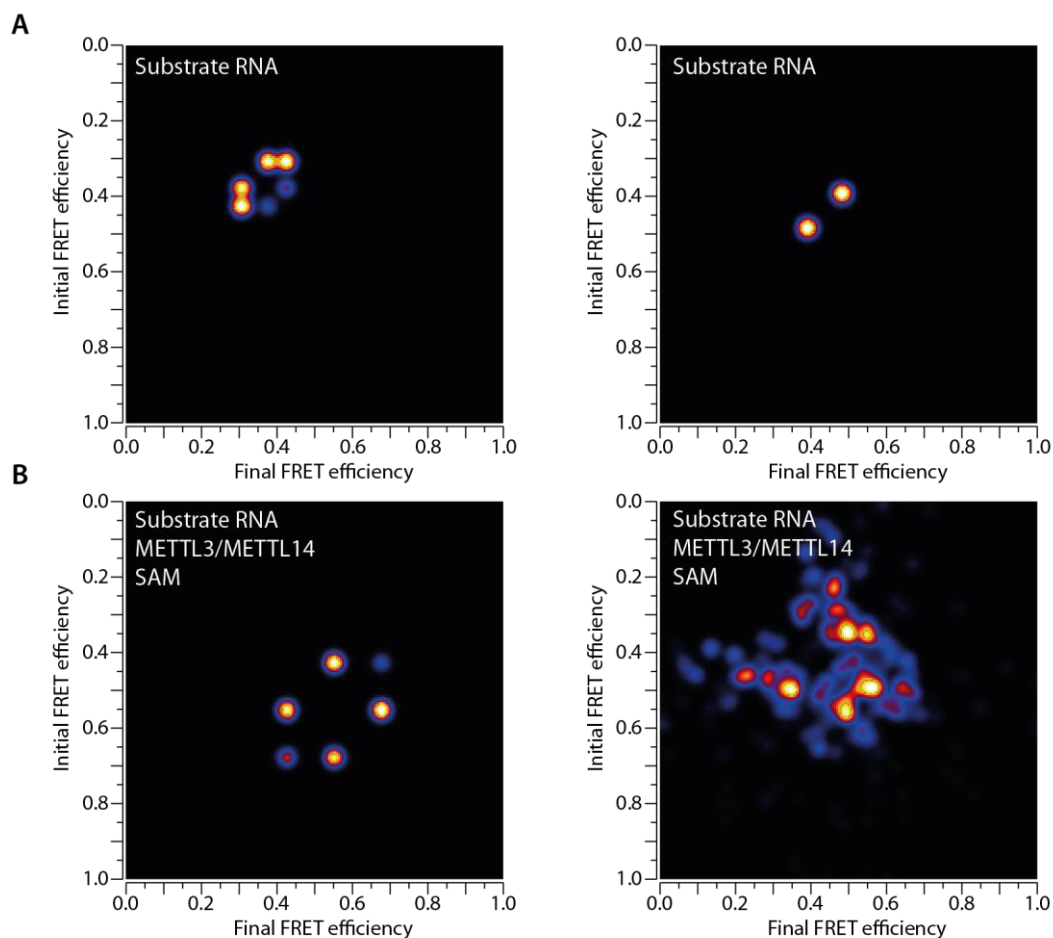


Figure 64: Transition density plots (weighted by number of transitions, “stitched” (left) and individual (right) analysis) of the movie analysis for the substrate RNA alone (A) and with the METTL3/METTL14 complex and SAM present in the imaging buffer (B).

many molecules are transitioning at all? This information is closely related to a well-founded decision on which transition between two states is considered valid for the given condition. Apart from this, identification of the dynamic molecules helps to check for single molecules with an unreasonable behavior that distort the whole picture.

All the problems and limitations of the previously applied analysis on the HaMMY output described above, prompted us to write suitable Matlab scripts (chapter 6.7.4 to 6.7.6) for a more detailed analysis of the data. The analysis was performed starting with the individual output files generated by HaMMY fitting for every molecule to be analyzed (approx. 200 molecules per condition).

However, as the HaMMY fitting of the individual molecules produces a huge variety of fitted states, often only with minor differences, the analysis was performed by clustering the distribution of fitted states into 4 main groups. This creates a much better comparability

between the different measurement conditions and renders the evaluation of the results manageable. The borders of these groups were determined from the TDPs of the stitched state fitting. The main advantage of the stitched analysis lays in finding the midpoint in FRET efficiency that fits the most of all molecules in the given state. The “stitched TDPs” of all protein containing measurement conditions were compared and all of them showed a very similar pattern with only minor differences in the centers of the shown transitions (see chapter 6.8.2, Figure 88 and Figure 89). The borders of the clustered groups were determined on a representative TDP shown in Figure 65 (see also Table 23 in chapter 5.10.5.2). The first group was clustered from molecules fitted to a FRET efficiency between 0.30 and 0.44. The following borders between the individual groups were set to 0.51, 0.60 and 0.80. Everything fitted below 0.30 and above 0.80 was considered as irrelevant for the analysis.

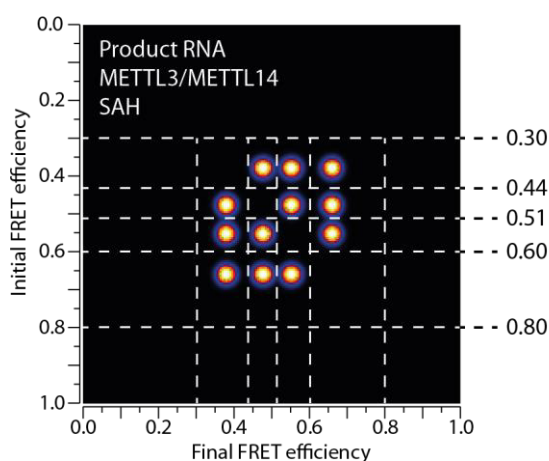


Figure 65: Representative transition density plot for the determination of the clustering borders for further analysis.

The Matlab scripts were programmed to first cluster the fitted HaMMY states and respective transitions accordingly and then generate a list of several key parameters per measurement condition:

1. The amount of dynamic molecules
2. The population distribution of static versus dynamic molecules
3. The population distribution between the fitted states
4. The total number and percentage of molecules showing a specific transition
5. The total number and percentage observed of a specific transition
6. The average number a specific transition observed per molecule
7. The transition rate for a specific transition

This list of key parameters was generated for all the different conditions summarized in Table 6. These conditions represent the catalytically relevant states for the METTL3/METTL14 complex with and without WTAP together with “reference” measurements for a catalytically incompetent RNP complex with the abasic RNA.

The evaluation of these key parameters will be described in the following sections.

4.2.3.2 Population distribution

To begin with, the analyzed molecules were divided into static molecules that showed no transition and dynamic molecules with identified transitions between the clustered states. The substrate RNA was analyzed first as a reference point. Here only 1 out of 85 molecules (1%)

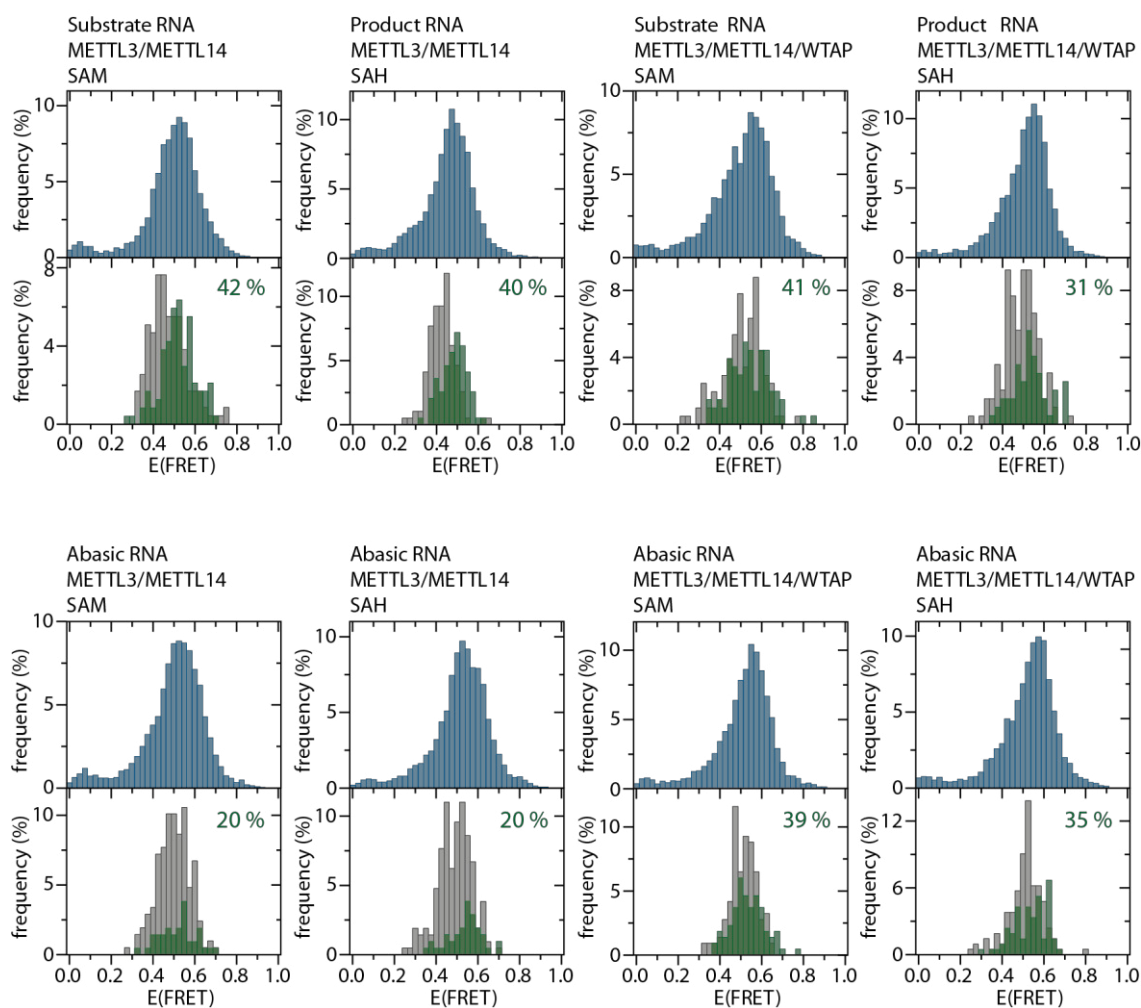


Figure 66: Comparison between previously acquired histogram data (blue) and the dynamic (green) and static (grey) molecules from the movie analysis. The proportion of dynamic molecules is indicated.

showed a dynamic behavior. This renders the substrate RNA only state a predominantly static state and that is why this condition was not analyzed further. Linking back to the TDP analysis (Figure 64A), this also underlines the importance to capture the proportion of molecules where the observed transitions originate from to render a specific transition valid.

For the protein related measurements, the highest proportion of dynamic molecules was identified for the METTL3/METTL14 complex with supplemented SAM on the substrate RNA (42 % of 237 molecules). The lowest proportion of dynamic molecules was observed for the RNPs formed by the METTL3/METTL14 complex with the abasic RNA (20 % each).

To confirm that the acquired movie data is representative for the given measurement condition, a separate histogram for the static and dynamic molecules from the movie analysis was plotted to be compared to the previously obtained smFRET histograms (Figure 66). The histograms were derived in the same way as for the histograms from the shorter traces (average of first 20 frames per trace, normalized to the number of total molecules). In all cases, the combined static and dynamic molecules represent the population distribution of the previous histograms very well.

Next, the population distribution of the dynamic and static molecules was compared. First the corresponding histograms were normalized to the respective number of molecules (not the total number) and thereafter the difference (dynamic versus static) of the resulting distribution

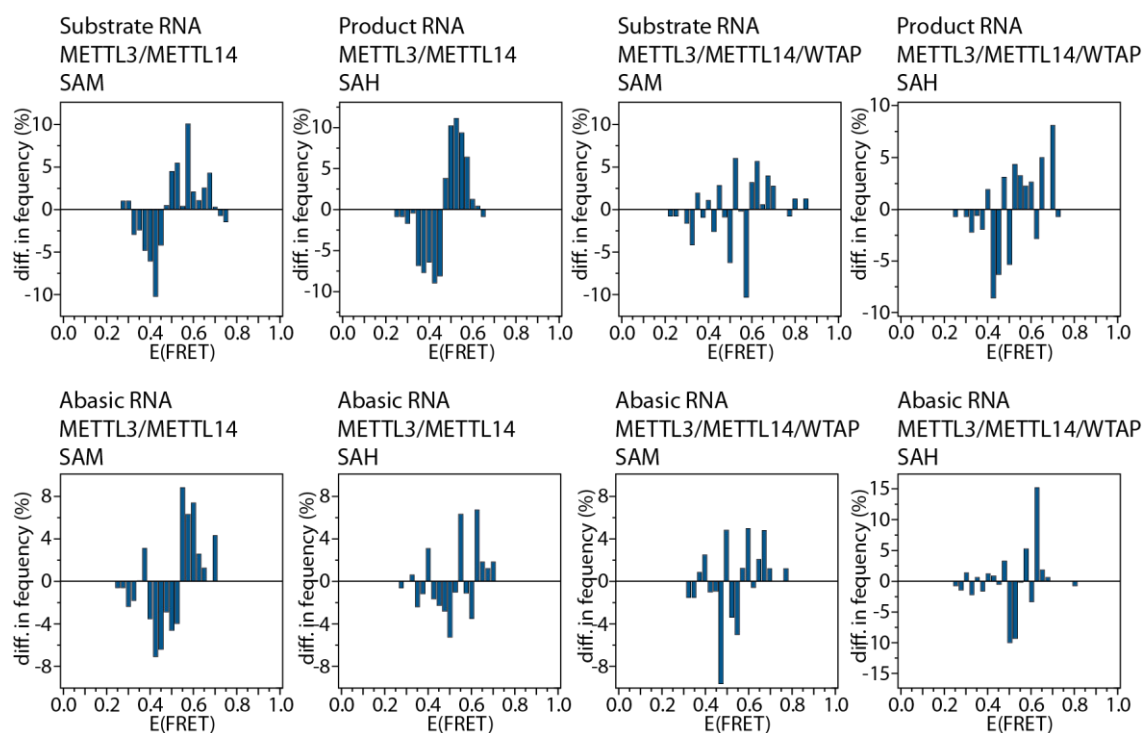


Figure 67: Difference plots for the distribution of dynamic and static molecules (dynamic -static).

was plotted. The result is shown in Figure 67. For all conditions, the dynamic molecules are shifted towards higher FRET efficiencies with different intensity. The most definite discrepancy between static and dynamic molecules was identified for the complex METTL3/METTL14 with SAH on the product RNA. In summary, this feature is more pronounced for measurements with the METTL3/METTL14 complex, but is also observed for the WTAP complex.

4.2.3.3 Cluster analysis

Prior to analyzing individual transitions, the average FRET efficiency of all fitted states was calculated per cluster (Table 7). With this, the four clusters were with this determined to represent the average FRET states at a FRET efficiency of 0.38, 0.48, 0.56 and 0.66 respectively. In the following analysis the clusters are named after their respective average FRET state.

Table 7: Average FRET efficiencies of the fitted states per cluster for every measurement condition.

Condition	Cluster 0.30 to 0.44	Cluster 0.44 to 0.51	Cluster 0.51 to 0.60	Cluster 0.60 to 0.80
METTL3/METTL14 with SAM on substrate RNA	0.38	0.48	0.55	0.65
METTL3/METTL14 with SAH on product RNA	0.39	0.49	0.55	0.66
METTL3/METTL14/WTAP with SAM on substrate RNA	0.39	0.48	0.56	0.66
METTL3/METTL14/WTAP with SAH on product RNA	0.37	0.47	0.55	0.65
METTL3/METTL14 with SAM on abasic RNA	0.39	0.47	0.56	0.69
METTL3/METTL14 with SAH on abasic RNA	0.40	0.47	0.55	0.65
METTL3/METTL14/WTAP with SAM on abasic RNA	0.38	0.47	0.56	0.65
METTL3/METTL14/WTAP with SAH on abasic RNA	0.37	0.47	0.56	0.64
Average	0.38	0.48	0.56	0.66

As a first step, the population distribution of the dynamic molecules between the four states was calculated for every measurement condition. This was performed on the full-length traces with a single frame resolution (see chapter 5.10.5.2). With this, the overall population distribution of the dynamic molecules was determined for the analyzed data as a whole. Table 8

shows the result of this population analysis per condition. For better comparison, the distribution was normalized to the total number of frames each.

The resulting population distribution shows significant differences between the individual RNP compositions. Catalytically competent RNPs show an increased population of the highest FRET state at 0.66 in comparison to the post-catalytic RNPs. Here, an increase in population for the lowest FRET state at 0.38 is visible. The population analysis will be discussed in more detail together with the analysis of the individual transitions in chapter 4.2.3.5.

Table 8: Population distribution of the dynamic molecules derived from the full-length traces.

Condition	0.38	0.48	0.56	0.66
METTL3/METTL14 with SAM on substrate RNA	15 %	26 %	36 %	24 %
METTL3/METTL14 with SAH on product RNA	28 %	42 %	24 %	5 %
METTL3/METTL14/WTAP with SAM on substrate RNA	17 %	33 %	24 %	26 %
METTL3/METTL14/WTAP with SAH on product RNA	23 %	24 %	37 %	16 %
METTL3/METTL14 with SAM on abasic RNA	17 %	43 %	25 %	14 %
METTL3/METTL14 with SAH on abasic RNA	15 %	29 %	34 %	22 %
METTL3/METTL14/WTAP with SAM on abasic RNA	20 %	29 %	30 %	20 %
METTL3/METTL14/WTAP with SAH on abasic RNA	14 %	31 %	33 %	22 %

4.2.3.4 Individual transitions

In the next step, the individual transitions of every measurement condition were analyzed separately. Here, a threshold was introduced rendering a transition as relevant only if more than 5 % of the analyzed molecules show the respective transition. As the transitions are also closely related to the average length of the analyzed traces, the average trace length was calculated for every condition to rule out major differences. The longest average trace length was calculated for METTL3/METTL14 with SAM on the substrate RNA (25.7 s) while the shortest length was identified for the respective condition with WTAP in the protein complex (16.2 s). The majority of conditions was analyzed with an average trace length of approximately 20 s, which results in a suitable comparability.

The resulting tables from this analysis for all conditions can be found in the appendix, chapter 6.8.3. The results were illustrated together (Figure 71 and Figure 72, chapter 4.2.3.5) for a

summarized and comprehensible characterization. In the following paragraphs, the analysis approach for the individual transitions is shown on the example of the complex METTL3/METTL14 with SAM on the substrate RNA (Table 9). Here, also the obstacles and limitations of this analysis are discussed. All other conditions were examined in an analogous way.

Table 9: METTL3/METTL14 with SAM on substrate RNA. Summarized analysis of movie data with calculated indication values for all observed transitions. The median of transitions per molecule are given with the covered range (1-max). Irrelevant transitions are greyed out. Transitioning molecules: 100 of 237 analyzed molecules (42%). Total number of detected transitions: 1465. Mean trace length: 25.7 s.

Transition (from ... to ...)		Total molecules	Percentage of molecules	Total transitions	Percentage of transitions	Transitions per molecule (max)
0.48	0.56	32	14	204	14	3 (31)
0.56	0.48	25	11	191	13	6 (35)
0.56	0.38	21	9	148	10	6 (35)
0.38	0.48	20	8	202	14	4 (31)
0.48	0.38	19	8	194	13	2 (43)
0.38	0.56	19	8	162	11	2 (29)
0.56	0.66	18	8	67	5	5 (28)
0.66	0.56	18	8	66	5	5 (28)
0.48	0.66	10	4	74	5	3 (14)
0.66	0.48	10	4	71	5	2 (15)
0.66	0.38	7	3	43	3	7 (19)
0.38	0.66	6	3	43	3	6 (18)

For the complex METTL3/METTL14 with SAM on the substrate RNA, the transitions between the two states at 0.48 and 0.55 are the most prominent ones (14 % and 11 % of analyzed molecules). The transitions between 0.38 and 0.66 as well as 0.48 and 0.66 in both directions were only populated to a minor percentage (3 % and 4 % of analyzed molecules) and therefore labeled as irrelevant. The number of transitions per molecule showed a very wide distribution, which was observed for nearly all transitions. This is also closely related to the length of the respective

traces with longer traces resulting in the opportunity for more transitions per molecule. In the given case however, this is not the major reason for this observation. Upon a closer look at the individual molecule traces, it becomes evident that molecules with roughly the same trace length often exhibit a completely different behavior. This can be shown at a prime example of three molecules only transitioning between the two states at a FRET efficiency of 0.48 and 0.56: The first molecule with trace length of 40.1 s shows two transitions from 0.48 to 0.56. The second molecule has a trace length of 50.7 s but shows 15 transitions from 0.48 to 0.56 and the third molecule with nearly the same trace length (51.5 s) displays 31 transitions of the same type. All of them exhibit at least one transition with a long dwell time but also (several) transitions with short dwell times (up to 3 s). This clearly indicates the nature of these transitions being very divergent already within a single two state transition. Besides ranking the specific transitions after the proportion of molecules displaying this transition, it is challenging to further analyze the heterogeneous nature of the transitions.

To get a rough estimate of the transition rates, a dwell time analysis was performed for every transition. A histogram of all dwell times was plotted and fitted with an exponential decay curve to derive the transition rate (k) (Figure 68). The transition rates of all relevant transitions for all measurement conditions are summarized in Table 32 in the appendix (page 204).

Depending on the respective transition, the quality of the fit (R-squared) was calculated to be between 0.84 and 0.97. The given errors in the transitions rates are derived from the fitting error. However, the actual inaccuracy can be estimated to be significantly higher in all cases.

One of the main limitations of the whole analysis of specific transitions is the statistical significance. For the number of transitions this is directly related to the goodness of fit as a high number of dwell times (and with that transitions) in the histogram helps during parameterization. Here, the number of transitions differed between 285 for the transition from 0.38 to 0.48 for the METTL3/METTL14 complex with SAH on the product RNA and 22 transitions for the transition from 0.38 to 0.66 for the METTL3/METTL14/WTAP complex with SAM on the substrate RNA. The exponential decay fits of the histograms derived from a low number of transitions were oftentimes not converging and are therefore given as not determined (n.d.).

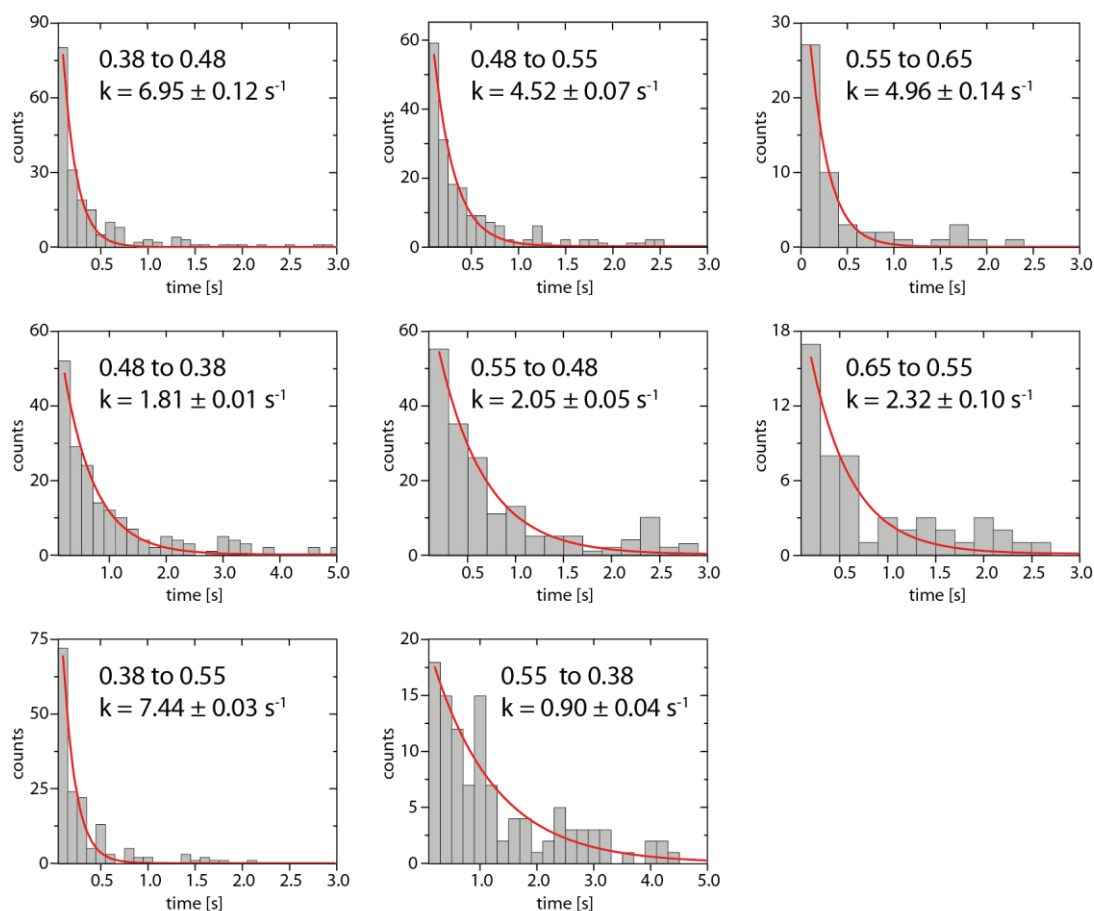


Figure 68: Dwell time analysis for all relevant transitions for METTL3/METTL14 with SAM on substrate RNA. The indicated transition rates (k) were obtained by exponential decay fit (red lines).

Moreover, transitions with longer dwell times are underrepresented in the exponential decay fit analysis. The derived fitting curve is mostly suitable within the window of the first dwell times up to 3 s. Longer dwell times are not captured in the asymptotic region of the fit. Figure 69 exemplarily shows this error source on the instance of the transition from 0.48 to 0.56 for the METTL3/METTL14 complex with SAM on the substrate RNA. The exponential decay fit of this data (204 data points) was rated with a goodness of fit of 0.97 as one of the best fitted data. However, when comparing the fitted curve with the data and evaluating by eye the fit clearly differs from the actual data. In the front area of the histogram (approx. up to 1 s), the fitted curve captures the data to a very good extent but is clearly not in line with the longer dwell times. This again demonstrates the indicated error in the transition rates derived from the fit is presumably underestimating the actual error to quite an extent.

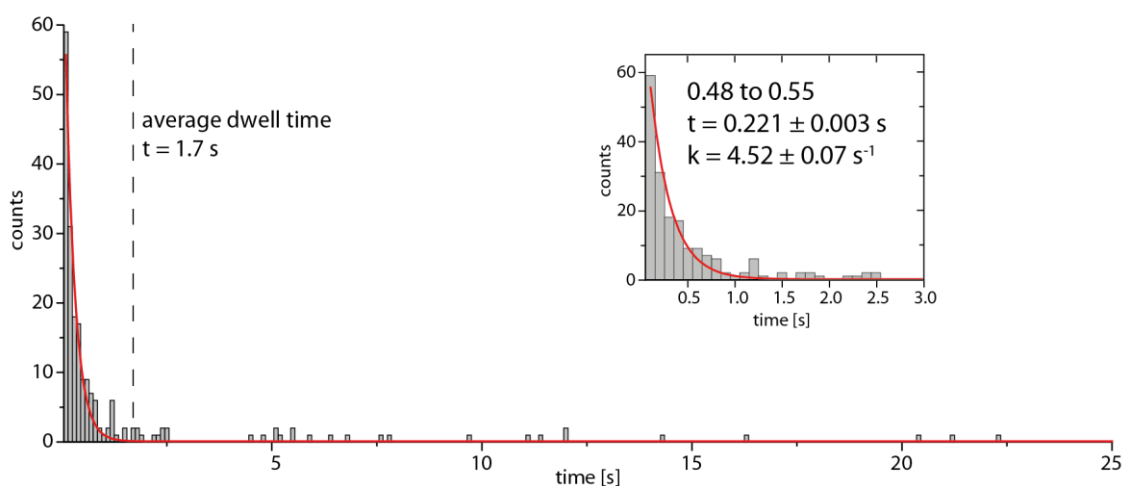


Figure 69: Histogram of the binned dwell times for the transition from 0.48 to 0.56 for the METTL3/METTL14 complex with SAM on the substrate RNA. The data is shown on two different scales. The exponential decay fit (red curve) was used to determine the transition rate (k).

To have a reference value that evenly weights the different dwell times, the average dwell time of every transition was calculated. As the dwell time distribution exhibits an exponential decay shape, this however leads to an overrepresentation of the longer dwell times during transition rate calculation ($1/\text{average dwell time}$) (see Table 33 in appendix page 205). Overall, both approaches to calculate the transition rates are error-prone with the given statistics. Nevertheless, the transition rates derived from the average dwell times are better balanced with respect to the population analysis. For both, the analysis was based on the full-length time traces with every frame weighted equally, independent from the transitions per molecule.

The limitation in statistical significance due to the number of analyzed molecules is also not sufficiently included in the goodness of fit of the exponential decay curve as it is not well represented within the number of transitions. In the best case, 42 % of 237 analyzed molecules show a dynamic behavior (100 dynamic molecules for the METTL3/METTL14 complex with SAM on the substrate RNA) and 14 % of them display a specific transition (0.48 to 0.56). This leaves 32 molecules to analyze for this specific transition. Yet in the given case, the majority (51.5 %) of the observed number of transitions originates from only 4 molecules (12.5 %) as displayed in Figure 70A. This originates from the large variety of transitions per molecule that was discussed above. A similar picture uncovers for the relevant transition displayed by the lowest number of molecules (11 molecules for transition from 0.38 to 0.56 in the METTL3/METTL14 complex with SAH on the product RNA). Here, a total of 11 molecules was analyzed, but only 3 molecules (27.3 %) are responsible for 66.7% of all transitions (Figure 70B).

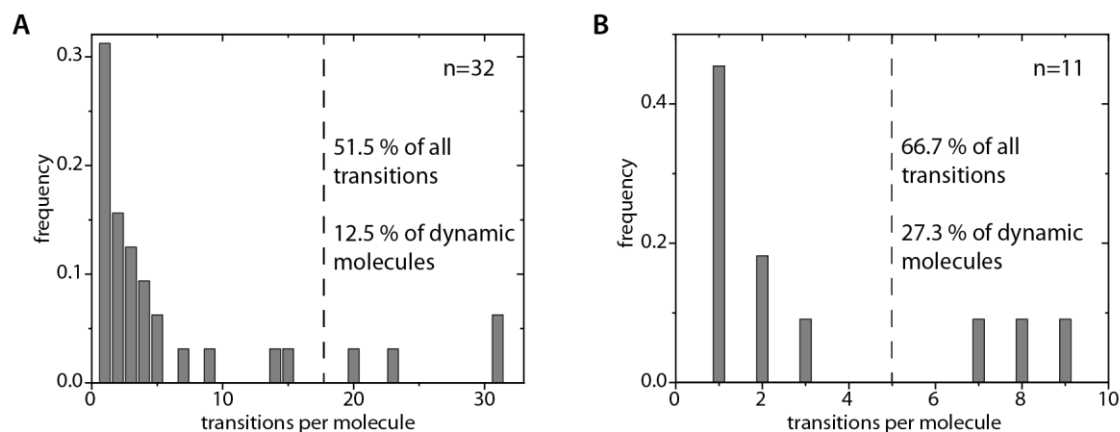


Figure 70: Distribution of the transitions per molecule for **(A)** the transition from 0.48 to 0.56 for the METTL3/METTL14 complex with SAM on the substrate RNA and **(B)** for transition from 0.38 to 0.56 in the METTL3/METTL14 complex with SAH on the product RNA. The total number of molecules is given with n.

This is also closely related to the dwell time distribution as molecules with a high number of transitions per molecule dominate the shorter dwell times. Referring back to the previous discussion this suggests an additional bias for the transition rates from the exponential decay fit. Based on the fit primarily to the shorter dwell times this indicates the fit being dominated by a small proportion of the dynamic molecules. This bias is not included in the average dwell time of the transition.

In summary, this renders the statistical significance of the analyzed transitions insufficient to calculate exact transition rates, and therefore use of these rates as basis for further interpretation. However, the population analysis together with the estimation of transition rates can give a valid insight into the molecular dynamics. Especially the comparison of the different measurement conditions allows for reasonable conclusions. The results for the different RNP compositions will be discussed and compared in the following chapter.

4.2.3.5 Summary of transition analysis

The comparison of the different RNP compositions was performed in two dimensions. First, the dwell times derived from the exponential decay fit (Table 32 on page 204) were examined. Strikingly, the derived transition rates per transition are in the same range for all the measurement conditions, where calculation was possible. The transition from state 0.48 to 0.56 was rated as relevant for every RNP composition and rates were calculated ranging from 3.4 s^{-1} (complex METTL3/METTL14/WTAP with SAM on the substrate RNA) to 6.5 s^{-1} (complex

METTL3/METTL14/WTAP with SAM on the abasic RNA). Moreover, independent of the specific transition the rates towards the higher state were always significantly larger than those of the reverse transition. This was observed also independent of the measurement conditions. Following these transition rates, the majority of dynamic molecules should be in the highest FRET efficiency state at 0.66. This is in contrast to the histogram data calculated for the dynamic molecules and also contradicting the histograms obtained from a larger number of molecules (Figure 66). Here, the main FRET population was observed at a FRET efficiency between 0.5 and 0.6. This inconsistency points towards a systematical error in the dwell time analysis and was already discussed in the preceding chapter. The interpretation based on this analysis is therefore not leading to substantiated results.

Therefore, the following detailed analysis was established on the population distribution of dynamic molecules between the four FRET states and the population of the respective transitions. The number of transitions and the transition rates calculated from the average dwell time were also taken into account. However, they were considered as side information due to their error-prone nature discussed in the previous chapter.

To begin with, the results for the different catalytically relevant RNP compositions with and without WTAP were analyzed (Figure 71). The FRET state at 0.38 was assumed to represent the free RNA state while the other states are protein bound, based on the histogram data from chapter 4.2.2. The catalytically competent RNP formed by the METTL3/METTL14 complex with SAM as cofactor and the substrate RNA as target shows the main population (36 %) at the FRET state of 0.56. The nearby FRET states at 0.48 and 0.66 are equally populated (26 and 24 %), while the FRET state at 0.38 is less populated (15 %). Analysis of the transitions revealed the participation of all states in the dynamics of a relevant number of molecules. The transition from 0.48 to 0.56 is populated the most with 14 % of all molecules, followed by the reverse transition (11 %). Comparison to the post-catalytic RNP formed by the METTL3/METTL14 complex with SAH as cofactor and the product RNA uncovers some major differences in population distribution and dynamics. Here, the highest population was determined for the FRET state at 0.48 (42 %). Most interestingly, the population distribution between the highest FRET state (0.66) and the free RNA state is inverted in comparison to the catalytically competent RNP. The free RNA state is populated with 28 % while the FRET state at 0.66 is populated only to 5 %. This displays a significant shift of the population distribution towards lower FRET efficiencies for the post-catalytic complex, which is in line with the previous histogram data (chapter 4.2.2.4). Moreover, this is supported by the observed transitions. In contrast to the catalytically

competent RNP, transitions to or from the highest FRET state at 0.66 were not identified for a relevant number of molecules. It cannot be ruled out that these transitions are not observable due to dwell times under 100 ms, however, this would influence the population distribution only to a very minor extent. The most populated transitions in this case are between the two lower FRET states at 0.38 and 0.48 in both directions (13 and 15 % of all molecules).

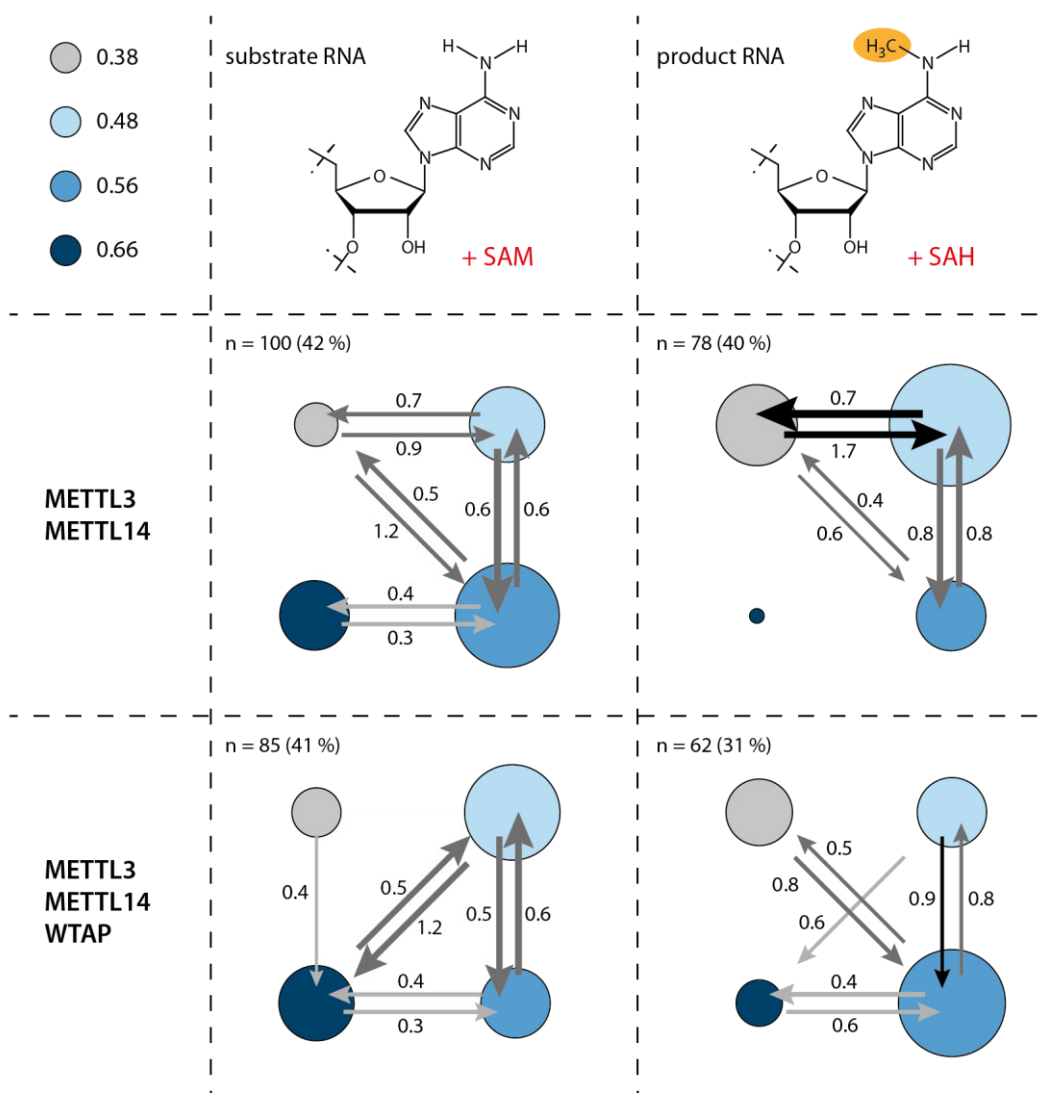


Figure 71: Summarized analysis of movie data for different RNP compositions. The size of the circles indicates the population of the respective FRET state. Arrow sizes correspond to the population of the transition and arrow color indicates the proportion of all transitions (up to 9 % light grey, 10 to 15 % grey, above 15 % black). Only relevant transitions are given. Transition rates derived from the average dwell time are noted besides the arrows and are given in s^{-1} . The number of dynamic molecules is given as n, together with the proportion of dynamic molecules in percent.

A very similar picture develops for the same RNP compositions with WTAP in the protein complex. Comparison of the catalytically competent RNP to the post-catalytic RNP shows a significant decrease in the population of the highest FRET state at 0.66 from 26 % to 16 %. An increase in population was observed for the free RNA state at 0.38, rising the population size from 0.17 % for the catalytically competent RNA to 23 % for the post-catalytic complex. This again indicates a population shift towards lower FRET efficiencies for the post-catalytic complex which is also represented in the transitions. For the catalytically competent RNP, the majority of molecules shows dynamics towards the FRET state at 0.66. Transitions from all other FRET states (0.38, 0.48 and 0.56) to 0.66 were identified for a relevant number of molecules with the transition from 0.48 to 0.66 being the most populated with 12 % of all molecules. Besides, the back-and-forth transition between 0.48 and 0.56 was also highly populated (12 and 13 %). In the post-catalytic RNP, transitions towards the FRET state at 0.66 were observed, however, to a significantly lesser extent. Upon comparison of the transition rates, the transition rate from 0.48 to 0.66 is doubled in the catalytically competent RNP (1.2 s^{-1} vs 0.6 s^{-1}) while the back transition from 0.66 to 0.56 is doubled in the post-catalytic RNP (0.3 s^{-1} vs 0.6 s^{-1}). Although the transition rates are limited in their significance, it supports the tendency of the catalytically competent complex for a shift in FRET efficiency towards the higher states. This is also in line with the histogram data (chapter 4.2.2.4) showing a shift of the main peak to higher FRET efficiencies for the catalytically competent RNP. A major difference between the two protein complexes is the significantly lower proportion of dynamic molecules for the post-catalytic state of the WTAP complex with respect to the catalytically competent RNP. For the METTL3/METTL14 complex, this was less pronounced.

Both protein complexes exhibit an analogous difference between the behavior of the catalytically competent and post-catalytic RNP. The catalytically competent RNP shifts the population distribution significantly towards the FRET state at 0.66, which is also mirrored in the transitions between the states and in the histogram data (chapter 4.2.2.4) in both cases. This tendency leads to the assumption of the FRET state at 0.66 playing an important role during catalysis. The increased FRET efficiency hereby hints towards a more compact RNA conformation than in the other protein bound state (0.48 and 0.56) or in the free RNA state (0.38).

Moreover, upon comparison of the two protein complexes, also addition of WTAP shifts the population distribution towards the FRET state at 0.66 and with that the catalytically important state. Again, this manifests not only in the population distribution but also in the transitions and

was already observed in the histogram data (chapter 4.2.2.4). One prime example for this besides the population of the 0.66 FRET state are the transitions towards this state. For METTL3/METTL14 alone, this state is only accessible from 0.56 in the catalytically competent RNP in a relevant way, while transitions from all states were observed for the WTAP containing complex.

In the next step, these differences were investigated with the abasic RNPs (Figure 72). Here, the influence of substrate and product nucleotide are eliminated and all RNPs are catalytically incompetent. For the METTL3/METTL14 RNPs, a significant reduction in dynamic molecules was observed in comparison to the catalytically relevant RNPs. Related to this, the only relevant transitions were observed between the FRET states at 0.48 and 0.56 independent from SAM or SAH. There were also no noteworthy population changes. For the WTAP containing abasic RNPs, a much more dynamic picture was observed for the transitions together with a rather equal population distribution. In this case, every FRET state was accessible in a relevant manner with a similar population of all shown transitions (6 to 10 %). This was also represented in the higher proportion of dynamic molecules than determined for the abasic RNPs without WTAP. Similar to the RNPs with the substrate and product RNAs, more dynamic molecules were observed for the RNPs with SAM. This leads to the hypothesis that addition of WTAP increases the dynamics on the RNA, which corresponds to increased protein complex dynamics.

Comparison of the abasic RNPs to the respective RNPs with substrate or product RNA reveals the influence of the substrate and product nucleotide. The strong differences in the population of the highest FRET state at 0.66 observed for the catalytically relevant RNPs is not present for the abasic RNPs. In fact, the population of the state at 0.66 for the abasic RNPs lays in between those determined for the corresponding catalytically competent and post-catalytic RNPs. For the METTL3/METTL14 RNPs this shows in a population decrease for state 0.66 from 24 % for the catalytically competent RNP to 14 % and 22 % for the abasic RNPs and down to the post-catalytic RNP with a population of only 5 %. A similar trend is observed for the WTAP containing complexes. This renders the substrate adenine and the product m⁶A important for the accessibility of the catalytically important state (0.66). With the presence of the substrate nucleotide adenine, the population of this state is significantly higher than for the presence of the product m⁶A. The mere presence of the methyl donor SAM or the post-catalytic SAH in the protein complex is not sufficient for this difference as shown with the abasic RNPs. In this case, the catalytically important RNA conformation is still moderately populated, however not as favored as for the presence of adenosine. With the presence of m⁶A the catalytically important

state is significantly less populated which renders formation of the product as “inhibitor” for the accessibility of this state.

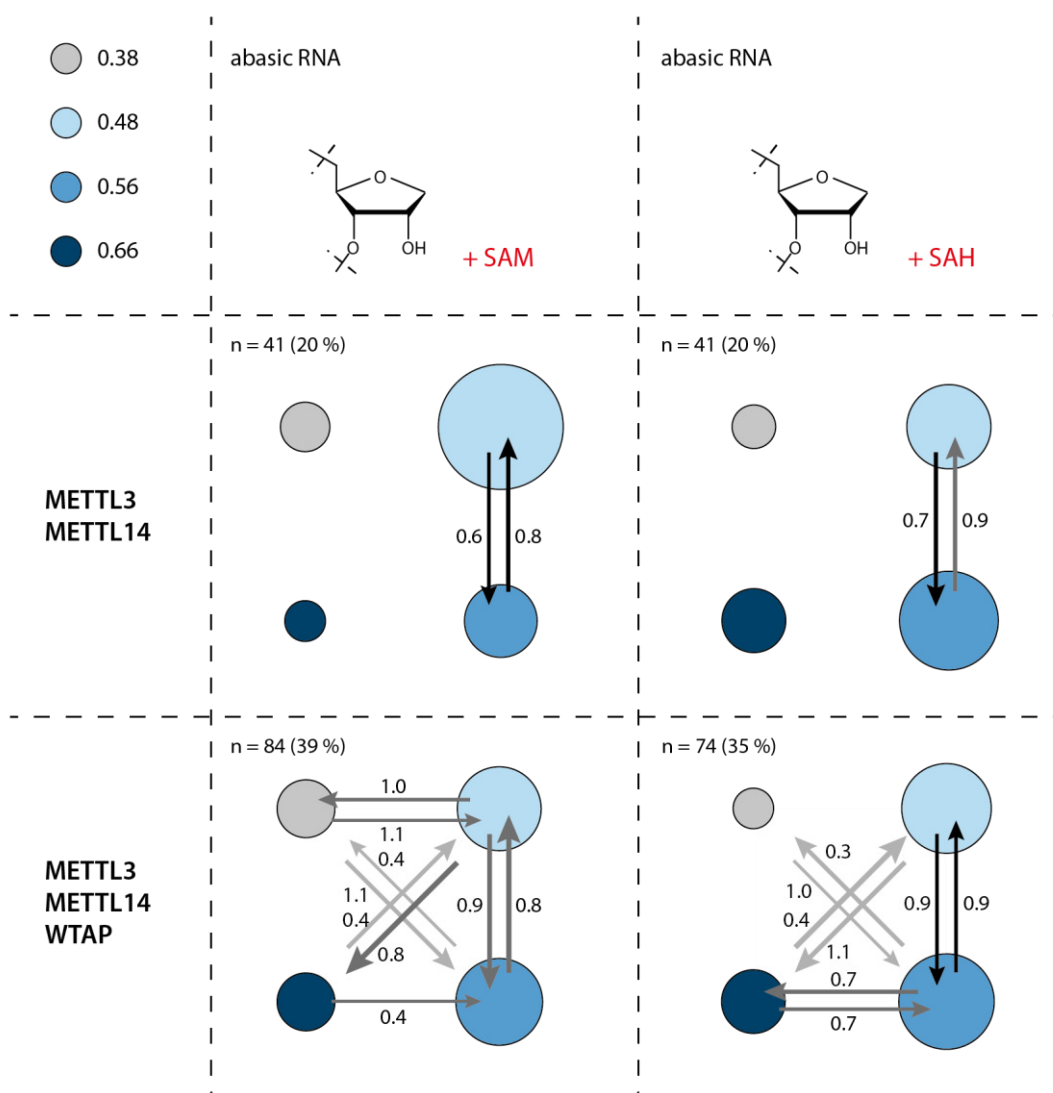


Figure 72: Summarized analysis of movie data for different RNP compositions with the abasic RNA. The size of the circles indicates the population of the respective FRET state. Arrow sizes correspond to the population of the transition and arrow color indicates the proportion of all transitions (up to 9 % light grey, 10 to 15 % grey, above 15 % black). Only relevant transitions are given. Transition rates derived from the average dwell time are noted besides the arrows and are given in s^{-1} . The number of dynamic molecules is given as n, together with the proportion of dynamic molecules in percent.

Taken together, the FRET state at 0.66 was identified as presumably important for catalysis. The accessibility of this state is dependent on the presence of adenine in the catalytically competent RNP composition and the formed m^6A in the post-catalytic RNP. Moreover, addition of WTAP

shifts the population towards the FRET state at 0.66 and significantly increases the RNP dynamics.

4.2.4 Conclusion and outlook

To gain insight into the dynamics of the METTL3/METTL14 complex, several smFRET experiments were conducted. First, three different target RNAs were designed that carry the substrate adenosine, the product m⁶A or an abasic site at the central position of the GGACU context. For all RNAs, the histogram data showed a similar shift towards higher FRET efficiencies (0.38 to ~0.52) and a widening of the peak upon addition of the protein complex, indicating a contraction of the RNA conformation and increased dynamics. Addition of SAM to the protein complex to form a catalytically competent RNP with the substrate RNA yielded a slightly higher FRET state ($E_{\text{FRET}}=0.55$) than for the post-catalytic complex (SAH, product RNA) ($E_{\text{FRET}}=0.50$). On the abasic RNA, addition of SAH or SAM to the complex did not alter the derived histograms. With WTAP added to the protein complex, the same behavior was observed in all cases but with the FRET efficiencies of the protein bound state shifted to slightly higher values in general.

In the following step, a more detailed picture was achieved by analysis of longer movie traces. Here, the dynamic behavior of the individual molecules for the different RNP compositions was evaluated. Although this was limited in statistical significance as discussed previously, there were still some key findings to report on. For all complexes four FRET states were identified that are involved in the dynamic behavior: an RNA only state at a FRET efficiency of 0.38 and three protein bound states ($E_{\text{FRET}}=0.48$, 0.56 and 0.66). Upon comparison of the catalytically competent RNPs (SAM/substrate RNA) with the respective post-catalytic RNP (SAH/product RNA), the highest FRET state ($E_{\text{FRET}}=0.66$) is significantly less populated for the latter. This is also mirrored in the accessibility by transitions and renders this state as presumably relevant for catalysis. The strong difference in population distribution between the SAM and SAH bound complexes was not observed for the catalytically incompetent RNPs with the abasic RNA. This hints towards a significant influence of the substrate and product nucleotide on the accessibility of the presumably catalytically relevant state ($E_{\text{FRET}}=0.66$). SAM or SAH addition alone was not sufficient to induce the observed differences in population distribution. Addition of WTAP to the complex increased the observed dynamics and shifts the population towards the FRET state at $E_{\text{FRET}}=0.66$.

These key findings together with the sensitivity of the complex towards RNA structure and the role of the individual RNA binding domains on catalysis give an estimation for a possible mechanism during catalysis. The FRET state at an $E_{\text{FRET}}=0.66$ was assumed to be relevant in the catalytic cycle. The successive increase in FRET efficiency from the RNA only state ($E_{\text{FRET}}=0.38$) over the other protein bound states ($E_{\text{FRET}}=0.48, 0.56$) to this state represents a step-by-step contraction of the RNA conformation. Based on the importance of both, the RGG domain (METTL14) and the ZFD (METTL3) in RNA binding and catalysis,^{17–20} this contraction might result from RNA interaction with these domains. In a first step, the RNA is bound by the complex through the ZFD and RGG domains and further on might be contracted by correlated dynamics of the two domains. This “squeezing” of the RNA presumably results in flipping of the adenosine into the active site where it is bound by the DPPW motif and the two active site loops for catalysis^{16–18} (Figure 73). After transfer of the methyl group, the formed m^6A most likely exhibits a significantly lower affinity to the active site than substrate adenine and flips back into the RNA structure. The rebinding of the methylation product in the active site is presumably sterically disfavored, which leads to a less effective squeezing motion and the significant population reduction for the $E_{\text{FRET}}=0.66$ state. This again hints towards a general RNA binding ability of the RNA binding domains, but a very specific architecture of the active site for catalysis on DRACH motifs. For the abasic RNA, this squeezing does not result in steric clashes with the active site due to the lacking base which enables this RNA to access the catalytic state. However, there is

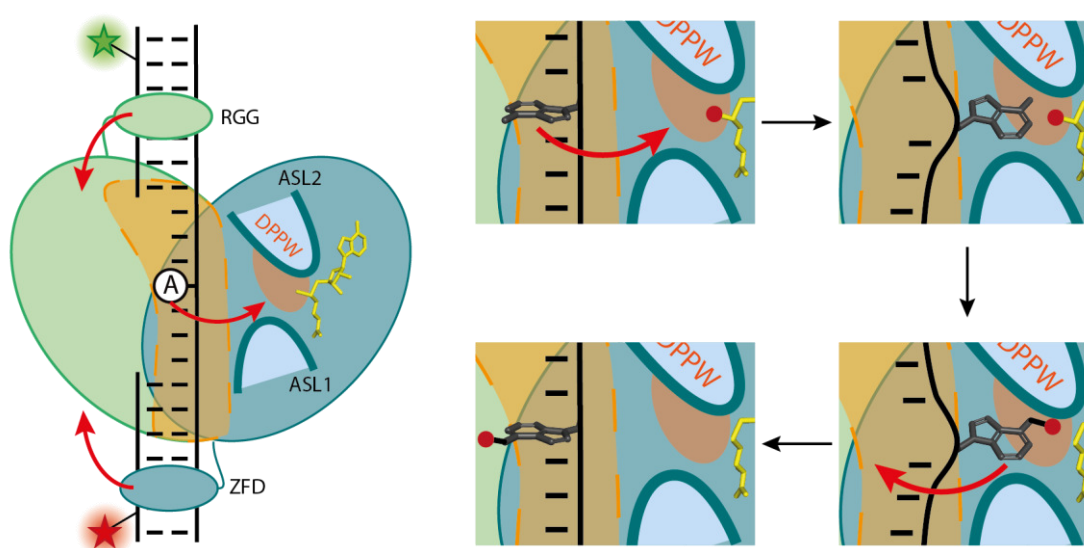


Figure 73: Hypothesized model for the RNA motion during catalysis. The RGG and zinc finger domains bind the RNA on both ends and “squeeze” the RNA. This bending flips out the target adenosine to be bound by the DPPW motif and the active site loops in the active site. After the methyl group transfer, the methylated adenosine is flipped out of the active site.

no effective binding of a base to stay in this compacted state which explains the reduced population of the 0.66 state in comparison to the catalytically competent RNPs. Moreover, this hypothesis is supported by the structural preference of the complex. The reduced catalytic activity on DRACH sequences in a hairpin motif in contrast to those in a loose duplex or unstructured RNA is presumably strongly related to the possibility for the RGG domain and the ZFD to anchor the RNA on both sides of the target sequence (see chapter 0). If this anchoring on both sides is restricted by the RNA secondary structure, this would also limit the squeezing motion of the complex and with that presumably catalytically important dynamics. For duplexes in full complementarity, this binding on two sides is possible, however the target adenine is tightly bound in WC base pairing. The strong base pairing itself disables flipping out the target adenosine without any helicase activity but moreover, the rigid form of a complementary duplex should also efficiently prevent squeezing of the RNA.

There are two additional protein bound RNA states ($E_{\text{FRET}}=0.48, 0.56$) that did not show any conclusive differences between the different RNP compositions. However, these states and the transitions between these states were highly populated in all cases. Moreover, the highest FRET state at $E_{\text{FRET}}=0.66$ was mostly inaccessible directly from the RNA only state ($E_{\text{FRET}}=0.38$) which shows the importance of these states to enter the catalytically relevant state. These intermediate steps might originate from several things. Most likely, the RNA binding through the ZFD as well as the RGG domain transfers the protein complex dynamics onto the RNA. This “bending and stretching” motions might help to further increase the RNA dynamics and to loosen the target adenine from any transient (WC) binding before catalysis. In this case, the transitions between the two states would be strongly dependent on the measurement temperature as there is no ATP consumption necessary for catalysis. Another possibility includes the stepwise binding of the ZFD and the RGG domain to subsequently contract the RNA. Moreover, this might also be related to a motion for correct positioning of the RNA near the active site after the initial binding of the protein complex.

The described hypothesis is very well in line with the findings reported in this thesis. However, especially the analysis of the RNP dynamics was strongly limited in its statistical significance. Besides the repetition of the smFRET experiments and analysis for a larger number of molecules, further experiments should be conducted to confirm or disprove the formulated hypothesis. On the side of follow up smFRET experiments, this could include measurements with an increased time resolution to detect a possible strict order of the observed states (0.38 to 0.48 to 0.56 to 0.66). Relatedly, potentially faster transitions that were not resolved in the given experiments

due to very short dwell times should become detectable. Moreover, measurements on the substrate RNA and the METTL3/METTL14 complex supplemented with the inhibitor STM2457 could be added. This RNP composition should efficiently prevent access to the FRET state at $E_{\text{FRET}}=0.66$ as a result of a blocked active site through binding of STM2457. Additionally, the truncated protein constructs should be tested on the different RNAs, focusing on the RGG domain and ZFD depleted complexes (see chapter 3.2).

The expression and purification of fluorescently labeled protein complexes would enable a set of presumably very valuable experiments. This could include the attachment of one or both fluorophores of a FRET pair to the METTL3/METTL14 complex on a variety of protein domains. To further validate the set-up hypothesis this should include labeling of the RGG domain, the ZFD, both active site loops as well as the MTDs on the upper and lower site of the binding groove. These labeling sites could be used in a variety of combinations or individually with a fluorescently labeled RNA. Especially labeling of the RGG domain and the ZFD individually or combined would be of great relevance to support the given hypothesis. For the RNA, addition of constructs with different secondary structures would be very interesting.

5 Materials and methods

5.1 Buffers and Media

5.1.1 Media

Media for bacterial and insect cell culture were supplemented with antibiotics. Insect cell culture media were supplemented with 0.5 % penicillin/streptomycin (10.000 U/mL, *ThermoFischer Scientific*) in addition to 5 % fetal bovine serum (*ThermoFischer Scientific*), unless stated otherwise. For bacterial cell cultures, antibiotics corresponding to the introduced resistance were used in the working concentration suggested by the manufacturer.

Table 10: Composition of media used within this thesis.

Medium name	Components
SB medium	32 g/L tryptone, 20 g/L yeast extract, 5 g/L NaCl, pH 7.0
LB medium	10 g/L tryptone, 5 g/L yeast extract, 5 g/L NaCl, pH 7.0
SOC outgrowth medium	<i>Purchased from New England Biolabs</i>
Sf-900 III medium	<i>Purchased from ThermoFischer</i>
Sf-900 II medium	<i>Purchased from ThermoFischer</i>
TC100 growth medium	<i>Purchased from ThermoFischer</i>

5.1.2 Buffers

All buffers were sterile filtered before use. Buffers used in protein purification or purification with an HPLC column were additionally degassed.

Table 11: Composition of buffers used within this project.

Buffer name	Components
Activity assay buffer	50 mM Tris, 20 mM KCl, 1 % glycerol, 10 μ M n-Dodecyl β -D-maltoside, 1.4 mM β -mercaptoethanol, pH 8.0
FA buffer	90 vol% formamide, 0.1 % bromphenol blue, 0.1 % xylene cyanol in TBE buffer
HisTrap buffer A	50 mM Tris/HCl, 200 mM KCl, 10 % glycerol, 100 μ M n-Dodecyl β -D-maltoside, 14 mM β -mercaptoethanol, 10 mM imidazole, pH 8.0 (at 4°C)

Materials and methods

Buffer name	Components
HisTrap buffer B	50 mM Tris/HCl, 200 mM KCl, 10 % glycerol, 100 μ M n-Dodecyl β -D-maltoside, 14 mM β -mercaptoethanol, 500 mM imidazole, pH 8.0 (at 4°C)
Imaging buffer	50 mM Tris, 20 mM KCl, 1 % glycerol, 10 μ M n-Dodecyl β -D-maltoside, 1.4 mM β -mercaptoethanol, Trolox saturated, pH 8.0
MES buffer	50 mM 2-(N-morpholino)ethanesulfonic acid, 50 mM Tris, 0.1 % SDS, 0.1 mM EDTA, pH 7.3
P1 buffer	20 mM NH ₄ OAc, 0.5 mM ZnCl ₂ , pH 5.3
PBS buffer	137 mM NaCl, 2.7 mM KCl, 10 mM Na ₂ HPO ₄ , 1.8 mM KH ₂ PO ₄ , pH 7.4
SDS buffer	25 mM Tris/HCl, 250 mM glycerol, 0.1 % SDS, pH 6.8
SEC buffer	50 mM Tris/HCl, 200 mM KCl, 10 % glycerol, 100 μ M n-Dodecyl β -D-maltoside, 14 mM β -mercaptoethanol, pH 8.0 (at 4°C)
T50 buffer	50 mM NaCl, 10 mM Tris, pH 8.0
TA buffer	50 mM Tris/OAc, 100 mM NaOAc, pH 8.3
TAE buffer	40 mM Tris, 20 mM acetic acid, 1 mM EDTA, pH 8.3
TBE buffer	90 mM Tris, 90 mM boric acid, 2 mM EDTA, pH 8.0
TEAA	100 mM Triethylammonium acetate, pH 7.0
TLC buffer A ¹³³	2-Methylpropionic acid / NH ₄ (25 % in H ₂ O) / H ₂ O (66 : 1 : 33 [v:v:v])
TLC buffer C ¹³³	2-propanol / HCl (37 % in H ₂ O) / H ₂ O (70 : 15 : 15 [v:v:v])
Transcription buffer	250 mM Tris/glutamate, 30 mM MgOAc, 10 mM DTT, 2 mM spermidine, pH 8.3
Transformation buffer	10 mM PIPES (pH 6.7), 250 mM KCl, 55 mM MnCl ₂ , 15 mM CaCl ₂
WB anode buffer	75 mM Tris, 20 % MeOH, pH 7.4
WB cathode buffer	40 mM ϵ -aminocaproic acid, 25 mM Tris, 20 % MeOH, pH 9.0

5.2 General biochemical methods

5.2.1 Concentration determination of biomolecules

Concentrations of nucleic acids and proteins were determined by UV/vis spectroscopy using a NanoDrop™ One (*ThermoFisher*). The absorbance was measured at a wavelength of 260 nm for nucleic acids and 280 nm for proteins. Measured protein concentrations were corrected for the calculated extinction coefficient (ProtParam, *Expasy*) of the protein.

5.2.1.1 Fluorescently labeled samples

In addition to the concentration of biomolecules, the fluorescent dye concentration for fluorescently labeled samples was determined. Using implemented programs of the NanoDrop™ One (*ThermoFisher*) for the specific dyes, if required, the proportion of labeled sample was calculated from the measurements.

5.2.1.2 Radioactively labeled samples

The concentration of radioactively labeled samples was determined via radioactivity incorporation. Radiation intensity was measured with a scintillation counter (Tri-Carb 2100TR, *Perkin Elmer*). Radioactively labeled samples were subject to UV/vis spectroscopy concentration determination only after radioactive decay. This was used to verify the preceding concentration determination from radioactivity.

5.2.2 Precipitation of nucleic acids

Nucleic acids were purified from other compounds in solution by precipitation with either ethanol or isopropanol. In both cases, the solution was adjusted to 0.3 M sodium acetate or 0.5 M ammonium acetate. Then 2.5 volume equivalents of ethanol or 1 volume equivalent of isopropanol were added and the mixture was incubated for at least 1 h at -20 °C. The precipitated nucleic acids were pelleted by centrifugation at 21000 g and 0 °C for at least 30 min, the supernatant was removed, the pellet dried in vacuum and afterwards dissolved in ddH₂O.

5.2.3 Electrophoretic methods

5.2.3.1 Analytical polyacrylamide gel electrophoresis

RNA and short DNA constructs were analyzed via denaturing polyacrylamide gel electrophoresis (PAGE) regarding size and charge. The gels were polymerized by adding 1 vol-% of 10 % ammonium persulfate solution (APS) and 0.1 vol-% of tetramethylethylenediamine (TEMED) to the gel solution (7 M urea, 10-20 % acryl-/bisacrylamide mix (29:1) in TBE buffer). Samples were mixed 1:1 with FA buffer, denatured at 95 °C for 5 min and 150 ng were loaded into the well. Separation was performed with TBE buffer as running buffer and a constant voltage of 240 V for 45 min. After separation, nucleic acids were stained with GelRed™ (1:10.000 in H₂O) and visualized with a Gel Imager (GelDoc™ XR+, Biorad). Additionally, fluorescently labeled samples were imaged with a Typhoon 9400 (*GE Healthcare life sciences*) beforehand.

5.2.3.2 Preparative polyacrylamide gel electrophoresis

RNA samples produced by *in vitro* transcription or ligation were separated from side products by preparative RNA-PAGE. Gels for this purpose were cast and run as described for analytical PAGE. However, samples were kept as concentrated as possible and loaded completely onto the gel. After separation, desired RNA bands were identified under UV light and cut out of the gel in thin slices. The RNA was then eluted with 0.3 M sodium acetate at room temperature and 600 rpm shaking overnight. After filtration of the supernatant through a filter tube (Nanosep MF, 0.45 µm, Pall), the RNA was precipitated with ethanol.

5.2.3.3 Native polyacrylamide gel electrophoresis

Native PAGE was used to analyze RNA and short DNA constructs (≤ 100 bp) in their native folding state. The gels were cast as described for analytical PAGE, however, TA buffer was used instead of TBE buffer. Samples were adjusted to a final concentration of 10 vol-% glycerol and 100 ng nucleic acid were loaded per well. For separation, TA buffer was used as running buffer and under water cooling a constant voltage of 50 V was applied. Depending on the used amount of acryl-/bisacrylamide mix, separation times between 1.5 h and 4 h were chosen. After separation, nucleic acids were stained as described for analytical PAGE.

5.2.3.4 Agarose gel electrophoresis

Large DNA constructs and plasmids were analyzed using agarose gel electrophoresis. Agarose gels were cast by heating agarose in TAE buffer and letting it solidify with a comb inserted. Depending on the desired separation range, between 0.8 % and 2 % agarose were used. Under a constant voltage of 120 V samples were separated for 40 min with TAE as running buffer.

Nucleic acids were stained with GelRed™ (1:10.000 in H₂O) and visualized with a Gel Imager (GelDoc™ XR+, Biorad).

5.2.3.5 Electrophoretic mobility shift assay

Electrophoretic mobility shift assays were performed to analyze the binding between RNA and the DNA handles used for FRET experiments or for the investigation of RNA-protein complexes. For all samples, at least one fluorescently labeled binding partner was present to enable a very sensitive analysis and with that low sample amounts. All results were visualized with a Typhoon 9400 imager (*GE Healthcare Life Sciences*).

For the smFRET construct, samples consisting of 1.8 μM RNA FU and 0.9 μM, 1.8 μM or 3.6 μM of one or both fluorescently labeled DNA handles (FHA, FHD) were mixed respectively and then analyzed via native PAGE (15 %, 60 V, 4 h, 100 ng RNA per lane).

Samples for analysis of the binding between RNA and the different protein constructs were prepared with 8 nM of RNA U1-Cy5. For all shortened protein constructs, the RNA was mixed with 80 nM and 500 nM protein each. The full-length complex was prepared in 80 nM, 250 nM, 500 nM and 800 nM concentration. Protein constructs M3ΔNTD and M3ΔZF1 were also tested on RNA D4-Cy5 and RNA OT-Cy5 with the already given concentrations. All samples were incubated for 30 min at 25 °C prior to analysis via native PAGE (8% prepared with 75:1 Acrylamide/Bisacrylamide mix, 50 V, 2.5 h).

5.2.3.6 SDS polyacrylamide gel electrophoresis

Protein samples were analyzed regarding their size and purity via SDS polyacrylamide gel electrophoresis (SDS-PAGE). Gels were either purchased as precast gradient gels (NuPAGE™ 4 to 12 %, Bis-Tris, 15 well, *Invitrogen*) or self-cast directly before running the gel. Self-cast gels were prepared as two layer gels with a stacking gel phase (upper ¼, 4 % acryl-/bisacrylamide mix (29:1) in SDS buffer) above the resolving gel phase (lower ¾, 12 % acryl-/bisacrylamide mix (29:1) in SDS buffer). The gel solution was polymerized with 1.25 vol-% of 10 % APS and 0.125 vol-% TEMED. The protein samples to analyze were mixed with dye (0.1 % bromophenol blue, 0.1 % xylene cyanol) containing SDS buffer and denatured for 5 min at 95 °C. Electrophoresis was carried out at 200 V for 45 min with the corresponding running buffer (MES buffer for precast gels and SDS buffer for self-cast gels). The protein bands were visualized with Coomassie staining (0.1 % Coomassie Blue, 10 % ethanol and 5 % acetic acid)

5.2.3.7 Western blot

The analysis of His-, FLAG- or Strep-tagged proteins was complemented by western blots. After SDS gel electrophoresis, the proteins from the unstained gel were transferred onto a PVDF membrane. Therefore, the PVDF membrane (Amersham Hybond, 0.45 μm , *GE Healthcare Life Sciences*) was first cut to the same size as the gel and equilibrated for 1 min in methanol, then 1 min in cathode buffer and placed onto of a filter paper for western blots (Extra Thick Western Blotting Filter Paper 20 cm x 20 cm, *ThermoFisher*, cut to the same size as the gel), soaked in anode buffer. The unstained protein gel was equilibrated in cathode buffer for 1 min, placed on top of the PVDF membrane and covered with a filter paper for western blots, soaked in cathode buffer. The proteins were transferred to the membrane by applying 1.1 mA current per 1 cm^2 gel for 55 min. Afterwards, the membrane was blocked for unspecific protein binding by incubation with 10 mL of a milk powder solution (5 % (w/v) in PBS) for 20 min. To specifically stain the tagged proteins, the membrane was covered (10 mL) with the corresponding primary antibody solution (from mouse, 1:3000 in milk powder solution) and incubated over night. Then, the antibody solution was removed and after three washing steps with PBS (10 mL for 5 min each, shaking), 10 mL of the secondary antibody (horseradish peroxidase conjugated goat / anti mouse, *Jackson Immuno Research*) in milk solution (1:10000) was added to the membrane. After incubation for 1 h, removal of the secondary antibody and three additional washing steps (5 min in 10 mL PBS each, shaking), a mixture of 1 mL luminol and 1 mL peroxide (SuperSignal West, *ThermoFisher*) was added to the membrane and incubated for 5 min. The protein bands were visualized by chemoluminescence detection with a fluorescence imager (Advanced Fluorescence Imager, *Intas*).

5.2.4 Polymerase chain reaction

DNA templates for RNA *in vitro* transcription or protein coding DNA sequences were amplified via polymerase chain reaction (PCR). The reaction mixture contained 0.1 μ M DNA template for non plasmid DNA or 100 ng for plasmid DNA, 0.5 μ M forward Primer, 0.5 μ M backward Primer and 200 μ M of each dNTP (Deoxynucleotide (dNTP) Solution Mix, *New England Biolabs*) in Phusion HF Buffer (*New England Biolabs*). Last, 1 μ L Phusion DNA Polymerase (home made by AK Schwalbe) was added per 100 μ L batch. The reaction was performed by using a temperature protocol described in Table 12. The reaction product was analyzed via native PAGE (5.2.3.3) for short DNA constructs (≤ 100 bp) and via agarose gel electrophoresis (5.2.3.4) for longer DNAs. The DNA templates and primers for PCR reactions can be found in the corresponding sections.

Table 12: Temperature protocol for PCR reaction.

Temperature	Time
98 °C	2 min
T_M	15 s
72 °C	30 s per kbp
98 °C	15 s
72 °C	1 min

5.2.5 *In vitro* transcription

Unlabeled RNA samples were produced by *in vitro* transcription. For this, 10 vol-% of corresponding PCR product and 4 mM of each NTP were mixed in transcription buffer and 32 ng/ μ L T7 polymerase (homemade by AK Schwalbe) were added last. Optionally, 5-20 vol-% dimethylsulfoxide were added to the reaction. The reaction was allowed to proceed for 4 h to overnight at 37 °C and 300 rpm. To get rid of the DNA template, 2 U Turbo DNase (*ThermoFisher*) were added to the mixture and the reaction was continued for another 30 min under the same conditions. The insoluble magnesium pyrophosphate ($Mg_2P_2O_7$) indicating a successful *in vitro* transcription was separated by centrifugation (1 min, 21000 g, 4 °C). The *in vitro* transcribed RNA in the supernatant was precipitated with ethanol. To separate the desired construct from side products, a preparative PAGE (5.2.3.2) was done. The purified product was analyzed with analytical PAGE (5.2.3.1).

Materials and methods

All primers and templates needed for *in vitro* transcription of the RNAs are given in Table 13. For final RNA sequences see Table 14.

Table 13: Sequences of all DNA templates used for *in vitro* transcription of RNAs. Most of the templates were also used as DNA splint during DNA splinted ligation.

Final RNA	DNA template (5' to 3')	Reverse primer (5' to 3')
U1	CCT CGT ATC CAG TCC TTG TGC ATC C TA TAG TGA GTC GTA TTA	CCT CGT ATC CAG TCC
U1.1	CCT CGT ATC CAG TTT TTG TGC ATC C TA TAG TGA GTC GTA TTA	CCT CGT ATC CAG TTT
U1.2	CCT CGT ATC CAG TCT TTG TGC ATC C TA TAG TGA GTC GTA TTA	CCT CGT ATC CAG TCT
U1.3	CCT CGT ATC CAG TTC TTG TGC ATC C TA TAG TGA GTC GTA TTA	CCT CGT ATC CAG TTC
D1 strand 1	GGA CGG ATG ACA AAG TCC CTC GCA GCC GCC TAT AGT GAG TCG TAT TA	GGA CGG ATG ACA A
D1 strand 2	GCG GCA AAG ACA GGT GGC CGC CCT ATA GTG AGT CGT ATT A	GCG GCA AAG ACA
D2 strand 1	GTT GGT AAA GGC AGT CCC CAG ATG GCA CC TAT AGT GAG TCG TAT TA	GTT GGT AAA GGC AGT
D2 strand 2	CCT TCT GGT GCC TCT CCT ATA GTG AGT CGT ATT A	CCT TCT GGT GCC
D3 strand 1	see U1	
D3 strand 2	purchased directly as RNA	
D4 strand 2	see D1 strand 2	
H1	GCA GAG AGA GAA GCA AAA CGA GTC CCT GCG GGG TG CC TAT AGT GAG TCG TAT TA	GCA GAG AGA GAA GCA
H2	GAT GCA CTC AGT CCT TGT GCA TCC TAT AGT GAG TCG TAT TA	GAT GCA CTC AGT CCT
H3	GGT CAG AGA AGG GTT CGC GTC CTT CCT TGA AG CC TAT AGT GAG TCG TAT TA	GGT CAG AGA AGG GTT
H4	ATA TTG TTA GTA GGT AGT CTG AAT GCG AGC TAT CTC TAA CC TAT AGT GAG TCG TAT TA	ATA TTG TTA GTA GGT AGT CTG
H4.1	ATA TTG TTA GTA GGT AGT CCG AAT GCG AGC TAT CTC TAA CC TAT AGT GAG TCG TAT TA	ATA TTG TTA GTA GGT AGT CCG
H5	CCA CAC TGT CAA ATT GTT TAA GCT TGG ATT TGA CAC C TAT AGT GAG TCG TAT TA	CCA CAC TGT CAA ATT
FU	GGT GGT TCG GCT CGC TTA GTC CTT CGC TCG GCT TGG TGG TCT CCC TA TAG TGA GTC GTA TTA	GGT GGT TCG GCT

5.2.6 Fluorescent labeling of nucleic acids

The nucleic acids to be fluorescently labeled were purchased (RNAs from *Dharmacon*, DNAs from *Eurofins Genomics*) with a C6-amino modification at different positions (for constructs see Table 15). First, 30 nmol of nucleic acids were precipitated with sodium acetate and ethanol and dissolved in 20 μ L fresh 0.1 M NaHCO₃. The NHS reactive dye to be coupled (Amersham CyDye™ mono-reactive NHS Ester, Cy3 or Cy5, *Cytiva*) were dissolved in 20 μ L DMSO and added to the nucleic acid. The reaction mixture was mixed, incubated for 90 min at room temperature and afterwards the nucleic acid was precipitated with ethanol. For RNA, a deprotection step following the manufacturer's instructions was added. To separate the fluorescently labeled nucleic acid from unreacted starting material, a chromatographic purification over a Kromasil C8 HPLC column (250 x 4.6 mm) was performed. The precipitated nucleic acid was dissolved in 400 μ L TEAA buffer and loaded onto the equilibrated column. The nucleic acids were separated over a 40 mL gradient from TEAA buffer to a mixture of 50 % TEAA buffer with 50 % acetonitrile. The absorption at 260 nm for nucleic acids and at 550 nm for Cy3 or at 650 nm for Cy5 samples was monitored. Fractions of 300 μ L were taken and analyzed via PAGE. Desired fractions were combined and the acetonitrile was removed in the vacuum centrifuge for 30 min. Then, the nucleic acid was precipitated with ethanol and resuspended in H₂O.

5.3 Nucleic acid constructs

All RNA constructs used within this project were either produced by run off *in vitro* transcription (5.2.5) or purchased with a modification or as fragment for ligation. Table 14 shows the sequences of all unmodified RNA constructs produced by *in vitro* transcription or ligation, while all purchased modified nucleic acid constructs are given in Table 15.

Table 14: Sequences of all RNA constructs used for methylation assays and melting point analysis. The methylation target sequence is printed in bold type. Some sequences were elongated for transcription (indicated in brackets) in comparison to their ligated version.

RNA construct	Sequence (5' to 3')
U1	GGA UGC ACA AGG ACU GGA UAC GAG G
U1.1	GGA UGC ACA AAA ACU GGA UAC GAG G
U1.2	GGA UGC ACA AAG ACU GGA UAC GAG G
U1.3	GGA UGC ACA AGA ACU GGA UAC GAG G
D1	strand 1: (G)GCG GCU GCG AGG GAC UUU GUC AUC CGU CC strand 2: GGG CGG CCA CCU GUC UUU GCC GC
D2	strand 1: (GG)UGC CAU CUG GGG ACU GCC UUU ACC AAC strand 2: AGA GGC ACC AGA AGG
D3	strand 1: GGA UGC ACA AGG ACU GGA UAC GAG G strand 2: CCU CGU AUC CAG UCC UUG UGC AUC C
H1	(GG)CAC CCC GCA GGG ACU CGU UUU GCU UCU CUC UCU GC
H2	GGA UGC ACA AGG ACU GAG UGC AUC
H3	(GG)CUU CAA GGA AGG ACG CGA ACC CUU CUC UGA CC
H4	(GG)UUA GAG AUA GCU CGC AUU CAG ACU ACC UAC UAA CAA UAU
H4.1	(GG)UUA GAG AUA GCU CGC AUU CGG ACU ACC UAC UAA CAA UAU
H5	(G)GUG UCA AAU CCA AGC UUA A ACA AUU UGA CAG UGU GG
FU	GGG AGA CCA CCA AGC CGA GCG AAG GAC UAA GCG AGC CGA ACC ACC

Table 15: Sequences of all RNA and DNA constructs purchased with modifications (for dye coupling) used in smFRET measurements or EMSAs. If present, the methylation target sequence is printed in bold type.

Construct	Type	Sequence (5' to 3')	Coupled dye
U1-Cy5	RNA	AmC6-GGA UGC ACA AGG ACU GGA UAC GAG G	Cy5
OT-Cy5	RNA	AmC6-GUG UAG UAU CUG UUC UUU UCA G	Cy5
D4-Cy5	RNA	strand 1: AmC6-GGA UGC ACA AGG ACU GGA UAC GAG G strand 2: GGG CGG CCA CCU GUC UUU GCC GC	Cy5
FUP	RNA	GGG AGA CCA CCA AGC CGA GCG AAG Gm⁶AC UAA GCG AGC CGA ACC ACC	
FUA	RNA	GGG AGA CCA CCA AGC CGA GCG AAG GXC UAA GCG AGC CGA ACC ACC; with X=ribo-abasic spacer (only sugar ring)	
FHA	DNA	CGC (AmC6dT)CG GCT TGG TGG TCT CCC-Biotin	Cy5
FHD	DNA	GGT GG(AmC6dT) TCG GCT CGC	Cy3

5.4 RNA melting point analysis

Melting point analysis was performed for several RNAs by CD spectroscopy. The RNAs to be analyzed were produced by *in vitro* transcription (5.2.5) and dissolved in assay buffer (50 mM Tris, pH 8 (4°C), 20 mM KCl, 1 % Glycerol, 1.4 mM β-Mercaptoethanol) to a final concentration of 10 μM. The measurements were performed in cuvettes (Hellma, absorption cuvette Quarz Suprasil®, d=1 mm, V=350 μL) on a spectropolarimeter (J-810, Jasco) between 15 °C and 95 °C in 0.1 °C increments (heating/cooling rate: 0.5 °C/min).

The measured data was normalized and plotted according to equation (10):¹⁴⁹

$$\theta_n(T) = \frac{\text{measured value}(T) - \text{lowest value}}{\text{highest value} - \text{lowest value}} \quad (10)$$

Θ_n: portion of folded RNA (1 = all RNA is folded)
 T: temperature in °C
 Highest value: average of first ten points measured
 Lowest value: average of last 10 points measured

The plotted data was fitted in Origin with a sigmoidal curve with three parameters (equation (11)):¹⁴⁹

$$\theta_n(T) = \frac{a}{1 + \exp\left(-\frac{T - T_M}{b}\right)} \quad (11)$$

a, b: variable parameter

T_M : melting temperature

5.5 Protein constructs

5.5.1 Amino acid sequences

All amino acid sequences are given without expression tags. The corresponding gene coding sequences can be found in chapter 6.1.

METTL3

1	MSDTWSSIQA	HKKQLDSLRE	RLQRRRKQDS	GHLDLRNPEA	ALSPTFRSDS
51	PVPTAPTSGG	PKPSTASAVP	ELATDPELEK	KLLHHLSDLA	LTLPTDAVSI
101	CLAISTPDAP	ATQDGVESLL	QKFQAAQELIE	VKRGLLQDDA	HPTLVTYADH
151	SKLSAMMGAV	AEKKGPGGEVA	GTVTGQKRRRA	EQDSTTVAAF	ASSLVSGLNS
201	SASEPAKEPA	KKSRKHAASD	VDLEIESLLN	QQSTKEQQSK	KVSQEILELL
251	NTTTAKEQSI	VEKFRSRGRA	QVQEFCDYGT	KEECMKASDA	DRPCRKLHFR
301	RIINKHTDES	LGDCSFLNTC	FHMDTCKYVH	YEIDACMDSE	APGSKDHTPS
351	QELALTQSVG	GDSSADRLFP	PQWICCDIRY	LDVSILGKFA	VVMADPPWDI
401	HMELPYGTLT	DDEMRRLNIP	VLQDDGFLFL	WVTGRAMELG	RECLNLWGYE
451	RVDEIIVVKT	NQLQRIIRTG	RTGHWLNHGK	EHCLVGKGN	PQGFNQGLDC
501	DVIVAEVRST	SHKPDEIYGM	IERLSPGTRK	IELFGRPHNV	QPNWITLGNQ
551	LDGIHLLDPD	VVARFKQRYP	DGIISKPKNL		

M3ΔNTD: aa 260 – aa 580

M3ΔZF1: aa 307 – aa 580

M3MTD: aa 352 – aa 580

METTL14

1	MDSRLQEIRE	RQKLRRQLLA	QQLGAESADS	IGAVLNSKDE	QREIAETRET
51	CRASYDTSAP	NAKRKYLDEG	ETDEDKMEEY	KDELEMQQDE	ENLPYEEIY
101	KDSSTFLKGT	QSLNPHNDYC	QHFVDTGHRP	QNFIRDVGLA	DRFEEYPKLR
151	ELIRLKDELI	AKSNTPPMYL	QADIEAFDIR	ELTPKFDVIL	LEPPLEEYR
201	ETGITANEKC	WTWDDIMKLE	IDEIAAPRSF	IFLWCGSGEG	LDLGRVCLRK
251	WGYRRCEDIC	WIKTNKNNPG	KTKTLDPKAV	FQRTKEHCLM	GIKGTVKRST
301	DGDFIHANVD	IDLITTEPE	IGNIEKPVEI	FHIIHFCLG	RRRLHLFGRD
351	STIRPGWLV	GPTLTNSNYN	AETYASYFSA	PNSYLTGCTE	EIERLRPKSP
401	PPKSKSDRGG	GAPRGGGRGG	TSAGRGRERN	RSNFRGERGG	FRGGRRGAHR
451	GGFPPR				

M14ΔRGG: aa 1 – aa 395**M14ΔNTD:** aa 164 – aa 456**M14MTD:** aa 164 – aa 395**WTAP**

1	MTNEEPLPKK	VRLSETDFKV	MARDELILRW	KQYEAYVQAL	EGKYTDLNSN
51	DVTGLRESEE	KLKQQQQESA	RRENILVMRL	ATKEQEMQEC	TTQIQYLKQV
101	QQPSVAQLRS	TMVDPAINLF	FLKMKGELEQ	TKDKLEQAQN	ELSAWKFTPD
151	SQTGKKLMAK	CRMLIQENQE	LGRQLSQGRI	AQLEAELALQ	KKYSEELKSS
201	QDELNDFIIQ	LDEEVEGMQS	TILVLQQQLK	ETRQQLAQYQ	QQQSASAPS
251	TSRTTASEPV	EQSEATSKDC	SRLTNGPSNG	SSSRQRTSGS	GFHREGNTTE
301	DDFPSSPGNG	NKSSNSSEER	TGRGGSGYVN	QLSAGYESVD	SPTGSENSLT
351	HQSNDTDSSH	DPQEEKAVSG	KGNRTVGSRH	VQNGLDSSVN	VOGSVL

shortWTAP: aa 1 – aa 151, but S151 mutated to R**ZC3H13**

1	MSKIRRKVTV	ENTKTISDST	SRRPSVFERL	GPSTGSTAET	QCRNWLKTGN
51	CLYGNTCRFV	HGPSRPGKGY	SSNYRRSPER	PTGDLRERMK	NKRQDVDTEP
101	QKRNTEESS	PVRKESSRGR	HREKEDIKIT	KERTPESEEE	NVEWETNRDD
151	SDNGDINYDY	VHELSEMKR	QKIQRELMKL	EQENMEKREE	IIKKEVSPE
201	VVRSKLSPSP	SLRKSSKSPK	RKSSPKSSSA	SKKDRKTSAV	SSPLLDQQRN
251	SKTNQSKKKG	PRTSPPPPPI	PEDIALGKKY	KEYYKVKDRI	EETRDRGKDR
301	GRDFERQREK	RDKPRSTSPA	GQHHSPISSR	HHSSSSQSGS	SIQRHSPSPR
351	RKRTPSPSYQ	RTLTPPLRRS	ASPYPHSLS	SPQRKQSPPR	HRSPMREKGR
401	HDHERTSQSH	DRRHERRREDT	RGKRDREKDS	REEREYEQDQ	SSSRDHRDDR
451	EPRDGRDRRD	ARDTRDRREL	RDSRDMRDSR	EMRDYSRDTK	ESRDPRDSRS
501	TRDAHDYRDR	EGRDTHRKED	TYPEESRSYG	RNHLREESSR	TEIRNESRNE

Materials and methods

551	SRSEIRNDRM	GRSRGRVPEL	PEKGSRGSRG	SQIDSHSSNS	NYHDSWETRS
601	SYPERDRYPE	RDNRDQARDS	SFERRHGERD	RRDNRERDQR	PSSPIRHQGR
651	NDELERDERR	EERRVDRVDD	RRDERARERD	RERERDRERE	RERERERDRE
701	REKERELERE	RARERERERE	KERDRERDRD	RDHDRERERE	RERDREKERE
751	REREERERER	ERERERERER	ERERERARER	DKERERQRDW	EDKDKGRDDR
801	REKREEIRED	RNPRDGHDER	KSKKRYRNEG	SPSPRQSPKR	RREHSPDSDA
851	YNSGDDKNEK	HRLLSQVVRP	QESRSLSPSH	LTEDRQGRWK	EEDRKPERKE
901	SSRRYEEQEL	KEKVSSVDKQ	REQTEILESS	RMRAQDIIGH	HQSEDRETS
951	RAHDENKKKA	KIQKKPIKKK	KEDDVGIERG	NIETTS	EDGQVFSPKKGQK
1001	KSIEKKRKS	KGSDISDEE	AAQSKKKRG	PRTPPITTK	ELVEMCNGKN
1051	GILEDSQKKE	DTAFSDWSDE	DVPDRTEVTE	AEHTATATTP	GSTPSPLSSL
1101	LPPPPPVATA	TATTVPATLA	ATTAATAATSF	STSAITISTS	ATPTNTTNT
1151	FANEDSHRKC	HRTRVEKVET	PHVTIEDAQH	RKPMQKRSS	SLGSNRSNRS
1201	HTSGRLRSPS	NDSAHRS	GDDQSGRKRVLHS	GSRDREKTKS	LEITGERKSR
1251	IDQLKRGEPS	RSTSSDRQDS	RSHSSRRSSP	ESDRQVHSRS	GSFDSRDRLQ
1301	ERDRYEHDRE	RERERRDTRQ	REWDRDADKD	WPRNRDRDRL	RERERERERD
1351	KRRDLRERE	RLISDSVERD	RDRDRDRTFE	SSQIESVKRC	EAKLEGEHER
1401	DLESTSRDSL	ALDKERMDKD	LGSVQGFEE	NKSERTESLE	GDDESKLDDA
1451	HSLGSGAGEG	YEPISDDELD	EILAGDAEKR	EDQQDEEKMP	DPLDVIDVDW
1501	SGLMPKHPKE	PREPGAALLK	FTPGAVMLRV	GISKKLAGSE	LFAKVKETCQ
1551	RLLEKPKDAD	NLFEHELAL	NMAALLRKEE	RASLLNLGP	CCKALCFRRD
1601	SAIRKQLVKN	EKGTIKQAYT	SAPMVDNELL	RLSLRLFKRK	TTCHAPGHEK
1651	TEDNKLSQSS	IQQELCVS			

RBM15

1	MRTAGRDPVP	RRSPRWRAV	PLCETSAGRR	VTQLRGDDL	RPATMKGKER
51	SPVKAKRSRG	GEDSTSRGER	SKKLGSGGS	NGSSSGKTDS	GGGSRSLHL
101	DKSSRGGSR	EYDTGGGSS	SRLHSYSSPS	TKNSSGGGES	RSSSRGGGGE
151	SRSSGAASSA	PGGGDGAEYK	TLKISELGSQ	LSDEAVEDGL	FHEFKRFGDV
201	SVKISHLSGS	GSGDERVAFV	NFRRPEDARA	AKHARGRLVL	YDRPLKIEAV
251	YVSRRRSRSP	LDKDTYPPSA	SVVGASVGGH	RHPPGGGGGQ	RSLSPGGAAL
301	GYRDYRLQQL	ALGRLPPPPP	PPLPRDLERE	RDYPFYERVR	PAYSLEPRVG
351	AGAGAAPFRE	VDEISPEDDQ	RANRTLFLGN	LDITVTESDL	RRAFDRFGVI
401	TEVDIKRPSR	GQTSTYGFLK	FENLDMSHRA	KLAMSGKIII	RNPIKIGYK
451	ATPTTRLWVG	GLGPWVPLAA	LAREFDRFGT	IRTIDYRKGD	SWAYIQYESL
501	DAAHAAWTHM	RGFPLGGPDR	RLRVDFADTE	HRYQQQYLQP	LPLTHYELVT
551	DAFGHRAPDP	LRGARDRTPP	LLYRDRDRDL	YPDSDWVPPP	PPVRERSTRT
601	AATSVPAYEP	LDSLDRRRDG	WSLDRDRGDR	DLPSSRDQPR	KRRLPEESGG
651	RHLDRSPESD	RPRKRHCAPS	PDRSPELSSS	RDRYNSDNDR	SSRLLLERPS
701	PIRDRRGSLE	KSQGDKRDRK	NSASAERDRK	HRTTAPTEGK	SPLKKEDRS
751	GSAPSTSTAS	SKLKSPSQKQ	DGGTAPVASA	SPKLCLAWQG	MLLLKNSNFP
801	SNMHLLQGD	QVASSLLVEG	STGGKVAQLK	ITQRLRLDQP	KLDEVTRRIK
851	VAGPNGYAIL	LAVPGSSDSR	SSSSAASDT	ATSTQRPLRN	LVSYLKQKQA

901	AGVISLPVGG	NKDKENTGVL	HAFPPCEFSQ	QFLDSPAKAL	AKSEEDYLVM
951	IIVRAKLVEQ	RMKIWNLSK			

VIRMA

1	MAVDSAMELL	FLDTFKHPSA	EQSSHIDVVR	FPCVVYINEV	RVIPPGVRAH
51	SSLPDNRAYG	ETSPHTFQLD	LFNNVSKPS	APVFDRLGSL	EYDENTSIF
101	RPNSKVNTDG	LVLRGWYNCL	TLAIYGSVDR	VISHDRDSPP	PPPPPPPPQ
151	PQPSLKRNP	HADGEKEDQF	NGSPRPQPR	GPRTPPGPPP	PDDDEDDPVP
201	LPVSGDKEED	APHREDYFEP	ISPDRNSVPQ	EGQYSDEGEV	EEEQQEEGEE
251	DEDDVDVEEE	EDEDEDDRRT	VDSIPEEEEE	DEEEEGEED	EGEGDDGYEQ
301	ISSDEDGIAD	LERETFKYPN	FDVEYTAEDL	ASVPPMTYDP	YDRELVPLLY
351	FSCPYKTTFE	IEISRMKDQG	PDKENSGAIE	ASVKLELLD	LYREDRGAKW
401	VTALEEIPSL	IIGLSYLQL	KNTKQDSLQ	LVDWTMQALN	LQVALRQPIA
451	LNVRQLKAGT	KLVSSLAECG	AQGVGTLLQA	GVISGLFELL	FADHVSSSLK
501	LNAFKALDSV	ISMTEGMEAF	LRGRQNEKSG	YQKLELILL	DQTVRVVTAG
551	SAILQKCHF	EVLSEIKRLG	DHLAECTSSL	PNHSEPDHDT	DAGLERTNPE
601	YENEVEASMD	MDLLESSNIS	EGEIERLINL	LEEVFHLMET	APHTMIQQPV
651	KSFPTMARIT	GPPERDDPYP	VLFRYLHSHH	FLELVTLLS	IPVTSAHPGV
701	LQATKDVLF	LAQSQKLLF	FMSEYEATNL	LIRALCHFID	QDEEEGLQSD
751	GVIDDAFALW	LQDSTQTLQC	ITELFSHFQR	CTASEETDHS	DLLGTLHNL
801	LITFNPVGRS	AVGHVFSLEK	NLQSLITLME	YYSKEALGDS	KSKKSVAYNY
851	ACILILVVVQ	SSSDVQMLEQ	HAASLLKCK	ADENNAKLQE	LGKWLEPLKN
901	LRFEINCIPN	LIEYVKQID	NLMTPEGVGL	TTALRVCNV	ACPPPPVEGQ
951	QKDLKWNLAV	IQLFSAEGMD	TFIRVLQKLN	SILTQPWRLH	VNMGTTLHRV
1001	TTISMARCTL	TLLKTMLTEL	LRGGSFEFKD	MRVPSALVTL	HMLLCSIPLS
1051	GRLDSDEQKI	QNDIIDILLT	FTQGVNEKLT	ISEETLANNT	WSLMLKEVLS
1101	SILKVEGFF	SGLILLSL	PLPLPMQTTQ	VIEPHDISVA	LNTRKLWSMH
1151	LHVQAKLLQE	IVRSFSGTTC	QPIQHMLRRI	CVQLCDLASP	TALLIMRTVL
1201	DLIVEDLQST	SEDKEKQYTS	QTTRLLALLD	ALASHKACKL	AILHLINGTI
1251	KGDERYAEIF	QDLLALVRSP	GDSVIRQQCV	EYVTSILQSL	CDQDIALILP
1301	SSSEGSISEL	EQLSNSLPNK	ELMTSICDCL	LATLANSESS	YNCLLTCVRT
1351	MMFLAEHDYG	LFHLKSSLRK	NSSALHSLK	RVVSTFSKDT	GELASSFLEF
1401	MRQILNSDTI	GCCGDDNGLM	EVEGAHTSRT	MSINAAELKQ	LLQSKEESPE
1451	NLFLELEKLV	LEHSKDDN	DSLSDSVVGL	KQMLESSGDP	LPLSDQDVEP
1501	VLSAPESLQN	LFNNRTAYVL	ADVMDLQKLS	MWFTPFQAE	IDTDLDLVKV
1551	DLIELSEKCC	SDFDLHSELE	RSFLSESSP	GRTKTKGFK	LGKHKHETFI
1601	TSSGKSEYIE	PAKRAHVVP	PRGRGRGGFG	QGIRPHDIFR	QRKQNTSRPP
1651	SMHVDDFVAA	ESKEVVPQDG	IPPKRPLKV	SQKISSRGGF	SGNRGGRGAF
1701	HSQNRFFTPP	ASKGNYSRRE	GTRGSSWSAQ	NTPRGNYNES	RGGQSNFNRG
1751	PLPLRPLSS	TGYRPSRDR	ASRGRGGLGP	SWASANSMSG	GSRGKFSVSGG
1801	SGRGRHVSF	TR			

5.5.2 Expression constructs

5.5.2.1 For recombinant baculovirus generation via the flashBAC system

Protein constructs that were individually expressed from the pTriEx vector carry a FLAG-His10-TEV-FLAG tag (5.47 kDa) at the N-terminal end and are shown in Table 16. The protein coding DNA plasmids were obtained from previous work¹²⁹ (METTL3, METTL14), purchased from *Genscript* (all other constructs except from M3MTD) or cloned via site directed mutagenesis (M3MTD). All protein coding genes were codon optimized for *Sf9* cells.

Table 16: List of protein constructs expressed from the pTriEx vector. Protein size and molecular weight are given without purification tags.

Protein construct	Protein size [aa]	Molecular weight
METTL3 constructs		
METTL3	580	64.47 kDa
M3 Δ NTD	320 (260 – 580)	36.65 kDa
M3 Δ ZF1	273 (307 – 580)	31.00 kDa
M3MTD	228 (352 – 580)	25.97 kDa
METTL14 constructs		
METTL14	456	52.15 kDa
M14 Δ RGG	395 (1 – 395)	45.89 kDa
M14 Δ NTD	292 (164 – 456)	32.90 kDa
M14MTD	231 (164 – 395)	26.64 kDa

5.5.2.2 For recombinant baculovirus generation via the Bac-to-Bac system

Protein constructs that were coexpressed from the same vector are shown in Table 17. The individual protein coding sequences were purchased from *Genscript* in a pLIB vector backbone with different purification tags and codon optimized for *Sf9* cells. The DNA plasmids with multiple protein coding sequences used for expression were formed via Gibson assembly. The resulting different protein complexes are given in Table 18.

Table 17: List of protein constructs expressed from the biGBac vector system.¹⁴⁰ Protein size and molecular weight are given without purification tags. Purification tags are given with location (N-terminal (Nterm) or C-terminal (Cterm)).

Protein construct	Protein size [aa]	Molecular weight	Purification tag
METTL3	580	64.47 kDa	FLAG (Nterm)
METTL14	456	52.15 kDa	His8-TEV (Nterm)
FLAG-WTAP	396	44.24 kDa	FLAG (Nterm)
FLAG-shortWTAP	151	17.67 kDa	FLAG (Nterm)
ZC3H13-FLAG	1668	196.64 kDa	TEV-FLAG (Cterm)
RBM15-Strep	969	106.37 kDa	TEV-Strep (Cterm)
VIRMA-Strep	1812	202.02 kDa	TEV-Strep (Cterm)

Table 18: Protein complexes that were coexpressed from the biGBac vector system.¹⁴⁰

Complex name	Coexpressed proteins
A1	METTL3, METTL14
A2	METTL3, METTL14, WTAP
A3	METTL3, METTL14, shortWTAP
AB-11	METTL3, METTL14, ZC3H13, RBM15
AB-21	METTL3, METTL14, WTAP, ZC3H13, RBM15
VIRMA	VIRMA

5.5.3 DNA plasmid preparation

5.5.3.1 Site directed mutagenesis

M3MTD

The protein coding plasmid for M3MTD was derived from the full-length plasmid of METTL3 (pTriEx) by IVA cloning.¹⁵⁰ The full-length plasmid was shortened by whole plasmid PCR reaction with the given primers. Afterwards, the starting plasmid was digested by adding DpnI (10 U, *New England Biolabs*) to the mixture and incubation at 37 °C for 30 min. 1 µL of the DpnI treated mixture was used for transformation into *E. coli* DH5α cells. After amplification and purification of the resulting plasmid, the sequence was verified by Sanger sequencing (*Eurofins Genomics*).

Primer 1: 5' GCCCTGAAAATAAAGATTCTCC 3'

Primer 2: 5' ATCTTTATTTTCAGGGCGCTTTACACAGAGTGTCGG 3'

shortWTAP

The protein coding plasmid for shortWTAP was prepared from the full length WTAP plasmid using the Q5® Site-Directed Mutagenesis Kit (*New England Biolabs*). Primers were designed following the given instructions and the reaction protocol was followed for a reaction volume of 25 µL. After isolation of the newly formed plasmid, the successful shortening of WTAP was proven by Sanger sequencing (*Eurofins Genomics*).

Primer 1: 5' GAGTTAACGGTCGGGGTGAACCTCCAGGCG 3'

Primer 2: 5' GAGCTGCGTAACTCGAGCATGCATCTAGAGCCTG 3'

5.5.3.2 Gibson assembly reaction

Protein coding DNA fragments were prepared by PCR with suitable DNA primers (Table 19) or cut out of a previous vector (for AB-11, AB-21 formation), following the biGBac method described by Weissmann *et al.*¹⁴⁰. This methods describes a toolbox of expression vectors together with a set of suitable primers that were optimized for Gibson assembly. First, matching primers were chosen for each set of gene assemblies that were then inserted via PCR, followed by Gibson assembly into the new vector. Already assembled genes were planned as part of a gene assembly in pBig2ab were cut out of the vector before assembly as whole gene cassette. Gibson assembly reactions were carried out using a Gibson Assembly Master Mix (*New England Biolabs*) and following the corresponding *New England Biolabs* protocol. The formed DNA plasmids with different protein combinations for coexpression are summarized in Table 20. All newly synthesized plasmids were verified by Sanger sequencing (*Eurofins Genomics*).

Table 19: DNA fragments for Gibson assembly reactions generated by PCR with used primers.

Fragment	Protein coding sequence	5' primer ¹⁴⁰	3' primer ¹⁴⁰
M3 $\alpha\beta$	METTL3	CasI_for	CasI_rev
M14 $\beta\omega$	METTL14	CasII_for	Cas ω _rev
M14 $\beta\gamma$	METTL14	CasII_for	CasII_rev
W $\gamma\omega$	WTAP	CasIII_for	Cas ω _rev
shortW $\gamma\omega$	shortWTAP	CasIII_for	Cas ω _rev
Z $\alpha\beta$	ZC3H13	CasI_for	CasII_rev
R $\beta\omega$	RBM15	CasII_for	Cas ω _rev
R $\beta\gamma$	RBM15	CasII_for	CasII_rev
V $\alpha\omega$	VIRMA	CasI_for	Cas ω _rev
V $\gamma\omega$	VIRMA	CasIII_for	Cas ω _rev

Table 20: DNA plasmids for protein expression, formed by Gibson Assembly reactions.

Plasmid name	Protein coding fragments	Vector backbone	Cutting enzyme
A1	M3 $\alpha\beta$, M14 $\beta\omega$	pBig1a	SwaI
A2	M3 $\alpha\beta$, M14 $\beta\gamma$, W $\gamma\omega$	pBig1a	SwaI
A3	M3 $\alpha\beta$, M14 $\beta\gamma$, shortW $\gamma\omega$	pBig1a	SwaI
B1	Z $\alpha\beta$, R $\beta\omega$	pBig1b	SwaI
AB-11	A1, B1	pBig2ab	PmeI
AB-21	A2, B1	pBig2ab	PmeI

5.5.3.3 Chemically competent *E. coli* cells

The *E. coli* cell lines DH5 α , NEB10beta (both *New England Biolabs*) and DH10Bac (*ThermoFisher Scientific*) were amplified and made chemically competent for DNA plasmid transformation (adapted from Untergasser *et al.*¹⁵¹). First, the cells were grown over night (37 °C, 120 rpm) in 3 mL LB medium. The next day, 200 mL fresh LB medium were inoculated with 1 mL of the overnight culture and the cells were grown to a OD₆₀₀ of 0.4 to 0.5. Aliquots of 50 mL were put on ice for 15 min in prechilled tubes, followed by centrifugation (15 min, 2500 g, 4 °C). The supernatant was discarded and each pellet resuspended in 15 mL ice cold transformation buffer. The cell suspension was again centrifuged and the supernatant was discarded. Then, each pellet was resuspended in 4 mL ice cold transformation buffer and thoroughly mixed with 300 μ L DMSO. The competent cells were aliquoted (100 μ L), flash frozen in liquid nitrogen and stored at -80 °C.

5.5.3.4 DNA plasmid transformation in *E. coli* cells

Transformation of plasmid DNA in DH5 α or NEB10beta *E. coli* cells (purchased from *New England Biolabs* or homemade (see 5.5.3.3)) was achieved following the *New England Biolabs* transformation protocol. First, the competent cells were thawed on ice, and > 10 ng plasmid DNA were added and mixed well. The mixture was incubated for 30 min on ice before a heat shock was performed at 42 °C for 30 s. After incubation on ice for another 5 min, 500 μ L SOC medium were added and the cells were incubated at 37 °C and 120 rpm for 1 h. Meanwhile, LB agar plates were poured with the suitable antibiotics.

5.5.3.5 DNA plasmid amplification

Plasmid DNA was amplified following either the QIAGEN Plasmid Mini Kit (*QIAGEN*) or the QIAGEN Plasmid Midi Kit (*QIAGEN*). For miniprep amplifications 5 mL overnight culture in SB medium were used, while 30 mL were used for midiprep amplifications. DNA plasmids to be used in transfection reactions were always amplified in midiprep quality.

5.5.3.6 Sanger sequencing

All DNA plasmids were verified at least once by Sanger sequencing. Therefore, samples with premixed primers were sent to *Eurofins Genomics* to be analyzed. The used primers can be found in Table 21.

Table 21: Sequencing primers.

Primer	Sequence 5' to 3'	Vector
TriEx_down	TCGATCTCAGTGGTATTTGTG	pTriEx
TriEx_up	GTTATTGTGCTGTCTCATCA	pTriEx
pMETTL3_fwd1	GGC AGC ATT GTC TCC AAC CTT C	pTriEx
5' pLIB	TCAACAGGTTGAACTGCTGATC	pLIB, pBig
3' pLIB	GGTGTAGCGTCGTAAGCTAATAC	pLIB, pBig
METTL3_fwd1	CTACCACTGTGGCTGCCTTCGC	pLIB, pBig
METTL3_fwd2	GCACCCGCAAGATCGAGCTGTT	pLIB, pBig
METTL14_fwd1	TGGACGAAGGCGAGACTGACGA	pLIB, pBig
METTL14_rev1	CCAGAGGTGGCTCCAGCAGGAT	pLIB, pBig
RBM15_fwd1	GAGGACTGCTGGACGTGACCCA	pLIB, pBig
RBM15_fwd2	CACTGGAACCTAGGGTCGGAGC	pLIB, pBig
RBM15_fwd3	GAACTGTCATCCAGCCGCGACC	pLIB, pBig
RBM15_rev1	GCCACCTCCGCTGGAGTTCTTG	pLIB, pBig
VIRMA_fwd1	CCCAAAGCACGCTGACGGTGAA	pLIB, pBig
VIRMA_fwd2	GTTTCGCTGACCACGTCTCCAGC	pLIB, pBig
VIRMA_fwd3	ACCTGCAGAGCCTGATCACTCT	pLIB, pBig
VIRMA_fwd4	GATCGTCCGCGAGTTCTCTGGC	pLIB, pBig
VIRMA_fwd5	GACGTCGAACCCGTGCTGTCAG	pLIB, pBig
VIRMA_rev1	TCGTCGCCTTCACCTCCTCGT	pLIB, pBig
WTAP_fwd1	AAGGTGCGCCTGTCCGAAACTG	pLIB, pBig
WTAP_rev1	TCGTTCTGAGCCTGTTCCAGCT	pLIB, pBig
ZC3H13_fwd1	GGACATCGCCCTGGGAAAGAAGT	pLIB, pBig
ZC3H13_fwd2	ACCGCGACCAGGCTAGAGACTC	pLIB, pBig
ZC3H13_fwd3	CCTCTGACAGGGCTCACGACGA	pLIB, pBig
ZC3H13_fwd4	AGCAGACGCTCTTCACCCGAGT	pLIB, pBig

5.6 Protein expression in *Spodoptera frugiperda* 9 (Sf9) cells

5.6.1 Handling of Sf9 cells

All recombinant protein constructs in this project were expressed in *Spodoptera frugiperda* 9 (Sf9) cells. The cells were maintained at 28 °C in Sf-900 III medium (*ThermoFischer Scientific*) supplemented with 0.5 % penicillin/streptomycin (10.000 U/mL, *ThermoFischer Scientific*) and 5 % fetal bovine serum (*ThermoFischer Scientific*). Cells in suspension (everything but small test expressions or transfections) were kept in Erlenmeyer flasks (Corning® Polycarbonate Erlenmeyer Flask with Vent Cap) with 86 rpm shaking. The cell density was kept between $0.5 \cdot 10^6$ and $10 \cdot 10^6$ cells/mL.

5.6.2 Recombinant baculovirus generation via the flashBAC system

Protein coding genes cloned into the pTriEx shuttle vector were cotransfected into Sf9 cells following the baculoCOMPLETE Protein Expression Kit (*Oxford Expression Technologies Ltd*, flashBAC ULTRA™ DNA) to generate a recombinant baculovirus. First, 1×10^6 Sf9 cells per well were seeded on a 6 well plate and let attach for 30 min. Before transfection, the medium was changed to unsupplemented Sf-900 II medium. The transfection complex was prepared by mixing 100 µL TC100 growth medium or Sf-900 II medium with 100 ng flashBAC ULTRA™ DNA, 500 ng protein coding plasmid DNA and 1.2 µL baculoFECTIN II, followed by incubation for 15 min at room temperature. The transfection complex was then added dropwise to the Sf9 cells and after overnight incubation, 1 mL supplemented Sf-900 II medium was added to each well. To generate a seed stock of recombinant baculovirus (P_0), the cells were incubated for another 5 days.

5.6.3 Recombinant baculovirus generation via the Bac-to-Bac system

Recombinant baculoviruses for protein expression from coding genes cloned into pLIB or pBig vectors were generated mainly following the manual of the Bac-to-Bac™ Baculovirus Expression System (*ThermoFisher Scientific*). Bacmid DNA was formed by transformation of protein coding plasmid DNA into DH10Bac *E. coli* (purchased from *ThermoFisher Scientific* or homemade (see 5.5.3.3)) cells. However, the heat shock was performed for 45 s at 42 °C and LB agar plates were poured with 50 µg/mL Kanamycin, 7 µg/mL Gentamicin, 10 µg/mL tetracycline, 100 µg/mL X-gal and 40 µg/mL IPTG. After 24-48h hours incubation at 37 °C, white colonies were identified and

the DNA was isolated following the instructions of the QIAGEN Plasmid Midiprep Kit. To confirm the colour of the picked colonies, they were restreaked on LB agar plates and incubated for 24 h at 37 °C. Before transfection, 1×10^6 Sf9 cells per well were seeded on a 6 well plate and after 30 min the medium was changed to unsupplemented Sf-900 II medium. Afterwards, 250 μ L Opti-MEM™ I Reduced Serum Medium was mixed with 10 μ L ExpiFectamine™ Sf Transfection Reagent and incubated for 5 min at room temperature. To form the final transfection complex, isolated bacmid DNA was added and the mix was incubated for another 5 min at room temperature. The mix was then added dropwise to the Sf9 cells that were incubated for 6 days after transfection to form a seed stock of recombinant baculovirus (P_0).

5.6.4 Virus amplification

From the initial recombinant baculovirus seed stock (P_0), high titre stocks were amplified. To generate the P_1 baculovirus stock, 50 mL of cells in suspension (2.5×10^6 cells/mL) were infected with 200 μ L of P_0 virus, incubated for 7 days (28 °C, 86 rpm) and harvested by separating the virus from the cells (centrifugation for 20 min at 1000 g). Amplification to P_2 baculovirus stock was done analogous, however, 200 mL cells were infected with 500 μ L P_1 virus stock.

5.6.5 Test expressions

For each construct, the P_1 and P_2 virus stock was tested for expression in a small scale. Test expressions were done in adhesion with 2 million log phase Sf9 cells seeded in a 35 mm tissue culture dish and covered with 2 mL supplemented Sf-900 III medium. After seeding the cells, 200 μ L recombinant baculovirus were added to the medium followed by incubation for 3 days. For constructs that did not show expression under these conditions, 500 μ L P_2 virus were tested in addition. The cells were harvested, separated from the medium (2500 rpm, 5 min) and resuspended in 250 μ L PBS buffer supplemented with protease inhibitors (EDTA free protease inhibitor cocktail, *cOmplete™*). Afterwards, the cells were lysed briefly (20 s ultrasonicator, 1 min on ice, 3 cycles) and 10 μ L of the cell lysate were used for analysis with a western blot (5.2.3.7).

5.6.6 Large scale expressions

Protein expressions for protein purification were performed with cells in suspension and larger volumes (250 mL to 1 L). The cell suspension was adjusted to a concentration between $5 \cdot 10^6$ and $6 \cdot 10^6$ cells/mL in supplemented Sf-900 III medium. Protein expression was induced by infection of the cells with 3 % P₂ baculovirus of each protein construct to be (co-)expressed. After 3 days of incubation, the cells were harvested (5000 g, 15 min), resuspended in PBS buffer and pelleted (5000 g, 10 min) into aliquots equal of 250 to 500 mL expression volume. The supernatant was discarded and the cell pellets were flash frozen in liquid nitrogen and stored at -80 °C for later purification.

5.7 Protein purification

After expression, the cell pellet was first resuspended in ice cold HisTrap Buffer A (at least 5 mL buffer per 1 g cell pellet) supplemented with protease inhibitors (EDTA free protease inhibitor cocktail, *cOmplete™*). Then, the cells were lysed with a microfluidizer (M110P, *Microfluidics*) or for small volumes by using ultrasound (30 s ultrasonic, 1 min pause, 6 cycles). In both cases, the cell suspension was always cooled on ice to keep a temperature around 4 °C. After lysis, Turbo DNase (*ThermoFisher*, 10 µL per 40 mL lysate) was added and the cell lysate was centrifuged (30000 g, 45 min, 4 °C) to separate the cell debris. The supernatant was filtered through a syringe filter (pore size 0.45 microns, *Corning*) and then used for further protein purification.

5.7.1 Ni-NTA affinity purification

The filtered cell lysate was loaded onto an equilibrated HisTrap™ FF column (*GE healthcare life sciences / Cytiva*). Depending on the volume of the cell lysate and the expression levels of the protein, a column volume of either 1 mL (up to 60 mL cell lysate or low expression levels) or 5 mL was chosen. The unbound proteins were washed out with HisTrap buffer A until a stable UV absorbance at 280 nm was reached. Afterwards, the protein of interest was eluted from the column applying an imidazole gradient up to a concentration of 500 mM (HisTrap buffer B) over 20 to 30 column volumes (cv). The eluate was collected in fractions of 1 cv and fractions with reasonable absorbance at 280 nm were analyzed via SDS PAGE. After analysis, fractions

containing the protein of interest were combined, concentrated (Vivaspin 2 or 20, MWCO approximately 1/3 of the complex size, *Sartorius*) to 500 μ L and then subject to size exclusion chromatography for further purification. During the concentration step, the protein solution was mixed at least every 15 min to prevent protein aggregation.

5.7.2 Size exclusion chromatography

To separate the assembled complex from monomeric proteins, a size exclusion chromatography was performed. A Superdex 200 (increase) 10/300 GL column (*GE healthcare life sciences / Cytiva*) was equilibrated in SEC buffer and the concentrated protein fractions from the Ni-NTA affinity purification were applied. The proteins were then separated by size with a flowrate of 0.5 mL/min and the absorption at 280 nm was monitored. Fractions (size 1 mL) of interest were analyzed via SDS PAGE and complex containing fractions were combined and concentrated (Vivaspin 2 or 20, MWCO approximately 1/3 of the complex size, *Sartorius*) to the desired volume (50 to 500 μ L) and concentration (at least 3 μ M). The concentrated protein complex was aliquoted, flash frozen in liquid nitrogen and stored at -80 °C. After experiments remaining protein solution from a thawed aliquot was discarded and not refrozen for later experiments.

5.8 Methyltransferase activity determination

5.8.1 Site-specific ³²P-labeling of substrate RNAs

RNA substrates for methyltransferase activity determination were site-specifically labeled with radioactive ³²P via 5'-phosphorylation followed by DNA splinted ligation. RNA constructs were purchased divided in 2 fragments with the 3' fragment carrying the target adenine at its 5' end.

5.8.1.1 5'-phosphorylation

Site-specific labelling was achieved by 5'-phosphorylation of the RNA 3' fragment with γ -³²P-ATP, after which only the target adenine carries a radioactive 5' phosphate. For this, the 3' fragment of the RNA (10 μ M) was mixed together with γ -³²P-ATP (2 μ M, 2.2 MBq/ μ L) and 0.4 U/ μ L T4 polynucleotide kinase (*New England Biolabs*) in T4 polynucleotide kinase reaction buffer (*New England Biolabs*). The reaction was incubated for 30 min at 37 °C before unlabeled

ATP in T4 polynucleotide kinase reaction buffer (*New England Biolabs*) was added to a final concentration of 10 mM. To finalize the phosphorylation, the reaction mixture was incubated for another 20 min at 37 °C. Then, EDTA was added to a final concentration of 50 mM and the enzyme was heat inactivated (10 min, 75 °C). To separate the RNA fragment from excess ATP and γ -³²P-ATP, the sample was purified over a MicroSpin G-25 column (*GE healthcare life sciences*) before DNA splint ligation.

5.8.1.2 DNA splinted ligation

The phosphorylated 3' RNA fragment was ligated to the 5' fragment to form the full-length RNA substrate via DNA splinted ligation. DNA splint, 5' RNA fragment and phosphorylated 3' RNA fragment were mixed in equal amounts (2.5 μ M each) in 0.5x T4 DNA ligase reaction buffer (*New England Biolabs*) and heated to 85 °C for 4 min. To correctly anneal the corresponding strands, the mixture was slowly cooled to room temperature again. Then, the T4 DNA ligase reaction buffer was adjusted (1x concentration) and 0.25 U/ μ L T4 DNA ligase (*New England Biolabs*) were added. The ligation was performed depending on the construct for 2 h at 37 °C or 25 °C and if required continued over night at 16 °C. After heat inactivation of the T4 ligase (10 min, 75 °C), 0.1 U/ μ L Turbo DNase (*ThermoFisher*) were added to digest the DNA splint. The digestion was incubated at 37 °C for 1 h and the Turbo DNase was then heat inactivated for 10 min at 75 °C. The remaining RNA was precipitated with ethanol and further purified via preparative PAGE. After separation, the gel was covered with plastic foil, exposed on a Storage Phosphor Screen (*GE healthcare life sciences*) which was then visualized with a Typhoon 9400 (*GE Healthcare life sciences*) to identify the desired RNA bands. The concentration of the purified RNA samples was determined via radioactivity incorporation.

The 5' and 3' fragments for all ligated RNAs are given in Table 22. As DNA splint, the corresponding transcription template (see Table 13) for each RNA was used.

Table 22: RNA sequences of all fragments used for DNA splinted ligation.

Final RNA	5' fragment (5' to 3')	3' fragment (5' to 3')
U1	GGA UGC ACA AGG	ACU GGA UAC GAG G
U1.1	GGA UGC ACA AAA	ACU GGA UAC GAG G
U1.2	GGA UGC ACA AAG	ACU GGA UAC GAG G
U1.3	GGA UGC ACA AGA	ACU GGA UAC GAG G
D1	(G)GCG GCU GCG AGG G	AC UUU GUC AUC CGU CC

Final RNA	5' fragment (5' to 3')	3' fragment (5' to 3')
D2	(GG)UGC CAU CUG GGG	ACU GCC UUU ACC AAC
D3	GGA UGC ACA AGG	ACU GGA UAC GAG G
H1	(GG)CAC CCC GCA GGG	ACU CGU UUU GCU UCU CUC UCU GC
H2	GGA UGC ACA AGG	ACU GAG UGC AUC
H3	(GG)CUU CAA GGA AGG ACG CGA	ACC CUU CUC UGA CC
H4	(GG)UUA GAG AUA GCU CGC AUU CAG	ACU ACC UAC UAA CAA UAU
H4.1	(GG)UUA GAG AUA GCU CGC AUU CGG	ACU ACC UAC UAA CAA UAU
H5	(G)GUG UCA AAU CCA AGC UUA A	ACA AUU UGA CAG UGU GG
FU	GGG AGA CCA CCA AGC CGA GCG A AGG	AC UAA GCG AGC CGA ACC ACC

5.8.2 Methyltransferase activity assay

After ³²P-labeling of the substrate RNAs, the enzymatic activity of the different complex compositions was tested. 50 nM of the substrate RNA were mixed with differing concentrations of SAM and protein complex in the activity assay buffer (see chapter 5.1.2). For experiments with full-length METTL3/METTL14 on the structurally different RNA substrates 1 μM protein complex and 200 μM SAM were used (results from chapter 3.1.4). The experiments with the different protein complexes were performed with a protein concentration of 250 nM and 100 μM SAM (results from chapter 3.2). For experiments on the comparison between the complexes with and without WTAP 250 nM protein complex and 64 μM SAM were used. Activity assays with the FRET RNA construct were performed with 200 nM RNA, 1 μM protein and 50 μM SAM.

In all experiments the reaction mixtures were prepared on ice just before the reaction. The reaction was performed at 37 °C except for experiments on the FRET RNA where the reaction took place also at room temperature (approx. 18 °C). The reaction mixture for most assays was incubated for 2.5 h. For the experiments with the METTL3/METTL14 complex and the METTL3/METTL14/WTAP complex on differently structured RNAs, incubation at 37 °C was performed for 1 h. The reaction was stopped by heat denaturation at 75 °C for 10 min followed by ethanol precipitation of the modified RNA in all experiments.

5.8.3 Modification analysis

5.8.3.1 P1 nuclease digestion

Samples from methyltransferase activity determination experiments were first digested to NMPs. The precipitated RNA was dissolved in 30 μ L P1 buffer and 0.2 U P1 endonuclease were added to digest the RNA sample. The digestion was allowed to proceed for 1 h at 55 °C. For negative control, unmodified, 32 P-labeled RNA was digested. Analysis of the resulting NMPs was done with thin layer chromatography.

5.8.3.2 Thin layer chromatography

NMPs from nucleic acid digestion after methyltransferase activity determination experiments were analyzed by thin layer chromatography (TLC). Accordingly, about 1.7 Bq of the radioactive sample were added dropwise onto a 10 cm long TLC cellulose plate (Merck) each and let dry. To separate N⁶-methyladenosine monophosphates (m⁶AMPs) from adenosine monophosphates (AMPs), the separation was performed with TLC buffer C for 5 h. Afterwards the cellulose plates were dried on air, stored in plastic bags and exposed to a Storage Phosphor Screen (*GE healthcare life sciences*) which was then visualized with a Typhoon 9400 (*GE Healthcare life sciences*). The resulting image was analyzed with ImageJ and from that, methyltransferase activity was calculated.

5.8.3.3 Calculation of methyltransferase activity

The methyltransferase activity was calculated from the TLC analysis of the RNA samples. Hence, absolute methylation yields were calculated comparing spot intensities for AMPs (I_A) and m⁶AMPs (I_{m^6A}) following equation (12). Spot intensities were determined with ImageJ.

$$\text{methylation yield} = \frac{I_{m^6A}}{I_A + I_{m^6A}} \quad (12)$$

The methylation yield was calculated from at least three independent replicates and standard mean deviations were determined.

5.9 Microscale thermophoresis

Microscale thermophoresis (MST) was used to investigate the binding affinity of the protein complex to its substrate RNA. MST is based on the diffusion behavior of fluorescently labeled molecules along a temperature gradient. The diffusion behavior is hereby dependent on the hydration shell and with that size, charge, and conformation of the molecule.^{152,153} In this case, the RNA is fluorescently labeled and upon protein binding, the diffusion behavior is changed.

As starting material for the measurement samples, a 100 nM sample of Cy5-labeled RNA U1 and a 1:1 dilution series of the METTL3/METTL14 complex alone (72 μ M, 16 samples) or together with SAM or SAH (1,6 mM) were prepared. Then, each protein sample was mixed 1:1 with RNA, mixed carefully and transferred to hydrophobic capillaries (*NanoTemper Technologies*). The measurements were carried out on a Monolith NT.115 (*NanoTemper Technologies*), at 24 °C and an excitation and MST power of 40 %. After the initial capillary scan, all samples were measured four times. The obtained sample data was analyzed with the software MO.Affinity Analysis v2.1.2 with Hill model fitting.

5.10 smFRET spectroscopy

5.10.1 Materials for construction of smFRET microscope setup

The detailed construction process is described in chapter 4.1 and detailed drawings of custom-made components can be found in chapter 6.6.

5.10.1.1 Supporting structure

- Optical Breadboard (1500 mm x 900 mm x 110 mm, PBI52516, Thorlabs)
- Passive Isolation Mounts (PWA074, Thorlabs)
- Passive Isolation Breadboard Support Frame (PFP90120, Thorlabs)

5.10.1.2 Basic components from Thorlabs

- \varnothing 12.7 mm standard post holders in various length (30 to 150 mm)
- \varnothing 12.7 mm stainless steel optical posts in various length (30 to 200 mm)
- Standard mounting bases in different shapes
- Right-angle brackets with counterbored slots

- Table clamps
- Solid Aluminum Breadboards 12.7 mm thick, different sizes
- Kinematic mirror mount (\varnothing 1.0")
- Kinematic mount for 1" tall rectangular optics
- Rotation mount for \varnothing 1" optics
- Lens mount with retaining ring
- Alignment plate
- Standard Iris, \varnothing 15 mm max aperture
- Black hardboard and tape

5.10.1.3 Excitation

- 637 nm laser (OBIS LX, 140 mW, Coherent)
- 532 nm laser (gem 532, 500 mW, Laser Quantum)
- 488 nm laser (LBX-488, 110 mW, OXXIUS)
- Optical beam shutter, \varnothing 1/2" aperture with controller (Thorlabs)
- Six station single filter wheel with 5 absorptive neutral density filters (Thorlabs)
- \varnothing 1" half-wave plate mounted in RSP1X15 indexing rotation mount (Thorlabs)
- 10 mm polarizing beam splitter cube, 420 – 680 nm, with holder (Thorlabs)
- Beam block (LB1, Thorlabs)
- Laser-Beamsplitter HC BS R532 (Chroma)
- Laser Beamsplitter H 488 LPXR superflat Vers. 2 (Chroma)
- Mounting post (\varnothing 1.5", 350 mm) with post base, post clamp and post platform (Thorlabs)
- Digital power and energy console, slim Si sensor, 400 – 1100 nm, 500 pW – 500 mW (Thorlabs)

5.10.1.4 Microscope

- Inverted microscope Olympus IX73 (Olympus)
- Objective: UPLSAPO 60XW (water immersion, NA 1.2, Olympus)
- 515/588/700 HC tripleband filter (Semrock)
- 532 LP edge basic longpass filter (Semrock)
- UVFS Pellin Broca prism, size 11 x 20 x 6.4 mm (EKSMA Optics)
- Lens f = 75 mm, \varnothing 1" achromatic doublet, ARC: 400 – 700 nm (Thorlabs)
- XYZ Linear Stage (M-460A-XYZ, Newport Corporation)

- custom-made aluminum breadboard, prism bridge and prism holder (detailed drawings in chapter 6.6)

5.10.1.5 Detection

- Camera: Photometrics Evolve 512 delta with CCD97 sensor (Photometrics)
- C-Mount extension tube, 50 mm long (Thorlabs)
- Laser beam splitter H 643 LPXR superflat, t=2 mm (Chroma)
- Adjustable mechanical slit, max slit width 0.24" (Thorlabs)
- Lens f = 100 mm, Ø1" achromatic doublet, ARC: 400 – 700 nm (Thorlabs)
- Lens f = 150 mm, Ø1" achromatic doublet, ARC: 400 – 700 nm (Thorlabs)
- 12.7 mm translation stage with standard micrometer (Thorlabs)
- XYZ translation stage (Thorlabs)

5.10.1.6 Remote control software

- Custom made control software was programmed in NI LabVIEW 2021 (64 bit) with Photometrics TPM Plugin (version 1.3.2.1)
- Stand-alone camera control: Micro-Manger 1.4.14^{154,155} with Photometrics PVCAM device adapter
- Stand-alone laser control: software included with individual lasers
- Stand-alone shutter control: Kinesis software (Thorlabs)

5.10.2 Sample slide preparation

Sample slides for FRET measurements were prepared to immobilize the sample on the glass surface. Quartz slides (Micro-Tec quartz microscope slide 76.2 x 25.4 x 1 mm, fused quartz, *Micro to Nano*) were thoroughly cleaned employing the following protocol. For this, the slides were washed by hand-scrubbing with Alconox and then sonicated in 10 % (w/v) Alconox in ddH₂O in a slide holder for 20 min. After rinsing with ddH₂O, the slides were sonicated for 5 min in ddH₂O, 15 min in acetone and 20 min in 1 M KOH. Then, the slides were rinsed with excessive amounts of ddH₂O and flamed with a propane torch.

To passivate the glass surface against unspecific protein binding, the surface was pegylated. A slide holder was cleaned (sonicated with methanol inside for 5 min) and filled with a solution of 5 % acetic acid and 1 % (3-Aminopropyl)triethoxysilane (*Merck*) in methanol. The clean quartz slides were inserted into the filled slide holder, sonicated for 1 min and incubated for another 20 min. After rinsing with ddH₂O, the slides were dried with nitrogen. A fresh PEG solution of

250 $\mu\text{g}/\mu\text{L}$ mPEG-NHS (MW 5000, *Nanocs*) and 10 $\mu\text{g}/\mu\text{L}$ Biotin-PEG-NHS (MW 5000, *Nanocs*) in 0.1 M NaHCO_3 (fresh, filtered) was prepared for the next step. The solution was mixed by vortexing and centrifuged for 1 min at 17000 g. 150 μL of the PEG solution were poured on each slide (on one side) and covered with a coverslip, avoiding the formation of air bubbles. The slides were stored in a dark moisture chamber overnight, cleaned with water the next day, dried with nitrogen and stored under argon gas at -80°C .

Shortly before FRET measurements (max. 1 week), the needed amount of pegylated quartz slides were prepared for measurement by creation of small channels on the slide (ensuring use of the pegylated side). The channels (approx. 2-3 mm wide) were separated by adhesive tape (double sided, *Scotch*) and covered with a coverslip.

The final slide preparations were done immediately before loading the sample onto the slide. For this, the channels were flushed with a 0.2 mg/mL streptavidin solution (*Thermo Fisher Scientific*) in T50-buffer and incubated for 5 min. Afterwards, the channels were rinsed with 45 μL T50 buffer and equilibrated with 45 μL sample buffer.

5.10.3 Sample preparation

RNA samples for smFRET measurements were annealed prior to each measurement. A solution of 1.05 μM RNA, 700 nM DNA FHA and 1.30 μM DNA FHD in annealing buffer (50 mM Tris, 10 mM MgCl_2 , pH 8.0) was heated to 85°C for 3 min, first cooled to room temperature and then stored on ice. The solution was diluted to a final DNA FHA concentration of 140 pM and loaded onto the measurement channel on the sample slide until a good molecule density was reached (typically 45 μL for approx. 300 molecules per picture). Unbound molecules were washed out with 45 μL imaging buffer. Directly before starting the measurement, the measurement channel was flushed with 45 μL imaging buffer supplemented with an oxygen scavenging system (5 mM protocatechuic acid, 50 nM protocatechuate-3,4-dioxygenase)¹¹⁴ and varying concentrations of protein (0 or 2 μM) and cofactor (SAM/SAH) (0 or 100 μM).

5.10.4 smFRET data acquisition

SmFRET data was acquired on the self-constructed microscope setup (see chapter 4.1) with remote control using the programmed software (see chapter 4.1.5). The measurements for histogram data were performed with green laser excitation (15 mW) for 25 frames, followed by automatic switching to red laser excitation (10 mW) for 25 frames with 100 ms exposure time and EM gain of 200. For each individual condition, measurements of 30 different image sections were taken (approximately 200 to 500 molecules per picture). Time traces were acquired on 600 frame long movies (green laser power 10 mW, 100 ms exposure time, EM gain 200). The different sample compositions were independently measured at least twice. On each measurement day, a 20 frame image stack of a TetraSpeck Microspheres (0.1 μm , blue/green/orange/dark red, *Thermo Fisher Scientific*) sample was acquired to be able to map the green onto the red channel during analysis.

5.10.5 smFRET data analysis

The acquired smFRET image data (saved as tif stacks) were first converted with the software Andor SOLIS (version 4.21) to raw format (16 bit integer) to be able to process the data with the given analysis software. The image analysis step was performed with IDL (student edition 6.2, *Exelis*) using IDL scripts for analysis written by Hazen Babcock and Greg Bokinsky¹⁵⁶ and provided by Dr. Martin Hengesbach. First, the TetraSpeck images were used to create a transformation map (script 'beads_co2' and 'beads_map', at least 100 found pairs), aligning the acceptor to the donor channel. Based on this transformation map, the sample image data was aligned and analyzed by gaussian point spread function fitting to localize the fluorophores and calculate the background corrected fluorescence intensity (script 'dat_all2', 'ffp_dat2', 'ap_dat2' and 'ave_tr_dat'). The generated *.traces file includes information on donor (I_D) and acceptor (I_A) fluorescence intensity over time (one value each per frame) for every detected FRET pair of the image stack.

Analysis of the extracted fluorescence intensities was completed in Matlab (Version R2021a, MathWorks) with scripts provided by Dr. Martin Hengesbach and edited by Christin Fuks and myself.

5.10.5.1 Histogram data

For regular smFRET histogram data, donor leakage into the acceptor channel was corrected ($I_{AC} = I_A - x * I_D$) by varying factor x until the center of the donor only peak results in E_{FRET} of 0 (with `red_threshold` set to 0). Moreover, the data was filtered for intact FRET pairs, eliminating FRET pairs below the set overall intensity threshold ($I_A + I_D > 2000$) and with an averaged acceptor intensity over the last 20 frames (acceptor excited data) below the set `red_threshold` (typically 2000). This yields FRET values for molecules where donor and acceptor fluorophores are present and intact over the whole 50 frames. FRET values were calculated ($E = I_{AC} / (I_{AC} + I_D)$) and averaged over the first 20 frames (see chapter 6.7.1). The calculated FRET efficiencies were combined into a histogram with binning into intervals of 0.025. As quality control, only histograms with at least 1000 molecules (after filtering) were assumed as high-quality data. The binned data was exported to Origin and normalized before plotting. Histogram data plots were created in Origin and (if reasonable) fitted with a gaussian function (equation (13)) for one to up to three peaks.

$$y = \frac{A}{w \sqrt{\frac{\pi}{2}}} e^{-2 \frac{(x-x_c)^2}{w^2}} \quad (13)$$

A = peak area

w = full width at half maximum

x_c = peak center

5.10.5.2 Movie analysis with HaMMMy

Time traces of the longer movies were first sorted by eye with the help of a Matlab script (see chapter 6.7.2), accepting only traces that show a typical FRET behavior. Only traces with visible bleaching of at least the acceptor fluorophore were chosen for further analysis. Afterwards, the traces were exported in a HaMMMy readable format where the traces were aborted 5 frames before reaching a FRET efficiency of 0.1 for the first time (see chapter 6.7.3). Additionally, a file containing all molecule traces in a 'stitched together' format was exported for HaMMMy and the average trace length was calculated. In the next step, the traces for all individual molecules and the stitched version were analyzed with hidden Markov modelling (HMM) by HaMMMy¹⁴⁸ and the reports visualized with a transition density plot (TDP). For RNA only data, 3 different FRET states were used for HaMMMy state fitting whereas 4 states were assumed for all protein related data.

The obtained HaMMY reports on fitted states and dwell times were further analyzed with Matlab scripts written by myself.

Analysis of fitted states

First, the states fitted by HaMMY were grouped to distinct FRET states and analyzed for detected transitions (see chapter 6.7.4). The group edges were hereby defined based on the initial TDP results. The output file contains a list of all transitions between relevant states with information on starting and ending of HaMMY fitted state, grouped state, number of transitions and molecule name where the transition was detected. Moreover, the portion of transitioning molecules was calculated and a list for the total number of transitions for each molecule was made. For the data analyzed here, the data was grouped as shown in Table 23.

Table 23: Conditions for FRET state grouping of HaMMY fitted FRET states.

Group name	Group range	Relevant group
-2	[-2;0.3[no
0.30	[0.3;0.44[yes
0.44	[0.44;0.51[yes
0.51	[0.51;0.60[yes
0.60	[0.60;0.80[yes
0.80	[0.80;2[no

From the output file of the summarized relevant transitions, a list of relevant values was calculated in Excel for each transition in every analyzed measurement condition:

1. Average HaMMY fitted state for starting and ending state
2. Total number observed for this transition
3. Number of molecules showing this transition
4. Average number this transition was observed per molecule
5. Percentage of point 2 in relation to total number of all observed transitions
6. Percentage of point 3 in relation to the total number of analyzed molecules

Based on these indication values, a transition was defined as sufficiently relevant for this condition if point 6 was calculated to be higher than 5 % for the given transition.

Analysis of dwell times

The analysis of HaMMY calculated dwell times was done in a similar way (see chapter 6.7.5). First, the corresponding HaMMY fitted states were grouped (according to Table 23) and irrelevant transitions were filtered out. The remaining transitions were binned for their dwell times with different bin sizes and binning ranges depending on the analyzed data (Table 24). The binned dwell times were then plotted as histogram in Origin and the resulting plot was fitted with a single exponential decay function (equation (14)) (based on Mundigala *et al*)¹⁵⁷. The binning of the data was chosen to yield the R² factor of the fit as close to 1 as possible. The transition rate (k) was then calculated with 1/t.

$$y = y_0 + Ae^{-\frac{x}{t}} \quad (14)$$

y_0 = offset; A = amplitude; t = time constant

Moreover, the transition rate was also calculated from the average dwell time (1/average dwell time) per transition for comparison purposes.

Table 24: Binning range and bins size for all analyzed measurement conditions. Transitions are indicated with the group names.

Analyzed data	Transition	Binning range	Bin size
METTL3 / METTL14 with SAM on substrate RNA	0.30 to 0.44	0.1 s to 60 s	0.1 s
	0.30 to 0.51	0.1 s to 60 s	0.1 s
	0.44 to 0.30	0.2 s to 60 s	0.2 s
	0.44 to 0.51	0.1 s to 60 s	0.1 s
	0.51 to 0.30	0.2 s to 60 s	0.2 s
	0.51 to 0.44	0.2 s to 60 s	0.2 s
	0.51 to 0.60	0.1 s to 60 s	0.2 s
METTL3 / METTL14 with SAH on product RNA	0.30 to 0.44	0.1 s to 60 s	0.1 s
	0.30 to 0.51	0.1 s to 60 s	0.1 s
	0.44 to 0.30	0.2 s to 60 s	0.2 s
	0.44 to 0.51	0.1 s to 60 s	0.1 s
	0.51 to 0.44	0.2 s to 60 s	0.2 s

Analyzed data	Transition	Binning range	Bin size
METTL3 / METTL14 / WTAP with SAM on substrate RNA	0.44 to 0.51	0.1 s to 60 s	0.1 s
	0.44 to 0.60	0.1 s to 60 s	0.2 s
	0.51 to 0.44	0.2 s to 60 s	0.2 s
METTL3 / METTL14 / WTAP with SAH on product RNA	0.30 to 0.51	0.1 s to 60 s	0.1 s
	0.44 to 0.51	0.1 s to 60 s	0.1 s
	0.44 to 0.60	0.1 s to 60 s	0.1 s
	0.51 to 0.44	0.2 s to 60 s	0.3 s
	0.51 to 0.60	0.1 s to 60 s	0.1 s
METTL3 / METTL14 with SAM on abasic RNA	0.44 to 0.51	0.1 s to 60 s	0.1 s
	0.51 to 0.44	0.2 s to 60 s	0.2 s
METTL3 / METTL14 with SAH on abasic RNA	0.44 to 0.51	0.1 s to 60 s	0.1 s
METTL3 / METTL14 / WTAP with SAM on abasic RNA	0.30 to 0.44	0.1 s to 60 s	0.1 s
	0.30 to 0.51	0.1 s to 60 s	0.1 s
	0.44 to 0.51	0.1 s to 60 s	0.1 s
	0.44 to 0.60	0.1 s to 60 s	0.1 s
	0.51 to 0.30	0.2 s to 60 s	0.3 s
	0.51 to 0.44	0.2 s to 60 s	0.2 s
METTL3 / METTL14 / WTAP with SAH on abasic RNA	0.44 to 0.51	0.1 s to 60 s	0.1 s
	0.44 to 0.60	0.1 s to 60 s	0.1 s
	0.51 to 0.60	0.1 s to 60 s	0.1 s
	0.60 to 0.44	0.2 s to 60 s	0.3 s
	0.60 to 0.51	0.2 s to 60 s	0.2 s

Analysis of populations

The population of the FRET states was analyzed in two ways (see chapter 6.7.6). To begin with, the FRET efficiency distribution of all dynamic and static molecules was calculated separately. For this, the FRET efficiency traces calculated by HaMMY (not fitted) were used to build the average of the first 20 data points of each molecule trace. The resulting distribution was then plotted into separate histograms for dynamic and static molecules to compare the FRET efficiency distribution of the analyzed molecules from the movies to the histogram data. To better visualize differences in the distribution of dynamic vs static molecules, both distributions were normalized individually and the difference (dynamic – static) was plotted. Moreover, the overall population distribution of the dynamic molecules over the full length of the traces was analyzed. For this, the HaMMY fitted molecule traces were divided into single frames and grouped into the FRET state groups as described in Table 23. This yields a population distribution of the individual FRET states with single frame resolution. For better comparability between the individual conditions, the population distribution was normalized to the total frame number.

6.1.3 WTAP

ATGGACTACAAAGATGACGACGACAAGACCAACGAGGAACCTCTGCCAAGAAGGTGCGCCTGTCCGAAACTGACTTC
 AAGGTCATGGCTCGCGACGAGCTGATCCTGCGTTGGAAGCAGTACGAGGCTTACGTGCAGGCCCTGGAGGGCAAGTAC
 ACCGACCTGAACTCTAACGACGTCCTGACCTGGCCCTGAGGGAATCAGAGGAAAAGCTGAAGCAGCAGCAGCAGGAAAGCGCT
 CGCCGTGAGAACATCCTGGTCATGAGACTGGCCACCAAGGAGCAGGAAATGCAGGAGTGCACCCTCAGATCCAGTAC
 CTGAAGCAGGTGCAGCAGCCTTCCGTGCTCAGCTGCGCTCCACTATGGTGGACCCTGCCATCAACCTGTTCTTCCCTG
 AAGATGAAGGGAGAGCTGGAACAGACTAAGGACAAGCTGGAACAGGCTCAGAACGAGCTGTCCGCCTGGAAGTTCACC
 CCCGACAGCCAGACTGGCAAGAAGCTGATGGCTAAGTGCCTGATGCTGATCCAGGAGAACCAGGAACTGGGTAGGCAG
 CTGTCTCAGGGCAGAATCGCCAGCTGGAGGCTGAACTGGCCCTGCAGAAGAAGTACTCAGAGGAACTGAAGTCCAGC
 CAGGACGAACCTGAACGACTTCATCATCCAGCTGGACGAGGAAGTGGAGGGCATGCAGTCCACCATCCTGGTCTGCAG
 CAGCAGCTGAAGGAGACTAGGCAGCAGCTGGCCAGTACCAGCAGCAGCAGTACAGGCTTCCGCTCCCTCAACCTCC
 CGCACCCTGCTTCCGAGCCAGTGAACAGAGCAGGCCACCTTAAGGACTGCTCACGTCTGACTAACCGGACCTAGC
 AACGGTTCTTCTATCCAGGCAGAGAACTAGCGGATCTGGTTTCCACCCTGAAGGAAACACCCTGAGGACGACTTCCCA
 AGCTCTCCTGGCAACGGAAACAAGTCATCCAACAGCTCTGAGGAACGCACCGGCCCTGGTGGCTCTGGATACGTGAAC
 CAGCTGTGAGCTGGCTACGAAAGCGTGGACTCTCCAAGTGGATCAGAGAACTCCCTGACCCACCAGTCTAACGACACT
 GACTCATCCCAGACCCTCAGGAAGAGAAGGCCGTGTCCGGCAAGGGCAACAGGACCCTGGGAAGCAGACACGTCCAG
 AACGGTCTGGACAGCTCTGTGAACGTCCAGGGTTCAGTCTGTAA

6.1.4 RBM15

ATGAGGACTGCTGGACGTGACCCAGTGCCTCGCCGTTCCCCACGCTGGAGGAGAGCCGTCCCTCTGTGCGAGACTTCT
 GCTGGTTCGCGGTGTGACTCAACTCCGCGGCGACGACTGAGGAGACTGCTACTATGAAGGGCAAGGAACGTTCCCCC
 GTCAAGGCTAAGCGCAGCCGTGGTGGCGAGGACTCAACCTCCCGCGGAGAACGTTCCAAGAAGCTGGGAGGTTCTGGC
 GGATCAAACGGTTCAGCTCTGGCAAGACAGACTCCGGTGGTGGTTCGCGCCGCTCCCTGCACCTGGACAAGTCTCTCC
 TCCCGCGGTGGTTCGCGGAGTACGACACTGGAGGTGGCTCTTCAATCCAGCCGCTGCACCTCTACTCTTACCAAGC
 ACCAAGAACTCCAGCGGAGGTGGCGAGTCTCGCTCTTCAATCCCGTGGAGGTGGCGGAGAATCACGTAGCTCTGGTGT
 GCCTCATCCGCTCCTGGTGGCGGAGACGGAGCTGAGTACAAGACCCTGAAGATCAGCGAACTGGGTTCTCAGCTGTCA
 GACGAGGCTGTGGAAGACGGCCTGTCCACGAGTCAAGCGTTTCGGAGACGTGTCCGTCAAGATCAGCCACCTGAGC
 GGATCTGGTTCAGCGCAGAGAGGGTGGCTTTCGTCAACTTCAGGAGACCAGAAGACGCCAGAGCTGCCAAGCAGCT
 AGGGGAAGACTGGTCTGTACGACAGGCCCTGAAGATCGAAGCCGTGTACGTGACCCGCGCTAGGTCCAGAAGCCCA
 CTGGACAAGGACACTTACCCTCCCTCTGCCTCAGTGGTGGTGTCTGTGGGTGGCCACAGGCACCCACCTGGAGGT
 GCGGAGGTCAAAGATCCCTGTCTCCTGGCGGAGCTGCTCTGGGTTACAGGGACTACAGACTGCAGCAGCTGGCTCTG
 GGAAGGCTGCCCCACCTCCCCACCTCCCCGTCACGCTGACCTGGAGCGCAACGTGACTACCCTTCTACGAGAGG
 GTGAGACCAGCTACTCACTGGAACCTAGGGTTCGAGCTGGAGCTGGAGCTGCTCCATTGAGAGAGGTGGACGAAATC
 TCCCCAGAGGACGACCAGAGGGCTAACCGTACTCTGTTTCCCTGGGAAACCTGGACATCACCGTACTGAGTCCGACCTG
 AGACGCGCTTTCGACCGCTTCGTTGTGATCACCGAAGTGCACATCAAGAGGCCCTCCCGCGGTCAGACCTCAACTTAC
 GGTTTCTGAAGTTCGAAAACCTGGACATGTCCACCGCGCCAAGCTGGCTATGAGCGGCAAGATCATCATCCGTAAC
 CCCATCAAGATCGGATACGGCAAGGCTACTCCTACCCTAGGCTGTGGGTGGGTGGCCTGGGACCTGGGTCCCCCTG
 GCTGCCCTGGCTAGGGAGTTCGACAGATTCCGAACTATCCGACCATCGACTACCCTAAGGGTACTCTTGGGCCTAC
 ATCCAGTACGAATCACTGGACGCTGCCACGCTGCCTGGACTCACATGCGTGGTTTCCCACTGGGAGGTCTGACCGT
 AGGCTGCGCGTGCAGTTCGCTGACACCGAGCACCCTTACCAGCAGCAGTACCTGCAGCCCTGCCACTGACTCACTAC
 GAACTGGTACTGACGCTTTCGGTTCACAGGGCTCCTGACCCCTGCGTGGCGCCCGGACAGGACTCCACCTCTGCTG
 TACAGGGACAGAGACCGGACCTGTACCCTGACTCTGACTGGGTGCCCCACCTCCCCAGTCCGTGAGAGGTCTACC
 CGACTGCTGCCACCTCAGTCCCCGCTTACGAACCACTGGACTCCCTGGACAGACGCCGTGACGGATGGAGCCTGGAC
 AGAGACCGCGGTGACCGTACCTGCCAGCTCTCGCGACCAGCTCGTAAGAGGAGACTGCCTGAGGAACTCTGGCGGA
 AGGCACCTGGACAGATCCCCGAGAGCGACCGTCCAGGAAGAGACTGCGCTCCATCTCCTGACCGCTCACCAGAA
 CTGTATCCAGCCGACCGTTACAACCTCTGACAACGACCGCTCTTACAGTCTGTGCTGGAGAGGCCCTCACCATC
 AGAGACCGCGGTGGCTCCCTGGAAAAGAGCCAGGGAGACAAGAGGGACAGAAAGAACTCTGCTTACGCCGAGCGCGAC
 CGTAAGCACAGGACCACTGCTCCCACTGAGGAAAAGAGCCCACTGAAGAAGGAAGACCGTCTGACGGTTCAGCCCCA
 TCCACCAGCACTGCTTCCAGCAAGCTGAAGTCTCCTTACAAAAGCAGGACCGTGGCACCGCTCCTGTGGCTTCCGCC
 AGCCCAAAGCTGTGCTGGCTTGGCAGGGCATGCTGCTGCTGAAGAAGTCTAAGTCCCTCAAACCTGACACTCCTC
 CAGGGGACTGCAGGTGGCTCTTCACTGGTTCGAGGATCAACCGGTGGCAAGTCTCAGTCAAGATCACC
 CAGAGGCTGAGACTGGACACGCAAGCTGGACGAAGTACTAGGAGAATCAAGGTGCTGGCCCTAACGGATACGCC
 ATCCTGCTGGCTGTGCCCGGTTCCAGCGACTCCCGTCTTCAATCCAGCTCTGCTGCCTCCGACACCGCCACTAGCACC
 CAGAGGCTCTGAGAAACCTGGTCTCCTACCTGAGCAGAAGCAAGCTGCTGGTGTGATCAGCCTGCTGTGCGCGGA
 AACAAAGACAAGGAGAACACTGGTGTGCTGCACGCTTCCCTCCCTGCGAGTTCTCTCAGCAGTCTCCTGGACTCACCT

GCTAAGCCCTGGCTAAGTCCGAGGAAGACTACCTGGTTCATGATCATCGTCAGGGCCAAGCTGGTGGAGCAGAGAATG
AAGATTTGGAACCTCCAAGCTGGAAAACCTGTACTTCCAGGGTGGCTGGTTCACACCCCCAGTTCGAGAAGTAA

6.1.5 VIRMA

ATGGCTGTCGACTCTGCTATGGAGCTGCTGTTTCTGGACACCTTCAAGCACCCCTCAGCCGAACAGTCCAGCCACATC
GACGTGGTCCGTTTTCCCATGCGTGGTCTACATCAACGAAGTCCGTGTGATCCCTCCCGGAGTGAGGGCCCACTCTTCA
CTGCCGACAAACAGGGCTTACGGTGAACCTCTCCACACACTTTCCAGCTGGACCTGTTCTTCAACAACGTCTCAAAG
CCATCTGCTCCTGTGTTTCGACAGGCTGGGTTCACTGGAGTACGACGAAAACACCTCAATCATCTTCCGTCCAAACTCC
AAGGTCAACACTGACGGCTGGTGTGAGGGGATGGTACAACCTGCCTGACCCTGGCCATCTACGGCTCCGTCGACAGA
GTGATCTCACACGACAGGACTCCCCACCTCCCCACCTCCCCACCTCCCCACCTCAGCCACAGCCTTCCCTGAAG
CGTAACCCAAAAGCAGCTGACGGTGAAGGAAAGACCAGTTCAACGGTTCCTCCTCGCCCTCAGCCAAGAGGTCCA
AGGACTCCTCCCGGCCACCTCCCCAGACGACGACGAAGACGACCCTGTCCCCTGCCAGTGTCCGGTGACAAGGAA
GAGGACGCTCCACACCGTGAGGACTACTTCGAACCAATCAGCCCTGACAGGAACTCTGTCCCTCAGGAAGGACAGTAC
AGCGACGAAGGAGAGGTGGAGGAAGAGCAGCAGGAAGAGGGAGAAGAGGACGAGGACGACGTGACGTGGAAGAGGAA
GAGGACGAAGACGAGGACGACCGCCGTACCGTCGACTCTATCCCTGAAGAGGAAGAGGAAGACGAGGAAGAGGAAGGT
GAGGAAGACGAGGAAGGTGAAGGCGACGACGGCTACGAACAGATCAGCTCCGACGAGGACGGAATCGCTGACCTGGAA
CGCGAGACTTTCAAGTACCCAACTTCGACGTGGAGTACACCGCTGAAGACCTGGCCAGCGTCCCTCCCATGACTTAC
GACCCCTACGACCGTGAGCTGGTGCCACTGCTGTACTTCTCCTGCCCTTACAAGACCCTTTCGAAATCGAGATCAGC
AGGATGAAGGACCAGGGCCAGACAAGGAAAACCTCAGGAGCTATCGAGGCTCCGTCAAGCTGACCGAAGTGTGGAC
CTGTACCGTGAGGACAGGGGTGCTAAGTGGGTGACTGCCCTGGAGGAAATCCCTAGCCTGATCATCAAGGGCTGTCT
TACCTGCAGCTGAAGAACACCAAGCAGGACAGCCTGGGACAGCTGGTGGACTGGACTATGCAGGCTCTGAACCTGCAG
GTGGCCCTGCGTCAGCCTATCGCTCTGAACGTGAGGACGCTGAAGGCTGGCACCAGCTGGTGTCTTCACTGGCTGAA
TGCGGTGCTCAGGGAGTACTGGTCTGCTGCAGGCCGAGTCACTCCTCGGTCTGTTTCGAGCTGCTGTTTCGCTGACCAC
GTCTCCAGCTCTCTGAAGCTGAACGCTTTCGAAGGCCCTGGACAGCGTGAATCTATGACCAGGGGTATGGAAGCCTTC
CTGAGGGGCAGACAGAACGAAAAGAGCGGATAACCAGAAGCTGCTGGAGCTGATCCTGCTGGACCAGACCCTCAGAGTG
GTCACCTGCTGGCTCTGCCATCCTGCAGAAGTGCCACTTCTACGAGGTGCTGTGAGAAATCAAGCGCCTGGGAGACCAC
CTGGCTGAAAAGACTTTCATCCCTGCCTAACCACTCTGAGCCTGACCACGACACTGACGCTGGTCTGGAAAGAACCAAC
CCCGAATACGAGAACGAAGTGGAGGCTTCTATGGACATGGACCTGCTGGAGAGCTCTAACATCTCAGAAGGCGAGATC
GAACGCTGATCAACCTGCTGGAGGAAGTCTTCCACCTGATGGAGACCGCCCCACACACTATGATCCAGCAGCCAGTG
AAGTCTTCCCTACTATGGCTAGAATCACTGGCCACCTGAACGCGACGACCCCTTACCCCGTCTGTTCGGCTACCTG
CACAGCCACCCTTCTGGAGCTGGTACTCTGCTGTGTCATCCAGTGACTTCTGCTCACCCCTGGTGTGCTGCAG
GCTACCAAGGACGCTCCTGAAGTTCCTGGCTCAGTCTCAGAAGGGCTGCTGTTCTTATGTCAGAGTACGAAGCCACT
AACCTGCTGATCCGCGCTCTGTGCCACTTCTACGACCAGGACGAGGAAGAGGGTCTGCAGTCCGACGGCGTCATCGAC
GACGCTTTCGCCCTGTGGCTGCAGGACTCCACCCAGACTCTGCAGTGCATACCGAGCTGTTTCAGCCACTTCCAGAGA
TGCACCGCTTCCGAAGAGACTGACCCTCTGACCTGCTGGGCACTCTGCACAACCTGTACCTGATCACCTTCAACCT
GTGGGACGCTCAGCCGTCCGTACGCTGTTCTCCCTGGAGAAGAACCTGCAGAGCCTGATCACTCTGATGGAGTACTAC
TCTAAGGAAGCTCTGGGAGACTCAAAATCCAAGAAGTCACTCGCTTACAACCTACGCTGCATCCTGATCCTGGTGGTC
GTGCAGTCAACGACGCTGCAGATGCTGGAGCAGCAGCTGCCTCCCTGCTGAAGCTGTGCAAGGCTGACGAAAAC
AACGCTAAGCTGCAGGAACCTGGGCAAGTGGCTGGAGCCACTGAAGAACCTGCGTTTCGAAATCAACTGCATCCCTAAC
CTGATCGAGTACGTCAAGCAGAACATCGACAACCTGATGACTCCTGAGGGAGTGGGTCTGACCACTGCTCTGAGGGTC
CTGTGCAACGCTGGCTTGCCTCCTCCTCCTGTCGAAGGACAGCAGAAGGACCTGAAGTGGAACTGGCCGTGATCCAG
CTGTTCTGCTGAGGGTATGGACACCTTTCATCAGAGTGTGCGAGAAGCTGAACCTCAATCCTGACTCAGCCTTGGCGC
CTGCAGCTGAACATGGGCACCACTCTGCACAGAGTACCCTATCTCTATGGCCCGCTGCACCCTGACTCTGCTGAAG
ACCATGCTGACTGAACTGCTGAGAGGTGGCTCCTTCGAGTTCGAAGGACATGCGCGTCCCTAGCGCTCTGGTGACCCTG
CACATGCTGCTGTGCTCAATCCCCCTGTCCGGACGCTCTGGACTCAGACGAACAGAAGATCCAGAACGACATCATCGAC
ATCCTGCTGACCTTCACTCAGGGTGTGAACGAGAAGCTGACCATCTCTGAAGAGACTCTGGCTAACAAACCTGGTCA
CTGATGCTGAAGGAAGTGTGCTTCTCAATCCTGAAGGTCCCTGAAGGCTTCTTCTCAGGACTGATCCTGCTGTCCGAG
CTGCTGCCACTGCCTCTGCCATGCAGACCCTCAGGTATCGAACCCACGACATCTCCGTGCGCCCTGAACACCAGA
AAGCTGTGGAGCATGCACCTGCACGTGCAGGCTAAGCTGCTGCAGGAGATCGTCCGCAGCTTCTCTGGCACCCTTGC
CAGCCATCCAGCACATGCTGAGGAGAATCTGCGTGCAGCTGTGCGACCTGGCTTCCCCAACTGCCCTGCTGATCATG
AGAACCGTCTGGACCTGATCGTGGAGGACCTGCAGTCAACTTCCGAGGACAAGGAAAAGCAGTACACTTCTCAAACCT
ACCAGGTGCTGGCCCTGCTGGACGCTCTGGCCAGCCACAAGGCTTGAAGCTGGCCATCCTGCACCTGATCAACGGC
ACCATCAAGGGCGACGAGCGTTACGCCGAAATCTTCCAGGACCTGCTGGCTCTGGTGCCTAGCCAGGCGACTCTGTC
ATCAGGCAGCAGTGCCTCGAGTACGTGACTTCAATCCTGCAGTCCCTGTGCGACCAGGACATCGCTCTGATCCTGCCT
TCCAGCTCTGAGGGAAGCATCTCTGAGCTGGAACAGCTGTCAAACCTCCCTGCCAACAAGGAACTGATGACTTCCATC
TGCGACTGCCTGCTGGCTACCTGGCCAAACGCGAGTCACTTACAACCTGCCTGCTGACCTGCGTGCACCTATGATG

TTCTTGCCGAGCAGACTACGGACTGTTCCACCTGAAGAGCTCTCTGAGAAAAGAACTCATCCGCTCTGCACTCCCTG
 CTGAAGCGCGTCTGTCAACTTTCTCCAAGGACACCGGTGAAGTGGCCAGCTCTTCTCTGGAGTTCATGCGTCAGATC
 CTGAACAGCGACACCATCGGCTGCTGCGGAGACGACAACGGCCTGATGGAGGTGGAAGGAGCTCACACCAGCAGGACT
 ATGTCTATCAACGCTGCCGAGCTGAAGCAGCTGCTGCAGTCAAAGGAAGAGTCCCCTGAGAACCCTGTTCTGGAGCTG
 GAAAAGCTGGTGTGGAACACTTAAGGACGACGACAACCTGGACAGCCTGCTGGACTCTGTGCTGGGACTGAAGCAG
 ATGCTGGAGTCAATCCGGTGACCCACTGCCTCTGTGTCAGACCAGGACGTCGAACCCGTGCTGTGTCAGCCCCAGAGTCCCTG
 CAGAACCCTGTTCAACAACCGCACCCGCTTACGTCTGGCCGACGCTGATGGACGACCAGCTGAAGTCTATGTGGTTCACT
 CCCTTCCAGGCTGAAGAGATCGACACCGACCTGGACCTGGTCAAGTGGACCTGATCGAGCTGAGCGAAAAGTGTGCTGC
 TCTGACTTCGACCTGCACTCAGAGCTGGAACGCTCCTTCTGTCTGAACCCAGCTCTCCAGGCCGACCAAGACCCT
 AAGGGATTCAAGTGGGCAAGCACAAGCAGCAGACCTTCATCATTTCATCCGAAAGTCCGAGTACATCGAACCCGCT
 AAGAGGGCTCACGTGCTGCCACCTCCCCGCGCCGTGGAAGGGGAGGTTTCGGTCAGGGCATCCGTCCACACGACATC
 TTCAGACAGCGCAAGCAGAACCAGCAGGCCACCTTCTATGCAGTGCAGACTTCGTGGCTGCCGAGAGCAAGGAA
 GTCGTGCCTCAGGACGGTATCCCTCCTTAAGCGTCCCCTGAAGTGTCTCAGAAGATCAGCTCTAGGGGCGGATTC
 TCTGGTAACAGAGGTGGCCGCGCGCCTTCCACTCTCAGAACAGATTCTTCACTCCCCAGCTTCAAAGGGTAACAC
 TCCC GCCGTGAAGTACTCGCGGCTCATCTGGTCAGCCAGAACACCCACGTTGCAACTACAACGAGAGCAGGGGA
 GGTCAATCTAACTCAACAGAGGACCTCTGCCTCCCCTGCGCCCTCTGAGCTCTACTGGCTACCGTCCAAGCCCAGT
 GACCGTGTCCCCGCGGTCGCGGTGGTCTGGGACCATCATGGGCTTCCGCCAACAGCGGTTCCGGTGGTTCGCCGCGC
 AAGTTCGTGAGCGAGGTTCTGGACGTGGTAGGACGTCGTTCTTCACTAGGGAGAACCCTGTACTTCCAGGGCGGA
 TGAGCCACCCCTCAGTTCGAAAAGTAA

6.1.6 ZC1H13

ATGTCAAAGATCCGCCGTAAAGTACTGTGAGAACACCAAGACTATCTCCGACAGCACCTCTAGGAGACCATCTGTG
 TTCGAGAGGCTGGGACCATCTACCGGATCAACTGCTGAAACCCAGTCCCGTAACTGGCTGAAGACTGGAAGTGCCTG
 TACGGTAACACCTGCAGGTTTCGTCCACGGTCCCTTCTCCAGAGGAAAGGGTTACTCCAGCAACTACCGCCGTTACCT
 GAGAGGCCAACTGGCGACCTGAGAGAACGCATGAGAACAAGCGCCAGGACGTTGACACTGAACCCAGAAGCGTAAAC
 ACCGAGGAATCTTCATCCCCAGTCAAGGAGAGCTCTCGTGGAAAGGCACAGAGAAAAGGAAGACATCAAGATCACT
 AAGGAACGTACCCAGAGTCCGAGGAAGAGAACGTTGGAATGGGAGACCAACAGGGACGACAGCGCAACCGGTGACATC
 AACTACGACTACGTCCACGAACGTCTCTGGAGATGAAGCGCCAGAAGATCCAGCGTGAGCTGATGAAGCTGGAACAG
 GAGAACATGGAGAAGCGCAAGAGATCATCATCAAGAAGGAAGTGTCCCCTGAGGTGGTCCGTAGCAAGCTGTACCA
 TCCCCTAGCCTGCGCAAGTCAATCAAGAGCCCTAAGCGTAAGAGCTCTCCAAGTCACTCCAGCGCTCCAAGAAGGAC
 CGTAAGACCAGCGCTGTCTCTTCACTCTGTGGACCAGCAGCGTAAACAGCAAGACTAACAGTCCAAGAAGAAGGGA
 CCTCGTACTCCATCCCCTCCCCACCTATCCCAGAGGACATCGCCCTGGGAAAAGTACAAGGAAAAGTACAAGGTG
 AAGGACAGGATCGAAGAGAAAAGTACAGACGGCAAGGACCGCGCCGCTGACTTCGAAAAGGCAGCGTGAGAAGCGCGAC
 AAGCCCCGTTCACTACCCAGCTGGTCAAGCACCCTCCCCAATCTCCAGCAGGCACCCTCTTCACTCCAGCCAGAGC
 GGCTCTTCAATCCAGCGCCACTCCCCAGCCCAAGGAGAAAGAGGACCCCTTCTCCCTCATAACCAGAGAACCCTGACT
 CCCCCACTGCGCCGTTCCGCCAGCCATACCCTTCTCACTACTGTCCAGCCCTCAGAGGAAGCAGTCTCCTCCCCGC
 CACCGTTACCCATGAGGGAAAAGGGTAGACACGACCACGAGCGCACTTCCAGAGCCACGACAGGAGACACGAACGC
 CGTGAGGACACCCGTGGCAAGAGGGACAGAGAGAAGGACTCCCGTGAAGAGAGGGAAATACGAGCAGGACCAGTCTTCA
 TCCAGAGACCACCGCAGACCCGTGAACCTAGGGACGGTAGAGACAGGAGAGACGCCCGCAGACCCCGTGAACCGCTG
 GAGCTGAGGGACTCCAGAGACATGCGCGACAGCCGTGAGATGAGGGACTACTCTCGTGACACTAAGGAATCAAGAGAC
 CCACGCGACTCTCGTTCAACCAGGGACGCTCAGACTACAGGGACAGGGAGGGAAGGGACACTCACCGTAAGGAAGAC
 ACCTACCCTGAAGAGTCCAGGAGCTACGGTAGAAACCCTTGCAGGAAAGAGAGCTCTCGTACTGAAAATCAGGAACGAG
 TCAAGAAACGAATCTCGCTCAGAGATCAGAAACGACCGCATGGGAAGGTCAGAGGTCGCGTGCCAGAACTGCCAGAG
 AAGGGTTCGCCGCGGTTCCCGCGGTTCCAGATCGACTCACACTCATCCAACCTCCAACCTACCACGACAGCTGGGAGACC
 CGCAGCTCTTACCCAGAACGTGACCGCTACCTGAGAGAGACAACCGCGACCAGGCTAGAGACTCATCCTTCGAAAAGG
 AGACACGGCGAGAGGGACCGCCGTGACAACAGAGAGCGCGACCAGCGTCCAGCTCTCCAATCCGTACCAGGGAAGG
 AACGACGAACCTGGAGAGGGACGAGAGGAGAGAAGAGCGCCGTGTGGACAGAGTCCAGCAGAGGAGAGACGAACGTGCC
 AGGGAGAGAGACCAGCAACGTGAGAGGGACAGAGAACCGAGCGTGAAGGGGAGAGAGAACCGACCCGTGAGAGGGAA
 AAGGAGCGTGAACCTGGAGAGAGAACGCGCTCGTGAGAGGGAAAGAGAGCGCGAAAAGGAGCGTGATCGTGAGAGAGAC
 CGCGACCGTGACCACGACCCGTGAGAGAGAGCGTGAGAGGGAAAGGGACAGAGAAAAGGAGCGCAACCTGAGAGGGAA
 GAGCGTGAGCGTGAAGGGGAGAGGGAGCGTGAGAGGGAAAGGGACAGAGAAAAGGAGCGCAACCTGAGAGGGAA
 GACAAGGAACCGAGCGTCAAGGGACTGGGAAGACAAGGACAAGGGTTCGCGACGACCCCGTGAAGAGCGTGAAGAG
 ATCCGTGAGGACAGGAACCCAGAGACGGACACGATGAGCGCAAGGACAAGAAGAGATACCGCAACGAAGGTTCCCCC
 AGCCACGTCAGTCTCCTAAGAGGAGACGCAACACTCACCCGACTCCGACGTTACAACCTCTGGAGACGACAAGAAC
 GAGAAGCACCCCTGCTGTACAGGTGGTCAAGGCCACAGGAGTCCCGCTCCCTGTCCCCTAGCCACCTGACCGAAGAC
 CGCCAGGGTTCGTTGGAAGGAAGAGGACAGAAAGCCAGAACGCAAGGAGTCAATCCCGTAGGTACGAGGAGCAGGAAC

Appendix

AAGGAGAAGGTGAGCTCTGTGCGACAAGCAGCGTGAACAGACTGAGATCCTGGAATCATCCCGTATGAGGGCTCAGGAC
ATCATCGCCACCACCAGTCTGAGGACAGGGAAACCTCTGACAGGGCTCACGACGAGAACAAGAAGAAGGCTAAGATC
CAGAAGAAGCCTATCAAGAAGAAGAAGGAAGACGACGTGGGCATCGAGCGCGGAAACATCGAAACCCTTCTGAGGAC
GGTCAGGTCTTCTCACCAAAGAAGGGCCAGAAGAAGAAGTCCATCGAGAAGAAGCGTAAGAAGAGCAAGGGCGACAGC
GACATCTCTGACGAAGAGGCTGCCAGCAGTCCAAGAAGAAGCGCGGACCCAGAACTCCACCTATCACCCTAAGGAA
GAGCTGGTGGAGATGTGCAACGGAAGAACGGTATCCTGGAAGACTCTCAGAAGAAGGAAGACACCGCCTTCTCAGAC
TGGTCCGACGAAGACGTGCCCGACCGCACTGAGGTACCGAAGCTGAGCACACCGCTACTGCCACCCTCCAGGTAGC
ACTCCTTCTCCCCTGAGCTCTGTGCTGCCCCACCTCCCCAGTGGCTACTGCTACTGCTACTACTGTCCCCGCTACC
CTGGCTGCTACCCTGCTGCCGCTGCCACTTCTTTCTCAACCTCCGCCATCACTATCAGCACCTCTGCTACCCCAACT
AACACCCTAACAACACTTTCGCCAACGAGGACTCCCACCGAAGTGCCACAGAACCCTCGGTGGAGAAGGTCGAAACT
CCACACGTGACCATCGAGGACGCTCAGCACAGGAAGCCTATGGACCAGAAGAGATCATCCAGCCTGGGATCAAACCGC
TCCAACCGTAGCCACACCTCTGGTCGTCTGAGGTACCTTCCAACGACTCAGCTCACCGTTCCGGCGACGACCAGTCC
GGAAGAAAGCGCTCTGACAGCGGCTCTCGTGACAGGGAGAAGACTAAGTCTCTGGAAATCACCGGCGAAAGGAAG
TCAAGAAATCGACCAGTGAAGAGGGGAGAGCCTTCAAGATCCACTTCTTTCAGACCGTCCAGGACAGCAGGTCTACTCC
AGCAGACGCTCTTCCACCGAGTCTGACCGCCAGGTCCACTCACGTTCCGGTTCCTTCGACTCAAGAGACCCTGCAG
GAACGTGACCGTTACGAGCACGATCGTGAACGCGAGCGTGAACGTAGGGACACTCGTCAGAGAGAGTGGGACAGGGAC
GCTGACAAGGACTGGCCACGCAACCGTGACAGGGACAGACTGAGGGAAAGGGAGCGTGAAAGAGAACCGGACAAGAGA
CGCGACCTGGACCGTGAGAGGGAAAGACTGATCTCTGACTCAGTGGAGCGGACCGTGACAGGGACAGAGACCGCACT
TTCGAATCCAGCCAGATCGAGTCCGTCAGCGTTGCGAGGCCAAGCTGGAGGGCGAACACGAGCGCGACCTGGAATCC
ACCAGCCGTGACTCACTGGCTCTGGACAAGGAGCGCATGGACAAGGACCTGGGTTCCGTGCAGGGCTTCGAAGAGACT
AACAAGTCTGAGCGTACCGAATCACTGGAGGGCGACGACGAGAGCAAGCTGGACGACGCTCACTCACTGGGCTCCGGA
GCTGGTGAAGGCTACGAGCCCATCTCTGACGACGAACCTGGACGAGATCCTGGCCGAGACGCTGAAAAGCGTGAGGAC
CAGCAGGACGAAGAGAAGATGCCAGACCCTCTGGACGTGATCGACGTCGACTGGTCCGGTCTGATGCCTAAGCACCT
AAGGAACCACGCGAGCCTGGAGCTGCTCTGCTGAAGTTCCTCCTGGTCCGTGATGCTGCGTGTGCGCATCTCTAAG
AAGCTGGCCGATCAGAACTGTTTCGCTAAGGTCAAGGAGACCTGCCAGCGCCTGCTGGAAAAGCCCAAGGACGCCGAC
AACCTGTTGCAACACGAGCTGGGAGCTCTGAACATGGCTGCCCTGCTGAGAAAAGGAAGAGCGCGCTTCCCTGCTGTCT
AACCTGGGTCCATGCTGCAAGGCCCTGTGCTTCCGTAGGGACTCCGCTATCCGTAAGCAGCTGGTGAAGAACGAAAAG
GGTACTATCAAGCAGGCCTACACCTCCGCTCCAATGGTCGACAACGAGCTGCTGCGTCTGAGCCTGAGGCTGTTCAAG
AGAAAGACCACTTGCCACGCTCCTGGCCACGAAAAGACCAGGACAACAAGCTGAGCCAGTCTTCAATCCAGCAGGAA
CTGTGCGTGTCCGAGAACCTGTACTTCCAGGGTGGCAGCATGGACTACAAGGACGACGACGACAAGTAA

6.2 RNA melting point analysis

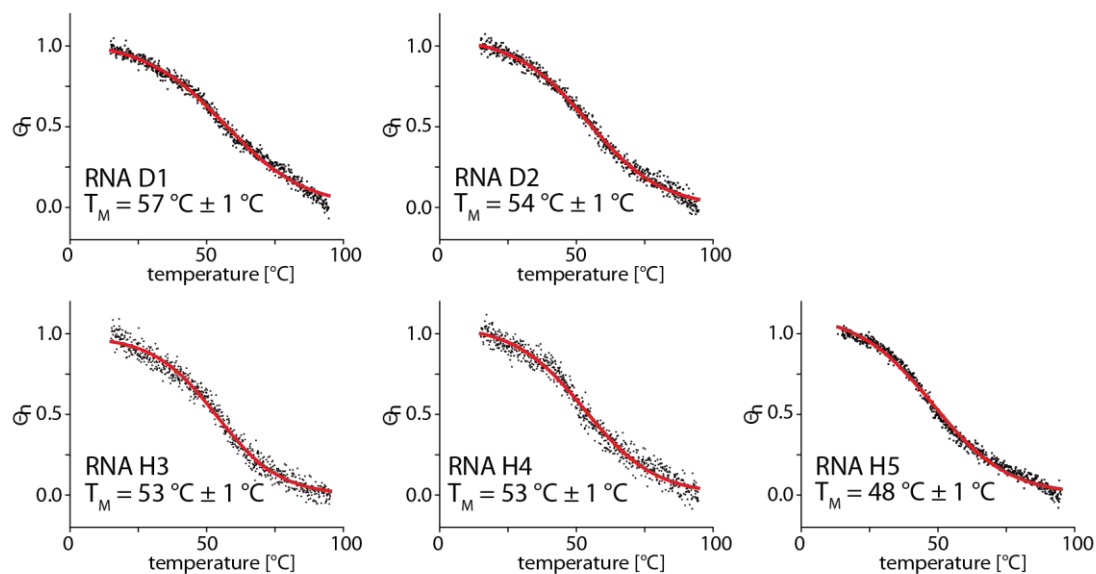


Figure 74: RNA melting point analysis with CD-spectroscopy. Fitted curves are indicated in red and determined melting points are given.

6.3 Optimization of assay conditions

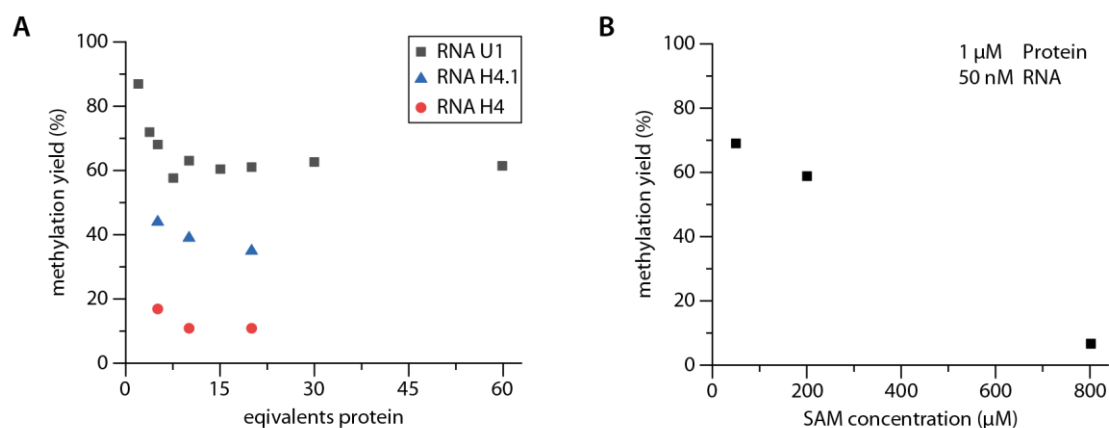


Figure 75: **A** Optimization of the protein to RNA ratio for single turnover conditions. The concentration of three different RNAs were varied while the concentration of the full-length METTL3/METTL14 complex was kept constant (750 nM, 150 μM SAM). For RNA U1 a concentration range of 12.5 to 800 nM and for RNAs H4 and H4.1 a concentration range of 50 to 200 nM was tested. **B** Optimization of the SAM concentration with 50 nM RNA U1 and 1 μM complex METTL3/METTL14.

6.4 Protein purification of shortened METTL3/METTL14 complexes

6.4.1 M3 Δ ZF1/METTL14

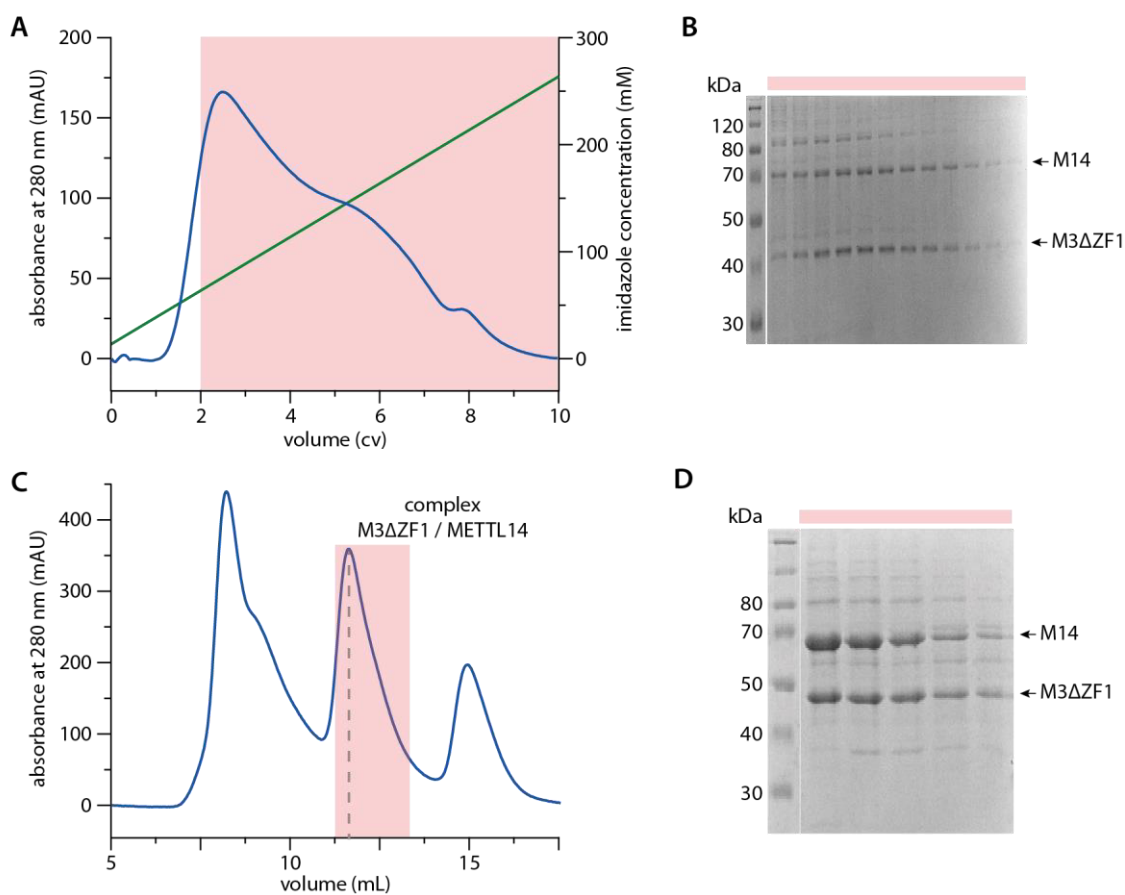


Figure 76: **A** Chromatogram of the HisTrap purification for complex M3 Δ ZF1/METTL14. Combined fractions for further purification are indicated. **B** SDS-PAGE analysis of the combined elution fractions from the HisTrap purification. **C** Chromatogram of the size exclusion purification for complex M3 Δ ZF1/METTL14. Combined fractions are indicated. **D** SDS-PAGE analysis of the combined elution fractions from the size exclusion purification.

6.4.2 M3MTD/METTL14

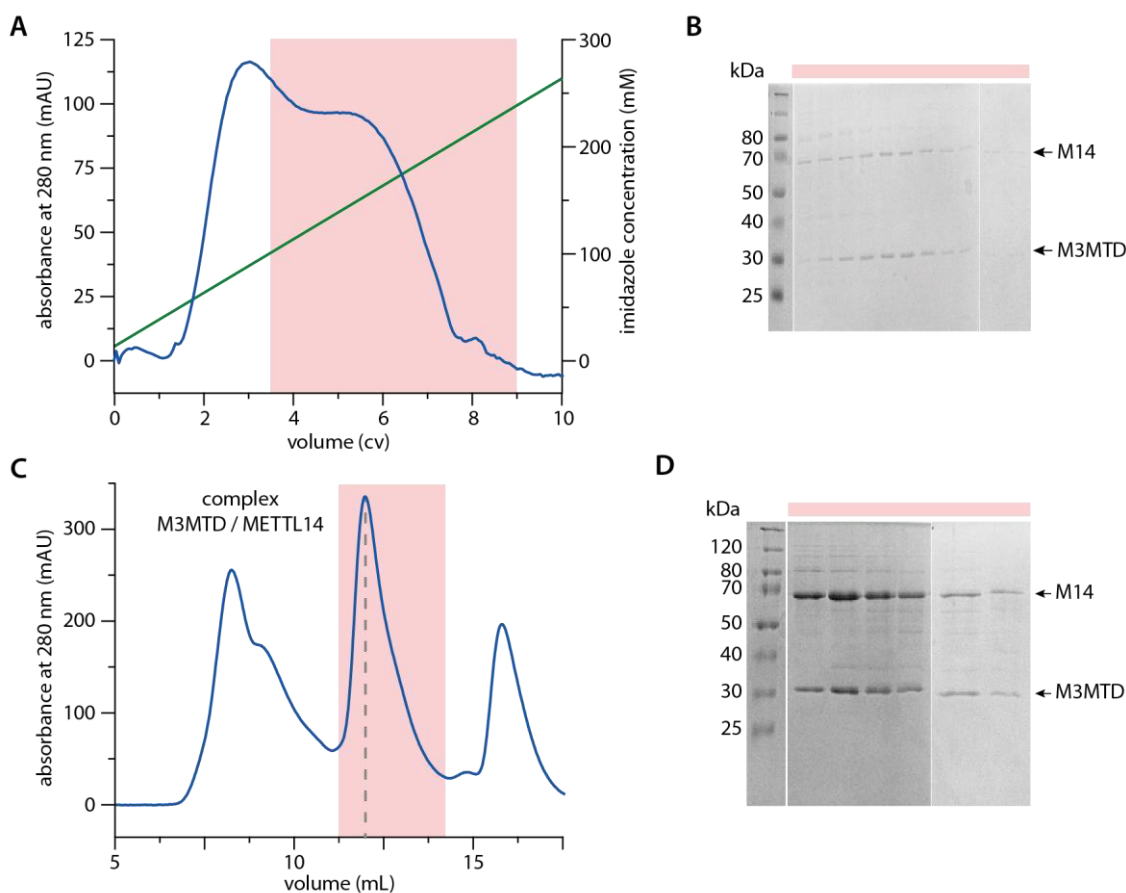


Figure 77: **A** Chromatogram of the HisTrap purification for complex M3MTD/METTL14. Combined fractions for further purification are indicated. **B** SDS-PAGE analysis of the combined elution fractions from the HisTrap purification. **C** Chromatogram of the size exclusion purification for complex M3MTD/METTL14. Combined fractions are indicated. **D** SDS-PAGE analysis of the combined elution fractions from the size exclusion purification.

6.4.3 METTL3/M14ΔRGG

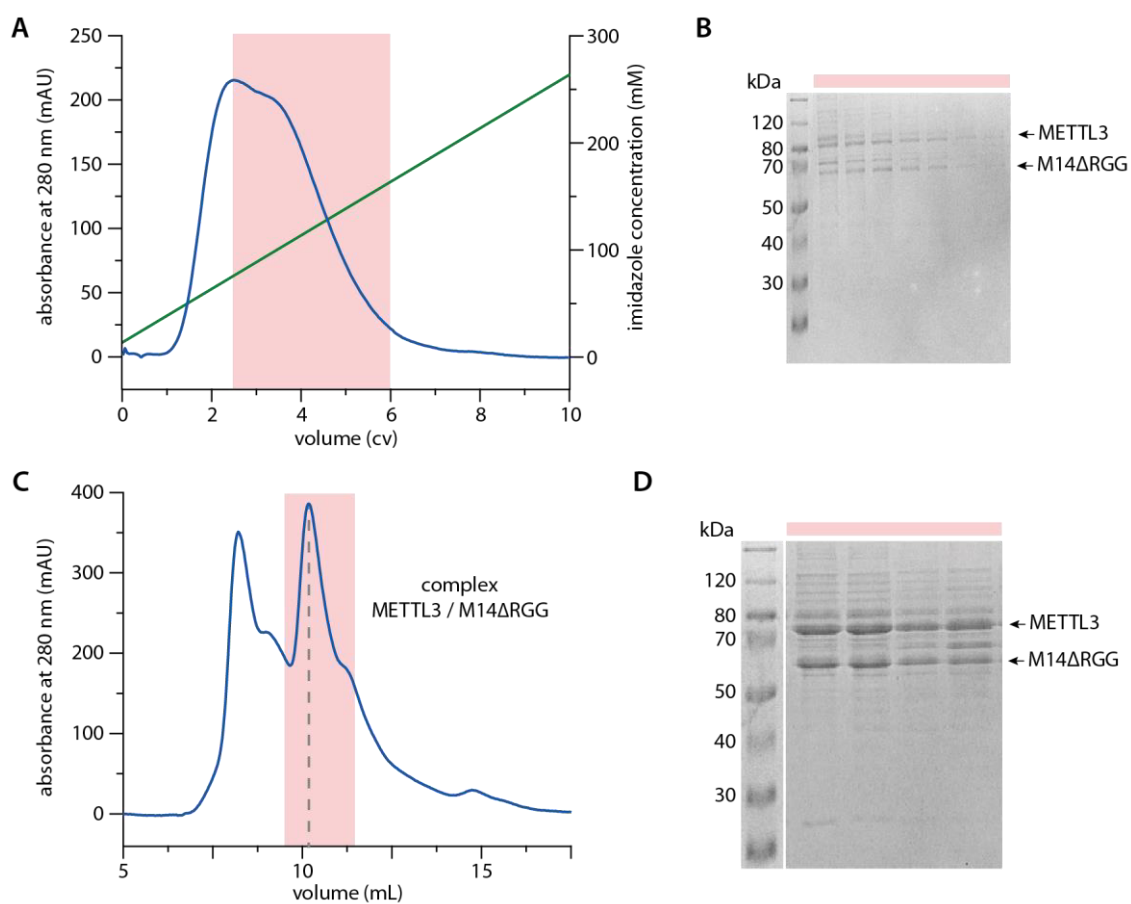


Figure 78: **A** Chromatogram of the HisTrap purification for complex METTL3/M14ΔRGG. Combined fractions for further purification are indicated. **B** SDS-PAGE analysis of the combined elution fractions from the HisTrap purification. **C** Chromatogram of the size exclusion purification for complex METTL3/M14ΔRGG. Combined fractions are indicated. **D** SDS-PAGE analysis of the combined elution fractions from the size exclusion purification.

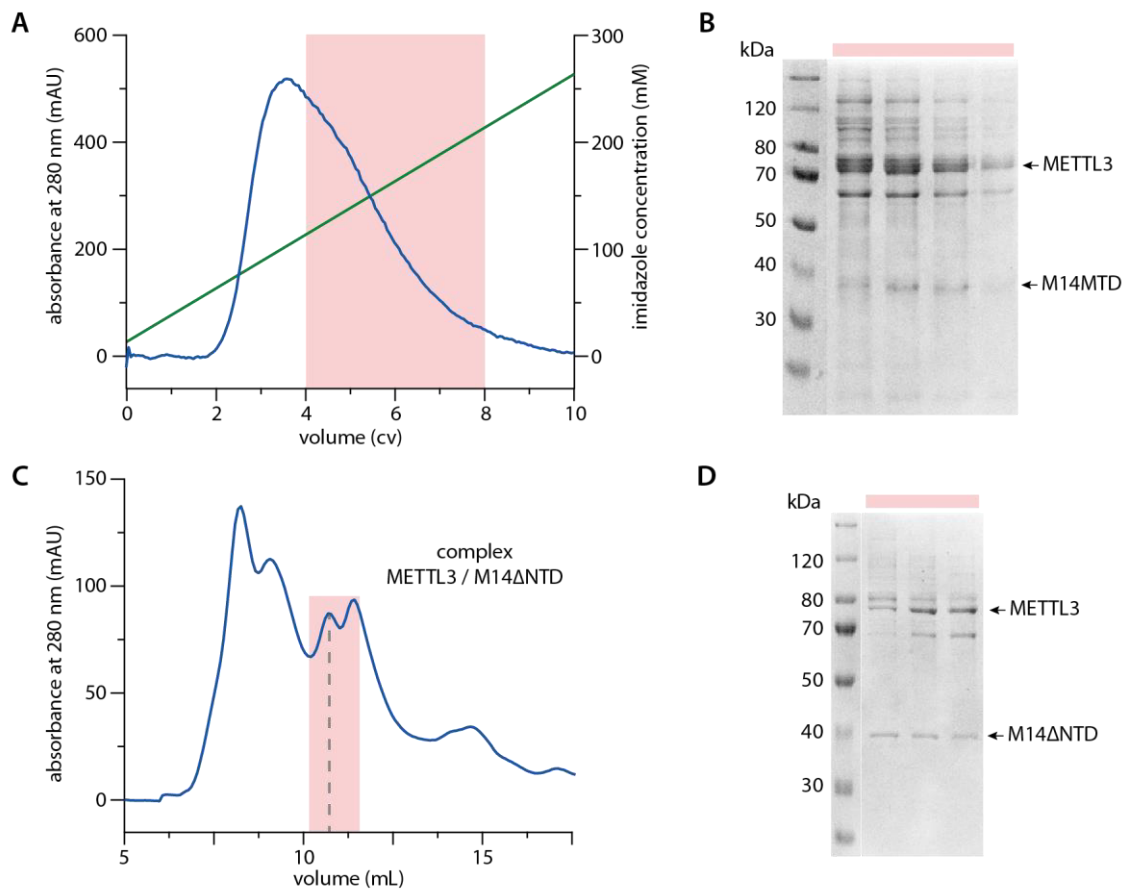
6.4.4 METTL3/M14 Δ NTD

Figure 79: **A** Chromatogram of the HisTrap purification for complex METTL3/M14 Δ NTD. Combined fractions for further purification are indicated. **B** SDS-PAGE analysis of the combined elution fractions from the HisTrap purification. **C** Chromatogram of the size exclusion purification for complex METTL3/M14 Δ NTD. Combined fractions are indicated. **D** SDS-PAGE analysis of the combined elution fractions from the size exclusion purification.

6.4.5 METTL3/M14MTD

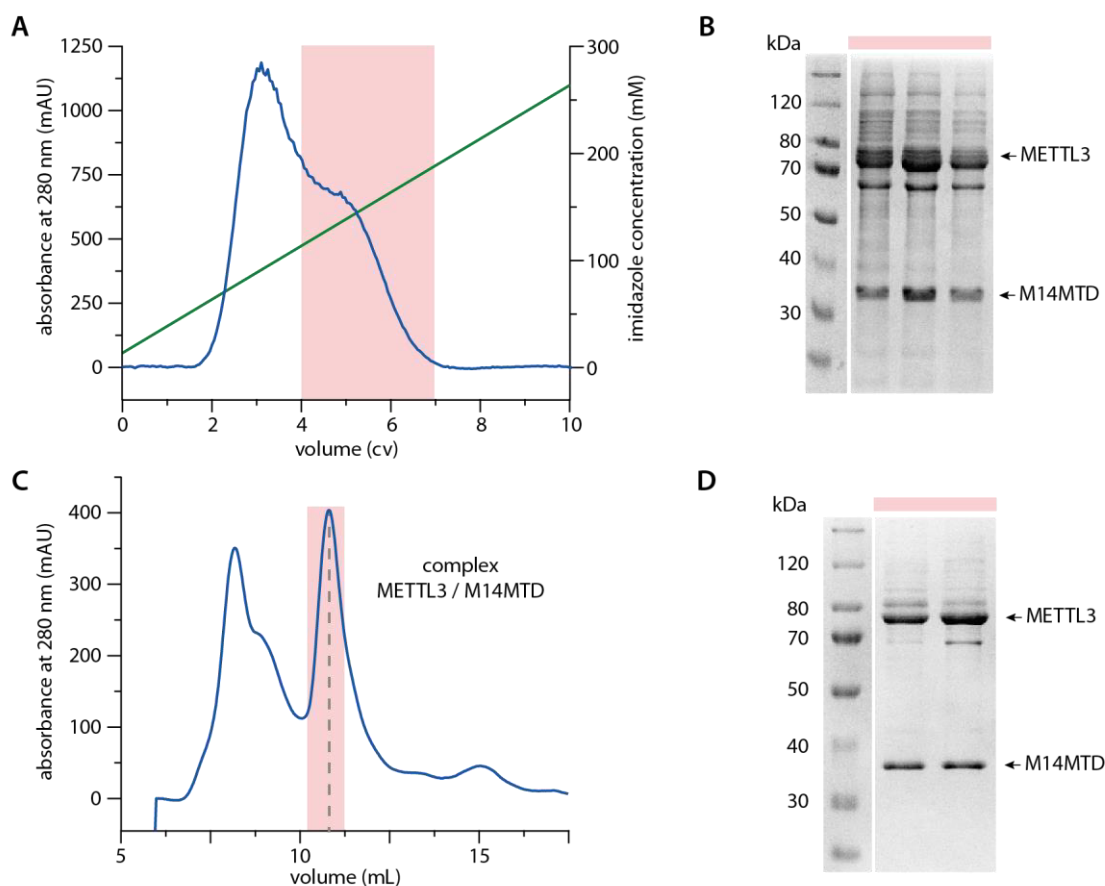


Figure 80: **A** Chromatogram of the HisTrap purification for complex METTL3/M14MTD. Combined fractions for further purification are indicated. **B** SDS-PAGE analysis of the combined elution fractions from the HisTrap purification. **C** Chromatogram of the size exclusion purification for complex METTL3/M14MTD. Combined fractions are indicated. **D** SDS-PAGE analysis of the combined elution fractions from the size exclusion purification.

6.4.6 M3MTD/M14MTD

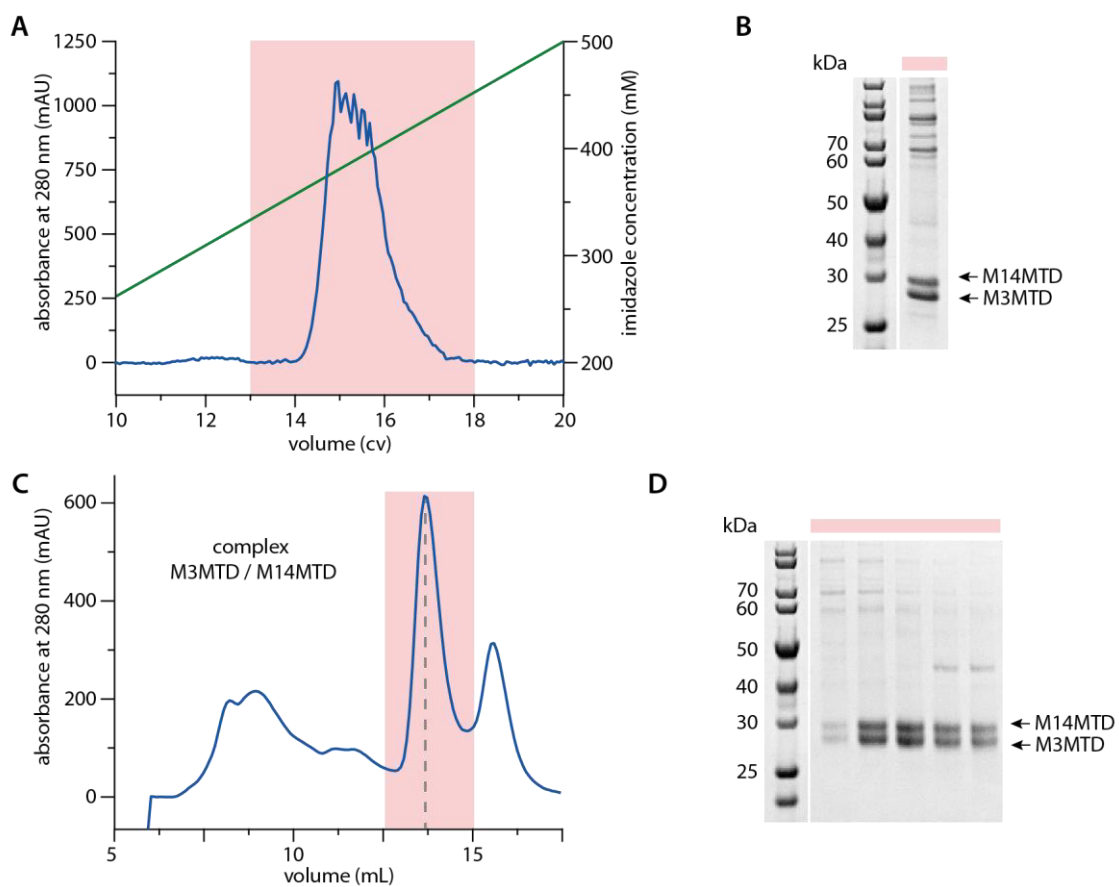


Figure 81: **A** Chromatogram of the HisTrap purification for complex M3MTD/M14MTD. Combined fractions for further purification are indicated. **B** SDS-PAGE analysis of the combined elution fractions from the HisTrap purification. **C** Chromatogram of the size exclusion purification for complex M3MTD/M14MTD. Combined fractions are indicated. **D** SDS-PAGE analysis of the combined elution fractions from the size exclusion purification.

6.5 Protein purification of the METTL3/METTL14/(short)WTAP complex

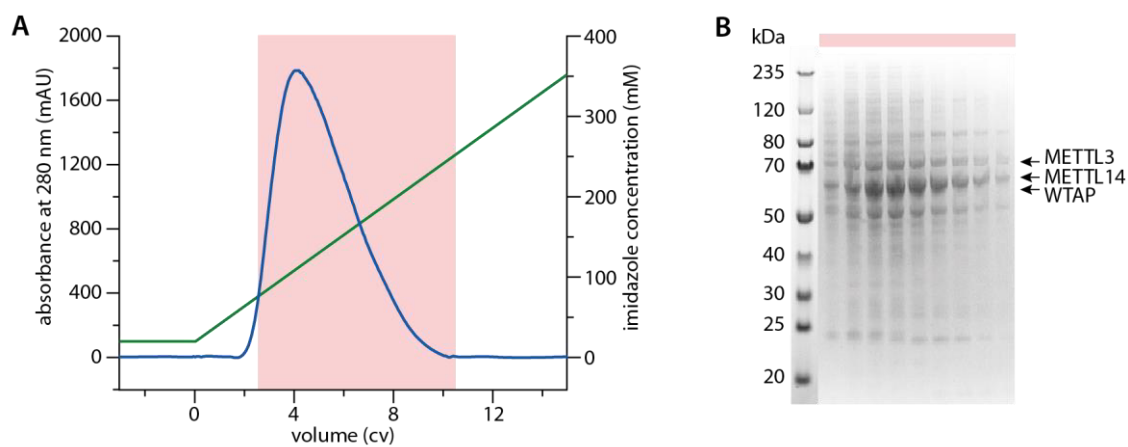


Figure 82: **A** Chromatogram of the HisTrap purification for complex METTL3/METTL14/WTAP. Combined fractions for further purification are indicated. **B** SDS-PAGE analysis of the fractions from the HisTrap purification.

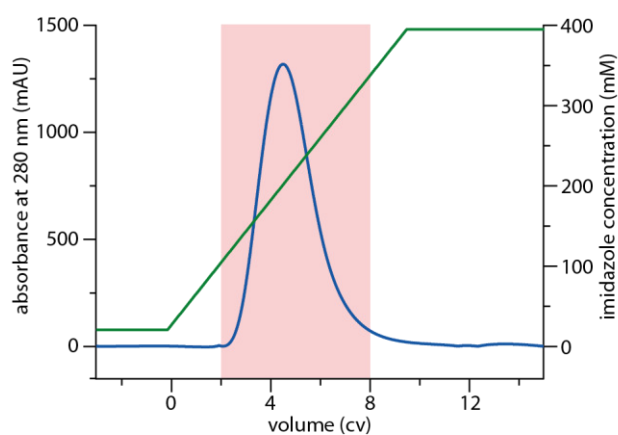


Figure 83: **A** Chromatogram of the HisTrap purification for complex METTL3/METTL14/shortWTAP. Combined fractions for further purification are indicated.

6.6 Custom made components for the smFRET setup

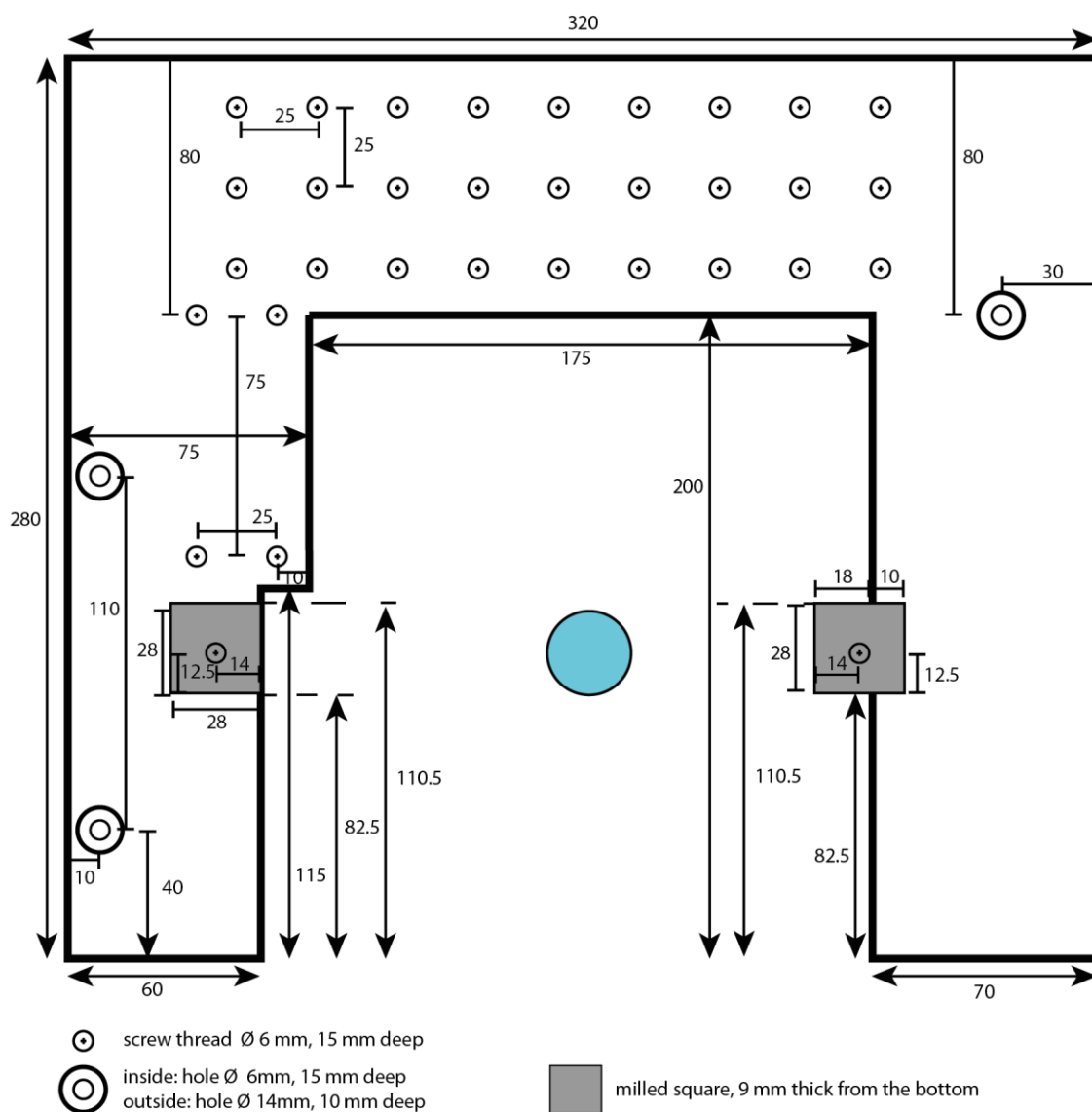


Figure 84: Detailed illustration of the custom-made aluminium breadboard (15 mm thick) for the microscope desk, designed to locate the prism above the center of the objective (blue circle). All dimensions are given in mm.

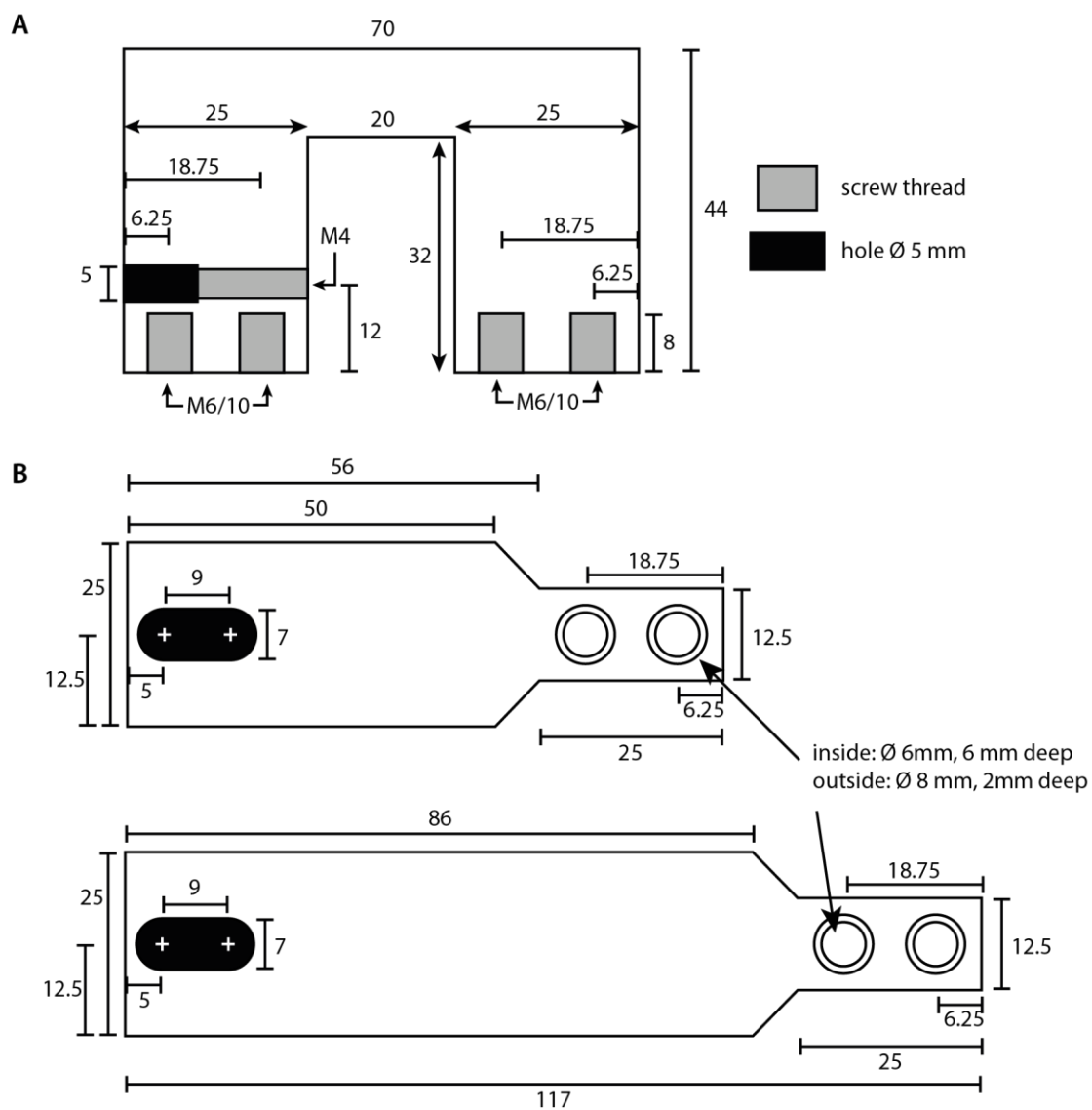


Figure 85: Prism bridge assembled from **(A)** the center of prism bridge (aluminium, 12.5 mm thick) and **(B)** two sides of supporting structure (aluminium, 6 mm thick). All dimensions are given in mm.

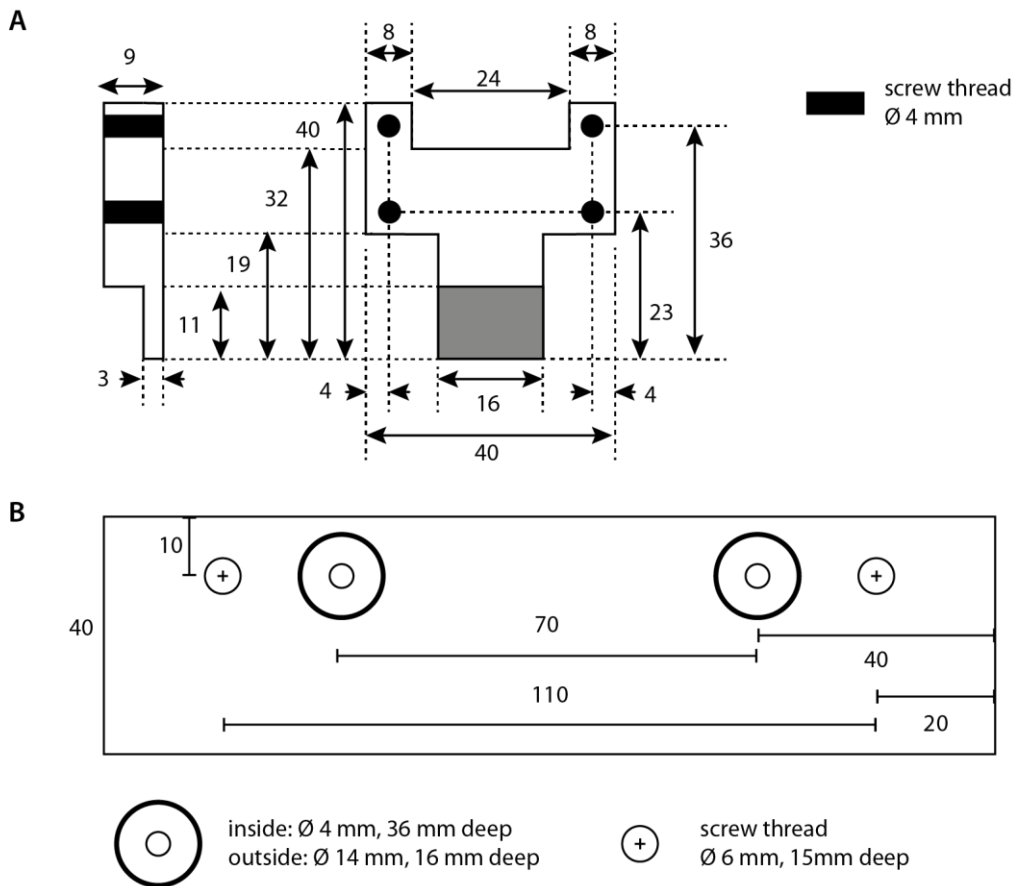


Figure 86: A Custom-made prism holder (aluminium, 9 mm thick). **B** Adaptor (aluminium, 36 mm thick) to secure the custom-made microscope desk onto the microscope. All dimensions are given in mm.

6.7 Matlab scripts for smFRET data analysis

6.7.1 Analysis of histogram data

Script name: readhisttracesandor_acceptorfilter

```
%Program to read binary traces file output by Hazens IDL program written by
%Michael Stone 01/17/2004
clear all;
total_frets = [];
n_molecules = [];
n_fr_analysis = 20;
red_on = 27; %input ('Which frame was red laser turned on?');
red_threshold = 2500 %input ('What is your acceptor threshold?');

s = pwd;
files = dir(s);

for i = 3:length(files)

fileName = files(i).name;
s = strcat(s,'\');
fullFileName = strcat(s,fileName);
FID = fopen(fullFileName);

% data = importdata(fullFileName);
accepted_traces = []; % an array to hold good molecule data for future
analysis
fretthreshold = 2000; % this is a threshold value to prevent crazy fret traces

%find number of frames and peaks from traces header file
[n_fr,z1] = fread(FID,1,'int32'); %z1 is just the number of indicated data
types read (should be 1)
[n_traces,z2] = fread(FID,1,'int16');
rate_fr = 10;%input('At what frame rate was the data collected?');
n_peaks = n_traces/2;
n_molecules = [n_molecules n_peaks];
exp_length = n_fr/rate_fr; %(seconds)
time = (1/rate_fr:(1/rate_fr):exp_length);

%% Now read in the rest of the data from the traces file
[total_data,z3] = fread(FID,[n_traces+1,n_fr],'int32');
fclose(FID);
total_data = total_data(2:n_traces+1,1:n_fr);
total_donor = [];
total_acceptor = [];
total_acceptor_corrected = [];
total_donor(1:n_peaks-1,1:n_fr) = total_data(2*(0:n_peaks-2)+1,1:n_fr);
%%remember matlab starts arrays at 1
total_acceptor(1:n_peaks-1,1:n_fr) = total_data(2*(0:n_peaks-2)+2,1:n_fr);

%correct the acceptor intensities for leakage from the donor channels
total_acceptor_corrected(1:n_peaks-1,1:n_fr)= total_acceptor(1:n_peaks-
1,1:n_fr)-0.17*total_donor(1:n_peaks-1,1:n_fr);

for i = 1:n_peaks-1
    acceptor_intensity = mean (total_acceptor_corrected (i,red_on:n_fr));
```

```

if acceptor_intensity > red_threshold; % only molecules that have acceptor
signal
    dyesum = total_donor(i,1:n_fr_analysis) +
        total_acceptor_corrected(i,1:n_fr_analysis);
    FRET = total_acceptor_corrected(i,1:n_fr_analysis)./dyesum;
    [rows] = find(dyesum <fretthreshold);
    FRET(rows) = [];
    total_frets = [total_frets mean(FRET)];

%remove FRET below 0.1 and above 1.1:
    [rowsoutsidelow] = find(total_frets <-0.1);
    total_frets(rowsoutsidelow) = [];
    [rowsoutsidehigh] = find(total_frets >1.1);
    total_frets(rowsoutsidehigh) = [];
end
end
end

%this creates the vector for binning into 0.025 FRET steps
bins = -0.1:0.025:1.125;
bins = bins';
% and this makes the vector for exporting the histogram figures (to Origin)
binned_data = hist(total_frets,bins);
binned_data = binned_data';
% finally, this gives the raw data in a vector for export
export_frets = total_frets';
figure
hist(total_frets,bins)
% plot A
title(fileName)

```

6.7.2 Selection of movie traces

Script name: readtracesandor

```

%Program to read binary traces file output by Hazens IDL program written by
%Michael Stone 01/17/2004
clear all;
accepted_traces = []; % an array to hold good molecule data for future
analysis
fretthreshold = 2000; % this is a threshold value to prevent crazy fret traces

%First read in file from directory
[fileName,path1] = uigetfile('*.traces','Read Traces File');
addpath(path1) %add path to path list each time to ensure proper file access
FID = fopen(fileName);
%find number of frames and peaks from traces header file
[n_fr,z1] = fread(FID,1,'int32'); %z1 is just the number of indicated
data types read (should be 1)
[n_traces,z2] = fread(FID,1,'int16');
rate_fr = input('At what frame rate was the data collected?');
n_peaks = n_traces/2;
exp_length = n_fr/rate_fr; %(seconds)
time = (1/rate_fr:(1/rate_fr):exp_length);

```

Appendix

```
%% Now read in the rest of the data from the traces file
[total_data,z3] = fread(FID,[n_traces+1,n_fr],'int32');
fclose(FID);
total_data = total_data(2:n_traces,1:n_fr);
total_donor = [];
total_acceptor = [];
total_donor(1:n_peaks-1,1:n_fr) = total_data(2*(0:n_peaks-2)+1,1:n_fr);
total_acceptor(1:n_peaks-1,1:n_fr) = total_data(2*(0:n_peaks-2)+2,1:n_fr);

%correct the acceptor intensities for leakage from the donor channels
total_acceptor_corrected(1:n_peaks-1,1:n_fr) = total_acceptor(1:n_peaks-1,1:n_fr) - 0.12*total_donor(1:n_peaks-1,1:n_fr);

i = input('What molecule would you like start on?');
h = figure('Position', [219 28 1015 689]);

%this while loop will continue for all traces
while i < n_peaks
    set(h,'Name',strcat('Molecule #',num2str(i), '/', num2str(n_peaks)));
    subplot(2,1,1)
    plot(time, total_donor(i,:), 'g');
    xlabel('Time, seconds')
    ylabel('Intensity')
    hold on
    plot(time, total_acceptor_corrected(i,:), 'r');
    ys = [0 max(total_donor(i,:))];
    xs = [10 10];
    plot(xs,ys, 'LineWidth', 2);
    hold off
    subplot(2,1,2)
    dyesum = total_donor(i,:)+total_acceptor_corrected(i,:);
    [rows] = find(dyesum < fretthreshold);
    FRET = total_acceptor_corrected(i,:)./dyesum;
    FRET(rows) = 0;
    plot(time, FRET)
    xlabel('Time, seconds')
    ylabel('FRET')
    ylim([-0.5 1.2])
    B = menu('analyze this molecule?', 'yes', 'next', 'previous', 'end
        program');
    if B == 1;
        A = menu('Which dye molecule bleached', 'acceptor', 'both', 'neither');
        if A == 1;
            [x,y] = ginput(2);
            x1 = floor(x(1))*rate_fr;
            x2 = floor(x(2))*rate_fr;
            norm = mean(total_acceptor_corrected(i,x1:x2)) - 0;
            acceptor_norm = total_acceptor_corrected(i,:)-norm;
            subplot(2,1,1)
            plot(time, total_donor(i,:), 'g');
            hold on
            xlabel('Time, seconds')
            ylabel('Intensity')
            plot(time, acceptor_norm, 'r');
            hold off
            subplot(2,1,2)
            dyesum = total_donor(i,:)+acceptor_norm;
            [rows] = find(dyesum < fretthreshold);
```



```

FRET = acceptor_norm./dyesum;
FRET(rows)=0;
plot(time,FRET)
xlabel('Time, seconds')
ylabel('FRET')
ylim([-0.5 1.2])
C = menu('Accept this trace', 'yes', 'no');
if C == 1;
    accepted_traces = [accepted_traces total_donor(i,:) ' acceptor_norm'];
end
elseif A==2;
    [x,y] = ginput(2);
    x1 = floor(x(1))*rate_fr;
    x2 = floor(x(2))*rate_fr;
    norm = mean(total_acceptor_corrected(i,x1:x2)) - 0;
    acceptor_norm = total_acceptor_corrected(i,:)-norm;
    norm = mean(total_donor(i,x1:x2)) - 0;
    donor_norm = total_donor(i,:)-norm;
    subplot(2,1,1)
    plot(time, donor_norm, 'g');
    xlabel('Time, seconds')
    ylabel('Intensity')
    hold on
    plot(time, acceptor_norm, 'r');
    hold off
    subplot(2,1,2)
    dyesum = donor_norm+acceptor_norm;
    [rows] = find(dyesum < fretthreshold);
    FRET = acceptor_norm./dyesum;
    FRET(rows) = 0;
    plot(time,FRET)
    xlabel('Time, seconds')
    ylabel('FRET')
    ylim([-0.5 1.2])
    C = menu('Accept this trace', 'yes', 'no');
    if C == 1;
        accepted_traces = [accepted_traces donor_norm ' acceptor_norm'];
    end
end
elseif A==3;
    C = menu('Accept this trace', 'yes', 'no');
    if C == 1
        accepted_traces = [accepted_traces total_donor(i,:) '
            total_acceptor_corrected(i,:)'];
    end
end
end
i = i+1;
elseif B == 2;
    i = i+1;
elseif B == 3;
    i = i-1;
elseif B == 4;
    break
end
end
close(h);

%save all data in workspace for later manipulation
saveFile = strcat(fileName, '.mat');
save(saveFile);

```

6.7.3 Export movie traces for HaMMMy

Script name: GetHistFromTracesAndExportHaMMMy_single_stitched_direct

```
[fileName2,path1] = uigetfile('*.mat','Read Traces File');
addpath(path1) %add path to path list each time to ensure proper file access
FID = fopen(fileName2);

load(fileName2, 'accepted_traces', 'time', 'exp_length')

data = accepted_traces;%rename matrix as 'data'
num_tr = size(data,2)/2;
runnum = 1;

fretthreshold = 1500; % a variable that should be just above sum of background
of D and A channel

% open variables for saving different stuff
fretvalues = [];
hammy_std= [];
start_time=0;
hammy_all_traces =[];
len_traces=[];

for runnum=1:num_tr
    index = runnum*2;
    donor = data(:,index-1);
    accep = data(:,index);
    aveDonor = medfilt1(donor',15);
    aveAccep = medfilt1(accep',15);
    dyesum = aveAccep+aveDonor;

    [rows] = find(dyesum <fretthreshold);
    FRET = aveAccep./dyesum;
    FRET(rows) = 0;
    % identify the end of the part of the trace that is suitable for HaMMMy
    [hammy_rows] = find(FRET <0.1);
    trace_end = min(hammy_rows)-1-5;
    if trace_end > 0
        inv_donor = donor';
        inv_accep = accep';
        %check for negative non mean FRETs in last 20 data points
        hammy_fret=[];
        for i=(trace_end-19):(trace_end)
            hammy_fret= [hammy_fret
inv_accep(i)/(inv_donor(i)+inv_accep(i))];
        end
        [hammy_rows_2] = find(hammy_fret < 0.1);
        if isempty(hammy_rows_2)
            trace_end=trace_end;
        else
            trace_end = trace_end-(21-min(hammy_rows_2));
        end
        len_traces=[len_traces; trace_end];
        hammy_trace = [time(1:trace_end); inv_donor(1:trace_end);
inv_accep(1:trace_end)];
        hammy_traceexport = hammy_trace';

        %std of donor and accpetor in Hammy_trace
        thistrace_std = std(hammy_traceexport,1);
        hammy_std=[hammy_std; thistrace_std(:,2:3)];

        %..and save the trace in a HaMMMy.compatible format
        file3=strcat('molecule_',num2str(runnum),'.dat');
        save (file3, 'hammy_traceexport', '-ascii', '-tabs');
```

```

        %prepare to save also in stitched format
        add_time=hammy_traceexport(:,1)+ start_time;
        add_trace=[add_time hammy_traceexport(:,2:3)];
        hammy_all_traces = [hammy_all_traces; add_trace];
        start_time=hammy_traceexport(trace_end,1)+ start_time;

        %prepare to save mean FRET as binned data
        fretvalues_thistrace = FRET(1:trace_end);
        fretvalues = [fretvalues; fretvalues_thistrace'];
    else
        continue
    end
    %...and continue with next trace
    runnum = runnum+1;
end

%bin the std data in 11 bins:
bins_std=400:200:2600;
binned_std=[bins_std(1:11)' histcounts(hammy_std, bins_std)'];

%bin the mean fretvalues
bins_FRET=(-0.1):0.025:1.125;
binned_FRET=[bins_FRET(1:49)' histcounts(fretvalues, bins_FRET)'];

%calculate the mean trace length
mean(len_traces)

% save combined data for:
fileName = input('Enter file name?','s') % ask for filename

%save all mean FRET values of full length
file2=strcat(num2str(fileName),'_fret.txt');
save (file2, 'binned_FRET', '-ascii');

%save std of donor and acceptor per molecules
file_std=strcat(num2str(fileName),'_std.txt');
save (file_std, 'binned_std', '-ascii');

%save hammy files als stiched format
file_all=strcat(num2str(fileName),'_all_traces.dat');
save (file_all, 'hammy_all_traces', '-ascii', '-tabs');

clearvars -except accepted_traces len_traces

```

6.7.4 Analysis of HaMMY identified states

Script name: combine_all_reports_afterHaMMY_and_analyze

```

% program to analyze single HaMMY reports of all molecules of same condition
% written by Nathalie Meiser 12/2022
%% open all reports and save in one matrix
s = pwd;
files = dir(s);
s = strcat(s,'\');

% Format for each line of text:
formatSpec = '%15f%15f%15f%15f%f%[\n\r]';
startRow = 5;

% some variables to store
n=1;
add_filenames=[];
all_molecule_reports=[];

```

Appendix

```
for i = 3:length(files)
    % Initialize variables.
    file = files(i).name;
    filename = strcat(s,file);

    % Open the text file.
    fileID = fopen(filename,'r');

    % Read columns of data according to the format.
    dataArray = textscan(fileID, formatSpec, 'Delimiter', ',', 'WhiteSpace', '',
    'TextType', 'string', 'HeaderLines', startRow-1, 'ReturnOnError', false,
    'EndOfLine', '\r\n');
    % Create output variable
    molecule_report = cell2mat(dataArray(:,1:5));
    % add filename to table
    for r= 1:length(molecule_report)
        add_filenames=[add_filenames; i];
    end
    fclose(fileID);

    % combine all reports.
    all_molecule_reports = [all_molecule_reports; molecule_report];
    n=n+1;
end

num_mol=n;
% add fileIDS
all_molecule_reports=[all_molecule_reports add_filenames];

%% bin molecules for the fitted FRET values
bins=(-0.1):0.05:1.1;
binned_FRET=[bins(1:24)' histcounts(all_molecule_reports(:,1), bins)'];
%histogram(all_molecule_reports(:,1), bins)

%% search only for transitions and bin
[no_transitions]=find(all_molecule_reports(:,5));
transitions=all_molecule_reports(no_transitions,1:6);
bins = [-2 0.3 0.44 0.51 0.6 0.8 2]; %change this to fit to your states

% bin for starting state FRET value
bins_trans= discretize(transitions(:,1), bins);
bins_trans(find(isnan(bins_trans)))=1;
[actual_bins]=bins(bins_trans)';
transitions=[actual_bins transitions];

% bin for ending FRET value
bins_trans= discretize(transitions(:,3), bins);
bins_trans(find(isnan(bins_trans)))=1;
[actual_bins]=bins(bins_trans)';
transitions=[transitions(:,1) actual_bins transitions(:,2:7)];
transitions=sortrows(transitions);

% replace I (fileID) with actual molecule name
transitions_char={transitions(:,1) transitions(:,2) transitions(:,3)
transitions(:,4) transitions(:,7) transitions(:,8)};
for i=1:length(transitions(:,8))
    transitions_char{7}(i,1)={files(transitions(i,8)).name};
end

%% compute total number of transitions, transition between states and add
molecule name
[number, transitioning_molecules]=groupcounts(transitions(:,8));
all=length(transitioning_molecules);
disp([num2str(all), ' of ', num2str(num_mol), ' molecules are transitioning'])
total_transitions=[];
molecule_names={};
for i=1:length(number(:,1))
```

```

    A=transitions(:,8)==transitioning_molecules(i);
    total_transitions(i,1) = sum(transitions(A,7));
    molecule_names{1}(i,1)={files(transitioning_molecules(i)).name};
end

%% pool for relevant transitions
rel_trans=[];
for i=2:5 %change to match relevant states in bin vector
    for n=2:5
        if n==i
            continue
        else
            relevant=find(transitions(:,1)==bins(i)
transitions(:,2)==bins(n));
            rel_trans=[rel_trans; relevant];
        end
    end
end
[number_rel,
transitioning_molecules_rel]=groupcounts(transitions(rel_trans,8));
all=length(transitioning_molecules_rel(:,1));
disp([num2str(all),' of ',num2str(num_mol),' molecules are transitioning in
relevant states'])
total_transitions_rel=[];
molecule_names_rel={};
for i=1:length(number_rel(:,1))
    A=transitions(rel_trans,8)==transitioning_molecules_rel(i);
    total_transitions_rel(i,1) = sum(transitions(rel_trans(A),7));
    molecule_names_rel{1}(i,1)={files(transitioning_molecules_rel(i)).name};
end

%% make result tables
transitions=table(transitions_char{1:5},transitions_char{7}, 'VariableNames',
{'Assigned FRET bin 1', 'Assigned FRET bin 2', 'State1', 'State2', 'Transitions',
'MoleculeName'});
total_transitions=table(molecule_names{1,1}(:,1), number, total_transitions,
'VariableNames', {'MoleculeName', 'State Transitions', 'Total Transitions'});
relevant_transitions=table(transitions_char{1}(rel_trans,:),
transitions_char{2}(rel_trans,:),
transitions_char{3}(rel_trans,:),transitions_char{4}(rel_trans,:),
transitions_char{5}(rel_trans,:),transitions_char{7}(rel_trans,:),
'VariableNames',{'Assigned FRET bin 1','Assigned FRET bin 2', 'State1',
'State2', 'Transitions', 'MoleculeName'});
relevant_total_transitions=table(molecule_names_rel{1,1}(:,1), number_rel,
total_transitions_rel, 'VariableNames', {'MoleculeName', 'State Transitions',
'Total Transitions'});

%% save as matlab workspace
file_mat = input('Enter file name?','s'); % ask for filename
file_mat= strcat(num2str(file_mat),'_all_reports.mat');
save(file_mat,
'transitions','total_transitions', 'bins',
'relevant_transitions', 'relevant_total_transitions');

%% clear all variables not needed
clearvars -except transitions total_transitions bins relevant_transitions
relevant_total_transitions

```

6.7.5 Analysis of HaMMY dwell times

Script name: combine_all_dwell_afterHaMMY_and_analyze

```
% program to analyze single HaMMY dwell times of all molecules of same condition
% written by Nathalie Meiser 12/2022
%% Import data from text file to combine dwells.
s = pwd;
files = dir(fullfile(s,'*.dat'));
s = strcat(s,'\');

% Format for each line of text:
formatSpec = '%10f%10f%f%[\n\r]';
startRow = 5;

% some variables to store
add_filenames=[];
all_molecule_dwells=[];

for i = 3:length(files)
    % Initialize variables.
    file = files(i).name;
    filename = strcat(s,file);

    % Open the text file.
    fileID = fopen(filename,'r');

    % Read columns of data according to the format.
    dataArray = textscan(fileID, formatSpec, 'Delimiter', ',', 'WhiteSpace', '',
        'TextType', 'string', 'ReturnOnError', false);

    % Create output variable
    molecule_dwell = cell2mat(dataArray(:,1:3));

    fclose(fileID);

    % combine all dwells.
    all_molecule_dwells = [all_molecule_dwells; molecule_dwell];

end

%% assign FRET values from report
bins = [-2 0.3 0.44 0.51 0.6 0.8 2]; %change this to fit to your states

% bin for starting state FRET value
bins_dwell= discretize(all_molecule_dwells(:,2), bins);
bins_dwell(find(isnan(bins_dwell)))=1;
[actual_bins]=bins(bins_dwell)';
assigned_dwells=[actual_bins all_molecule_dwells];

% bin for ending FRET value
bins_dwell= discretize(all_molecule_dwells(:,1), bins);
bins_dwell(find(isnan(bins_dwell)))=1;
[actual_bins]=bins(bins_dwell)';
assigned_dwells=[actual_bins assigned_dwells];
assigned_dwells=sortrows(assigned_dwells);
```

```

%% bin for dwell times in discrete transitions
bins_dwells= 1:2:600; % change this to make different binning
binned_dwelltimes=[0 0 bins_dwells(1:end-1)];

for i=2:5 %1:length(bins) %loop through all individual transitions and bin for
dwelltimes
    for n=2:5 %1:length(bins)
        if i==n
            continue
        else
            trans=[bins(i); bins(n)];
            relevant=find(assigned_dwells(:,1)==bins(i) &
assigned_dwells(:,2)==bins(n));
            binned_rel=histcounts(assigned_dwells(relevant,5), bins_dwells);
            %rel_dwells=[bins_dwells(1:end-1); binned_rel];
            rel_dwells=[trans; binned_rel'];
            binned_dwelltimes=[binned_dwelltimes rel_dwells];
        end
    end
end

%% save as workspace
file_mat = input('Enter file name?','s'); % ask for filename
file_mat= strcat(num2str(file_mat),'_all_dwells.mat');
save(file_mat, 'bins', 'assigned_dwells','binned_dwelltimes');

%% clear all variables not needed
clearvars -except bins assigned_dwells binned_dwelltimes

```

6.7.6 Population analysis of dynamic molecules

Script name: histogram_dynamics_afterHaMMMy_frompath

```

% program to determine populations and make histograms of static vs dynamic
molecules of same condition from "path" calculated by Hammy (fitted and not
fitted)
% written by Nathalie Meiser 12/2022

[fileName2,path1] = uigetfile('*.mat','Read MATLAB File');
addpath(path1) %add path to path list each time to ensure proper file access
FID = fopen(fileName2);
load(fileName2, 'relevant_total_transitions');
s = pwd;
files = dir(fullfile(s,'*.dat'));
s = strcat(s,'\');
last_size=1;
all_paths_forhist=[];
all_paths=[];

%load list of all transitioning molecules
relevant_names=relevant_total_transitions.MoleculeName;
found_names={};
trans_paths_forhist=[];
trans_fit=[];

for i = 1:length(files)
    fileName = files(i).name;
    fullFileName = strcat(s,fileName);

```

Appendix

```
formatSpec = '%15f%15f%15f%15f%f%[\n\r]';
fileID = fopen(fullFileName,'r');
searching=strcat(fileName(1:end-8),'report.dat');
isthere=strfind(relevant_names,searching);
checkCells=sum(cellfun('isempty', isthere));
%% Read columns of data according to the format.
dataArray = textscan(fileID, formatSpec, 'Delimiter', '', 'WhiteSpace',
'', 'TextType', 'string', 'ReturnOnError', false);
fclose(fileID);
%% Create output variable for dynamic histograms
molecule = table(dataArray{1:end-1}, 'VariableNames',
{'number','donor','acceptor','FRET','fitted_path'});
molecule_path_forhist=mean(molecule.FRET(1:20));
molecule_fit_forhist=mean(molecule.fitted_path(1:20));
extend_data_forhist=[molecule_path_forhist molecule_fit_forhist];
all_paths_forhist=[all_paths_forhist; extend_data_forhist];
%% Create output variable for all dynamic fitted paths for popcount
molecule_fit=molecule.fitted_path(1:end);
all_paths=[all_paths; molecule_fit];
if checkCells==length(isthere)
    continue
else
    trans_paths_forhist=[trans_paths_forhist; extend_data_forhist];
    trans_fit=[trans_fit; molecule_fit];
end
end
%% make histograms from first 20 frames of data
bins=(-0.1):0.05:1.1;
all_binned_FRET_path=[histcounts(all_paths_forhist(:,1), bins)'];
trans_binned_FRET_path=[histcounts(trans_paths_forhist(:,1), bins)'];
all_binned_FRET_fit=[histcounts(all_paths_forhist(:,2), bins)'];
trans_binned_FRET_fit=[histcounts(trans_paths_forhist(:,2), bins)'];
diff_path=all_binned_FRET_path(:,1)-trans_binned_FRET_path(:,1);
diff_fit=all_binned_FRET_fit(:,1)-trans_binned_FRET_fit(:,1);
binned_data=table(bins(1:end-1)', all_binned_FRET_path,
trans_binned_FRET_path, diff_path, all_binned_FRET_fit, trans_binned_FRET_fit,
diff_fit,'VariableNames',
{'bin','all_path','trans_path','diff_path','all_fit',
'trans_fit','diff_fit'});
%% make population distribution for fitted states
% first bin into same bins like transitions
bins = [-1 0 0.3 0.44 0.51 0.6 0.8 1]; %change this to fit to your states
% bin for FRET value
bins_fit= discretize(trans_fit(1:end), bins);
bins_fit(find(isnan(bins_fit)))=1;
[actual_bins]=bins(bins_fit)';
assigned_fits=[actual_bins trans_fit];
% take out only relevant states
frames=0;
pop_dist=[];
for i=3:6
    relevant=find(assigned_fits(:,1)==bins(i));
    pop_frames=assigned_fits(relevant,1);
    popcount=length(assigned_fits(relevant,1));
    frames=frames+length(pop_frames);
    extend=[bins(i); popcount];
    pop_dist=[pop_dist extend];
end
norm_pop=[pop_dist(1,1:end); pop_dist(2,1:end)/frames];
% save as matlab workspace
file_mat = input('Enter file name?','s'); % ask for filename
file_mat= strcat(num2str(file_mat),'_all_paths.mat');
save(file_mat, 'all_paths_forhist', 'bins', 'binned_data', 'norm_pop',
'pop_dist','frames');
```

```
clearvars -except norm_pop pop_dist frames binned_data trans_paths
all_paths_forhist bins
```


6.8 smFRET data

6.8.1 EMSA of the smFRET construct

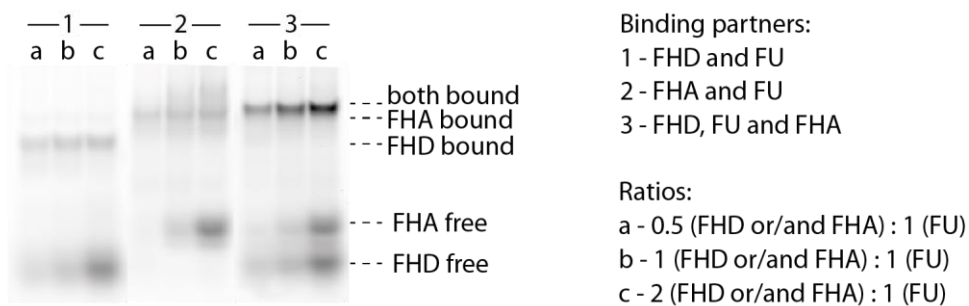


Figure 87: Electrophoretic mobility shift assay of the substrate FRET RNA with the two DNA handles, each coupled to a fluorophore. The gel was scanned for fluorescence (free RNA not visible).

6.8.2 Transition density plots

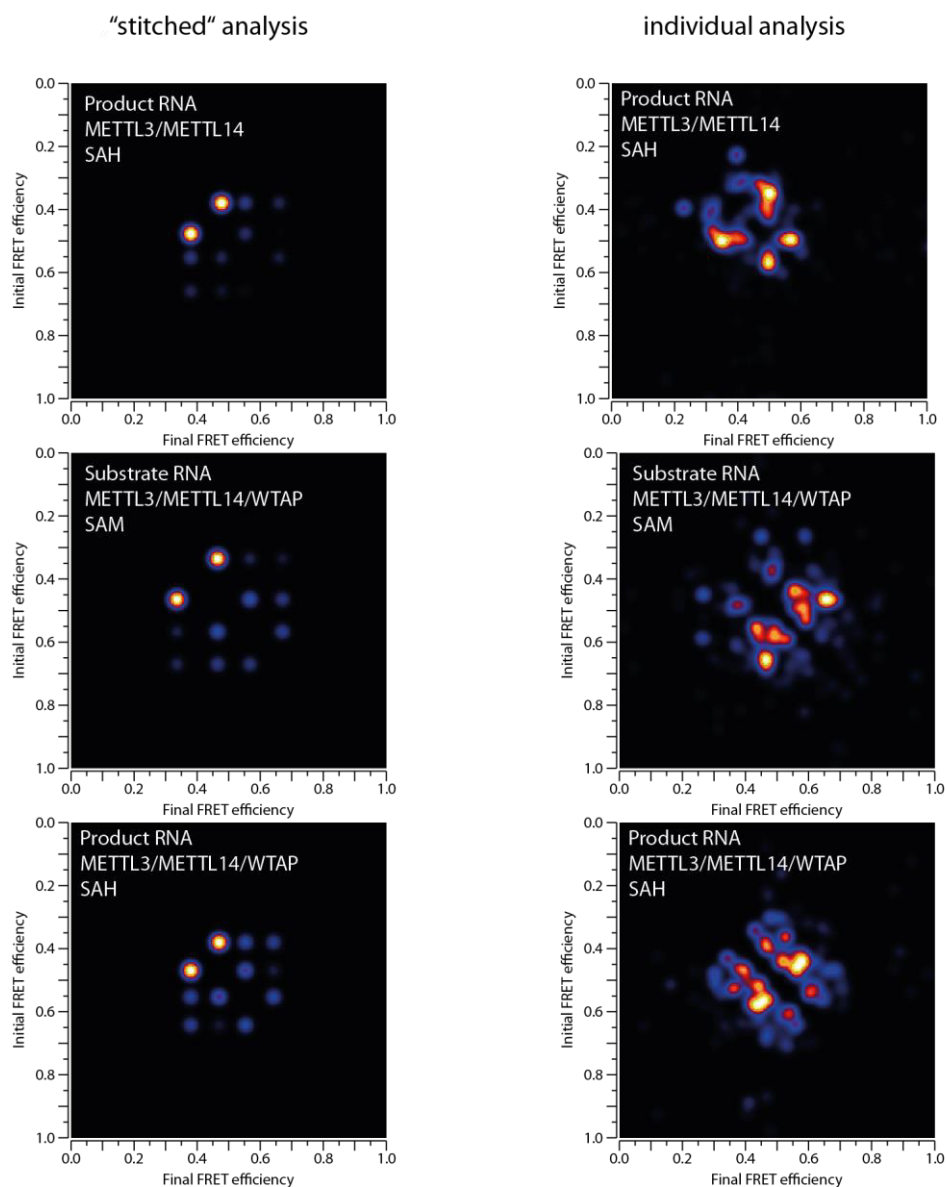


Figure 88: Transition density plots for different measurement conditions. The spots are weighted by number of transitions.

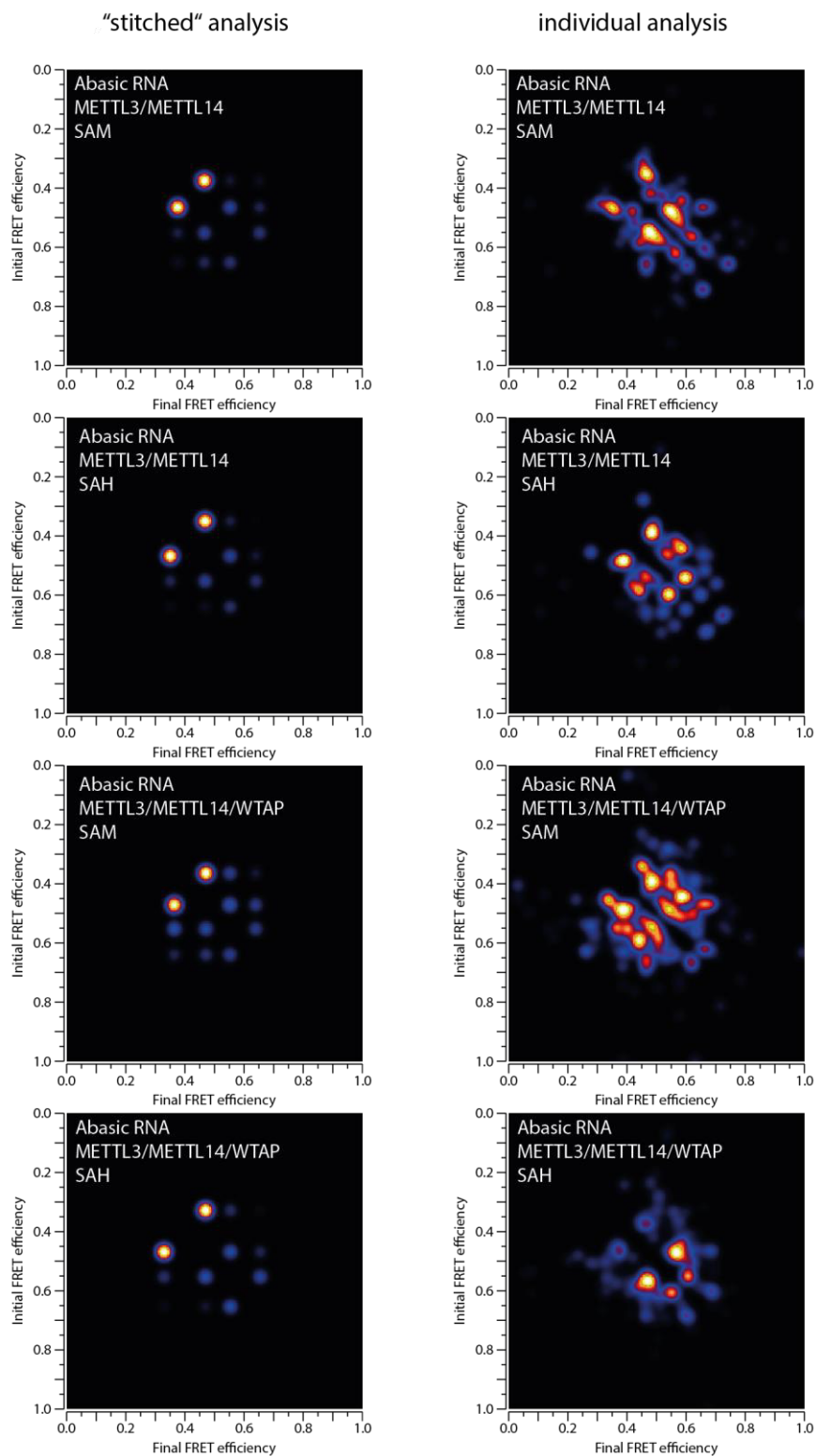


Figure 89: Transition density plots for different measurement conditions with the abasic RNA. The spots are weighted by number of transitions.

6.8.3 Movie analysis by individual transitions

Table 25: METTL3/METTL14 with SAH on product RNA. Summarized analysis of movie data with calculated indication values for all observed transitions. The median of transitions per molecule are given with the covered range (1-max). Transitioning molecules: 78 of 196 analyzed molecules (40%). Total number of detected transitions: 982. Mean trace length: 20.6 s.

Transition (from ... to ...)		Total molecules	Percentage of molecules	Total transitions	Percentage of transitions	Transitions per molecule (max)
0.48	0.38	30	15	285	29	7 (26)
0.38	0.48	26	13	285	29	10 (26)
0.48	0.56	24	12	152	15	5 (24)
0.56	0.48	24	12	140	14	4 (24)
0.56	0.38	13	7	34	3	1 (8)
0.38	0.56	11	6	36	4	2 (9)
0.56	0.66	8	4	16	2	1 (9)
0.66	0.38	6	3	8	1	1 (3)
0.66	0.48	5	3	5	1	1 (1)
0.66	0.56	4	2	12	1	1 (9)
0.38	0.66	3	2	6	1	1 (4)
0.48	0.66	3	2	3	0	1 (1)

Table 26: METTL3/METTL14/WTAP with SAM on substrate RNA. Summarized analysis of movie data with calculated indication values for all observed transitions. The median of transitions per molecule are given with the covered range (1-max). Transitioning molecules: 85 of 206 analyzed molecules (41%). Total number of detected transitions: 652. Mean trace length: 16.2 s.

Transition (from ... to ...)		Total molecules	Percentage of molecules	Total transitions	Percentage of transitions	Transitions per molecule (max)
0.56	0.48	26	13	100	15	3 (13)
0.48	0.56	24	12	99	15	3 (14)
0.48	0.66	24	12	98	15	3 (18)
0.66	0.48	22	11	92	14	3 (17)
0.66	0.56	17	8	27	4	1 (6)
0.56	0.66	16	8	26	4	1 (6)
0.39	0.66	12	6	22	3	2 (4)
0.48	0.39	10	5	44	7	4 (10)
0.66	0.39	10	5	18	3	2 (3)
0.39	0.48	9	4	45	7	5 (10)
0.56	0.39	8	4	42	6	2 (21)
0.39	0.56	6	3	39	6	2 (21)

Table 27: METTL3/METTL14/WTAP with SAH on product RNA. Summarized analysis of movie data with calculated indication values for all observed transitions. The median of transitions per molecule are given with the covered range (1-max). Transitioning molecules: 62 of 197 analyzed molecules (31%). Total number of detected transitions: 635. Mean trace length: 17.2 s.

Transition (from ... to ...)		Total molecules	Percentage of molecules	Total transitions	Percentage of transitions	Transitions per molecule (max)
0.56	0.66	17	9	51	8	2 (14)
0.66	0.56	16	8	49	8	2 (14)
0.38	0.56	15	8	73	11	2 (22)
0.48	0.66	15	8	29	5	1 (6)
0.48	0.56	14	7	102	16	5 (24)
0.56	0.38	14	7	73	11	4 (22)
0.56	0.48	14	7	96	15	4 (24)
0.66	0.48	10	5	25	4	2 (6)
0.66	0.38	9	5	9	1	1 (1)
0.38	0.48	8	4	61	10	6 (19)
0.48	0.38	8	4	60	9	7 (19)
0.38	0.66	5	3	7	1	1 (2)

Table 28: METTL3/METTL14 with SAM on abasic RNA. Summarized analysis of movie data with calculated indication values for all observed transitions. The median of transitions per molecule are given with the covered range (1-max). Transitioning molecules: 41 of 209 analyzed molecules (20%). Total number of detected transitions: 542. Mean trace length: 17.5 s.

Transition (from ... to ...)		Total molecules	Percentage of molecules	Total transitions	Percentage of transitions	Transitions per molecule (max)
0.56	0.48	16	8	104	19	5 (22)
0.48	0.56	14	7	105	19	6 (22)
0.38	0.48	10	5	96	18	7 (27)
0.48	0.38	9	4	93	17	9 (27)
0.48	0.66	9	4	24	4	3 (4)
0.66	0.48	9	4	22	4	2 (4)
0.56	0.66	6	3	28	5	2 (20)
0.66	0.56	6	3	29	5	2 (21)
0.38	0.56	4	2	20	4	4 (12)
0.56	0.38	4	2	21	4	4 (11)

Table 29: METTL3/METTL14 with SAH on abasic RNA. Summarized analysis of movie data with calculated indication values for all observed transitions. The median of transitions per molecule are given with the covered range (1-max). Transitioning molecules: 41 of 210 analyzed molecules (20%). Total number of detected transitions: 454. Mean trace length: 18.5 s.

Transition (from ... to ...)		Total molecules	Percentage of molecules	Total transitions	Percentage of transitions	Transitions per molecule (max)
0.48	0.56	16	8	73	16	5 (8)
0.56	0.48	16	8	66	15	5 (7)
0.56	0.66	11	5	62	14	6 (13)
0.66	0.56	9	4	59	13	6 (14)
0.38	0.48	5	2	61	13	12 (20)
0.48	0.38	5	2	61	13	12 (20)
0.66	0.48	5	2	10	2	2 (4)
0.48	0.66	4	2	11	2	2 (5)
0.38	0.56	3	1	24	5	7 (16)
0.56	0.38	3	1	22	5	6 (15)
0.66	0.38	3	1	3	1	1 (1)
0.38	0.66	2	1	2	0	1 (1)

Table 30: METTL3/METTL14/WTAP with SAM on abasic RNA. Summarized analysis of movie data with calculated indication values for all observed transitions. The median of transitions per molecule are given with the covered range (1-max). Transitioning molecules: 84 of 216 analyzed molecules (39%). Total number of detected transitions: 856. Mean trace length: 20.6 s.

Transition (from ... to ...)		Total molecules	Percentage of molecules	Total transitions	Percentage of transitions	Transitions per molecule (max)
0.48	0.66	22	10	84	10	2 (16)
0.56	0.48	21	10	113	13	4 (26)
0.48	0.56	19	9	116	14	4 (26)
0.38	0.56	17	8	76	9	2 (21)
0.66	0.48	17	8	66	8	2 (16)
0.48	0.38	16	7	120	14	6 (29)
0.56	0.38	14	6	69	8	3 (20)
0.38	0.48	13	6	115	13	7 (30)
0.66	0.56	13	6	29	3	1 (6)
0.66	0.38	10	5	22	3	2 (7)
0.38	0.66	9	4	20	2	1 (9)
0.56	0.66	8	4	26	3	3 (6)

Table 31: METTL3/METTL14/WTAP with SAH on abasic RNA. Summarized analysis of movie data with calculated indication values for all observed transitions. The median of transitions per molecule are given with the covered range (1-max). Transitioning molecules: 74 of 210 analyzed molecules (35%). Total number of detected transitions: 894. Mean trace length: 21.3 s.

Transition (from ... to ...)		Total molecules	Percentage of molecules	Total transitions	Percentage of transitions	Transitions per molecule (max)
0.66	0.48	20	10	76	9	3 (11)
0.48	0.66	18	9	78	9	4 (11)
0.56	0.38	17	8	41	5	1 (9)
0.56	0.66	16	8	94	11	3 (21)
0.48	0.56	15	7	159	18	2 (42)
0.66	0.56	15	7	96	11	4 (21)
0.56	0.48	14	7	154	17	7 (42)
0.38	0.56	13	6	40	4	2 (9)
0.38	0.48	11	5	67	7	6 (15)
0.48	0.38	9	4	65	7	7 (14)
0.66	0.38	6	3	12	1	1 (4)
0.38	0.66	4	2	12	1	3 (5)

6.8.4 Exponential decay fits for dwell time distribution

The following figures (Figure 90 to Figure 95) show the binned dwell times that were successfully fitted with a single exponential decay function (red line) for different complex compositions. The transition and the determined transition rates are indicated respectively.

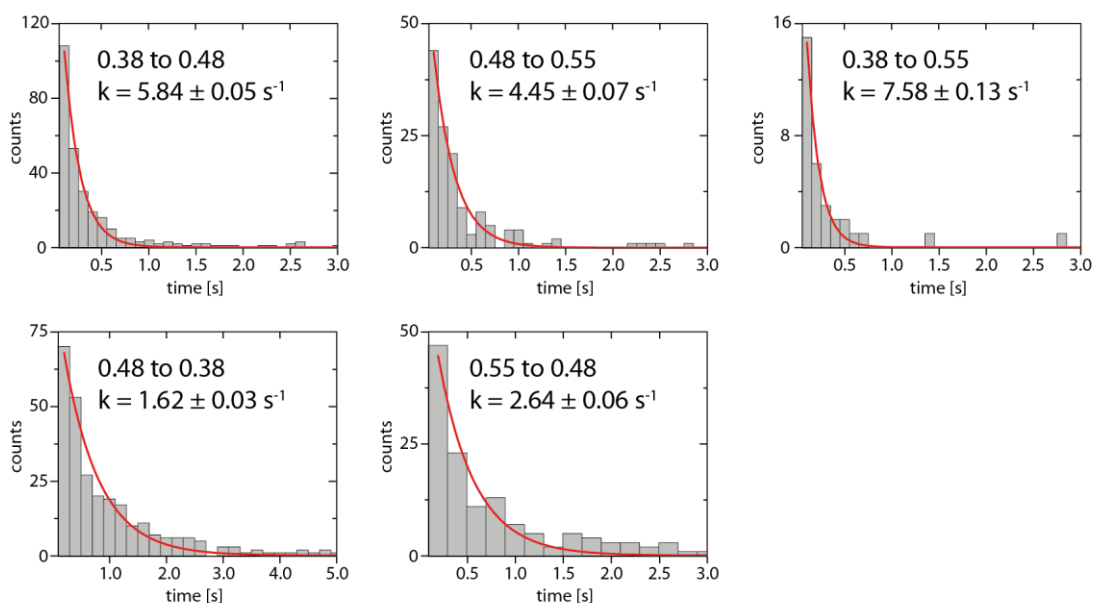


Figure 90: Exponential decay fits for the product RNA with complex METTL3/METTL14 and SAH.

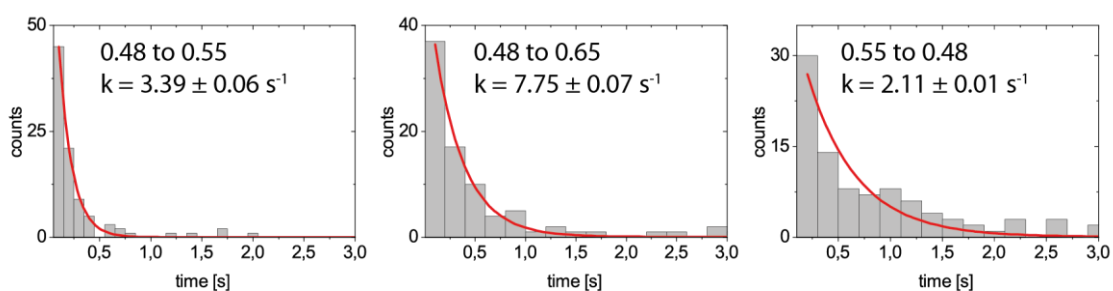


Figure 91: Exponential decay fits for the substrate RNA with complex METTL3/METTL14/WTAP and SAM.

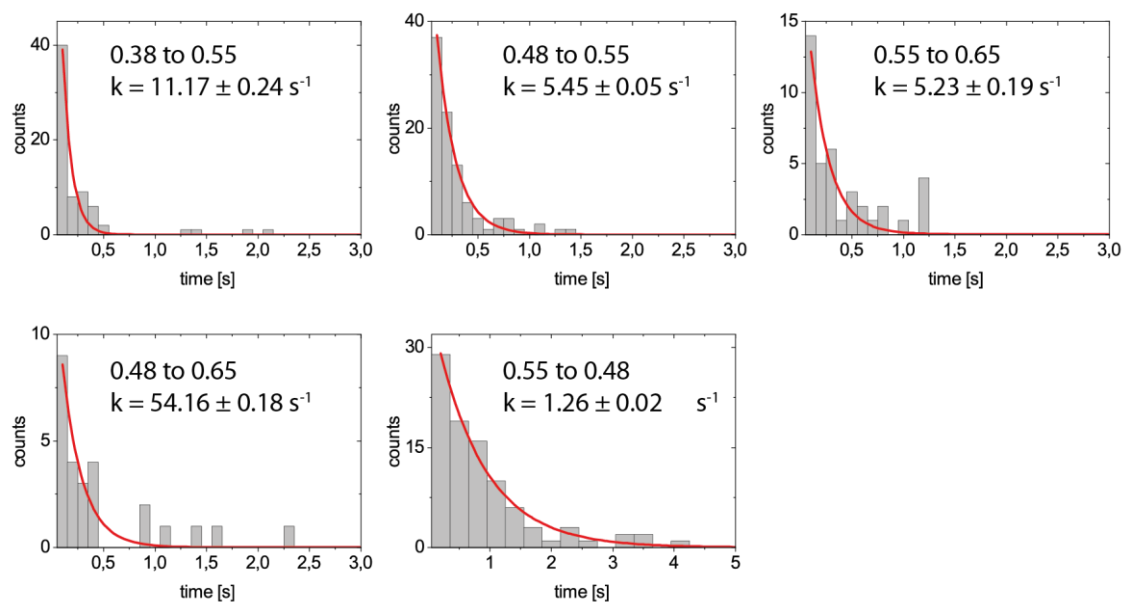


Figure 92: Exponential decay fits for the product RNA with complex METTL3/METTL14/WTAP and SAH.

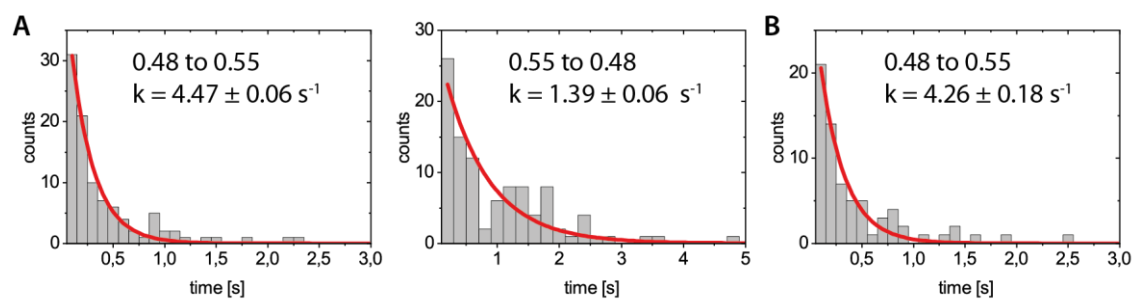


Figure 93: Exponential decay fits for the abasic RNA with complex METTL3/METTL14 and SAM (A) or SAH (B).

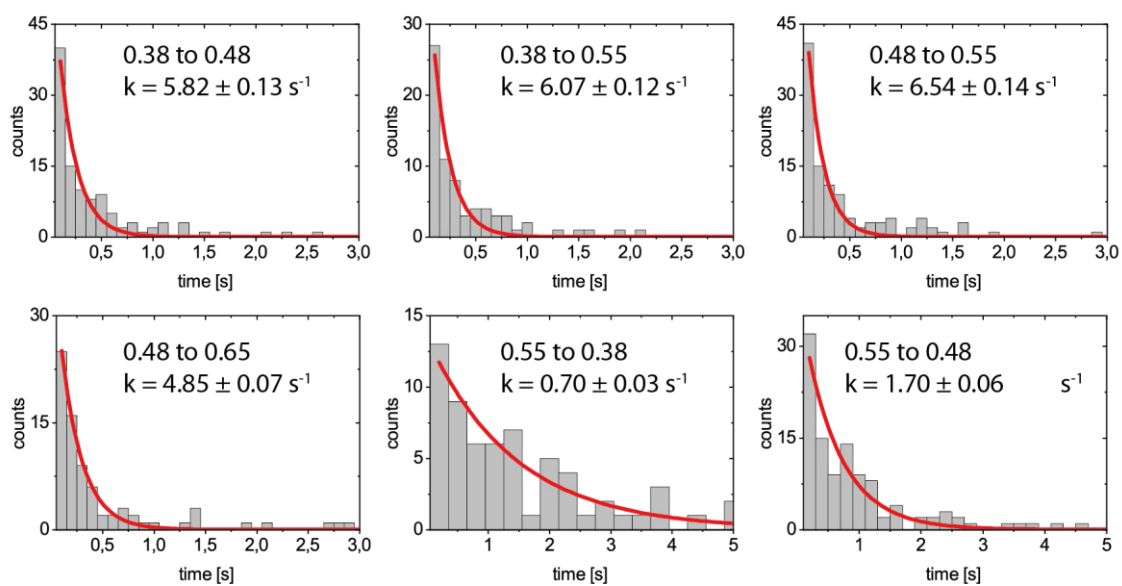


Figure 94: Exponential decay fits for the abasic RNA with complex METTL3/METTL14/WTAP and SAM.

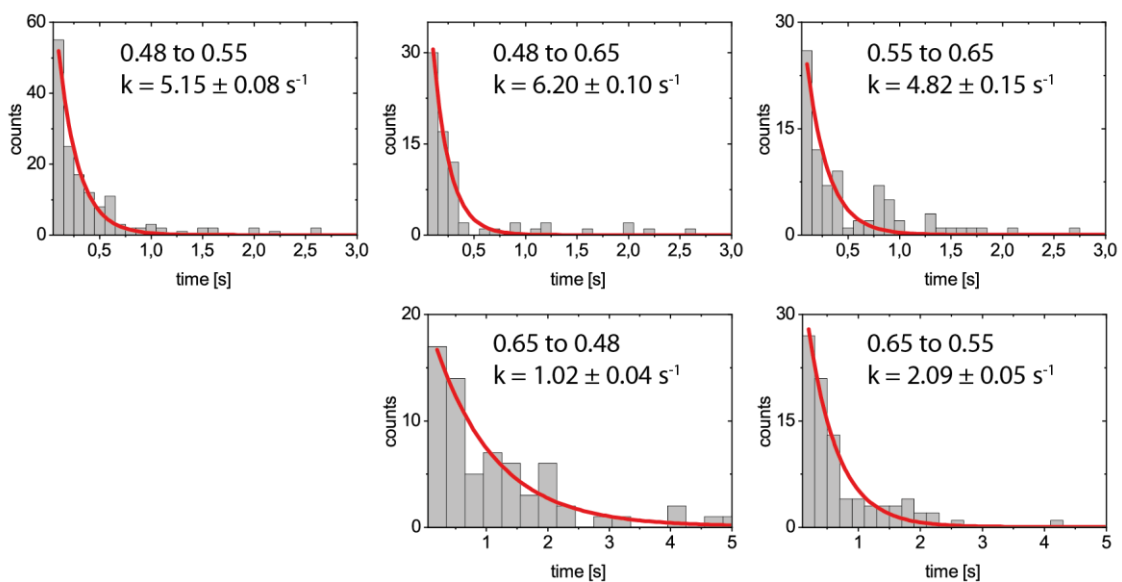


Figure 95: Exponential decay fits for the abasic RNA with complex METTL3/METTL14/WTAP and SAH.

6.8.5 Transition rates from exponential decay fit

Table 32: Summary of all transition rates obtained by exponential decay fits after dwell time analysis. The transition rates are given in s^{-1} . Relevant transitions with no successful exponential decay fit are marked with 'n.d.'. Irrelevant transitions are marked with '/'.

Condition	0.38 to 0.48	0.38 to 0.56	0.38 to 0.66	0.48 to 0.56	0.48 to 0.66	0.56 to 0.66
METTL3/METTL14 with SAM on substrate RNA	6.95 ± 0.12	7.44 ± 0.03	/	4.52 ± 0.07	/	4.96 ± 0.14
METTL3/METTL14 with SAH on product RNA	5.84 ± 0.05	7.58 ± 0.13	/	4.45 ± 0.07	/	/
METTL3/METTL14/WTAP with SAM on substrate RNA	/	/	n.d.	3.39 ± 0.06	7.75 ± 0.07	n.d.
METTL3/METTL14/WTAP with SAH on product RNA	/	11.17 ± 0.24	/	5.45 ± 0.05	5.16 ± 0.18	5.23 ± 0.19
METTL3/METTL14 with SAM on abasic RNA	/	/	/	4.47 ± 0.06	/	/
METTL3/METTL14 with SAH on abasic RNA	/	/	/	4.26 ± 0.18	/	/
METTL3/METTL14/WTAP with SAM on abasic RNA	5.82 ± 0.13	6.07 ± 0.12	/	6.54 ± 0.14	4.85 ± 0.07	/
METTL3/METTL14/WTAP with SAH on abasic RNA	/	n.d.	/	5.15 ± 0.08	6.20 ± 0.10	4.82 ± 0.15
Condition	0.48 to 0.38	0.56 to 0.38	0.66 to 0.38	0.56 to 0.48	0.66 to 0.48	0.66 to 0.56
METTL3/METTL14 with SAM on substrate RNA	1.81 ± 0.01	0.90 ± 0.04	/	2.05 ± 0.05	/	2.32 ± 0.10
METTL3/METTL14 with SAH on product RNA	1.62 ± 0.03	n.d.	/	2.64 ± 0.06	/	/
METTL3/METTL14/WTAP with SAM on substrate RNA	/	/	/	2.11 ± 0.01	n.d.	n.d.
METTL3/METTL14/WTAP with SAH on product RNA	/	n.d.	/	1.26 ± 0.02	/	n.d.
METTL3/METTL14 with SAM on abasic RNA	/	/	/	1.39 ± 0.06	/	/
METTL3/METTL14 with SAH on abasic RNA	/	/	/	n.d.	/	/
METTL3/METTL14/WTAP with SAM on abasic RNA	n.d.	0.70 ± 0.03	/	1.70 ± 0.06	n.d.	n.d.
METTL3/METTL14/WTAP with SAH on abasic RNA	/	n.d.	/	n.d.	1.02 ± 0.04	2.09 ± 0.05

6.8.6 Transition rates from average dwell times

Table 33: Summary of all transition rates obtained by average dwell times. The transition rates are given in s^{-1} . Irrelevant transitions are marked with '/'.

Condition	0.38 to 0.48	0.38 to 0.56	0.38 to 0.66	0.48 to 0.56	0.48 to 0.66	0.56 to 0.66
METTL3/METTL14 with SAM on substrate RNA	0.9	1.2	/	0.6	/	0.4
METTL3/METTL14 with SAH on product RNA	1.7	0.6	/	0.8	/	/
METTL3/METTL14/WTAP with SAM on substrate RNA	/	/	0.4	0.5	1.2	0.4
METTL3/METTL14/WTAP with SAH on product RNA	/	0.8	/	0.9	0.6	0.4
METTL3/METTL14 with SAM on abasic RNA	/	/	/	0.6	/	/
METTL3/METTL14 with SAH on abasic RNA	/	/	/	0.7	/	/
METTL3/METTL14/WTAP with SAM on abasic RNA	1.1	1.1	/	0.9	0.8	/
METTL3/METTL14/WTAP with SAH on abasic RNA	/	1.0	/	0.9	1.1	0.7
Condition	0.48 to 0.38	0.56 to 0.38	0.66 to 0.38	0.56 to 0.48	0.66 to 0.48	0.66 to 0.56
METTL3/METTL14 with SAM on substrate RNA	0.7	0.5	/	0.6	/	0.3
METTL3/METTL14 with SAH on product RNA	0.7	0.4	/	0.8	/	/
METTL3/METTL14/WTAP with SAM on substrate RNA	/	/	/	0.6	0.5	0.3
METTL3/METTL14/WTAP with SAH on product RNA	/	0.5	/	0.8	/	0.6
METTL3/METTL14 with SAM on abasic RNA	/	/	/	0.8	/	/
METTL3/METTL14 with SAH on abasic RNA	/	/	/	0.9	/	/
METTL3/METTL14/WTAP with SAM on abasic RNA	1.0	0.4	/	0.8	0.4	0.4
METTL3/METTL14/WTAP with SAH on abasic RNA	/	0.3	/	0.9	0.4	0.7

References

1. Desrosiers, R., Friderici, K. & Rottman, F. Identification of methylated nucleosides in messenger RNA from Novikoff hepatoma cells. *Proceedings of the National Academy of Sciences* **71**, 3971–3975; 10.1073/pnas.71.10.3971 (1974).
2. Adams, J. M. & Cory, S. Modified nucleosides and bizarre 5'-termini in mouse myeloma mRNA. *Nature* **255**, 28–33; 10.1038/255028a0 (1975).
3. Dominissini, D. *et al.* Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. *Nature* **485**, 201–206; 10.1038/nature11112 (2012).
4. Horowitz, S., Horowitz, A., Nilsen, T. W., Munns, T. W. & Rottman, F. M. Mapping of N6-methyladenosine residues in bovine prolactin mRNA. *Proceedings of the National Academy of Sciences* **81**, 5667–5671; 10.1073/pnas.81.18.5667 (1984).
5. Meyer, K. D. *et al.* Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. *Cell* **149**, 1635–1646; 10.1016/j.cell.2012.05.003 (2012).
6. Schibler, U., Kelley, D. E. & Perry, R. P. Comparison of methylated sequences in messenger RNA and heterogeneous nuclear RNA from mouse L cells. *Journal of molecular biology* **115**, 695–714; 10.1016/0022-2836(77)90110-3 (1977).
7. Wei, C. M. & Moss, B. Nucleotide sequences at the N6-methyladenosine sites of HeLa cell messenger ribonucleic acid. *Biochemistry* **16**, 1672–1676; 10.1021/bi00627a023 (1977).
8. Zhao, B. S., Roundtree, I. A. & He, C. Post-transcriptional gene regulation by mRNA modifications. *Nature reviews. Molecular cell biology* **18**, 31–42; 10.1038/nrm.2016.132 (2017).
9. Bokar, J. A., Shambaugh, M. E., Polayes, D., Matera, A. G. & Rottman, F. M. Purification and cDNA cloning of the AdoMet-binding subunit of the human mRNA (N6-adenosine)-methyltransferase. *Rna* **3**, 1233–1247 (1997).
10. Bokar, J. A., Rath-Shambaugh, M. E., Ludwiczak, R., Narayan, P. & Rottman, F. Characterization and partial purification of mRNA N6-adenosine methyltransferase from HeLa cell nuclei. Internal mRNA methylation requires a multisubunit complex. *The Journal of biological chemistry* **269**, 17697–17704 (1994).
11. Harper, J. E., Miceli, S. M., Roberts, R. J. & Manley, J. L. Sequence specificity of the human mRNA N6-adenosine methylase in vitro. *Nucleic Acids Research* **18**, 5735–5741; 10.1093/nar/18.19.5735 (1990).
12. Knuckles, P. *et al.* RNA fate determination through cotranscriptional adenosine methylation and microprocessor binding. *Nature Structural & Molecular Biology* **24**, 561–569; 10.1038/nsmb.3419 (2017).
13. Liu, J. *et al.* A METTL3-METTL14 complex mediates mammalian nuclear RNA N6-adenosine methylation. *Nature chemical biology* **10**, 93–95; 10.1038/nchembio.1432 (2014).
14. Louloui, A., Ntini, E., Conrad, T. & Ørom, U. A. V. Transient N-6-Methyladenosine Transcriptome Sequencing Reveals a Regulatory Role of m6A in Splicing Efficiency. *Cell reports* **23**, 3429–3437; 10.1016/j.celrep.2018.05.077 (2018).
15. Slobodin, B. *et al.* Transcription Impacts the Efficiency of mRNA Translation via Co-transcriptional N6-adenosine Methylation. *Cell* **169**, 326–337.e12; 10.1016/j.cell.2017.03.031 (2017).

16. Śledź, P. & Jinek, M. Structural insights into the molecular mechanism of the m(6)A writer complex. *eLife* **5**, 1–16; 10.7554/eLife.18434 (2016).
17. Wang, P., Doxtader, K. A. & Nam, Y. Structural Basis for Cooperative Function of Mettl3 and Mettl14 Methyltransferases. *Molecular Cell* **63**, 306–317; 10.1016/j.molcel.2016.05.041 (2016).
18. Wang, X. *et al.* Structural basis of N(6)-adenosine methylation by the METTL3-METTL14 complex. *Nature* **534**, 575–578; 10.1038/nature18298 (2016).
19. Huang, J. *et al.* Solution structure of the RNA recognition domain of METTL3-METTL14 N6-methyladenosine methyltransferase. *Protein & Cell* **10**, 272–284; 10.1007/s13238-018-0518-7 (2019).
20. Schöller, E. *et al.* Interactions, localization, and phosphorylation of the m6A generating METTL3-METTL14-WTAP complex. *Rna* **24**, 499–512; 10.1261/rna.064063.117 (2018).
21. Ping, X.-L. *et al.* Mammalian WTAP is a regulatory subunit of the RNA N6-methyladenosine methyltransferase. *Cell Research* **24**, 177–189; 10.1038/cr.2014.3 (2014).
22. Schwartz, S. *et al.* Perturbation of m6A writers reveals two distinct classes of mRNA methylation at internal and 5' sites. *Cell reports* **8**, 284–296; 10.1016/j.celrep.2014.05.048 (2014).
23. Knuckles, P. *et al.* Zc3h13/Flacc is required for adenosine methylation by bridging the mRNA-binding factor Rbm15/Spenito to the m6A machinery component Wtap/FI(2)d. *Genes & Development* **32**, 415–429; 10.1101/gad.309146.117 (2018).
24. Patil, D. P. *et al.* m(6)A RNA methylation promotes XIST-mediated transcriptional repression. *Nature* **537**, 369–373; 10.1038/nature19342 (2016).
25. Yue, Y. *et al.* VIRMA mediates preferential m6A mRNA methylation in 3'UTR and near stop codon and associates with alternative polyadenylation. *Cell Discovery* **4**, 10; 10.1038/s41421-018-0019-0 (2018).
26. Bawankar, P. *et al.* Hakai is required for stabilization of core components of the m6A mRNA methylation machinery. *Nature Communications* **12**, 3778; 10.1038/s41467-021-23892-5 (2021).
27. Boccaletto, P. *et al.* MODOMICS: a database of RNA modification pathways. 2021 update. *Nucleic Acids Research* **50**, D231–D235; 10.1093/nar/gkab1083 (2022).
28. Boccaletto, P. *et al.* MODOMICS: a database of RNA modification pathways. 2017 update. *Nucleic Acids Research* **46**, D303–D307; 10.1093/nar/gkx1030 (2018).
29. Czerwoniec, A. *et al.* MODOMICS: a database of RNA modification pathways. 2008 update. *Nucleic Acids Research* **37**, D118–21; 10.1093/nar/gkn710 (2009).
30. Dunin-Horkawicz, S. *et al.* MODOMICS: a database of RNA modification pathways. *Nucleic Acids Research* **34**, D145–9; 10.1093/nar/gkj084 (2006).
31. Machnicka, M. A. *et al.* MODOMICS: a database of RNA modification pathways--2013 update. *Nucleic Acids Research* **41**, D262–7; 10.1093/nar/gks1007 (2013).
32. Lorenz, C., Lünse, C. E. & Mörl, M. tRNA Modifications: Impact on Structure and Thermal Adaptation. *Biomolecules* **7**, 35; 10.3390/biom7020035 (2017).
33. Charette, M. & Gray, M. W. Pseudouridine in RNA: what, where, how, and why. *IUBMB Life (International Union of Biochemistry and Molecular Biology: Life)* **49**, 341–351; 10.1080/152165400410182 (2000).

References

34. Kierzek, R., Burkard, M. E. & Turner, D. H. Thermodynamics of single mismatches in RNA duplexes. *Biochemistry* **38**, 14214–14223; 10.1021/bi991186l (1999).
35. Zhou, H. *et al.* m(1)A and m(1)G disrupt A-RNA structure through the intrinsic instability of Hoogsteen base pairs. *Nature Structural & Molecular Biology* **23**, 803–810; 10.1038/nsmb.3270 (2016).
36. Berg, M. D. & Brandl, C. J. Transfer RNAs: diversity in form and function. *RNA Biology* **18**, 316–339; 10.1080/15476286.2020.1809197 (2021).
37. Björk, G. R., Wikström, P. M. & Byström, A. S. Prevention of translational frameshifting by the modified nucleoside 1-methylguanosine. *Science (New York, N.Y.)* **244**, 986–989; 10.1126/science.2471265 (1989).
38. Hagervall, T. G., Tuohy, T. M., Atkins, J. F. & Björk, G. R. Deficiency of 1-methylguanosine in tRNA from *Salmonella typhimurium* induces frameshifting by quadruplet translocation. *Journal of molecular biology* **232**, 756–765; 10.1006/jmbi.1993.1429 (1993).
39. Perry, R. P. & Kelley, D. E. Existence of methylated messenger RNA in mouse L cells. *Cell* **1**, 37–42; 10.1016/0092-8674(74)90153-6 (1974).
40. Maravić, G. Macrolide resistance based on the Erm-mediated rRNA methylation. *Current Drug Target -Infectious Disorders* **4**, 193–202; 10.2174/1568005043340777 (2004).
41. Aly, A., Scott, G., Calderon, M. & Haghghi, A. P. N6-Adenosine Methylation of SARS-CoV-2 5'-UTR Regulates Translation. *bioRxiv*; 10.1101/2022.10.17.512569 (2022).
42. Burgess, H. M. *et al.* Targeting the m6A RNA modification pathway blocks SARS-CoV-2 and HCoV-OC43 replication. *Genes & Development* **35**, 1005–1019; 10.1101/gad.348320.121 (2021).
43. Li, N. *et al.* METTL3 regulates viral m6A RNA modification and host cell innate immune responses during SARS-CoV-2 infection. *Cell reports* **35**, 109091; 10.1016/j.celrep.2021.109091 (2021).
44. Jia, G. *et al.* N6-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. *Nature chemical biology* **7**, 885–887; 10.1038/nchembio.687 (2011).
45. Zheng, G. *et al.* ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. *Molecular Cell* **49**, 18–29; 10.1016/j.molcel.2012.10.015 (2013).
46. Garcia-Campos, M. A. *et al.* Deciphering the "m6A Code" via Antibody-Independent Quantitative Profiling. *Cell* **178**, 731–747.e16; 10.1016/j.cell.2019.06.013 (2019).
47. Barbieri, I. *et al.* Promoter-bound METTL3 maintains myeloid leukaemia by m6A-dependent translation control. *Nature* **552**, 126–131; 10.1038/nature24678 (2017).
48. Bertero, A. *et al.* The SMAD2/3 interactome reveals that TGFβ controls m6A mRNA methylation in pluripotency. *Nature* **555**, 256–259; 10.1038/nature25784 (2018).
49. Fish, L. *et al.* Nuclear TARBP2 Drives Oncogenic Dysregulation of RNA Splicing and Decay. *Molecular Cell* **75**, 967–981.e9; 10.1016/j.molcel.2019.06.001 (2019).
50. Uzonyi, A. *et al.* Exclusion of m6A from splice-site proximal regions by the exon junction complex dictates m6A topologies and mRNA stability. *Molecular Cell* **83**, 237–251.e7; 10.1016/j.molcel.2022.12.026 (2023).
51. He, P. C. & He, C. m6 A RNA methylation: from mechanisms to therapeutic potential. *The EMBO Journal* **40**, e105977; 10.15252/EMBJ.2020105977 (2021).

52. Xiao, W. *et al.* Nuclear m(6)A Reader YTHDC1 Regulates mRNA Splicing. *Molecular Cell* **61**, 507–519; 10.1016/j.molcel.2016.01.012 (2016).
53. Roundtree, I. A. *et al.* YTHDC1 mediates nuclear export of N6-methyladenosine methylated mRNAs. *eLife* **6**; 10.7554/eLife.31311 (2017).
54. Wang, X. *et al.* N(6)-methyladenosine Modulates Messenger RNA Translation Efficiency. *Cell* **161**, 1388–1399; 10.1016/j.cell.2015.05.014 (2015).
55. Huang, H. *et al.* Recognition of RNA N6-methyladenosine by IGF2BP proteins enhances mRNA stability and translation. *Nature Cell Biology* **20**, 285–295; 10.1038/s41556-018-0045-z (2018).
56. Wang, X. *et al.* N6-methyladenosine-dependent regulation of messenger RNA stability. *Nature* **505**, 117–120; 10.1038/nature12730 (2014).
57. Huang, L., Ashraf, S., Wang, J. & Lilley, D. M. Control of box C/D snoRNP assembly by N6-methylation of adenine. *EMBO reports* **18**, 1631–1645; 10.15252/embr.201743967 (2017).
58. Moore, T., Zhang, Y., Fenley, M. O. & Li, H. Molecular basis of box C/D RNA-protein interactions; cocrystal structure of archaeal L7Ae and a box C/D RNA. *Structure* **12**, 807–818; 10.1016/j.str.2004.02.033 (2004).
59. Watkins, N. J., Dickmanns, A. & Lührmann, R. Conserved stem II of the box C/D motif is essential for nucleolar localization and is required, along with the 15.5K protein, for the hierarchical assembly of the box C/D snoRNP. *Molecular and cellular biology* **22**, 8342–8352; 10.1128/MCB.22.23.8342-8352.2002 (2002).
60. Liu, N. *et al.* N(6)-methyladenosine-dependent RNA structural switches regulate RNA-protein interactions. *Nature* **518**, 560–564; 10.1038/nature14234 (2015).
61. Chen, M. *et al.* RNA N6-methyladenosine methyltransferase-like 3 promotes liver cancer progression through YTHDF2-dependent posttranscriptional silencing of SOCS2. *Hepatology* **67**, 2254–2270; 10.1002/hep.29683 (2018).
62. Vu, L. P. *et al.* The N6-methyladenosine (m6A)-forming enzyme METTL3 controls myeloid differentiation of normal hematopoietic and leukemia cells. *Nature Medicine* **23**, 1369–1376; 10.1038/nm.4416 (2017).
63. Wang, H., Xu, B. & Shi, J. N6-methyladenosine METTL3 promotes the breast cancer progression via targeting Bcl-2. *Gene* **722**, 144076; 10.1016/j.gene.2019.144076 (2020).
64. STORM Therapeutics. Available at <https://www.stormtherapeutics.com/science/stc-15/>.
65. Yankova, E. *et al.* Small-molecule inhibition of METTL3 as a strategy against myeloid leukaemia. *Nature* **593**, 597–601; 10.1038/s41586-021-03536-w (2021).
66. Liu, N. *et al.* Probing N6-methyladenosine RNA modification status at single nucleotide resolution in mRNA and long noncoding RNA. *Rna* **19**, 1848–1856; 10.1261/rna.041178.113 (2013).
67. Meyer, K. D. & Jaffrey, S. R. The dynamic epitranscriptome: N6-methyladenosine and gene expression control. *Nature reviews. Molecular cell biology* **15**, 313–326; 10.1038/nrm3785 (2014).
68. Bringmann, P. & Lührmann, R. Antibodies specific for N6-methyladenosine react with intact snRNPs U2 and U4/U6. *FEBS Letters* **213**, 309–315; 10.1016/0014-5793(87)81512-0 (1987).

References

69. Schwartz, S. *et al.* High-resolution mapping reveals a conserved, widespread, dynamic mRNA methylation program in yeast meiosis. *Cell* **155**, 1409–1421; 10.1016/j.cell.2013.10.047 (2013).
70. Linder, B. *et al.* Single-nucleotide-resolution mapping of m6A and m6Am throughout the transcriptome. *Nature Methods* **12**, 767–772; 10.1038/nmeth.3453 (2015).
71. König, J. *et al.* iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. *Nature Structural & Molecular Biology* **17**, 909–915; 10.1038/nsmb.1838 (2010).
72. Körtel, N. *et al.* Deep and accurate detection of m6A RNA modifications using miCLIP2 and m6Aboost machine learning. *Nucleic Acids Research* **49**, e92; 10.1093/nar/gkab485 (2021).
73. Huang, H. *et al.* Histone H3 trimethylation at lysine 36 guides m6A RNA modification co-transcriptionally. *Nature* **567**, 414–419; 10.1038/s41586-019-1016-7 (2019).
74. Iyer, L. M., Zhang, D. & Aravind, L. Adenine methylation in eukaryotes: Apprehending the complex evolutionary history and functional potential of an epigenetic modification. *BioEssays : news and reviews in molecular, cellular and developmental biology* **38**, 27–40; 10.1002/bies.201500104 (2016).
75. Gupta, Y. K., Chan, S.-H., Xu, S.-Y. & Aggarwal, A. K. Structural basis of asymmetric DNA methylation and ATP-triggered long-range diffusion by EcoP15I. *Nature Communications* **6**, 7363; 10.1038/ncomms8363 (2015).
76. Goedecke, K., Pignot, M., Goody, R. S., Scheidig, A. J. & Weinhold, E. Structure of the N6-adenine DNA methyltransferase M.TaqI in complex with DNA and a cofactor analog. *Nature structural biology* **8**, 121–125; 10.1038/84104 (2001).
77. Horiuchi, K. *et al.* Identification of Wilms' tumor 1-associating protein complex and its role in alternative splicing and the cell cycle. *The Journal of biological chemistry* **288**, 33292–33302; 10.1074/jbc.M113.500397 (2013).
78. Wen, J. *et al.* Zc3h13 Regulates Nuclear RNA m6A Methylation and Mouse Embryonic Stem Cell Self-Renewal. *Molecular Cell* **69**, 1028–1038.e6; 10.1016/j.molcel.2018.02.015 (2018).
79. Yan, D. & Perrimon, N. spenito is required for sex determination in *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences of the United States of America* **112**, 11606–11611; 10.1073/pnas.1515891112 (2015).
80. Su, S. *et al.* Cryo-EM structures of human m6A writer complexes. *Cell Research* **2022** **32**, 982–994; 10.1038/s41422-022-00725-8 (2022).
81. Sir John Frederick William Herschel. IV. 'ΑμόρφωYa, no. I.— on a case of superficial colour presented by a homogeneous liquid internally colourless. *Phil. Trans. R. Soc.* **135**, 143–145; 10.1098/rstl.1845.0004 (1845).
82. Atkins, P. W. & Paula, J. de. *Physikalische Chemie*. 5th ed. (Wiley-VCH, Weinheim, 2013).
83. Lakowicz, J. R. *Principles of Fluorescence Spectroscopy* (Springer US, Boston, MA, 2006).
84. Roy, R., Hohng, S. & Ha, T. A practical guide to single-molecule FRET. *Nature Methods* **5**, 507–516; 10.1038/nmeth.1208 (2008).
85. Gell, C., Brockwell, D. & Smith, A. *Handbook of single molecule fluorescence spectroscopy* (Oxford University Press, Oxford, 2006).

86. Moerner, W. E. & Fromm, D. P. Methods of single-molecule fluorescence spectroscopy and microscopy. *Review of Scientific Instruments* **74**, 3597–3619; 10.1063/1.1589587 (2003).
87. Ha, T. *et al.* Probing the interaction between two single molecules: fluorescence resonance energy transfer between a single donor and a single acceptor. *Proceedings of the National Academy of Sciences* **93**, 6264–6268; 10.1073/pnas.93.13.6264 (1996).
88. Suddala, K. C., Wang, J., Hou, Q. & Walter, N. G. Mg(2+) shifts ligand-mediated folding of a riboswitch from induced-fit to conformational selection. *Journal of the American Chemical Society* **137**, 14075–14083; 10.1021/jacs.5b09740 (2015).
89. Warhaut, S. *et al.* Ligand-modulated folding of the full-length adenine riboswitch probed by NMR and single-molecule FRET spectroscopy. *Nucleic Acids Research* **45**, 5512–5522; 10.1093/nar/gkx110 (2017).
90. Deniz, A. A. *et al.* Single-molecule protein folding: diffusion fluorescence resonance energy transfer studies of the denaturation of chymotrypsin inhibitor 2. *Proceedings of the National Academy of Sciences* **97**, 5179–5184; 10.1073/pnas.090104997 (2000).
91. Jia, Y. *et al.* Folding dynamics of single GCN-4 peptides by fluorescence resonant energy transfer confocal microscopy. *Chemical Physics* **247**, 69–83; 10.1016/S0301-0104(99)00127-5 (1999).
92. Talaga, D. S. *et al.* Dynamics and folding of single two-stranded coiled-coil peptides studied by fluorescent energy transfer confocal microscopy. *Proceedings of the National Academy of Sciences* **97**, 13021–13026; 10.1073/pnas.97.24.13021 (2000).
93. Schuler, B. & Eaton, W. A. Protein folding studied by single-molecule FRET. *Current Opinion in Structural Biology* **18**, 16–26; 10.1016/J.SBI.2007.12.003 (2008).
94. Meiser, N., Fuks, C. & Hengesbach, M. Cooperative Analysis of Structural Dynamics in RNA-Protein Complexes by Single-Molecule Förster Resonance Energy Transfer Spectroscopy. *Molecules* **25**; 10.3390/molecules25092057 (2020).
95. Bhattacharjee, A., Wang, Y., Diao, J. & Price, C. M. Dynamic DNA binding, junction recognition and G4 melting activity underlie the telomeric and genome-wide roles of human CST. *Nucleic Acids Research* **45**, 12311–12324; 10.1093/nar/gkx878 (2017).
96. Feng, X. *et al.* CTC1-STN1 terminates telomerase while STN1-TEN1 enables C-strand synthesis during telomere replication in colon cancer cells. *Nature Communications* **9**, 2827; 10.1038/s41467-018-05154-z (2018).
97. Hengesbach, M., Akiyama, B. M. & Stone, M. D. Single-molecule analysis of telomerase structure and function. *Current opinion in chemical biology* **15**, 845–852; 10.1016/j.cbpa.2011.10.008 (2011).
98. Parks, J. W., Kappel, K., Das, R. & Stone, M. D. Single-molecule FRET-Rosetta reveals RNA structural rearrangements during human telomerase catalysis. *Rna* **23**, 175–188; 10.1261/rna.058743.116 (2017).
99. Jakob, L., Gust, A. & Grohmann, D. Evaluation and optimisation of unnatural amino acid incorporation and bioorthogonal bioconjugation for site-specific fluorescent labelling of proteins expressed in mammalian cells. *Biochemistry and Biophysics Reports* **17**, 1–9; 10.1016/J.BBREP.2018.10.011 (2019).
100. Kramm, K., Endesfelder, U. & Grohmann, D. A Single-Molecule View of Archaeal Transcription. *Journal of molecular biology* **431**, 4116–4131; 10.1016/j.jmb.2019.06.009 (2019).

References

101. Beier, D. H. *et al.* Dynamics of the DEAD-box ATPase Prp5 RecA-like domains provide a conformational switch during spliceosome assembly. *Nucleic Acids Research* **47**, 10842–10851; 10.1093/nar/gkz765 (2019).
102. Blanco, M. R. *et al.* Single Molecule Cluster Analysis dissects splicing pathway conformational dynamics. *Nature Methods* **12**, 1077–1084; 10.1038/nmeth.3602 (2015).
103. Rodgers, M. L. *et al.* Conformational dynamics of stem II of the U2 snRNA. *Rna* **22**, 225–236; 10.1261/rna.052233.115 (2016).
104. van der Feltz, C., DeHaven, A. C. & Hoskins, A. A. Stress-induced Pseudouridylation Alters the Structural Equilibrium of Yeast U2 snRNA Stem II. *Journal of molecular biology* **430**, 524–536; 10.1016/j.jmb.2017.10.021 (2018).
105. Munro, J. B., Altman, R. B., O'Connor, N. & Blanchard, S. C. Identification of two distinct hybrid state intermediates on the ribosome. *Molecular Cell* **25**, 505–517; 10.1016/j.molcel.2007.01.022 (2007).
106. Prokhorova, I. *et al.* Aminoglycoside interactions and impacts on the eukaryotic ribosome. *Proceedings of the National Academy of Sciences of the United States of America* **114**, E10899–E10908; 10.1073/pnas.1715501114 (2017).
107. Wang, L., Wasserman, M. R., Feldman, M. B., Altman, R. B. & Blanchard, S. C. Mechanistic insights into antibiotic action on the ribosome through single-molecule fluorescence imaging. *Annals of the New York Academy of Sciences* **1241**, E1-16; 10.1111/j.1749-6632.2012.06839.x (2011).
108. Joo, C., Ha, T. & Selvin, P. R. Single-molecule FRET with total internal reflection microscopy. *Single-molecule techniques : a laboratory manual* (2008).
109. Hecht, E. *Optik*. 7th ed. (De Gruyter, 2018).
110. Michalet, X. *et al.* Detectors for single-molecule fluorescence imaging and spectroscopy. *Journal of Modern Optics* **54**, 239; 10.1080/09500340600769067 (2007).
111. Mullan, A. & Cooper, J. Single-Molecule Studies - EMCCD or sCMOS? Available at <https://andor.oxinst.com/learning/view/article/what-is-the-best-detector-for-single-molecule-studies>.
112. Mullan, A. & Wise, A. Why EMCCD remains the most sensitive sensor technology - Separating the Facts from the Fiction. Available at <https://andor.oxinst.com/learning/view/article/emccd-separating-the-facts-from-the-fiction>.
113. Englander, S. W., Calhoun, D. B. & Englander, J. J. Biochemistry without oxygen. *Analytical Biochemistry* **161**, 300–306; 10.1016/0003-2697(87)90454-4 (1987).
114. Aitken, C. E., Marshall, R. A. & Puglisi, J. D. An oxygen scavenging system for improvement of dye stability in single-molecule fluorescence experiments. *Biophysical journal* **94**, 1826–1835; 10.1529/BIOPHYSJ.107.117689 (2008).
115. Patil, P. V. & Ballou, D. P. The use of protocatechuate dioxygenase for maintaining anaerobic conditions in biochemical experiments. *Analytical Biochemistry* **286**, 187–192; 10.1006/abio.2000.4802 (2000).
116. Rasnik, I., McKinney, S. A. & Ha, T. Nonblinking and long-lasting single-molecule fluorescence imaging. *Nature Methods* **3**, 891–893; 10.1038/NMETH934 (2006).

117. Hengesbach, M. *et al.* RNA intramolecular dynamics by single-molecule FRET. *Current Protocols in Nucleic Acid Chemistry* **Chapter 11**, Unit 11.12; 10.1002/0471142700.nc1112s34 (2008).
118. Hanspach, G., Trucks, S. & Hengesbach, M. Strategic labelling approaches for RNA single-molecule spectroscopy. *RNA Biology* **16**, 1119–1132; 10.1080/15476286.2019.1593093 (2019).
119. Schmidt, A., Altincekic, N., Gustmann, H., Wachtveitl, J. & Hengesbach, M. The Protein Microenvironment Governs the Suitability of Labeling Sites for Single-Molecule Spectroscopy of RNP Complexes. *ACS Chemical Biology* **13**, 2472–2483; 10.1021/acscchembio.8b00348 (2018).
120. Jalihal, A. P., Lund, P. E. & Walter, N. G. Coming Together: RNAs and Proteins Assemble under the Single-Molecule Fluorescence Microscope. *Cold Spring Harbor Perspectives in Biology* **11**, a032441; 10.1101/cshperspect.a032441 (2019).
121. Rinaldi, A. J., Suddala, K. C. & Walter, N. G. *Native purification and labeling of RNA for single molecule fluorescence studies* (2015).
122. Smith, G. J., Sosnick, T. R., Scherer, N. F. & Pan, T. Efficient fluorescence labeling of a large RNA through oligonucleotide hybridization. *Rna* **11**, 234–239; 10.1261/rna.7180305 (2005).
123. Schmitz, A. G., Zelger-Paulus, S., Gasser, G. & Sigel, R. K. O. Strategy for Internal Labeling of Large RNAs with Minimal Perturbation by Using Fluorescent PNA. *ChemBioChem* **16**, 1302–1306; 10.1002/cbic.201500180 (2015).
124. Xi, C., Balberg, M., Boppart, S. A. & Raskin, L. Use of DNA and peptide nucleic acid molecular beacons for detection and quantification of rRNA in solution and in whole cells. *Applied and Environmental Microbiology* **69**, 5673–5678; 10.1128/AEM.69.9.5673-5678.2003 (2003).
125. Kim, Y. *et al.* Efficient site-specific labeling of proteins via cysteines. *Bioconjugate Chemistry* **19**, 786–791; 10.1021/bc7002499 (2008).
126. Müller, D., Trucks, S., Schwalbe, H. & Hengesbach, M. Genetic Code Expansion Facilitates Position-Selective Modification of Nucleic Acids and Proteins. *ChemPlusChem* **85**, 1233–1243; 10.1002/cplu.202000150 (2020).
127. Ha, T. Single-molecule fluorescence resonance energy transfer. *Methods* **25**, 78–86; 10.1006/meth.2001.1217 (2001).
128. Sofia, S. J., Premnath, V. & Merrill, E. W. Poly(ethylene oxide) Grafted to Silicon Surfaces: Grafting Density and Protein Adsorption. *Macromolecules* **31**, 5059–5070; 10.1021/ma971016l (1998).
129. Meiser, N. Reconstitution of RNP complexes in RNA epigenetics. Masterthesis, 2018.
130. Glasel, J. A. Validity of nucleic acid purities monitored by 260nm/280nm absorbance ratios. *BioTechniques* **18**, 62–63 (1995).
131. Xuan, J.-J. *et al.* RMBase v2.0: deciphering the map of RNA modifications from epitranscriptome sequencing data. *Nucleic Acids Research* **46**, D327–D334; 10.1093/nar/gkx934 (2018).
132. Trucks, S., Hanspach, G. & Hengesbach, M. Eukaryote specific RNA and protein features facilitate assembly and catalysis of H/ACA snoRNPs. *Nucleic Acids Research* **49**, 4629–4642; 10.1093/nar/gkab177 (2021).

References

133. Grosjean, H., Keith, G. & Droogmans, L. Detection and quantification of modified nucleotides in RNA using thin-layer chromatography. *Methods Mol Biol* **265**, 357–391; 10.1385/1-59259-775-0:357 (2004).
134. Keith, G. Mobilities of modified ribonucleotides on two-dimensional cellulose thin-layer chromatography. *Biochimie* **77**, 142–144; 10.1016/0300-9084(96)88118-1 (1995).
135. L'Annunziata, M. F. *Handbook of Radioactivity Analysis* (Academic Press, 2012).
136. Zuker, M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Research* **31**, 3406–3415; 10.1093/nar/gkg595 (2003).
137. Woodcock, C. B. *et al.* Human MettL3-MettL14 complex is a sequence-specific DNA adenine methyltransferase active on single-strand and unpaired DNA in vitro. *Cell Discovery* **5**, 63; 10.1038/s41421-019-0136-4 (2019).
138. Helm, M., Lyko, F. & Motorin, Y. Limited antibody specificity compromises epitranscriptomic analyses. *Nature Communications* **10**, 5669; 10.1038/s41467-019-13684-3 (2019).
139. Ota, T. *et al.* Complete sequencing and characterization of 21,243 full-length human cDNAs. *Nature genetics* **36**, 40–45; 10.1038/ng1285. (2004).
140. Weissmann, F. *et al.* biGBac enables rapid gene assembly for the expression of large multisubunit protein complexes. *Proceedings of the National Academy of Sciences of the United States of America* **113**, E2564-9; 10.1073/pnas.1604935113 (2016).
141. Larson, J. *et al.* Design and construction of a multiwavelength, micromirror total internal reflectance fluorescence microscope. *Nature protocols* **9**, 2317–2328; 10.1038/nprot.2014.155 (2014).
142. Hale, G. M. & Querry, M. R. Optical Constants of Water in the 200-nm to 200-microm Wavelength Region. *Applied Optics* **12**, 555–563; 10.1364/AO.12.000555 (1973).
143. Ciddor, P. E. Refractive index of air: new equations for the visible and near infrared. *Applied Optics* **35**, 1566–1573; 10.1364/AO.35.001566 (1996).
144. Malitson, I. H. Interspecimen Comparison of the Refractive Index of Fused Silica*,†. *Journal of the Optical Society of America* **55**, 1205; 10.1364/JOSA.55.001205 (1965).
145. Kapanidis, A. N. *et al.* Alternating-laser excitation of single molecules. *Accounts of chemical research* **38**, 523–533; 10.1021/AR0401348 (2005).
146. Gautam, Y. & Bade, M. Effect of Auditory Interference on Visual Simple Reaction Time. *Kathmandu University medical journal (KUMJ)* **15**, 329–331 (2017).
147. Stein, I. H., Schüller, V., Böhm, P., Tinnefeld, P. & Liedl, T. Single-molecule FRET ruler based on rigid DNA origami blocks. *Chemphyschem : a European journal of chemical physics and physical chemistry* **12**, 689–695; 10.1002/cphc.201000781 (2011).
148. McKinney, S. A., Joo, C. & Ha, T. Analysis of single-molecule FRET trajectories using hidden Markov modeling. *Biophysical journal* **91**, 1941–1951; 10.1529/biophysj.106.082487 (2006).
149. Mergny, J.-L. & Lacroix, L. Analysis of thermal melting curves. *Oligonucleotides* **13**, 515–537; 10.1089/154545703322860825 (2003).
150. García-Nafría, J., Watson, J. F. & Greger, I. H. IVA cloning: A single-tube universal cloning system exploiting bacterial In Vivo Assembly. *Scientific Reports* **6**, 27459; 10.1038/srep27459 (2016).

151. Untergasser, A. Preparation of Chemical Competent Cells. Available at http://www.untergasser.de/lab/protocols/competent_cells_chemical_v1_0.htm (2008).
152. Jerabek-Willemsen, M. *et al.* MicroScale Thermophoresis: Interaction analysis and beyond. *Journal of Molecular Structure* **1077**, 101–113; 10.1016/J.MOLSTRUC.2014.03.009 (2014).
153. Jerabek-Willemsen, M., Wienken, C. J., Braun, D., Baaske, P. & Duhr, S. Molecular interaction studies using microscale thermophoresis. *Assay and drug development technologies* **9**, 342–353; 10.1089/adt.2011.0380 (2011).
154. Edelstein, A. D. *et al.* Advanced methods of microscope control using μ Manager software. *Journal of Biological Methods* **1**, e10; 10.14440/jbm.2014.36 (2014).
155. Edelstein, A., Amodaj, N., Hoover, K., Vale, R. & Stuurman, N. Computer control of microscopes using μ Manager. *Current Protocols in Molecular Biology* **Chapter 14**, Unit14.20; 10.1002/0471142727.MB1420S92 (2010).
156. Zhuang, X. *et al.* A single-molecule study of RNA catalysis and folding. *Science (New York, N.Y.)* **288**, 2048–2051; 10.1126/science.288.5473.2048 (2000).
157. Mundigala, H., Michaux, J. B., Feig, A. L., Ennifar, E. & Rueda, D. HIV-1 DIS stem loop forms an obligatory bent kissing intermediate in the dimerization pathway. *Nucleic Acids Research* **42**, 7281–7289; 10.1093/nar/gku332 (2014).

List of abbreviations

aa	Amino acid
Am	2'-O-dimethyladenosine
APS	Ammonium persulfate
ASL	Active site loop
ATP	Adenosine triphosphate
bp	Base pairs
cv	Column volumes
DMF	Dimethylformamide
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleoside triphosphate
DTT	Dithiothreitol
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	Ethylenediaminetetraacetic acid
E	FRET efficiency
FBS	Fetal bovine serum
FRET	Förster resonance energy transfer
FTO	Fat mass and obesity-associated protein
HG	Hoogsteen
I_A	Acceptor intensity
I_D	Donor intensity
LB	Lysogeny broth
m6A	N6-methyladenosine
mRNA	Messenger RNA
MCS	Multiple cloning site
MeRIP-Seq	Methylated RNA immunoprecipitation with next generation sequencing
METTL14	Methyltransferase like 14
METTL3	Methyltransferase like 3
NHS	N-Hydroxysuccinimide
nt	Nucleotides
n.d.	Not determined
NTA	Nitrilotriacetic acid
OAc	Acetate
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PNA	Peptide nucleic acid
PNK	Polynucleotide kinase
pTriEx	pTriEx 1.1 TEV
PVDF	Polyvinylidene difluoride
RNA	Ribonucleic acid
RNP	Ribonucleoprotein
rNTP	Deoxyribonucleoside triphosphate

rSAP	Shrimp Alkaline Phosphatase
SAH	S-adenosyl homocysteine
SAM	S-adenosyl methionine
SB	Super broth
sd	Standard deviation
SDS	Sodium dodecyl sulfate
smFRET	Single molecule Förster resonance energy transfer
snoRNP	Small nucleolar ribonucleoprotein
TDP	Transition density plot
TEMED	Tetramethylethylenediamine
TIRF	Total internal reflection fluorescence
TLC	Thin Layer Chromatography
Tris	Tris(hydroxymethyl)aminomethane
tRNA	Transfer RNA
WTAP	Wilms' tumour 1 associated protein
WC	Watson-Crick

Deutsche Zusammenfassung

N6-Methyladenosin (m^6A) ist die häufigste und am besten erforschte Modifikation in eukaryotischer mRNA und wurde zuerst in polyadenylierten Teilen der mRNA identifiziert.^{1,2} Die definierte Verteilung von m^6A im Transkriptom wurde durch frühe biochemische Arbeiten und später durch Antikörper-basierte Sequenzierungstechniken aufgedeckt und zeigt eine besonders hohe m^6A Dichte in langen internen Exons, 3'UTRs und um Stoppcodons.³⁻⁷ Die so genannte m^6A -Schreiber-, Leser- und Entferner-Maschinerie ist für die dynamische und damit regulatorische Natur der m^6A -Modifikation verantwortlich.⁸ Der menschliche N6-Methyltransferase-Komplex (MTC) methyliert als m^6A -Schreiber cotranskriptionell das zentrale Adenin innerhalb eines RRACH-Sequenzkontextes (vorzugsweise GGACU) in der naszierenden RNA.⁹⁻¹⁵ Der katalytische Kern dieses Komplexes wird von den beiden Proteinen METTL3 und METTL14 gebildet, wobei sich das aktive Zentrum in der Methyltransferase-Domäne (MTD) von METTL3 befindet.¹⁶⁻¹⁸ Es wurde postuliert, dass das DPPW-Motiv in diesem aktiven Zentrum in

der Nähe der Bindungsstelle des Methylsponders S-Adenosylmethionin (SAM) das Zieladenin während der Katalyse bindet. Darüber hinaus wurden eine positiv geladene Furche in der METTL3-METTL14-Grenzfläche, die C-terminale RGG-Domäne in METTL14 und die Zinkfinger motive in METTL3 als wichtige Domänen für die RNA-Bindung identifiziert.^{19,20} Bis heute gibt es jedoch keine Strukturen des katalytischen METTL3-METTL14-Komplexes in voller Länge oder mit gebundener Substrat-RNA.

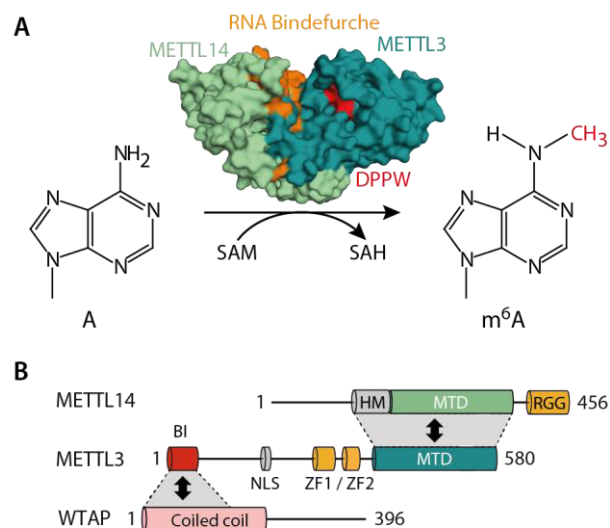


Figure 96: **A** m^6A -Bildung durch den METTL3/METTL14-Komplex (Kristallstruktur der MTDs, PDB 5IL1). **B** Schematische Darstellung der METTL3/METTL14/WTAP-Domänen und Interaktionen.

Des Weiteren bindet eine Reihe von akzessorischen Proteinen an das zentrale METTL3-METTL14-Heterodimer um den vollständigen MTC zu bilden. WTAP fungiert hierbei als Bindeplattform für die weiteren Proteine und bindet selbst an die N-terminale Leader-Helix in METTL3.²⁰ Es wurde gezeigt, dass WTAP den gesamten Komplex an die *nuclear speckles* lokalisiert und die m^6A -Ablagerung an bestimmten Stellen in der mRNA modulieren kann.^{13,21,22} Die weiteren akzessorische Proteine wie VIRMA, RBM15, ZC3H13 und HAKAI wurden zumeist

als Vermittler der positionsspezifischen Methylierung identifiziert.^{23–26} So wurde beispielsweise gezeigt, dass RBM15 eine regionsselektive Methylierung in U-reiche Sequenzen in Abhängigkeit von WTAP vermittelt.²⁴

Die beobachtete Spezifität des Methyltransferase-Komplexes, nur ortsspezifische DRACH-Sequenzen zu methylieren, ist noch wenig verstanden. Einige mögliche Modulatoren wie die Rolle der akzessorischen Proteine werden derzeit untersucht, doch der strukturelle Kontext der RNA-Methylierungsstellen oder eine strukturelle Präferenz des Komplexes wurden bisher weitgehend vernachlässigt. Darüber hinaus ist die strukturelle Dynamik des Komplexes während dieses Methylierungsprozesses noch nicht geklärt. Diese Arbeit leistet einen Beitrag zu den eben genannten Aspekten, indem sie den Methylierungsprozess im Hinblick auf den Einfluss der RNA-Struktur mit enzymatischen Aktivitätsassays und die Dynamiken der Katalyse mit Hilfe eines smFRET-Ansatzes analysiert.

Auf der Grundlage der RNA-Bindungselemente des Komplexes (positiv geladene Bindefurche, Zinkfingerdomäne, RGG-Domäne) und der vermuteten Bindung des Zieladenins im aktiven Zentrum, stellten wir die Hypothese auf, dass die Sekundärstruktur der Ziel-RNA ein weiterer wichtiger Modulator der Methylierungseffizienz ist. Hier postulierten wir die Möglichkeit, dass ein rausgedrehtes Adenin von besonderer Bedeutung sein könnte, was eng mit der lokalen Stabilität der Zielstruktur zusammenhängt. Außerdem sollte eine effiziente Bindung des Proteinkomplexes an die RNA die Fähigkeit voraussetzen, die RNA auf beiden Seiten der Zielsequenz zu verankern.

Um diese Hypothese zu untersuchen, wurden fünf GGACU-haltige RNAs mit unterschiedlicher Sekundärstruktur entworfen und in enzymatische Aktivitätstests mit dem METTL3/METTL14-Komplex als Substrate verwendet. Die resultierenden Methylierungsausbeuten zeigten signifikante Unterschiede zwischen den jeweiligen Sekundärstrukturen, wobei die effizienteste m⁶A-Bildung in der unstrukturierten RNA erfolgte. Bei einem Duplex mit eher lockerer Basenpaarung in der Zielregion war die Methylierungsausbeute im Vergleich zur unstrukturierten RNA deutlich auf 75 % gesunken. Im Gegensatz dazu wurde keine Methylierungsaktivität an einem Zieladenosin

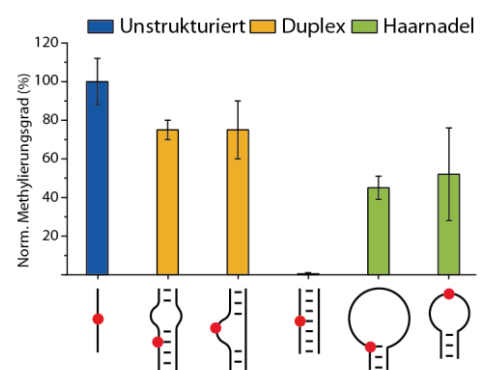


Figure 97: Ergebnisse des Aktivitätsassays für unterschiedlich strukturierte RNAs. Lokale Sekundärstrukturen sind angegeben (nicht die RNA-Konstrukte in voller Länge).

festgestellt, welches durch enge Watson-Crick-Basenpaarung (WC) im Zentrum eines vollständig komplementären Duplexes gebunden war. Dies beweist, dass die lokale Stabilität der Zieladenin enthaltenden RNA-Struktur invers mit der Katalyseeffizienz verbunden ist (hohe Stabilität = geringe Methylierungsausbeute). Der weitere Vergleich des losen Duplexes mit einer Haarnadelstruktur mit ähnlicher lokaler Stabilität ergab eine deutlich geringere Methylierungsausbeute für die Haarnadel. Diese verringerte Methylierungsausbeute bei einer eher gebogenen RNA-Struktur bestätigt vermutlich die Notwendigkeit einer effektiven Proteinbindung auf beiden Seiten der Ziel-RNA-Sequenz für eine effiziente Methylierung.

Darüber hinaus haben wir unstrukturierte und Haarnadel-RNAs mit unterschiedlichen RRACH-Sequenzkontexten getestet. Änderungen an der präferierten GGACU-Sequenz führten zu einer verringerten Methylierungsausbeute, wobei das zentrale GAC ein wichtiger Faktor für eine produktive Katalyse ist. Bei den untersuchten RNA-Konstrukten wurde beobachtet, dass der Einfluss des RRACH-Sequenzkontextes bei Zielnukleotiden in einer Haarnadelstruktur stärker ausgeprägt ist. Insgesamt zeigt dies, dass die Methylierungseffizienz des METTL3/METTL14-Komplexes in hohem Maße von der Kombination aus Sekundärstruktur und Methylierungssequenz der Ziel-RNA abhängt.

Im nächsten Schritt wurde die Rolle der verschiedenen Proteindomänen bei der RNA-Bindung und Katalyse untersucht, wobei der Schwerpunkt auf den RNA-Bindedomänen lag. Dazu wurden sechs verkürzte Proteinkomplexe erfolgreich exprimiert und aufgereinigt. Bei METTL3 wurde der Komplex nach und nach von der N-terminalen Domäne (NTD) verkürzt, so dass beide Zinkfinger (M3 Δ NTD), ein Zinkfinger (M3 Δ ZF1) oder nur die MTD (M3MTD) noch vorhanden waren. Den in METTL14 verkürzten Komplexen fehlt die RGG-Domäne (M14 Δ RGG), die NTD (M14 Δ NTD) oder beide (M14MTD). Ein Minimalkomplex, welcher nur aus den beiden MTDs besteht, wurde als Negativkontrolle verwendet (M3MTD/M14MTD). Nach der Aufreinigung wurden die enzymatische Aktivität sowie die RNA-Bindungseigenschaften aller verkürzten Komplexe an einer unstrukturierten RNA (GGACU) getestet. Die Ergebnisse zeigen eine leicht erhöhte Methylierungseffizienz für den M3 Δ NTD-Komplex im Vergleich zum Vollängenkomplex. Bei allen Komplexen, denen eine RNA-Bindedomäne fehlt, wurde eine deutlich geringere Methylierungsausbeute beobachtet. Dies deutet darauf hin, dass alle RNA-Bindedomänen für eine effiziente Katalyse erforderlich sind, was weitestgehend mit den jeweils ermittelten RNA-Bindungseigenschaften übereinstimmt. Darüber hinaus führt die Deletion des strukturell wichtigen α -helikalen Motivs in METTL14 zu einer verringerten Methylierungsaktivität und einer reduzierten RNA-Bindung. Interessanterweise wurde für den

Komplex mit nur noch einem der beiden Zinkfinger in METTL3 zwar eine geringere Methylierungsausbeute, aber trotzdem eine gute RNA-Bindefähigkeit beobachtet. Weitere RNA-Bindungsexperimente mit einem Off-Target-Konstrukt führten zu der Annahme einer unproduktiven RNA-Bindung dieses Komplexes. Hierbei könnte die RNA-Bindung zwar erfolgen, jedoch in der falschen Orientierung für die Katalyse. Darüber hinaus wurde die Bindung an die Off-Target-RNA auch für den Volllängenkomplex und den Komplex M3 Δ NTD gezeigt, was auf eine generelle RNA-Bindungsfähigkeit des METTL3/METTL14-Komplexes hindeutet und die Sequenzspezifität vor allem durch das aktive Zentrum erfolgt.

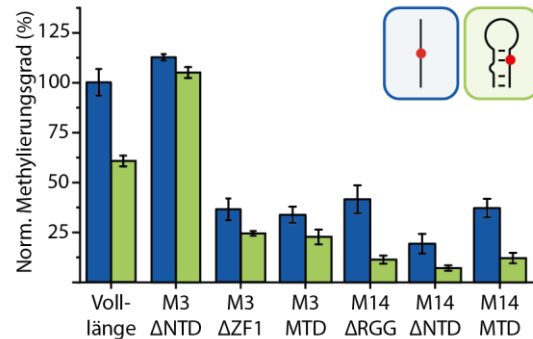





Figure 98: Ergebnisse der Aktivitätstests für die verkürzten Proteinkomplexe an zwei verschiedenen RNAs (farbcodiert). Die lokalen Sekundärstrukturen der RNAs sind angegeben (nicht die vollständige Struktur der Konstrukte).

Alle verkürzten Proteinkomplexe wurden zusätzlich mit einer haarnadelstrukturierten Ziel-RNA getestet, um ihre Empfindlichkeit gegenüber RNA-Sekundärstrukturen zu beurteilen. Die meisten Komplexe zeigten die erwartete verringerte Methylierungsaktivität auf der Haarnadel-RNA. Die relative Abnahme der Methylierungsausbeute war bei den Komplexen denen die RGG-Domäne fehlte ausgeprägter, als bei den ZFD-defizienten Komplexen. Dies deutet auf eine bevorzugte RNA-Orientierung zu Gunsten der RGG-Domäne während der Katalyse an diesem RNA-Konstrukt hin. Auch hier zeigte der Komplex M3 Δ NTD ein einzigartiges Verhalten, wobei keine signifikante Verringerung der Methylierung für die Haarnadel-RNA festgestellt werden konnte. Um diese Beobachtung zu bestätigen, wurden für diesen Komplex enzymatische Aktivitätstests mit zwei weiteren strukturierten RNAs (Haarnadel, loser Duplex) durchgeführt. Die Ergebnisse bestätigten die bereits beobachtete fehlende Empfindlichkeit gegenüber der RNA-Sekundärstruktur. Die NTD in METTL3 enthält die Bindungsstelle für WTAP, welches seinerseits als Bindungsplattform für andere akzessorische Proteine dient.

In den folgenden Experimenten wurde daher die Empfindlichkeit gegenüber der Sekundärstruktur der Ziel-RNA für den METTL3/METTL14/WTAP-Komplex untersucht. Zunächst wurde dieser Komplex mit einem anderen Ansatz exprimiert und aufgereinigt, bei dem nur METTL14 einen His-Tag trägt. Nach erfolgreicher Aufreinigung des Komplexes wurden enzymatische Aktivitätstests mit unterschiedlich strukturierten RNAs durchgeführt. Die

Methylierungsergebnisse für die unstrukturierte RNA waren im Hinblick auf den METTL3/METTL14-Komplex ähnlich. Bei den strukturierten RNAs wurde ein deutlicher Rückgang der Methylierungseffizienz für den WTAP-haltigen Komplex beobachtet. Dies zeigt eine erhöhte strukturelle Empfindlichkeit für diesen Komplex im Vergleich zum METTL3/METTL14-Kernkomplex. Die erzielten Ergebnisse wurden zum besseren Vergleich mit den früheren Daten referenziert. Der Vergleich der drei Komplexe bestätigte die besondere Rolle der NTD in METTL3 für die Empfindlichkeit gegenüber der Sekundärstruktur der Ziel-RNA. Die Deletion der NTD verringerte diese Sensitivität erheblich, im Gegensatz zur Bereitstellung des NTD-Bindungspartners WTAP, wo eine erhöhte Sensitivität beobachtet wurde. Dies könnte mit Unterschieden in der Fähigkeit zur lokalen Destabilisierung der Zielstruktur, dem allgemeinen RNA-Bindungsmodus, der Dynamik im Komplex oder der Zugänglichkeit der RNA-Bindedomänen zusammenhängen.

Table 34: Methylierungsausbeuten für verschiedene Proteinkomplexe an diversen RNAs. (*referenzierte Ausbeuten)

Normalisierter Methylierungsgrad			
M3ΔNTD / METTL14	112 %	118 %	109 %
METTL3 / METTL14	100 %	75 %	45 %
METTL3 / METTL14 / WTAP	102 %	57 %*	28 %*

Im zweiten Teil dieses Projekts wurde die strukturelle Dynamik des METTL3/METTL14-Komplexes in verschiedenen katalytischen Phasen mittels eines smFRET-Ansatzes untersucht. Zunächst wurde ein TIRF-Mikroskop (Total Internal Reflection Fluorescence) gebaut, um die smFRET-Messungen durchzuführen. Der Anregungsstrahl wurde mit drei verschiedenen Laserquellen (637 nm, 532 nm, 488 nm) versehen, die zu einem einzigen Strahlengang parallelisiert wurden. Für jeden Laser wurde direkt nach der Laserquelle ein zusätzlicher Verschluss angebracht, um ein schnelles Ein- und Ausschalten des Anregungslichts auf dem Probenträger zu ermöglichen. Der Laserstrahl wurde auf ein inverses Lichtmikroskop gelenkt und auf ein Prisma fokussiert, um eine interne Totalreflexion auf dem darunter liegenden Objektträger zu erzeugen. Von dort aus sammelt das Wasserimmersionsobjektiv die Probenfluoreszenz und leitet sie an den Detektionspfad weiter. Hier wird das emittierte Licht vom Anregungslicht gefiltert und in zwei Kanäle aufgeteilt (Wellenlänge oberhalb/unterhalb von 642 nm). Die Kanäle wurden dann auf separate Bereiche der EMCCD-Kamera fokussiert, die für die Detektion verwendet wird. Der gesamte Aufbau kann durch individuelle Softwares für die drei Hauptkomponenten (Laser, Verschluss, Kamera) ferngesteuert werden. Um eine wesentlich schnellere und einfachere Kontrolle über diese Komponenten zu erhalten, wurde eine eigene Fernsteuerungssoftware programmiert. Mit dieser können alle erforderlichen Einstellungen in einer einzigen Software vorgenommen werden. Am wichtigsten ist, dass Anregung und

Aufnahme auf Wunsch gekoppelt werden können, was eine große Vielfalt von Anwendungen eröffnet, einschließlich der Filterung aller Moleküle mit nur einem Donorfluorophor.

Für smFRET-Messungen wurde die Ziel-RNA durch Hybridisierung mit kurzen DNA-Strängen fluoreszenzmarkiert und immobilisiert. Hier wurden drei RNA-Konstrukte hergestellt, die sich nur in der Mitte der Zielsequenz unterscheiden: eines mit dem Substrat-Nukleotid (GGACU), eines mit dem Produkt-Nukleotid (GGm6ACU) und eines mit einer abasischen Stelle (GGXCU). Bei der Proteinbindung und Katalyse werden die Abstandsunterschiede der an die RNA gebundenen Farbstoffe mit Änderungen der FRET-Effizienz erfasst. Die katalytische Aktivität der Proteinkomplexe mit

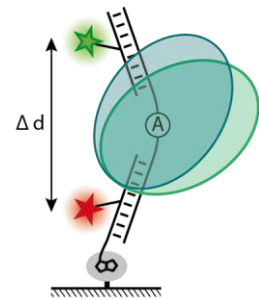


Figure 99: Schematische Darstellung des smFRET Kontrukts.

und ohne WTAP an der Substrat-RNA wurde nachgewiesen. Zunächst wurden die smFRET-Daten aufgenommen, zu Histogrammen verarbeitet und die Ergebnisse für die verschiedenen Ribonukleoproteine (RNPs), bestehend aus RNA (Substrat/Produkt/abasisch), Proteinkomplex (METTL3/METTL14 ohne WTAP) und Cofaktor (SAM/SAH), verglichen. Das Zentrum des proteingebundenen FRET-Zustands für die katalytisch kompetenten RNPs (Substrat-RNA, Proteinkomplex und SAM) war im Vergleich zu den postkatalytischen RNPs (Produkt-RNA, Proteinkomplex und SAH) zu höheren FRET-Effizienzen verschoben. Dies deutet auf eine verstärkte Kontraktion der RNA-Konformation bei den katalytisch kompetenten RNPs hin. Im weiteren Verlauf wurden diese Unterschiede durch die Analyse von Zeitspuren auf ihr dynamisches Verhalten hin untersucht. Die einzelnen Moleküle wurden individuell durch Hidden-Markov-Modellierung angepasst (was zu 4 Zuständen bei $E_{\text{FRET}}=0,38, 0,48, 0,56$ und $0,66$ führte) und einer umfangreichen zusammenfassenden Analyse unterzogen. Die dynamischen Moleküle wurden identifiziert und die Populationsverteilung sowie die Übergänge zwischen den identifizierten Zuständen wurden bestimmt. Außerdem

wurden die Übergangsraten berechnet. Diese Analyse ergab deutliche Unterschiede zwischen den katalytisch kompetenten, postkatalytischen und katalytisch inkompetenten (abasischen) RNPs. Die Population des höchsten FRET-Zustands ($E_{\text{FRET}}=0,66$) war bei den katalytisch kompetenten RNPs im Vergleich zu den postkatalytischen RNPs deutlich erhöht. Dies zeigt, dass dieser Zustand für die Katalyse relevant ist und deutet auf eine Kontraktion der RNA während der Methylierung hin.

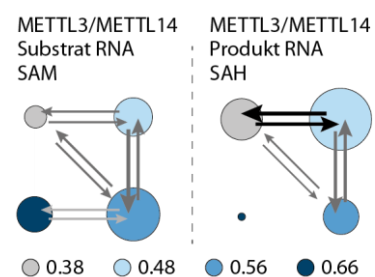


Figure 100: Population von FRET-Zuständen (Kreisgrößen) und Übergängen der dynamischen Moleküle.

Der Vergleich mit den abasischen RNPs ergab, dass dieser Unterschied auf das Vorhandensein von Substrat- oder Produktnukleotid und nicht von SAM oder SAH zurückzuführen ist. Diese Unterschiede spiegeln sich auch in der Population der entsprechenden Übergänge wider. Das gleiche Verhalten wurde für die entsprechenden WTAP-haltigen Komplexe beobachtet, allerdings mit einem allgemeinen Anstieg der Gesamtdynamik und der Population des FRET-Zustands bei $E_{\text{FRET}}=0,66$.

Auf der Grundlage der in dieser Arbeit vorgestellten Ergebnisse wurde eine Hypothese für den katalytischen Zyklus postuliert. Der beobachtete katalytisch relevante FRET-Zustand ($E_{\text{FRET}}=0,66$) ist mit einer Kontraktion der RNA-Konformation im Vergleich zur freien RNA verbunden. Es wurde gezeigt, dass die RNA-Bindung im METTL3/METTL14-Komplex von der Anwesenheit der RGG- und Zinkfinger-Domänen sowie der positiv geladenen Bindungsfurche abhängt.¹⁷⁻²⁰ Darüber hinaus zeigte die hier berichtete Empfindlichkeit des Proteinkomplexes gegenüber der RNA-Sekundärstruktur, wie wichtig eine effektive Proteinverankerung auf beiden Seiten der Ziel-RNA ist. Dies hängt höchstwahrscheinlich mit einer RNA-Bindung durch die RGG-Domäne auf der einen Seite und der ZFD auf der anderen Seite der Zielsequenz zusammen. Zusammengefasst deutet dies auf eine mögliche "Quetschbewegung" der beiden RNA-bindenden Proteindomänen hin, um die gebundene RNA während des Methylierungsprozesses zu kontrahieren. Beim Kontrahieren könnte das Substratadenosin aus der RNA-Struktur in Richtung der aktiven Stelle rotiert werden und wird vermutlich durch das DPPW-Motiv und die beiden Schleifen des aktiven Zentrum¹⁶⁻¹⁸ zur Katalyse gebunden. Nach der Methylierung weist das gebildete m⁶A höchstwahrscheinlich eine geringere Affinität zum aktiven Zentrum auf und wird daher zurück in die RNA-Struktur gedreht. Die erneute Bindung des Produktnukleotids an das aktive Zentrum wird wahrscheinlich sterisch gehindert.

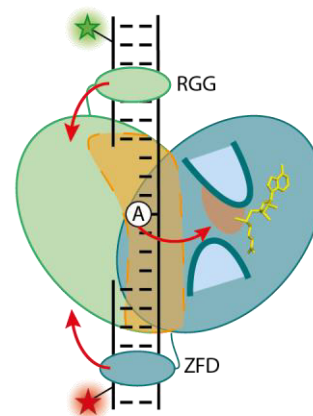


Figure 101: Hypothetisches Modell für die RNA-Bewegung während der Katalyse.

Insgesamt lieferte diese Arbeit neue Einblicke in die RNA-Methylierung durch den METTL3/METTL14-Komplex. Es wurde gezeigt, dass der Methylierungsprozess empfindlich auf die Sekundärstruktur der Ziel-RNAs reagiert, wobei die Stabilität der lokalen Struktur und die Proteinbindung auf beiden Seiten des Ziel-Adenosins besonders wichtig sind. Diese Empfindlichkeit hängt eng mit dem Vorhandensein der NTD in METTL3 zusammen und ist für

den Komplex mit WTAP stärker ausgeprägt. Außerdem wurde beobachtet, dass alle RNA-Bindungsdomänen des Komplexes für eine effiziente Methylierung unerlässlich sind. Die Ergebnisse deuten auf eine allgemeine RNA-Bindung durch die RGG- und Zinkfingerdomänen hin, wobei die Sequenzspezifität vermutlich vom aktiven Zentrum ausgeht. SmFRET-Messungen mit dem selbstgebauten TIRF-Mikroskop deuteten auf eine Kontraktion der gebundenen RNA im katalytisch aktiven Zustand hin. Zusammen mit den vorherigen Ergebnissen ermöglichte dies die Aufstellung einer Hypothese für die Methylierungskatalyse. Diese Hypothese beinhaltet einen möglichen Quetschmechanismus der gebundenen RGG- und Zinkfingerdomänen, welcher die RNA kontrahiert und damit vermutlich das Zieladenosin für die Katalyse in das aktive Zentrum lokalisiert.