Structure, Assembly and Dynamics of H/ACA-Ribonucleoproteins

Dissertation

zur Erlangung des Doktorgrades

der Naturwissenschaften

vorgelegt beim Fachbereich Biochemie, Chemie und Pharmazie

am Institut für Organische Chemie und Chemische Biologie

der Johann Wolfgang Goethe - Universität

in Frankfurt am Main

von

Gerd Werner Günter Hanspach

aus Frankfurt am Main

Frankfurt 2020

(D 30)

Die vorliegende Dissertation wurde im Zeitraum von April 2015 bis Februar 2020 in der Arbeitsgruppe von Dr. Martin Hengesbach am Institut für Organische Chemie und Chemische Biologie der Goethe-Universität Frankfurt am Main angefertigt.

Vom Fachbereich Biochemie, Chemie und Pharmazie der

Johann Wolfgang Goethe - Universität angenommen.

Dekan: Prof. Dr. Clemens Glaubitz

Gutachter: Dr. Martin Hengesbach

Prof. Dr. Harald Schwalbe

Datum der Disputation:

Table of Contents

Chapter I

General Introduction

1.1. Role and Function of RNAs in the Cell

In the middle of the 20th century, the classical dogma in molecular biology stated that deoxyribonucleic acid (DNA) functions as a host for the genetic information, proteins are the machineries of the cell - performing enzymatic reactions and regulation of cellular processes - and ribonucleic acid (RNA) passively mediates the protein expression. The mediation between genetic host and active protein requires three types of RNA: 1) the messenger RNA (mRNA) carrying the genetic information from the DNA to the ribosome which is the cellular protein synthesizer; 2) the transfer RNA (tRNA) carrying the required amino acid to the ribosome for peptide synthesis; and 3) the ribosomal RNA (rRNA) being the main component of the protein synthesis machinery. In recent decades, acquired knowledge about the role and function of cellular RNAs has promoted the RNA from its passive mediator between DNA and proteins to a versatile bio macromolecule with highly important enzymatic and regulatory tasks in all kingdoms of life.

With the discovery of small nuclear RNAs (snRNA) around the 1970s, a new type of non-coding RNA was identified which requires no direct involvement in protein biosynthesis[1]. snRNAs are found in the cell nucleus and form RNA-protein complexes, also called small nuclear ribonucleoprotein (snRNP). Nowadays it is known that snRNPs are involved in RNA processing; especially in splicing, and hence play a key role in gene regulation[2,3]. In the following decades many more types of non-coding RNAs (ncRNAs) were discovered with various functions and tasks in all organisms. Even today new RNAs are discovered, showing that RNAs are much more than just a passive mediator for protein synthesis. Some of these RNAs and their functions will be introduced in the next paragraphs.

A special type of a short ncRNA is called microRNA (miRNA) with a typical length between 21 to 26 nucleotides (nt). After their initial discovery in the well-studied model organism *Caenorhabditis elegans* miRNAs have been found in different higher organisms like flies, plants and mammals[4,5]. miRNAs interrupt gene expression by base pairing to complementary sequences within mRNAs, following up with e.g. less efficient mRNA translation by the ribosomes or mRNA cleavage into two strands due to miRNA base-pairing[6,7]. An example for an evolutionary advantage of this mRNA silencing mechanism can be found in *C. elegans*: In an early stage of larval development, specific miRNAs bind to the 3'-untranslated region of the target RNA,

preventing expression of proteins that are only required in a later stage of the worms life cycle[8]. Another example for gene regulatory RNAs are riboswitches. Riboswitches are located within the untranslated region of mRNAs and can generally be divided into two parts: an aptamer domain that is capable of binding a specific metabolite and an expression platform for protein synthesis. The aptamer binds the metabolite and subsequently the expression platform undergoes conformational changes that either lead to protein biosynthesis 'on-switch' or expression inhibition 'off-switch'[9]. This type of self-regulation can happen during transcription or translation[10,11], and can be responsive to environmental triggers like Mg^{2+}-concentration as well as temperature[12,13], making riboswitches a highly versatile bio macromolecule. Riboswitches and miRNAs are just two examples for RNAs that influence gene expression but are not directly involved in protein biosynthesis (like mRNAs, tRNAs and rRNAs).

Besides gene regulation RNAs can also function as site-specific guides for modifications or perform enzymatic reactions. A class of RNA that is responsible for modifications of other RNAs is called small nucleolar RNA (snoRNA). They are present in the nucleolus of all eukaryotes, range from ~60 to over 200 nucleotides 'nt' in length and take a crucial part in ribosome biogenesis. The main function of snoRNAs is to promote posttranscriptional base modification in rRNAs and snRNAs[14], which is crucial for accurate ribosomal activity and precise splicing. snoRNAs are not capable of performing the modification themselves, but rather transport the active enzyme consisting of the snoRNA and a set of different proteins to the target RNA via sequence specific base pairing, making snoRNAs a highly accurate guide molecule for enzymatic modification. A more detailed look into this enzyme class can be found in Chapter 1.3.

RNAs are not only capable of guiding reactions, but also performing enzymatic reactions in absence of proteins. Such catalytically active RNAs are called ribozymes. The first to be discovered enzymatic stand-alone RNA was ribonuclease (RNase) P , a ribozyme able to cleave of overhangs from tRNA molecules[15]. This finding represents a landmark in RNA biology, since at that time it was generally accepted that biological catalysis was exclusively performed by proteins. Today many more ribozymes are known, like the hammerhead or the hepatitis delta virus ribozyme, which are capable of self-cleavage at specific positions[16,17]. RNAs like these two ribozymes have been engineered for generating self-cleaving constructs. Both are commonly used in *in vitro* transcriptions or other applications for precise RNA synthesis.

The fact that RNAs are directly involved in many cellular processes like gene expression, regulation, RNA modification and processing shows that they are much more than just passive mediators for protein expression. To perform such a variety of functions, RNAs must be able to adopt many different structures and own an intrinsic ability for structural dynamics within itself. With the first solved crystal structure of a tRNA molecule it was apparent that RNA seems to have more structural flexibility than the closely related DNA[18]. Even though, RNA and DNA share many chemical and structural features. RNA folding is more analogous to protein folding than to the highly repetitive double helix DNA folding. Figure 1.1 shows some common cellular RNA motifs.

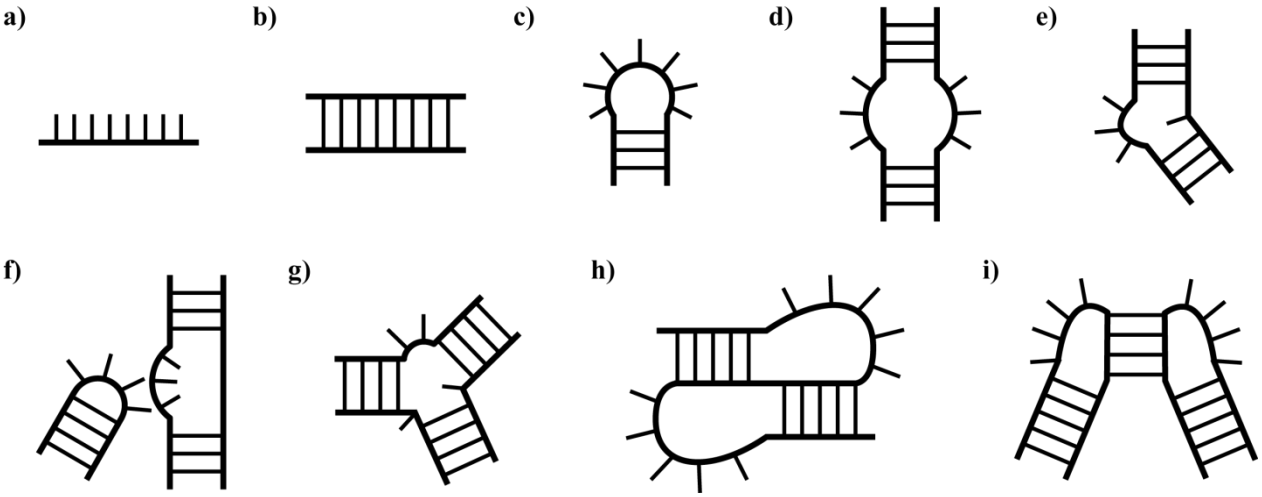


Figure 1.1: Exemplary cellular RNA structures: *a)* Single stranded RNA, *b)* double stranded RNA, *c)* hairpin loop, *d)* internal loop, *e)* bulge, *f)* hairpin loop-bulge interaction, *g)* three stem junction, *h)* pseudoknot and *i)* hairpin-hairpin interaction.

The above mentioned functions; the ability of RNAs to adopt various 3-dimensional structures, the ability to undergo significant conformational changes (e.g. riboswitches) and the fact even today many questions in their cellular roles and functions remain unanswered, show that RNAs are biomolecules of high importance, making them a crucial target for scientists to obtain a full understanding about the cellular processes of life.

1.2. RNA Modifications

The previous chapter was about the biological purpose of different RNAs in the cell. One reason for the high structural and functional variety of these macromolecules can be attributed to the many naturally occurring modifications that can be found in RNAs. As commonly known, RNAs consists of the four consecutive bases adenine (A), guanine (G), uridine (U) and cytidine (C). To date over 170 distinct RNA modifications are known that are inserted enzymatically into RNAs co- or posttranscriptionally[19]. Nucleic acid modifications can be found in all kingdoms of life and range in variety from isomers, simple insertion of methyl groups at sugar or base, to the point of highly chemically modified derivatives of the actual base (see Figure 1.2).

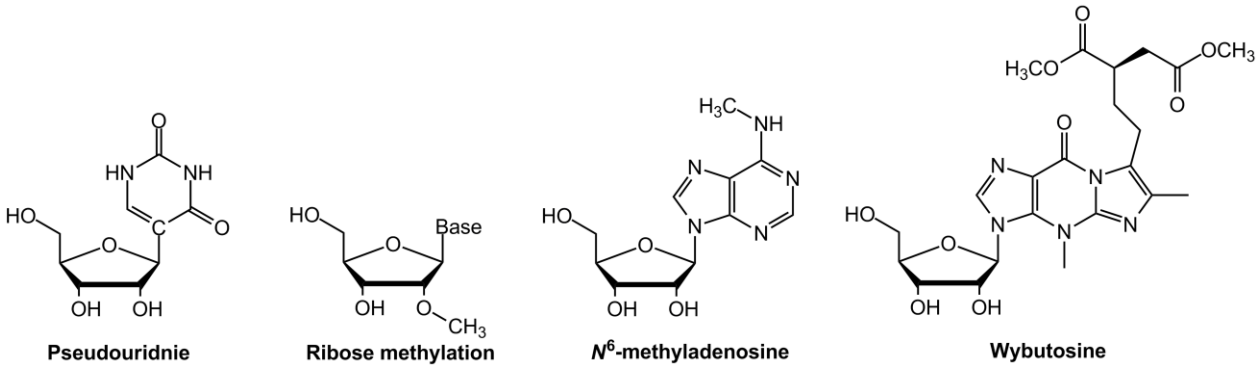


Figure 1.2: Exemplary naturally occurring cellular RNA modifications.

Modified nucleotides are highly distributed in many important RNAs like tRNAs, rRNAs, mRNAs etc. Modification type along with its position can be well conserved even across different domains of life[20-22]. Even though numerous modifications are identified, the biochemical purpose of a certain modification is often unclear. In some cases however, the impact of missing modifications in certain RNAs could be studied. For tRNAs, it was reported that certain modifications, frequently found at position 34 or 37 directly in or next to the codon-anticodon site, facilitate the tRNA-mRNA-interaction by improving base-stacking and ionic interactions with the ribosome, and absence of these modifications may lead to severe translational defects such as frameshifting[23,24]. In eukaryotic mRNAs the most frequent modification is N^6-methyladenosine (m^6A). Recent studies suggest that m^6A has essential roles in pre-mRNA processing. For example, depletion of certain m^6A-writing enzymes results in severe splicing defects for hundreds of genes[25,26], leading to falsely translated proteins. In rRNAs modified nucleotides – predominantly pseudouridine and 2'-O-methylated bases – are clustered in functionally important regions of the ribosome[27,28]. This spatial distribution and the fact that it

is well conserved throughout evolution suggest that these modifications significantly contribute to ribosomal activity. It has been shown that disruption of pseudouridine-formation and 2'-O-methylation leads to severe growth effects in yeast[29,30]. Some of these modifications seem to be dispensable for ribosome function, while others are essential[31]. Despite today's knowledge about RNA modifications, the various types of modified nucleotides and their conserved nature in many cellular RNAs, in most cases the exact function of modifications in specific RNAs remains elusive.

Pseudouridine

The by far most abundant and also the first to be discovered nucleoside modification of all RNAs is pseudouridine (Ψ)[19]. Ψ is a regioisomere of uridine and introduced into the RNA in an enzymatic manner (see Figure 1.3). The glycosidic C-N-bond between sugar and base is broken; the base is kept at its place and turned by 180°. Then the C5 forms a new glycosidic C-C-bond, resulting in the uridine isomer with the N1 becoming a secondary amine. This extra amine group changes the electrochemical properties of the base and slightly increases the polarity of the nucleoside, due to which Ψ and U can easily be separated by thin layer chromatography (TLC)[32]. Additionally the secondary amine yields a new H-bond donor, while the Watson-Crick base pairing pattern is preserved.

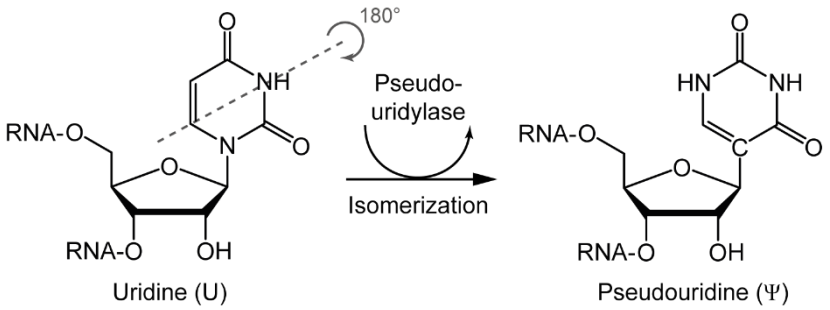


Figure 1.3: Enzymatic isomerization of uridine to pseudouridine.

Pseudouridine is present in all kingdoms of life and can be found in all cellular RNA species, with a strong correlation between organism complexity and Ψ-frequency[33]. It was first discovered in rRNA – more than 60 years ago – and more recently in mRNA[34,35]. Even though the modification has been known for decades, its exact biological function often remains unclear until today. It is well known that Ψ stabilizes RNA stacking of adjacent bases in single stranded RNAs[36], and recently it could be shown that base pairing of Ψ to all four major bases is more stable than U base pairing in RNA duplexes[37,38]. These positive effects on stability are assumed

to origin from N1-imino proton, which coordinates via a water molecule to the phosphate backbone[39]. As a result of its stabilizing properties it is not surprising that Ψ frequently appears in regions where the structure of the RNA is important for its function. For example, the tertiary structure of tRNAs is stabilized by conserved Ψs and prevents malfunction during protein biosynthesis[40]. Furthermore, the high abundance of Ψ in rRNA (~1.2% of total nucleotides in yeast rRNA) and spliceosomal snRNAs (~7% in vertebrate U2 snRNA) suggests that the uridine isomer is important for correct function of the ribosome and spliceosome, two cellular essential machineries. It has been shown that inactivation of the enzymes that catalyze pseudouridylation in rRNAs and snRNAs may lead to malfunctions in protein translation or mRNA splicing, respectively[41,42]. Generally, the contribution of Ψ to ribosome or spliceosome function seems to be in a synergistic manner, meaning that in most cases one single modified nucleotide is not essential for correct function, but the negative influence on precise enzymatic activity growths with increasing number of missing modifications[43].

Enzymatic insertion of Ψ is done by pseudouridylases (Ψases) and can be performed by stand-alone proteins or in RNA guided fashion. Enzymes responsible for pseudouridylation can be divided into six families[44]: TruA, TruB, TruD, RluA, RsuA and Pus10p. All Ψases share an overall fold with a similar catalytic domain and a strictly conserved aspartate residue for enzymatic Ψ formation. These similarities among all Ψases suggest a common catalytic mechanism, which is not yet fully understood, even though several reaction mechanisms are proposed[40]. Stand-alone proteins achieve site specificity for Ψ insertion without any additional factors, but exclusively through differences in protein structures. In contrast, RNA guided uridine isomerization is much more versatile and will be explained in the next chapter.

1.3. Small Nucleolar RNPs

The RNAs directing pseudouridylation belong to the family of snoRNAs, which can be divided into two categories; Box C/D- and box H/ACA-RNAs. Box C/D-RNAs take part in 2'-O-methylation of the ribose, while H/ACA-RNAs facilitate Ψ-formation. In both cases a RNP particle consisting of the snoRNA and several proteins has to be formed prior to the site specific modification of the substrate. Small nucleolar RNAs carry the canonically conserved box sequences as well as unique guide sequences responsible for substrate recognition. In archaea, C/D- and H/ACA-RNAs are termed small RNA (sRNA), sharing the same conserved box motifs as their eukaryotic counterparts. The snoRNA/sRNA guides the modification by base-pairing an individual target RNA to the snoRNP/sRNP. After substrate recruitment, the target base gets modified by one of the proteins and subsequently the product gets released. To ensure correct modification, the snoRNA/sRNA undergoes specific folding and protein recruitment. In archaea, the C/D-sRNP consists of the guide RNA carrying the genetically conserved box C/D and C'/D' motifs, which fold into kink-turn and kink-loop structures, and two sets of three core proteins (Fibrillarin, L7Ae and Nop5) that bind to the C/D-RNA[45,46].

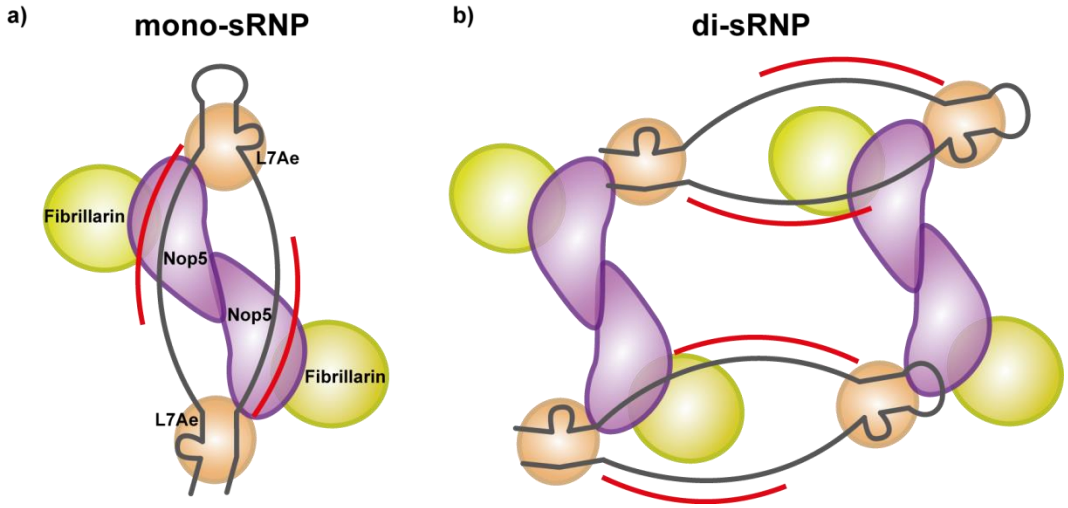


Figure 1.4: Schematic representation of two archaeal box C/D-RNPs: *a)* monomeric sRNP and *b)* dimeric sRNP. Protein colors are indicated. The box C/D-RNA is shown in grey and the target RNA (ribosomal RNA) in red. L7Ae binds to the kink-turn/loop at the box C/D and C'/D' sequences.

Assembly of the RNP is initiated by the kink-turn/loop binding protein L7Ae with further recruitment of the two other proteins, while the methyltransferase Fibrillarin is the actual active enzyme requiring the cofactor S-adenosyl-L-methionine as methyl group donor. The guide RNA inherits a variable hairpin structure in which the C/D- as well as C'/D'-motifs form base pairs. In

between those boxes, there are the guiding sequences for two target RNAs, enabling C/D-RNPs to bind and methylate two substrates. Several structures of archaeal and prokaryotic box C/D-RNPs have been reported and today it is considered that these RNA-protein complexes can either exist as mono-particles (see Figure 1.4a) or as dimers (see Figure 1.4b) formed by two C/D-RNPs[47-50]. Since there is no case known where a single C/D-RNP can form both monomers and dimers, it is assumed that each box C/D-RNP adopts either a monomeric or dimeric state suitable to its biological function.[51]

H/ACA-RNPs

In eukaryotes, pseudouridylation is guided by the other type of snoRNA, the H/ACA-RNA. Likewise to ribose-methylation, a snoRNP has to be formed before the isomerization of uridine can occur. Therefore, the guide RNA typically adopts a hairpin-bulge-hairpin motif with its genetically conserved H- (ANANNA) and ACA-boxes. In archaea, one to three hairpins can be formed by the sRNA, while for snoRNAs in eukaryotes almost exclusively double hairpin structures are known (see Figure 1.5). Each hairpin binds a set of four proteins: Non-Histone Protein 2 'Nhp2' (L7Ae in archaea), Nucleolar Protein 10 'Nop10', Glycine Arginine Rich 1 'Gar1' and Centromere Binding Factor 5 'Cbf5' (Dyskerin or NAP57 in higher eukaryotes). Herein, Cbf5 harbors the catalytically active center, while the other proteins function as cofactors In the following eukaryotic versions of H/ACA-proteins are termed scNop10, scCbf5 and scGar1.

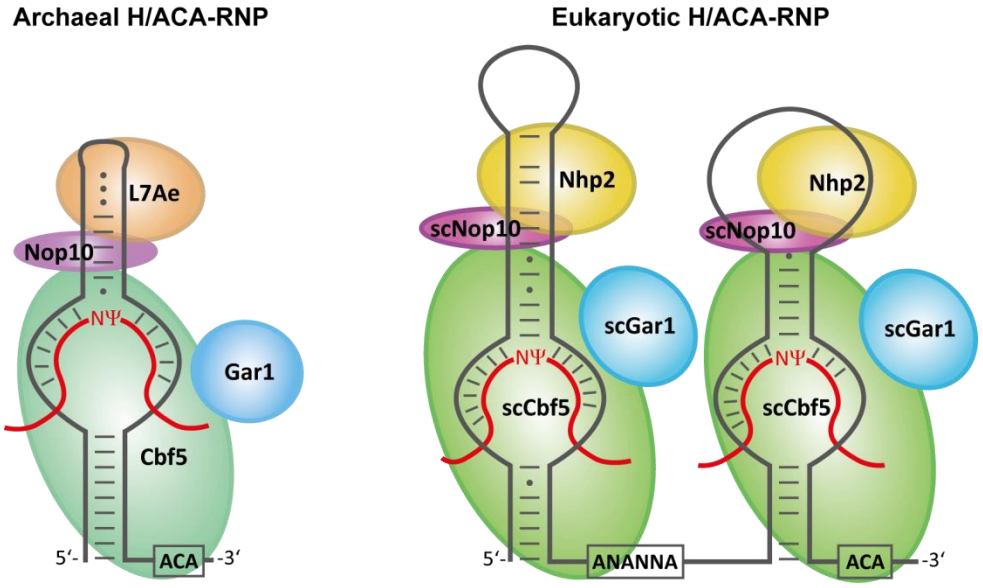


Figure 1.5: Schematic representation of an archaeal sRNP from *Pyrococcus furiosus* (left) and of a eukaryotic snoRNP from *Saccharomyces cerevisiae* (right). The H/ACA-RNA is shown in grey, the target RNA is shown in red, and the genetically conserved H- and ACA-boxes are highlighted.

Today's structural knowledge about H/ACA-RNPs mostly derives from X-ray studies of archaeal single hairpin complexes. Many partial as well as complete crystal structures of different archaeal complexes have been published and all show the same overall RNP architecture[52-56]. The H/ACA-RNA can form up to three hairpins, each containing a pseudouridylation pocket where the target RNA is recruited and released after the isomerization reaction. In the sRNP the three proteins L7Ae, Nop10 and Cbf5 directly interact to the guide RNA. L7Ae binds to the kink-turn motif located in the upper stem of the H/ACA-RNA, close to or within the loop structure[57]. Nop10, the smallest of all proteins, is anchored in between L7Ae, Cbf5 and the sRNA, while Gar1 is only attached to Cbf5. Nop10, Cbf5 and Gar1 are known to form a stable ternary protein sub-complex in archaea 'NCG' and eukaryotes 'scNCG'[58,59], but these complexes no catalytic activity in the absence of the H/ACA-RNA.

The actual Ψase Cbf5 inherits two major structural domains; the larger catalytic domain and the pseudouridine synthase archaeosine transglycosylase (PUA) domain at the C-terminus. The PUA domain is responsible for recognition and binding of the ACA-motif in the lower stem of the guide RNA and facilitates pseudouridylation[52,60]. In the catalytic domain the highly conserved aspartate – essential for Ψ-formation – and a so called 'thumb loop' region are sustained in archaeal and eukaryotic Ψases[59,60]. Due to the importance of the conserved aspartate, several different molecular reaction mechanisms for uridine isomerization have been proposed. One is based on a Michael-addition-like attack of the nucleophilic aspartate at the C6-atom of uridine that is polarized by one of the carbonyl groups. Subsequently the glycosidic bond is broken and after base rotation and removal of the aspartate residue the Ψ is formed[61]. The alternative would be an 'acylal- or glycal-mechanism' in which the nucleophilic aspartate attacks the C1'-atom of the ribose under removal of negatively charged uracil, and production of an acylal- or glycal-ribose-intermediate. Again the isomer is formed by rotation of the base, reconnection to the ribose and removal of the aspartate residue[62]. So far it is unclear which of these mechanisms is actually occurring during the modification[40]. However, before pseudouridylation can be performed the target uridine of the substrate RNA has to be placed into the correct position, which is the presumed purpose of the highly flexible thumb loop domain in Cbf5[63].

A critical role in the process of catalytic turnover plays Gar1, the only H/ACA-protein for which no direct interaction to the sno- or sRNA is reported until now. As the name suggests it inherits

many consecutive arginine-glycine-glycine (RGG) repeats. Archaeal Gar1 proteins lack these RGG-repeats, but the core structure of the protein is comparable to the eukaryotic version 'scGar1'[52,54,59]. Gar1 is attached to the catalytic domain of Cbf5 close to the functionally highly important thumb loop and binds it in the absence of substrate RNA[52]. Presumably, this interaction promotes recruitment of the substrate RNA into the pseudouridylation pocket. By use of fluorescence correlation spectroscopy it could be shown that Gar1 significantly accelerates product release[64]. Thus, it is assumed that after substrate binding the thumb loop of Cbf5 undocks from Gar1 and positions the target base into the active site. Subsequently, uridine isomerization leads to a decrease in thumb loop-substrate affinity and the loop rebinds to Gar1, and therefore facilitates product release[63]. Hence, it is not surprising that removal of this protein can lead to significant loss in catalytic activity[59].

Not directly involved in catalytic turnover but important for complex assembly are the loop binding protein Nhp2 in eukaryotes and its archaeal analogue the kink-turn binding protein L7Ae. Even though both proteins show high structural homology, they bind in different manners to the H/ACA-RNA[65]. L7Ae anchors the upper stem of the guide RNA by binding to a kink-turn motif which is either located in the loop ('kink-loop') of the sRNA or within close distance to the loop. The highly specific interaction between kink-turn and L7Ae – dissociation constants (K_d) in the low nM regime[57] – enables the protein to independently bind the guide RNA, and hence initiate H/ACA-RNP assembly[66]. Furthermore, different crystal structures of full and partial complexes show that the presence of L7Ae promotes accurate substrate placement in the sRNP[53,67], and absence of this protein leads to a substantial loss in pseudouridylation activity of the Ψase[59]. In contrast to L7Ae, the eukaryotic counterpart Nhp2 binds in a more unspecific way to the H/ACA-RNA and to RNA stem-loops in general. The loop of snoRNAs is typically larger than the loop of the archaeal sRNAs and does not inherit a specific binding motif like the kink-turn. The purpose of these relatively well conserved large loops in snoRNAs is not yet fully understood. But it is known that a sub class of snoRNAs called small Cajal body-specific RNAs contain a special binding motif the CAB box (UGAG) located in the terminal loops of 5'- and 3'-hairpins[68]. These CAB boxes direct modification of RNA polymerase II-transcribed spliceosomal snRNAs in the Cajal bodies[69,70]. Besides, a further difference of Nhp2 to L7Ae is the ability of the eukaryotic protein to form an Nhp2-scNop10-scCbf5 'WNC' sub-complex before RNA binding[71]. This sub-complex seems to promote proper RNP assembly, since a significantly higher binding

affinity (by three orders of magnitude) to the snR34 guide RNA of yeast could be observed for Nhp2 in presence of the other H/ACA-proteins[72]. The role of the smallest protein Nop10 in RNA guided pseudouridylation remains largely unknown. In archaea the binding interface from Nop10 to Cbf5 is more than double as big as to L7Ae, which is only bound via an N-terminal zinc ribbon (lacking in eukaryotic scNop10[73]) and a C-terminal α-helix of Nop10[52,54]. This difference in binding surfaces may account for the lack of stable L7Ae-Nop10 interaction, while the ternary NCG protein complex is readily formed.

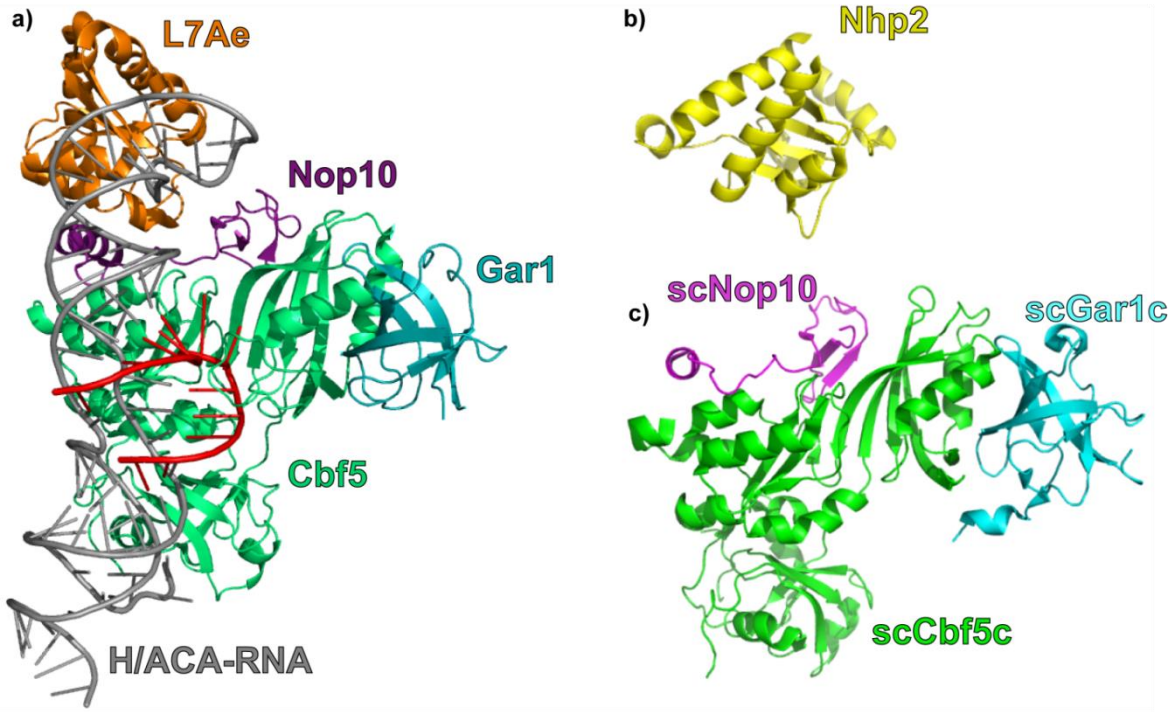


Figure 1.6: ***a)*** **Crystal structure of a complete substrate (red) bound archaeal H/ACA-RNP from** ***Pyrococcus furiosus*** **(PDB: 3HAY)[54].** ***b)*** **NMR structure of eukaryotic Nhp2 from** ***Saccharomyces cerevisiae*** **(PDB: 2LBX)[65].** ***c)*** **Crystal structure of the eukaryotic ternary scNCG complex, containing only the core domains scCbf5c and scGar1c from yeast (PDB: 3U28)[59].**

As mentioned above, knowledge about RNA guided pseudouridylation is mostly derived from X-ray studies of single hairpin H/ACA-RNPs from archaea like *Pyrococcus furiosus* (*P. furiosus*) '*pf*'. Figure 6a displays such a structure of a single hairpin complex. In eukaryotes, however, exclusively double hairpin snoRNPs are known. Structural data available for eukaryotic H/ACA-RNPs stems from: 1) a high resolution NMR structure of Nhp2 from *Saccharomyces cerevisiae* (yeast)[65] (see Figure 1.6b); 2) a 1.9 Å resolution crystal structure of the scNCG ternary protein complex from yeast; however, not the full-length constructs, but only the core domains of scCbf5 (scCbf5c) and scGar1 (scGar1c), which share high homologies to their archaeal relatives[59] (see Figure 1.6c); and most recently, 3) a cryo-EM structure (8.2 Å resolution) of a human telomerase

attached to a complete H/ACA-RNP (see Figure 1.7)[74]. The modest resolution of the cryo-EM structure allowed the authors to model already existing structure of an archaeal sRNP (Protein Data Base 'PDB': 2HVY[52]) into the observed electron density. This cryo-EM structure is the first to directly show an interaction of a chaperone protein to the snoRNP. The chaperone protein TCAB1 interacts with the 3'-hairpin, Dyskerin (Cbf5 analogue) and human GAR1. Chaperones are proteins that contribute to folding/unfolding or assembly/disassembly of macromolecular structures. It is known that chaperones like NAF1 assist in human snoRNP assembly[75,76], but are not essential for catalytic activity at least in lower eukaryotes. The evolutionary advantage of chaperones in snoRNP assembly remains to be unraveled. Furthermore, in this human telomerase bound H/ACA-RNP complex it was observed that the four core proteins bind the individual hairpins in a different manner. The 3'-hairpin of the snoRNA is anchored by human NHP2, human NOP10, Dyskerin and GAR1, similar to the archaeal H/ACA-RNP. The 5'-hairpin only binds to Dyskerin, remaining unclear if this assembly suffices for pseudouridylation activity. Besides, the authors could show a novel Dyskerin-Dyskerin interaction site, in which mutations have been linked to different diseases.[74]

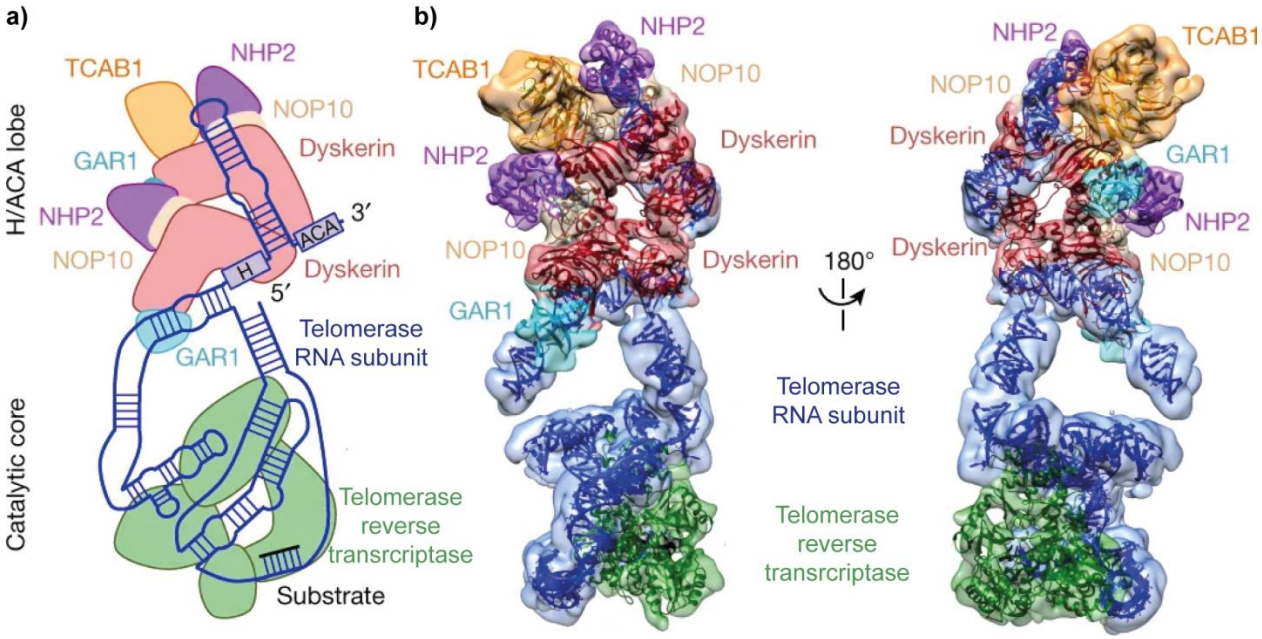


Figure 1.7: ***a)*** **Schematic representation of a human telomerase containing a complete H/ACA-RNP. b) Front (left) and back (right) views of cryo-EM structure from the catalytic core and H/ACA-lobe with fitted subunits (color coded). This Figure was modified from Nguyen *et al.*[74]**

Although H/ACA-RNPs are primarily known for pseudouridylation, more catalytic and structural functionalities of this enzyme class were reported: such as telomere maintenance, pre-mRNA splicing and ribosome biogenesis[77,78]. Hence, it is not surprising that malfunction of snoRNPs can

lead to lethal physiological effects. As mentioned above mutations in Dyskerin are correlated to diseases like Dyskeratosis congenita (DC) or the more severe Hoyeraal-Hreidarsson syndrome (HHS)[74,79,80]. DC and HHS are bone marrow malfunctions which can cause premature mortality[81]. Furthermore, snoRNAs and Ψ-formation have been linked to different types of cancer, viral diseases and neurodegenerative disorders[82]; highlighting the importance of this RNA class and its cellular functions. Hence, a full understanding of this enzyme class and of the processes during RNA guided pseudouridylation may contribute to find possible cures for these diseases and is of great scientific relevance. Modern techniques facilitating the investigation of these H/ACA-RNPs – as well as many other biomolecules – on a single molecular base will be introduced in the next chapter.

1.4. Methods in Single Molecule Spectroscopy

In the last century many effective spectroscopic methods like crystallography (X-ray), nuclear magnetic resonance (NMR), electron paramagnetic resonance (EPR) and fluorescence based approaches have been established to analyze biological systems and have significantly contributed to our knowledge of life. However, to some degree these techniques are based on ensemble type experiments, meaning that individual states and minor populations will inevitably be averaged over the full ensemble of observed molecules. Thus, a complete understanding of all biological processes is not achievable only by ensemble based measurements. Since large biomolecules often have an intrinsic instability or heterogeneity that might lead to different molecular behavior, methods are required that allow the observation of individual molecules to see how even small populations contribute to the entirety of cellular processes.

In the recent decade a lot of technical progress in single molecule observation has been made, especially in the field of fluorescence spectroscopy. The development of affordable super sensitive photo detectors and the ability to optically resolve molecular structures below the Abbe-limit (< 200 nm) by usage of different physical effects opened completely new possibilities for researchers to track molecular processes *in vitro* and directly in cells[83]. Resolution below the diffraction limit of 200 nm has been achieved by electron microscopy (EM) nearly 100 years ago. In EM, the sample is evaporated and imaged by electron irradiation. For structural analysis of biomolecules cryo-EM – a technique where the samples are cooled down to -150 °C – has become a popular tool among scientists[84]. Until today the resolution limit achievable with EM has been improved to even lower than 0.05 nm for inorganic crystals[85] and down to ~0.2 nm for large biomolecules[86]. Despite the high resolved images that can be obtained by cryo-EM, the harsh conditions needed for this method make it suboptimal for investigation of dynamics within the complex of interest and the freezing of the sample to very low temperatures prohibits the direct observation of biological processes *in vitro* and in living cells. Some of these problems can be overcome with single molecule fluorescence spectroscopy, as experiments can be performed at physiological conditions. Among others two popular approaches have been established in the recent two decades in the field of single molecule fluorescence spectroscopy; super resolution microscopy (SRM) can be used to track RNAs as well as other biomolecules directly in a cellular environment and to observe structural changes within the complex of interest or between several

biomolecules Förster Resonance Energy Transfer (FRET) is the method of choice for single molecule detection.

Super Resolution Microscopy

In SRM an optical resolution below the refraction limit of conventional light microscopy (~200 nm) is accessible by labeling the biomolecule of interest with bright organic fluorophores and usage of specific photophysical effects, which allow highly resolved images below 200 nm. Two general ways to perform SRM are commonly used among researchers. One way is to scan the sample with a confocal laser, while the effective excitation volume is reduced by a depletion laser[87]. This method is called Stimulated Emission Depletion Microscopy (STED). Another way is to illuminate the sample on a wide-field, while the fluorophores stochastically switch between a bright and a dark state. This switching of states is observed over a certain period of time and afterwards the center of each fluorophore can be determined precisely by computational evaluation. Popular methods utilizing these effects are Stochastic Optical Reconstruction Microscopy (STORM) and Photo-Activated Localization Microscopy (PALM)[88,89]. STED allows its user to obtain an optical resolution down to 20-70 nm with a millisecond to second temporal resolution, while STORM and PALM achieve a slightly higher spatial resolution with 10-30 nm, at the expense of a slower temporal resolution (seconds to minutes)[90]. However, each approach allows its user to track and image biomolecules in their native conditions *in situ* as well as in living cells, making them powerful tools for life scientists. Therefore, Stefan Hell and Eric Betzig as well as William Moerner were awarded with the Nobel Prize for Chemistry in 2014 for their key contributions in the development of SRM techniques[91-93].

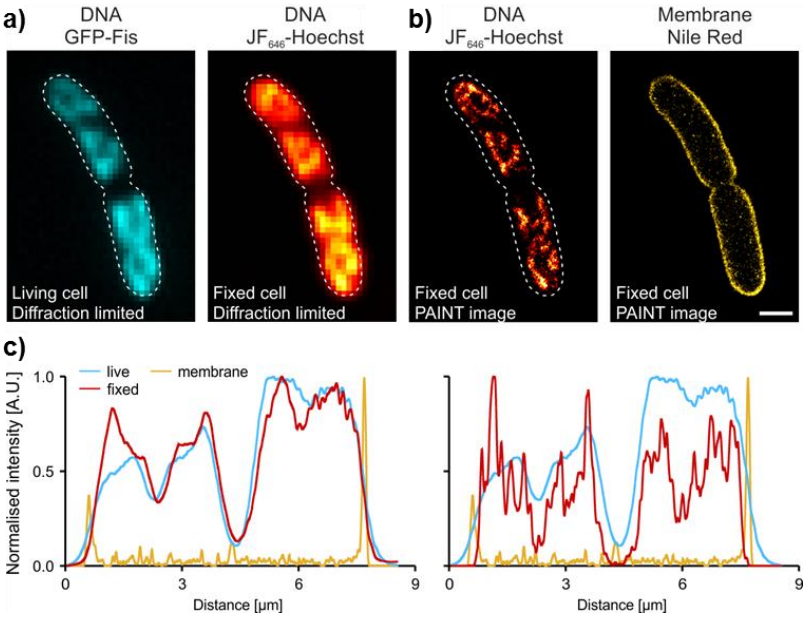


Figure 1.8: Comparison of diffraction limited microscopy versus super resolved STORM. *a)* Wide-field microscopy of GFP-Fis signal prior to fixation of *E. coli* (left) and JF_{646}-Hoechst signal using point accumulation for imaging in nanoscale topography 'PAINT' for DNA visualization after cell fixation (right). *b)* Super resolved STORM pictures of visualized DNA in fixed cell (left) and of the membrane using Nile Red (right). The scale bar (white) represents 1 µm. *c)* Normalized intensity profiles of the cell shown in the diffraction limited images (left) and of the same cell in the super resolved image (right). This Figure was modified from C. Spahn *et al.*[94]

Single Molecule FRET

FRET is a well-studied effect named after its discoverer Theodor Förster, which can be used to observe dynamics and interactions in biomolecular complexes. It is based on a radiationless energy transfer usually between two fluorophores called the FRET-pair, a FRET-donor and a FRET-acceptor. When the donor gets excited it transfers its energy to the FRET-partner resulting in detectable emission of the acceptor dye (see Figure 1.9a). The acceptor can also be a quencher molecule. There are two major requirements for this radiation less energy transfer to occur: One is a spectral overlap of the emissions spectrum of the FRET-donor to the absorption spectrum of the FRET-acceptor (see Figure 1.9b), the other requirement is a limited distance between both fluorophores. FRET can only occur in a defined range, usually 1-10 nm, depending on the FRET-pair and their electrochemical environment. The FRET-efficiency (E-FRET) strongly correlates with the intramolecular dye distance *r*, as well as other factors summarized in the Förster-Radius R_0 (Figure 1.9c). R_0 represents the distance a FRET-pair requires to obtain a 50% energy transfer (E-FRET = 0.5) from donor to acceptor fluorophore. The closer the fluorophores are the higher E-FRET is. Hence, determination of the FRET-efficiency allows its user to draw conclusions about structural changes and dynamics of labeled biomolecules. For decades FRET has been used

based on ensemble type experiments in biophysical assays, but in recent years single molecule FRET (smFRET) approaches have become popular among scientists to unravel cellular processes on the level of individual molecules.

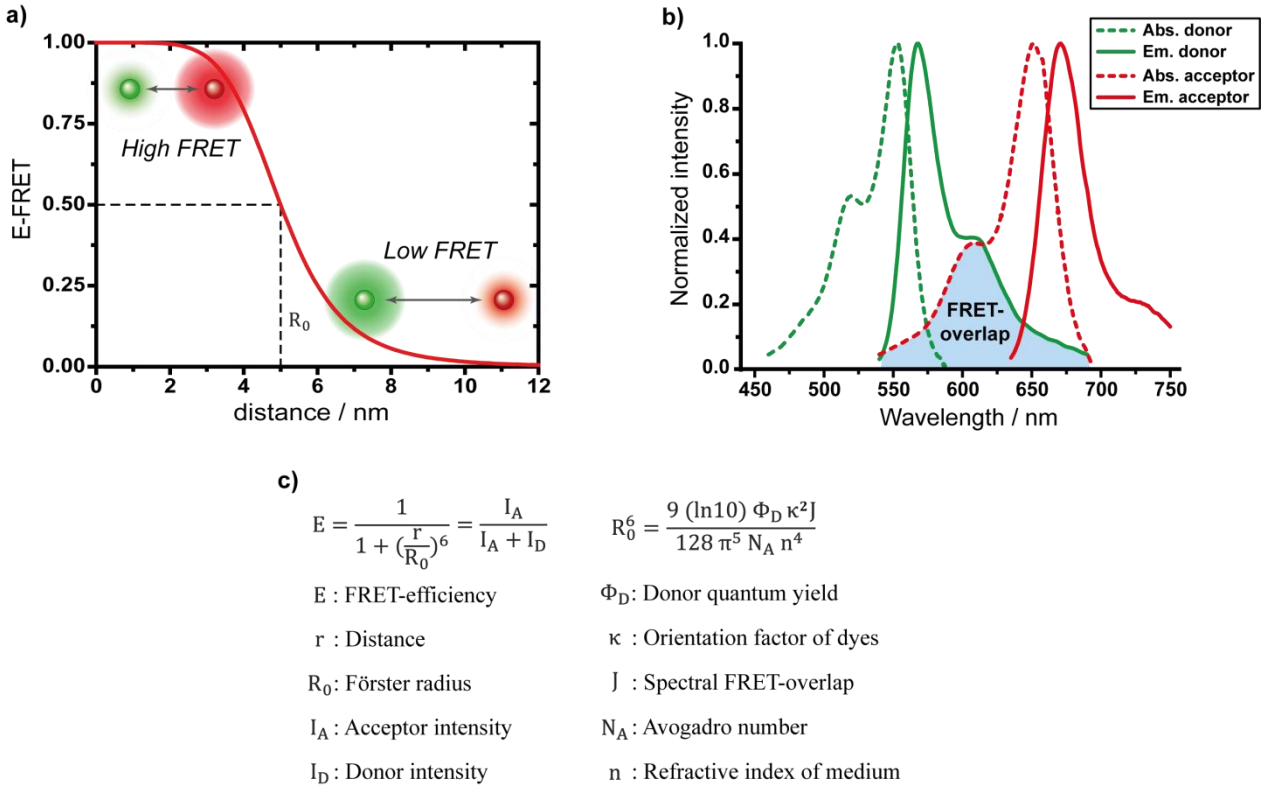


Figure 1.9: ***a)*** **E-FRET as function of fluorophore distance (r) for R_0 = 5 nm. A high FRET state occurs when donor (green) and acceptor (red) dyes are in close proximity. A low FRET state occurs when donor and acceptor dyes are in longer distance.** ***b)*** **Normalized emission and absorption spectra of donor and acceptor molecules. The required spectral overlap of donor emission and acceptor absorption is colored in blue.** ***c)*** **Equations for FRET-efficiency and the Förster radius.**

Two general methods have been developed to perform smFRET measurements: solution type or surface based microscopy. In confocal based solution type smFRET microscopy, an excitation laser beam is focused into a freely diffusing sample to illuminate a small volume (~10^{-15} L)[95]. Randomly transiting molecules get excited and the resulting fluorescence intensities of donor and acceptor dye are detected during the short time window in which the labeled complex diffuses through the excitation volume, allowing observation of molecules one at a time. The other approach is surface based smFRET microscopy, where the complex of interest is immobilized on a glass surface and illuminated in a wide-field setup. To improve the signal to noise ratio, total internal reflection fluorescence (TIRF) microscopy is commonly used with this method. In TIRF microscopy, the excitation beam is focused on the glass surface in a critical angle in a way that total reflection occurs. At this position an evanescent field is created that extends roughly 200 nm

from the surface into the sample medium, which is sufficient to excite the immobilized and labeled molecules at the illuminated spot. A major advantage of TIRF microscopy is the background suppression of the exciting laser, which leads to a significantly higher signal to noise ratio. After excitation, the fluorescence of FRET-donor and acceptor can be filtered into separate channels to locate each individual dye. With this technique several hundred molecules can be observed at the same time, which is the major advantage compared to confocal based solution type smFRET microscopy. Due to the immobilization of the complex of interest one can observe single molecules for several seconds or minutes and even up to hours[96], allowing its user to detect conformational changes and complex dynamics within the distance of the FRET-pair. However, the temporal resolution of TIRF microscopy is worse (limited to the millisecond range) compared to confocal solution type smFRET with sub millisecond time resolution[95]. Another disadvantage of wide-field based smFRET microscopy is the required attachment of a functional group into the macromolecule for immobilization purposes. In many cases this is not trivial because the extra modification must be added at a position where no undesired interactions between the modification and the complex may arise. This criterion also applies for the insertion of the two fluorophores. Hence, choosing labeling sites as well as labeling chemistry is not trivial. Site specific labeling at regions of interest is crucial to obtain reliable structural and dynamical data with smFRET. These sites have to be well chosen to avoid undesired interactions with the complex of interest along with reduction of biological activity. Additionally, both fluorophores have to be in a well-defined range sensitive to changes in FRET-efficiency. Furthermore, insertion of two or more modifications requires orthogonal labeling approaches, if the modifications should not be attached randomly. In the next chapter some considerations regarding choice of label, labeling sites and available labeling methods focused on RNA modification will be discussed.

1.5. Labeling Methods for RNAs

For any fluorescence based single molecule technique the RNA has to be labeled with suitable fluorophores. For biophysical applications these fluorophores have to fulfill several requirements like high brightness (extinction coefficient > 50,000 M^{-1} cm^{-1}; quantum yield > 0.1), photostability, minimal photophysical effects and susceptibility for chemical aggregation, high water solubility and small size to avoid steric hindrance[97]. For some applications usage of luminous fusion proteins like the famous green fluorescent protein (GFP) is well established, even for *in vivo* experiments[98]. However, low photostability and especially their relatively large size (> 4 nm) make fusion proteins unsuitable for RNA labeling. Usage of comparably small (< 1 nm) synthetic organic dyes for RNA labeling has become standard in single molecule fluorescence spectroscopy. The most popular organic fluorophores fulfilling the required attributes are the commercially available Alexa, Atto and Cy dyes (see Figure 1.10). High stability, spectral selectivity and accessibility to a multitude of chemical coupling reactions make these dyes a prime choice for labeling of RNAs and other biomolecules for fluorescence based approaches.

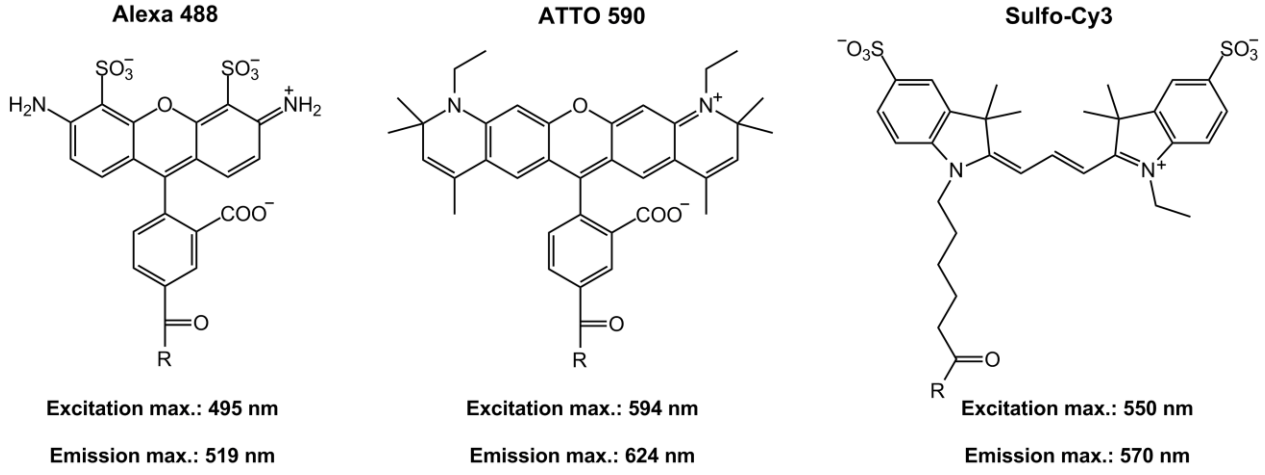


Figure 1.10: Chemical structures of representative commercially available dyes. Excitation and emission properties can be altered by variation of the delocalized π-system and substituents. Reactive functional groups for chemical coupling are typically introduced at R.

When labeling RNAs with fluorophores two major considerations have to be made:

1) Which type of labeling is necessary? Is covalent attachment of the label necessary or is a non-covalent hybridization approach of a small oligomer to the RNA of interest sufficient? While for smFRET experiments covalently attached dyes are obligatory in

some cases, many SRM applications require only the base pairing of a fluorophore carrying oligomer to the target RNA.

2) Which labeling sites can be chosen? This point is important for several reasons. For instance, structural and functional impacts of the label to the RNA should be minimized. If structural data of the RNA is available it should be utilized to specify the possible labeling sites. This is not only necessary to avoid negative influences on biological function, but also to avoid perturbations of photophysical properties of the fluorophore. Especially for smFRET measurements the exact positions and distance of the labels are essential for reliable conclusions about structure and dynamics of the observed complex.

In the following several well established, along with new labeling techniques, will be introduced and discussed.

Chemical Coupling Reactions

Covalent attachment of the dyes directly to the RNA of interest or to an oligomer for transient labeling is necessary in most cases for single molecule fluorescence spectroscopy. To ensure highly specific coupling between RNA and fluorophore, functional groups with a certain chemical selectivity are required at both reaction partners. Several bioorthogonal labeling approaches compatible with RNA folding and stability have been developed by chemists to ensure coupling without undesired side products (reviewed by Hanspach *et al.*)[99].

A well-established labeling reaction in chemical biology is the copper-catalyzed azide-alkyne cycloaddition 'CuAAC', one of the so called 'click reactions'[100]. Both functional groups can be attached to either the fluorophore or the RNA (see Figure 1.11a). The driving force in this reaction is the thermodynamic stability of the coupling product, a 1,2,3-triazole. Despites its bipolarity the azide is not very active at room temperature. To perform the reaction at moderate conditions an initiator is needed to activate the 1,3-dipole (azide). A common possibility to start the reaction is the usage of an *in situ* generated Cu(I)-catalyst[101]. However, to generate stable Cu(I) in the reaction mixture, it has to be either reduced from Cu(II) or oxidized from Cu(0). These redox conditions might lead to undesired side reactions at the biomolecule, and the fact the Cu is cytotoxic in the needed concentrations, limit the range of applications for biological systems. Another possibility to initiate coupling between azide and alkyne is the strain promoted [3+2]-cycloaddition 'SPAAC'[102] (see Figure 1.11b). This reaction requires the alkyne to be in a

closed ring system; usually 8-rings like cyclooctynes are used. Due to the linear geometry of the triple bond the cyclic ring is put under steric strain. This tension suffices to activate the triazole formation at physiological conditions. Since no cytotoxic metal catalysts as well as oxidative reaction conditions are needed, this labeling approach is compatible for *in vivo* applications[103,104]. However, since synthesis of cyclooctynes typically gives low yields and is more complex compared to regular alkynes, the strain promoted labeling is not as well as established as the Cu-catalyzed approach.

A different strain promoted coupling method is the reaction of 1,2,4,5-tetrazines or 1,2,4-triazines with cyclic alkenes and alkynes[105] (see Figure 1.11c). The reaction mechanism follows a hetero Diels-Alder [4+2]-cycloaddition with inverse electron demand 'IEDDA', where the tetra-/triazine is the diene and the cyclic alkene/alkyne the dienophile, followed by a retro Diels-Alder reaction with elimination of nitrogen gas. The omission of N_2 leads to the formation of a stable hetero 6-ring, prevents reversibility of the Diels-Alder reaction, and the absence of metal catalyst and redox agents make this labeling approach suitable for in cell experiments[106,107].

Another well-established labeling method is based on the introduction of primary amine groups into the RNA and usage of active *N*-Hydroxysuccinimide (NHS)-esters as coupling reagent[108] (see Figure 1.11d). The NHS-ester does not react with the natural occurring amines of the nucleotides and is selective for the introduced functional group. The free electron pair of the primary amine donates into the partially positive carbon atom of the carboxyl group and forms a highly stable amide bond under elimination of NHS. Considering NHS is a good leaving group and not as acidic as other potential leaving groups like protonated halides (HCl, HBr, etc.), the reaction can be performed under physiological conditions and proceeds quickly. However, since primary amines occur frequently in cellular components (e.g. the amino acid lysine in proteins), this reaction is not suitable for specific *in vivo* labeling.

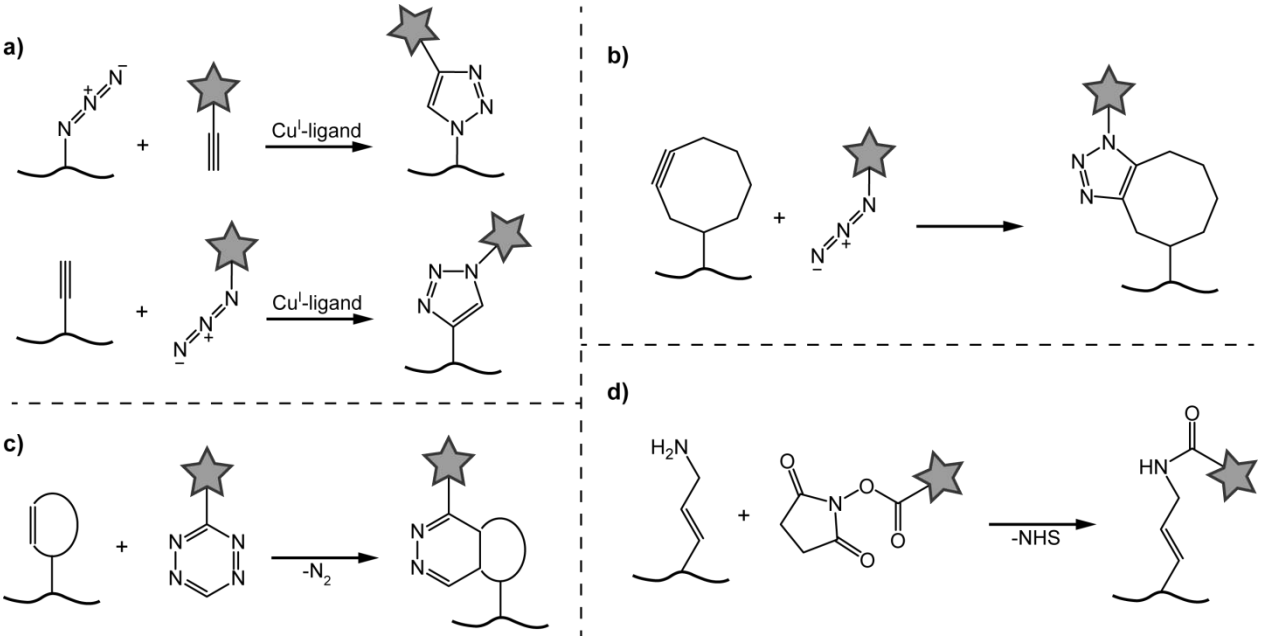


Figure 1.11: Chemoselective labeling reactions: *a)* copper catalyzed alkyne-azide cycloaddition, *b)* strain promoted alkyne-azide cycloaddition, *c)* tetrazine-alkene hetero Diels-Alder cycloaddition and *d)* reaction of a primary amine to an NHS-ester.

The above mentioned labeling reactions give scientist a range of possibilities to site-specifically insert several modifications into one RNA or other biomolecules. Many different fluorophores with a variety of functional groups (e.g.: alkyne, azide, NHS-ester, tetrazine, cyclooctynes, etc.) are commercially available, removing the effort of self-synthesizing the reactive dyes[99]. The most common way to introduce modification into DNAs or RNAs is by solid phase synthesis. This technique is based on step by step incorporation of modified nucleosides via, e.g. phosphoramidite chemistry, and allows site-specific introduction of labels at sugar, base or phosphate backbone of the nucleic acid. Ongoing side reactions during the actual synthesis limit DNA and RNA length to ca. 100 nt[109,110]. Hence, for longer nucleic acids other labeling approaches have to be considered.

Splinted Ligation

A convenient and extensively investigated enzymatic method to overcome the size limit of solid phase synthesis is splinted ligation of two or more oligomers with T4-DNA-ligase or T4-RNA-ligase 2[111,112]. For effective enzymatic ligation the RNA oligomers have to be hybridized to a complementary DNA strand to form a stable DNA:RNA-hybrid. Then, the hydroxyl at the 3'-end of one oligomer is attached enzymatically to a monophosphate at the 5'-end of the neighboring oligomer. By combination of solid phase synthesis and of splinted

ligation highly modified RNAs can be obtained. Several considerations have to be made to guarantee high ligation efficiencies. The overlap of the complementary DNA strand at the linking site should be long enough (> 10 base pairs 'bp') for both RNAs to form a stable double stranded DNA:RNA-hybrid. Large labels like organic fluorophores should not be attached too closely – ideally more than 3 nt away – to the ligation site, since this might reduce catalytic activity of the ligation enzyme, due to steric hindrance. T4-DNA-ligase is more common than T4-RNA-ligase 2 as ligation enzyme. However, both ligases have individual advantages and disadvantages. Which ligase works best oftentimes has to be tested for each ligation system individually.

5'-End Functionalization

A chemical approach to label RNAs is based on the functionalization of a monophosphate at the 5'-end by usage of 1-Ethyl-3(3-dimethylaminopropyl)carbodiimide (EDC). As hydrochloride EDC is highly water soluble and ready to activate carboxylic acids as well as organic phosphates[113]. The OH-group of the phosphate reacts with the carbodiimide to form an active ester intermediate. By addition of imidazole and ethylenediamine under mild acidic conditions, this intermediate is converted into a phosphoramidite with a primary amine group, which can be used for further modification with NHS-ester chemistry. For this type of functionalization RNAs carrying a monophosphate the 5'-end are required. However, standard T7 RNA-polymerase *in vitro* transcription with nucleotide triphosphates (NTPs) yields RNAs with a triphosphate at the 5'-end. The monophosphate can be introduced either by *in vitro* transcription in excess of guanosine monophosphate, when the first nucleotide of the RNA of interest is guanosine, or by enzymatic dephosphorylation of the triphosphate at the 5'-end and subsequent phosphorylation. This method allows convenient labeling of larger quantities for small and large RNAs, but only at the 5'-end. Additionally, further purification steps have to be considered since degradation of the RNA might be an issue during functionalization.

Enzymatic 3'-Labeling

Modification of the 3'-end of a transcribed RNA can be achieved by usage of T4 RNA ligase 1 and a nucleoside 3',5'-bisphosphates. T4 RNA ligase 1 is capable of attaching nucleotides and oligomers to single stranded RNAs[114,115]. Usage of 3',5'-bisphosphates ensures single nucleotide extension of the RNA, since the resulting 3'-phosphate will block further incorporation of other nucleotides. The attached nucleotide can carry a wide range of

modifications at phosphate, sugar or base-site for further labeling approaches. Further modification of the RNA can be achieved by removal of the 3'-phosphate with shrimp alkaline phosphatase, subsequent purification and addition of a second RNA strand using splinted ligation[115]. This way the 3'-modification can be turned into an internal modification, facilitating a high scope of internally labeled RNAs in various lengths.

Cotranscriptional Labeling

So far the above-mentioned approaches for RNA labeling are using chemical or enzymatic modification of the RNA after its synthesis. A relatively new method called Position-Selective Labeling of RNA (PLOR) allows cotranscriptional insertion of labeled nucleotides, overcoming the size limit (~100 nt) of solid phase synthesis. PLOR is based on the 'pause-restart' ability of T7 RNA polymerase. One way to achieve transcription stalling is omission of one NTP, and can be restarted by addition of the missing NTP[116]. By PLOR this ability can be exploited to synthesize high amounts (up to milligrams) of site and position-specifically labeled RNAs[117]. The process of RNA synthesis is divided into three general stages: initiation, elongation and termination. For initiation, the DNA template is immobilized on beads (solid phase) via biotin-streptavidin interaction. Transcription is started in the absence of one NTP, in which case a stable DNA-RNA-polymerase complex is formed at the position of the missing NTP. The solid phase with its stalled DNA-RNA-protein complex is extracted several times to remove all initial NTPs. After the first pausing the elongation phase starts. A mixture of modified and/or unmodified NTPs can be added, still missing one certain NTP for later transcription pausing. The modified NTPs will be inserted by the polymerase at the respective position or positions according to Watson-Crick base pairing until the polymerase comes to the point where the missing NTP should be inserted and pauses again. By repetition of the washing procedure and a wise choice of NTPs, the elongation step can be repeated several times until all of the desired modifications are inserted. The final step is the termination of the transcription and purification of the RNA. For the final elongation a reaction mixture without GTP – to avoid reinitiation of T7 RNA polymerase – is added, incubated and cooled down to stop transcription. Finally, the RNA can be extracted from the solid phase, collected and purified. The key step for this method is the washing procedure. If it is not performed thoroughly even minor remains of NTPs will lead to undesired transcription elongation, yielding unlabeled RNAs. PLOR is powerful tool to insert a large variety of modifications for many applications into RNAs of any size[118]. If becoming

commercially available, PLOR might develop to the method of choice for RNA labeling in the near future.

Chapter II

Motivation and Objectives

H/ACA-RNPs are RNA-protein complexes which can be found in archaea and eukaryotes. Their major function is the isomerization of U to Ψ in rRNAs and snRNAs. Ψ-introduction of H/ACA-RNPs occurs in a sequence-specific manner, in which the guide RNA binds the substrate and a certain target uridine is then isomerized. This modification is known to be important for proper ribosome and spliceosome function[43], and thus crucial for protein biosynthesis. Hence, it is not surprising that malfunction of H/ACA-RNPs has been linked to severe diseases like different types of bone marrow failures[74,79,80], and overexpression of snoRNAs was associated to various kinds of cancer[82]. Therefore, a deeper understanding of this enzyme class might help researchers to find potential cures for these lethal diseases.

The structure of archaeal single hairpin H/ACA-RNPs is well studied. Several partial and full crystal structures yielded many insights about the biological function of H/ACA-RNAs and proteins[52-56]. Less is known about the double hairpin structured eukaryotic H/ACA-RNP (see Chapter 1.3). Furthermore, knowledge about the processes of complex assembly, reactional dynamics and structure in solution is scarce. This holds true for the eukaryotic as well as for the archaeal RNA based Ψase. The aim of this thesis was to gain more knowledge about structural dynamics of H/ACA-RNPs and RNA guided Ψ-formation, which cannot be obtained by x-ray studies, since this method lacks the ability to observe conformational changes in solution. For that reason and in order to observe potential dynamics, surface based smFRET spectroscopy was the method of choice. Additionally, this technique allows the differentiation of individual states that might be lost by averaging in ensemble based methods.

This thesis is divided into two projects: Preparation and investigation of

1) an archaeal H/ACA-RNP from *Pyrococcus furiosus*, and of
2) a eukaryotic H/ACA-RNP from *Saccharomyces cerevisiae*.

The focus in these projects was on structural changes within the sRNA or snoRNA. Therefore, guide RNAs and substrate RNAs were individually marked with fluorophores. Starting with the thermophile RNA guided Ψase from archaea; RNP-assembly, substrate recruitment and substrate turnover were to be elucidated with smFRET. Complex formation was to be investigated in a step-by-step fashion via addition of individual proteins to differently labeled sRNAs. In that way, observation of overall structural rearrangements in the guide RNA should be made accessible.

Furthermore, impacts of individual proteins on substrate recruitment were to be investigated via fluorophore labeled target RNAs and omission of single proteins. And, in order to examine different states during catalytic turnover, various target RNAs with chemically modified target nucleotide were used. These target RNAs were then either labeled with a fluorophore or biotin to observe structural changes at the substrate or guide RNA site. For these purposes, preparation of labeled and active H/ACA-RNPs had to be established and optimized, and was followed up on previous results and achievements of my master thesis[119].

Following the successful establishment of both preparation and smFRET experiments of the archaeal constructs, it was aimed to start investigation of the second project. The eukaryotic snoRNP, consisting of the snR81 RNA and H/ACA-proteins from *Saccharomyces cerevisiae*, was used as model system. The snR81 RNA is known to direct the formation of Ψ42 – as well as Ψ93 under stress conditions – in the U2 snRNA (5'-hairpin target) and of Ψ1051 in the 25S rRNA (3'-hairpin target)[120,121]. Reconstitution of snoRNPs is not trivial, and successful *in vitro* formation of catalytically active eukaryotic H/ACA-RNPs has been reported only twice[59,72]. Moreover, the lack of high resolution structures even of any snoRNP substructure complicates the search for adequate fluorophore attachment sites to the snoRNA. Thus, preparation and reconstitution of active wildtype and labeled H/ACA-RNPs from yeast had to be established at first. The impacts of individual proteins and RNA domains on pseudouridylation activity were of interest. snoRNP formation should be compared to archaeal complex assembly via smFRET. And, due to the fact that the eukaryotic Ψase is functional at modest temperatures, initial activity experiments allowing direct observation of pseudouridylation events under the microscope were aimed for. To begin with this project, single hairpin constructs (5'- and 3'-hairpin) carrying only one set of proteins instead of the full-length bipartite complex were to be used for initial studies, since it is known that also partial snoRNPs are capable of Ψ-formation[59,72]. Here, emphasis was placed on the 5'-hairpin of the snR81 complex. Another focus was set on the comparably large and well conserved loop sequences of snoRNAs. To investigate the impact of these loops on complex formation and activity, two constructs for each stand-alone hairpin were designed; a full-length version, containing the complete loop sequence, and a truncated hairpin, in which the large loop was substituted by a tetra loop. For all these purposes; each construct had to be designed, synthesized and tested for biological function in order to gain new insights about eukaryotic pseudouridylation.

Chapter III

Material and Methods

3.1. General Methods

3.1.1. Buffers and Media

A selection of frequently used buffers and media is given in this chapter. All buffers were sterile filtered (0.22 µm) before use. Buffers for chromatographic applications were degassed via ultra sound (15 - 30 min) prior usage. Media for cell culturing were autoclaved (121 °C, 15 min, 2 bar) after preparation. Antibiotics (Ampicillin 'Amp' and Kanamycin 'Kan') were prepared as 1000x stock solution in H_2O and sterile filtered (0.22 µm). Working concentrations were 100 µg/mL in case of Amp and 30 µg/mL in case of Kan.

Media

LB		**SB**		**TB**	
5 g/L	Yeast extract	20 g/L	Yeast extract	24 g/L	Yeast extract
10 g/L	Tryptone	35 g/L	Tryptone	12 g/L	Tryptone
10 g/L	NaCl	5 g/L	NaCl	5 g/L	Glycerol
				17 mM	KH_2PO_4
				72 mM	K_2HPO_4
				(phosphates added after autoclaving)	

Buffers for Preparation

IMAC-A		**IMAC-B**		**nSEC**	
1 M	NaCl	1 M	NaCl	1 M	NaCl
20 mM	Imidazole	500 mM	Imidazole	40 mM	NaH_2PO_4
50 mM	NaH_2PO_4	50 mM	NaH_2PO_4	10 mM	Na_2HPO_4
14 mM	β-ME	14 mM	β-ME	pH 7.6	(HCl adjusted)
pH 6.0	(HCl adjusted)	pH 6.0	(HCl adjusted)		

oSEC		**sSEC**		**TEAA**	
500 mM	KCl	500 mM	KCl	100 mM	Triethylamine
50 mM	HEPES	20 mM	HEPES	100 mM	AcOH
10% (v/v)	Glycerol	20% (v/v)	Glycerol	pH 7.0	(AcOH adjusted)
pH 7.0	(KOH adjusted)	pH 7.5	(KOH adjusted)		

scA		scB		scSEC	
500 mM	NaCl	500 mM	NaCl	500 mM	NaCl
50 mM	Tris	50 mM	Tris	50 mM	Tris
20 mM	Imidazole	500 mM	Imidazole	10% (v/v)	Glycerol
10% (v/v)	Glycerol	10% (v/v)	Glycerol	pH 8.0	(HCl adjusted)
pH 8.0	(HCl adjusted)	pH 8.0	(HCl adjusted)		

Assay buffers

Ψ-assay		scΨ-assay		P1-digestion	
100 mM	NH_4OAc	250 mM	NaCl	20 mM	NH_4OAc
100 mM	Tris	12.5 mM	Tris	0.5 mM	$ZnCl_2$
5 mM	$MgCl_2$	pH 8.0	(HCl adjusted)	pH 5.3	(AcOH adjusted)
pH 8.0					

50x Gloxy		Imaging		scImaging	
50 mM	NaCl	2% (v/v)	50x Gloxy	2% (v/v)	50x Gloxy
10 mM	Tris	10% (w/v)	D-Glucose	10% (w/v)	D-Glucose
4 mg/mL	Glucose oxidase	0.1 mM	Dodecylmaltoside	saturated	Trolox
1 mg/mL	Catalase	saturated	Trolox	in	1x scΨ-assay
pH 8.0	(HCl adjusted)	in	1x Ψ-assay		

3.1.2. Standard Protocols

Plasmid Transformation

For transformation of plasmids into competent *Escherichia coli* (*E. coli*) cells, 50 µL frozen cells were thawed for 20 min on ice. Plasmid (1 – 100 ng) were added to the cells and incubated for 15 min on ice. Then a heat shock was performed for exactly 45 sec at 42 °C. Afterwards, cells were put on ice for at least two minutes, 500 µL SOC media was added and the suspension was shaken at 400 rpm at 37 °C for 1 h. Typically 150 – 300 µL of the cell suspension were spread on LB-agar plates (25 g/L LB media and 15 g/L agar) containing the required antibiotics and incubated at 37 °C overnight or at 20 °C for 48 h. Cell colonies were identified per eye and marked for further experiments.

Ethanol Precipitation

Precipitation of nucleic acids was done by adding NH_4OAc or NaOAc to the aqueous RNA solution to a final concentration (c_{final}) of 0.5 M or 0.3 M, respectively. Then 2.5 volume equivalents (vol. eq.) pure EtOH were added to the RNA solution and the mixture was stored at -80 °C for 30-60 min or at -20 °C overnight. Samples were centrifuged for 30-45 min at 17000 g, the supernatant was removed and the remaining pellet was dried in a vacuum concentrator.

Phenol Extraction

Separation of nucleic acids and proteins was done by phenol extraction. Therefore, one vol. eq. water-saturated aqua-phenol or one vol. eq. TE-buffer-saturated (pH 7.5-8.0) phenol/chloroform/isoamylalcohol (25:24:1) 'PCI' was added to the aqueous RNA solution. The mixture was vortexed thoroughly and centrifuged for 1 min at 2000 g. The aqueous phase was separated from the organic phase, 3 vol. eq. water-saturated Et_2O were added, vortexed thoroughly and centrifuged for 30 sec at 2000 g. The organic phase was removed and with the remaining aqueous solution an ethanol precipitation was carried out.

3.1.3. Gel Electrophoresis

Analytical Agarose Gel (DNA)

Analytical scale agarose gel electrophoresis was performed on a Mini-Horizontal Unit Electrophoresis Systems (FisherBiotech). Typically, 25 mL 1.0 – 2.5% (w/v) agarose solution in TAE-buffer (40 mM Tris, 20 mM acetic acid, 1 mM ethylenediamintetraacetic acid (EDTA), pH 8.0) were prepared. For analytical scale electrophoresis 10 µL samples containing ~100 ng DNA were mixed with 2 µL 6x gel loading dye (New England Biolabs 'NEB') and loaded adjacent to 3 µL 2-log DNA ladder (1 µg/µL, 0.1-10.0 kb, NEB). Typically, the gel was run at constant 120 V for 45 min at room temperature. DNAs were stained using GelRed™ (Biotium) and visualized via UV-illumination.

Preparative Agarose Gel (DNA)

Preparative scale agarose gel electrophoresis was performed on a Mini-Horizontal Unit Electrophoresis Systems (FisherBiotech). Typically 150 mL 1.0 – 2.5% (w/v) agarose solution in TAE-buffer containing 0.001% (v/v) 10000x GelRed™ (Biotium) were prepared. For preparative

scale 100 µL samples were mixed with 20 µL 6x gel loading dye (New England Biolabs). Typically, the gel was run at constant 120 V for 45 min at room temperature. Desired DNA bands were detected by eye or visualized with a Dark Reader Blue Light Transilluminator (LABGENE Scientific), excised by hand and purified using the QIAquick Gel Extraction Kit (QIAGEN).

Denaturing Polyacrylamide Gel (RNA)

Analytical denaturing PAGE was performed on a Multigel system (Biometra). For analytical electrophoresis 3.5 – 15% (w/v) polyacrylamide (PAA) solutions were prepared from a 30% acryl-/bisacrylamide(29:1) stock solution, 10xTBE-buffer (c_{final} (1x): 90 mM Tris, 90 mM boric acid, 2 mM EDTA, pH 8) and urea (c_{final} = 7 M). 10 mL of the respective solution were polymerized by addition of 100 µL 10% (w/v) ammonium persulfate and 10 µL TEMED. Typically, 5 µL samples containing ~100 ng RNA were mixed with 5 µL FA-buffer (90% (v/v) formamide, 90 mM Tris, 90 mM boric acid, 2 mM EDTA, 0.1% (w/v) bromophenol blue and 0.1% (w/v) xylene cyanol) and heated to 98 °C for 3 min. For samples containing fluorophore labeled RNAs bromophenol blue and xylene cyanol were omitted from the formamide buffer. The analytical denaturing PAGE was performed in TBE-buffer at typically 200 V for 30 min. Unlabeled RNAs were stained using GelRed™ (Biotium) and visualized via UV-illumination. Fluorophore labeled RNAs were visualized via fluorescence detection on a Typhoon 9400 (GE) imager. Applied Typhoon settings are given in Table 3.1.

Typhoon settings	Laser excitation	Emission filter
Cy3-Scan	532 nm	580 nm
Cy5-Scan	633 nm	670 nm
FRET-Scan	532 nm	670 nm
Phosphor screen	633 nm	390 nm

Table 3.1: Typhoon settings for detection of fluorophores or phosphor storage screen scanning. Sensitivity and PMT voltage were adjusted individually.

Native Polyacrylamide Gel (RNA/RNP)

Analytical native PAGE was performed on a Multigel system (Biometra). For analytical electrophoresis 8% (v/v) polyacrylamide solutions were prepared from a 30% acryl-/bisacrylamide(29:1) stock solution and 10xTBE-buffer (final concentration: 1x). 10 mL of the respective solution were polymerized by addition of 100 µL 10% (w/v) ammonium persulfate and 10 µL TEMED. After sample annealing, 0.2 vol. eq. of 60% (v/v) glycerol was added to each

sample and the samples were loaded onto the gel. The analytical native PAGE was performed in TBE-buffer at either constant 70 V for 90 min at room temperature (*Pyrococcus furiosus* constructs) or at constant 60 V for 180 min at 4 °C (*Saccharomyces cerevisiae* constructs). Unlabeled RNAs were stained using GelRed™ (Biotium) and visualized via UV-illumination. Fluorophore labeled RNAs were visualized via fluorescence detection (see Table 3.1).

Sodium Dodecyl Sulfate Polyacrylamide Gel (Protein)

For electrophoretic analysis of proteins sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) was performed on Mini-PROTEAN Tetra Cell-chambers (BioRad). The gel was casted in a discontinuous manner: 1) the loading gel (5% acryl-/bisacrylamide(29:1), 250 mM Tris/HCl, 0.1% (w/v) SDS, pH 6.8) and 2) the resolving gel (15% acryl-/bisacrylamide(29:1), 375 mM Tris/HCl, 0.1% (w/v) SDS, pH 8.0). Gel polymerization was achieved by addition of 0.1% (v/v) TEMED and 0.1% (w/v) ammonium persulfate. The upper third of the gel consisted of the loading gel and the lower 2/3 of the resolving gel. Typically, 5 µL samples were mixed with 5 µL SDS-loading-buffer (100 mM Tris/HCl, 4% SDS, 2 mM β-ME, 20% Glycerol, 0.1% (w/v) bromophenol blue, pH 6.8), heated for 3 min at 98 °C and loaded adjacent to 7.5 µL LMW-SDS Marker Kit (MW/kDa: 14.4, 20.1, 30.0, 45.0, 66.0, 97.0) (GE). The electrophoretic protein separation was performed in SDS-running-buffer (25 mM Tris/HCl, 250 mM glycine, 0.1% SDS, pH 6.8) at constant 220 V for 40 min. Protein visualization was achieved by microwaving the gel for 2 min in ~30 mL SDS-staining solution (0.1% (w/v) Coomassie Brilliant Blue, 10% (v/v) EtOH, 5% (v/v) AcOH) and waiting for an additional 15 min. For better contrast the gel was destained in H_2O overnight.

3.1.4. Concentration Determination of Biomolecules via UV/Vis-Spectroscopy

Concentration of proteins, DNAs, RNAs and fluorophores were determined by UV/Vis-spectroscopy on a NanoDrop 1000 or NanoDrop One spectrophotometer (ThermoFisher) according to the Beer-Lambert law. The wavelength at absorbance maximum (λ_{max}) was used for concentration calculation. Utilized extinction coefficients (ε) are given in Table 3.2.

Nucleic acids	**λ_{max} / nm**	**ε / (µL•ng^{-1}•cm^{-1})**
DNA	260	1/50
RNA	260	1/40

Proteins*	**λ_{max} / nm**	**ε / (M·cm^{-1})**
Archaeal NCG	280	67840
Eukaryotic scNCG	280	63550
Nhp2	280	9970
Fluorophores	**λ_{max} / nm**	**ε / (M·cm^{-1})**
Cy3	550	150000
Cy5	650	250000

Table 3.2: Extinction coefficients for concentration determination by UV/Vis-spectroscopy. *ε of proteins was determined via the online tool ProtParam (ExPASy).

3.1.5. Preparation of Radioactively Labeled RNAs

Target RNA transcription with stochastically incorporated α-^{32}P-UTP for preliminary single turnover experiments was performed as previously described[119]. For site specific target nucleotide labeling (with the radioactive isotope ^{32}P), the target RNAs were divided into two oligomers. Sequences of the purchased oligomers (Eurofins Genomics) are given in Table 3.3.

Archaeal substrate		**5'-3'-Sequence**
pf Sub	5'-oligo	GAUGGAGCG
	3'-oligo	***U***GCGGUUUAAU
Eukaryotic substrates		**5'-3'-Sequence**
5'-Sub	5'-oligo	GGGAGUAGUAUC
	3'-oligo	***U***GUUCUUUUCAG
3'-Sub	5'-oligo	GGGAACUU
	3'-oligo	***U***AAAUAUGUAAGAA

Table 3.3: Substrate oligomers for site specific isotope labeling. The target uridine for radioactive labeling is highlighted in bold and italic.

Isotope Labeling

Radioactive 5'-end phosphorylation of the target uridine (first nucleotide of the 3'-oligo) was done with ^{32}P-γ-ATP (Hartmann Analytic) and T4-Polynucleotidekinase (NEB). Therefore, 5 nmol of the 3'-oligo of each construct (*pf* Sub, 5'-Sub and 3'-Sub) were incubated individually with 3 – 4 µL 3.33 µM ^{32}P-γ-ATP (c_{final} = 0.4 µM) and 20 units T4-Polynucleotidekinase (c_{final} = 0.67 units/µL) in 1xT4-PNK-buffer for 1 h at 37 °C in a 30 µL reaction scale. Then, non-radioactive ATP was added to a final concentration of 10 mM (final volume: 40 µL) and incubated for 1 h at 37 °C. Afterwards, unincorporated ATP was removed from the 3'-oligos using illustra MicroSpin G-25 columns (GE).

Ligation & purification

The purified 3'-oligos were then attached to the respective 5'-oligo in a DNA splinted enzymatic ligation. Therefore, the radioactively labeled 3'-oligos (~ 5 nmol) were annealed with 5 nmol of its corresponding 5'-oligo and DNA ligation template at equimolar concentration (50 µM) in 0.5xT4-DNA-ligase-buffer by heating the sample to 85 °C for 5 min, cooling down to room temperature for 15 min, and were then put on ice. After the annealing, the buffer was set adjusted 1-fold concentration, T4-DNA-ligase (NEB, 2000 units/µL) was added to a final concentration of 80 units/µL and the reaction mixture was incubated at 16 °C and 500 rpm overnight. TURBO^{TM} DNase (Invitrogen) was added (c_{final} = 0.11 units/µL) and incubated at 37 °C for 1 h. One vol. eq. of FA-buffer was added; samples were heated to 98 °C for 3 min and loaded on a denaturing 15% polyacrylamide gel (7 M Urea, TBE-buffer) in a Maxigel system (Biometra). The gel was typically run at 300 V for 3 h. Radioactive RNAs were visualized by autoradiography with a storage phosphor screen (GE) by exposing the gel for 2 min to the screen (*Note*: The gel was covered with plastic wrap to ensure that the gel does not directly touch the surface of the screen). The screen was scanned with a Typhoon 9400 (GE) imager (see Table 3.1). Ligated products were excised from the gel and eluted in 0.3 M NaOAc at 20 °C and 600 rpm overnight. Gel residues were removed with centrifugal filter units (0.45 µm) and eluted RNAs were ethanol precipitated (0.3 M NaOAc). Finally, target RNAs were resuspended in H_2O and the concentration was determined by UV/Vis-spectroscopy.

3.2. The Archaeal H/ACA-Ribonucleoprotein Complex

3.2.1. Cloning of Codon Optimized Genes into pET-Expression Platforms

Plasmid amplification and digestion

E. coli codon optimized synthetic genes of archaeal Nop10, Cbf5 and Gar1 (Sequences are shown in Appendix A1 – Genes) were ordered from Eurofins Genomics in pEX-A2 (Nop10 and Gar1) or pEX-K4 (Cbf5) vector backbones. DNA recognition sites for restriction enzymes were: NdeI and XhoI for Nop10, NcoI and XhoI for Gar1, and NcoI and NotI for Cbf5. Vectors were transformed into competent DH5α cells (see Chapter 3.1.2), and colonies were picked for inoculation in 25 mL SB media overnight cultures – containing the respective antibiotics: pEX-A2 (Amp), pEX-K4 (Kan) – and incubated at 37 °C and 150 rpm for DNA amplification. Afterwards, plasmid DNA was extracted using the QIAGEN Plasmid Midi Kit. Enzymatic digestion of the plasmids (5 – 35 µg) was carried out with the two corresponding restriction enzymes (NEB) for each gene (c_{final} = 0.4 units/µL per enzyme) in 1x CutSmart-buffer (NEB) in a 100 µL reaction scale at 37 °C overnight. Cbf5 genes and digested pET-vectors were purified using the QIAquick PCR Purification Kit, and Nop10 as well as Gar1 genes were purified by preparative agarose gel electrophoresis (see Chapter 3.1.3).

Vector ligation and transformation

Nop10 and Cbf5 were inserted into the pET-Duet-1 vector, and Gar1 into the pET-28b(+) vector. Nop10 insertion was carried out prior to Cbf5 insertion. Insertion of codon optimized genes (9 – 30 ng, 5 eq) into pET-vectors (80 ng, 1 eq) was done with NEB T4-DNA-ligase (c_{final} = 20 units/µL) in a 20 µL reaction. The reaction mixture was incubated for 10 min at room temperature. After that, 5 µL of the ligation reaction were used for transfection into competent cells (DH5α 'high efficiency') according to Chapter 3.1.2. For plasmid sequencing, 5 mL (1xLB) overnight cultures containing the respective antibiotics (pET-Duet-1 (Amp), pET-28b (Kan)) were inoculated with colonies and incubated at 37 °C and 150 rpm overnight. Plasmid DNA was extracted using the QIAGEN Plasmid Mini Kit and send to Eurofins Genomics for sequencing. Plasmids containing the correct genes were transformed into BL21(DE3) cells optimized for protein expression (protocol in Chapter 3.1.2). For NCG a cotransformation of pET-Duet-1 (Nop10 and Cbf5) and pET-28b (Gar1) into *E. coli* cells was performed. Glycerol stocks were

prepared from 5 mL (1xLB) overnight cultures. Therefore, 500 µL of 50% glycerol were added to 500 µL cell suspensions, frozen in liquid N_2 and stored at -80 °C.

3.2.2. Protein Expression and Purification

Amino acid sequences of codon optimized archaeal H/ACA-proteins used in this thesis are given in Table 3.4.

Protein	Amino acid sequence (N- to C-terminus)	Size
Cbf5	MAHHHHHHGSARDEVRRILPADIKREVLIKDENAETNPDWGFPPEK RPIEMHIQFGVINLDKPPGPTSHEVVAWIKKILNLEKAGHGGTLDPKV SGVLPVALEKATRVVQALLPAGKEYVALMHLHGDVPEDKIIQVMKEF EGEIIQRPPLRSAVKRRLRTRKVYYIEVLEIEGRDVLFRVGVEAGTYIR SLIHHIGLALGVGAHMSELRRTRSGPFKEDETLITLHDLVDYYYFWKE DGIEEYFRKAIQPMEKAVEHLPKVWIKDSAVAAVTHGADLAVPGIAKL HAGIKRGDLVAIMTLKDELVALGKAMMTSQEMLEKTKGIAVDVEKVF MPRDWYPKLWEKRDRS	39.6 kDa
Gar1	MERLGKVLHYAKQGFLIVRTNWVPSLNDRVVDKRLQFVGIVKDVFG PVKMPYVAIKPKVSNPEIYVGEVLYVDERKRKESPKKNKEKRMKKKK RLNRAAALE	11.9 kDa
Nop10	MADLFRIRKCPKCGRYTLKEVCPVCGEKTKVAHPPRFSPEDPYGEY RRRWKREVLGIGRKEK	7.4 kDa
L7Ae	MHHHHHHGSMAKPSYVKFEVPKELAEKALQAVEIARDTGKIRKGTN ETTKAVERGQAKLVIIAEDVDPEEIVAHLPPLCEEKEIPYIYVPSKKEL GAAAGIEVAAASVAIIEPGKARDLVEEIAMKVKELMK	14.5 kDa

Table 3.4: Amino acid sequences of archaeal H/ACA-proteins.

NCG Expression

In order to express the archaeal NCG proteins 2xLB media was used for cell growth containing the respective antibiotics: Amp (100 µg/mL) for Nop10 and Cbf5, and Kan (30 µg/mL) for Gar1. 25 mL pre-cultures were inoculated from glycerol stocks and incubated at 37 °C and 150 rpm overnight. 2 L main-cultures (in 5 L baffled flasks) were inoculated from the pre-culture to an optical density (600 nm) of 0.1. Cell growth was carried out at 37 °C and 120 rpm to an optical density (600 nm) of 0.5 to 0.9. Protein expression was started by addition of Isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 0.1 mM, and 20 µL of Antifoam Y-30 emulsion (Sigma Aldrich) were added. Cell cultures were incubated at 37 °C

and 120 rpm for 4 h and harvested at 4000 g for 20 min at 4 °C. Cell pellets were divided into four samples, frozen with liquid N_2 and stored at -80 °C.

NCG Purification

For protein purification one cell pellet was resuspended in 30 mL IMAC-A-buffer and one pill of EDTA-free Protease Inhibitor Cocktail (Roche) was added. Resuspended cells were lysed by sonification using a Bandelin Sonopuls under ice water cooling (6x30 sec at 60% Cycles with 50-60% power). Cell fragments were removed by centrifugation for 40 min at 4 °C and 15000 rpm (average 27600 g) at a Beckman Coulter Optima L-90K Ultracentrifuge using a SW 32 Ti rotor. The supernatant was mixed with 500 µL 5% (w/v) polyethyleneimine, incubated at 4 °C for 15 min and centrifuged at 9000 g for 20 min at 4 °C. The remaining supernatant was filtered (0.22 µm) and applied on a 1 mL HisTrap^{TM} (GE). A gradient from 100% IMAC-A to 100% IMAC-B-buffer was applied over 20 mL at a flowrate of 1 mL/min. The eluate was collected in 2 mL fractions. Fractions showing high UV-absorbance were analyzed by SDS-PAGE and fractions containing the desired proteins were combined and concentrated (ca. 1 mL) for further purification. Typically, size exclusion was performed on a Superdex 200 increase 10/300 GL column (GE) in nSEC-buffer at a flow rate of 0.75 mL/min. The eluate was collected in 1 mL fractions. Fractions containing the desired proteins were pooled, rebuffered in sSEC-buffer. Protein concentration was determined by UV/Vis-spectroscopy. For long term storage the samples were aliquoted, frozen in liquid N_2 and stored at -80 °C.

3.2.3. Preparation of Unlabeled Archaeal H/ACA-RNA

DNA template for transcription of the archaeal H/ACA-RNA and primers for DNA amplification were ordered from Eurofins Genomics (see Table 3.5).

DNA	5'-3'-Sequence
pf H/ACA-Template	TTGTGGGCCACGGGAGCGAACGCGGCTCATTGATCAC CGCGCGCGGTTTCCGTGGCCCTATAGTGAGTCGTATTA
Forward primer	TTGTGGGCCACGGGAG
Reverse primer	TAATACGACTCACTATAGGG

Table 3.5: DNA template and primers for archaeal H/ACA-RNA synthesis.

Polymerase Chain Reaction

The DNA template was amplified by polymerase chain reaction (PCR) using the Phusion® HF PCR Kit (NEB) and homemade Phusion polymerase. Reactions were carried out in a 100 µL scale under following conditions: 1xHF-buffer, 3 mM dNTP-Mix, 0.3 mM of each primer, 200 ng DNA template and 0.05 µg/µL homemade Phusion polymerase. The applied temperature program is given in Table 3.6.

Step	Temperature	Time	Repeats
1.	98 °C	120 s	/
2.	54 °C	15 s	31
3.	68 °C	15 s	
4.	98 °C	30 s	
5.	4 °C	Pause	/

Table 3.6: Temperature program for template amplification by PCR.

In vitro Transcription

The PCR reaction mixture was used for transcription without additional purification. RNA transcription was done at 37 °C for 5 h using homemade T7-RNA polymerase (P266L mutant) under following conditions: 10% (v/v) DNA template (PCR), 100 mM Tris (pH 8.1), 30 mM $Mg(OAc)_2$, 5 mM NTPs, 3 mM spermidine, 10 mM DTT and 0.03 µg/µL homemade T7-RNA polymerase. Next, TURBO™ DNase (Invitrogen) was added (c_{final} = 0.03 units/µL) and further incubated for 1 h at 37 °C. The reaction was ended by addition of EDTA to a final concentration of 30 mM. One volume equivalent FA-buffer was added to the sample for purification by denaturing PAGE (see Chapter 3.1.3). The RNA was visualized by UV-shadowing, excised using a surgical scalpel and eluted from the gel in 0.5 M NH_4OAc at 20 °C and 400 rpm overnight. Gel residues were removed with centrifugal filter units (0.45 µm). Finally, the RNA was ethanol precipitated (0.5 M NH_4OAc) and resuspended in H_2O. The concentration of unlabeled H/ACA-RNA was determined by UV/Vis-spectroscopy.

3.2.4. Preparation of Labeled H/ACA-RNAs and Target RNAs

Oligomers for H/ACA-RNA Synthesis

For generation of site specifically labeled H/ACA-RNAs amino modifiers were introduced into RNA oligomers via solid phase synthesis. These oligomers were purchased from Dharmacon

in 2'-bis(2-Acetoxyethoxy)methyl-(ACE)-protected form, received as lyophilized powders and resuspended in H_2O (c_{final} = 1 mM). Sequences of the oligomers are given in Table 3.7.

Oligomer	5'-3'-Sequence
5'-G1	***G***GGCCACGGAAACCGCGCGCGGUGAUCAAU
5'-U26	GGGCCACGGAAACCGCGCGCGGUGA***U***CAAU
bio-3'-C37	pGAGCCG***C***GUUCGCUCCCGUGGCCCACAA-bio
bio-3'-U39	pGAGCCGCG***U***UCGCUCCCGUGGCCCACAA-bio
3'-U39	pGAGCCGCG***U***UCGCUCCCGUGGCCCACAA

Table 3.7: Purchased oligomers for synthesis of labeled archaeal H/ACA-RNAs. Amino modified nucleotides are highlighted as italic and bold letters. 5'-phosphates 'p' were inserted for ligation. 3'-Biotin '-bio' modifications were inserted for immobilization.

To enable site specific labeling, a 6-aminohexyl phosphate was introduced at G1 (5'-oligo), a 2'-amino group was introduced at C37 and aminoallyl groups were introduced at U26 and U39 (3'-oligos). Chemical structures of the functional groups and reactive cyanine dyes are shown in Figure 4.10 on page 83.

Labeling and Deprotection

For fluorophore labeling of individual oligos one tube of the Amersham CyDye Mono-Reactive Dye Pack' (GE) containing the respective dye (Cy3/Cy5) was used. The dye was diluted in 20 µL DMSO. 30 nmol of the oligomer were ethanol precipitated (0.3 M NaOAc) and diluted in freshly prepared 0.1 M $NaHCO_3$. Dilutions of RNA and dye were mixed and incubated for 90 min at room temperature in the dark. Afterwards, the RNA was ethanol precipitated (0.3 M NaOAc), resuspended in 300 µL deprotection-buffer (supplied by Dharmacon: 100 mM AcOH, pH 3.4-3.8 adjusted with TEMED) and heated to 60 °C for 30 min (90 min for oligos containing a biotin modification). The RNA was ethanol precipitated (0.5 M NH_4OAc) and resuspended in 300 µL TEAA-buffer for HPLC purification. Unmodified oligomers were directly deprotected in deprotection-buffer (10 nmol in 100 µL), ethanol precipitated (0.5 M NH_4OAc) and resuspended in H_2O. The concentration of unlabeled oligomers was determined by UV/Vis-spectroscopy.

HPLC Purification

Chromatographic purifications of labeling reactions were done using a Kromasil C8-HPLC-column (250 x 4.6 mm) on an Äkta Basic^{TM} (GE). The column was equilibrated with

TEAA-buffer and the sample was applied to the column via a 500 µL loop. A gradient was set over 80 mL from 100% TEAA-buffer to 100% MeCN at a flow rate of 1 mL/min. Absorbance was measured at 260 nm (RNA) and 550/650 nm (Cy3/Cy5). Fractions exhibiting high dye absorption were collected, concentrated for 30 min in a vacuum centrifuge and ethanol precipitated (0.5 M NH_4OAc). Pellets were resuspended in H_2O and the concentration of labeled RNAs was determined by UV/Vis-spectroscopy.

Enzymatic Ligation

To obtain the full-length H/ACA-RNA, splinted ligation was performed. Equimolar amounts (500 or 1000 pmol for preparative scale and 40 pmol for analytical scale) of 5'- and 3'- RNA oligos and the DNA-splint (H/ACA-template) in 0.5xT4-DNA-ligase-buffer (NEB) were heated to 85 °C for 5 min and allowed to cool down to room temperature for at least 15 min. Then, buffer was adjusted to 1x-concentration and either a mixture of T4-DNA-Ligases from ThermoFisher (c_{final} = 0.125 Weiss units/µL) and NEB (c_{final} = 50 units/µL) or only the homemade T4-RNA-ligase 2 (c_{final} = 0.04 µg/µL) was added to a final oligomer concentration of 2.5 µM. The reaction mixture was incubated at 37 °C for 2.5 h. Afterwards, TURBO^{TM} DNase (Invitrogen) was added (c_{final} = 0.05 units/µL) and further incubated for 1 h at 37 °C. Proteins were phenol extracted and RNAs were ethanol precipitated (0.5 M NH_4OAc). Pellets was resuspended in 45 µL H_2O for preparative denaturing PAGE. Fluorophore labeled RNAs were visualized by eye, excised from the gel and eluted from the gel in 0.5 M NH_4OAc at 20 °C and 400 rpm overnight. Gel residues were removed with centrifugal filter units (0.45 µm). Finally, the purified RNA was ethanol precipitated (0.5 M NH_4OAc) and resuspended in H_2O. The concentration of labeled H/ACA-RNAs was determined by UV/Vis-spectroscopy.

Target RNAs

Amino modifiers for site specific labeling were introduced into target RNAs via solid phase synthesis. These RNAs were purchased from Dharmacon in the ACE-protected form, received as lyophilized powders and resuspended in H_2O (c_{final} = 1 mM). Sequences of the oligomers are given in Table 3.8.

Target RNA	5'-3'-Sequence
U-Sub	GAUGGAGCG***U***GCGGUUUAAU-NH_2
Ψ-Sub	GAUGGAGCG***Ψ***GCGGUUUAAU-NH_2
5fU-Sub	GAUGGAGCG***5fU***GCGGUUUAAU-NH_2
C-Sub	GAUGGAGCG***C***GCGGUUUAA-NH_2
rab-Sub	GAUGGAGCG***(rab)***GCGGUUUAAU-NH_2

Table 3.8: Various archaeal target RNAs with different nucleotides at the target site. 3'-ends are functionalized with an aminohexyl phosphate group (-NH_2). The target bases are highlighted as bold and italic letters.

Fluorophore labeling and subsequent HPLC purification of the target RNAs were performed as described above. For insertion of a biotin modification at the 3'-end one tube of EZ-Link™ NHS-PEG4-Biotin (ThermoFisher) was used instead of a reactive cyanine dye, without changing the further procedure. Ligation was not necessary. HPLC purified and labeled target RNAs were ethanol precipitated (0.5 M NH_4OAc), resuspended in H_2O and the concentration was determined by UV/Vis-spectroscopy.

3.2.5. RNP Reconstitution and Fluorescence Anisotropy

RNA-Protein Binding

RNP formation was tested with Electrophoretic Mobility Shift Assays (EMSAs). Samples were prepared with constant RNA and increasing protein concentration in a 10 µL scale in PCR-tubes. Different measurement rows for L7Ae and NCG were prepared. For unlabeled H/ACA-RNAs a 5 µM sRNA stock solution was heated to 98 °C for 2.5 min and immediately put on ice afterwards. Unlabeled sRNA concentration was adjusted to 1 µM in nSEC-buffer for each sample. NCG concentrations were adjusted to 0.0, 1.0, 2.0, 5.0, or 10 µM. L7Ae concentrations were adjusted to 0.0, 2.0, 5.0, 10, or 20 µM. All samples were annealed at 70 °C for 15 min and then native polyacrylamide PAGE was performed (see Chapter 3.1.3).

For labeled H/ACA-RNAs, a 25 nM stock solution was heated to 98 °C for 2.5 min and immediately put on ice. Labeled sRNA concentration was adjusted to 5 nM in all samples. Individual measurement rows for L7Ae, NCG, CG and Cbf5 were performed; each containing of five samples with increasing protein concentrations. The L7Ae measurement row was performed in 0.5x oSEC-buffer, the NCG measurement row in 0.8x nSEC-buffer, and the CG and Cbf5 measurement rows in 1x Ψ-buffer. L7Ae concentrations were adjusted to 0.0, 2.0, 5.0, 7.0, or

10 µM, NCG and CG concentrations to 0.0, 0.1, 0.2, 0.5, or 1.0 µM, and Cbf5 concentrations to 0.0, 0.2, 0.4, 0.8, or 1.4 µM. All samples were annealed at 70 °C for 60 min (NCG and L7Ae) or 15 min (CG and Cbf5). Finally, native polyacrylamide PAGE was performed (see Chapter 3.1.3).

Fluorescence Anisotropy

Bulk fluorescence anisotropy measurements of archaeal H/ACA-RNPs were performed on FluoroMax-4 spectrophotometer (Horiba Scientific) in L-format geometry with a clear 1 mL (10x10 mm) cuvette at 21 °C. Stock solutions of the labeled guide RNAs in water (1.1 to 6.0 µM) were heated to 98 °C for 2.5 min and immediately put on ice. Samples were prepared in an 800 µL scale containing 5 nM of the labeled H/ACA-RNA or free dye in 1x Ψ-buffer. For RNP samples L7Ae and NCG concentration were adjusted to 5 µM and 0.5 µM, respectively. All samples (free dyes, H/ACA-RNA and H/ACA-RNP) were heated to 70 °C for 5 min and stored on ice until the actual measurement was carried out. Excitation was set to 515 nm for Cy3 and 615 nm for Cy5. Emission was detected at 561 nm for Cy3 and 661 nm for Cy5. The bandwidth was set to 8 nm for Cy3 and 10 nm for Cy5. For each sample 20 scans with an integration time of 100 ms were performed. The fluorescence anisotropy (*A*) was calculated from the intensities detected when the polarizer was in parallel direction to the polarized emission (I_p), and when the polarizer was in vertical direction to the polarized emission (I_v), according to this equation: $A = \frac{I_p - I_v}{I_p + 2I_v}$.

3.2.6. Pseudouridylation Activity Assays

Enzymatic activity of labeled and unlabeled archaeal H/ACA-RNPs was tested with radioactive ^{32}P-labeled substrates. Single turnover assays were performed with transcribed target RNA, containing stochastically incorporated α-^{32}P-U. Multiple turnover assays were performed with a mixture of ligated target RNA '*pf* Sub' (see Table 3.3, page 46), in which the target uridine was labeled with γ-^{32}P-ATP, and of deprotected U-Sub (see Table 3.8, page 54). The preparation of radioactively labeled target RNAs is described in Chapter 3.1.5. All activity assays were performed in Ψ-buffer.

Single Turnover Assays

Full and partial sRNPs were tested for single turnover activity. Therefore, the full H/ACA-RNP was compared to incomplete sRNPs; lacking the guide RNA 'NCG-only' (negative control), or individual proteins: ΔGar1, ΔL7Ae and ΔNop10. All reactions were performed in a 40 µL scale. For sample preparation, a 40 µM H/ACA-RNA stock solution was heated to 95 °C for 5 min and immediately put on ice. If present, final concentrations were: 1 µM guide RNA, 5 µM L7Ae, and 2 µM NCG/CG/NC. The samples were annealed for 15 min at 70 °C in absence of the target RNA. Then the transcribed ^{32}P-substrate RNA was added (c_{final} = 0.8 µM). Samples were incubated at 70 °C for 1 h. One additional full H/ACA-RNP sample was incubated at 37 °C for 14 h. Reactions were stopped by addition of 160 µL H_2O and storage on ice. Samples were phenol extracted with 1 vol. eq. water saturated aqua phenol, ethanol precipitated (0.5 M NH_4OAc) and resuspended in 57 µL P1-buffer. 3.5 µL P1-endonuclease (0.25 units/µL) were added and the solution was incubated at 55 °C for 3 h. Next, ^{32}P-isotope counts per minute of each sample were determined with a TRI-CARB 2100 TR Liquid Scintillation Analyzer (Packard). 400 – 1000 counts per minute for each sample were spotted on a 20 cm long thin layer chromatography cellulose plate (Merck) at a distance of 15 mm to the neighboring spots. Digested samples were separated by TLC using a mixture of isopropanol/H_2O/conc. HCl (70:15:15) as eluent at room temperature overnight (adapted from Keith[32]). Radioactive nucleotides were visualized by autoradiography with a storage phosphor screen (GE) by exposing the TLC plate for 72 h to the screen. The screen was scanned with a Typhoon 9400 (GE) imager. Evaluation of the separated spots was done by densitometric analysis using the ImageJ software[122]. Pseudouridylation yields were calculated by comparison of integrated uridine (I_U) and pseudouridine (I_Ψ) spots, using the equation: $Yield_{uniformly\ U\ labeled} = I_\Psi \cdot \left(\frac{I_\Psi + I_U}{6}\right)^{-1}$.

Multiple Turnover Assays

Labeled FRET-RNAs were compared to the wt-H/ACA-RNA for catalytic activities in multiple turnover assays with a 20-fold excess of substrate RNA. Before sample preparation guide RNAs were diluted (3.5 – 5.0 µM) in H_2O, heated to 98 °C for 2.5 min and immediately put on ice. For each measurement row a stock solution in Ψ-buffer containing 0.05 µM guide RNA (labeled or unlabeled), 0.4 µM NCG, 5 µM L7Ae, 1.9 µM U-Sub and 0.1 µM ^{32}P-substrate was prepared. The radioactive target RNA was added last. The stock solutions were heated to

70 °C. For each construct a 20 µL sample was taken after 1, 5, 15, 30, 60, and 180 min, added to a suspension of 80 µL H_2O and 100 µL PCI and phenol extracted to stop the reaction. RNAs were ethanol precipitated (0.3 M NaOAc), the pellets were resuspended in a mixture of 39 µL P1-buffer and 1 µL P1-endonuclease (0.25 units/µL). The reaction mixture was incubated at 55 °C for 4 h and 3 µL (2 x 1.5 µL) of each sample were spotted on a 20 cm long thin layer chromatography cellulose plate (Merck) at a distance of 15 mm to the neighboring spots. Digested samples were separated by TLC using a mixture of isopropanol/H_2O/conc. HCl (70:15:15) as eluent at room temperature overnight. Radioactive nucleotides were visualized by autoradiography with a storage phosphor screen (GE) by exposing the TLC plate for 48 – 72 h to the screen. The screen was scanned with a Typhoon 9400 (GE) imager. Evaluation of the separated spots was done by densitometric analysis using the ImageJ software[122]. Pseudouridylation yields were calculated by comparison of integrated uridine (I_U) and pseudouridine (I_Ψ) spots, using the equation: $Yield_{target\ U\ labeled} = \frac{I_\Psi}{I_\Psi + I_U}$.

3.3. The Eukaryotic H/ACA-Ribonucleoprotein Complex

3.3.1. Protein Expression and Purification

BL21(DE3) cells optimized for protein expression containing the plasmids for protein biosynthesis (scNop10 and scCbf5 in pET-Duet-1, Gar1 in pET-28b and Nhp2 in pET-24b) were prepared and provided by Sven Trucks (Goethe University). scNop10, scCbf5 and scGar1 were coexpressed in *E. coli* and copurified as ternary protein complex 'NCG'. Amino acid sequences of eukaryotic H/ACA-proteins used in this thesis are given in Table 3.9.

Protein	**Amino acid sequence (N- to C-terminus)**	**Size**
Nhp2	MGKDNKEHKESKESKTVDNYEARMPAVLPFAKPLASKKLNKKVLKTV KKASKAKNVKRGVKEVVKALRKGEKGLVVIAGDIWPADVISHIPVLCE DHSVPYIFIPSKQDLGAAGATKRPTSVVFIVPGSNKKKDGKNKEEEYK ESFNEVVKEVQALGLEHHHHHH	18.3 kDa
scCbf5	MAKEDFVIKPEAAGASTDTSEWPLLLKNFDKLLVRSGHYTPIPAGSSP LKRDLKSYISSGVINLDKPSNPSSHEVVAWIKRILRCEKTGHSGTLDPK VTGCLIVCIDRATRLVKSQQGAGKEYVCIVRLHDALKDEKDLGRSLEN LTGALFQRPPLISAVKRQLRVRTIYESNLIEFDNKRNLGVFWASCEAG TYMRTLCVHLGMLLGVGGHMQELRRVRSGALSENDNMVTLHDVMD AQWVYDNTRDESYLRSIIQPLETLLVGYKRIVVKDSAVNAVCYGAKLMI PGLLRYEEGIELYDEIVLITTKGEAIAVAIAQMSTVDLASCDHGVVASVK RCIMERDLYPRRWGLGPVAQKKKQMKADGKLDKYGRVNENTPEQW KKEYVPLDNAEQSTSSSQETKETEEEPKKAKEDSLIKEVETEKEEVKE DDSKKEKKEKKDKKEKKEKKEKKDKKEKKEKKEKKRKSEDGDSEEKK SKKSKKHHHHHH	55.5 kDa
scGar1	MGSSHHHHHHSSGLVPRGSHMSFRGGNRGGRGGFRGGFRGGRTG SARSFQQGPPDTVLEMGAFLHPCEGDIVCRSINTKIPYFNAPIYLENKT QVGKVDEILGPLNEVFFTIKCGDGVQATSFKEGDKFYIAADKLLPIERF LPKPKVVGPPKPKNKKKRSGAPGGRGGASMGRGGSRGGFRGGRG GSSFRGGRGGSSFRGGSRGGSFRGGSRGGSRGGFRGGRR	23.6 kDa
scNop10	MHLMYTLGPDGKRIYTLKKVTESGEITKSAHPARFSPDDKYSRQRVTL KKRFGLVPGQ	6.6 kDa
scGar1-Ctag	MSFRGGNRGGRGGFRGGFRGGRTGSARSFQQGPPDTVLEMGAFL HPCEGDIVCRSINTKIPYFNAPIYLENKTQVGKVDEILGPLNEVFFTIKC GDGVQATSFKEGDKFYIAADKLLPIERFLPKPKVVGPPKPKNKKKRSG APGGRGGASMGRGGSRGGFRGGRGGSSFRGGRGGSSFRGGSRG GSFRGGSRGGSRGGFRGGRRHHHHHH	22.3 kDa

scGar1-notag	MSFRGGNRGGRGGFRGGFRGGRTGSARSFQQGPPDTVLEMGAFL HPCEGDIVCRSINTKIPYFNAPIYLENKTQVGKVDEILGPLNEVFFTIKC GDGVQATSFKEGDKFYIAADKLLPIERFLPKPKVVGPPKPKNKKKRSG APGGRGGASMGRGGSRGGFRGGRGGSSFRGGRGGSSFRGGSRG GSFRGGSRGGSRGGFRGGRR	21.5 kDa

Table 3.9: Amino acid sequences of eukaryotic H/ACA-proteins.

scNCG and Nhp2 Expression

Eukaryotic scNCG and Nhp2 were expressed in TB media containing the respective antibiotics (Amp and Kan for scNCG, and Amp for Nhp2). 25 mL pre-cultures were inoculated from glycerol stocks and incubated at 37 °C and 150 rpm overnight. 2 L main-cultures (in 5 L baffled flasks) were inoculated from the pre-culture to an optical density (600 nm) of 0.1. Cell growth was carried out at 37 °C and 120 rpm to an optical density (600 nm) of 0.5 to 0.9. Then the main-cultures were stored on ice to cool down the cell suspension for ca. 15 min. Protein expression was initiated by adding IPTG (c_{final} = 0.1 mM) to the cooled cell culture. 20 µL of Antifoam Y-30 emulsion (Sigma Aldrich) were added. Cell cultures were then incubated at 20 °C and 120 rpm for 12-18 h and harvested at 4000 g for 20 min at 4 °C. Cell pellets were divided into four samples, frozen with liquid N_2 and stored at -80 °C.

scNCG and Nhp2 Purification

For protein purification one cell pellet was resuspended in 30 mL scA-buffer. 10 µL TURBO™ DNase (Invitrogen, 2 units/µL), 10 µL PureLink™ RNase A (Invitrogen, 20 mg/mL) and one pill of EDTA-free Protease Inhibitor Cocktail (Roche) were added. Resuspended cells were lysed by sonification using a Bandelin Sonopuls under ice water cooling (6x45 sec at 60% Cycles with 50-60% power). Cell fragments were removed by centrifugation for 40 min at 4 °C and 15000 rpm (27600 g) at a Beckman Coulter Optima L-90K Ultracentrifuge using a SW 32 Ti rotor. The supernatant was mixed with 500 µL 5% (w/v) polyethyleneimine, incubated for 15 min at 4 °C min and centrifuged at 9000 g for 20 min at 4 °C. The remaining supernatant was filtered (0.22 µm) and applied on a 5 mL HisTrap™ (GE). *Note*: In case of scNCG-Ctag and scNCG-notag the column bound proteins were additionally washed with 30 mL high-salt buffer (2 M LiCl, 50 mM Tris, 20 mM imidazole, 10% (v/v) glycerol, pH 8.0 adjusted with HCl) and rebuffered to scA. Then, a gradient from 100% scA- to 100% scB-buffer was applied over 100 mL at a flow rate of 3 mL/min. The eluate was collected in 5 mL fractions. Fractions showing high UV-absorbance were analyzed by SDS-PAGE and checked for RNA

contamination. Fractions containing the desired proteins were combined and concentrated to a final volume of 10 mL. 20 µL PureLink™ RNase A (Invitrogen, 20 mg/mL) were added to the concentrated samples and incubated at 4 - 15 °C (scNCG) or 20 °C (Nhp2) overnight. Samples were diluted with scSEC-buffer (final imidazole concentration < 20 mM) and applied on a 1 mL HisTrap™ (GE). A gradient from 100% scA- to 100% scB-buffer was applied over 20 mL at a flowrate of 1 mL/min. Fractions showing high UV-absorbance were combined and concentrated (ca. 1 mL) for further purification. Size exclusion was performed on a Superdex 200 increase 10/300 GL column (GE) in scSEC-buffer at a flowrate of 0.75 mL/min. The eluate was collected in 1 mL fractions. Fractions containing the desired proteins were pooled and concentrated. Protein concentration was determined by UV/Vis-spectroscopy. For long term storage the samples were aliquoted, frozen in liquid N_2 and stored at -80 °C.

3.3.2. Preparation of Unlabeled Eukaryotic H/ACA-RNA Constructs

DNA templates for transcription of eukaryotic H/ACA-RNA constructs and respective primers for DNA amplification were ordered from Eurofins Genomics (see Table 3.10).

Construct	DNA	5'-3'-Sequence
H5	H5-template	AGGATTGCTCTTGGGACCGTAGTATCGCGACAAGGAAGTCTGTGTAGTTCCACTTGATGTGGGCTGCCTCGCCGCTTCTTTTGCAGTCCCTATAGTGAGTCGTATTA
	Forward primer	AGGATTGCTCTTGGGACC
	Reverse primer	TAATACGACTCACTATAGGG
H5d	H5d-template	AGGATTGCTCTTGGGACCGTAGTATCGCGACCGAAGCCGCTTCTTTTGCAGTCCCTATAGTGAGTCGTATTA
	Forward primer	AGGATTGCTCTTGGGACC
	Reverse primer	TAATACGACTCACTATAGGG

Table 3.10: DNA templates and primers for eukaryotic H/ACA-RNA construct synthesis.

Polymerase Chain Reaction

The DNA template was amplified PCR using the Phusion® HF PCR Kit (NEB) and homemade phusion polymerase. Typically, reactions were carried out in a 100 µL scale under following conditions: 1xHF-buffer, 200 µM dNTP-Mix, 3 µM of each primer, 0.1 µM DNA

template and 0.05 µg/µL homemade Phusion polymerase. The applied temperature program is shown in Table 3.11.

Step	Temperature	Time	Repeats
1.	98 °C	120 sec	/
2.	54 °C	15 sec	31
3.	68 °C	15 sec	
4.	98 °C	30 sec	
5.	4 °C	Pause	/

Table 3.11: Temperature program for template amplification by PCR.

In vitro Transcription

The PCR reaction mixture was used for transcription without additional purification. RNA transcriptions were done at 37 °C for 4 - 5 h using homemade T7-RNA polymerase (P266L mutant) under following conditions: 10% (v/v) DNA template (PCR), 100 mM Tris (pH 8.1), 30 mM $Mg(OAc)_2$, 3.75 mM NTPs, 15 mM guanosine monophosphate, 3 mM spermidine, 10 mM DTT and 0.03 µg/µL homemade T7-RNA polymerase. Next, TURBO™ DNase (Invitrogen) was added (c_{final} = 0.03 units/µL) and further incubated for 1 h at 37 °C. Reactions were ended by addition of EDTA to a final concentration of 30 mM. One volume equivalent FA-buffer was added to one sample for purification by denaturing PAGE (see Chapter 3.1.3). The RNA was visualized by UV-shadowing, excised using a surgical scalpel and eluted from the gel in 0.3 M NaOAc at 20 °C and 400 rpm overnight. Gel residues were removed with centrifugal filter units (0.45 µm). Finally, the RNA was ethanol precipitated (0.3 M NaOAc) and resuspended in H_2O. The concentration of unlabeled H/ACA-RNA constructs was determined by UV/Vis-spectroscopy.

3.3.3. Preparation of Labeled H/ACA-RNAs and Target RNAs

Oligomers for H/ACA-RNA Synthesis

For generation of site specifically labeled H/ACA-RNAs amino modifiers were introduced into RNA oligomers via solid phase synthesis. These oligomers were purchased from Dharmacon in the ACE-protected form, received as lyophilized powders and resuspended in H_2O (c_{final} = 1 mM). The 5'-hairpin constructs *H5*- and *H5d-FRET* were divided into three oligomers (5`-, middle- and 3'-oligo). The 3'-hairpin constructs were divided into three oligomers for

H3-FRET (5`-, middle- and 3'-oligo) and into two oligomers for *H3d-FRET* (5`-, and 3'-oligo). DNA ligation splints were ordered from Eurofins, received as lyophilized powders and resuspended in H_2O (c_{final} = 100 µM). For *H5*- and *H5d-FRET* the respective transcription templates were used as DNA-splints (see Table 3.10, page 60). The other nucleic acid strands are given in Table 3.12.

Constructs	Oligomer	5'-3'-Sequence
5'-hairpin H5/H5d	5'-oligo	***G***GGACUGCAAAAGAAGCGG
	3'-oligo	pGAUACUACGGUCCCAAGAGCAAUCCU-bio
	H5-middle-oligo	pCGAGGCAGCCCACAUCAAGUGG AACUACACAGACUUCCUUG***U***CGC
	H5d-middle-oligo	pCUUCGG***U***CGC
3'-hairpin H3/H3d	5'-oligo	pAACAAGCAAUUACAUAUUCC***U***CCGCU
	3'-oligo	pUGGGGGAAGACGCUU***U***UUC ACAUCUUCUUGCAUGAUAA-bio
	H3-middle-oligo	pGAACCUGUACAGUCCACGGAU GGUGCAGAAGUUAUAUGAUU
DNA-splints	H3-splint	GTGAAAAAGCGTCTTCCCCCAAATCATATAACTTCTG CACCATCCGTGGACTGTACAGGTTCAGCGGAGGAAT ATGTAATTGC
	H3d- splint	GTGAAAAAGCGTCTTCCCCCAAGCGGAGGAATATGT AATTGC

Table 3.12: Purchased oligomers for synthesis of labeled eukaryotic H/ACA-RNA constructs. Amino modified nucleotides are highlighted as italic and bold letters. 5'-phosphates 'p' were inserted for ligation. 3'-Biotin '-bio' modifications were inserted for immobilization.

For the 5'-hairpin constructs a 6-aminohexyl phosphate was introduced at the 5'-oligo (G1) and an aminoallyl uridine was introduced at the middle-oligo (U61 for H5 and U26 for H5d) for site specific fluorophore coupling. For the 3'-hairpin constructs aminohexylacrylamino groups were introduced at the 5'-oligo (U21) and at the 3'-oligo (U83 for H3 and U42 for H3d). Chemical structures of the functional groups and reactive cyanine dyes are shown in Figure 4.10 on page 83.

Labeling and Deprotection

For fluorophore labeling of individual oligos, one tube of the Amersham CyDye Mono-Reactive Dye Pack (GE) containing the respective dye (Cy3/Cy5) was used. The dye was diluted

in 20 µL DMSO. 30 nmol of the oligomer were ethanol precipitated (0.3 M NaOAc) and diluted in freshly prepared 0.1 M $NaHCO_3$ (*Note*: The H5d-middle-oligo was deprotected prior to ethanol precipitation). Dilutions of RNA and dye were mixed and incubated for 90 min at room temperature in the dark. Afterwards, the RNA was ethanol precipitated (0.3 M NaOAc), resuspended in 300 µL deprotection-buffer (supplied by Dharmacon: 100 mM AcOH, pH 3.4-3.8 adjusted with TEMED) and heated to 60 °C for 30 min (90 min for oligos containing a biotin modification). The RNA was ethanol precipitated (0.3 M NaOAc) and resuspended in 300 µL TEAA-buffer for HPLC purification. Unmodified oligomers were directly deprotected in deprotection-buffer (10 nmol in 100 µL), ethanol precipitated (0.3 M NaOAc) and resuspended in H_2O. The concentration of unlabeled oligomers was determined by UV/Vis-spectroscopy.

HPLC Purification

Chromatographic purifications of labeling reactions were done using a Kromasil C8-HPLC-column (250 x 4.6 mm) on an Äkta Basic™ (GE). The column was equilibrated with TEAA-buffer and the sample was applied to the column via a 500 µL loop. A gradient was set over 80 mL from 100% TEAA-buffer to a mixture of MeCN/TEAA-buffer (1:1) at a flow rate of 1 mL/min. Absorbance was measured at 260 nm (RNA) and 550/650 nm (Cy3/Cy5). Fractions exhibiting high dye absorption were collected, concentrated for 30 min in a vacuum centrifuge and ethanol precipitated (0.3 M NaOAc). Pellets were resuspended in H_2O and the concentration of labeled RNAs was determined by UV/Vis-spectroscopy.

Enzymatic Ligation

To obtain H/ACA-RNA constructs of desired lengths, splinted ligation was performed. Equimolar amounts (1000 pmol for preparative scale and 40 pmol for analytical scale) of the required oligos and the respective DNA-splint in 0.5xT4-DNA-ligase-buffer (NEB) were heated to 85 °C for 5 min and allowed to cool down to room temperature for at least 15 min. Then, buffer was adjusted to 1x-concentration and the homemade T4-RNA-ligase 2 (c_{final} = 0.04 µg/µL) was added with a final oligomer concentration of 2.5 µM. The reaction mixture was incubated at 37 °C for 2.5 h. Then, TURBO™ DNase (Invitrogen) was added (c_{final} = 0.05 units/µL) and further incubated for 1 h at 37 °C. Proteins were phenol extracted and RNAs were ethanol precipitated (0.3 M NaOAc). RNA pellets were resuspended in 45 µL H_2O for preparative denaturing PAGE. Fluorophore labeled RNAs were visualized by eye, excised from the gel and

eluted from the gel in 0.3 M NaOAc at 20 °C and 400 rpm overnight. Gel residues were removed with centrifugal filter units (0.45 µm). Finally, the purified RNA was ethanol precipitated (0.3 M NaOAc) and resuspended in H_2O. The concentration of labeled H/ACA-RNAs was determined by UV/Vis-spectroscopy.

Target RNAs

Amino modifiers for site specific labeling were introduced into target RNAs via solid phase synthesis. These RNAs (U-, Ψ- and 5fU-scSub) were purchased from Dharmacon in the ACE-protected form, received as lyophilized powders and resuspended in H_2O (c_{final} = 1 mM). Sequences of these RNAs are given in Table 3.13.

Target RNA	5'-3'-Sequence
U-scSub	GUGUAGUAUC***U***GUUCUUUUCAG-NH_2
Ψ-scSub	GUGUAGUAUC***Ψ***GUUCUUUUCAG-NH_2
5fU-scSub	GUGUAGUAUC***5fU***GUUCUUUUCAG-NH_2

Table 3.13: Various eukaryotic target RNAs with different nucleotides at the target site. 3'-ends are functionalized with an aminohexyl phosphate group (-NH_2). The target bases are highlighted as bold and italic letters.

Fluorophore labeling and subsequent HPLC purification of the target RNAs were performed as described above. Purified and labeled target RNAs were ethanol precipitated (0.3 M NaOAc), resuspended in H_2O and the concentration was determined by UV/Vis-spectroscopy.

3.3.4. RNP Reconstitution

RNP formation was tested with EMSAs. Samples were prepared in PCR-Tubes with constant RNA and increasing protein concentration in a 10 µL scale for unlabeled guide RNAs and in 4 µL scale for labeled guide RNAs. Different measurement rows for Nhp2 and scNCG were prepared. A 10 µM stock solution of H5 and H5d in water was heated to 98 °C for 2.5 min and immediately put on ice. Unlabeled guide RNA concentration was adjusted to 2 µM in each sample. Nhp2 concentrations were adjusted to 0.0, 1.0, 2.0, 4.0, or 8.0 µM. scNCG concentrations were adjusted to 0.0, 4.0, 8.0, 16, or 32 µM. All samples were annealed at 30 °C for 15 min in 0.8x scSEC-buffer. Finally, native polyacrylamide PAGE was performed (see Chapter 3.1.3).

For labeled guide RNAs, a 0.5 µM stock solution of *H5*- and *H5d-FRET* was heated to 98 °C for 2.5 min and immediately put on ice. Labeled RNA concentration was adjusted to 0.1 µM in each sample. Nhp2 concentrations were adjusted to 0.0, 0.5, 1.0, 2.0, 4.0, or 8.0 µM. scNCG concentrations were adjusted to 0.0, 0.25, 0.5, 1.0, 2.0, or 4.0 µM. All samples were annealed at 30 °C for 15 min in scΨ-buffer. Finally, native polyacrylamide PAGE was performed (see Chapter 3.1.3).

3.3.5. Pseudouridylation Activity Assays

Enzymatic activity of fluorophore labeled and unlabeled eukaryotic H/ACA-RNPs was tested with radioactive ^{32}P-labeled substrates. All activity assays were performed with radioactive target RNAs – 5'-Sub and 3'-Sub (see Table 3.3, page 46) –, which where selectively labeled at the target uridine with γ-^{32}P-ATP. The preparation of these radioactively labeled target RNAs is described in Chapter 3.1.5. In multiple turnover assays non-radioactive target RNAs (deprotected U-scSub/5fU-scSub for 5'-hairpin constructs; see Table 3.13, page 62) were added to the respective experiments. All activity assays were performed in scΨ-buffer.

Activity assays of unlabeled and labeled eukaryotic snoRNPs were performed in the same manner. In multiple turnover experiments a 40-fold excess of substrate RNA over the H/ACA-RNP was used. In single turnover experiments equimolar amounts of target RNA and H/ACA-RNPs were used. Before sample preparation, guide RNA stock solutions were diluted (1.3 – 5.0 µM) in H_2O, heated to 98 °C for 2.5 min and immediately put on ice. Concentrations were adjusted to 0.1 µM guide RNA (labeled or unlabeled), 1.0 µM of the used proteins (Nhp2, scNop10, scCbf5 and scGar1), 3.9 µM nonradioactive target RNAs (deprotected U-scSub) in case of multiple turnover assays, and 0.1 µM of the ^{32}P-substrate (5'-Sub or 3'-Sub). The radioactive target RNA was added last. Single turnover experiments were incubated at 30 °C for 1 h. In multiple turnover assays 20 µL samples were taken after 5, 20, 60, and 240 min (in case of H3 and *H3-FRET* an additional sample after 22 h was taken). Samples of the single and multiple turnover reactions were stopped by adding a mixture of 80 µL H_2O and 100 µL PCI and subsequent phenol extraction. RNAs were ethanol precipitated (0.3 M NaOAc), the pellets were resuspended in a mixture of 39 µL P1-buffer and 1 µL P1-endonuclease (0.25 units/µL). The reaction mixture was incubated for 2 h at 55 °C and for 30 min at 70 °C. 3 µL (2 x 1.5 µL) of each sample were spotted on a 20 cm long thin layer chromatography cellulose plate (Merck) at a

distance of 15 mm to the neighboring spots. Digested samples were separated by TLC using a mixture of isopropanol/H_2O/conc. HCl (70:15:15) as eluent at room temperature overnight (adapted from Keith[32]). Radioactive nucleotides were visualized by autoradiography with a storage phosphor screen (GE) by exposing the TLC plate for 24 – 72 h to the screen. The screen was scanned with a Typhoon 9400 (GE) imager. Evaluation of the separated spots was done by densitometric analysis using the ImageJ software[122]. Pseudouridylation yields were calculated by comparison of integrated uridine (I_U) and pseudouridine (I_Ψ) spots, using the equation: $Yield_{target\ U\ labeled} = \frac{I_\Psi}{I_\Psi + I_U}$.

3.4. Single molecule FRET Spectroscopy

3.4.1. Slide Preparation

In order to immobilize samples for smFRET experiments PEG-passivated glass sides were prepared. Therefore, 24x60 mm glass coverslips (Carl Roth) were plasma cleaned (Diener Electronic) in O_2-flow (~ 0.2 mbar) for 10 min. Glass slide holders filled with MeOH were sonicated for 5 min. The methanol in the glass slide holders was replaced with a 100:5:1-mixture (v/v) of MeOH, AcOH and (3-aminopropyl)ethoxysilane (Sigma-Aldrich). The plasma cleaned coverslips were put into the glass slide holder containing the aminosilane mixture, sonicated for 1 min and let sit for 15 – 20 min. Glass slides were rinsed with H_2O thoroughly, and dried under N_2-flow. For 14 slides, 200 mg PEG-NHS (MW 5000, Rapp Polymere) and 2 mg Biotin-PEG-NHS (MW 5000, Nanocs) were dissolved in 1 mL of a freshly prepared 0.1 M $NaHCO_3$ solution. 150 µL of the PEG-solution were put onto the surface of one coverslip. A second coverslip was placed on top, such that the PEG-solution was sandwiched in between two coverslips (*Note*: The surfaces not in contact with PEG-solution were marked for identification). The PEGylation reaction was allowed to incubate at room temperature overnight in a self-prepared moisture chamber in the dark. Next, the coverslips were rinsed with H_2O thoroughly, dried under N_2-flow and stored for further usage under inert gas (N_2 or argon) at -80 °C.

The final glass slides for smFRET experiments were prepared on the same day of the actual measurement. Therefore, the required number of PEGylated coverslips was thawed, and channels (ca. 3 mm wide) were prepared with double-sided sticky tape (Scotch) on plasma cleaned (N_2-flow) 26x76 mm microscope slides (Carl Roth). Finally, the passivated coverslip was put on the sticky tape, in a way that the PEGylated side faced the microscope slide. Typically, a prepared glass slide for smFRET microscopy contained seven channels for sample immobilization.

3.4.2. Sample Preparation

Archaeal smFRET Constructs

Annealing of archaeal samples for smFRET experiments was performed in a 4 µL scale in Ψ-buffer. If present, concentrations were adjusted to: 50 nM H/ACA-RNA, 500 nM NCG/NC/CG, 5 µM L7Ae, and 70 nM target RNAs. Annealing was done at 70 °C. Samples

containing target RNAs were annealed for 5 min, or 1 min in case of U-Sub-Cy5 and U-Sub-biotin. Samples without target RNAs were annealed for 15 min. The annealed samples were stored on ice until they were used for the actual measurement. Per sample, one channel of the surface passivated glass slides was washed with 15 µL 0.2 mg/mL streptavidin in T50-buffer (50 mM NaCl, 10 mM Tris, pH 8.0 adjusted with HCl), incubated for 1 min and then washed with 50 µL Ψ-buffer. The sample was diluted 1:100 in Ψ-buffer (c_{final}: 100 pM H/ACA-RNA) and, typically, 15 µL of the dilution were flushed through the channel. Density of immobilized molecules was checked on the microscope, and in case of satisfying molecule density, the channel was flushed with 30 µL of freshly prepared Imaging-buffer (see Chapter 3.1.1) and the measurement was started.

Eukaryotic smFRET Constructs

Annealing of archaeal samples for smFRET experiments was performed in a 4 µL scale in scΨ-buffer. If present, concentrations were adjusted to: 200 nM H/ACA-RNA, 2000 nM scNCG, 2000 nM Nhp2. In experiments containing target RNAs, the respective RNA was added to the scImaging-buffer (c_{final} = 10 nM). Samples were annealed for 5 min at 30 °C and stored on ice until they were used for the actual measurement. Per sample, one channel of the surface passivated glass slides was washed with 15 µL 0.2 mg/mL streptavidin in T50-buffer, incubated for 1 min and then washed with 50 µL scΨ-buffer. The sample was diluted to 250 – 500 pM H/ACA-RNA in scΨ-buffer and, typically, 15 µL of the dilution were flushed through the channel. Density of immobilized molecules was checked on the microscope, and in case of satisfying molecule density, the channel was flushed with 30 µL of freshly prepared scImaging-buffer (see Chapter 3.1.1) and the measurement was started.

3.4.3. smFRET Data Acquisition and Analysis

Single molecule experiments were performed on a TIRF-based microscopy setup (see Figure 3.1). The setup was equipped with a green (532 nm, 150 mW) and a red (637 nm, 140 mW) diode laser (Coherent OBIS), an acousto-optical tunable filter (AOTFnC-400.650-TN, AA Opto-Electronic), a NA 1.4 100x oil immersion objective (Olympus), an objective-type microscope (IX71, Olympus), an dual emission image splitter (OPTOSPLIT II, Cairn Research) and an electron multiplying charge coupled device (EMCCD) camera (iXon 897, Andor Technology). *Note*: The AOTF can be used for single molecule alternating-laser excitation,

which was not done in this thesis. All measurements were performed with continuous laser (532 nm) excitation. The red laser (637 nm) was only used for sample checking. Data acquisition was done with the software Andor SOLIS (version 4.21) at an integration time of 100 ms with maximum EM gain of 300. Recorded data was saved in *.sif-format and converted into 16-bit-integer raw-format (*.dat) with Andor Solis 4.2 for further processing and evaluation.

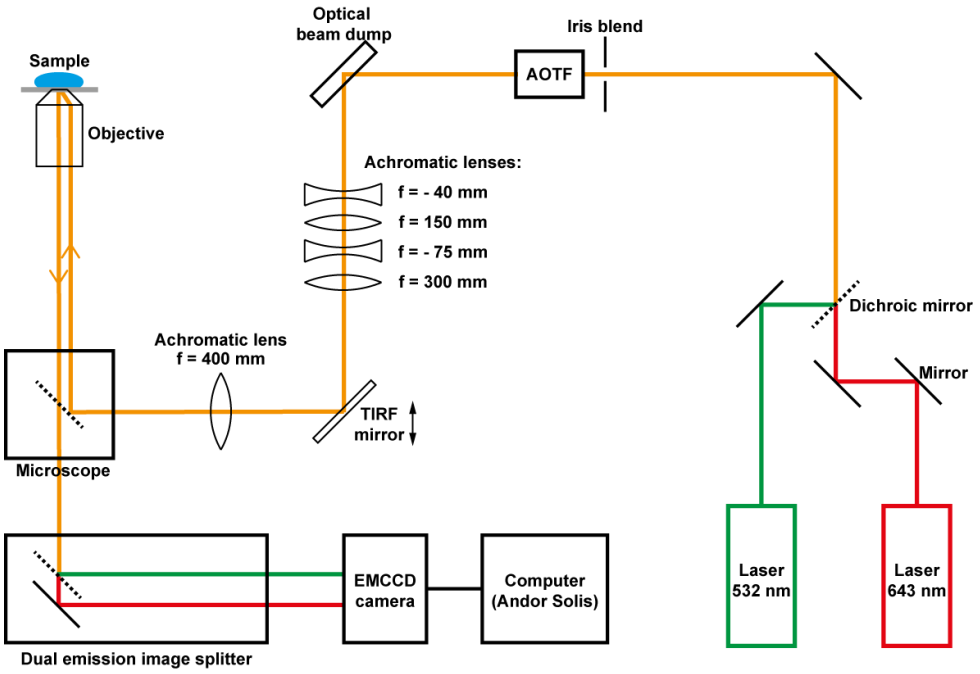


Figure 3.1: Schematic representation of the TIRF-based laser setup and the optical path for smFRET microscopy.

For sample measurement, the setup was started, adjusted to TIRF mode and run with the green laser set to 10 mW for at least 30 min. On each day smFRET experiments were performed, a 20-frame calibration movie of a 1:250 dilution of 0.1 µm TetraSpeck™ Microspheres (ThermoFisher) in T50-buffer was recorded. Typically, twenty 20-frames (2 sec) and five 1200-frames (2 min) movies were acquired per sample.

Computational data processing was done with IDL (Version 6.2, Exelis; written by Hazen Babcock and Greg Bokinsky[123]). IDL software and scripts (modified from Hazen Babcock) were kindly provided by Martin Hengesbach (Goethe University). For donor and acceptor channel alignment a transformation map was created from the calibration movie of the TetraSpecks with the IDL-scripts; *beads_co2* and *beads_map*. With the created transformation map at hand, the IDL-scripts *dat_all2*, *ffp_dat2*, *ap_dat2* and *ave_tr_dat* were used to generate background corrected smFRET fluorescence intensities for each movie by Gaussian point spread function

fitting of donor and acceptor channel. In the script *ffp_dat2*, the STD threshold was set to 3. For each recorded movie the program exported a raw-format (.dat) file into a single .traces file, which was used for further analysis in Matlab (Version R2018b, MathWorks). Matlab-scripts were kindly provided by Martin Hengesbach (Goethe University). In order to generate histograms from the sum of *.traces-files of one sample the Matlab-script *readhisttracesandor3* was used. Donor (I_D) and acceptor (I_A) intensity were averaged over 20 frames, the FRET-efficiency was calculated as E-FRET = $I_A/(I_A+I_D)$ and binned into intervals of 0.025. A threshold value 'fretthreshold' (line 20) for minimum dye-sum intensity ($I_A + I_D$) – to accept single molecules – was typically set to 5000 or 10000. The center of a donor-only peak in the histogram was adjusted to E-FRET = 0 by varying a value compensating for the acceptor-to-donor-channel fluorescence leakage (line 60). The software Origin (OriginPro 2017G, OriginLab) was used to plot the exported data and for Gaussian fitting. In order to observe trajectories from the *.traces-files of single molecules the Matlab-script *readtracesandor* was used. The fretthreshold (line 5) was set to 2000 and the value for acceptor-to-donor-channel leakage was adapted from *readhisttracesandor3* script used for the same sample. Trajectories of single molecules were evaluated individually, and a final background correction was made.

Chapter IV

Archaeal H/ACA-RNP from *Pyrococcus furiosus*

4.1. Protein-Complex Preparation and Purification

4.1.1. Preparation of the Archaeal NCG Subcomplex

Starting with the archaeal project, pure and functional H/ACA-proteins had to be prepared, in order to obtain reliable data on structure, assembly and function of H/ACA-RNPs. Previous studies on these RNA-protein complexes in this and in other working groups have laid the foundations for reconstitution of active H/ACA-RNPs in this thesis. All proteins were expressed in *E. coli* and purified chromatographically. Initial expression and purification of the ternary protein sub-complex NCG were already established in our lab. However, several challenges like complex precipitation, degradation of i.e. Gar1 and low Nop10 abundance remained, which were approached first. Therefore, *E. coli* codon optimized genes for the archaeal Nop10, Cbf5 and Gar1 proteins of *P. furiosus* were ordered from Eurofins Genomics and cloned into the pET-Duet-1 (Nop10 and Cbf5) as well as into pET-28b(+) (Gar1) vectors. Both plasmids were cotransformed into BL21(DE3) Competent *E.coli* cells from NEB for further protein coexpression and copurification. L7Ae – the last of the four archaeal H/ACA-proteins – was provided pure and RNA-free by Andreas Schmidt (Goethe University). For purification reasons some of the proteins carried six consecutive histidine residues – namely a 'HisTag' – at either the C- or N-terminus.

At the beginning of this thesis, non-codon optimized genes were used for expression of the NCG complex, as described[119]. After IPTG induced protein over expression and subsequent cell lysis, the lysate was heated to 70 °C, in order to denature some of the prokaryotic proteins, while the archaeal proteins from the thermophilic organism *P. furiosus* remained in solution. Then polyethylenimine was added to the lysate for nucleic acid precipitation. For the first chromatographic purification step the two incorporated HisTags at Cbf5 (N-terminal) and Gar1 (C-terminal) were used in a nickel(II) nitrilotriacetic acid (Ni-NTA) affinity chromatography. The consecutive histidine residues of the HisTag carrying proteins bind to Ni^{2+} in the column resin, while most other proteins do not interact with the column. Elution of bound proteins is achieved by an increase of imidazole concentration, which competes with the HisTag to the nickel. A gradient from 20 mM to 500 mM imidazole was used for the affinity chromatography. The resulting fractions are than typically analyzed by UV-detection and subsequent SDS-PAGE (see Figure 4.1).

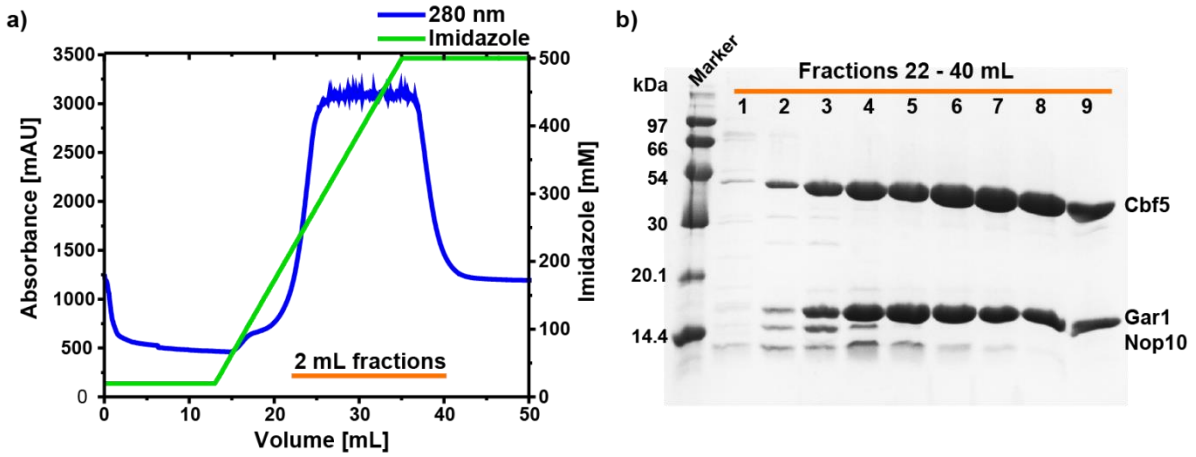


Figure 4.1: ***a)*** **Ni-NTA-chromatogram of non-codon-optimized NCG expression on a 1 mL Ni-NTA-column.** ***b)*** **Resulting 15% SDS-PAGE of fractions showing high UV-absorption.**

The strong absorption detected in the gradient elution indicates high NCG expression (see Figure 4.1a). In the analytical gel, it is visible that all three proteins Cbf5, Gar1 and Nop10 were eluted from the column (see Figure 4.1b). However, Nop10 expression seems to be not as effective as for the other two proteins, since from fraction 6 onwards the band of the smallest protein is significantly less present or completely vanished compared to Gar1 and to Cbf5. In fractions 2 to 4 an additional protein band is observable, which might be due to the above mentioned Gar1 degradation. For previous experiments the two fractions showing the highest Nop10 presence were combined and used for further purification; fractions 4 and 5 in this case. The united fractions were concentrated and loaded onto a size exclusion chromatography (SEC) column. SEC is used to separate macromolecules according to their hydrodynamic diameter under a permanent flow of eluent and pressure. Large molecules are eluted prior to small molecules. Figure 4.2 shows a representative chromatogram of such a SEC of an old NCG construct and a typical SDS-gel of the analyzed fractions.

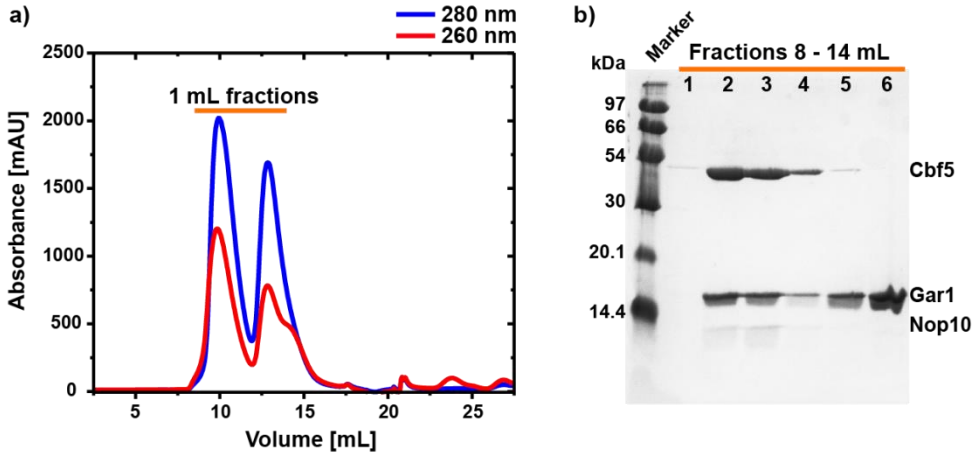


Figure 4.2: ***a)*** **NCG size exclusion chromatogram of combined fractions 4 and 5 of the Ni-NTA-chromatography on a Superdex 75 10/300 GL.** ***b)*** **Resulting 15% SDS-PAGE of fractions showing high UV-absorption.**

In the size exclusion chromatogram, two distinctive peaks can be seen (see Figure 4.2a). Their electrophoretic analysis showed the NCG complex in fraction 2 to 4 and single Gar1 protein in fractions 5 and 6 (see Figure 4.2b). The Nop10 band in fraction 2 and 3 is significantly weaker compared to the other proteins. This could be due to the fact that Nop10 as the smallest protein (7.4 kDa) is not as well stained as Cbf5 (39.6 kDa) and Gar1 (12.8 kDa with HisTag), or due to the possibility that expression of Nop10 from the pET-Duet-1 vector does not work as well as for Cbf5. The latter would indicate that Nop10 is not equally present in the purified protein complex and some complexes would only consist of Cbf5 and Gar1. The size exclusion was done in oSEC-buffer (with additional 1 mM DTT and 0.5 mM EDTA) and yielded typically 1 mg NCG proteins out of 500 mL cell culture.

To solve the problem of low Nop10 abundance in the purified protein complex, NCG expression was performed from *E. coli* codon optimized genes in the same expression platform. For the new constructs Gar1 (11.9 kDa without HisTag) did not carry an additional HisTag. Hence, pulldown of the ternary complex from the Ni-NTA-column must be achieved over Cbf5 only. After successful test expressions, the cell cultures were up scaled and stored for further procedures. The first steps of purification for the new codon optimized NCG proteins were performed in a similar manner: heat-shock, nucleic acid precipitation and Ni-NTA-chromatography.

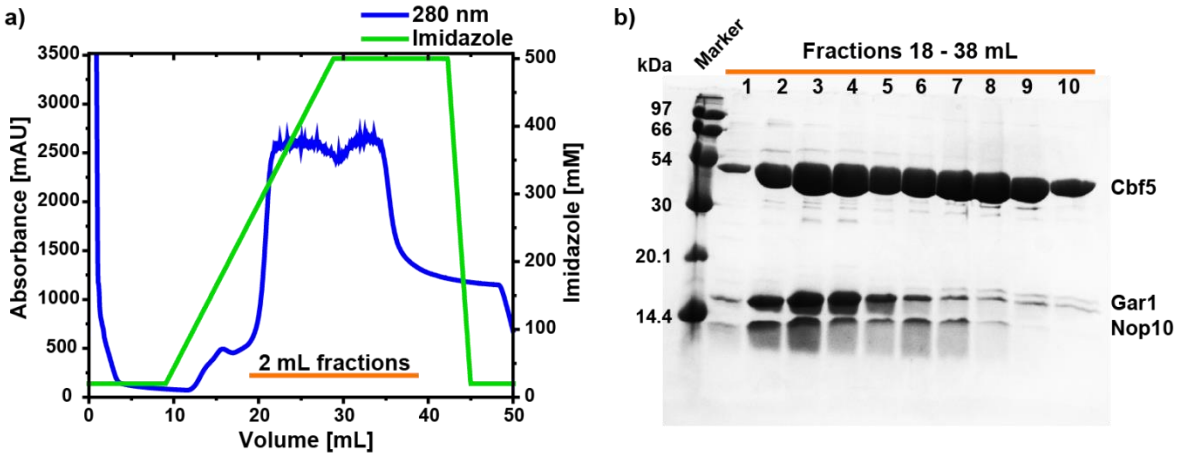


Figure 4.3: ***a)*** **Ni-NTA-chromatogram of codon-optimized NCG expression on a 1 mL Ni-NTA-column.** ***b)*** **Resulting 15% SDS-PAGE of fractions showing high UV-absorption.**

In the affinity chromatogram it is apparent that high amounts of protein could be purified by the Ni-NTA-column (see Figure 4.3a). The gel analysis shows high expression of all three proteins: Cbf5, Gar1 and Nop10 (see Figure 4.3b). For the next purification step fraction 3 was applied to the gel filtration column. A new high salt buffer containing 1 M NaCl (nSEC-buffer) was used, adapted from Yang *et al.*[64].

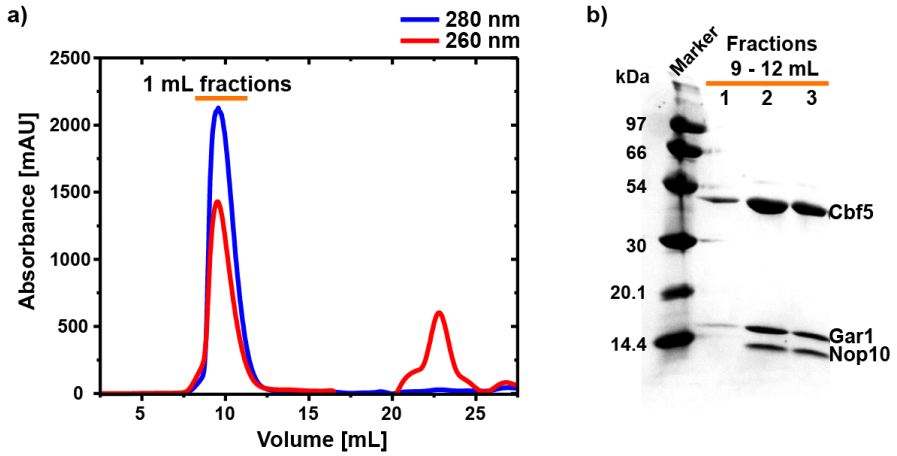


Figure 4.4: ***a)*** **NCG size exclusion chromatogram of fraction 3 of the Ni-NTA-chromatography on a Superdex 75 10/300 GL.** ***b)*** **Resulting 15% SDS-PAGE of fractions showing high UV-absorption.**

The size exclusion of the codon optimized NCG expression showed one distinct peak with the presence of all three proteins Nop10, Cbf5 and Gar1 (see Figure 4.4a). For the new constructs the Nop10 band is significantly more present compared to the old constructs (see Figure 4.4b). Only fractions 2 and 3 were combined, concentrated and stored at 4 °C, as fraction 1 showed a higher 260/280 nm ratio in the absorption spectrum than the other two fractions. The protein preparation of the codon optimized constructs yielded approximately 1 mg NCG per 500 mL expression

culture. After several days of storage in the high salt nSEC-buffer neither visible precipitation of the NCG complex, nor distinguishable degradation of Gar1 were detected.

Another approach to purify the ternary NCG complex was the use of ion-exchange chromatography (IEC) instead of SEC. Therefore, a cation exchange column based on a negatively charged heparin resin was used. The NCG complex has a positively charged surface and thus is capable of binding RNAs and other negatively charged molecules. Hence, the positive surface of the protein complex should bind to negatively charged heparin column. Elution can be achieved by adding high salt buffers. The high amount of cations in the buffer competes with the protein-heparin interaction and leads to dissociation of the complex at a certain salt concentration. For the IEC, a gradient from 500 mM to 1200 mM NaCl was used. The test was performed according to Li *et al.*[52] with fraction 7 of the Ni-NTA-chromatography (Figure 4.3) after more than two weeks storage in nSEC-buffer at 4 °C. Before the sample was applied to the column the sample was diluted to a NaCl concentration of 500 mM.

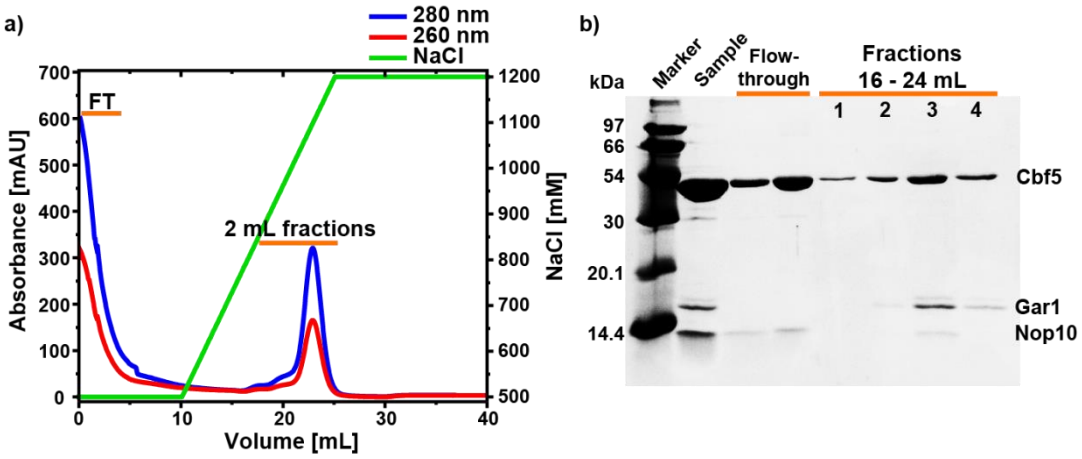


Figure 4.5: ***a)*** **NCG ion-exchange chromatogram of fraction 7 of the Ni-NTA-chromatography on a 1 mL HiTrap Heparin column.** ***b)*** **Resulting 15% SDS-PAGE of fractions showing high UV-absorption**

In the IEC two major peaks can be detected (see Figure 4.5a). The first peak basically resembles the flow-through (FT) of the sample application, which is unbound protein. The second major peak results from the gradient elution at an estimated concentration of 1000 mM NaCl. The SDS-gel shows the distinct Cbf5, Gar and Nop10 bands. Neither of the proteins seems to be degraded. However, Gar1 is missing in the FT fractions, so the NCG complex appears to be partially dissociated after being stored for several days in nSEC at 4 °C. Nevertheless, fractions 3 to 4 contain the ternary NCG complex with all three proteins and the 260/280 nm ratio suggests an RNA-free protein. Hence, pure NCG can be obtained by SEC or IEC after initial affinity

chromatography. After the final purification step the fraction with the highest UV absorption was aliquoted and stored at -80 °C.

4.1.2. Preparation of NC and CG Subcomplexes

For experiments lacking one or more of the three proteins, individual expression platforms were designed. Removal of Nop10 was achieved by using a pET-Duet-1 vector carrying only the Cbf5 gene for the double transformation of both plasmids into *E. coli*. Exclusion of Gar1 was accomplished by single transformation of the pET-Duet-1 vector – carrying Cbf5 and Nop10 genes – into competent expression cells.

Expression and purification of the partial Nop10-Cbf5 'NC' and Cbf5-Gar1 'CG' subcomplexes was performed in the same way as for the ternary NCG complex. After heat and polyethylenimine treatment, the proteins were purified with affinity and subsequent size exclusion chromatography. In the following only the resulting SDS-PAGE gels of the Ni-NTA-chromatograms and the gel filtration chromatogram will be shown, starting with the purification of the NC complex.

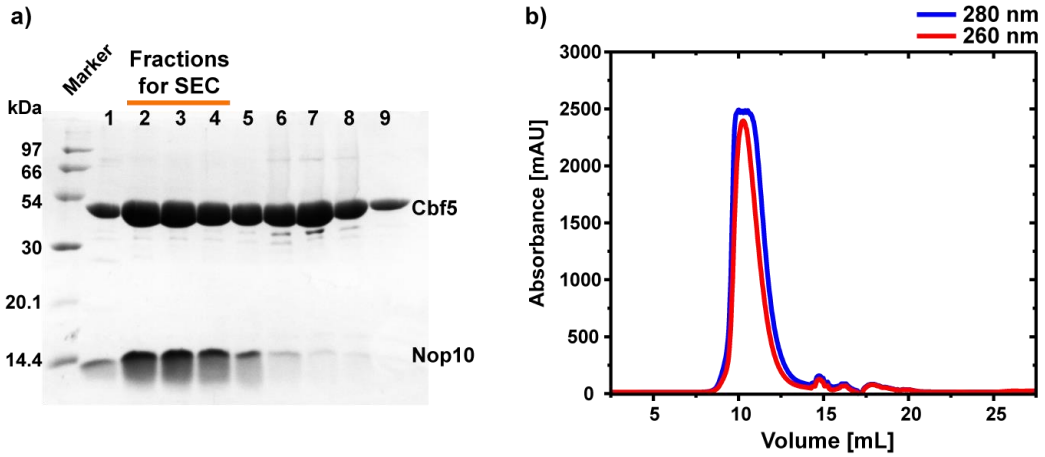


Figure 4.6: ***a)*** **15% SDS-PAGE of Ni-NTA-chromatography for the NC complex on a 1 mL Ni-NTA-column.** ***b)*** **NC size exclusion chromatogram of combined fractions 2 to 4 on a Superdex 75 10/300 GL.**

The analytic gel shows high abundance of both proteins (see Figure 4.6a). Fractions 2 to 4 showed high amounts of both proteins and were therefore used for further purification. In the size exclusion chromatogram of NC one distinct peak is observable (see Figure 4.6b). The 260/280 nm ratio in might suggest RNA impurities, but most probably the absorbance at 280 nm reached a maximum (indicated by the fluctuation at the top of the peak), due to high amount of proteins.

Thus, the detected absorption is out of the linear range where the Beer-Lambert law applies. The fraction with the highest UV absorption was aliquoted and stored at -80 °C

Preparation of the Cbf5-Gar1 complex was performed analogous as previous complexes (see Figure 4.7). The SDS-PAGE of the gradient elution showed high abundance of both proteins. Fraction 3 showed the least impurities and was used for SEC. In the size exclusion chromatogram only one distinct peak with a good 260/280 nm ratio was observed. The fraction with the highest UV absorption was aliquoted and stored at -80 °C.

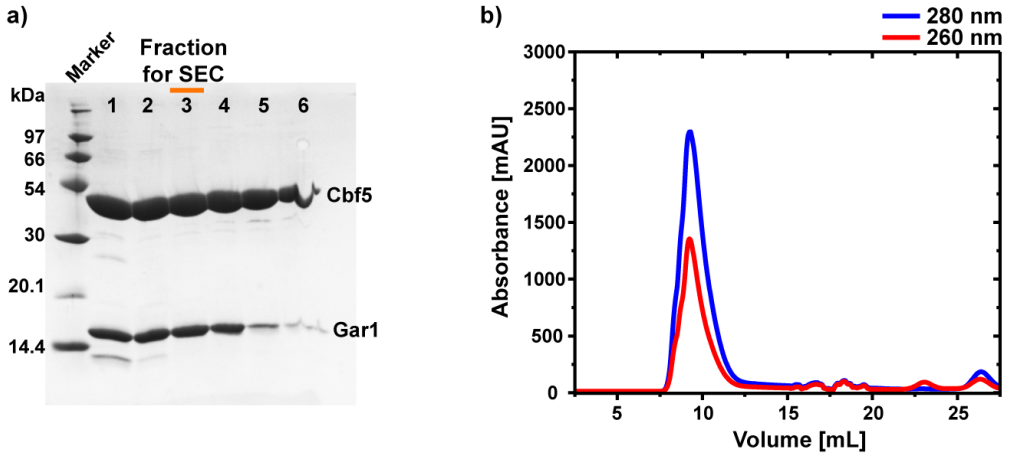


Figure 4.7: ***a)*** **15% SDS-PAGE of Ni-NTA-chromatography for the CG complex on a 1 mL Ni-NTA-column.** ***b)*** **CG size exclusion of fraction 3 of the Ni-NTA-chromatography on a Superdex 75 10/300 GL.**

In summary, expression and purification of the codon optimized protein constructs was successful for all archaeal H/ACA-proteins. The ternary NCG as well as the partial NC and CG complexes were obtained highly pure and free of significant RNA contaminations.

4.2. Synthesis, Labeling and Purification of sRNA Constructs

4.2.1. Wildtype Archaeal H/ACA-RNA

For single molecule FRET spectroscopy fluorophore labeled constructs are required, and in life science, typically small organic dyes are used to label the biomolecule of interest. As previously mentioned, the introduction of such dyes might lead to undesired interactions, such as docking of the fluorophore to the RNP, which might interfere with complex activity and/or FRET-experiments. To test if the obtained data by smFRET is reliable, the biochemical integrity (e.g. enzymatic activity or complex assembly) of the labeled constructs has to be compared to the unlabeled wildtype 'wt' complex. Thus, sufficient wt-H/ACA-RNA has to be synthesized and purified for a comparison of unlabeled to labeled H/ACA-RNPs. Another way to determine the significance of interactions between fluorophores and biomolecules is to measure the rotational freedom of the dyes. This is discussed particularly in Chapter 4.3.2.

Enzymatic synthesis of the wt-H/ACA-RNA was achieved by amplification of the corresponding DNA template (75mer) by standard PCR and subsequent T7-RNA polymerase *in vitro* transcription. The resulting H/ACA-RNA has an expected length of 58 nt (see Figure 4.8a). The transcription reaction was purified according to Chapter 3.1.3, and analyzed via denaturing PAGE by comparison to a previously synthesized H/ACA-RNA 'Ref. H/ACA' (see Figure 4.8b), of which the correct size was verified by mass spectrometry.

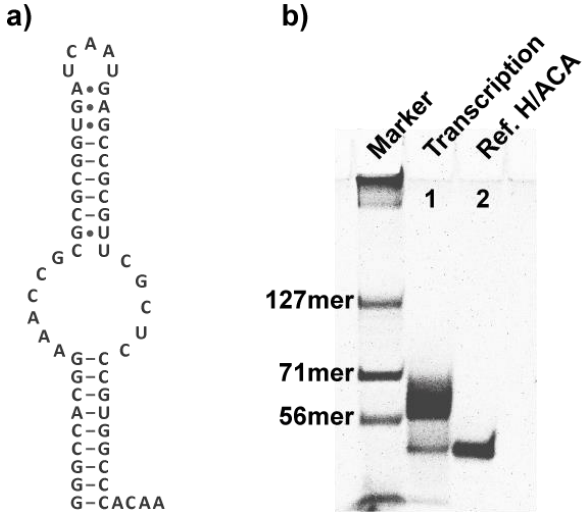


Figure 4.8: ***a)*** **Secondary structure of** ***pf*** **wt-H/ACA-RNA.** ***b)*** **10% denaturing PAGE of the wt-H/ACA-RNA transcription. RNAs were stained with GelRed.**

In Figure 4.8b two bands – a small band below and a big band above the 56mer of the marker (provided by Heidi Zetsche from the Goethe University) – for the wt-H/ACA-RNA transcription (sample 1) can be observed. The small band runs at the same height as the reference 'Ref. H/ACA' (sample 2), suggesting it to be the desired RNA. However, the larger band – which is probably an overlay of several other transcription products – is located slightly above the 56mer of the reference sample, indicating this to be the correct RNA. Hence, a clear characterization of the transcribed sRNAs could not be obtained by PAGE. Thus, both RNAs were compared by mass spectrometry (Spectrograms are shown in Appendix A2 – Mass Spectrometry Data). The expected size of the wt-H/ACA-RNA is 18.9 kDa and detected masses are given in Table 4.1.

Construct	Mass (m/z)
Transcription (1)	18973
Ref. H/ACA (2)	18928, 9875

Table 4.1: Detected masses for *in vitro* transcribed H/ACA-RNA constructs.

In both samples the expected mass of 18.9 kDa was detected and represented the main peak. For the reference sample (2) an additional peak at mass-to-charge ratio (m/z) of 9875 was found, resembling the two-fold charged RNA. Overall, the detected peaks in the spectrograms were not homogeneous and the signal-to-noise ratio of sample 1 was significantly lower compared to sample 2. A possible explanation for these observations might be a high salt concentration in both samples, which impedes with mass spectrometry. A series of tests with varying PCR, transcription and purification conditions gave no further insights into the appearance of the additional gel bands of sample 1, or into the fact that the RNA band of sample 2 (58mer) migrates below the 56mer of the applied marker. Several *in vitro* transcription reactions with the same conditions yielded different outcomes. One possible explanation for this behavior may be the high GC content (69%) of the thermophile H/ACA-RNA, resulting in a high melting temperature (89.4 °C according to the online tool 'OligoCalc'[124]). Formed secondary structures (e.g. hairpin motif), or dimers, might be sufficiently stable to withstand the denaturing conditions of the analytical PAGE. Hence, these RNA-structures might migrate differently than anticipated in the gel.

4.2.2. Fluorophore Labeled H/ACA-RNAs

Site specific labeling of the complex of the H/ACA-RNP is essential to obtain reliable structural information during biochemical processes; such as complex assembly, substrate recruitment or the actual isomerization. For this purpose four labeling sites in the sRNA were chosen from a published crystal structure of an archaeal H/ACA-RNP (PDB: 3HJW). The selected sites were located at the 5'-end (G1), at the beginning of the hairpin loop (U26), in the upper stem (C37) and next to the pseudouridylation pocket (U39) of the H/ACA-RNA (see Figure 4.9). For site specific RNA labeling, nucleotides U26 and U39 were functionalized at the C5-position of the base with an aminoallyl group, G1 was functionalized at the phosphate group with a 5'-amino-C6-modifier, and the 2'-OH of C37 was substituted with an amino-group (see Figure 4.10). In the early stages of this project the C37 position did not prove to be a good choice. In preliminary FRET-measurements constructs labeled at this position yielded significantly lower observable molecules. Thus, this labeling site will not be discussed any further.

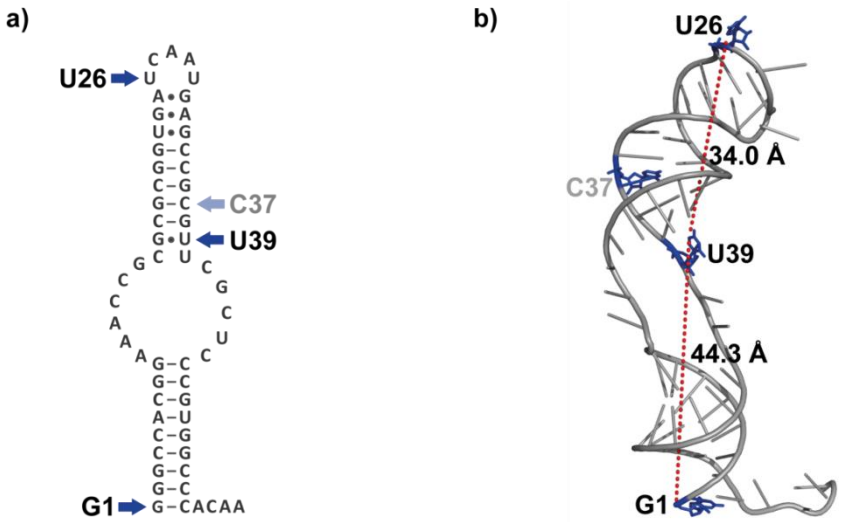


Figure 4.9: ***a)*** **Secondary structure of archaeal H/ACA-RNA with selected labeling sites (blue arrows).** ***b)*** **Crystal structure of archaeal H/ACA-RNA (PDB: 3HJW)[55] with highlighted modified nucleotides (blue) and intramolecular distances of G1-U39 and U26-U39 labeling sites (red). Proteins from the crystal structure were omitted in this figure.**

Fluorophore attachment to the RNA was achieved by usage of primary amines and NHS-esters according to the previously introduced coupling reaction (see Chapter 1.5). For site specific labeling of the H/ACA-RNA, the 58mer was divided into two oligomers: a 5'-strand (30mer) and a 3'-strand (28mer). These purchased short RNA strands were synthesized via solid phase synthesis and preserved by 2'-ACE protecting groups. Hence, nucleotide modifications for G1

and U26 (5'-oligos) or for U39 and C37 (3'-oligos) were inserted site specifically with each oligomer carrying only one functional group for dye coupling. 3'-Oligomers were either purchased with or without a biotin linker at the 3'-end (bio-3'-oligo), depending on the desired complex immobilization for smFRET measurements. The above mentioned RNA modifications and reactive fluorophores for dye coupling are shown in Figure 4.10.

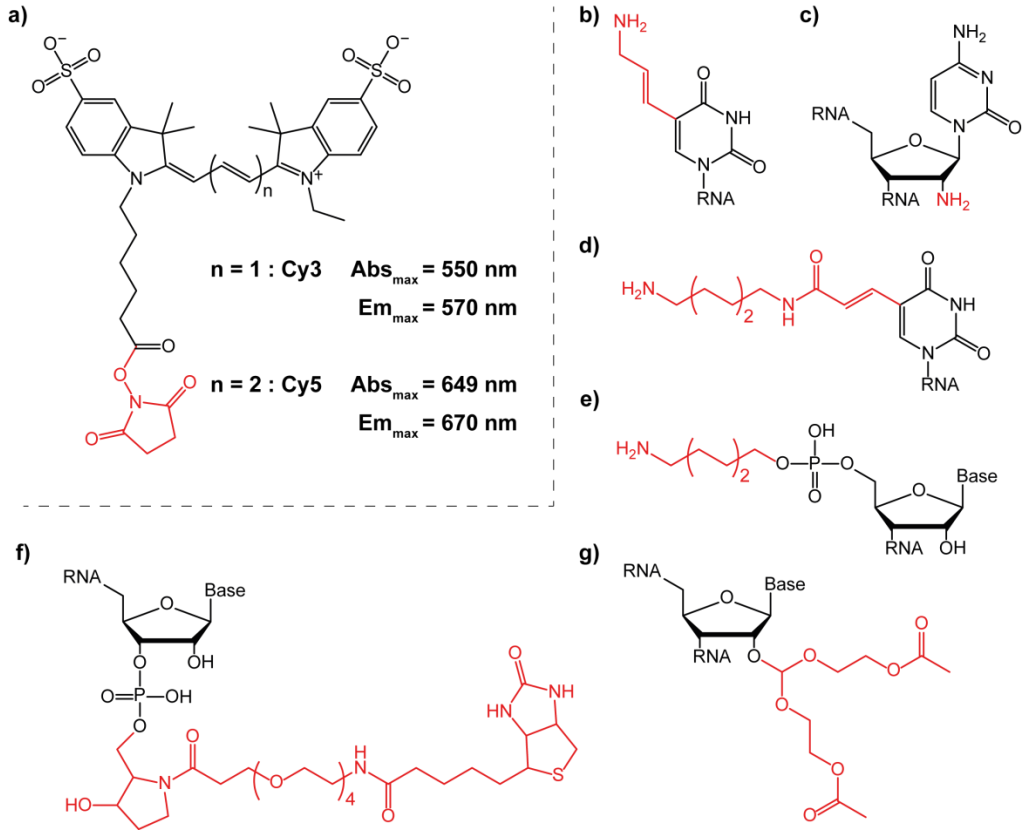


Figure 4.10: Synthetic RNA modifications and organic CyDyes. Functional groups are highlighted in red. ***a)*** **Reactive Cy Dyes with their respective absorption and emission maximum.** ***b)*** **Aminoallyl uridine.** ***c)*** **2'-Amino cytidine.** ***d)*** **Aminohexylacrylamino uridine.** ***e)*** **5'-6-Aminohexyl phosphate.** ***f)*** **Biotin modified 3'-end.** ***g)*** **2'-bis(2-Acetoxyethoxy)methyl 'ACE' protecting group.**

Preparation of the labeled H/ACA-RNAs can be divided into three steps: 1) fluorophore labeling and deprotection of the individual oligomers, 2) HPLC purification of the coupling reaction, and 3) enzymatic splinted ligation of the modified RNA strands to the full-length labeled H/ACA-RNA. To obtain dual labeled FRET-H/ACA-constructs, one oligomer was labeled with either Cy3 or Cy5, and vice versa. For single labeled Cy-H/ACA-constructs, one of the two synthesized RNA strands was left unmodified, and instead directly deprotected and used for splinted ligation. The labeling reaction was performed at room temperature, and afterwards the oligomers were deprotected and prepared for reversed phased HPLC purification. In this technique the RNA is bound to a hydrophobic stationary phase of the column. Elution can be achieved by applying a

gradient from a polar (e.g. TEAA-buffer in H_2O) to a less polar solvent (e.g. MeCN). For the dye labeled RNAs, the elution is expected at a higher percentage of the less polar solvent, since the organic dyes increase the hydrophilic properties of the RNA. Representative HPLC chromatograms for the labeling reactions of the biotinylated oligomer *bio-3'-U39* and unbiotinylated oligomer *3'-U39* to Cy3 and Cy5 respectively are shown in Figure 4.11.

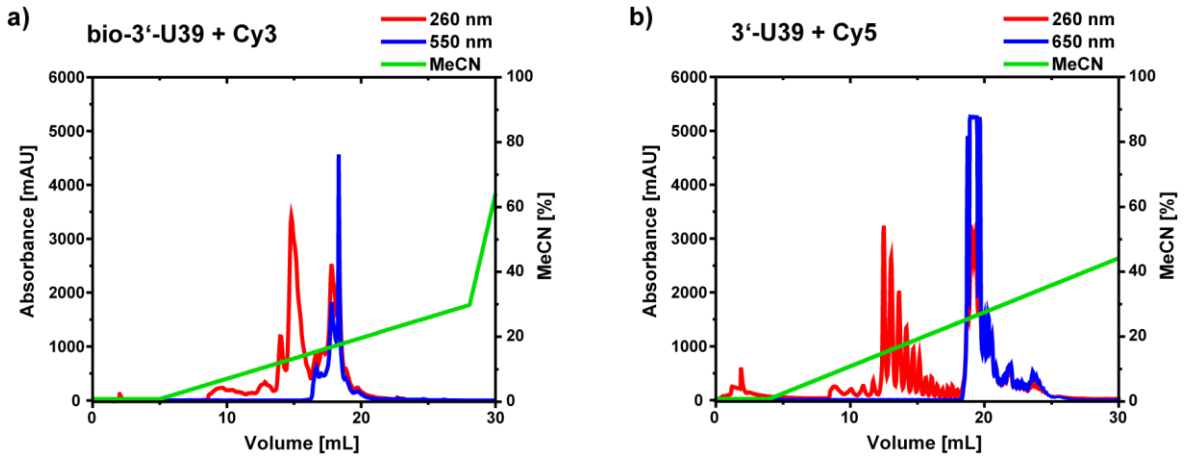


Figure 4.11: ***a)*** **HPLC chromatogram of the labeling reaction of oligomer 3'-U39-bio to Cy3 on a reversed phased C8-column.** ***b)*** **HPLC chromatogram of the labeling reaction of oligomer 3'-U39 to Cy5 on a reversed phased C8-column.**

Both chromatograms show a clear separation of labeled from unlabeled RNAs. In Figure 4.11b wave packeting of the eluted peaks can be observed. This phenomenon was due to air in one of the two pumps of the chromatography system, which lead to a detectable and periodical loss of pressure at the column during the HPLC run. Nonetheless, purification of the applied oligomer was successful. Elution of the biotinylated oligomer *bio-3'-U39-Cy3* started at 16% MeCN (see Figure 4.11a), while the unbiotinylated *3'-U39-Cy5* eluted from 24% MeCN upwards (see Figure 4.11b). Labeling was efficient for the two constructs, but not quantitative. Both chromatograms show a significant amount (~50%) of unlabeled RNA (absorbance at 260 nm only). Nevertheless, the obtained amount of labeled oligomers was sufficient for further construct synthesis. Fractions containing high dye-specific absorption (Cy3: 550 nm / Cy5: 650 nm) were pooled and used for analysis and further enzymatic modification.

Splinted ligation of the modified or unmodified oligomers to the full-length H/ACA-RNA was done with either T4-DNA-ligase (using a mix of NEB and ThermoFisher enzymes) or homemade T4-RNA-ligase 2. Prior to the ligation equimolar amounts of the DNA-splint and both RNA oligomers were annealed to form a RNA:DNA hybrid. In order to obtain specific base pairing, the reaction mixture was heated to 85 °C and slowly cooled down to room temperature. Once the

annealing process was completed, the ligase was added for the actual reaction, and the mixture was heated to 37 °C for optimal enzymatic activity. After two to three hours DNase was added for digestion of the DNA template, to be able to purify DNA-free H/ACA-constructs. As final purification step and to separate the ligated constructs from the free oligomers, PAGE was performed. The product band could be detected by eye and was excised and eluted from the polyacrylamide gel. In a series of test ligations for double and single labeled H/ACA-RNAs the enzyme activities of T4-DNA-ligase (DnL) and T4-RNA-ligase 2 (RnL) were compared (see Figure 4.12).

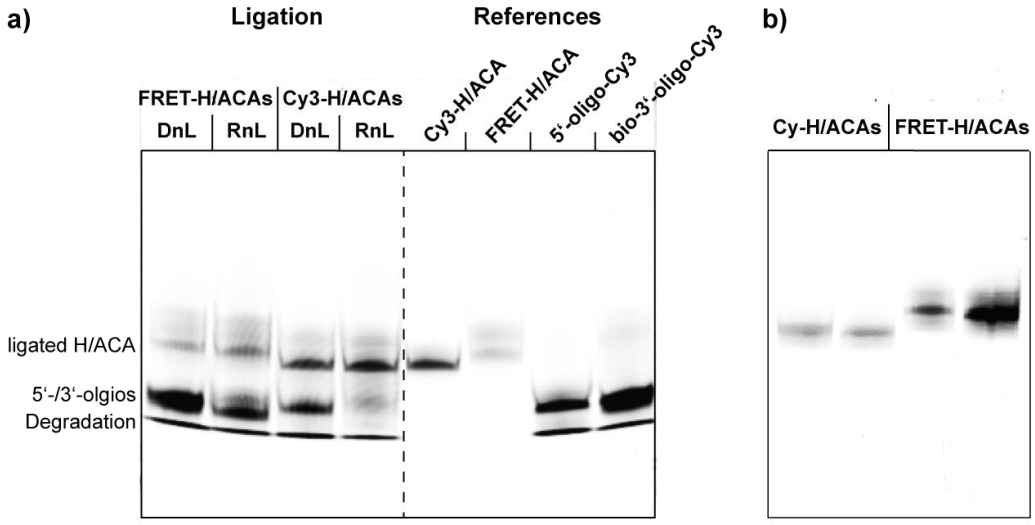


Figure 4.12: ***a)*** **10% denaturing PAGE of H/ACA-RNA test ligations and comparisons of DnL versus RnL for two different constructs (Typhoon settings: Cy3-scan).** ***b)*** **10% denaturing PAGE of purified and labeled H/ACA-constructs (Typhoon settings: FRET-scan).**

From the depicted gel it becomes clear, that the splinted ligation works well for both ligation enzymes, but the RNA ligase shows slightly better ligation efficiencies for both constructs. This is indicated by lower fluorescence intensities for the 5'-/3'-oligonucleotide bands in case of the RnL compared to the DnL lanes (see Figure 4.12a). A fluorescence scan of the excised RNAs' analytical denaturing PAGE shows one distinct band for each H/ACA-construct (see Figure 4.12b). The migration difference in the gel for FRET-H/ACAs compared to the Cy-H/ACAs results from the number of dyes attached to the RNA: The FRET constructs carry two fluorophores, while Cy-H/ACAs are single labeled RNAs. Due to size and charge of the fluorophores the migration of the RNAs is altered in the gel. Typically, 1000 pmol of each nucleic acid were used for one labeling reaction, and yielded around 140 pmol (14%) of purified H/ACA-constructs. However, ligation yields were subject to change, mostly depending on the

used quantities of starting materials (nucleic acids and ligases). But also RNA degradation due to RNase contamination may decrease the amount of purified and labeled H/ACA-RNAs. In general, higher reaction scales lead to higher overall yields (up to 25%), presumably because a specific amount of sample is always lost for each purification step (e.g. ethanol precipitation and gel elution). Schematic depictions of synthesized FRET- and Cy-H/ACA-constructs for smFRET experiments are displayed in Figure 4.13.

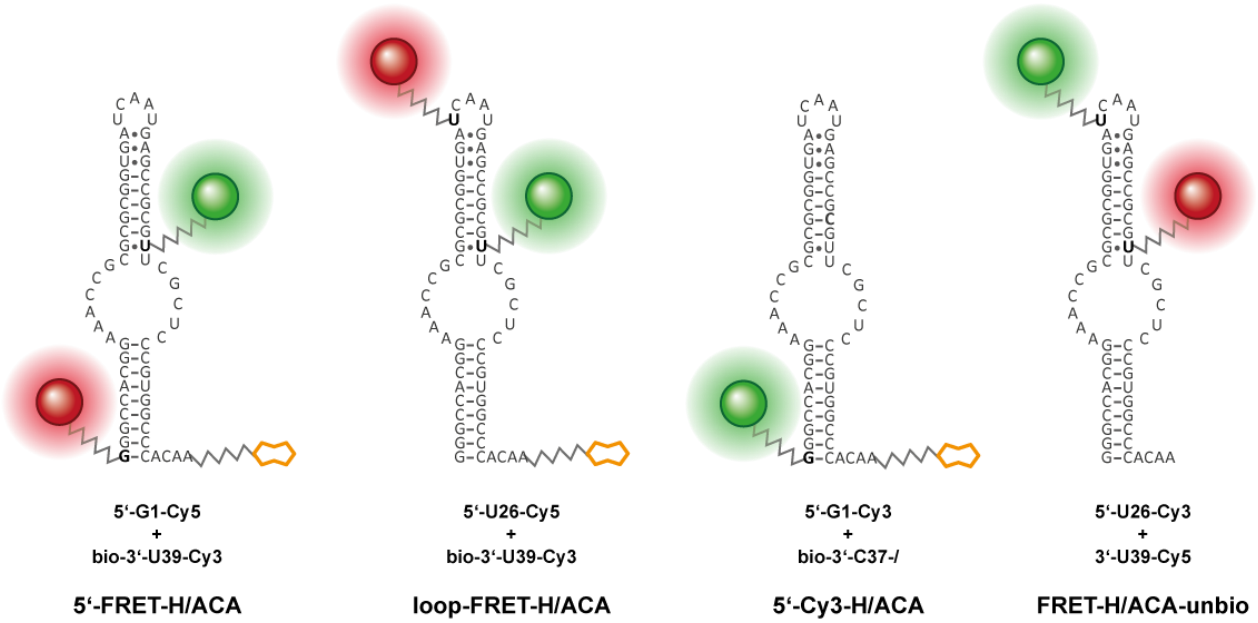


Figure 4.13: Secondary structure of synthesized *pf* H/ACA-constructs for smFRET experiments. RNA-labels are marked in different colors: Cy3 (green), Cy5 (red) and biotin (orange). The amino modified nucleotides are highlighted in bold black characters and linkers are indicated as jagged lines.

4.2.3. Target RNAs

Besides labeling of the guide RNA, chemical modification of substrate like RNAs was performed to gain further insights into the processes of RNA guided pseudouridylation. Therefore, a set of chemically synthesized target RNAs (X-Sub) with altered target nucleotides (see Figure 4.14a) was purchased. These substrate analogues carried a C6-amino modifier at the terminal 3'-phosphate of the RNA for further labeling reactions (see Figure 4.14b). This modification was used to modify the target RNAs with either a biotin linker (see Figure 4.14c) or with a fluorophore, in both cases using NHS-amine-coupling chemistry.

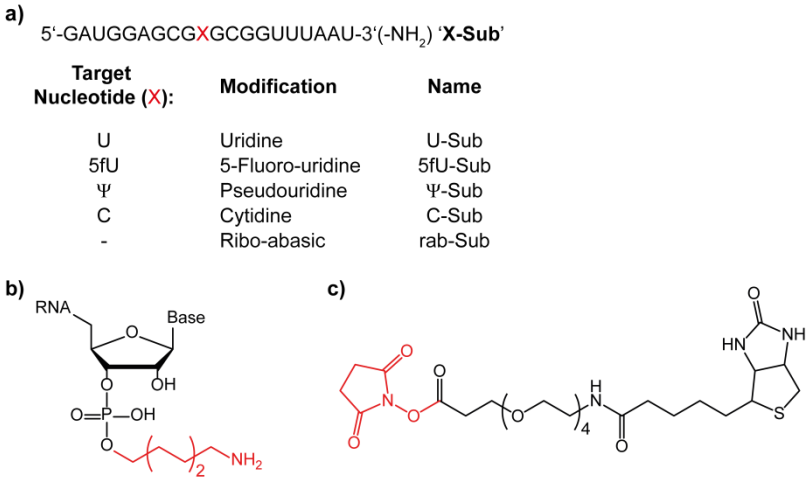


Figure 4.14: ***a)*** **Sequence of the target RNA and target nucleotide (red) base modifications.** ***b)*** **Functionalized 3'-end with an aminohexyl phosphate group (red).** ***c)*** **NHS-biotin.**

Labeling reactions and deprotection of the target RNAs were executed according to the procedure of the oligomers for the fluorophore labeled H/ACA-constructs. After chemical coupling HPLC was performed to separate labeled from unlabeled RNAs. Representative HPLC chromatograms for Cy5 or biotin labeled X-Subs are shown in Figure 4.15.

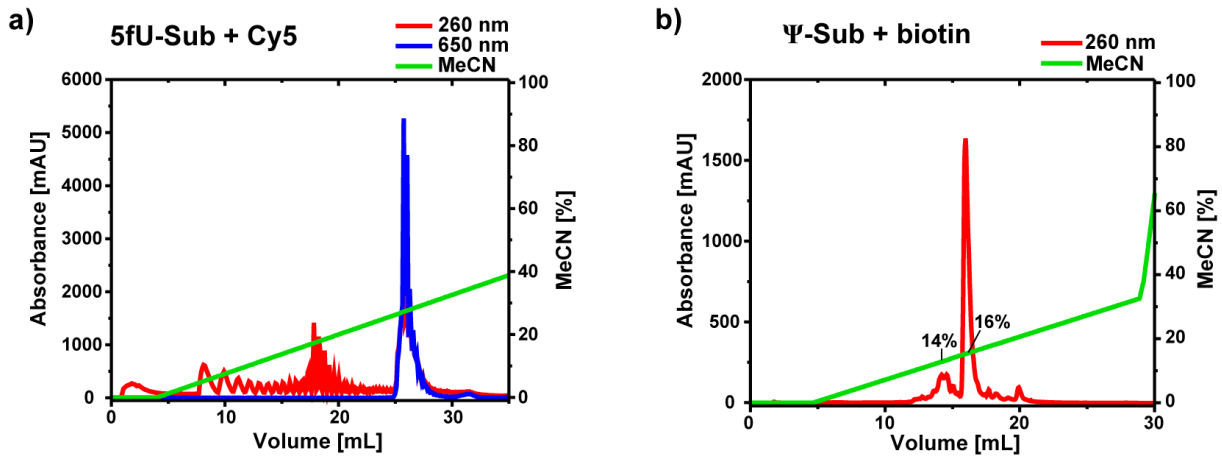


Figure 4.15: ***a)*** **HPLC chromatogram of the labeling reaction of 5fU-sub to Cy5 on a reversed phased C8-column.** ***b)*** **HPLC chromatogram of the labeling reaction of Ψ-Sub to a biotin linker on a reversed phased C8-column.**

For the fluorophore labeled 5fU-Sub-Cy5, a clear separation of unlabeled (elution at 17% MeCN) to labeled (elution at 26% MeCN) RNA was achieved by HPLC (see Figure 4.15a). Here again, wave packeting for the eluted peaks can be observed, for the same reasons as described above (compare Figure 4.11b, page 84). Fractions containing high Cy5 (650 nm) and RNA (260 nm) absorbance were pooled and used for further experiments. The remaining X-Sub-Cy5 constructs were provided by Andreas Schmidt (Goethe University). Labeling and purification of these RNAs was done with the same protocol.

For the biotinylated Ψ-Sub-bio, a separation of one minor (elution at 14% MeCN) to one major peak (elution at 16% MeCN) could be observed (see Figure 4.15b). Since successful labeling of the target RNA could not be determined by UV/Vis-spectroscopy, the fractions of the major peak were pooled and analyzed by mass spectrometry (Spectrograms are shown in Appendix A2 – Mass Spectrometry Data). The results for the various target RNAs are shown in Table 4.2.

Construct	Mass (m/z)	Expected (g/mol)
U-Sub-bio	7230.6	7116.1
5fU-Sub-bio	7187.5	7134.1
Ψ-Sub-bio	7119.8	7116.1
C-Sub-bio	7326.7	7115.1
rab-Sub-bio	n.d.	7006.0

Table 4.2: Measured and calculated masses of biotinylated target RNAs. For rab-Sub-bio m/z was not determined (n. d.).

The detected m/z for Ψ-Sub-bio was in good agreement with its expected mass. For the other biotinylated target RNAs, small differences in calculated and measured masses were observed. These differences could not be assigned to common deviations in mass spectrometry like salts, degradation or incomplete deprotection of the nucleotides in particular. However, since in no spectrum detectable signals of unbiotinylated target RNAs – calculated mass minus 473.2 g/mol – were found, labeling and purification were deemed successful for all X-Sub-bio constructs.

4.3. Reconstitution of Labeled and Functional Archaeal H/ACA-RNPs

4.3.1. RNA-Protein Binding

Unlabeled H/ACA-RNPs

After successful preparation of all components for the archaeal H/ACA-RNP, the biochemical function of the labeled constructs had to be compared to the unlabeled wt-RNAs. In order to assess whether the dyes interfere with H/ACA-RNP assembly, RNA-protein binding was tested via an electrophoretic mobility shift assay (EMSA). EMSA is an affinity electrophoresis, which allows identifying DNA- or RNA-binding proteins by performing a native PAGE. In general, several samples with constant concentration of one of the potential binding partners (e.g. RNA) and with increasing concentration of the other binding partner (e.g. protein) are prepared and applied to the gel. If the protein binds to the RNA one would expect a different migration pattern in the gel electrophoresis for the free RNA compared to the RNP. Hence, RNA-protein complexes can be visualized and distinguished from free RNAs.

The binding affinity of the labeled H/ACA-constructs was compared to the wt-H/ACA-RNA, starting with EMSAs for the unlabeled guide RNA. Therefore, a series of measurements for L7Ae and the ternary protein complex NCG were performed. For each series five samples with constant RNA concentration (1 µM) and increasing protein concentrations (0 – 20 µM for L7Ae and 0 – 10 µM for NCG) were prepared, annealed for 15 min at 70 °C and analyzed by native PAGE (see Figure 4.16). Furthermore, for NCG and L7Ae the H/ACA-RNA was taken from different transcription reactions; transcription product 1 – showing several bands in the denaturing PAGE – for NCG, and transcription product 2 – showing one band in the denaturing PAGE – for L7Ae (compare Figure 4.8b on page 80).

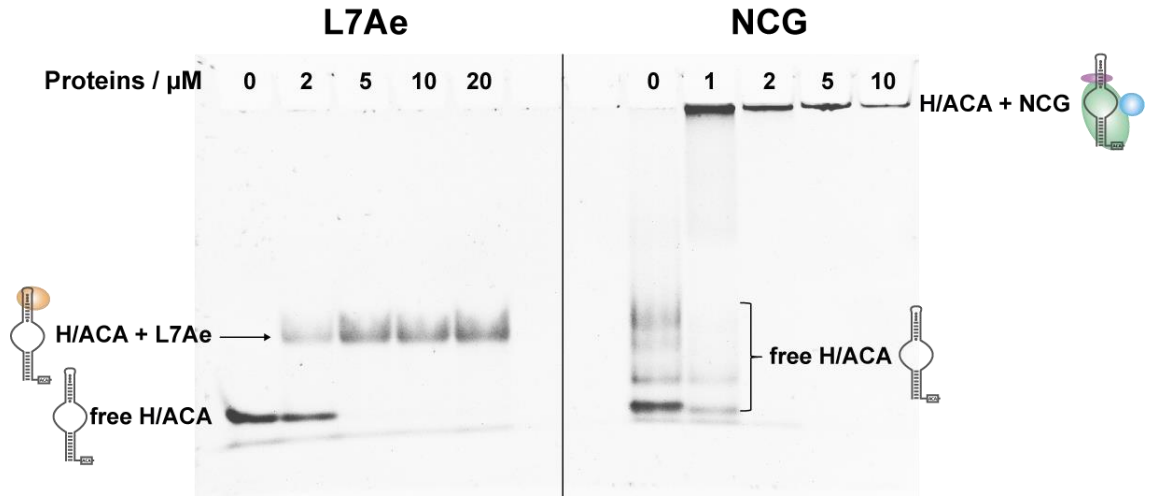


Figure 4.16: Native 8% PAGE of L7Ae and NCG with wt-H/ACA-RNA in nSEC-buffer. The RNA concentration was adjusted to 1 µM in all samples. L7Ae concentration ranged from 0 to 20 µM and NCG concentration ranged from 0 to 10 µM. RNAs were stained with GelRed.

In both EMSAs, the formation of an RNA-protein complex can be observed. For L7Ae, the native PAGE shows a clear distinguishable RNP-band in the gel. At a protein concentration of 5 µM no free RNA can be detected. The distinct H/ACA-L7Ae band also indicates that L7Ae binds the archaeal guide RNA only once and presumably site-specific at the kink-turn, as is expected from previous studies[57,66]. Furthermore addition of more L7Ae (up to 20 µM) does not change the migration pattern of the H/ACA-L7Ae complex, pointing to the perception that no form of unspecific aggregation between protein and RNA occurs.

For NCG a similar behavior is apparent. The bands of the free H/ACA-RNA decrease with a higher protein-to-RNA-ratio. At a one-to-one ratio, the guide RNA is almost completely bound, and at a twofold excess no free RNA remains. In contrast to the L7Ae-RNA complex, the H/ACA-NCG complex migrates only poorly or not at all into the gel. Tests with lower polyacrylamide concentrations (down to 4%) – in case the RNP might be too large for the gel – did not affect this behavior. With increasing protein concentration from 1 µM to 10 µM NCG, the observable migration of the RNA-protein particles into the gel decreases. This might indicate an aggregation of the complex under gel electrophoretic conditions, which agrees with previous findings of the archaeal NCG being prone to precipitation[119]. Another explanation might be that the surface of NCG is positively charged[52], and the complex (58.9 kDa) itself is much larger compared to L7Ae (14.5 kDa), leading to a significantly less overall negative charge at the RNA-protein complex, which is required for migration to the positively charged electrode (anode in

this case). Due to the mentioned reasons, it cannot be deduced from these EMSAs, whether NCG binds to the guide RNA in a specific or unspecific manner.

Tests using a different loading-buffer containing 0.08% SDS (6x Gel Loading Dye (NEB) instead of 60% glycerol) showed a migration of an RNA-NCG-complex into the gel (see Figure A3.1 in Appendix A3 – Supplementary Figures), suggesting that the SDS binds to the complex and its negative charge promotes RNP migration to the anode. However, since SDS is a strong detergent and capable of denaturing proteins, which may lead to unexpected behavior. Hence, these tests have to be considered with care and will not be discussed any further.

Labeled H/ACA-RNPs

For the wt-H/ACA-RNA, protein binding to L7Ae and NCG could be verified. The same was done with the fluorophore labeled guide RNAs. Therefore, the EMSAs were repeated with the synthesized FRET-H/ACA-constructs. In order to reduce the usage of labeled samples, to achieve assembly conditions as near as possible to smFRET experiments, but also to have a sufficient signal-to-noise ratio for fluorophore scanning, the guide RNA concentration was decreased to 5 nM (50 fmol labeled RNA per sample were applied to the native gel). Figure 4.17 shows the EMSAs for L7Ae and NCG with *5'*- and *Loop-FRET-H/ACA*s.

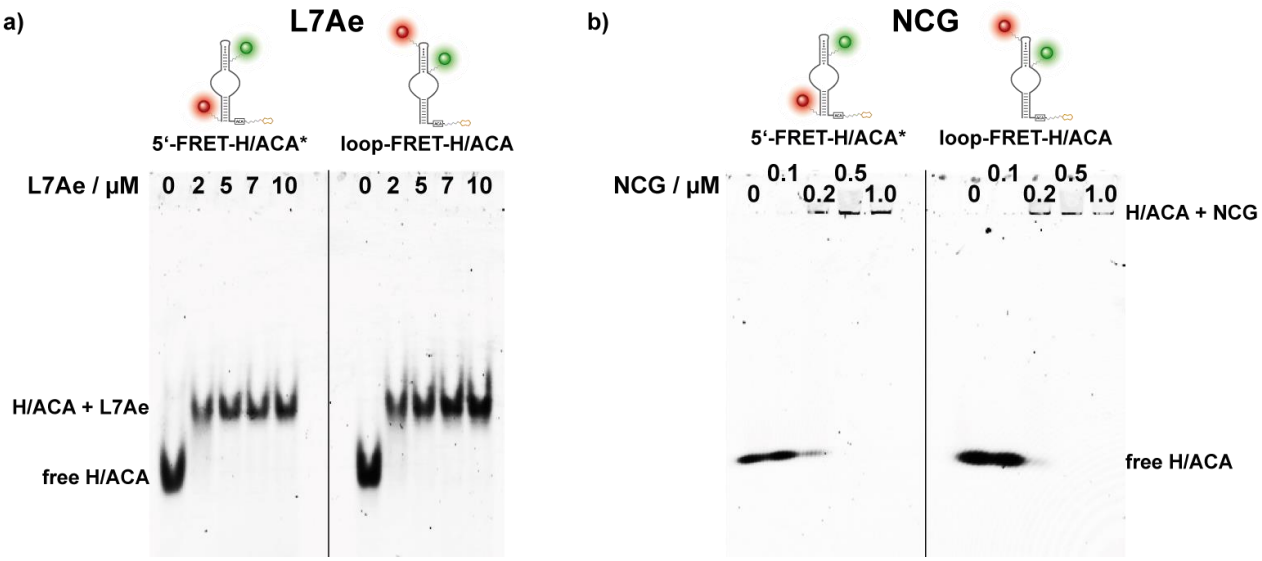


Figure 4.17: *a)* 8% PAA EMSA of L7Ae (0 – 10 µM) with 5 nM 5'-* (left) and Loop-FRET-H/ACA (right) in 0.5x oSEC-buffer (Typhoon settings: Cy5-scan). *b)* 8% PAA EMSA of NCG (0 – 1 µM) with 5 nM 5'-* (left) and Loop-FRET-H/ACA (right) in 0.8x nSEC-buffer (Typhoon settings: Cy5-scan). *This construct was labeled at position C37 instead of U39.

For the kink-turn binding protein L7Ae, both FRET-RNAs showed distinct RNP formation (see Figure 4.17a) and a clear migration of the H/ACA-L7Ae complex into the gel. At a 5 µM L7Ae

concentration, the RNAs (5 nM) were completely bound for both constructs. A similar behavior could be observed for the FRET-RNAs in case of NCG. Both double labeled constructs showed high affinities to the ternary protein complex (see Figure 4.17b). At 500 nM NCG the free RNA band vanished completely. As for the unlabeled H/ACA-RNA, no migration of FRET-H/ACA-NCG particles into the gel was observed.

As shown above RNA-protein binding for the NCG subcomplex was effective. Nevertheless, for certain experiments it is important to know whether the archaeal H/ACA-proteins bind to the guide RNA even in the absence of some proteins, like Nop10 and/or Gar1. This was also tested in EMSAs with the CG subcomplex and with Cbf5-only (see Figure 4.18).

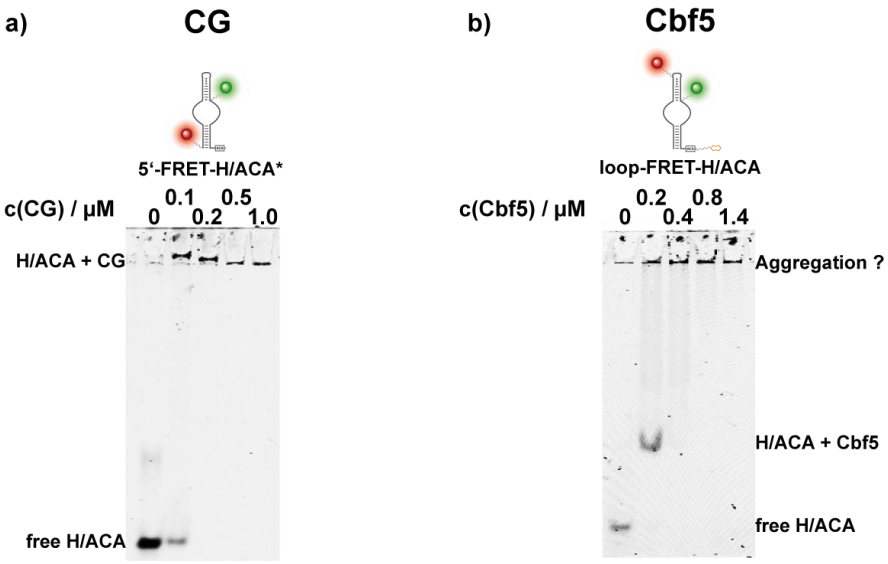


Figure 4.18: ***a)*** **8% PAA EMSA of CG (0 – 1 µM) with 5 nM 5'-FRET-H/ACA* in nSEC-buffer (Typhoon settings: Cy5-scan).** ***b)*** **8% PAA EMSA of Cbf5 (0 – 1.4 µM) with 5 nM Loop-FRET-H/ACA in nSEC-buffer (Typhoon settings: Cy5-scan. *This construct was not biotinylated.**

In Figure 4.18a a high affinity for CG to the H/ACA-RNA was observed, even in absence of Nop10. At 200 nM protein concentration, the free RNA band (5 nM) could not be detected anymore. However, as for NCG, no clear statements about the specificity of the RNP formation for CG can be made. In case of Cbf5 (see Figure 4.18b) a distinct H/ACA-Cbf5 complex was observed for the 200 nM Cbf5 sample, but with increasing protein concentrations this band disappeared and only the presumed aggregates in the gel pockets remained.

Overall, the EMSAs showed that the prepared proteins are capable of binding the unlabeled as well as the labeled H/ACA-RNAs quantitatively. The EMSAs for NCG, CG and Cbf5 provide no clear hint on a specific binding mode for the RNA-protein complex. For L7Ae however, a clear

binding mode can be distinguished in the native gels. Furthermore, no difference for the selected fluorophore sites in regard to RNA-protein binding capability could be detected, suggesting that the labels do not interfere with H/ACA-RNP assembly.

4.3.2. Fluorescence Anisotropy of the Dyes

Besides complex formation capability, the impact on the rotational freedom of the dyes has to be considered as well, when introducing organic fluorophores into biomolecules, since interactions between label and the complex of interest might reduce the fluorophores rotating flexibility. This might lead to changes in the orientation factor (κ^2) of the dyes, which directly affects the FRET-efficiency (see Figure 1.9c, page 27). κ^2 is dependent of the relative orientation of donor and acceptor dipoles and can inherit values from 0 to 4. Generally, κ^2 is assumed to be equal to 2/3 for FRET experiments, in which the rotation of donor and acceptor is fast compared to the energy transfer[125]. A strong deviation of this value would lead to an error in the calculated distances from the detected E-FRET. In the case of κ^2 values in between 1 (head-to-tail dipole orientation) and 4 (parallel dipole orientation) the deviation could be as high as 35%[125], while for $\kappa^2 = 0$ (vertical dipole orientation) severe errors for expected distances from the observed E-FRET values would occur[126]. Since the relative orientation of the dipoles cannot be determined directly, the fluorescence anisotropy '*A*' – a parameter representing the rotational freedom of fluorophores – is commonly used to check on dye flexibility in FRET experiments. *A* can be determined by exciting a fluorophore with polarized light and measuring the extent of polarization of the emitted light. Thereby, the polarization strongly depends on the rotational rate of the dye. Assuming a small fluorophore in a solution of low viscosity rotates much faster than its intrinsic fluorescence lifetime, the emitted light would be depolarized. The resulting anisotropy would be close to zero. Hence, a high *A* value (highest theoretical value is $A = 1$) implicates a strong loss of rotational freedom, which is not desired for FRET experiments. To test if the fluorophores' rotational freedom is strongly influenced by H/ACA-RNP assembly, bulk anisotropy was measured for the free dyes (Cy3/Cy5) and for each dye at each labeling site (G1, U26 and U39) in absence and presence of all four proteins (LNCG).

a)

$$A = \frac{I_p - I_v}{S} = \frac{I_p - I_v}{I_p + 2I_v}$$

A : Fluorescence Anisotropy

I_p : Intensity when polarizer and polarized excitation are parallel

I_v : Intensity when polarizer and polarized excitation are vertical

S : Summed intensity

b)

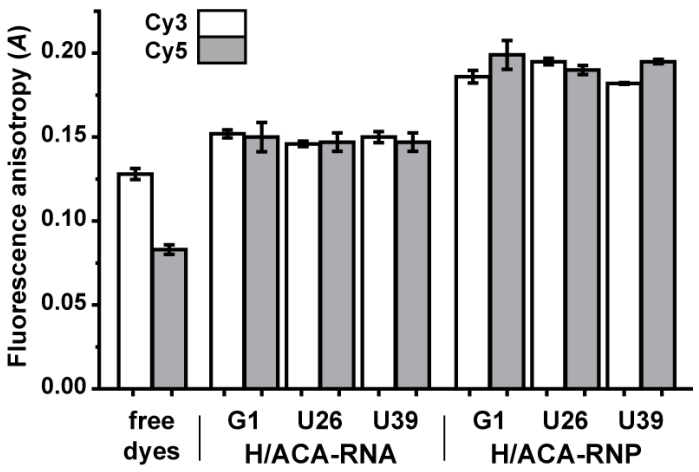


Figure 4.19: ***a)*** **Definition fluorescence anisotropy** ***A*****.** ***b)*** **Determined bulk fluorescence anisotropies for the free dyes Cy3 (white) and Cy5 (grey), labeled H/ACA-RNAs and labeled H/ACA-RNPs. Concentrations: RNA/dye 5 nM, NCG 500 nM and L7Ae 5000 nM.**

Figure 4.19 illustrates the results of the bulk anisotropy measurements. As expected the free dyes experience the least steric hindrance and show the lowest *A* values. Attachment to the guide RNA shows an increase in anisotropy to around 0.15 for all fluorophores and labeling sites. Addition of all four proteins (LNCG) further decreases the rotational freedom of the labels for all samples. For the assembled H/ACA-RNP the detected anisotropy *A* lies in between 0.18 and 0.20, which is considered a reasonable value for smFRET experiments[97].

4.4. Enzymatic Activity of Archaeal H/ACA-RNPs

4.4.1. Activity Tests of Unlabeled H/ACA-RNPs

The major consideration when attaching fluorophores to enzymes is; if they interfere with the biological function of the macromolecule. To test the catalytic functionality of the designed and synthesized guide RNAs, several pseudouridylation assays were performed. Therefore, radioactive ^{32}P-labeled substrate RNA was incubated with full or partial H/ACA-RNPs in Ψ-buffer. After the reaction was finished, proteins were removed by phenol extraction and the remaining RNAs were digested to nucleotide monophosphates with P1-endonuclease. The resulting nucleotides were separated by TLC and visualized by autoradiography. Hence, only ^{32}P-labeled nucleotides can be detected.

For initial activity tests, radioactive substrate labeling was accomplished by *in vitro* transcription of the RNA with α-^{32}P-UTP, as was established previously[119]. The substrate RNA was designed to form Watson-Crick base pairing at each possible site in the pseudouridylation pocket and an overhang of five nt at its 5'- and four nt at its 3'-end (adapted from Liang *et al.*)[55], resulting in six Us total per substrate RNA. Since standard T7 transcription was performed, the incorporation of radioactive uridine into the substrate RNA occurred stochastically at each of the six potential positions, leading to an excess of unisomerized radioactive U for each assay. At first, unlabeled constructs were tested in single turnover assays; the full complex (LNCG) at 70 °C and 37 °C, and the partial complexes with omitted L7ae (NCG), Gar1 (LNC) or Nop10 (LCG) at 70 °C. Additionally, a negative control without the sRNA (only NCG) was carried out, to check whether the proteins exhibit basal pseudouridylation activity. Incubation at 70 °C was done for 1 h and incubation at 37 °C for 14 hours. For evaluation reasons the yield of the full H/ACA-RNP was normalized to 100% and the other samples were compared accordingly (see Figure 4.20).

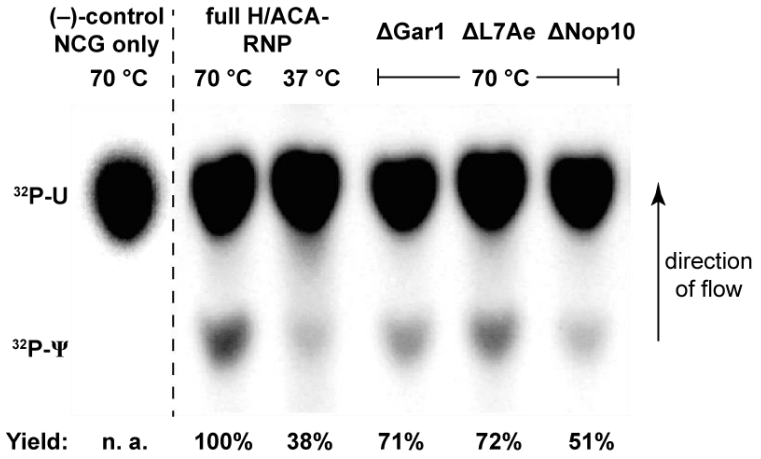


Figure 4.20: TLC of single turnover activity assay with full and partial unlabeled H/ACA-RNPs (Typhoon settings: Phosphor screen). Lanes for samples with omitted proteins are marked with Δ and the respective protein. The negative control (NCG only) showed no significant activity (n. a.).

Removals of individual proteins lead to a significant decrease in uridine turnover. The influence on Ψase activity for Gar1 or L7Ae omission can be considered to be similar in this single turnover experiment. Depletion of the smallest protein Nop10 showed the biggest impact on RNA guided pseudouridylation, with a near to 50% loss of yield. Furthermore, it could be shown that the prepared NCG subcomplex has no basal activity in the absence of the sRNA, which is in agreement with existing literature[55]. Reduction of the reaction temperature to 37 °C resulted in a major decrease of the single turnover yield, down to 38% even though the incubation was extended from one hour to 14 hours. Further decrease in temperature down to 20 °C for 20 hours yielded a complete loss of enzymatic activity for the reconstituted thermophile H/ACA-RNP from *P. furiosus*. Since smFRET measurements are typically performed at room temperature, a series of experiments was performed to increase Ψase activity at 20 °C. Therefore, it was attempted to mimic the required high temperatures (> 70 °C) by addition of denaturing agents (urea or guanidine hydrochloride), which are known to destabilize interactions of biomolecules. These attempts did not prove successful and were not pursued any further. The results of these tests are shown in Figure A3.2 in Appendix A3 – Supplementary Figures.

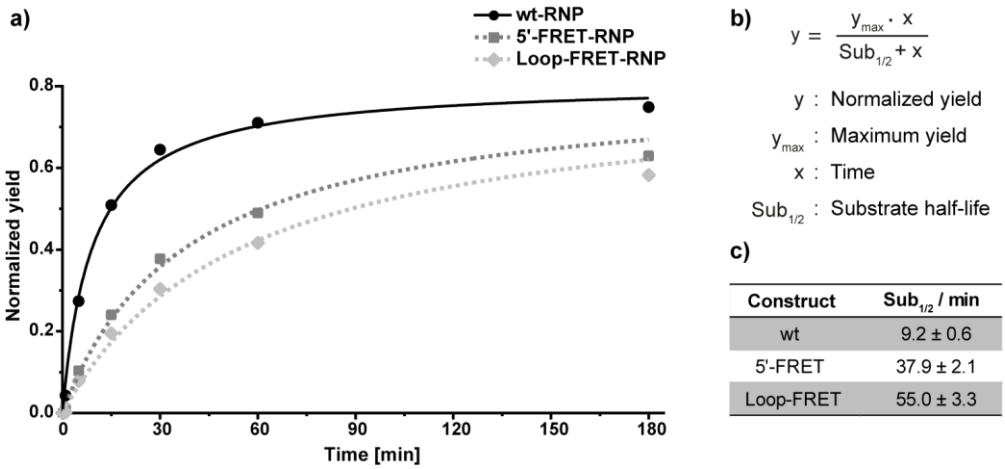
As mentioned above, in the beginning of this thesis radioactive labeling of the substrate RNA was achieved by stochastic incorporation of α-^{32}P-UTP via *in vitro* transcription. Due to the sequence of the substrate, this leads to an inherent excess of detectable ^{32}P-U compared to ^{32}P-Ψ, even when target uridine is quantitatively isomerized. Furthermore, T7 RNA polymerase is prone to randomly add nucleotides at the 3'-end of the transcription product[127], and in case of random uridine insertion, the evaluation of these pseudouridylation assays would intrinsically be falsified.

To overcome these issues a different labeling method for radioactive substrate RNAs was established. The substrate RNA was divided into two oligomers for splinted ligation. The target U was located at the 5'-end of the 3'-oligonucleotide for radioactive labeling with γ-^{32}P-ATP by phosphorylation. Afterwards, splinted ligation of the 5'-RNA strand to the radioactive 3'-oligomer was performed, and the radioactive target RNA was purified by PAGE. This way selective ^{32}P-labeling and exclusive visualization of the target nucleotide could be ensured.

Multiple turnover activity assays of the wt-H/ACA-RNPs with selectively ^{32}P-labeled target RNA were performed by Andreas Schmidt (Goethe University)[128]. The complete archaeal RNP showed high catalytic activity (62% Ψ-formation), while removal of individual proteins like Gar1 or L7Ae drastically reduced pseudouridylation yield, down to 8% or 6% respectively. Hence, all archaeal RNP components are required for efficient RNA guided uridine isomerization.

4.4.2. Activity Tests of Labeled H/ACA-RNPs

Effects on pseudouridylation activity of the H/ACA-RNP upon dye insertion into sRNAs were tested with the double labeled *5'*- and *Loop-FRET-H/ACA* constructs. Both FRET-RNPs were compared to the wt-RNP in a multiple turnover assay with a 20-fold excess of substrate RNA (1 µM). The RNP concentration was referred to the guide RNA concentration (50 nM), since the stand-alone proteins inherited no catalytic activity. Proteins were added in excess according to the EMSA conditions (500 nM NCG and 5 µM L7Ae) in order to ensure complete RNP formation. To obtain high amounts of substrate concentration, non-radioactive target RNA was added to the reaction mixture. Incubation was performed at 70 °C and in order to achieve kinetic information of the reaction an aliquot of each construct was taken after 1, 5, 15, 30, 60, and 180 minutes (see Figure 4.21a). These samples were treated and prepared for visualization by autoradiography according to the previous assays.



Construct	$Sub_{1/2}$ / min
wt	9.2 ± 0.6
5'-FRET	37.9 ± 2.1
Loop-FRET	55.0 ± 3.3

Figure 4.21: ***a)*** **Time dependent pseudouridylation yields of wt- and FRET-H/ACA-RNPs for multiple turnover activity assays.** ***b)*** **Michealis-Menten based equation used for fitting.** ***c)*** **Table of reaction kinetics. The errors were obtained from fitting results.**

To gain kinetic information, the data points for each construct were fitted to a Michaelis-Menten like equation (see Figure 4.21b). For each fit y_{max} was set to 0.81, the detected maximum yield of a previously performed single turnover assay for the unlabeled H/ACA-RNP. The time required by the enzyme to pseudouridylate half of the remaining substrate RNA ($Sub_{1/2}$) is presented for the tested constructs in Figure 4.21c. As expected, the unlabeled wt-RNP showed the highest catalytic activity with $Sub_{1/2}$ = 9.2 min. For the *5'-FRET-H/ACA* construct the reaction speed decreased by a factor of four to $Sub_{1/2}$ = 37.9 min, and for the *Loop-FRET-H/ACA* RNA the reactivity was reduced even further down by a factor of six ($Sub_{1/2}$ = 55.0 min) compared to the unmodified RNP. Thus, attachment of the dyes to the sRNA at the selected labeling sites leads to noticeable loss of RNA guided Ψase activity. Nevertheless, both FRET-constructs are still capable of performing the uridine isomerization effectively and in a multiple turnover manner. Hence, the designed and synthesized fluorophore labeled H/ACA-RNAs are deemed to be well suiting for smFRET experiments, and the states observed are considered to be of biological relevance.

4.5. smFRET Analysis of the Archaeal H/ACA-RNP

4.5.1. General Considerations

In the previous chapters it has been shown that the designed and fluorophore labeled H/ACA-RNAs are capable of RNP assembly as well as of efficient pseudouridylation, and that the dyes' rotational freedom is not significantly confined by RNA-protein complex formation. These are important requirements for FRET studies. The results of the smFRET experiments of the archaeal H/ACA-RNP from *P. furiosus* will be presented in the next subchapters.

Prior to the spectroscopic measurement individual samples were prepared and annealed in a small scale (4 µL) at 70 °C for one to 15 minutes depending on the sample and experiment. Afterwards, the assembled complexes were immediately put on ice to prevent any further biochemical processes, and cooled until the samples were used for the actual measurements. Complex reconstitution was performed at the same concentrations that were used for the pseudouridylation activity assays; 50 nM H/ACA-RNA, 500 nM NCG and 5 µM L7Ae. In case of experiments containing Cy5- or biotin-labeled target RNAs, 1.4 equivalents (70 nM) of labeled X-Sub were added to the sample for annealing. Detection of fluorophore labeled biomolecules was done by TIRF microscopy, which required complex immobilization. Therefore, glass slides were passivated with a PEG/PEG-biotin mixture to prevent undesired glass-protein interactions. Specific pull down of the biotinylated H/ACA- or substrate RNA was achieved by the highly stable biotin-streptavidin interaction ($K_d \approx 10^{-14}$ M)[129]. Therefore, streptavidin was added to the passivated glass surface to enable complex immobilization. Then, the annealed sample was diluted to a final H/ACA-RNA concentration of 100 pM, to avoid overloaded glass slides and to allow for a distinct identification of individual molecules. After the sample was loaded onto the glass slide, particle density was checked and adjusted if necessary. When sufficient molecules could be detected the channel was rinsed with Imaging-buffer (IB), containing an oxygen-scavenging system (glucose oxidase, catalase and glucose) as well as a triplet state quencher (a Vitamin E analogue 'Trolox'), to generate long-lasting and nonblinking fluorophores[130]. Finally, the actual smFRET measurements were performed at room temperature – where activity of the thermophile Ψase is drastically reduced (see Chapter 4.4.1) – to generate histograms (distributional plots of the sum of all detected molecules per sample) or traces (videos following dye intensities of individual molecules) of each sample. In some histograms molecules

experiencing E-FRET < 0.2 were removed from the plot for better visualization. Regarding the well suited anisotropy values of the chosen labeling sites for smFRET experiments and according to published structural data, molecules with such a low FRET-efficiency most likely belong to one of the following options; donor-only constructs in which the FRET-acceptor bleached, impurities on the slide, or a minor fraction of improperly folded H/ACA-RNAs, which presumably have no biological relevance.

4.5.2. H/ACA-RNP Assembly

To gain new insights into H/ACA-RNP assembly, the two synthesized FRET-constructs (*5'-FRET-H/ACA* and *Loop-FRET-H/ACA*), were annealed and investigated in absence and/or presence of the four proteins; L7Ae, Nop10, Cbf5 and Gar1. Since the FRET-acceptor (Cy5) is attached at an exterior spot – at either the 5'- (G1) or at the loop (U26) of the H/ACA-RNA – and the FRET-donor (Cy3) is located in a central position – at U39, next to the pseudouridylation pocket – of the guide RNA (see Figure 4.9, page 82), an overall overview of structural changes within the sRNA should be obtainable.

5'-FRET-H/ACA

For detection of conformational rearrangements in the H/ACA-RNA between its lower stem and the substrate binding pocket *5'-FRET-H/ACA* was used. smFRET was used to generate histograms and traces of the free RNA, partial (L, LNC, LCG, NCG) or full (LNCG) H/ACA-RNPs in absence of the target RNA. Histograms of these experiments are shown in Figure 4.22.

For *5'-FRET-H/ACA* two populations were observed; an open state at E-FRET = 0.45 and a closed state at E-FRET = 0.59. For the free RNA (sRNA), as well as in presence of L7Ae (sRNA + L), the closed conformation was predominantly populated (> 80%). Addition of the ternary NCG complex to the guide RNA (sRNA + NCG) led to a minor increase of the lower E-FRET population of up to 30%. Assembly of partial RNPs lacking Gar1 or Nop10 showed a nearly even distribution between both states, with a slight dominance for the closed state. Only when all four H/ACA-proteins were added to the guide RNA the open conformation (E-FRET = 0.45) became the major population (> 80%). The decrease in FRET-efficiency signals an increase in the intramolecular dye distance, due to structural changes of the sRNA.

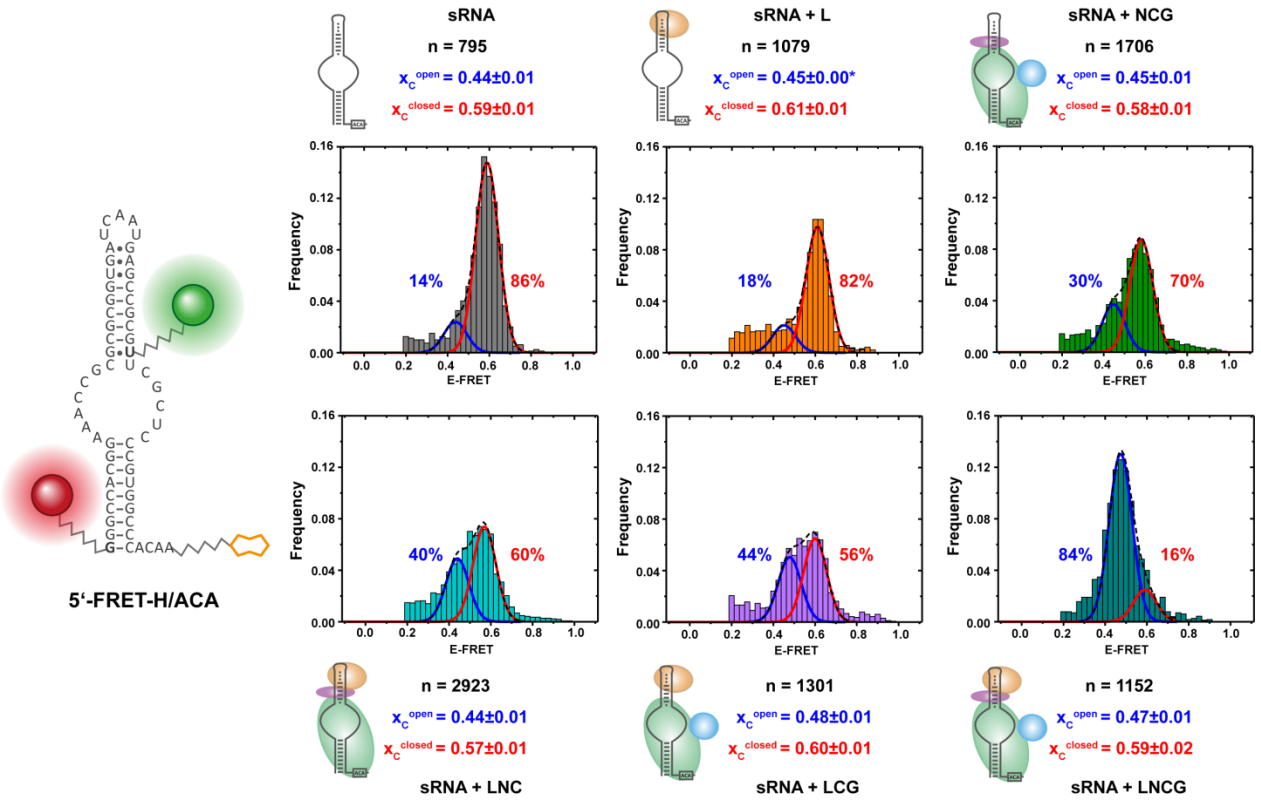


Figure 4.22: smFRET histograms of 5'-FRET-H/ACA assembly. Histograms of individual complexes are marked by color: free sRNA (grey), sRNA plus L7Ae (orange), sRNA plus Nop10-Cbf5-Gar1 (green), sRNA plus L7Ae-Nop10-Cbf5 (cyan), sRNA plus L7Ae-Cbf5-Gar1 (violet) and sRNA plus L7Ae-Nop10-Cbf5-Gar1 (teal). The absolute number of detected molecules is 'n', 'x_C^{open}' (blue) represents the peak center of the open state and 'x_C^{closed}' (red) represents the peak center of the closed state. For proper fitting, the peak width of each Gaussian fit was set to 0.11. Errors result from fitting. *For sRNA + L x_C^{open} was set to 0.45.

Evaluation of single molecule traces showed that once the sRNA adopted one of the two conformations, it remained in this state, since no relevant transitions between the two FRET-states could be observed for the *5'-FRET-H/ACA* construct (see Figure 4.23). In the first part of the movies no significant change of donor and acceptor fluorescence can be detected. Hence, the calculated FRET-efficiency remains constant as well. In this exemplary trace of an RNA in the closed state the acceptor dye bleaches after 44 seconds and correspondingly the donor intensity increases (Figure 4.23a). For the same molecule donor bleaching follows after 117 seconds. This single step photo bleaching indicates that only one single molecule carrying two fluorophores and no aggregates or other form of multimers were observed. For the exemplary trace of an opened RNA acceptor bleaching was detected after 61 seconds (Figure 4.23b). Donor bleaching did not follow for this molecule within the 120 second movie.

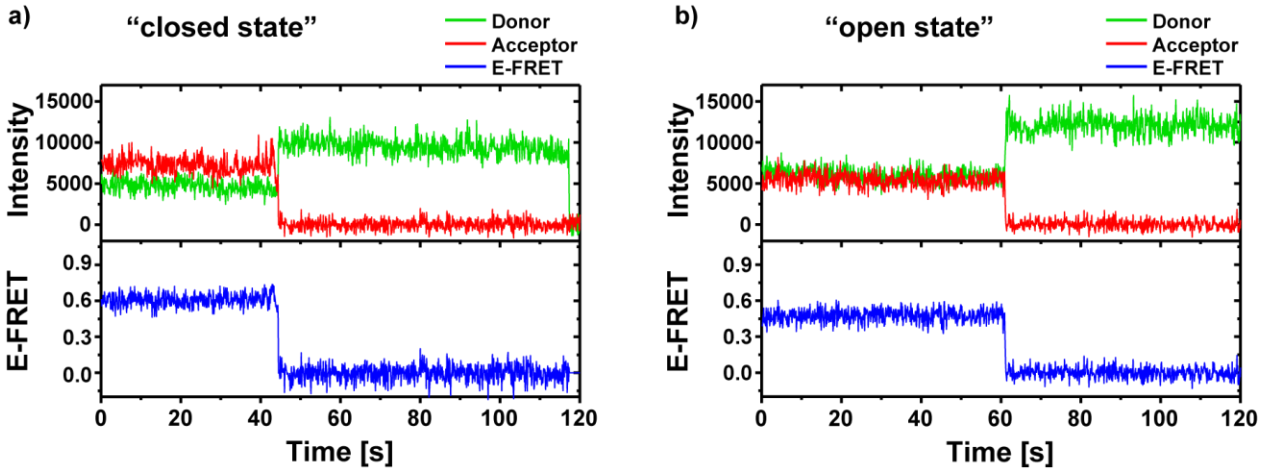


Figure 4.23: ***a)*** **Exemplary trace of the closed state (E-FRET = 0.59) from sRNA + L.** ***b)*** **Exemplary trace of the open state (E-FRET = 0.45) from sRNA + LNCG.**

Loop-FRET-H/ACA

The same experiments as for *5'-FRET-H/ACA* were repeated for the other construct *Loop-FRET-H/ACA*, labeled at the loop (U26) and next to the pseudouridylation pocket (U39). This construct enables detection of structural changes in the upper stem of the guide RNA. Histograms of these experiments are shown in Figure 4.24.

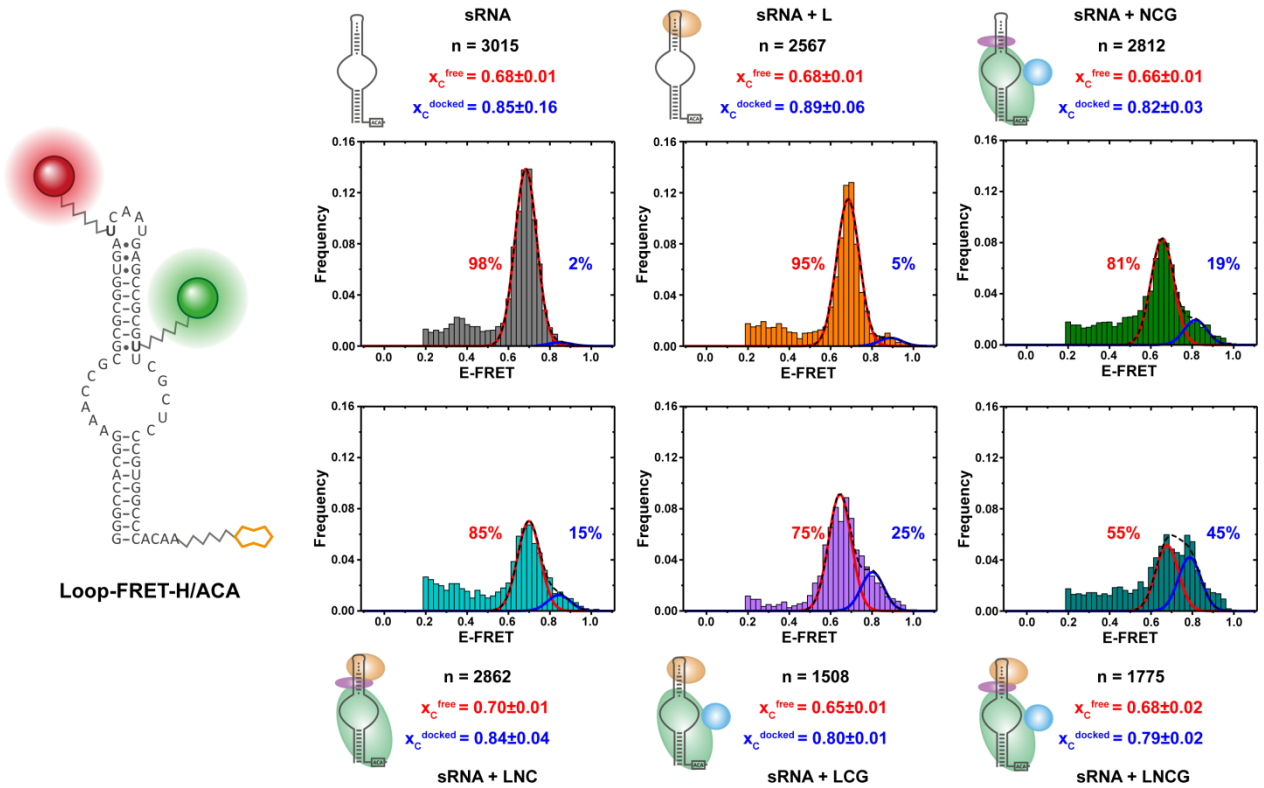


Figure 4.24: smFRET histograms of loop-FRET-H/ACA assembly. Histograms of individual complexes are marked by color: free sRNA (grey), sRNA plus L7Ae (orange), sRNA plus Nop10-Cbf5-Gar1 (green), sRNA plus L7Ae-Nop10-Cbf5 (cyan), sRNA plus L7Ae-Cbf5-Gar1 (violet) and sRNA plus L7Ae-Nop10-Cbf5-Gar1 (teal). The absolute number of

detected molecules is n. x_c^{free} (red) represents the peak center of the free conformation and x_c^{docked} (blue) represents the peak center of the docked conformation. For proper fitting, the peak width of each Gaussian fit was set to 0.11. Errors result from fitting.

For *Loop-FRET-H/ACA* again two populations could be observed; one state at E-FRET = 0.68, termed 'free state', and a high FRET-state at E-FRET = 0.80, termed 'compact state'. In case of the free H/ACA-RNA (sRNA) and in presence of L7Ae (sRNA + L), the free conformation was almost quantitatively populated. Addition of NCG or LNC slightly increased the relative number of detected molecules for the compacted state, to 19% and 15% respectively. For the partial RNP missing Nop10 (sRNA + LCG), 25% of the observed molecules showed the high FRET state. The highest population of the compacted state (45%) could be detected in presence of all four H/ACA proteins (sRNA + LNCG). However, a quantitative compaction of the guide RNA in the upper stem could not be observed for any of the annealed samples.

Again, intramolecular transitions between both states could not be observed for the assembled H/ACA-RNP. Representative traces for the free state (see Figure 4.25a) and for the docked state (see Figure 4.25b) show that the *Loop-FRET-H/ACA* remains in one conformation after complex assembly.

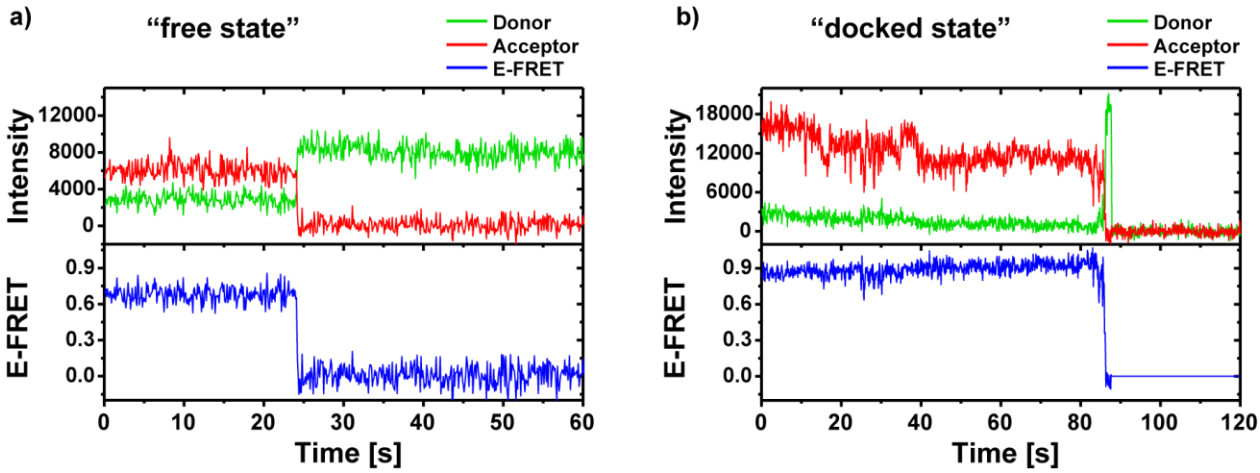


Figure 4.25: *a)* Exemplary trace of the free state (E-FRET = 0.67) from sRNA + LNCG. *b)* Exemplary trace of the docked state (E-FRET = 0.8) from sRNA + LNCG.

4.5.3. Substrate Recruitment

In the above mentioned assembly experiments conformational rearrangements of the sRNA were observed upon H/ACA-protein binding. To investigate the purpose of these structural changes substrate recruitment tests were performed. Therefore, the labeling scheme was altered. The guide RNA was only labeled with the FRET-donor (Cy3) at the 5'-end (*5'-Cy3-H/ACA*) and the FRET-acceptor (Cy5) was attached to the 3'-end of the substrate Sub-Cy5. To ensure the

target RNA remains bound to the RNP after complex annealing a substrate analogue was used, where the target base is substituted by 5-fluoro-uracil (5fU). This modification is known to potently inhibit pseudouridylation of several Ψases[131-134]. In order to observe substrate bound H/ACA-RNP complexes, the labeled RNAs were annealed for 15 minutes at 70 °C in presence of all four proteins or with one of the proteins missing (see Figure 4.26).

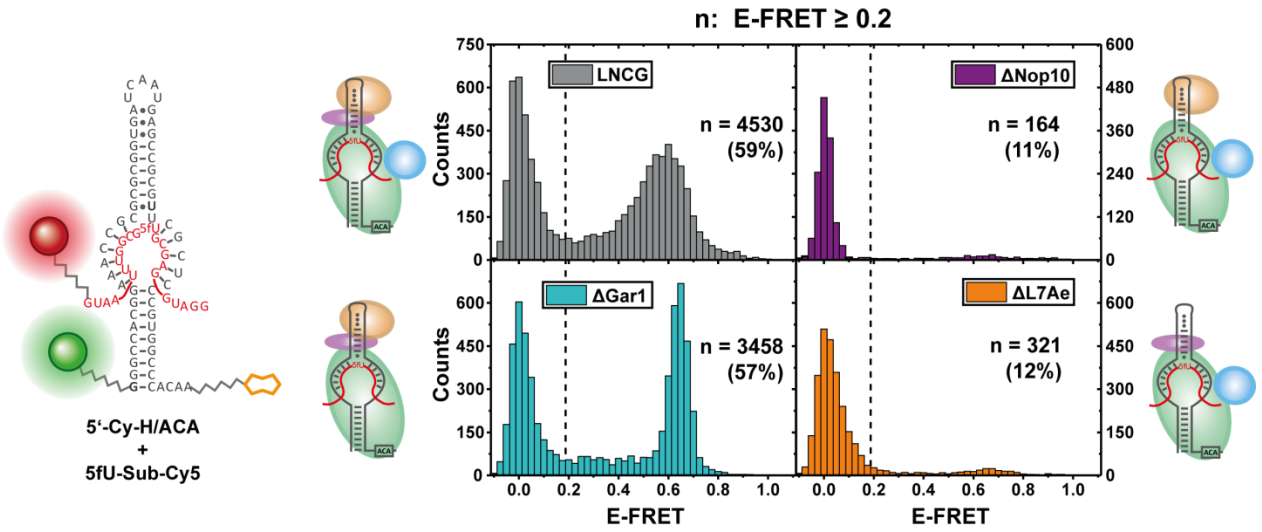


Figure 4.26: smFRET histograms of substrate recruitment tests of 5'-Cy3-H/ACA plus 5fU-Sub-Cy5. The individual H/ACA-RNPs are marked in different colors: full RNP (grey), lack of Nop10 (purple), lack of Gar1 (cyan) and lack of L7Ae (orange). n represents the number of molecules detected for E-FRET ≥ 0.2 (the ratio of n compared to all observed for the respective sample molecules is given in %).

Histograms of all substrate recruitment tests show a distinct state for E-FRET = 0. This population is corresponding to molecules that lack the FRET-acceptor. Reasons for that could either be bleached Cy5, unbound substrate RNA or fluorescent impurities on the glass slide. Thus, this population only serves as reference for states showing FRET and was therefore included in the histogram. For the full H/ACA-RNP (LNCG) a broad population of target-RNA bound complexes (at E-FRET ~ 0.6) was detected. The RNP lacking Gar1 (ΔGar1) also showed a substrate bound fraction at similar FRET-efficiency, but the observed population was much sharper in full width at half maximum (w = 0.08) compared to the full complex (w = 0.23), which is an indication for a more dynamic system in case of the full RNP. The ratio of substrate bound to unbound RNPs is almost identical for the full complex (59%) compared to the ΔGar1 particle (57%). On the other side, the other two samples (ΔNop10 and ΔL7Ae) showed significantly less substrate RNA bound H/ACA-RNPs, 11% of all observed molecules for ΔNop10 and 12% for ΔL7Ae. Thus, the simultaneous presence of L7Ae and Nop10 in the RNP appears to be essential for effective substrate binding to the sRNA.

4.5.4. Characterization of Reactive States by Atomic Mutagenesis

In the previous chapter, the impact of Gar1 on substrate bound H/ACA-RNPs could be observed. Absence of this protein significantly reduced the peak width of the obtained FRET state (see Figure 4.26). The experiment was repeated for the Cy5-labeled target RNA with altered nucleobases at the isomerization site (U, 5fU, Ψ, C, or rab). Atomic mutagenesis on the target nucleotide was used to differentiate distinct states during the isomerization process. U and 5fU were used for a reactive state, Ψ and C for a release state and rab (abasic site) to mimic an uncatalytic state. The labeled RNAs were reconstituted in presence of L7Ae and either the NCG or NC subcomplex. Annealing time was decreased to five minutes for 5fU-, Ψ-, C- and rab-Sub-Cy5, and down to one minute for U-Sub-Cy5, in order to reduce the amount of pseudouridylated target RNAs. For comparison of the peak center (x_c) and peak width (w), the obtained FRET states in each histogram were described by one Gaussian fit. Figure 4.27 shows the summary of these experiments.

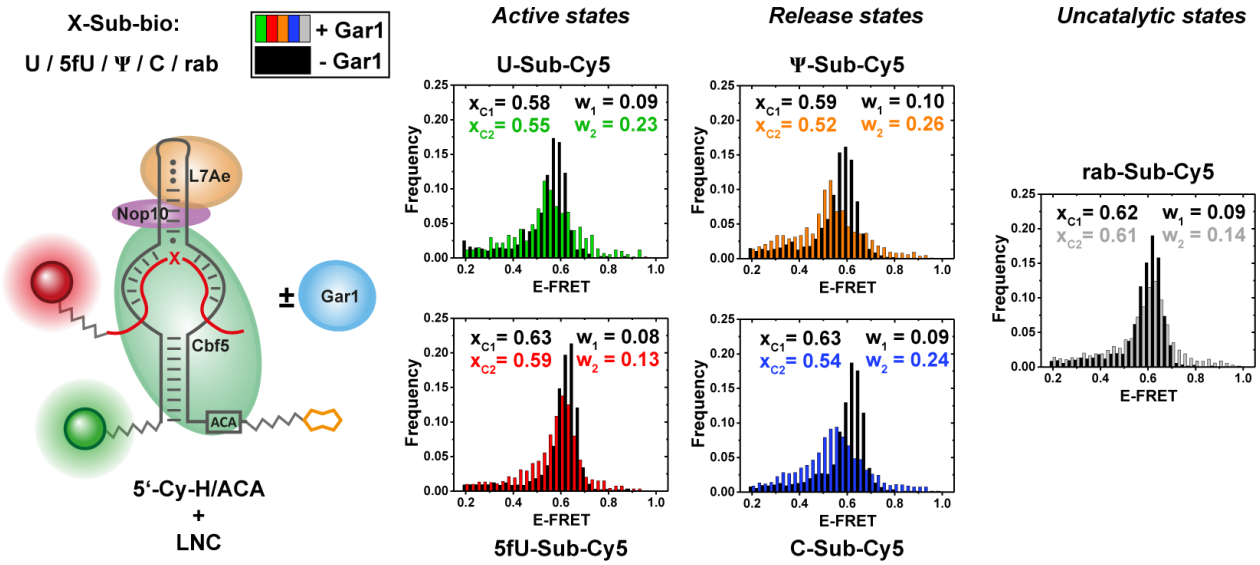


Figure 4.27: smFRET histograms of X-Sub-Cy5 and 5'-Cy3-H/ACA with LNC (black) or LNCG (red). Peak center (x_c) and full width at half maximum (w) were obtained by fitting each histogram with one Gaussian distribution.

For the constructs missing Gar1 (black), an overall similarity for the detected FRET states could be observed. The peak center of all target nucleotides showed only minor deviations, varying from 0.58 to 0.63 E-FRET. This also holds true for the peak width, here, only sharp FRET distributions were found, with values of w in range from 0.08 to 0.10. These distinct populations indicate a rigid binding of the target RNA in absence of Gar1. In contrast, RNPs containing Gar1 showed a more variable behavior of the FRET states. The substrate analogues that were used to

represent an active state (U- and 5fU-Sub-Cy5) of the RNA guided Ψase exhibited an increase in peak width (w = 0.23 for U and w = 0.13 for 5fU), and at least for U-Sub-Cy5 a minor shift to lower FRET-values (E-FRET = 0.55) occurred. For Ψ and C (release state) this trend was even more pronounced: Here, a drop of nearly 0.1 E-FRET for x_c could be observed for the two RNAs, and the width more than doubled in both cases. This major expansion in w indicates an increase in intrinsic dynamics of the system. Still, evaluation of the traces yielded no distinct transitions. In contrast to active and release states, the abasic substrate analogue rab-Sub-Cy5 – representing the uncatalytic state – showed no significant changes for x_c in absence or presence of Gar1, and only a minor increase in full width at half maximum for the full RNP.

After examination of reactive states at the site of the substrate RNA, conformational changes of the H/ACA-RNA during pseudouridylation were to be investigated in a similar manner. Therefore, the immobilization mechanism was altered. The double labeled guide RNA construct lacking biotin '*FRET-H/ACA-unbio*' was annealed separately with each of the five biotinylated substrate RNAs (X-Sub-bio) in presence of all four proteins (LNCG). With this immobilization scheme, detection of substrate bound H/ACA-RNPs could be assured. Sample annealing was carried out likewise to the previously described experiments.

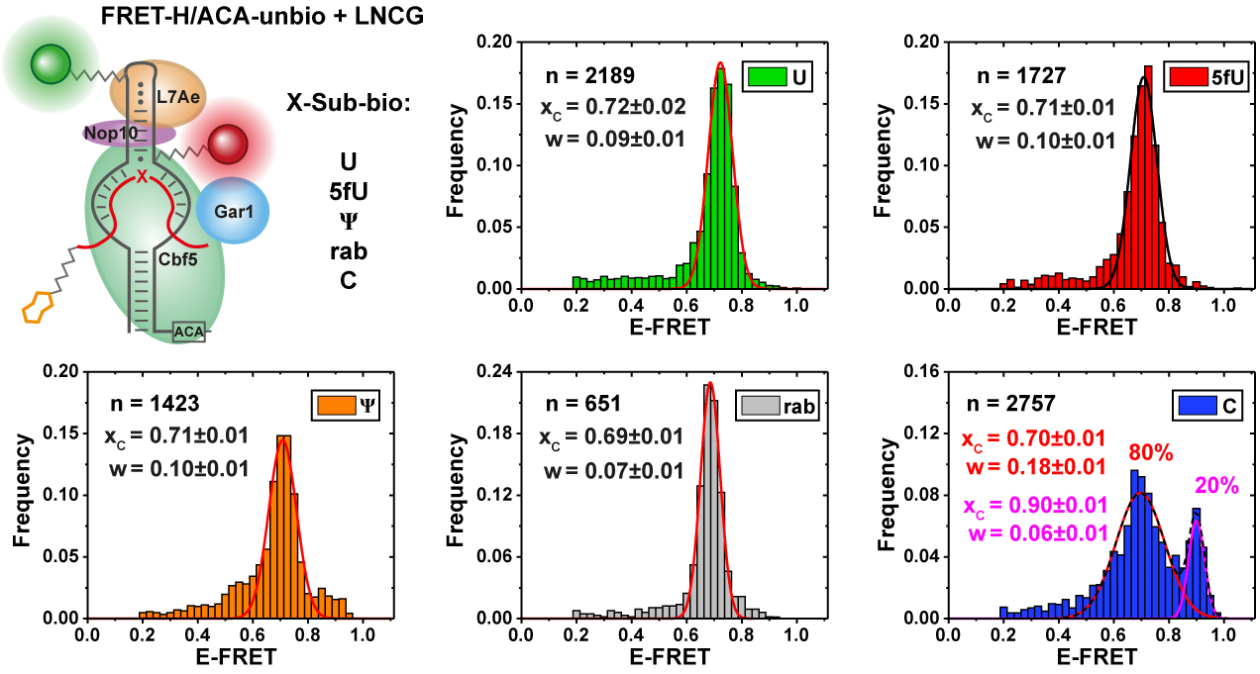


Figure 4.28: smFRET histograms of FRET-H/ACA-unbio plus X-Sub-bio in presence of LNCG. Different species of X-Sub-bio are color coded: U (green), 5fU (red), Ψ (orange), rab (grey) and C (blue). The absolute number of observed molecules with E-FRET ≥ 0.2 is represented by 'n', the peak centers by 'xc' and peak widths by 'w'. For C-Sub-bio x_c and w are color coded corresponding to the fits. Errors were obtained by Gaussian fitting.

Figure 4.28 shows distinct FRET-populations at around E-FRET = 0.71 for each of the five substrate analogues, probably resembling the docked conformation from the assembly experiments. Only for the mismatched target base C-Sub-bio (blue) an additional state at E-FRET = 0.90 (termed ‘compact state') was observed. U-, 5fU, Ψ- and rab-Sub-bio all show a similar FRET-population in peak width as well as peak center. Hence, no conformational changes in the upper stem of the sRNA could be detected for these target RNAs. Furthermore, the sharp peak width (w = 0.09) indicates a rigid structure of the RNP between loop and pseudouridylation pocket. Only for C-Sub-bio – which is supposed to mimic a release state, since the Ψase has to remove the mismatched RNA in order to perform pseudouridylation – a significant peak broadening (w = 0.18) is observable for the population at E-FRET = 0.71. Representative traces for C-Sub-bio are shown in Figure 4.29.

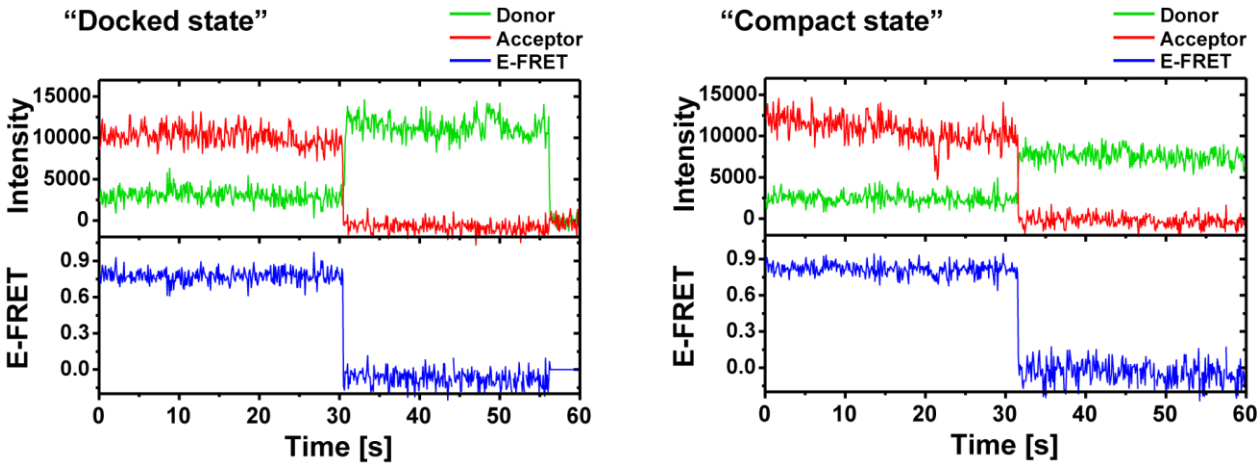


Figure 4.29: Exemplary fluorescence trajectories of FRET-H/ACA-unbio immobilized via C-Sub-bio.

Here again, no transitions between the docked and compact states could be observed for the H/ACA-RNA at room temperature. Once a conformation was formed after cool down the sRNP remained in its conformation. Even though, the appearance of a new high-FRET state for C-Sub-bio indicates a certain function of the upper stem during pseudouridylation.

4.6. Discussion

4.6.1. Reconstitution of H/ACA-RNPs for smFRET

Protein and RNA Synthesis and Purification

In order to investigate structure, assembly and dynamics of the archaeal H/ACA-RNP by smFRET, fluorophore labeled and enzymatically active RNA guided Ψases had to be reconstituted. Based on previous accomplishments[119], preparation of the ternary protein complex NCG had to be improved for two main reasons: first, the smallest protein Nop10 showed low abundance compared to Cbf5 and Gar1 in analytical gels after complex purification (see Figure 4.2b, page 75), and second, precipitation of NCG occurred rapidly in the former buffer conditions. Insufficient Nop10 expression – which might lead to complexes lacking this protein – could be overcome by usage of *E. coli* codon optimized synthetic genes. After vector cloning and plasmid transformation into *E. coli* cells competent for protein expression, Nop10 abundance increased significantly in analytical gels, and the distinct peak in the SEC chromatogram of NCG indicated a homogenous complex formation (see Figure 4.2, page 76). The problem of protein precipitation could be solved by usage of a high salt buffer according to Yang *et al.*[64] for the size exclusion chromatography. Apparently, the high salt concentration (1 M NaCl) stabilizes the surface of the thermophilic proteins and thereby prevents complex aggregation without interfering with NCG stability. Furthermore, Gar1 degradation – which was previously reported[119] – could not be observed over a period of 2 weeks of storage at 4 °C in the new buffer. Overall, the optimization of NCG preparation as well as for other partial protein complexes (NC or CG) was successful and showed no major contaminations, such as unspecifically bound RNAs.

Besides protein optimization, smFRET suitable RNA constructs had to be designed and synthesized. The nucleotide G1 (5'-end), U26 (loop) and U39 (pseudouridylation pocket) were chosen for site specific fluorophore coupling to the sRNA (see Figure 4.9, page 82). With these labeling sites at hand, overall structural information of the guide RNA during complex assembly and Ψ-formation can be obtained. Fluorophore labeled H/ACA-RNAs were synthesized by dye coupling of different oligomers, HPLC purification of the labeling reaction and subsequent splinted ligation of these oligomers to the full-length guide RNA. All these steps were successful and yielded different constructs for distinct purposes (see Figure 4.13, page 86). The double labeled and biotinylated RNAs *5'-FRET-H/ACA* and *Loop-FRET-H/ACA* were designed for

assembly experiments and to observe changes in the structure of the guide RNA at the lower and upper stem, respectively. The single dye labeled *5'-Cy3-H/ACA* and the unbiotinylated *FRET-H/ACA-unbio* were used for substrate binding experiments. Substrate analogues could be modified at the 3'-end with a fluorophore or biotin linker, depending on the desired experiment. To study substrate bound RNPs, biotin was introduced at the target RNA, while for analyzing the effects on the substrate RNA, the FRET acceptor was attached to the 3'-end. Dye or biotin labeling and purification of these constructs were carried out in the same way as for the H/ACA-oligonucleotides, and since effective biotin coupling could not be verified by UV absorption in the HPLC chromatogram, mass spectrometry was performed for the biotinylated oligomers (see Table 4.2, page 81). Labeling as well as purification was successful for all constructs, yielding suitable target RNAs for different smFRET experiments.

RNP Formation

Introduction of fluorophores and other modifications like biotin is essential for TIRF based smFRET spectroscopy. In order to see if these labels interfere with H/ACA-RNP function several assays were performed: 1) binding capability of the modified H/ACA-RNAs was checked with EMSAs, 2) interactions between fluorophores and the RNA-protein complex were tested by fluorescence anisotropy measurements, and 3) complex activity of labeled RNPs was compared to wildtype RNPs.

Regarding the first point, binding of the prepared H/ACA-proteins to modified or wildtype sRNAs could be detected in several EMSAs (see Chapter 4.3.1). However, only in the case of L7Ae a distinct RNA-protein complex migrated into the gel, indicating specific binding of L7Ae to the kink-loop of the H/ACA-RNA. In case of the other proteins (NCG, CG and Cbf5) RNA-protein formation was observed, but the formed RNPs remained in the loading pockets of the native gel. Hence, a definite conclusion about binding specificity of these complexes cannot be made from these EMSAs. Complete binding of the labeled H/ACA-RNA (5 nM) was detected at 5 µM L7Ae or 500 nM NCG sample concentration. Thus, these protein concentrations were used for complex assembly in further experiments.

Secondly, to check whether the RNP formation competes with rotational freedom of the dyes, bulk fluorescence anisotropy of each labeling site was determined in absence and presence of LNCG and compared to the free fluorophores (see Figure 4.19, page 94). The impact of dye

orientation on E-FRET is represented by κ^2 and was already discussed in Chapter 4.3.2. Overall, an increase in fluorescence anisotropy – thus, a loss of rotational freedom – of the dyes can be observed if the label is attached to the RNA or RNP, as could be expected. The average increase in *A* for free Cy3 and Cy5 when bound to the sRNA were 0.02 and 0.06, respectively. When bound to the RNP an average increase of *A* by 0.06 or 0.11 for Cy3 or Cy5 was observed. The higher *A* values can mainly be contributed to two effects: one being the covalent attachment of the dyes linker to the comparably large and rigid RNA, which implicates a decrease in rotational speed of the whole molecule, due to the gain in molecular weight. This effect can be minimized by usage of flexible linkers, as was done in these experiments (compare Figure 4.10, page 83). Such a linker enables the fluorophore to rotate more or less independently from the biomolecule. The second major reason for loss of rotational freedom is potential steric hindrance by neighboring domains, which in the worst case might lead to a docking of the fluorophore to the surface of the macromolecule. However, in the bulk anisotropy experiments *A* ranged from 0.18 to 0.20 in case of the assembled RNPs, which are considered adequate values for FRET experiments[97]. Hence, global interactions between fluorophores and biomolecules as mentioned above, resulting in a major loss of dye movement with severe impacts on the FRET experiments, can be excluded.

Enzymatic Activity

Last, but most importantly (3), the enzymatic activity of the labeled complexes was checked in radioactive pseudouridylation assays. For this purpose, a labeling scheme, where only the target nucleotide is radioactively labeled, was established. The double labeled FRET-H/ACA-RNAs were compared to the transcribed guide RNA in multiple turnover assays with a 20-fold excess of substrate RNA (1 µM) over the H/ACA-RNP (0.05 µM). Samples of each construct were taken at different time points (1 to 180 min) and in order to gain kinetic information of the isomerization rate these data points were fitted to a Michaelis-Menten equation (see Figure 4.21, page 94), yielding the time required for the enzyme to isomerize half of the remaining substrate ($Sub_{1/2}$). For kinetic comparison of individual complexes, the initial velocity of Ψ-formation (v_{start}) in s^{-1} was approximated via linear regression of the early reaction phase. Therefore, the data point at 1 min from the obtained fit was taken (*Note*: this point was used for approximation since it resembled the first physically taken data point, and the reaction progress can be estimated to be linear with a decent certainty for this time window). Errors were not calculated for this

method, since only one data point was used for the linear approximation. This approach was used to make an educated guess for the actual maximum reaction rate of RNA guided pseudouridylation, which could not be determined by the applied conditions; significantly higher substrate excess (> 1000-fold) would be required to mimic ideal settings for Michealis-Menten kinetics. The comparably small excess of substrate RNA over enzyme allow the hypothesis that the real initial velocities should be higher than the approximated values, since binding of already pseudouridylated substrates might occur more frequently and A summary of fitted ($Sub_{1/2}$) and approximated (v_{start}) kinetics of labeled and unlabeled sRNPs under multiple turnover conditions is given in Table 4.3.

RNP	**WT**	**5'-FRET**	**Loop-FRET**
Proteins	LNCG		
Temperature	70 °C		
$Sub_{1/2}$ / min	9.2 ± 0.6	37.9 ± 2.1	55.0 ± 3.3
v_{start} / min^{-1}	1.6	0.42	0.29

Table 4.3: Summarized pseudouridylation kinetics of archaeal H/ACA-RNPs.

The unmodified wt-RNP showed the fastest catalytic turnover with an approximated initial rate of v_{start} = 1.6 min^{-1} ($2.7 \cdot 10^{-2}$ s^{-1}) at the applied conditions, which is around one order of magnitude faster than a previously reported reaction rate of $3.9 \cdot 10^{-3}$ s^{-1} of an H/ACA-RNP from *P. furiosus*[64]. The variation in reactivity is probably due to the different reaction conditions; lower reaction temperature and single turnover settings. However, the high activity of the wt-RNP shows that the purified proteins and transcribed RNA are functional. Insertion of the two fluorophores led to detectable reductions of catalytic activity. A 4-fold decrease in the substrate turnover rate was observed for *5'-FRET-H/ACA* (0.42 min^{-1}) and a slightly greater decrease for *Loop-FRET-H/ACA* down to 0.29 min^{-1}. Nevertheless, despite the slight reduction in pseudouridylation speed, both labeled RNPs were highly capable of Ψ-formation in a multiple turnover fashion. Taking together all the results on RNA-protein binding, rotational freedom of the dyes and especially complex activity, the designed and synthesized RNA constructs are suitable for smFRET experiments and potentially observed states are considered to be of biological relevance in RNA guided pseudouridylation.

4.6.2. smFRET Analysis of the Archaeal H/ACA-RNP

Investigation of the archaeal H/ACA-RNP from *P. furiosus* was done by smFRET spectroscopy. The focus in this project is the structural dynamics of the sRNA during complex assembly, uridine isomerization and product release. As described in the previous chapter, various RNA constructs were successfully synthesized, carrying fluorophore labels and/or biotin linkers at the guide RNA and/or target RNAs. Thus, the possibility to detect conformational changes within the sRNA, or between substrate and guide RNA, was accomplished by either direct immobilization of the H/ACA-RNA or by immobilizing substrate bound RNPs via the target RNA. All samples were annealed at 70 °C, and put on ice immediately after incubation. This procedure was done to freeze individual states for the actual smFRET measurements. Since no enzymatic activity of the thermophile Ψase could be observed at room temperature (discussed in Chapter 4.4.1), these trapped states can be assumed to yield snap shots of the catalytic process. The results of these experiments can be divided into three categories:

1) H/ACA-RNP assembly,

2) Substrate recruitment, and

3) Reactive states during pseudouridylation.

For the following discussion it is important to mention that protein-binding to the RNA cannot be observed directly, since only RNAs were fluorophore labeled in this thesis. However, in experiments carried out by Andreas Schmidt with fluorophore labeled proteins it could be shown that H/ACA-RNPs were readily formed under the given conditions[128]. Thus, estimating that at least 80% of the observed molecules (for samples containing RNPs) were bound by proteins is reasonable.

H/ACA-RNP Assembly

Assembly experiments of the archaeal sRNP were performed with FRET labeled and biotinylated H/ACA-RNAs. Therefore, both guide RNAs were annealed with or without the proteins L7Ae, Nop10, Cbf5 and Gar1 in absence of the target RNA.

Structural changes of the sRNA in between 5'-end and substrate binding site were observed with the *5'-FRET-H/ACA* construct, labeled at bases G1 and U39. Overall, two conformations of the sRNA, an open (E-FRET = 0.45) and a closed state (E-FRET = 0.59), could be observed in the

lower stem of the guide RNA (see Figure 4.22, page 101). The open state represents a distance increase of the attached fluorophores between 5'-terminus and upper end of the pseudouridylation pocket. Assuming a Förster radius of R_0 = 6 nm, as is commonly done for Cy3 and Cy5[97,135,136], the calculated dye distance is 62.0 Å for the open state and 56.5 Å for the closed state, resembling an average distance increase of 5.5 Å between the two labeling sites when the sRNA switches from the closed into the open conformation. It is important to mention that these calculated distances have to be considered with care and should only be used as approximations. The real R_0 values of the experiments were not determined and are strongly depend of the electrochemical environment of both dyes. The summarized assembly experiments for *5'-FRET-H/ACA* are shown in Figure 4.30.

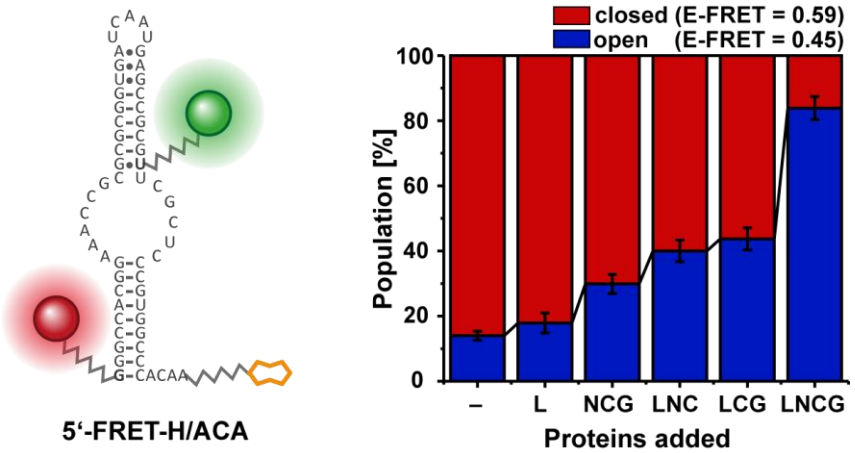


Figure 4.30: Summary of 5'-H/ACA-FRET assembly experiments.

For the free RNA as well as for the L7Ae bound complex, the closed state was far more populated (> 80%). Addition of only three of the four proteins (NCG, LNC or LCG) led to an increase of the open population, but only in presence of all four proteins (LNCG) the opened conformation became the mainly occupied state of the sRNA (> 80%). Most likely, the distance increase of 5.5 Å upon full assembly of the *5'-FRET-H/ACA* labels resembles an opening of the pseudouridylation pocket rather than a conformational change in the lower stem, since the binding pocket should be far more flexible than the base paired 5'-end. In circular dichroism studies of an archaeal Ψase from *Pyrococcus abyssi* it has been reported that the H/ACA-RNA undergoes structural changes upon RNA-protein binding[137]. Apparently, these rearrangements are at least partially associated to the observed opening of the substrate binding site, and all four proteins seem to be required for an efficient opening of the guide RNA. It is not surprising that binding of L7Ae-only – to the kink-loop – shows no significant structural changes in the lower

stem of the sRNA, since the hairpin cannot be investigated with this construct. Omission of Nop10 (LCG) or L7Ae (NCG) yielded also an inefficient opening of the H/ACA-RNA. The most likely explanation for this observation would be the missing interactions between the RNA-binding proteins: L7Ae, Nop10 and Cbf5. The largest of all proteins (Cbf5) covers the pseudouridylation pocket, while L7Ae binds to the kink-loop and Nop10 is positioned in between, binding both proteins and the H/ACA-RNA[52,54,55]. This might explain the necessity of Nop10 and L7Ae for an efficient opening of the substrate binding site. The smFRET data suggest however, that opening of the pseudouridylation pocket is also dependent on Gar1 presence, since removal of this protein (40% open state) has a similar effect on the sRNA structure like removal of Nop10 (44% open state). This behavior was unexpected, since Gar1 does not directly interact with the H/ACA-RNA, but only to Cbf5[52,54]. One explanation for the observed Gar1-dependency on the RNA structure might be a Cbf5-Gar1 interaction that influences Cbf5 binding to the sRNA. Possibly, the absence of Gar1 decreases the binding affinity of Nop10-Cbf5 to the sRNA, which might result in incomplete binding of NC to the H/ACA-RNA, considering the highly diluted samples required in TIRF-based smFRET experiments. Furthermore, the absence of substrate RNA should be considered as well. It is likely that the H/ACA-RNP is more dynamical when no target RNA is bound. The high annealing temperatures (70 °C) might induce switching between different states, and only in presence of all four proteins the open conformation of the sRNA is stabilized.

For the other RNA '*Loop-FRET-H/ACA*', which is labeled at bases U26 and U39, a different behavior was observed. This construct was used to detect structural dynamics in the upper stem of the sRNA. Again, two conformations were detected; termed 'free state' (E-FRET = 0.67) and 'docked state' (E-FRET = 0.80). Under the same assumption as for the other construct, the average distance of Cy3 and Cy5 was 53.3 Å in the free and 47.6 Å in the docked conformation. The summarized assembly experiments for this construct are given in Figure 4.31.

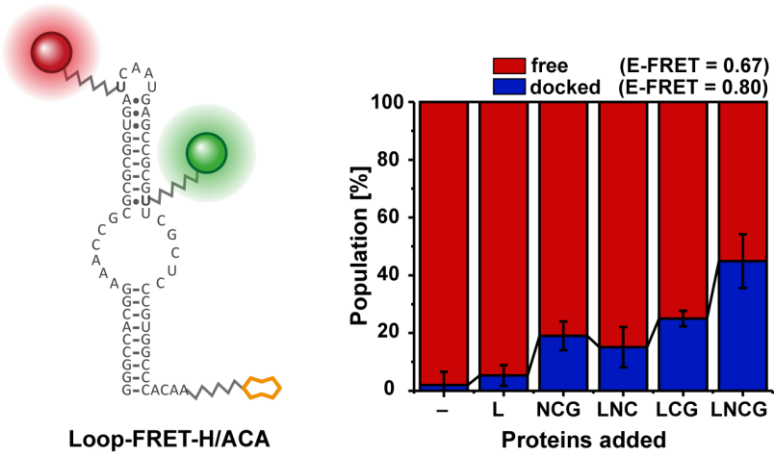


Figure 4.31: Summary of Loop-H/ACA-FRET assembly experiments.

This time addition of the proteins led to an increase in FRET-efficiency, representing an estimated decrease in fluorophore distance of 5.7 Å. However, the observed transition to the docked state was not near quantitative, as observed for *5'-FRET-H/ACA* in presence of LNCG (see Figure 4.30, page 113). Upon addition of all four proteins only 45% of the detected molecules showed docking in the upper stem of the sRNA. Removal of one (minus L7Ae, Nop10 or Gar) or multiple (minus NCG) archaeal H/ACA-proteins led to a significant loss in particles populating the docked conformation, ranging from 5% (RNA+L) to 25% (RNA+LCG) of RNPs in the high-FRET state. Interestingly, no significant changes in FRET-efficiency were observed upon L7Ae addition, even though this protein is known to independently bind to the kink-loop of the sRNA[57,66]. This could either mean that no major structural rearrangement happens at the sRNA upon L7Ae binding, or that it could not be detected in the used construct by smFRET, which might be the case if the two fluorophores inherit the same vectorial distance in both conformations (unbound and bound to L7Ae). Addition of the other proteins however yielded differences in E-FRET. The partially observed rearrangement of the upper stem in the H/ACA-RNA might have two possible explanations. One being: some of the detected RNA molecules might not be bound by LNCG, since presence of the proteins cannot be confirmed with only the guide carrying fluorophores. This however is unlikely, since the other construct showed an almost quantitative shift from the closed to the open conformation, which strongly indicates that most sRNAs are bound by proteins. Second and more likely: the upper stem of the H/ACA-RNA inherits high conformational flexibility upon complex assembly at 70 °C. Here again, the high temperatures may induce structural rearrangements which can be trapped by fast cool down of

the sample. The existing of the docked state indicates a certain functional role of the upper stem in the H/ACA-RNA, which cannot be assigned by the performed experiments.

Substrate Recruitment

In the assembly experiments with substrate free H/ACA-RNPs the impact of individual proteins on guide RNA conformation was studied. To investigate the impact of these proteins on substrate recruitment the 5fU-substrate analogue – known for inhibiting the pseudouridylation process – was used and the labeling scheme was altered. Construct *5'-Cy3-H/ACA*, where the FRET-donor is attached to the 5'-end of the guide RNA, and substrate analogues labeled at the 3'-end (FRET-acceptor) were used. Hence, FRET can only be observed when the substrate is in close proximity to the sRNA. It is important to mention that the discussed data in this subchapter should only be considered on a qualitative scale, since effects like acceptor bleaching or inconsistent numbers of immobilized molecules per channel – due to varying channel widths during slide preparation – might severely affect the amount of detected FRET-events. Nevertheless, the relative of number of particles showing FRET over donor-only molecules (E-FRET < 0.2) should not be affected by these two effects and allow for a qualitative comparison of the different samples. The full RNP consisting of guide RNA, target RNA and LNCG was compared to partial RNPs, where either Nop10, L7Ae or Gar1 were omitted (see Figure 4.26, page 104). For the full complex LNCG and in absence of Gar1 distinct populations above E-FRET > 0.2 were detected. For both complexes similar substrate binding efficiencies were observed; with 59% of particles carrying a target RNA in case of LNCG (n = 4530) and 57% of particles for the ΔGar1 complex (n = 3458). The Gaussian distribution of both high-FRET populations indicates a specific binding mode of the target RNA to the H/ACA-RNP, as would be expected for the base pairing RNAs. For the RNPs missing Nop10 or L7Ae considerably less substrate bound molecules were detected; 11% (n = 164) for ΔNop10 and 12% (n = 321) for ΔL7Ae particles. Thus, these proteins are required for effective and correct substrate recruitment into the pseudouridylation pocket of the guide RNA and the smFRET data suggest that both proteins contribute equally to target RNA binding capability of the H/ACA-RNP. Hence, the L7Ae-Nop10-Cbf5 interaction plays a key role in this process, since both L7Ae and Nop10 are both required for efficient recruiting of the target RNA, even though they are not in direct contact with the substrate. This hypothesis is supported by the results of the assembly experiments, especially by these of construct *5'-FRET-H/ACA* where a quantitative opening of the

pseudouridylation pocket could only be observed in presence of all proteins. Furthermore, by combining crystallography and computational modelling Liang *et al*. showed that positioning of the target base in a substrate assembled RNP may vary in absence or presence of L7Ae and they proposed that L7Ae has an important role in target RNA placement[53]. This proposal is now directly supported by the substrate recruitment assay and apparently, the smallest protein Nop10 is the key mediator in between the kink-loop binding L7Ae, the reactive enzyme Cbf5 and the guiding sRNA for proper substrate positioning.

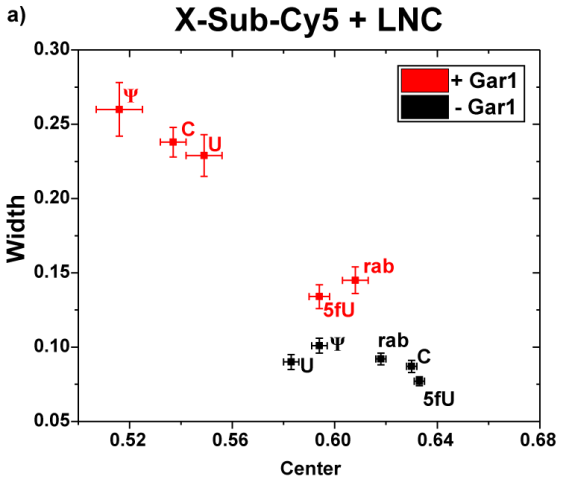
Besides the strongly reduced substrate binding ability of the H/ACA-RNP upon L7Ae or Nop10 removal, Figure 4.26 shows also a clear difference in the observed FRET population of the full complex (LNCG) in contrast to the Gar1-missing RNP (ΔGar1). For LNCG, a peak width of w = 0.23 was observed, while for ΔGar1 this value is significantly reduced down to w = 0.08 FRET-efficiency, indicating that the presence of the protein induces structural dynamics in the system. Since Gar1 does not directly interact with the target RNA – but only with Cbf5 – these dynamics are probably triggered by the thumb loop domain of Cbf5, which is in close proximity to the Gar1 binding site[52,54] and is proposed to regulate recruitment and release of the target RNA in the sRNP[63]. In fluorescence correlation spectroscopy studies it was also suggested that Gar1 facilitates product release[64], giving rise to the assumption that the larger peak width of the complete RNP (LNCG) results from the attempt of the Ψase to remove the target RNA. This however cannot occur due to the fact that the thermophile complex is inactive at the measurement conditions (room temperature). Furthermore, direct evidence for these Gar1 induced dynamics were reported by smFRET studies from Andreas Schmidt, in which he fluorescently labeled the thumb loop of Cbf5 and different substrate analogues. In absence of Gar1 the thumb loop predominantly remained in an open conformation, while Gar1 presence enables the thumb loop to populate other high-FRET states, which are required for catalytic turnover[128].

Reactive States

The experiment described above was repeated with different substrate RNAs. Atomic mutagenesis on the target base allowed isolation of individual reactive states. Reactive states were probed with the *bona fide* substrate ‘U’ and the inhibitor ‘5fU’ at the target site, release states were mimicked by the product ‘Ψ’ and the mismatch ‘C’, and an uncatalytic state was imitated by the abasic nucleotide ‘rab’. Each of the different target RNAs ‘X-Sub-Cy5’ carried

the FRET-acceptor at the 3'-end. Hence, distance changes between the lower stem of H/ACA-RNA and the substrate RNA were detectable. It is assumed that the major contribution of deviations in E-FRET can be attributed to conformational rearrangements at the target RNA, since the lower stem of the H/ACA-RNA forms a rigid double helix structure with 9 base pairs of which eight represent G-C-pairing. Influences of Gar1 were investigated by its presence or absence in the designed experiment (see Figure 4.27, page 104). Overall removal of this protein led to the same behavior for the other substrate analogues as was observed for 5fU-Sub-Cy5 in the above mentioned experiment. The full width at half maximum ranged from 0.08 to 0.10 E-FRET for ΔGar1 RNPs, indicating a rigid complex at both RNA sites. Addition of Gar1 led to an overall increase in peak width from 0.13 up to 0.26 E-FRET. Moreover, for the release state mimicking Ψ- and C-Sub-Cy5 shifts of the peak center x_c to lower FRET-values in the presence of Gar1 could be observed. The summary of these results is shown in Figure 4.32.

a)



b)

X-Sub-Cy5	Gar1	Peak center ± Error	Peak width ± Error
U	+	0.549 ± 0.007	0.229 ± 0.014
U	-	0.583 ± 0.003	0.090 ± 0.005
5fU	+	0.594 ± 0.004	0.134 ± 0.008
5fU	-	0.633 ± 0.002	0.077 ± 0.003
Ψ	+	0.516 ± 0.009	0.260 ± 0.018
Ψ	-	0.594 ± 0.003	0.101 ± 0.005
C	+	0.537 ± 0.005	0.238 ± 0.010
C	-	0.630 ± 0.002	0.087 ± 0.004
rab	+	0.608 ± 0.005	0.145 ± 0.009
rab	-	0.618 ± 0.002	0.092 ± 0.004

Figure 4.32: ***a)*** **Plot of peak center x_c against peak width w of smFRET experiments for 5'-Cy3-H/ACA with X-Sub-Cy5 plus LNC in presence (red) or absence (black) of Gar1.** ***b)*** **Table of obtained data by Gaussian fits from the smFRET experiments in Figure 4.27. Errors were obtained from fitting results.**

In Figure 4.32a the peak centers (x_c) are plotted against the peak widths (w) for each individual X-Sub-Cy5 RNAs in presence (red) or absence (black) of Gar1. Several trends are discernible from this plot: Gar1 systematically induces movement on the target RNA site, indicated by the increasing peak width, especially for Ψ-, C- and U-Sub-Cy5. Furthermore, these three constructs also show a reduced FRET-efficiency, indicated by the lower peak centers. The smaller E-FRET value are most significantly observed for the release state mimicking RNAs (Ψ- and C-Sub-Cy5) and for the actual substrate U-Sub-Cy5, which might as well be partially isomerized to Ψ due to preincubation of the sample. These results suggest a release state and function of the H/ACA-

RNP which is dependent on Gar1, agreeing with reports that this protein promotes product release[64]. Another indirect evidence for the proposed release state is the behavior of the rab-Sub-Cy5, used to mimic an uncatalytic state. Of all the tested substrate analogues the abasic target RNA shows the least Gar1-dependency, as would be expected for a non-reactive state. Conclusively, the Ψase recognizes the target nucleotide in case of the product (Ψ) or a mismatch (C), and subsequently the target RNA experiences structural alterations to be released from the H/ACA-RNP.

After investigation of reactive states on the target site, structural dynamics within the guide RNA during pseudouridylation were also studied. For this, the immobilization strategy was altered and construct *FRET-H/ACA-unbio* was used for these experiments. The biotin label was removed from the H/ACA-RNA, and complex pull down was achieved via 3'-biotinylated substrate RNAs. FRET labels were attached to the sRNA at position U26 (donor) and U39 (acceptor), allowing the observation of conformational rearrangements in the upper stem of the guide RNA. For the investigation of reactive states during pseudouridylation again five different target RNAs (U-, 5fU-, Ψ-, C-, and rab-Sub-bio) were used. For four of the five different substrate analogues only one major FRET population was observed at E-FRET = 0.72, probably representing the docked conformation from the assembly experiments. Interestingly, there is a shift in E-FRET values for the docked states of *loop-FRET-H/ACA* (used in the assembly experiments; see Figure 4.24, page 102) and of *FRET-H/ACA-unbio* (used for characterization of reactive states; see Figure 4.28, page 106). Two likely reasons might explain this observation. One being the swapping of the dyes (Cy3 and Cy5) at the labeling positions (U26 and U39). Due to the different size of the fluorophores and the change of the steric environment minor changes in FRET-efficiencies might occur, and more importantly the presence of the bound target RNA. In the assembly experiments compaction of the upper stem was far from quantitative (45%) and conformational changes might occur at high temperatures – that could not be detected at smFRET conditions. The presence of the target RNA might stabilize the overall H/ACA-RNA structure in this region leading to a quantitative docking of the upper stem. However, a new high-FRET state at E-FRET = 0.90 could be detected for the mismatched target RNA C-Sub-bio (see Figure 4.28, page 106). An overlay of 5fU-, rab- and C-sub-bio histograms emphasizes the formation of this new high-FRET conformation in the upper stem for *FRET-H/ACA-unbio* (see Figure 4.33).

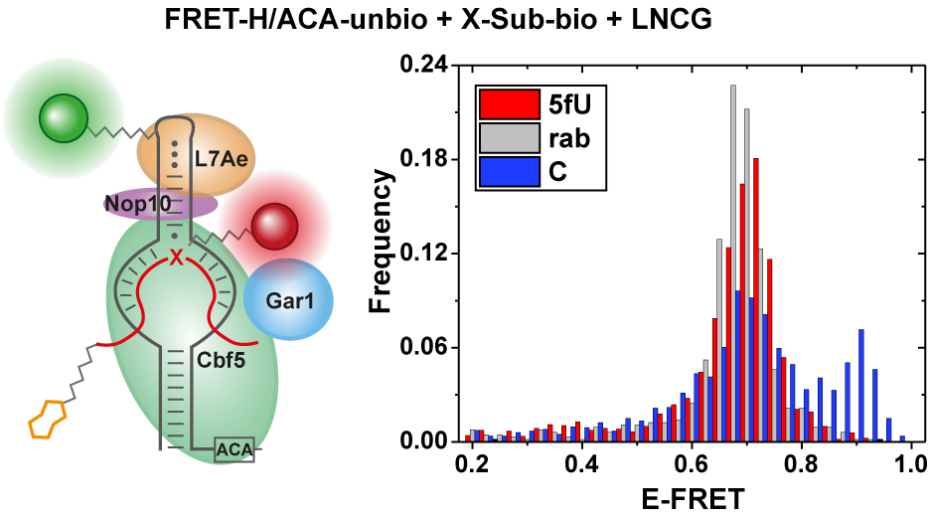


Figure 4.33: Overlaid smFRET histograms of FRET-H/ACA-unbio plus LNCG for 5fU/rab/C-Sub-bio.

The high-FRET state was only observed for the fully assembled H/ACA-RNP, indicating a tight compaction of the upper stem for the mismatched C-Sub-bio. Omission of Gar1 led to a complete disappearance of the new state (see Figure A3.4 in Appendix A3 – Supplementary Figures), showing that this protein is essential for effective stem compaction. A quantitative population of this high-FRET state could not be observed (with varying annealing times from 1 to 15 minutes). At room temperature transitions between the compact (E-FRET = 0.72) and high-FRET state (E-FRET = 0.90) were not detectable. Only in case of the substrate analogue carrying a C at the target site the upper stem of the guide RNA showed this tight compaction, indicating a possible discrimination function of the RNA guided Ψase for RNAs that were bound by mistake. Another likely explanation for the observed high-FRET population is a possible RNA release state of the H/ACA-RNP, which could be trapped by C-Sub-bio during the reaction pathway. Yang *et al.* reported a three to four fold decrease for the off-rates of C-Sub compared to Ψ-Sub[64]. Hence, the mismatch at the target site reduces the kinetics of product release – possibly by altering the H-bond interactions between target RNA and sRNP – and therefore might enable trapping of a release state by fast cooling. A clear statement on which of these two possibilities is true cannot be made by smFRET experiments. However, this tight compaction of the upper stem has not been reported yet and might be of interest for further crystallization studies with a target RNA containing C at the modification site, which has not been reported yet. So far, only crystal structures with 5fU, 3-methyluridine, 4-thiouridine or 2'-deoxyuridine at the target site are known[54,56].

The same experiments were also performed with an H/ACA-RNA labeled and at U39 and G1 instead of U26 (corresponding to *5'-FRET-H/ACA* construct without the biotin-linker at the 3'-end) to see whether structural rearrangements occur in between the lower stem of the guide RNA. For this construct however, only one FRET-population – representing the open conformation – was observed for all tested target RNAs (see Figure A3.5 in Appendix A3 – Supplementary Figures), pointing to the assumption that the lower stem of the sRNA remains passive and does not experience any major rearrangements during pseudouridylation.

4.7. Conclusion

RNP Assembly

In the following the smFRET data for the archaeal H/ACA-RNP from *P. furiosus* will be summarized and evaluated. An overview of sRNA rearrangements during complex assembly and pseudouridylation is given in Figure 4.34.

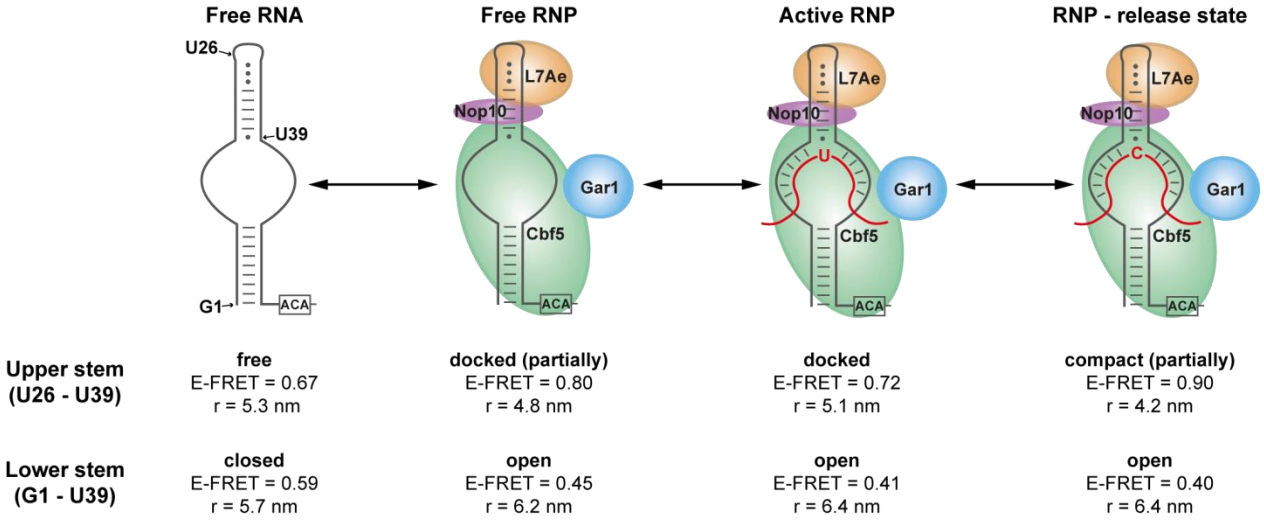


Figure 4.34: Summarized data of detected FRET-states for FRET-labeled H/ACA-RNAs. Obtained E-FRET values of free RNA, substrate-free RNP, substrate-bound RNP and mismatched target RNA bound RNP (from left to right) of upper and lower stem are depicted. Nucleotides used for labeling are highlighted in the free RNA. The average intramolecular dye distance 'r' was calculated assuming a Förster radius of R_0 = 6.0 nm.

For the lower stem of the H/ACA-RNA two overall conformations were observed: a closed and an open state (see Figure 4.30, page 113). The closed state was predominantly occupied by the free sRNA. Upon protein binding (i.e. addition of LNCG) the guide RNA switched to the open state. Presumably, H/ACA-RNP formation leads to a rearrangement of the pseudouridylation pocket, promoting substrate recruitment. This opening of the guide RNA was only efficient when all four proteins were present. Once the open state was occupied no significant conformational changes were observed between 5'-end and substrate binding pocket, indicating that the lower stem of the sRNA remains intact during substrate recruitment, turnover and product release.

The upper stem (between loop and substrate binding pocket) however indicated a more dynamical behavior during H/ACA-RNP assembly and catalysis in the smFRET experiments. Overall three different conformations were identified in this region: a free, a docked and a compact state. In absence of proteins the sRNA occupied the free state. Addition of LNCG led to the formation of the compact state, which however was only partially populated (45%) after

H/ACA-RNP assembly (see Figure 4.31, page 115), in contrast to the open state (84%) of the lower stem. This partial occupation of two states (free and docked) from the substrate free sRNP indicates structural dynamics between loop and pseudouridylation pocket occurring at high temperatures (> 70 °C) and might be associated to a reversible compaction of loop and hairpin structure. Upon substrate binding only one distinct conformation (presumably the compact state) was observable for the upper stem, indicating that the presence of the target RNA stabilizes this part of the sRNA to a certain degree.

Substrate Recruitment

In substrate recruitment assays in which the FRET-acceptor was attached to the target (5fU-Sub-Cy5) it was shown that the presence of L7Ae and Nop10 are essential for efficient target RNA binding to the Ψase (see Figure 4.26, page 104). These results, as well as the fact that an opening of the substrate binding site only occurred readily in presence of these two proteins, are direct evidence for the role of L7Ae proposed by Liang *et al.* in correct target RNA placement for pseudouridylation[53]. Furthermore, the smallest of all proteins Nop10, seems to play a key role in this process by binding to and mediating the interactions between L7Ae (the kink-loop binding protein), Cbf5 (the reactive enzyme) and the H/ACA-RNA. Gar1 did not show such a significant effect on substrate recruitment, but its presence indicated dynamics in between target RNA and the sRNP by a significantly higher peak width of the FRET-population for the full H/ACA-RNP. The impact of Gar1 on RNP-dynamics was studied by atomic mutagenesis of the target nucleotide in labeled substrate analogues (X-Sub-Cy5), mimicking different states during pseudouridylation: reactive- (U and 5fU), release- (Ψ and C) and uncatalytic-state (rab). Presence of Gar1 induced dynamics at the target RNA for all samples; represented by the greater peak width (see Figure 4.32, page 118). Furthermore, substrate analogues carrying Ψ and C, as well as U (which might be partially isomerized to Ψ due to preincubation of the samples) at the target site showed a significant shift of the peak center towards lower FRET-values in comparison to 5fU- and rab-Sub-Cy5 when Gar1 was present. These findings suggest a release function of the Ψase in which target RNAs are rearranged in order to be removed from the H/ACA-RNP and this function seems to be Gar1-dependent.

Reactive States

Potential dynamics within the guide RNA during pseudouridylation were also studied by atomic mutagenesis of the target nucleotide in a biotinylated substrate RNA 'X-Sub-bio' (see Figure 4.28, page 106). Immobilization of FRET-labeled guide RNAs via X-Sub-bio permitted the detection of various states within the sRNA and ensured that only target RNA-bound RNPs were observed, since it was shown that L7Ae, Nop10 and Cbf5 are required for efficient substrate recruitment. In the above described assay it was shown that the proteins (L7Ae, Nop10 and Cbf5) are required for effective substrate binding. Between lower stem and pseudouridylation pocket of the guide RNA no apparent changes were observed for any of the target RNAs, indicating that once the substrate binding sites opens no significant structural rearrangements occur during pseudouridylation. Regarding the upper stem of the sRNA, also no significant changes in FRET-populations for U, Ψ, 5fU and rab at the target site were observed, suggesting that the upper stem remains in its original conformation during pseudouridylation as well. For C-Sub-bio however, a new high-FRET population termed 'compact state' (E-FRET = 0.90) occurred. Since this state was only observed for the mismatched target RNA, this new FRET-state could be an indication for a discrimination function of the H/ACA-RNP to effectively release mistakenly bound RNAs of similar sequence that cannot be pseudouridylated. Another explanation for this compaction of the upper stem might be the existence of a release state for isomerized substrates in the H/ACA-RNP, which could be trapped by fast cooldown after sample annealing. In case of the latter assumption, the compact state should have been visible for Ψ-Sub-bio as well, which was not the case. A possible explanation therefore might be the increase in dissociation kinetics of C-Sub in comparison to the product Ψ-Sub, as was reported by Yang *et al.*[64]. Which of these two possibilities is correct cannot be proven by the obtained data. The results from investigation of reactive states via X-Sub-Cy5 (see Figure 4.32, page 118) support the assumption of a release state over the discrimination function, since similar behavior of x_c and w were observed for U-, Ψ- and C-Sub-Cy5 in presence and absence of Gar1.

Chapter V

Eukaryotic H/ACA-RNP from *Saccharomyces cerevisiae*

5.1. Protein Preparation and Purification

5.1.1. Preparation of Nhp2

The second project investigates the biological function of the eukaryotic H/ACA-RNP. Equally to the archaeal project, eukaryotic proteins had to be prepared and tested for enzymatic function, in order to produce reliable smFRET data. For investigation of the snoRNP all four H/ACA-proteins were obtained by recombinant expression in *E. coli* and subsequent purification. Competent BL21(DE3) cells containing the plasmids with the respective genes (scCbf5 and scNop10 in pET-Duet-1, scGar1 in pET-28b, and Nhp2 in pET-24b) for protein expression were provided by Sven Trucks (Goethe University). Protein expression had to be carried out at 20 °C overnight, since higher incubation temperatures led to visible formation of inclusion bodies. The purification of Nhp2 – the eukaryotic analogue of archaeal L7Ae – was divided into several steps: 1) Cell lysis and centrifugation, 2) nucleic acid precipitation via polyethylenimine, 3) initial affinity chromatography, and finally 4) size exclusion chromatography. A HisTag at the C-terminus of Nhp2 was introduced for affinity based purification steps. Furthermore, a mutant of this protein was used, in which the amino acid serine at position 82 is substituted by tryptophan (S82W). This modification stabilizes an intrinsic *cis* conformation in the protein and promotes the association of Nhp2 to the snoRNP[65]. Additionally, the aromatic indole ring of the tryptophan – the only one in the protein – increases the absorbance at 280 nm, simplifying detection and quantification of the protein during purification. In the following, the procedure of Nhp2 purification will be shown and discussed.

Likewise to the archaeal proteins, cells containing over expressed proteins were lysed by sonication. Prior to cell lysis DNase and RNase were added to the mixture to reduce the amount of long nucleic acids, since Nhp2 is expected to bind endogenous *E. coli* RNAs. After centrifugation of the cell debris, polyethylenimine (c_{final} = 0.8 µg/mL) was added to the supernatant to remove further nucleic acids. Initial chromatic purification was done using a Ni-NTA-column exploiting the C-terminal HisTag. Therefore, a gradient elution from 20 mM to 500 mM imidazole was performed and fractions showing high UV-absorption were analyzed by SDS-PAGE (see Figure 5.1).

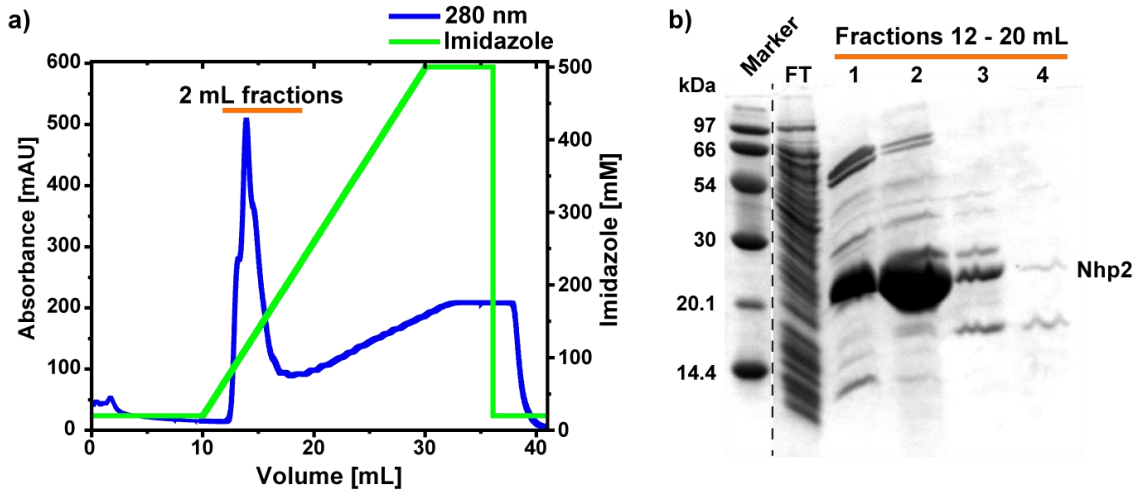


Figure 5.1: ***a)*** **Ni-NTA-chromatogram of Nhp2 expression on a 1 mL Ni-NTA-column.** ***b)*** **Resulting 15% SDS-PAGE of fractions showing high UV-absorption.**

The affinity chromatogram shows a clear absorbance peak for the gradient elution (see Figure 5.1a). In the analytical gel a major protein band is apparent (see Figure 5.1b), indicating successful Nhp2 expression. The presumed Nhp2 (18.3 kDa) band migrates slightly over the 20.1 kDa reference from the marker, but minor deviations in migration pattern due to charged amino acids in the protein are possible in SDS gel electrophoresis[53]. Fractions 1 to 3 showed high amounts of the desired protein, but also considerable impurities. Therefore, these fractions were pooled and further purified by SEC.

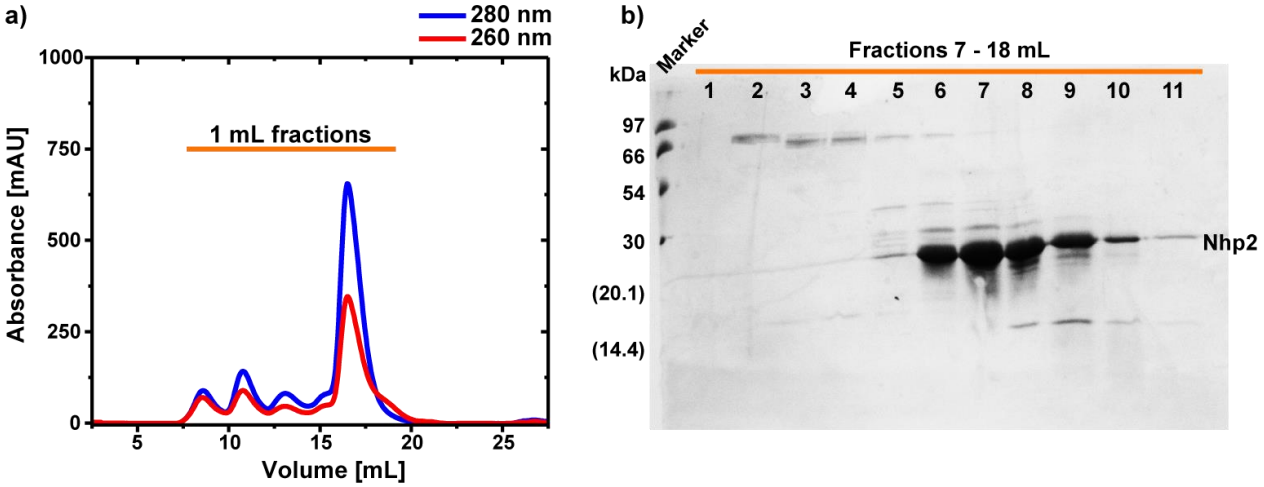


Figure 5.2: ***a)*** **Nhp2 size exclusion chromatogram of combined fractions 1 to 3 of the Ni-NTA-chromatography on a Superdex 75 10/300 GL.** ***b)*** **Resulting 15% SDS-PAGE of fractions showing high UV-absorption.**

The size exclusion chromatogram displays effective separation of three minor peaks, between 6 to 15 mL, and one major peak at around 16 mL (see Figure 5.2a). All peaks were analyzed by SDS-PAGE. In the analytical gel a dominant Nhp2 band is visible for fractions 6 to 10 (see

Figure 5.2b), resulting from the major peak in the gel filtration. The 260/280 nm absorption ratio of 0.53 for the major peak indicates RNA-free protein, demonstrating that the applied purification strategy is suitable for Nhp2 preparation. The total yield was ca. 1.2 mg of purified protein for a 500 mL expression scale. However, in some Nhp2 preparations RNA-contaminations were not completely removed by this protocol. An additional RNase treatment (20 °C, 16 h, c_{final} = 0.04 µg/µL) after the first affinity chromatography ensured efficient RNA removal. The RNase was then removed by a second Ni-NTA-purification step. RNA-free Nhp2 used in the further mentioned experiments was provided by Sven Trucks (Goethe University), applying the expended protocol (additional RNase treatment) for removal of RNA contaminations.

5.1.2. Preparation of the Eukaryotic scNCG Subcomplex

Eukaryotic scNop10, scCbf5 and scGar1 from the model organism yeast were prepared in a comparable manner as the archaeal proteins. Coexpression of the ternary scNCG complex was performed from codon optimized genes in *E. coli* as described above. Full-length constructs of the three proteins and histidine tagged scCbf5 (C-terminal) and scGar1 (N-terminal) were used. The purification was divided into several steps: 1) Cell lysis, 2) initial affinity chromatography, 3) RNA digestion, 4) repeated affinity chromatography, and 5) size exclusion chromatography. RNase and DNase were added to the cell lysis by sonication, as was done for Nhp2. After centrifugation of the cell suspension nucleic acids in the supernatant were precipitated via polyethylenimine (c_{final} = 0.8 µg/mL). Since the proteins did not stem from a thermophile organism, the denaturing process by heating the cell lysate to 70 °C – as was done for the archaeal scNCG – was omitted. Initial chromatographic purification was done with a Ni-NTA-column exploiting the HisTags in the expressed proteins. A representative chromatogram and an analytical gel are shown in Figure 5.3.

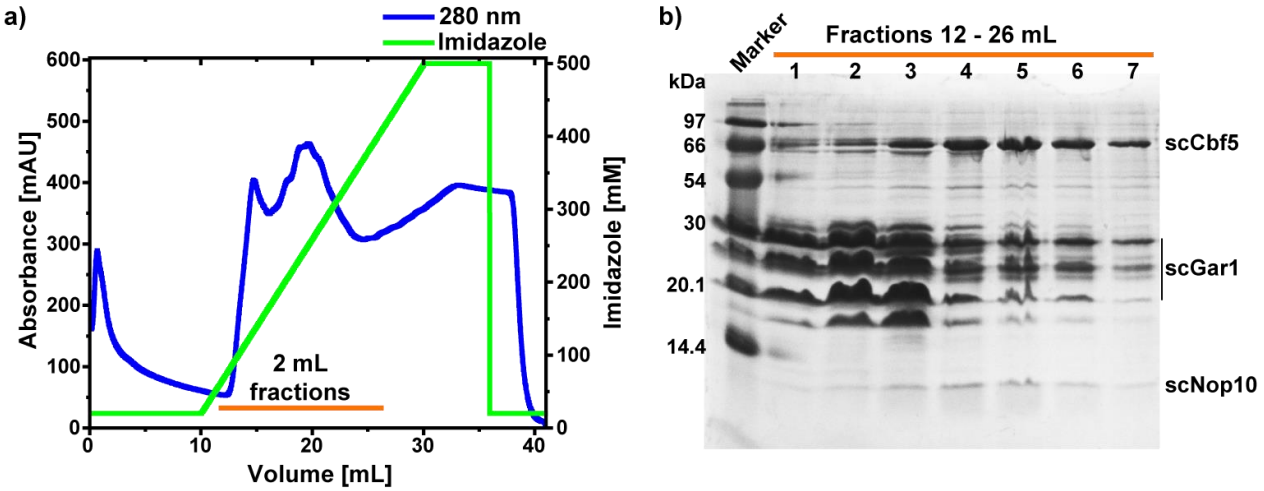


Figure 5.3: ***a)*** **Ni-NTA-chromatogram of eukaryotic scNCG expression on a 1 mL Ni-NTA-column.** ***b)*** **Resulting 15% SDS-PAGE of fractions showing high UV-absorption. Chromatogram and gel were performed by Jason Nigel Martins within the scope of a bachelor thesis.**

The affinity chromatogram of scNCG shows several peaks for the gradient elution; one at the beginning of the run, resulting from sample application, and one broad inhomogeneous peak between 12 and 26 mL, resulting from an increase in imidazole concentration (see Figure 5.3a). The gel electrophoretic analysis of the broad peak shows the successful elution of the three proteins scCbf5 (55.5 kDa), scGar1 (23.6 kDa) and scNop10 (6.6 kDa). As previously reported, full-length eukaryotic scGar1 is prone to degradation[59]. Thus, several bands are visible in the SDS-PAGE analysis for this protein. Fractions 1 to 3 showed major impurities, therefore fractions 4 to 7 were combined and used for further purification. Determination of the 260/280 nm ratio (> 1.8) indicated significant RNA contamination. Hence, RNase was added to the pooled fractions and incubated at 20 °C for 30 minutes. Afterwards, a second affinity chromatography was performed for RNase removal, which was also analyzed by SDS-PAGE (see Figure A3.6 in Appendix A3 – Supplementary Figures). Fractions containing scNCG were combined and used for SEC.

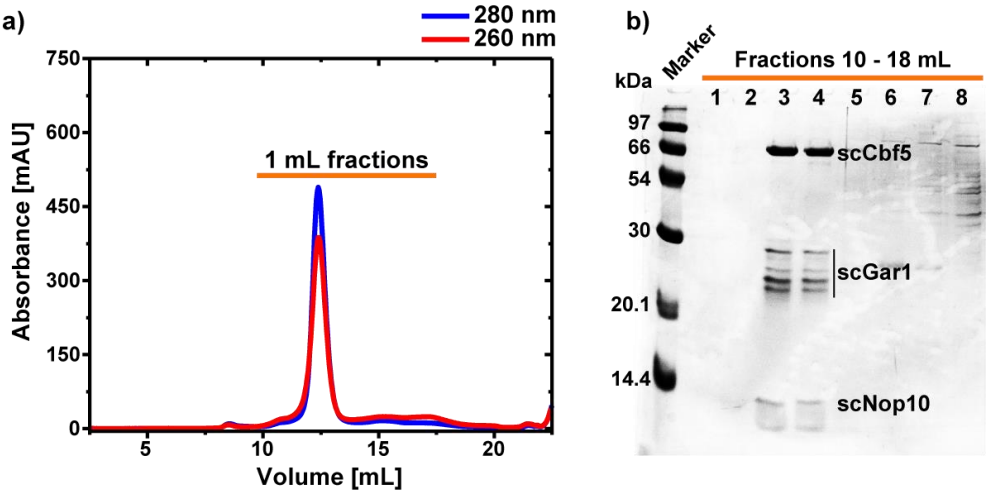


Figure 5.4: ***a)*** **Eukaryotic scNCG size exclusion chromatogram of combined fractions 4 to 6 of the Ni-NTA-chromatography on a Superdex 200 10/300 GL.** ***b)*** **Resulting 15% SDS-PAGE of fractions showing high UV-absorption. Chromatogram and gel were performed by Jason Nigel Martins within the scope of a bachelor thesis.**

The size exclusion yielded one distinct peak for the ternary scNCG complex (see Figure 5.4a). In the resulting SDS-PAGE it is apparent that all three proteins were successfully copurified (see Figure 5.4b). For scCbf5 and scNop10 one homogeneous band was detected in the gel, while scGar1 is partially degraded within the scNCG-complex. Nevertheless, it remains associated to the ternary protein complex. The 260/280 nm absorbance ratio of 0.79 in the SEC chromatogram indicates minor RNA impurities, despite the RNase treatment (a 260/280 nm ratio of < 0.60 would indicate RNA-free protein). Purification of a 1 L expression culture yielded 0.7 mg of protein.

Preparation of stable Gar1 variants in the scNCG complex

In the later stages of this thesis, it was attempted to prepare stable full-length scGar1 protein within the scNCG complex. Therefore, two additional eukaryotic scGar1 constructs were designed: scGar1-notag without any HisTag ‘scNCG-notag’, and scGar1-Ctag with a C-terminal instead of the N-terminal HisTag ‘scNCG-Ctag’ (*Note*: In all constructs scCbf5 still carries the C-terminal HisTag). For clarity, the above mentioned eukaryotic scNCG complex will be termed ‘scNCG-Ntag’ in this subchapter. For scNCG-notag and scNCG-Ctag a part of the N-terminus – containing the HisTag in scNCG-Ntag – was removed (amino acid sequences are given in Table 3.9 on page 59). Hence, the number of HisTags in the scNCG complex was decreased down to one (at scCbf5) for scNCG-notag, or remained two for scNCG-Ctag and -Ntag (at scCbf5 and scGar1). Table 5.1 summarizes the applied constructs and the respective positioning of intramolecular HisTags.

Protein	HisTag	scNCG-Ctag	scNCG-notag	scNCG-Ntag
scGar1	C-terminal	+	-	-
	N-terminal	-	-	+
scCbf5	C-terminal	+	+	+

Table 5.1: Summary of HisTag positioning of the various scNCG-constructs used in this thesis.

In order to obtain RNA-free proteins, the purification strategy was optimized. An additional washing step after sample loading onto the first Ni-NTA-column was introduced. By washing the column bound proteins with a high salt (2 M LiCl) buffer, unspecific RNA-protein interactions should be dissociated, leading to an elution of free RNAs while the scNCG complex remains attached to the column. Afterwards the purification strategy remained as described above; gradient elution, RNA digestion, repeated affinity chromatography and final size exclusion. The adapted protocol was tested for scNCG-Ctag.

After application of the sample to the affinity column, the usage of the high salt buffer led to a significant increase in UV detection. An analytical gel of this peak showed no significant elution of scNCG. The washing step, the SDS-gel and the subsequent affinity chromatogram are shown in the appendix (see Figure A3.7 in Appendix A3 – Supplementary Figures). Fractions containing all three proteins were combined for RNase treatment (incubation for five hours at room temperature) and then again Ni-NTA purified. The gradient elution was analyzed by SDS-PAGE, and fractions containing the scNCG complex combined for gel filtration. The SEC chromatogram and the resulting analytical gel are displayed below.

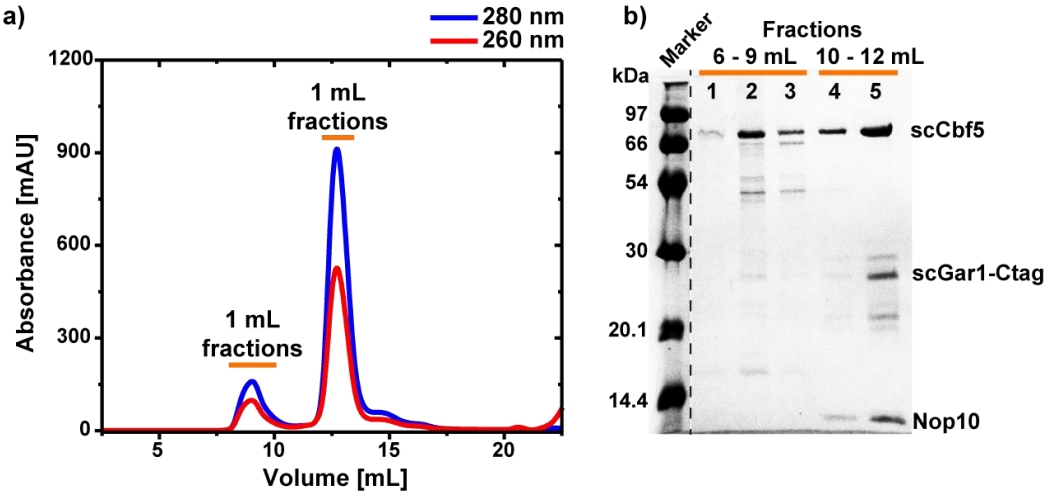


Figure 5.5: ***a)*** **Size exclusion chromatogram of scNCG-Ctag after high salt wash and RNA digestion on a Superdex 200 Increase 10/300 GL.** ***b)*** **Resulting 15% SDS-PAGE of fractions showing high UV-absorption.**

Figure 5.5a shows the effective purification of the ternary scNCG-Ctag complex. The 260/280 nm ratio of 0.58 indicates RNA-free protein. Thus, the new strategic approach to efficiently free the protein complex of RNA by adding a high salt washing step proved successful. Furthermore, the 2 M LiCl concentration did not show any major effects on the stability of the scNop10-scCbf5-scGar1 complex, since one distinct peak in the size exclusion chromatogram for scNCG-Ctag was detected. However, purification of homogeneous scGar1 could not be observed in the analytical gel, and degradation bands were still visible after the SEC in fraction 5 (see Figure 5.5b). Nevertheless, comparing the SDS-PAGEs of purified scNCG-Ctag to purified scNCG (see Figure 5.4b); scGar1-degradation appears to be less. Furthermore, the degradation process seems to be time dependent, since homogeneous scGar1 bands could be observed for scNCG-Ctag and scNCG-notag during the purification procedure of the ternary complex, but not for scNCG-Ntag (see Figure 5.6). Thus, introduction of a C-terminal HisTag or removals of the HisTag in scGar1 are promising attempts to prevent intramolecular scission of the protein.

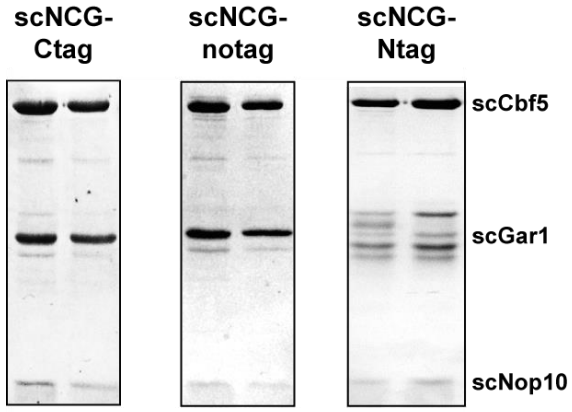


Figure 5.6: Stability comparison of scGar1 in different scNCG constructs; C-terminal HisTag (left), no HisTag (center) and N-terminal HisTag (right) at scGar1. For each construct two fractions of a SDS-PAGE after the second Ni-NTA-purification (after RNase digestion) were compared.

5.2. Synthesis, Labeling and Purification of snoRNA Constructs

5.2.1. Unlabeled Eukaryotic H/ACA-RNAs

In order to reconstitute snoRNPs and to perform activity assays, unlabeled guide RNAs had to be synthesized. For functional studies of the eukaryotic H/ACA-RNP, different RNA constructs were designed derived from the full-length snR81 snoRNA of yeast. Initial investigation of the bipartite Ψase was to be performed on the partial 5'-hairpin of the snoRNA. Structural similarities to the archaeal H/ACA-RNP from *P. furiosus* and the knowledge that this standalone hairpin of the snoRNP inherits catalytic activity on its own[59,72], qualified it as an interesting starting complex for comparison of eukaryotic to archaeal pseudouridylation. Furthermore, the influences of well conserved and comparably large loop sequences of the eukaryotic H/ACA-RNAs on pseudouridylation were to be investigated as well. For these purposes, two constructs were designed; the full-length 5'-hairpin 'H5' with the native nucleotide sequence of the loop, and a shortened version 'H5d', in which the upper stem of the hairpin is truncated and replaced by a UUCG tetra loop (see Figure 5.7). In nucleotide length and loop structure the H5d construct (55 nt) is more similar to the archaeal H/ACA-RNA (58 nt) than the H5 construct (90 nt), allowing for a better comparison of eukaryotic to archaeal RNA guided pseudouridylation.

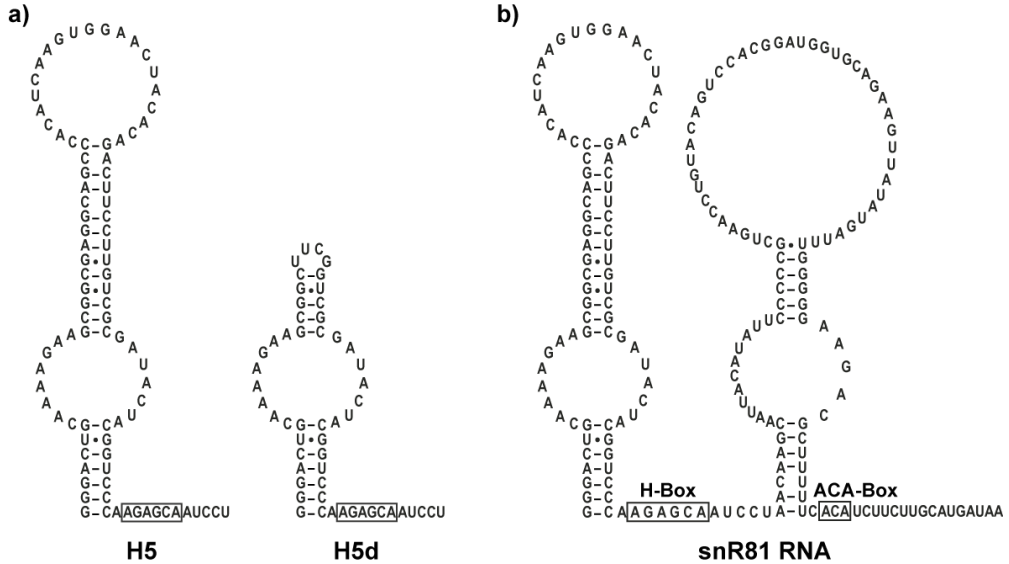


Figure 5.7: ***a)*** **Secondary structures of designed constructs H5 and H5d for investigation of the standalone 5'-hairpin.** ***b)*** **Native snR81 RNA secondary structure for comparison.**

Unmodified H5 and H5d RNAs were synthesized by *in vitro* transcription using T7 RNA polymerase and a PCR amplified DNA template. Purification of both constructs was achieved by denaturing PAGE, excision of the RNA bands from the gel and overnight elution. To identify whether these produced RNAs were successfully purified, an analytical denaturing gel was performed (see Figure 5.8a). Both constructs showed distinguishable RNA bands – H5 running higher in the gel than H5d –, but in case of H5d minor impurities were observable, indicating an overlay of RNAs with different sizes. Therefore, the purified H5d sample was analyzed by mass spectrometry. Due to its length and size H5 (29.2 kDa) could not be analyzed by this method. For H5d (expected mass 17.8 kDa) three major peaks with an m/z of 17818, 18124 and 18736 were detected (see Appendix A2 – Mass Spectrometry Data). Thus, the desired RNA was described but could not be separated from slightly larger RNAs (18124: +1 nt, 18736: +3 nt), which most likely stem from the ability of T7 RNA polymerase to randomly insert extra nucleotides at the 3'-end of the transcribed RNA during polymerization[127]. Known attempts to improve the homogeneity of RNA products are the usage of DMSO during transcription and 2'-O-methoxy primers during PCR amplification of the transcription template[138,139]. These adaptions, however, showed no positive effects on H5d transcripts. No further attempts for a defined H5d transcription were performed, since addition of a few nucleotides to the 3'-end is presumably negligible for RNA guided pseudouridylation.

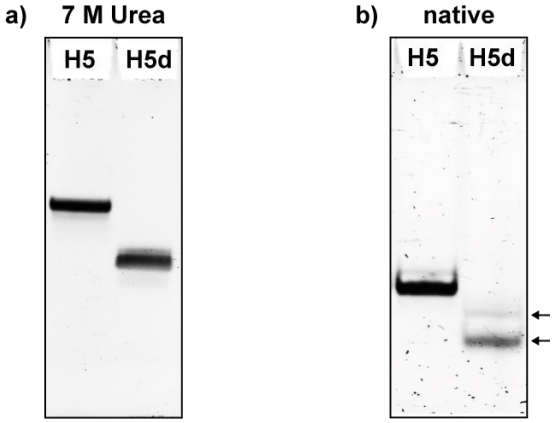


Figure 5.8: ***a)*** **Denaturing 10% PAGE of purified H5 and H5d.** ***b)*** **Native 8% PAGE of snap cooled H5 and H5d. The two appearing bands for H5d are indicated by arrows.**

While the presence of some extra nucleotides can be neglected, correct folding of the H/ACA-RNA is highly important for snoRNP formation as well as enzymatic activity. Therefore, the formation of secondary structures for H5 and H5d were tested by native PAGE. To promote correct folding the samples were heated to 98 °C and cooled down rapidly on ice. The idea

behind this approach is to completely denature secondary structures or potential multimers of the RNA at high temperatures in the first step, and in the second step kinetic formation of intramolecular base pairs is promoted by fast cool down to low temperatures. This way the development of the hairpin structure should be promoted over potential dimer formation or the like. Figure 5.8b shows a native gel of H5 and H5d after the above mentioned temperature treatment. For H5 one distinct band is visible, indicating one overall fold for the transcribed RNA. H5d shows two bands; one major band and one minor band running slightly higher. Presumably, this minor population arises from the inhomogeneous H5d transcription. The random extension of the 3'-end might lead to the formation of unexpected secondary structures, since the extra nucleotides might enable new base pairing possibilities within the RNA. An optimization by varying temperature treatment and buffer conditions showed no improvements. Hence, both constructs were always folded as mentioned above before performing further experiments.

5.2.2. Fluorophore Labeled Eukaryotic H/ACA-RNAs

In order to perform smFRET experiments labeling sites had to be selected for the RNA. Since there is no complete structural data of the snR81 RNP available, positions for chemoselective dye attachment were adapted from the preceding and successful investigation of the archaeal sRNA. For the 5'-hairpin constructs, functional groups were inserted in the lower stem at nucleotide G1 and in the upper stem above the pseudouridylation pocket at base U61 for H5, and respectively at base U26 for H5d. At G1 an aminohexyl group was introduced at the 5'-phosphate and at bases U61 or U26 an aminoallyl group was attached to the C5-atom for site specific RNA labeling (molecular structures of reactive amines are shown in Figure 4.10 on page 83). In the later stages of this thesis, investigation of the 3'-hairpin was also performed to compare the individual pseudouridylation domains of the bipartite snoRNP. For this purpose two 3'-constructs were designed in a similar fashion as for the 5'-hairpin: The complete 3'-hairpin with the full loop sequence 'H3' and a truncated version of the hairpin in which the large loop is also substituted by a tetra loop 'H3d'. For these constructs reactive groups for site specific labeling were inserted in the upper stem at base U21 – at this position C was mutated to U for labeling purposes – and in the lower stem at base U83 for H3 and respectively base U42 for H3d. For the modification sites of the 3'-hairpin constructs an aminohexylacrylamino group had to be introduced at the C5-position instead of the previously used aminoallyl group, since the latter was

no longer purchasable. The point mutation at position 21 (base 111 in the native snoRNA) was introduced because insertion of functional groups at the C5-atom of uridine proved to be reliable for RNA labeling previous constructs and is commonly used for smFRET experiments with RNAs[12,96,140]. Thereby, the canonical C-G Watson-Crick base pair was substituted by a non-canonical U-G pair. The influence of this mutation was estimated to be of low impact on snoRNA function, since it is flanked by two C-G pairs at each site, stabilizing the modified stem. All constructs and labeling sites are shown in Figure 5.9.

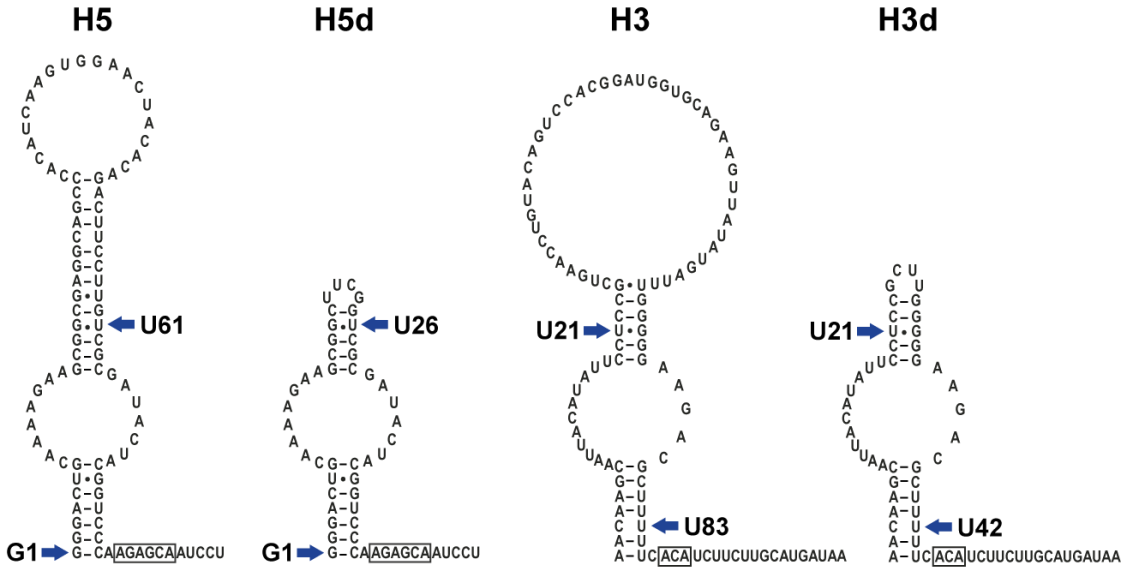


Figure 5.9: Secondary structures with highlighted labeling sites of 5'- and 3'-hairpin constructs H5, H5d, H3 and H3d.

In order to obtain fluorophore labeled snoRNAs, each construct was divided into three (H5, H5d and H3) or two oligomers (H3d) of varying lengths. They were designed in a construction-kit way where full-length H5 and H5d derived from the same 5'- and 3'-oligomers, but a different oligomer for the middle part (see Table 3.12, page 62). The same accounted for H3 and H3d constructs. All 3'-strands were modified with a biotin-linker for immobilization purpose. The purchased and chemically synthesized short RNA strands carried one or no functional group for fluorophore attachment and were protected by ACE chemistry. Preparation of labeled eukaryotic H/ACA-RNAs was divided into three steps: 1) fluorophore labeling and deprotection of the individual oligomers, 2) HPLC purification of the coupling reaction, and 3) enzymatic splinted ligation of modified and unmodified RNA strands to the final construct. Step 1) and step 2) were performed likewise to the archaeal constructs; by using reactive NHS-esters for coupling to the functional primary amines, and subsequent HPLC purification of the reaction mix. For the chromatographic purification of the eukaryotic RNA strands an mixture of 50% MeCN and 50%

aqueous 100 mM TEAA (pH 7) was used for elution, instead of pure MeCN (as was done for the archaeal constructs). Representative chromatograms are shown in Figure 5.10.

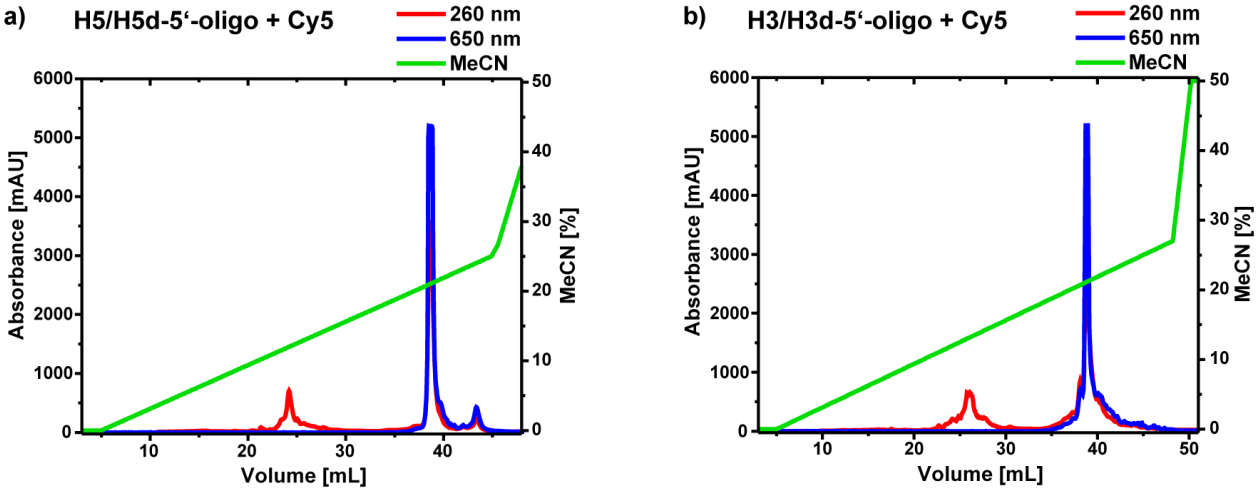


Figure 5.10: ***a)*** **HPLC chromatogram of the labeling reaction of the 5'-oligos for H5 and H5d to Cy3.** ***b)*** **HPLC chromatogram of the labeling reaction of the 5'-oligos for H3 and H3d to Cy5.**

In both chromatograms a clear separation of unlabeled from labeled oligomers is discernable. Elution of dye-modified RNAs were observed at 20% MeCN for H5/H5d-5'-oligo-Cy5 (see Figure 5.10a) and at 22% MeCN for H3/H3d-5'-oligo-Cy5 (see Figure 5.10b). The use of a MeCN/TEAA mixture instead of pure MeCN improved the separation of labeled and unlabeled RNAs significantly (compare Figure 4.11 on page 84), and was therefore applied in the HPLC purifications of the other eukaryotic constructs as well. Fractions showing high UV absorption of dye (650 nm) and RNA (260 nm) were pooled and used for further RNA preparation. Oligonucleotides that were not labeled with fluorophores were directly deprotected, ethanol precipitated and resuspended in water.

After successful purification of modified oligomers the final hairpin was to be obtained by splinted ligation and subsequent gel purification (step 3). Ligation templates were designed individually for each construct (H5, H5d, H3 and H3d). Annealing of these nucleic acids was done with equimolar amounts of respective RNA oligomers and the according DNA-splint for the corresponding hairpin construct at 85 °C for 5 minutes and slow cool down to room temperature. Afterwards the ligases were added and the mixture was incubated for 2.5 hours at 37 °C. For H5 and H5d a mixture of NEB and ThermoFisher T4-DNA-ligases was sufficient for good ligation yields. For H3 however, homemade T4-RNA-ligase 2 had to be used to gain the complete construct. Otherwise, no efficient ligation of the 3'-oligomer to the other RNA strands of H3

could be detected. In order to check if the formation of RNA-DNA-hybrids and if the enzymatic connection of RNA oligomers occurred specifically several ligation tests were carried out, in which single nucleic acid strands were omitted from the reaction mixture (see Figure 5.10). Omission of individual RNA strands was done for all constructs consisting of three different oligomers (H5, H5d and H3).

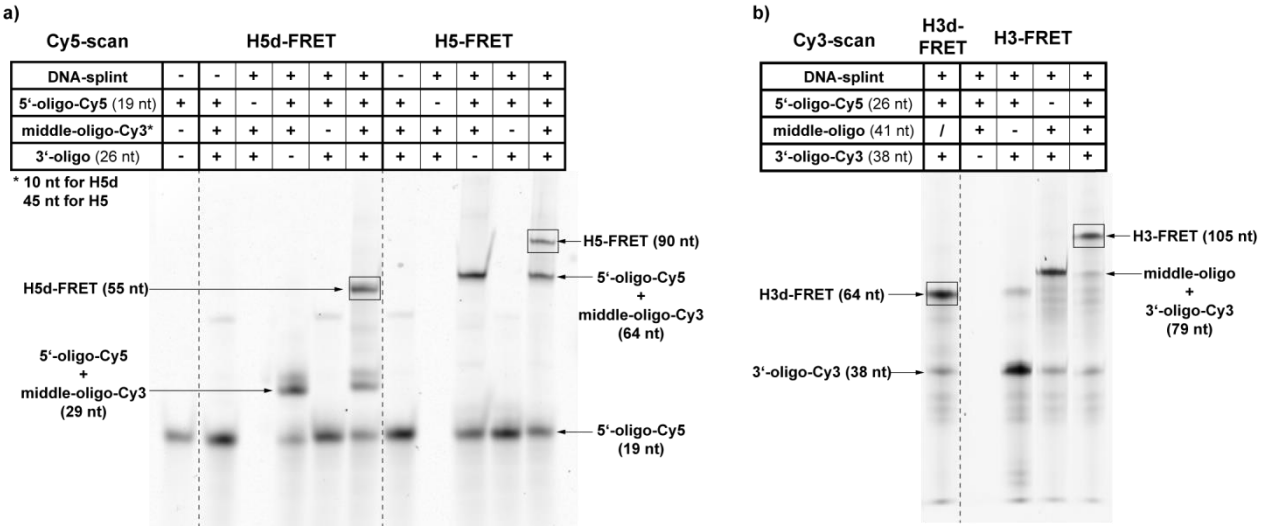


Figure 5.11: ***a)*** **15% denaturing PAGE of the test ligations for H5d- and H5-FRET (Typhoon settings: Cy5-scan). Ligation and PAGE were performed by David Mentrup in the scope of a bachelor thesis.** ***b)*** **10% denaturing PAGE of the test ligations for H3d- and H3-FRET (Typhoon settings: Cy3-scan). The fully ligated product bands are framed.**

The analytical gels show that only in presence of all required oligomers, as well as the DNA-splint, the enzymatic ligation yields the desired products. In the absence of the middle-oligomer in case of *H5*-, *H5d*- and *H3-FRET* no ligation bands were detected. Furthermore, no major amounts of unspecifically ligated side products were observed. Thus, annealing of the RNA-oligonucleotides to the DNA-splint and subsequent ligation were specific and successful. The test reactions were performed in small, analytical scale (10-50 pmol) and scaled up (1000 pmol) for preparative ligations. Yields of these preparative approaches after the final gel purification were typically around 10%. Constructs prepared for activity and/or FRET-experiments are shown in Figure 5.12.

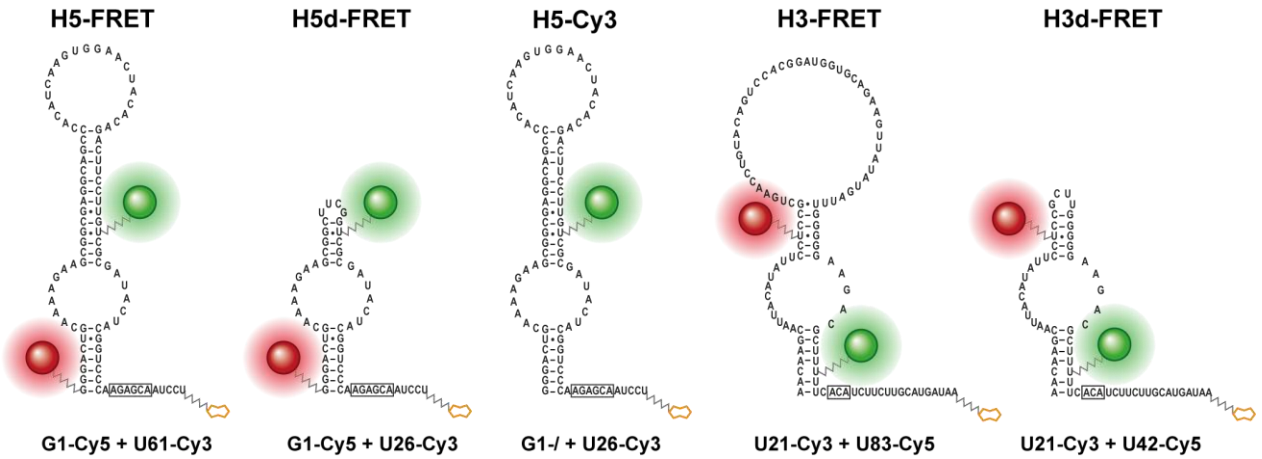


Figure 5.12: Secondary structure of synthesized yeast H/ACA-constructs for smFRET experiments. RNA-labels are marked in different colors: Cy3 (green), Cy5 (red) and biotin (orange). Linkers are indicated as jagged lines.

As in the archaeal project, for investigation of pseudouridylation in eukaryotic H/ACA-RNPs substrate analogues (X-scSub), carrying one of three different nucleotides (X: U, 5fU or Ψ) at the target site, were used for certain experiments. These target RNAs (24mers) were purchased and either labeled at the 3'-end with Cy5 (X-scSub-Cy5), or directly deprotected and used without any further modification. Labeling and purification of eukaryotic target RNAs were performed in the same way as described above.

5.3. snoRNP Reconstitution

5.3.1. EMSAs of Unlabeled H5 and H5d

In order to investigate whether the produced H/ACA-proteins and snoRNA constructs are capable of complex formation EMSAs were performed. Binding affinity of full-length and truncated 5'-hairpins was compared for the unlabeled (H5 and H5d) and labeled (*H5-FRET* and *H5d-FRET*) RNAs. Thus, a binding assay was conducted for each construct with either scNCG or Nhp2. For each measurement row the RNA concentration was held constant, while the protein concentration increased. For the unlabeled RNAs the concentration was adjusted to 2 µM each, while the protein concentrations ranged from 0 – 8 µM for Nhp2, and 0 – 16 µM for scNCG. All samples were annealed at 30 °C and analyzed by native PAGE.

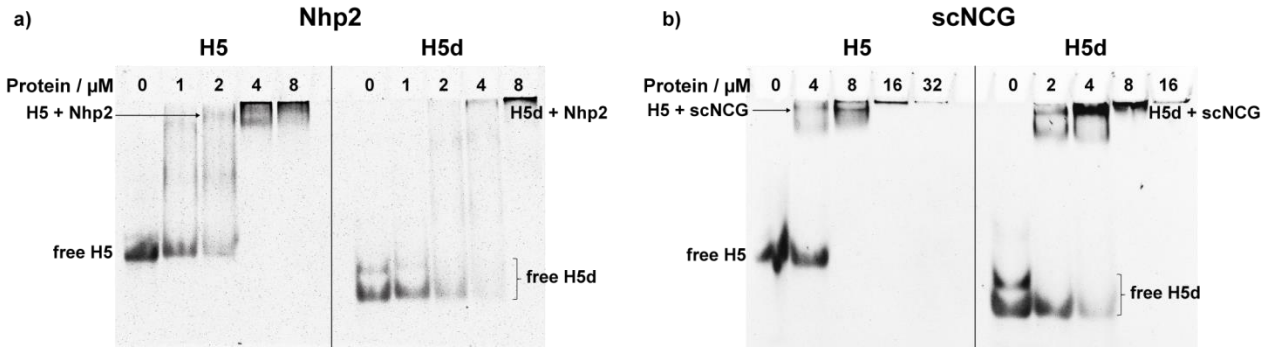


Figure 5.13: ***a)*** **Native 8% PAGE of 2 µM H5 (left) or H5d (right) with increasing Nhp2 concentration.** ***b)*** **Native 8% PAGE of 2 µM H5 (left) or H5d (right) with increasing scNCG concentration. Annealing was done in 0.8x scSEC-buffer. RNAs were stained with GelRed. The EMSAs were performed by David Mentrup in the scope of a bachelor thesis.**

The EMSAs show that Nhp2 and scNCG are able to bind the individual hairpins. With increasing protein concentrations the formation of RNA-protein bands is visible. However, for none of these constructs a clear and distinct RNP band migrated into the gels. For higher protein concentrations the observed RNA-protein complex did not migrate into the gel, pointing towards to possible aggregation under gel electrophoretic conditions. Nevertheless, some trends from the EMSAs are discernible. In case of Nhp2-binding, no free H5 band was observable at a two-fold excess of protein, while for the shortened H5d minor amounts of unbound RNA remained visible at a 4 µM protein concentration (see Figure 5.13a). This indicates a slightly higher affinity of Nhp2 to the complete hairpin than to the truncated RNA. Furthermore, some minor blurred bands in between free RNA and RNP are visible in the native gels. They might be due to unspecific Nhp2-RNA binding, since this protein is known to unspecifically bind nucleic acids. However, clear statements about H5/H5d-Nhp2 binding modes cannot be made from these EMSAs.

For scNCG a different behavior could be detected. Binding of scNCG to H5d was observably more efficient than to H5 at a two-fold excess (4 µM) of the ternary protein complex over the guide RNAs (see Figure 5.13b), indicated by the more dominant free H5 band at 4 µM protein concentration. Thus, truncation of the loop and upper stem of the 5'-hairpin does not interfere with scNCG binding. Here again, no distinct statements about RNA-protein binding specificity can be made. From the EMSAs it is not discernible how many proteins are attached to the RNA and at which position binding occurs. Thus, these experiments were not repeated for H3 and H3d. Nevertheless, binding capability of the labeled 5'-hairpin constructs was further investigated.

5.3.2. EMSAs of Labeled H5- and H5d-FRET

To verify RNP-formation of *H5-FRET* and *H5d-FRET* a binding assay with Nhp2 and scNCG was performed for each construct. The concentration of the guide RNA was reduced to 100 nM and remained constant in each sample. Protein concentrations ranged from 0 – 4 µM for scNCG, and from 0 – 8 µM for Nhp2. Complex annealing was done at 30 °C for 15 min.

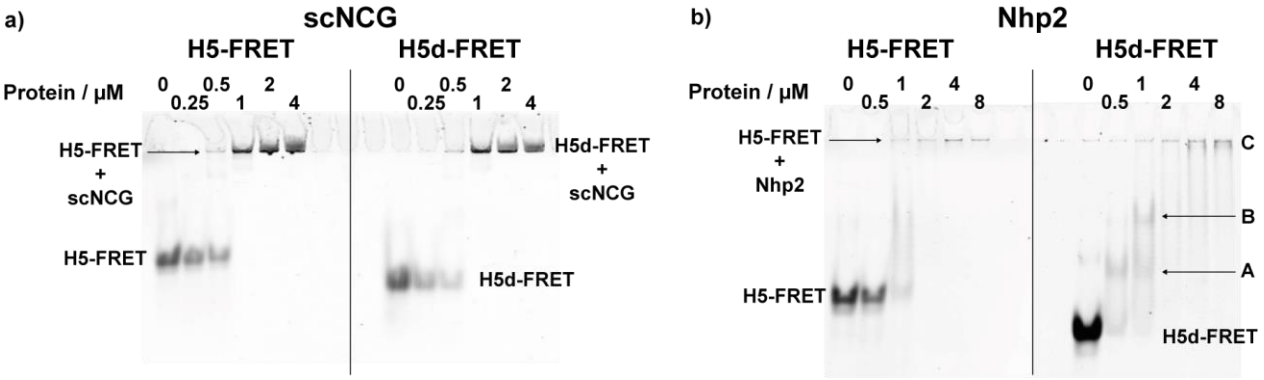


Figure 5.14: ***a)*** **Native 8% PAGE of 100 nM H5 (left) or H5d (right) with increasing scNCG concentration (Typhoon settings: Cy5-scan).** ***b)*** **Native 8% PAGE of 100 nM H5 (left) or H5d (right) with increasing Nhp2 concentration. Annealing was done in scΨ-buffer (Typhoon settings: Cy5-scan).**

In the EMSA of *H5-FRET* and *H5d-FRET* with the ternary scNCG subcomplex RNP formation can be detected for both constructs (see Figure 5.14a). The full and the truncated hairpin showed no significant difference in binding affinity. For both samples the free RNA bands disappeared at a 1 µM scNCG concentration. Here again, the RNP is stuck in the top of the gel. Thus, no statements about binding specificity can be made. For Nhp2, a different behavior in RNA-protein binding could be observed for the two FRET constructs (see Figure 5.14b). In case of *H5-FRET* the free RNA band completely vanished at 2 µM Nhp2 concentration and with increasing protein concentration the RNA-protein remained in the loading pocket of the gel. For *H5d-FRET* the free

RNA band disappeared almost completely at 0.5 µM Nhp2 concentration. But more strikingly, new distinct RNP bands in the gel appeared (termed A and B) at 0.5 µM and 1 µM protein concentration. With increasing protein concentration up to 8 µM bands A and B vanished, and only one band in the loading pocket (C) remained. Most likely, bands A and B represent an unspecific binding mode of Nhp2 to H5d, which may be easier to distinguish with fluorescently labeled RNAs. For *H5-FRET* these bands are not as distinct, which indicates that binding of Nhp2 to H5 is more specific. However, clear statements about binding region and specificity of scNCG and Nhp2 cannot be made from these EMSAs. Nevertheless, it could be shown, that the labeled guide RNAs are completely bound at a 10-fold excess of scNCG. Due to these results and to the fact that the binding affinity of Nhp2 to the H/ACA-RNP is significantly increased in presence of the other proteins, complex annealing in the further experiments was performed with a 10-fold excess of proteins over RNA.

5.4. Enzymatic Activity of Eukaryotic H/ACA-RNPs

5.4.1. Single Turnover Assays

Catalytic functionality of the designed and synthesized guide RNAs was tested in initial single turnover pseudouridylation assays. Labeled and unlabeled H/ACA-RNAS were compared directly. That way, the effect of dye incorporation on enzymatic activity of the RNPs was to be investigated. Substrates (5'-Sub and 3'-Sub) for the eukaryotic H/ACA-RNP were designed for each hairpin individually. Target RNAs were expected to form base pairs at all possible sites within the pseudouridylation pocket (according to the natural substrate), containing overhangs of a few nucleotides at every end. The assays were performed with site specifically radioactively labeled target RNAs. The H/ACA-RNAs were folded by snap cooling prior to sample preparation. Then, equimolar amounts (0.1 µM) of guide RNA and respective substrate (5'-Sub or 3'-Sub) were incubated in presence of all four proteins WNCG (1.0 µM) in scΨ-buffer for one hour at moderate temperatures (30 °C), and purified by phenol extraction. The RNAs were digested into mononucleotides for TLC and visualized by autoradiography, so that only the ^{32}P-labeled nucleotides could be detected. Figure 5.15 shows the summary of the activity experiments for the labeled and unlabeled standalone hairpins.

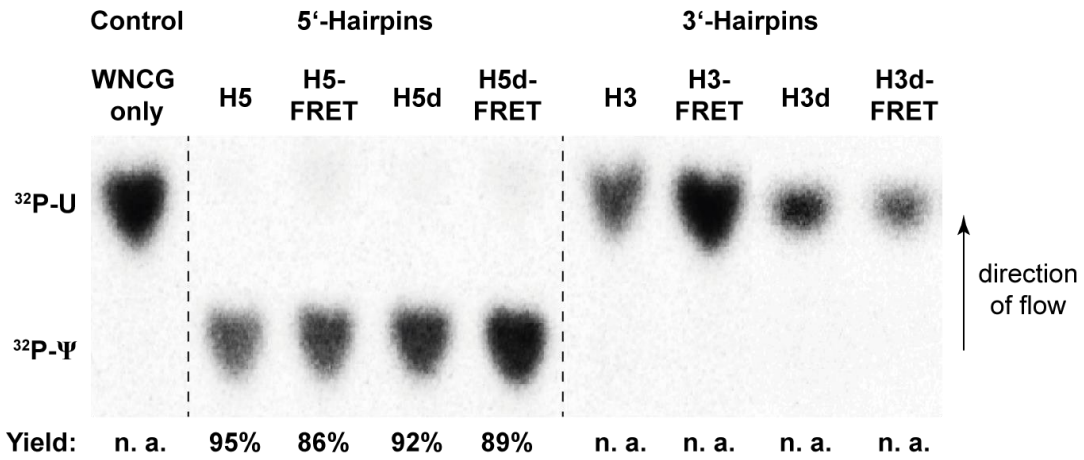


Figure 5.15: TLC of summarized single turnover activity assays for unlabeled and labeled standalone hairpins from the eukaryotic snR81 H/ACA-RNP (Typhoon settings: Phosphor screen).

From the single turnover experiments one can clearly see that all 5'-hairpin constructs (H5, H5d, *H5*- and *H5d-FRET*) show near to quantitative pseudouridylation (> 85%) after one hour of incubation. No significant deviations in activity between the native (H5) or truncated (H5d) hairpin could be observed in these assays. Both constructs showed over 90% Ψ-formation. The

labeled constructs showed slightly less activity than their unlabeled counterparts, but were still highly active with 86% and 89% for *H5*- and *H5d-FRET*. On the other hand, for the standalone 3'-hairpins no apparent isomerization activity (n. a.) was detected, neither for the full-length (H3) nor for the truncated (H3d) hairpin. The labeled *H3*- and *H3d-FRET* constructs showed no pseudouridylation as well. The control sample – missing the guide RNA – showed that the proteins do not inherit any catalytic activity on their own.

5.4.2. Multiple Turnover Assays

After enzymatic activity of the synthesized H/ACA-RNPs could be verified for the 5'-hairpin, multiple turnover experiments were performed to gain information about the kinetics of eukaryotic Ψ-formation. Impacts of individual proteins and attached fluorophores on pseudouridylation rates were to be studied. H5 and H5d were tested in presence and absence of Nhp2, to investigate how the upper stem of the snoRNA, and the protein that is assumed to bind this stem, affects pseudouridylation. scGar1 was also removed in a separate measurement row to study its influence. Additionally, a target RNA containing the anti-cancer drug 5fU at the target side (5fU-scSub) – known to interfere with Ψase activity[131-134] – was added in some experiments, to see how this nucleotide affects pseudouridylation kinetics. Furthermore, multiple turnover assays were performed at different temperatures (30 °C, 20 °C and on ice) with labeled and unlabeled snoRNAs, to see if eukaryotic uridine isomerization can generally be investigated by smFRET spectroscopy at moderate conditions.

Multiple turnover assays were performed with a 40-fold excess of substrate RNA over the RNP. Sample preparation was done in the same manner as for the single turnover experiments. Labeled or unlabeled guide RNAs (0.1 µM) were assembled in the presence of the respective proteins (1.0 µM). Radioactive substrate (5'-Sub; 0.1 µM) and non-radioactive target RNAs (U-scSub or 5fU-scSub; 3.9 µM) were added just before incubation. Samples of each measurement row were taken after 5, 20, 60, and 240 min and directly inactivated. Of each sample the RNAs were digested into mononucleotides for TLC and visualized by autoradiography. To gain kinetic information, the data points of each experiment were fitted to a Michaelis-Menten equation (see Figure 4.21b, page 98). For comparison of different constructs, in each fit y_{max} was set to 0.95, the detected maximum yield of the previously performed single turnover assay for the H5 snoRNP. For the unlabeled RNAs H5 and H5d two separate measurement rows at 30 °C were

performed in presence of WNCG and scNCG in order to obtain mean values and for statistical verification. In Figure 5.16 the above described experiments are summarized.

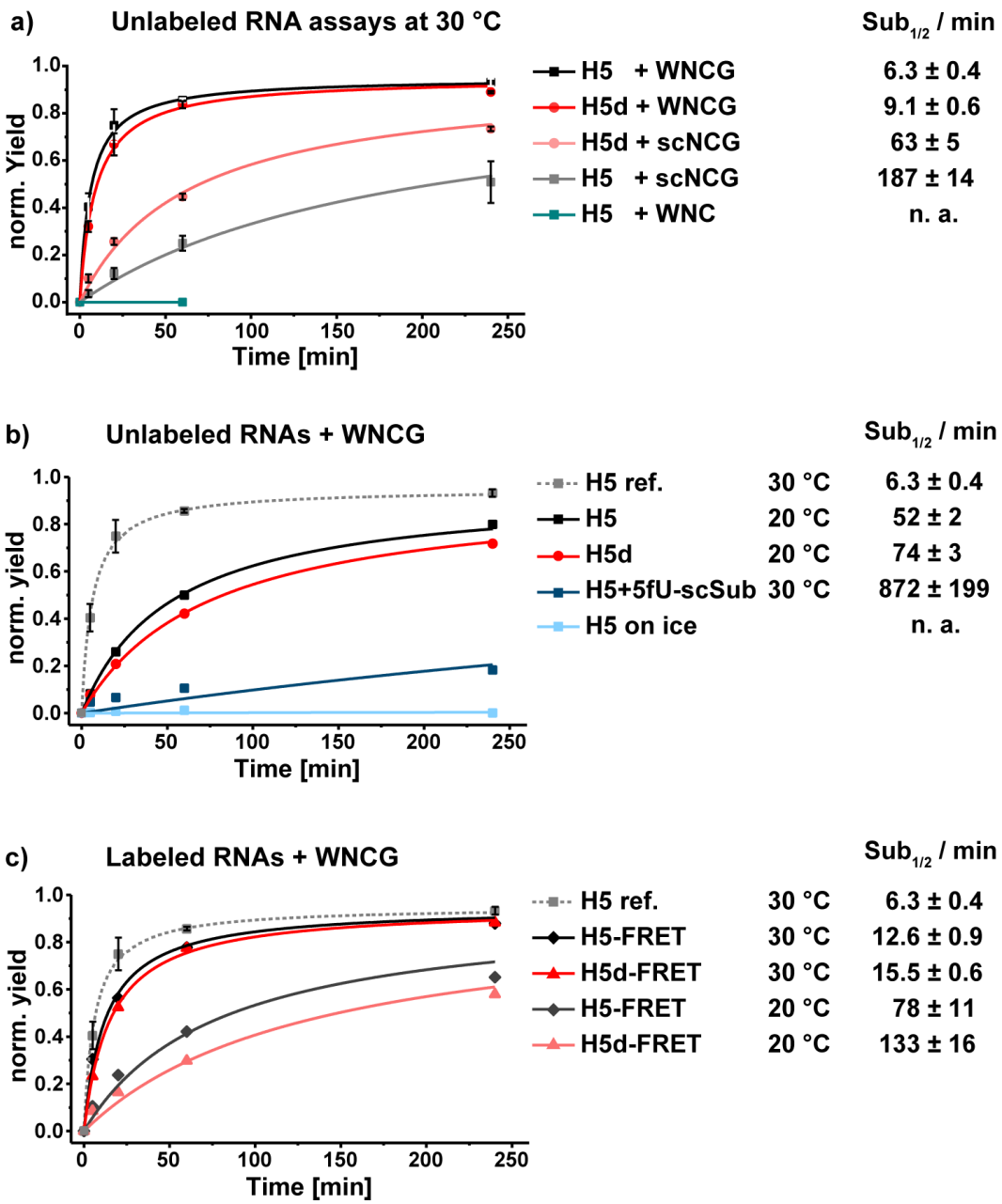


Figure 5.16: Time dependent pseudouridylation yields of multiple turnover assays for various 5'-hairpin constructs from the snR81 RNP. Errors for samples with two data sets were obtained by standard deviation. *a)* Assays of unlabeled guide RNAs with varying proteins at 30 °C. *b)* Assays of unlabeled guide RNAs with WNCG at varying temperatures. *c)* Assays of FRET-labeled guide RNAs with WNCG.

Figure 5.16 displays the multiple turnover experiments for full (WNCG) or partial (scNCG/WNC) H/ACA-RNPs with unlabeled H5 and H5d at 30 °C. The complete 5'-hairpin (H5+WNCG) exhibits the fastest turnover rate with a substrate half-life time of $Sub_{1/2}$ = 6.3 ± 0.4 min. With a $Sub_{1/2}$ = 9.1 ± 0.6 min for H5d+WNCG, the truncated construct is almost as active as

the native hairpin. Removal of Nhp2 led to significant reductions of pseudouridylation speed. In case of H5+scNCG, a 30-fold decrease in the reaction rate ($Sub_{1/2}$ = 187 ± 14 min) was observed, while for H5d+scNCG the reaction rate was only decreased by a factor of seven ($Sub_{1/2}$ = 63 ± 5 min), when the upper stem binding protein is missing. Thus, omission of Nhp2 has a bigger influence on enzymatic activity for the complete 5'-hairpin than for the truncated hairpin. In absence of scGar1 (H5+WNC), the eukaryotic H/ACA-RNP experienced no activity at all.

The effects on pseudouridylation speed for different incubation temperatures (20 °C and on ice), as well as for addition of 5fU-scSub instead of U-scSub (at 30 °C), are shown in Figure 5.16b. The sample H5+WNCG at 30 °C (see Figure 5.16a) was added to this plot as reference. Reduction of the reaction temperature down to 20 °C led to a decrease in turnover rate by a factor of eight for both constructs: H5+WNCG ($Sub_{1/2}$ = 52 ± 2 min) and H5d+WNCG ($Sub_{1/2}$ = 74 ± 3 min), compared to incubation at 30 °C. A further drop of the reaction temperature down to ca. 4 °C led to a complete loss of enzymatic activity. Addition of modified target RNA 5fU-scSub instead of the substrate U-scSub showed resulted in a major loss in pseudouridylation (~30-fold decrease), but showed ongoing catalytic activity ($Sub_{1/2}$ = 872 ± 199 min).

Figure 5.16c displays the multiple turnover experiments for fluorophore labeled 5'-hairpin constructs in presence of all four proteins at 30 °C and 20 °C. From the graph it is apparent that the reactivity of *H5-FRET* ($Sub_{1/2}$ = 12.6 ± 0.9 min) and *H5d-FRET* ($Sub_{1/2}$ = 15.5 ± 0.6 min) at 30 °C is almost as high as for their unlabeled counterparts; the turnover rate decreased only by a factor of two for H5 ($Sub_{1/2}$ = 6.3 ± 0.4 min) and by a factor of 1.5 for H5d ($Sub_{1/2}$ = 9.1 ± 0.6 min). At 20 °C reaction temperature the activity of both constructs dropped notably; by a factor of six for *H5-FRET* ($Sub_{1/2}$ = 78 ± 11 min) and a factor of nine for *H5d-FRET* ($Sub_{1/2}$ = 133 ± 16 min) compared to the higher incubation temperature. Nevertheless, the tested constructs inhibited high catalytic activity in a multiple turnover manner at room temperature, and are, thus, applicable for observation of RNA guided pseudouridylation by smFRET.

5.5. smFRET Analysis of Eukaryotic H/ACA-RNPs

5.5.1. General Considerations

Before the results of the smFRET measurements of the eukaryotic H/ACA-RNPs will be shown and discussed, some general aspects about the execution of these experiments will be mentioned. Prior to every measurement, RNAs and proteins were annealed in a 10-fold excess of protein over RNA, according to the experimental results of the EMSA. Afterwards, the samples were stored on ice to prevent undesired processes, in particular complex dissociation. Particle immobilization was achieved in a manner similar to the archaeal H/ACA-RNP. Biotinylated constructs were pulled down on a passivated glass surface via the highly stable biotin-streptavidin interaction. Just before the measurement, samples were diluted in order to prevent an overloading of the channel. The guide RNA was diluted to 0.5 nM in scΨ-buffer, immobilized on the glass slide and particle density was checked before the oxygen-scavenging system was added. The actual measurements were performed at room temperature, at which partial eukaryotic snoRNPs proved to be catalytically active (see Chapter 5.4.2), unlike the archaeal Ψase. Short (2 seconds) and long movies (120/200 seconds) were recorded of each sample, in order to plot apparent states in histograms and observe potential dynamics in traces of single molecules. Evaluation of the histograms was done using Gaussian distribution. States were fitted individually, and, in order to allow comparison of different samples within a measurement row, the peak width ‘w’ was fixed for individual states and samples. If necessary, donor-only populations were removed by Gaussian fitting.

5.5.2. Eukaryotic H/ACA-RNP Assembly

5’-Hairpin Constructs

Initial assembly experiments of the eukaryotic H/ACA-RNP were performed with the 5’-hairpin constructs; *H5*- and *H5d-FRET*. Fluorophores were attached at the 5’-end (G1) and slightly above the pseudouridylation pocket (U61 for *H5-FRET* and U26 for *H5d-FRET*) of each guide RNA (see Figure 5.9, page 137). With these constructs, structural changes between the lower and the upper stem of the partial snoRNAs can be detected. In order to compare RNP formation of the eukaryotic to the archaeal complex, the stepwise assembly experiments were repeated. Therefore, the FRET-RNAs were annealed in absence of any protein (RNA), in

presence of only Nhp2 (RNA+W), in presence of the ternary scNCG complex (RNA+scNCG), or in presence of all four proteins (RNA+WNCG).

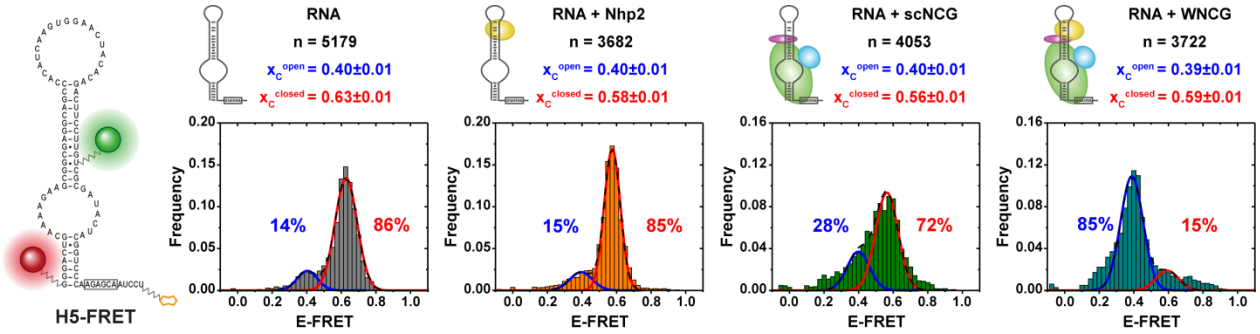


Figure 5.17: smFRET histograms of H5-FRET assembly. Histograms of individual complexes are marked by color: free RNA (grey), RNA plus Nhp2 (orange), RNA plus scNop10-scCbf5-scGar1 (green), and RNA plus Nhp2-scNop10-scCbf5-scGar1 (teal). The absolute number of detected molecules for E-FRET ≥ 0.1 is given as 'n'. 'x_C^{open}' represents the peak center of the open state (blue) and 'x_C^{closed}' represents the peak center of the closed state (red). For proper fitting, the peak width of each Gaussian fit was set to w = 0.13. Errors result from fitting. Donor only peaks were removed by fitting.

Figure 5.17 summarizes the results of RNP-assembly experiment for *H5-FRET*. Overall, two distinct populations were observed in this experiment; a lower FRET population at E-FRET = 0.40, termed 'open state', and a higher FRET population at E-FRET = 0.59, termed 'closed state'. For the free 5'-hairpin and when Nhp2 was added to the labeled construct, the closed state was the preferred conformation (≥ 85%) of the partial snoRNA. Addition of scNCG led to a minor increase of molecules occupying the lower FRET state (up to 28%). Only in presence of all four proteins an almost quantitative shift from the higher to the lower FRET state was observed (85% of detected molecules in the open conformation). This overall decrease in FRET-efficiency – when all four proteins are present – indicates an increase in the intramolecular dye distance upon snoRNP formation.

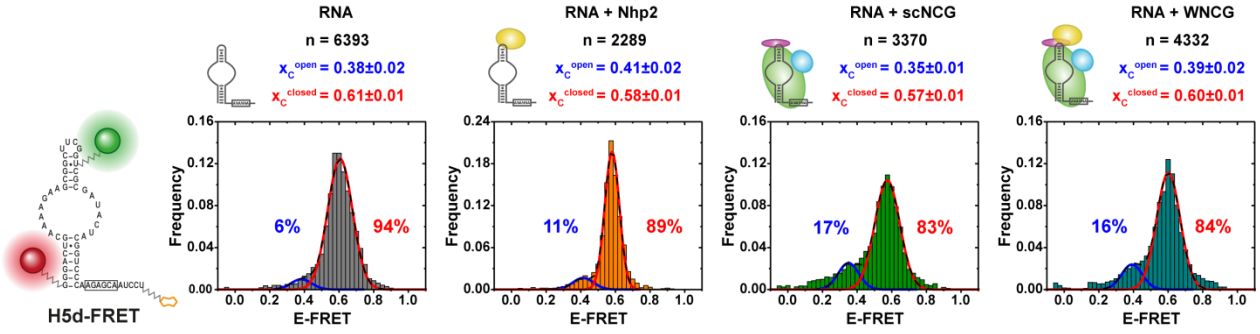


Figure 5.18: smFRET histograms of H5d-FRET assembly. Histograms of individual complexes are marked by color: free RNA (grey), RNA plus Nhp2 (orange), RNA plus scNop10-scCbf5-scGar1 (green), and RNA plus Nhp2-scNop10-scCbf5-scGar1 (teal). The absolute number of detected molecules for E-FRET ≥ 0.1 is given as 'n'. 'x_C^{open}' represents the peak center of the open state (blue) and 'x_C^{closed}' represents the peak center of the closed state (red). For proper fitting, the peak width of the open state was set to w = 0.12. Errors result from fitting. Donor only peaks were removed by fitting.

In Figure 5.18 the stepwise assembly experiment for *H5d-FRET* is shown. Here again, two different FRET populations were observed; an open state at E-FRET = 0.38 and a closed state at E-FRET = 0.59. In absence of any proteins the truncated 5'-hairpin almost quantitatively inherits the closed state (94%). Neither addition of Nhp2, scNCG nor all four proteins (WNCG) showed significant changes in the observed E-FRET pattern. Even though, a minor increase of molecules populating the open state could be observed for scNCG (17%) and WNCG (16%), the closed state remained the predominant population. Thus, no major structural rearrangements of the labels could be observed for *H5d-FRET* upon complex assembly.

In order to check for potential transitions between the open and the closed state, traces (2 min videos) of single molecules were evaluated. Representative traces of each state are of the sample *H5-FRET*+WNCG are given in Figure 5.19. Both trajectories show no transitions between the two states while both fluorophores are emitting light. Hence, no conformational changes between the two labels were observed during the time of measurement.

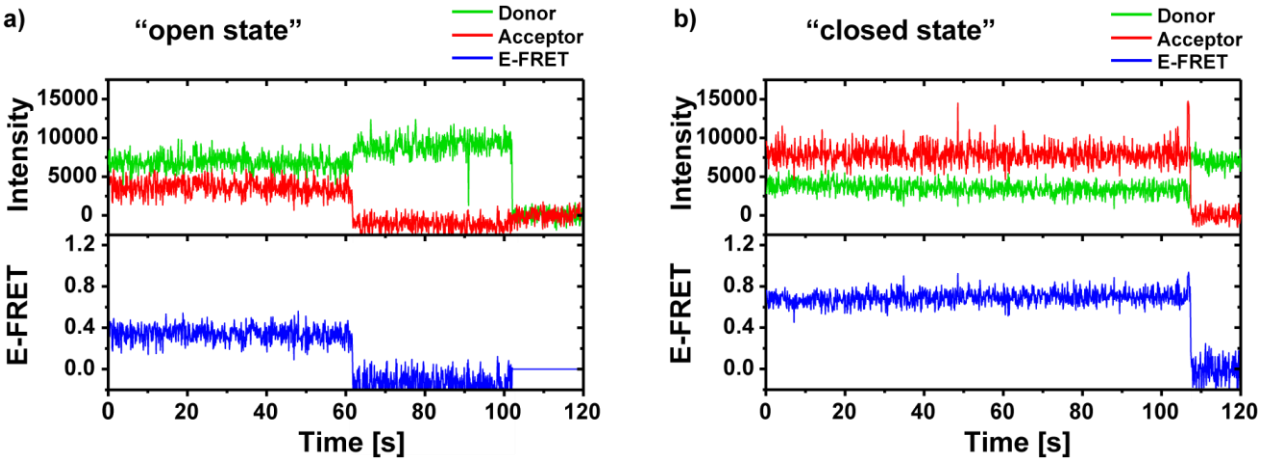


Figure 5.19: ***a)*** **Exemplary trace of the open state (E-FRET = 0.40) from H5-FRET+WNCG.** ***b)*** **Exemplary trace of the closed state (E-FRET = 0.59) from H5-FRET+WNCG.**

3'-hairpin constructs

The stepwise assembly experiment was repeated for the synthesized 3'-hairpin constructs *H3-FRET* and *H3d-FRET*. In regard to the 5'-hairpin constructs the fluorophore positioning differed. Labels were attached to the point mutated base U21 in the upper stem and to the unpaired base U83 in *H3-FRET*, or respectively U42 in *H3d-FRET*, in the lower stem (see Figure 5.9, page 137). With these labeling sites structural changes within the two stems, flanking the pseudouridylation pocket of the 3'-hairpin, should be detectable.

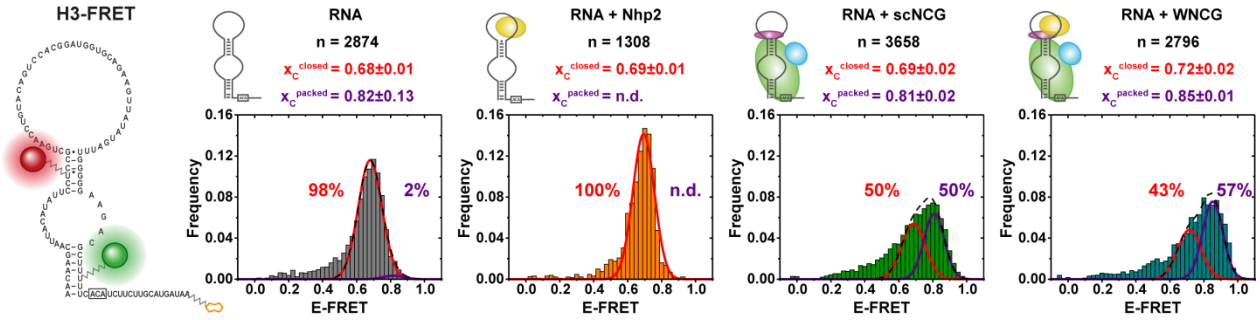


Figure 5.20: smFRET histograms of H3-FRET assembly. Histograms of individual complexes are marked by color: free RNA (grey), RNA plus Nhp2 (orange), RNA plus scNop10-scCbf5-scGar1 (green), and RNA plus Nhp2-scNop10-scCbf5-scGar1 (teal). The absolute number of detected molecules for E-FRET ≥ 0.1 is given as 'n'. 'x_C^{closed}' represents the peak center of the closed state (red) and 'x_C^{packed}' represents the peak center of the packed state (purple). For proper fitting, the peak width of the closed state was set to w = 0.14, and for the packed state set to w = 0.12. Errors result from fitting. Donor only peaks were removed by fitting.

As for *H5-FRET*, two different FRET populations were observed in the step-by-step assembly for *H3-FRET*; a 'closed state' at E-FRET = 0.69 and a 'packed state' at E-FRET = 0.82 (see Figure 5.20). For proper fitting, the peak widths were fixed differently for the two states: w = 0.14 for the closed state, and w = 0.12 for the packed state. The width of the packed state was adjusted to a lower value in order to compensate the loss of distance sensitivity, when reaching higher FRET-efficiencies. For the sample 'RNA + Nhp2' a packed state could not be defined. The free RNA as well as in presence of Nhp2 one major population – the closed state – was observed. Addition of scNCG or WNCG led to the formation of the packed states with 50% or 57% occupation, respectively. This increase in E-FRET suggests a compaction of the standalone 3'-hairpin upon RNP assembly.

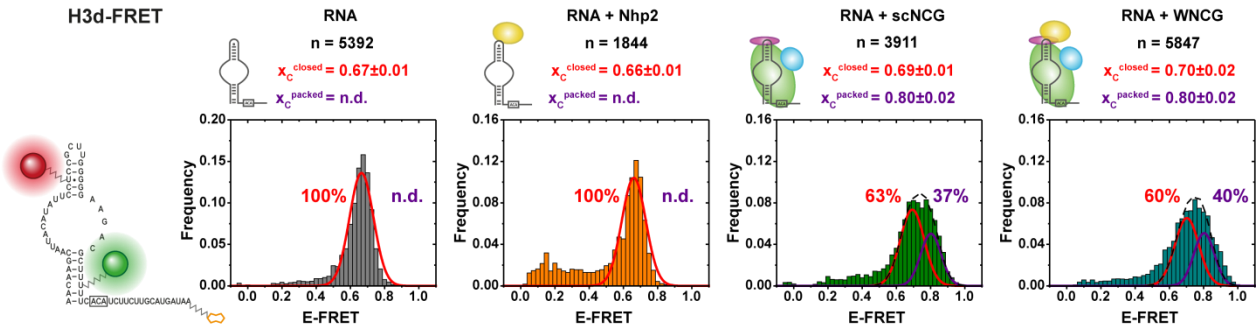


Figure 5.21: smFRET histograms of H3d-FRET assembly. Histograms of individual complexes are marked by color: free RNA (grey), RNA plus Nhp2 (orange), RNA plus scNop10-scCbf5-scGar1 (green), and RNA plus Nhp2-scNop10-scCbf5-scGar1 (teal). The absolute number of detected molecules for E-FRET ≥ 0.1 is given as 'n'. 'x_C^{closed}' represents the peak center of the closed state (red) and 'x_C^{packed}' represents the peak center of the packed state (purple). For proper fitting, the peak width of the closed state was set to w = 0.14, and for the packed state set to w = 0.12. Errors result from fitting. Donor only peaks were removed by fitting.

In the assembly experiment of *H3d-FRET* a similar behavior as for *H3-FRET* was observed. The same states were distinguished; a 'closed state' at E-FRET = 0.69 and a 'packed state' at E-FRET = 0.82 (see Figure 5.21). Peak fitting was performed as for the previous measurements in

order to compare results. For the free RNA and for RNA+Nhp2 no molecules in the packed state were observed. Only in presence of the ternary protein complex (scNCG or WNCG) a minor population for the high-FRET state could be observed; 37% for scNCG and 40% for the full complex. As for *H3-FRET*, this increase indicates a further compaction of the pseudouridylation pocket upon scNCG binding.

As for the previous constructs, single molecules movies of different samples (*H3-FRET* and *H3-FRET*+WNCG) were evaluated to identify possible transitions between the closed and packed state. Representative traces are shown in Figure 5.22. For the free RNA the dye intensities remained constant until the acceptor bleached (see Figure 5.22a). In the case of the protein-bound construct, larger fluctuations in fluorescence intensity could be observed, especially for the acceptor (see Figure 5.22b). However, these variations in acceptor intensity appeared irregularly and are most likely associated to biophysical artefacts of the fluorophore, which may occur due to short intramolecular label distance or interactions to surrounding proteins. Nevertheless, no statistically relevant transitions were observed.

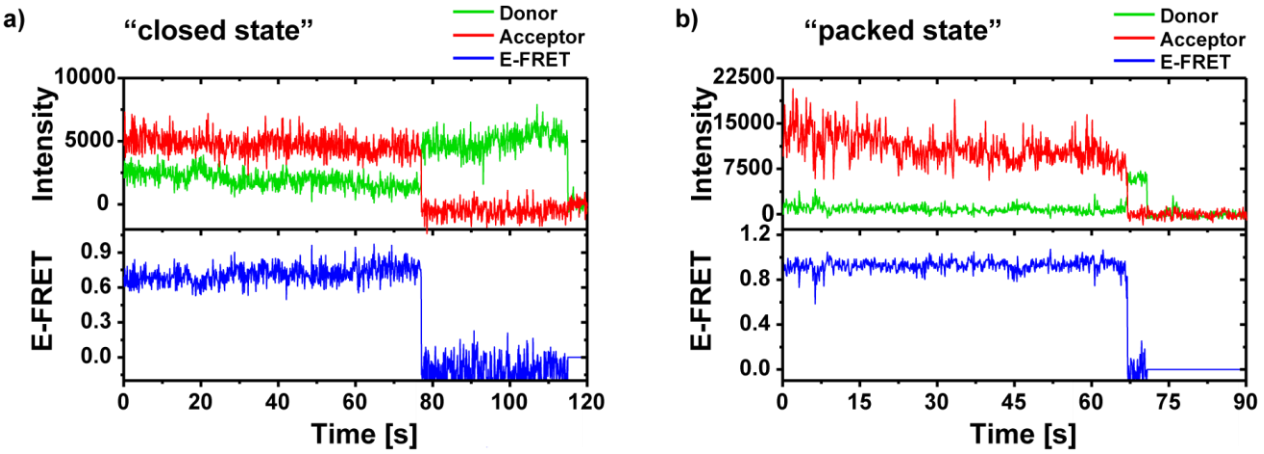


Figure 5.22: ***a)*** **Exemplary trace of the closed state (E-FRET = 0.69) from H3-FRET.** ***b)*** **Exemplary trace of the packed state (E-FRET = 0.82) from H3-FRET+WNCG.**

5.5.3. Substrate Turnover

In multiple turnover assays it could be observed that the synthesized constructs *H5*- and *H5d-FRET* are capable of efficient pseudouridylation at ambient temperatures, at which single molecule experiments are typically performed. Hence, conformational changes between the 5'-end and the upper stem of the guide RNA during substrate loading, isomerization and product release should be detectable by smFRET. In order to see such potential dynamics within the guide RNA during U-isomerization, the target RNA U-scSub was added to scImaging-buffer in

high excess (c_{final} = 2 µM) over the immobilized H/ACA-RNPs. Sample annealing and immobilization were performed as in the assembly experiments for the complete complexes (RNA+WNCG). The imaging buffer containing the substrate RNA (c_{final} = 2 µM) was added after a sufficient molecule density was observed, and several short movies (2 sec) for histogram plotting as well as long videos (200 sec instead of 120 sec) for single molecule trajectories were recorded. The summarized results for *H5-FRET* (see Figure 5.23a) and *H5d-FRET* (see Figure 5.23b) are shown below.

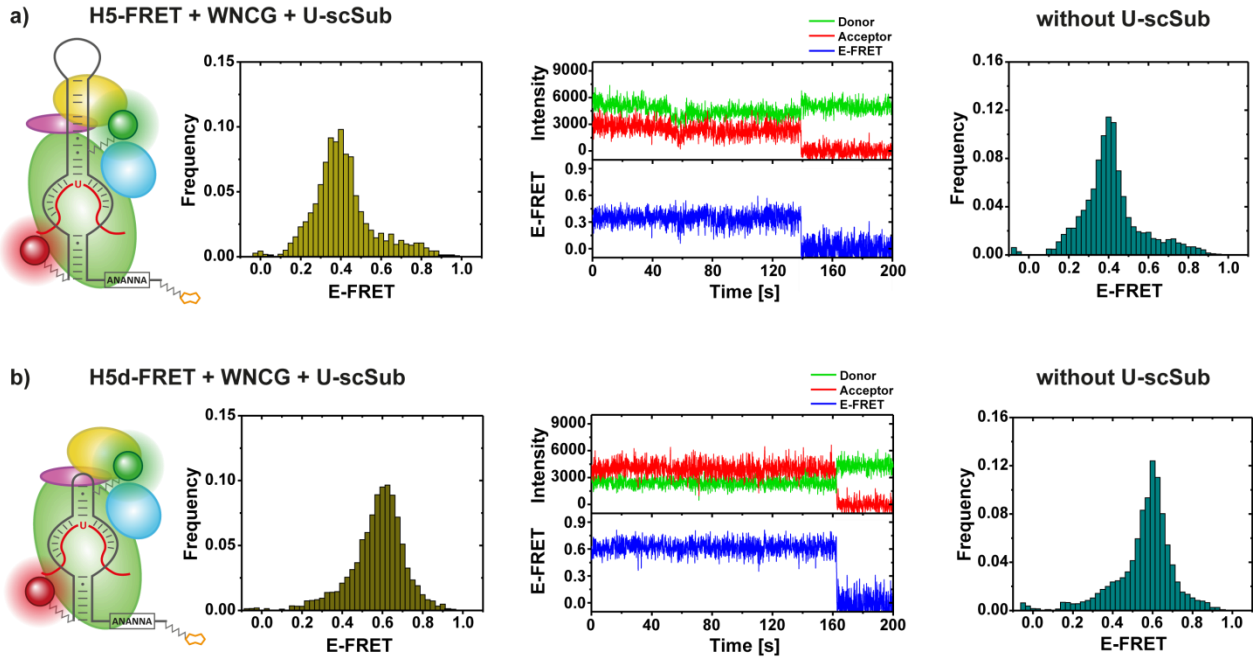


Figure 5.23: ***a)*** **Histogram of H5-FRET+WNCG in presence of 2 µM U-scSub (left), an exemplary trace (center) and the histogram of H5-FRET+WNCG from the assembly experiment (right) for comparison.** ***b)*** **Histogram of H5d-FRET+ WNCG in presence of 2 µM U-scSub (left), an exemplary trace (center) and the histogram of H5d-FRET+WNCG from the assembly experiment (right) for comparison. Donor only peaks were removed by fitting.**

The histograms of both constructs showed no significant differences when substrate RNA was added to the imaging buffer, compared to the assembled substrate free H/ACA-RNPs. As previously observed, *H5-FRET* showed a major population around 0.4 FRET-efficiency, and *H5d-FRET* at around 0.6 E-FRET. Furthermore, evaluation of the traces yielded no distinct transitions. Once a state was occupied the guide RNA remained in this conformation. These results indicate that the guide RNA does not undergo major structural rearrangements during pseudouridylation.

From the experiments shown above no clear statement can be made, on whether target RNA is bound to the H/ACA-RNP or not. In order to check for substrate recruitment the labeling scheme

was altered. The FRET-donor remained at the guide RNA at position U61 for the complete 5'-hairpin (*H5-Cy3*), or at the respective U26 for the truncated 5'-hairpin (*H5d-Cy3*). The FRET-acceptor was attached to a substrate analogue containing 5fU as target nucleotide (5fU-scSub-Cy5), which decreases substrate turnover rates according to the activity data from Chapter 5.4.2. With this labeling scheme, FRET could only be detected when the target RNA is within close distance (< 10 nm) of the snoRNA. Samples containing labeled guide RNA (1.0 eq) and 5fU-scSub-Cy5 (1.3 eq) were annealed at 30 °C for 15 min in presence of either all four proteins (WNCG) or in absence of Nhp2 (scNCG). Figure 5.24 summarizes the results of these experiments.

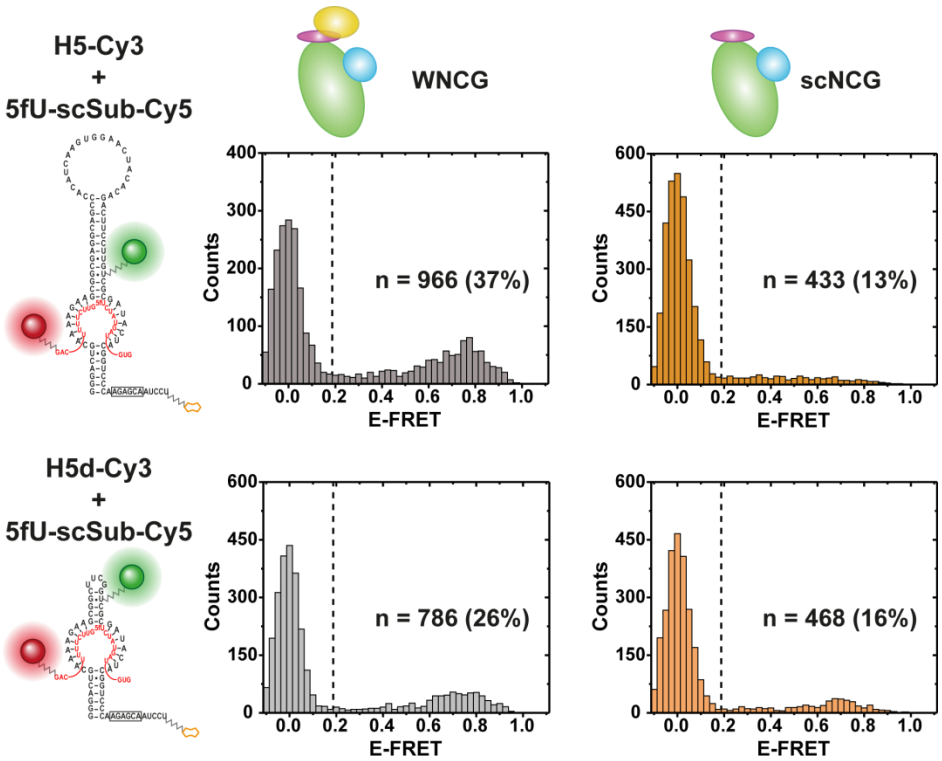


Figure 5.24: Histograms of H5-Cy3 (top) or H5d-Cy3 (bottom) plus 5fU-scSub-Cy5 with WNCG (left) or scNCG (right). The absolute number of detected molecules for E-FRET ≥ 0.2 is given as 'n' (the ratio of n compared to all observed molecules for the respective sample is given in %).

For the complete RNPs (WNCG) two broad high-FRET populations around E-FRET = 0.75 were detected. *H5-Cy3* showed 37% and *H5d-Cy3* 26% molecules with E-FRET ≥ 0.2. Removal of Nhp2 in the H/ACA-RNP (scNCG) reduced the relative number of target RNA bound molecules down to 13% for *H5-Cy3*. For *H5d-Cy3* the overall loss was comparably smaller (16%). These results indicate that Nhp2 affects substrate recruitment differentially for the complete and the partial 5'-hairpin.

In order to identify various states during substrate turnover and to directly detect pseudouridylation, the scope of the experiment shown above was expanded to other target RNAs. Three different target RNAs with varying bases at the target site were used: the previously mentioned 5fU-scSub-Cy5 – expected to extend the duration of pseudouridylation –, a target RNA to observe the actual isomerization (U-scSub-Cy5), and an RNA carrying Ψ at the target site (Ψ-scSub-Cy5) to elucidate possible release states. *H5-Cy3* was annealed with all proteins (WNCG) or in absence of scGar1 (WNC), diluted to 0.5 nM guide RNA concentration and immobilized. The labeled substrates RNAs (X-scSub-Cy5) were not pre-annealed as the guide RNAs and proteins, but added to the scImaging-buffer (c_{final} = 10 nM). Higher concentrations of Cy5-labeled target RNAs in the imaging buffer led to a significant increase in background fluorescence in the red channel, and were therefore avoided. The histograms for these experiments are summarized in Figure 5.25.

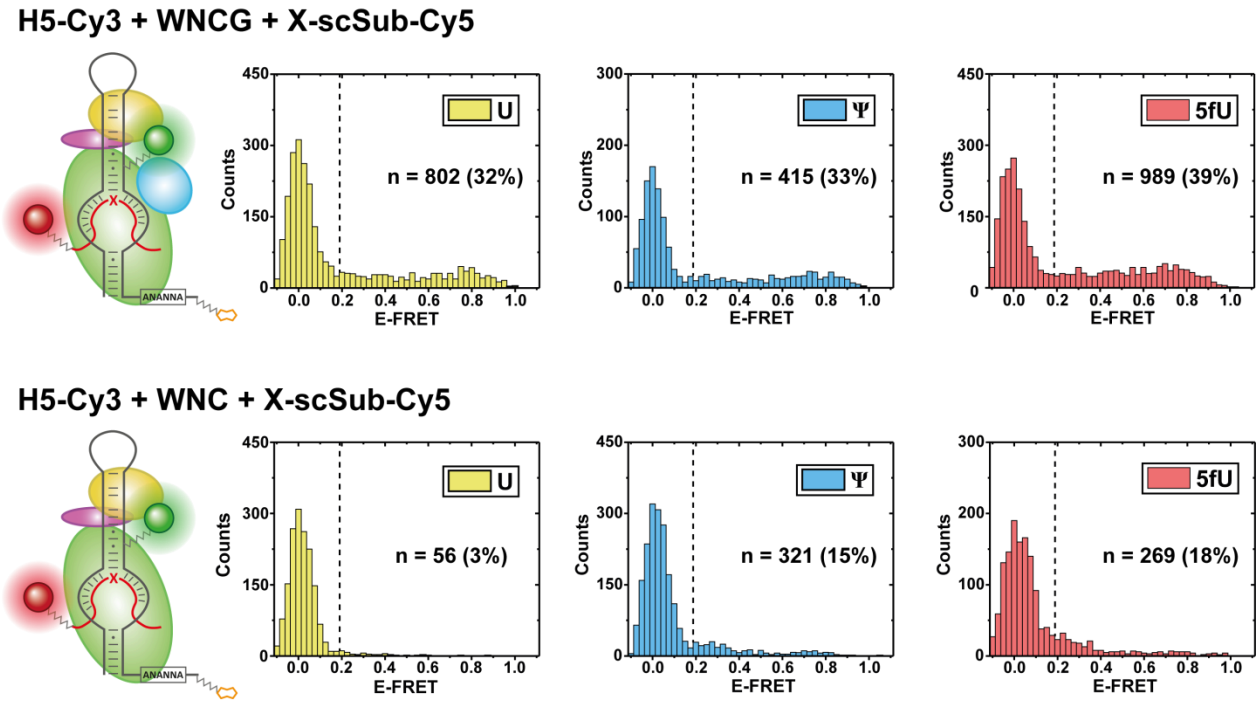


Figure 5.25: Histograms of H5-Cy3+X-scSub-Cy5 in presence (top) or absence (bottom) of scGar1 plus WNC. Plots of the different target nucleotides in X-scSub-Cy5 are color coded: U (light yellow), Ψ (light blue) and 5fU (light red). The absolute number of detected molecules for E-FRET ≥ 0.2 is given as 'n'.

From the results presented above several trends are discernible. First of all, in presence of scGar1 (top row of histograms) the relative number of molecules experiencing FRET ≥ 0.2 (given in percent) is notably higher than in the absence of scGar1 (bottom row of histograms). In case of *H5-Cy3*+WNCG over 30% of detected molecules experienced FRET for each construct, with a

maximum of 39% for 5fU-scSub-Cy5. Omission of scGar1 (*H5-Cy3*+WNC) led to a notable reduction in the relative number of observed FRET-events. While for Ψ- and 5fU-scSub-Cy5 molecules experiencing FRET was reduced by around a half, the actual substrate (U) showed by far the least FRET-events, with only 3% of all molecules. Furthermore, other trends for the overall observed FRET-efficiencies can be assumed. Despite the dominant donor-only peak – representing molecules that do not carry Cy5-labeled target RNA or molecules where the acceptor fluorophore bleached – the formation of a second major FRET population is not clearly visible. Nevertheless, there seems to be a small trend to higher FRET-efficiencies (around E-FRET = 0.75) for the fully assembled RNP (*H5-Cy3*+WNCG+X-scSub-Cy5). This broad distribution in E-FRET might either indicate unspecific binding modes of the target RNA to the snoRNP, or a highly dynamic system. To identify which of these two assumptions is more likely, the traces of single molecules were analyzed and evaluated. Two representative movies for *H5-Cy3*+WNCG are shown in Figure 5.26.

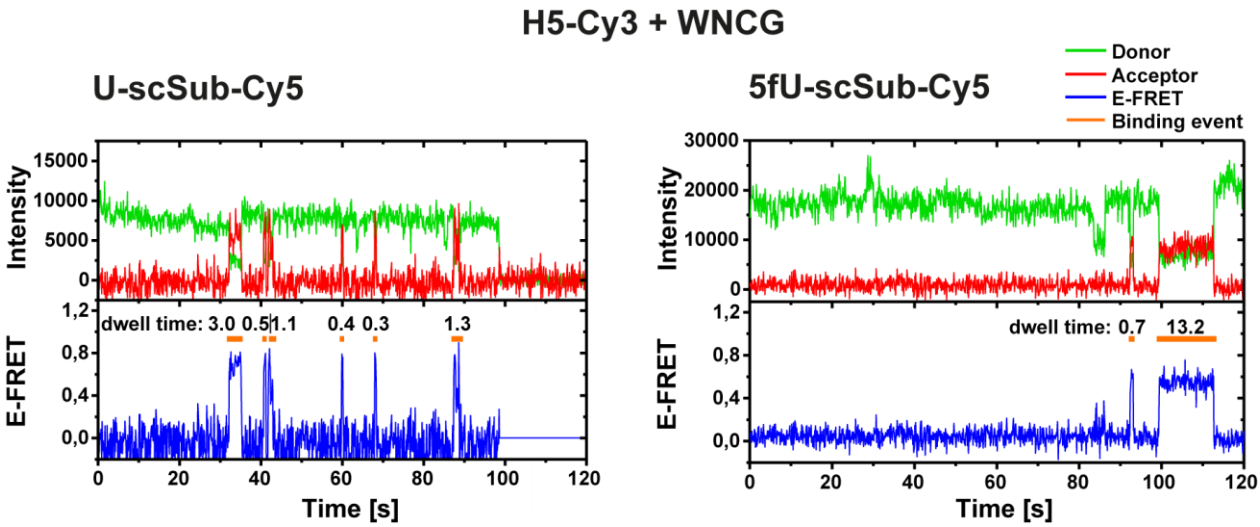


Figure 5.26: Representative traces of H5-Cy3+WNCG in presence of U-scSub-Cy5 (left) and 5fU-scSub-Cy5 (right).

In the fluorescence trajectories of single molecules spontaneous acceptor intensity bursts, with corresponding donor intensity decrease, can be observed. These FRET-events are highlighted with orange bars and their length is given in seconds. These spontaneous rises in E-FRET from 0 to typically 0.5 to 0.8 represent target RNA binding to snoRNP. The dwell time of these binding events could last from a few hundred milliseconds to several seconds and in some cases even for over a minute. Generally, short and long binding events could be observed for all samples – all three target RNAs and in presence and absence of scGar1 –, but with varying frequencies. Single traces showed individual behavior, while some molecules had frequent FRET-bursts, other

showed only one or two binding events, and many showed none at all. As the available evaluation software could not be successfully applied to this analysis, evaluation of all traces from all these constructs was done manually and is summarized below.

Binding events	**Dwell time / s**	**WNCG** *H5-Cy3* + **X**-scSub-Cy5			**WNC** *H5-Cy3* + **X**-scSub-Cy5		
		U	**Ψ**	**5fU**	**U**	**Ψ**	**5fU**
ultra-short	< 0.5	71	55	32	6	16	5
short	0.5 – 1.0	78	49	31	1	23	13
intermediate	1.1 – 3.0	58	35	45	4	49	18
long	3.1 – 10.0	55	51	65	5	46	29
very long	> 10.0	84	47	178	13	143	61
Sum of binding events		346	237	351	29	277	126
# of molecules with binding events		159	157	298	29	217	99
# of detected molecules		605	571	962	640	806	575

Table 5.2: Summarized results of observed binding events for H5-Cy3+WNCG/WNC+X-scSub-Cy5.

The number of detected binding events is given in Table 5.2. They were divided into five categories: ultra-short (< 0.5 sec), short (0.5 – 1.0 sec), intermediate (1.1 – 3.0 sec), long (3.1 – 10.0 sec) and very long (> 10.0 sec). Several trends from this table can be argued. In absence of scGar1, the relative quantity of molecules experiencing target RNA binding decreased. For U-scSub-Cy5 the ratio of molecules experiencing binding events dropped from 26% (159 of 605 molecules) to 5% (29 of 640 molecules) when scGar1 was omitted. For 5fU-scSub-Cy5 this decrease was less drastic but still noteworthy; from 31% (298 of 962 molecules) to 17% (99 of 575 molecules). Only for Ψ-scSub-Cy5 the ratio of molecules experiencing FRET-bursts remained constant at 27% (157 of 571 molecules in case of WNCG and 217 of 806 molecules for WNC). Furthermore, the relative number of long and very long dwell times increased when scGar1 was removed or when 5fU was at the target site. The ratio of binding events for the different constructs will be discussed particularly in Chapter 5.6.3.

5.6. Discussion

5.6.1. Eukaryotic H/ACA-RNP Reconstitution

Protein Preparation

To acquire enzymatic and structural data of the snR81 H/ACA-RNP from *Saccharomyces cerevisiae*, snoRNAs and the four H/ACA-proteins were synthesized, purified and tested for biological function. Proteins were obtained by recombinant expression in *E. coli* and purified using affinity and size exclusion chromatography, similar as for the archaeal particles. Buffers were adapted from Li *et al.*[59]. Expression of Nhp2 and the ternary scNCG complex was successfully carried out at 20 °C, even though with reduced yields; otherwise aggregation and precipitation of the proteins after cell lysis were observed. Nhp2 was expressed and purified separately, while the other proteins – scNop10, scCbf5 and scGar1 – were obtained simultaneously as ternary subcomplex 'scNCG'. For Nhp2 the S82W mutant with a C-terminal HisTag was used. This modification was shown to increase snoRNA binding affinity[65] and simplified detection as well as concentration determination of the protein by measuring the UV-absorbance at 280 nm, since no other tryptophan occurs in the wildtype amino acid sequence. In case of scNCG the full-length constructs were used in this thesis. scCbf5 and scGar1 carried a C- and N-terminal HisTag, respectively. As both Nhp2 and the scNCG subcomplex are capable of RNA-binding; binding of endogenous *E. coli* RNA during protein expression and purification led to undesired RNA contamination. To remove these prokaryotic RNAs two RNA digestion steps via RNase A were included during the purification procedure; one directly during cell lysis and a second after the first Ni-NTA-chromatography step. Incubation for the second step could either be done at 4 °C overnight or at 15 – 20 °C for 30 mins to several hours. Higher incubation temperatures led to visible precipitation of the scNCG complex, and were thus avoided. In order to remove the RNase a second affinity chromatography was performed and then gel filtration was done as a final purification step. Nhp2 (see Figure 5.2, page 128) and scNCG (see Figure 5.4, page 131) yielded homogeneous peaks in the size exclusion. For Nhp2 minor protein impurities were still visible, but the 260/280 nm ratio (0.53) indicated RNA-free protein. In case of scNCG minor RNA impurities could not be completely removed, indicated by a 260/280 nm ratio of 0.79. Furthermore, for scGar1 not a single but multiple bands were observed in the SDS-gel were detected. Thus, eukaryotic scGar1 is partially degraded within the scNCG-complex. This

behavior has been reported previously[59]. Nevertheless, the homogeneity of the peak from the gel filtration showed that coexpression and purification of ternary subcomplex were successful. RNA-free proteins used for activity and smFRET experiments were provided by Sven Trucks (Goethe University).

In the final stages of this thesis, attempts were made to purify and obtain homogeneous scGar1 in the scNCG subcomplex. Therefore, two new scGar1-constructs were designed: one without any HisTag (scNCG-notag) and one with a C-terminal instead of the N-terminal HisTag (scNCG-Ctag). In both cases a part of the N-terminus containing the HisTag was removed. Furthermore, the purification protocol was optimized to ensure RNA-free proteins. After sample application to the Ni-NTA-column and before the initial chromatography the bound complexes were washed with a high-salt buffer containing 2 M LiCl. The high-salt conditions led to the dissociation of RNAs bound with low affinity from the scNCG complex. RNAs exhibiting higher affinity to the subcomplex were removed with RNase after the initial chromatographic purification. SDS-gels of the three scNCG constructs (scNCG-notag, scNCG-Ctag and the actual scNCG 'scNCG-Ntag') after the second Ni-NTA-column showed that significantly less scGar1-degradation occurs for scNCG-notag and –Ctag compared to the previously used complex (see Figure 5.6, page 133). Furthermore, subsequent size exclusion of scNCG-Ctag yielded a homogeneous peak containing all three proteins and an improved 260/280 nm ratio (0.58 instead of 0.79) showing that the applied LiCl-washing step does not interfere with complex stability, and reduces RNA contaminations in the sample. Thus, these changes in scGar1 yielded promising constructs for future experiments with recombinantly purified snoRNPs.

RNA Preparation

For the beginning of the project, not the full-length snoRNA, consisting of a double hairpin structure, but standalone hairpins (5' and 3') were used for initial studies on eukaryotic RNA guided pseudouridylation. It has been shown that such single hairpins are capable of uridine isomerization[59,72]. The focus was laid on the 5'-hairpin. Two constructs were designed and prepared; a complete (H5) and a truncated (H5d) hairpin (see Figure 5.7, page 134). With these constructs it was aimed to investigate the impact of the well conserved and comparably large loop structure of eukaryotic H/ACA-RNAs. Unlabeled guide RNAs were synthesized via *in vitro* transcription. Analytical gels showed the successful purification and an overall homogeneous

folding of H5 and H5d (see Figure 5.8, page 135). Thus, the prepared RNAs were used for complex reconstitution and activity assays. Unlabeled 3'-hairpin constructs (H3 and H3d) for activity assays were provided by Sven Trucks and Christin Fuks (Goethe University).

Labeled snoRNAs were prepared via enzymatic splinted ligation of chemically synthesized oligonucleotides, modified with reactive amines for individual fluorophore attachment. Since no structural data of eukaryotic H/ACA-RNPs was known, labeling sites in the snoRNAs were adapted from the previously and successfully tested archaeal constructs. Dyes could be attached at the 5'-end (G1) and in the upper stem (U61 or U26) above the pseudouridylation pocket for the 5'-hairpin constructs (*H5-FRET* and *H5d-FRET*), and directly in the upper (U21) and lower (U83 or U42) stem for *H3-FRET* and *H3d-FRET* (see Figure 5.9, page 137). Oligomer labeling and purification as well as RNA ligation and purification were successful. Overall, the yields for the eukaryotic FRET constructs were lower (~10%) compared to the archaeal RNAs (~20%), which is probably due to the fact that three instead of two oligomers had to be annealed to the DNA splint for complete enzymatic ligation. Nevertheless, the amount of purified and labeled RNAs was sufficient for reconstitution, activity and smFRET experiments. Final constructs are shown in Figure 5.12 (page 140). These were used to investigate eukaryotic complex assembly, RNA guided pseudouridylation and for comparison to the archaeal H/ACA-RNP.

Complex Reconstitution

RNA-protein binding of the prepared biomolecules was tested with EMSAs. Unlabeled and labeled 5'-hairpin RNAs were annealed to Nhp2 and scNCG in separate measurement rows with increasing protein concentrations. RNP formation was observed for all constructs, but distinct migration of the RNA-protein particles into the gels could not be seen (discussed in Chapter 5.3). A possible explanation for this observation might be eventual precipitation/aggregation of the partial snoRNP at EMSA conditions, as was observed for archaeal complexes. Apparently, size and net-charge of the eukaryotic RNPs is not adequate for effective PAGE analysis. Nevertheless, qualitative statements from the RNA-protein binding experiments can be made. EMSAs with *H5*- and *H5d-FRET* were performed at 100 nM RNA concentration. Both showed complete RNA binding at a 10-fold excess (1.0 µM) of scNCG, or at a 20-fold excess (2.0 µM) of Nhp2. These results show that scNCG appears to have a higher affinity to the 5'-hairpin than Nhp2, and is in agreement with recent literature. For a different snoRNA (snR34) Caton *et al.* could show that the

K_d of single Nhp2 (~600 nM) is significantly higher – by three orders of magnitude – than that of scNCG (0.3 nM)[72]. This drastic variation in affinity is not seen in the EMSAs, but the overall trend is the same. Possible reasons for these deviations might be the usage of a dissimilar snoRNAs, and the use of a different type of assay (surface affinity versus gel migration). However, since the free RNA was completely bound at 1.0 µM scNCG (10-fold excess), and Nhp2 to snoRNA affinity increases with the presence of scNCG[72], complex annealing for activity and smFRET experiments was performed with 10-fold excess of each protein.

5.6.2. Eukaryotic Pseudouridylation

Single Turnover Assays

The designed and synthesized H/ACA-RNPs were for their function. Especially the activity of the labeled RNAs was of high interest, since the introduction of large organic dyes into enzymes may lead to interferences with the biomolecules function. For initial activity tests, pseudouridylation assays with radioactive substrate RNAs were performed at equimolar amounts of guide and target RNAs (0.1 µM). Unlabeled and labeled 5'- and 3'-hairpins were tested in presence of all four proteins (WNCG, 1 µM) at 30 °C for 1 h. Additionally a control experiment was performed to see whether the eukaryotic H/ACA-proteins inherit catalytic function in the absence of the snoRNA. For the 5'-hairpins (H5, H5d, *H5-FRET* and *H5d-FRET*) nearly quantitative Ψ-formation could be seen, while the 3'-hairpin constructs (H3, H3d, *H3-FRET* and *H3d-FRET*) showed no product at all (see Figure 5.15, page 144). Furthermore, the standalone proteins inherited no pseudouridylation ability in the absence of a snoRNA. Overall, the labeled 5'-hairpins showed slightly less product formation compared to their unlabeled counterparts, but were still highly active. Interestingly, the standalone 3'-hairpins were not capable of uridine isomerization under the given conditions. Folding tests via native PAGE of *H3*- and *H3d-FRET* yielded one distinct RNA-band for each construct with comparable migration pattern as *H5*- and *H5d-FRET* (see Figure A3.8 in Appendix A3 – Supplementary Figures). Thus, misfolding of the guide RNA can be excluded as explanation for the observed inactivity. In single turnover experiments with the full-length snR81 RNA – performed by Christin Fuks (Goethe University) – high pseudouridylation activity was observed for the 3'-hairpin, indicating that its function depends on the presence of the 5'-hairpin. Furthermore, in multiple turnover assays with H3 – performed by Sven Trucks (Goethe University) – Ψ-formation could be detected after several

hours of incubation. These results give rise to the assumption that proper RNP assembly of H3 is hindered in the absence of the 5'-hairpin, and the latter one is required for correct function of the 3'-hairpin.

Multiple Turnover Assays

After the unlabeled and labeled 5'-constructs have been tested successfully for enzymatic function, several multiple turnover assays with a 40-fold excess of substrate RNA over enzyme were performed to investigate the function of individual proteins. For this, Nhp2 or scGar1 were omitted in individual experiments. Furthermore, 5fU-scSub instead of U-scSub was added in one experiment to see whether this modification affects catalytic function of the H/ACA-RNP. Most importantly, catalytic uridine turnover of the labeled guide RNAs was to be tested at moderate temperatures (20 °C) to check whether pseudouridylation can be observed in smFRET spectroscopy, which is typically performed at room temperature. In order to obtain kinetic information, Ψ-yields of each experiment were determined after 5, 20, 60, and 240 min (for the ΔscGar1-RNP only after 60 min). The obtained data points were fitted to a Michaelis-Menten like equation (see Figure 4.21, page 98) yielding the time required for the enzyme to isomerize half of the remaining substrate ($Sub_{1/2}$). For kinetic comparison of individual complexes, the initial velocity of Ψ-formation (v_{start}) in min^{-1} was approximated via linear regression of the early reaction phase. Therefore, the data point at 5 min from the obtained fit was taken (*Note*: this point was used for approximation since it resembled the first physically taken data point, and the reaction progress can be estimated to be linear with a decent certainty for this time window). Errors were not calculated for this method, since only one data point was used for the linear approximation. This approach was used to make an educated guess for the actual maximum reaction rate of RNA guided pseudouridylation, which could not be determined by the applied conditions; significantly higher substrate excess (> 1000-fold) would be required to mimic ideal settings for Michealis-Menten kinetics. A summary of fitted ($Sub_{1/2}$) and approximated (v_{start}) kinetics of the performed activity assays under multiple turnover conditions is given in Table 5.3.

RNA	H5				(5fU-scSub)	H5d			H5-FRET		H5d-FRET	
Proteins	full	ΔNhp2	ΔGar1	full	full	full	ΔNhp2	full	full		full	
Temperature	30 °C	30 °C	30 °C	20 °C	30 °C	30 °C	30 °C	20 °C	30 °C	20 °C	30 °C	20 °C
$Sub_{1/2}$ / min	6.3±0.4	187±14	n. a.	52±2	872±199	9.1±0.6	63±5	74±3	12.6±0.9	78±11	15.5±0.6	133±16
v_{start} / min^{-1}	3.4	0.20	-	0.67	0.04	2.7	0.56	0.48	2.2	0.46	1.9	0.28

Table 5.3: Summarized pseudouridylation kinetics of the 5'-hairpin constructs.

Overall, the highest activity was observed for the fully assembled H5-RNP at 30 °C with an estimated starting rate v_{start} = 3.4 min^{-1}. Comparing these results to already existing *in vitro* rates of a eukaryotic H/ACA-RNP composed of the snR5 RNA by Li *et al*. – 0.073 min^{-1} for the complete snR5-RNP and 0.028 min^{-1} for the standalone snR5-5'-hairpin[59] –, RNA guided pseudouridylation of the standalone H5-RNP prepared in this thesis is faster by two orders of magnitude. One reason for this tremendous deviation in turnover speed might be the use of a different guide RNA, since individual RNAs seem to inherit differential catalytic activity[59,72]. Another reason might be that complex reconstitution of Li *et al.* generated less active Ψases in comparison to this thesis. Hence, the high initial velocities of the standalone 5'-hairpins show that RNA and protein preparation, as well as *in vitro* complex reconstitution, yielded highly active eukaryotic H/ACA-RNPs. Addition of 5fU-scSub in the activity test led to a significant decrease in the reaction rate of H5+WNCG (v_{start} = 0.04 min^{-1}), but not to complete loss of activity. Thus, the modified target base does not inhibit enzymatic function of snR81 RNP completely, but slows down substrate turnover by around two orders of magnitude. From this experiment it cannot be stated whether 5fU is isomerized to 5fΨ, or if the target RNA binds to the complex and is released unmodified thereafter.

The truncated guide RNA H5d showed a similar activity at the same conditions with v_{start} = 2.7 min^{-1} as the full-length 5'-hairpin. It can be inferred that the well conserved and large loops in eukaryotic H/ACA-RNAs are not necessarily essential for efficient pseudouridylation and their cellular function is not within the process of Ψ-formation. Interestingly, removal of Nhp2 has a stronger effect on the H5-RNP than on the Ψase made of the truncated guide RNA. Pseudouridylation of H5d+scNCG (v_{start} = 0.56 min^{-1}) is almost three times faster than of H5+scNCG (v_{start} = 0.20 min^{-1}). These results point towards a stabilization function of Nhp2 in eukaryotic H/ACA-RNPs. Possibly, Nhp2 anchors the upper stem of the snoRNA by binding to the scNCG complex and the guide RNA. In case of H5d, where the large loop sequence is

substituted by a highly stable tetra loop UUCG (see Figure 5.7, page 134), this anchoring function of Nhp2 seems to be less important for enzyme function. Nevertheless, removal of this protein led to a notable decrease in turnover rate for the truncated RNP. Thus, Nhp2 is still interacting with the guide RNA and/or scNCG subcomplex and promoting catalytic function. Omission of scGar1 on the other hand led to a complete catalytic inactivation of the H/ACA-RNP. This is consistent with existing literature, since uridine isomerization in absence of Nhp2 could be detected for several snoRNPs, but not for ΔscGar1 complexes[59,72]. Apparently, scGar1 inherits an essential function and is more important than Nhp2 for effective uridine isomerization. Comparing these results to archaeal RNA guided pseudouridylation; the kink-turn binding protein L7Ae seems to be more relevant for Ψ-formation than its eukaryotic counterpart Nhp2, and scGar1 is not as essential in archaea as in yeast. This was shown in multiple turnover assays for the thermophile H/ACA-RNP from *P. furiosus*, also studied in another thesis[128]. In single turnover experiments, removal of archaeal Gar1 or L7Ae had a similar impact (see Figure 4.20, page 96). A possible explanation for the higher importance of L7Ae in pseudouridylation compared to Nhp2 could be found in the native temperatures of the host organisms (*Pyrococcus furiosus* or *Saccharomyces cerevisiae*). The specific L7Ae-kink turn formation might be essential at high temperatures (70 - 100 °C) to stabilize the RNP. While the product releasing functionality, which could be shown for archaeal Gar1[63,64], is not as relevant as stability for the thermophile Ψase, due to the higher thermal energy in the system. Conclusively, scGar1-presence is more critical in eukaryotic pseudouridylation at moderate temperatures where less thermal energy for molecular flexibility is available, than in archaeal pseudouridylation of the thermophile organism *P. furiosus*.

Catalytic activity of FRET-labeled 5'-hairpins was tested at 30°C and 20 °C. Both constructs showed substrate turnover rates at 30 °C –2.2 min^{-1} for *H5-FRET* and 1.9 min^{-1} for *H5d-FRET* – with only minimal losses related to the unlabeled guide RNAs. A decrease in incubation temperature yielded the same trend for labeled guide RNAs as for the unlabeled ones. v_{start} was reduced approximately by a factor of five for each construct when the temperature was decreased by 10 °C: to 0.48 min^{-1} for *H5-FRET* and to 0.28 min^{-1} for *H5d-FRET*. Compared to the calculated v_{start} of the unlabeled guide RNAs (0.67 min^{-1} for H5 and 0.48 min^{-1} for H5d), these are only minor losses in catalytic activity. Thus, fluorophore insertion did not have a major impact on H/ACA-RNP function, which in turn means that the described observation of states

during complex assembly and pseudouridylation via smFRET is highly likely to be of biological relevance.

5.6.3. smFRET Analysis of Eukaryotic H/ACA-RNPs

H/ACA-RNP Assembly

Assembly experiments of standalone hairpins from eukaryotic snR81 snoRNP were performed with FRET-labeled, biotinylated guide RNAs. To this end, full-length (*H5-FRET* and *H3-FRET*) and truncated (*H5d-FRET* and *H3d-FRET*) constructs were annealed with or without Nhp2, scNCG or WNCG in absence of substrate RNA. Fluorophores were attached above and below the pseudouridylation pocket at the upper and lower stem of each guide RNA (see Figure 5.12, page 140). In this area – where substrate binding, turnover and product release occur – the most conformational rearrangements within the guide RNA were expected. Furthermore, the labeling sites were similar to the *5'-FRET-H/ACA* construct from *P. furiosus* in order to compare archaeal and eukaryotic RNP assembly. A summary of the obtained histograms for *H5-FRET* (see Figure 5.17, page 149) and *H5d-FRET* (see Figure 5.18, page 149) is given below.

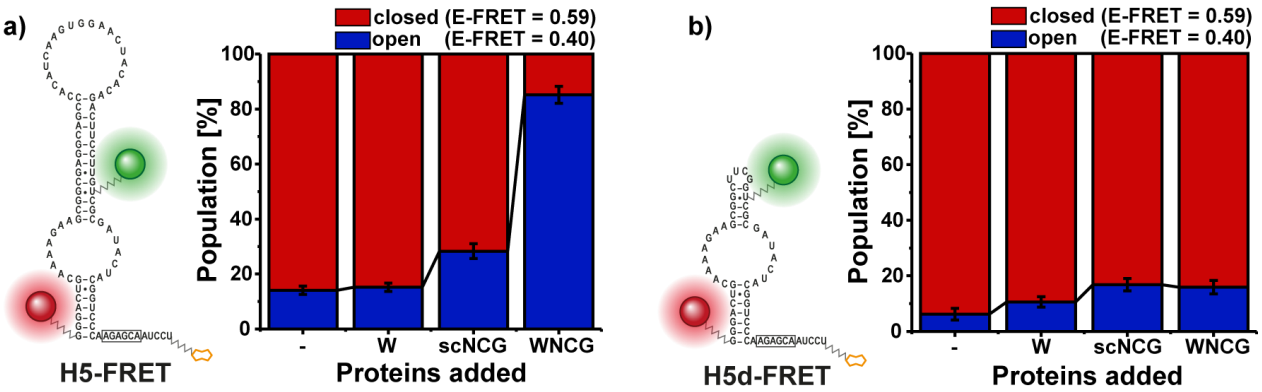


Figure 5.27: ***a)*** **Summary of H5-FRET assembly experiments.** ***b)*** **Summary of H5d-FRET assembly experiments.**

Overall, two states were observed for the 5'-hairpin constructs; a closed state at E-FRET = 0.59 (red) and an open state at E-FRET = 0.40 (blue). Assuming a Förster radius of R_0 = 6 nm, as was done for the archaeal constructs (discussed in Chapter 4.6.2), the average dye distance is 64.2 Å in the open and 56.5 Å in the closed state, representing an distance deviation of 7.7 Å. Free *H5-FRET* remained mainly in the closed conformation (86%). Addition of Nhp2 or scNCG showed near to none or only minor changes in the E-FRET (see Figure 5.27a). Only in presence of all four proteins (WNCG) the open state was predominantly populated (85%) for the full-length

standalone hairpin *H5-FRET*, causing an increase in fluorophore distance of 7.7 Å. A similar behavior was observed for the archaeal H/ACA-RNP. For the sRNA an average distance increase of 5.5 Å was observed between 5'-end and pseudouridylation pocket when the guide RNA switched from the closed to the open conformation. As for the archaeal H/ACA-RNA, it is assumed that the purpose of this rearrangement is to open the pseudouridylation pocket and thereby promote correct substrate loading into the RNP. The observed resemblances in *H5-FRET* and *5'-FRET-H/ACA* during complex assembly show that eukaryotic and archaeal H/ACA-RNP formation seem to follow the same mechanism.

For the truncated construct *H5d-FRET* this structural rearrangement was not observed. The guide RNA remained mainly in the closed population (> 80% for all samples), even when all proteins were added. One possible explanation is that protein binding does not initiate opening of the shortened RNA, as was observed for the previously mentioned constructs. In this case, the removed domains of the upper stem and loop would somehow influence RNA rearrangement between lower stem and pseudouridylation pocket during RNP assembly. Furthermore, it would mean that the open state is not essential for efficient pseudouridylation, since the truncated guide RNA showed nearly as high reaction rates as the native 5'-hairpin. Evidence for this explanation can be obtained from the substrate recruitment assays shown in Figure 5.24 on page 154. In this experiment an effect of 5fU-scSub-Cy5 binding was observed upon Nhp2 removal. The overall number of molecules binding the target RNA was reduced from 26% to 16% when Nhp2 was omitted from the system, showing that the presence of the protein promotes substrate recruitment and indicates the formation of an RNP.

Another reason might be that protein binding did not occur for *H5d-FRET* in the assembly tests. Since the fluorophores are only attached to the guide RNA, there is no direct evidence for protein presence in the performed assembly experiments. Tests performed by Sven Trucks with Cy3-labeled Nhp2, Cy5-labeled H5/H5d and scNCG support this assumption. Significantly less FRET-events were observed in case of the truncated guide RNA in comparison to the complete 5'-hairpin. This indicates that the binding affinity of the snoRNA to the H/ACA-proteins decreases when parts of the upper stem and the conserved loop are missing. For the 5'-hairpin of the snR34 H/ACA-RNA a K_d of 1.5 ± 0.8 nM to WNCG was observed. Samples in the assembly experiments were diluted to 0.5 nM guide RNA and 5.0 nM protein concentration prior to

immobilization, meaning that the smFRET experiments are performed in the regime of the dissociations constants. Omission of potential RNA binding sites within the guide RNA could lead to a decrease in binding affinity of the proteins and might prevent RNP formation in the given conditions. This could be tested by increasing RNA and/or protein concentration. However, increasing the concentration of labeled RNA to 1 nM led to a visible overloading of the glass slide and was therefore avoided. An increase in protein concentration led to the visible formation of aggregates on the glass slide and could also not be tested. In the end, a clear statement about the efficiency of RNP formation for *H5d-FRET* cannot be made.

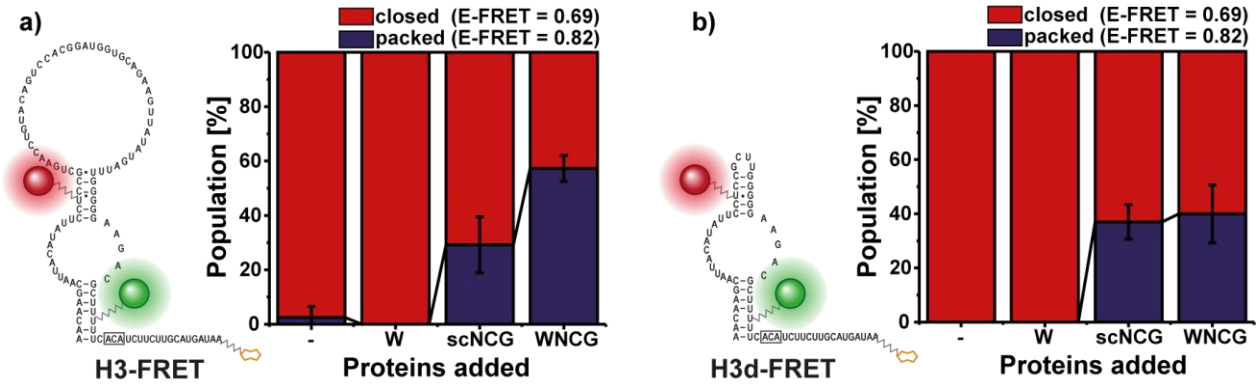


Figure 5.28: ***a)*** **Summary of H3-FRET assembly experiments.** ***b)*** **Summary of H3d-FRET assembly experiments.**

The summary of the assembly experiments for *H3-FRET* (see Figure 5.20, page 151) and *H3d-FRET* (see Figure 5.21, page 151) is given in Figure 5.28. Overall, two FRET populations were detected in the for the labeled 3'-hairpins; the closed state, for these constructs at E-FRET = 0.69, and a new high-FRET population at E-FRET = 0.82, termed packed state. The higher FRET-efficiency of the closed state for the 3'-hairpins is due to the fact that the number of base pairs in between fluorophores and pseudouridylation pocket is reduced from 11 bp for the 5'-hairpin constructs down to 8 bp (including the labeled nucleotides) for *H3-FRET* and *H3d-FRET*. Both constructs showed near to quantitative occupation of the closed conformation in case of the free RNA and in presence of Nhp2. Addition of scNCG or WNCG led to formation of the packed state for both RNAs. However, the closed state remained predominantly occupied except for *H3-FRET*+WNCG. The increase in E-FRET from 0.69 to 0.82 resembles a calculated average fluorophore distance decrease of 5.9 Å – from 52.5 Å (closed) to 46.4 Å (packed). Considering that the standalone 3'-hairpins showed no catalytic activity at single turnover conditions (see Figure 5.15, page 144), the observed packing between lower and upper stem of the guide RNA

most likely resembles a misfolded state of the H/ACA-RNP which prevents substrate RNA recruitment and thus inhibits enzymatic function of the partial Ψase.

Substrate Turnover

After investigation of eukaryotic H/ACA-RNP assembly with the partial snoRNAs, RNA guided pseudouridylation was investigated. All experiments regarding substrate turnover were performed with the 5'-constructs only, since the standalone 3'-hairpins inherited no catalytic activity. To see whether substrate binding, uridine isomerization or product release induce conformational changes within the guide RNA, unlabeled substrate (U-scSub, c_{final} = 2 µM) was added to immobilized *H5-FRET* and *H5d-FRET* RNPs (+WNCG). The addition of the target RNA showed no significant changes in the histograms of both constructs (see Figure 5.23, page 153). Furthermore, evaluation of time trajectories of single molecules showed no dynamics. Once a state was populated, E-FRET remained constant. Hence, no structural rearrangements in the guide RNA were not observed, indicating that the upper and lower stem remain passive during pseudouridylation. However, due to the labeling scheme (both fluorophores attached to the guide RNA) it cannot be said whether substrate RNA was bound to the complex, during the time course of the experiment, or not. Consequently, the labeling scheme was altered and the FRET-acceptor was attached to a substrate analogue carrying 5fU at the target site. In a multiple turnover experiment it was shown that this modification decreases the initial reaction rate by two orders of magnitude (see Table 5.3, page 163) and should therefore facilitate the detection of substrate bound H/ACA-RNPs by smFRET.

In order to assess the substrate recruitment process, guide RNAs carrying the donor-dye (*H5-Cy3* and *H5d-Cy3*) were annealed either in presence of all four proteins (WNCG) or in absence of Nhp2 (scNCG) with an excess of labeled 5fU-scSub-Cy5 (see Figure 5.24, page 154). In general, the histograms showed distinct substrate binding to the immobilized Ψase. In case of *H5-Cy3* 37% of all detected molecules carried a target RNA, and 26% in case of *H5d-Cy3,* when all proteins were present. Interestingly, removal of Nhp2 showed a more significant effect on substrate recruitment for the complete 5'-hairpin (down to 13%) in contrast to the truncated construct (down to 16%). These observations are in agreement with the activity results from the multiple turnover experiments. Omission of Nhp2 led to a stronger reduction in catalytic activity for H5 than for the shortened guide RNA H5d (see Table 5.3, page 163). Thus, the results from

the substrate recruitment experiments support the previously mentioned theory that Nhp2 has a stabilizing function within the snoRNP, but is not essential for complex activity. Statements on whether Nhp2 is directly bound to the guide RNA or only via protein-protein interactions, as was previously reported[71], cannot be made for the H5d-RNP from these experiments. Moreover, lack of structural data of functional eukaryotic H/ACA-RNPs does not allow making of well-founded assumptions. Nevertheless, it can be stated that Nhp2 does interact with the truncated RNP and promotes Ψ-formation, even though parts of the presumed binding site within the snoRNA – half of the upper stem and the conserved loop structure – are missing.

After target RNA binding was shown for the partial snoRNPs, substrate turnover was to be investigated. For this, the same labeling scheme as in the substrate recruitment assay was applied. In order to probe different states during pseudouridylation, labeled target RNAs carrying either U, Ψ or 5fU at the target site were added to the full-length 5'-hairpin in presence of all proteins (WNCG) or in absence of scGar1 (WNC) after complex immobilization. scGar1 was omitted in individual samples to investigate the effect of this protein on substrate recruitment, turnover and product release, since this protein is essential for H/ACA-RNP activity. In the summarized histograms of the performed experiments a significant decrease in molecules exhibiting FRET was observed upon scGar1 removal for *H5-Cy3* (see Figure 5.25, page 155), indicating that scGar1 does not only accelerate product release, as was shown for the archaeal protein[63,64], but also promotes substrate recruitment in a direct or indirect manner. A closer look into the fluorescence trajectories of single molecules showed that some RNPs experienced spontaneous FRET-bursts, representing binding of labeled target RNAs (see Figure 5.26, page 156). Dwell times of these binding events ranged from a few hundred milliseconds to over a minute and were divided into different categories: ultra-short (< 0.5 s), short (0.5 – 1.0 s), intermediate (1.1 – 3.0 s), long (3.1 – 10.0 s) and very long (> 10.0 s). A visual representation of the summarized results in Table 5.2 (page 157) is given in Figure 5.29.

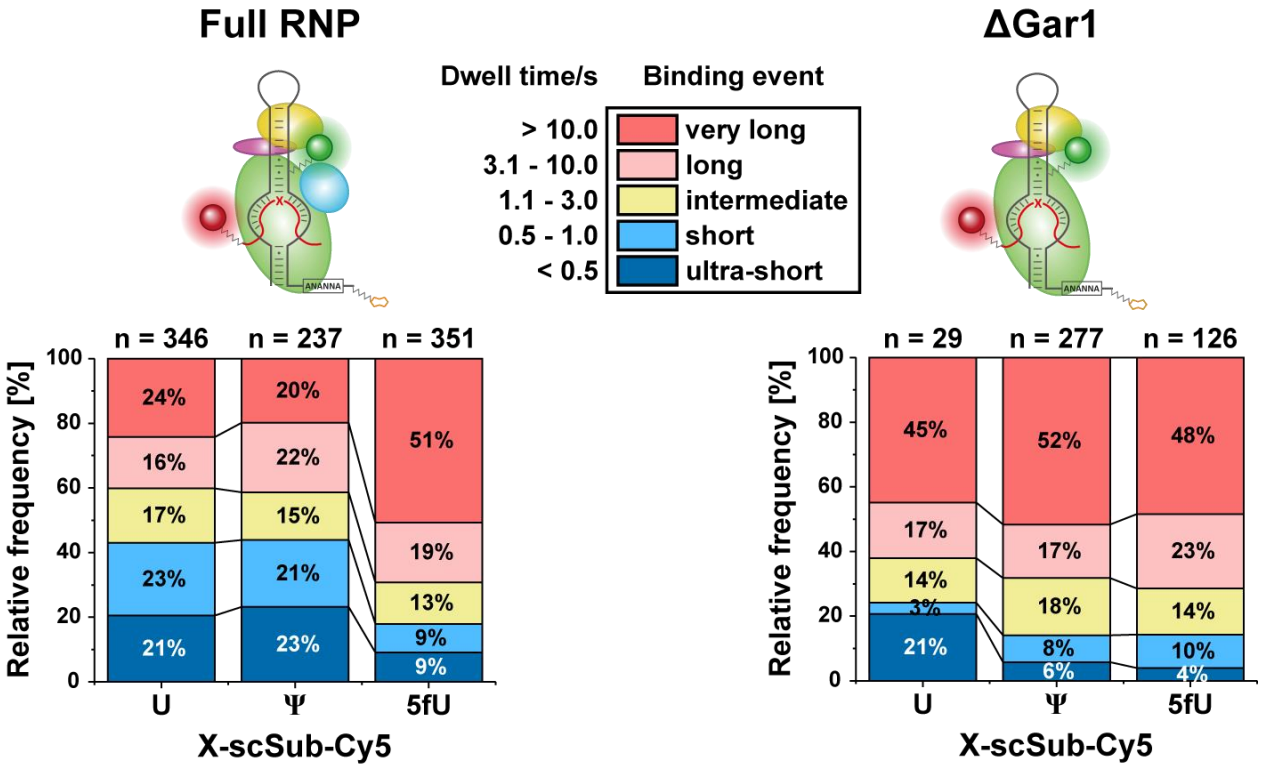


Figure 5.29: Summarized results of binding events for H5-Cy3+X-scSub-Cy5 in presence 'full RNP' (left) or absence 'ΔscGar1' (right) of scGar1. The total number of observed binding events per sample is given in 'n'.

For the full RNP the relative frequency of observed binding events was distributed more or less equally for U- and Ψ-substrate. Only in case of 5fU-scSub-Cy5 the very long dwell times occurred more frequently (51%), while the relative number of short and ultra-short bindings was reduced to 9% each. This trend could be expected, since 5fU at the target site significantly reduces the substrate turnover rate (see Table 5.3, page 163). Besides, removal of scGar1 had a similar effect on the dwell time of bound substrates. For all three target RNAs (U-, Ψ- and 5fU-scSub-Cy5) most of the FRET-events occurred longer than 10 seconds (very long). Overall the anticipated trends for 5fU-scSub and the ΔscGar1-complex were observed, indicating that binding and releasing of target RNAs by functional RNPs was observed. Considering the experimentally obtained initial velocity of 0.46 min^{-1} for *H5-FRET* at 20 °C, theoretically up to two pseudouridylation events per RNP may be observable every minute. However, conditions in the activity assays and smFRET tests cannot be compared directly. The reduction of guide RNA (from 100 nM to 0.5 nM) as well as target RNA (from 4000 nM to 10 nM) concentration, and the immobilization itself might have strong impacts on pseudouridylation kinetics in the smFRET experiments, most likely resulting in a significant decrease of v_{start} under the applied conditions. This might explain why the majority of detected molecules did not show any binding events (see

Table 5.2, page 157). Furthermore, it is important to note that the identity of target nucleotide cannot be investigated by smFRET. Hence, no statement on whether a target RNA was isomerized or released without modification can be made. Nevertheless, the results allow some educated guesses about function and kinetics of eukaryotic RNA guided Ψ-formation.

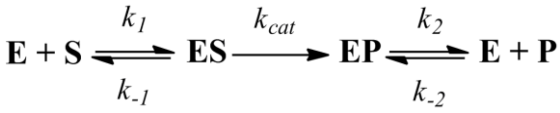


Figure 5.30: Schematic reaction kinetics of RNA guided pseudouridylation. The H/ACA-RNP is represented as 'E', the target RNA as 'S' and the isomerization product as 'P'. On and off rates are given as 'k_1' and 'k_{-1}' for the target RNA and as 'k_2' and 'k_{-2}' for the product. The actual isomerization process is represented by 'k_{cat}'.

In order to assign the categorized FRET-events to individual processes in pseudouridylation several considerations have to be made. Figure 5.30 displays the relevant kinetic steps during RNA guided Ψ-formation: enzyme (E) and substrate (S) binding (described by the on- and off-rates k_1 and k_{-1}), the isomerization of U to Ψ (described by k_{cat}) and the release of product (P) from the H/ACA-RNP (described by the on- and off-rates k_2 and k_{-2}). It is assumed that the actual modification step k_{cat} is irreversible, since no experimental evidence of reversed pseudouridylation events can be found. Given the fact that no significant difference in dwell times could be observed whether substrate (U) or product (Ψ) were present at the target nucleotide for the full RNP, the protein catalyzed isomerization (k_{cat}) must be a fast process in the millisecond to sub millisecond regime – faster than the applied time resolution in the smFRET measurements (100 ms). Furthermore, for the complete RNP the on-rate of substrate (k_1) and product (k_{-2}) must be within the same order of magnitude – as was shown for a thermophile H/ACA-RNP from *P. furiosus*[64] –, otherwise differences in dwell times for U and Ψ would have been observable. Conclusions about the substrate off-rate k_{-1} for the full RNP cannot be made from these experiments, since a correctly bound substrate is most likely isomerized too rapidly for dissociation of the enzyme-substrate complex. If 5fU is used instead of U or Ψ as a target nucleotide a decrease of k_{cat} by two orders of magnitude can be expected according to the results of the multiple turnover assays (see Table 5.3, page 163). This prolongation in reaction time is most likely represented by the increased frequency of molecules experiencing long (19%) and very long binding (51%) events for the full RNP.

In case of the ΔscGar1 complex, the reaction scheme is reduced to the first step: enzyme and substrate binding and dissociation, since the partial H/ACA-RNP is not capable of Ψ-formation.

Samples missing scGar1 also showed an increase in FRET-events with very long dwell times for all target RNAs, indicating that target RNA dissociation of substrate (k_{-1}) or product (k_{-2}) from the partial RNP are notably slower compared to the full RNP. This finding indicates that eukaryotic scGar1 also promotes product release, likewise to its archaeal equivalent. However, binding events of intermediate, short and ultra-short dwell times were observed for all samples and it is not possible to clearly attribute one of these categories to a pseudouridylation event. It may be that the ultra-short FRET-events (< 0.5 s) represent spontaneous association and dissociation of the target RNA to the RNP due to incomplete base pairing. Considering an estimated melting temperature of 26 °C for the target RNA (for the calculation the online tool 'OligoCalc'[124] was applied and only nucleotides contributing to RNA-RNA pairing were considered), spontaneous dissociation events of incompletely base paired substrate RNAs might occur. In this case intermediate or long binding events might represent a complete pseudouridylation process, including: substrate binding (k_1), isomerization (k_{cat}) and product release (k_2), which would probably take place in the lower second timescale. Consequently, formation of an active enzyme-substrate complex (ES) would be the rate limiting step in RNA guided pseudouridylation, since isomerization and product release are protein catalyzed steps, occurring much faster than RNA-RNA binding.

However, the above described results should only be considered with care, since several error sources are immanent. First of all, it cannot be distinguished whether a labeled target RNA dissociated from the complex or the FRET-acceptor bleached. Additionally, the applied labeling scheme does not allow ensuring whether the detected RNAs carry all of the added proteins. Lack of any of the four proteins would significantly alter function and kinetics of the H/ACA-RNP and also the expected frequency of occurring binding events. And most importantly, the scGar1 used in these experiments was partially degraded (see Figure 5.4b, page 131) and it might be possible that any of the observed scGar1 fragments is not functional, resulting in potential catalytically inactive H/ACA-RNPs despite complete assembly. Thus, further optimization is required to make well-founded assumptions about which of these binding events represent actual pseudouridylation processes. Nevertheless, the obtained data from the substrate turnover experiments show the anticipated trends and can be put into context with existing literature. Hence, it is very likely that pseudouridylation events of functional H/ACA-RNPs were observed and the assumptions about the different kinetic steps for RNA guided Ψ-formation are correct.

5.7. Conclusion

RNP Activity

In the following activity and smFRET data of the H/ACA-RNP constructs from *Saccharomyces cerevisiae* will be concluded and compared to the results of the archaeal Ψase. Full-length and shortened 5'- and 3'-hairpins were tested for enzymatic function and investigated for structural dynamics. Single turnover experiments showed high activity for the 5'-constructs (labeled and unlabeled alike), while for none of the 3'-RNAs formation of Ψ could be detected (see Figure 5.15, page 144). Activity tests (performed by Christin Fuks and Sven Trucks) of the full-length snR81 RNA showed efficient pseudouridylation of the 3'-target RNA. Hence, the 3'-hairpin requires presence of the 5'-hairpin to effectively guide uridine isomerization. In multiple turnover experiments the impact of guide RNA labeling, as well as protein (Nhp2 or scGar1) omission and absence of the loop-structure (construct H5d) on enzymatic activity was tested (see Table 5.3, page 163). Most importantly, insertion of the organic fluorophores had no major impact on pseudouridylation efficiency of the designed construct at the temperatures tested (20 °C and 30 °C), validating that the dyes did not interfere with enzymatic function significantly. Full-length (H5) and truncated (H5d) guide RNAs showed similar initial velocities; 3.4 min^{-1} and 2.7 min^{-1} respectively, pointing towards the assumption that the conserved loop structures in snoRNAs are not essential for pseudouridylation. Interestingly, removal of Nhp2 had a stronger effect on catalytic activity for H5 (v_{start} = 0.20 min^{-1}) in comparison to the partial hairpin H5d (v_{start} = 0.56 min^{-1}), suggesting that the main function of Nhp2 is to stabilize the upper domain of the H/ACA-RNP. Omission of scGar1 led to a complete loss of catalytic activity, demonstrating that this protein is essential for enzyme function in eukaryotes. For the thermophile Ψases from *P. furiosus* the opposite behavior was observed: the stabilization effect of L7Ae (eukaryotic counterpart to Nhp2) seemed to have a more significant influence on pseudouridylation than the proposed product releasing function of archaeal Gar1. Considering the vast deviation in native temperatures of both organisms (20 - 30 °C for yeast and 70 - 100 °C for *P. furiosus*), the high importance of scGar1 in eukaryotic Ψ-formation presumably originates from its ability to promote target nucleotide positioning and product release[63,64], which seems to be more vital at modest temperatures in RNA guided pseudouridylation than the stabilizing effect of Nhp2.

RNP Assembly

Eukaryotic H/ACA-RNP assembly was investigated in the same manner as was done for the archaeal sRNA. FRET labels were attached in a similar fashion as for the archaeal construct *5'-FRET-H/ACA* (in between lower and upper stem of the guide RNA), in order to compare eukaryotic and archaeal RNP formation. As for the archaeal H/ACA-RNA, two states were observed for the full-length 5'-hairpin (*H5-FRET*); a closed state (E-FRET = 0.59) and an open state (E-FRET = 0.40). Here again, efficient population of the open conformation was only observed in presence of all four proteins 'WNCG' (see Figure 5.27a, page 165). The comparable FRET-behavior in the assembly experiments for the partial snoRNA and the sRNA suggests a similar H/ACA-RNP formation mechanism throughout Eukarya and Archaea. Apparently, an efficient opening of the substrate binding site in the guide RNA, which was proposed to be necessary in archaeal Ψases for effective substrate recruitment, is only achieved for complete H/ACA-RNPs.

For the truncated 5'-hairpin (*H5d-FRET*) an apparent population of the open state was not observed, as even in presence of WNCG only 16% of detected particles populated the open conformation (see Figure 5.27b, page 165). Two likely explanations for this observation come into question, one being that removal of parts of the upper stem and the loop structure reduces protein binding affinity to the snoRNA. The experiments were performed in single molecule conditions – meaning low concentrations (500 pM guide RNA). Hence, even minor changes in RNA-protein binding affinity might lead to detectable losses of protein bound RNAs, and since only the RNA is labeled in this assay, a direct quantification of protein free versus protein bound RNAs cannot be made. The other possibility might be that the upper domains are required for the structural rearrangement of the guide RNA. This however would also mean that the open state is not required for effective pseudouridylation, since H5d showed near as high activity as H5. Which of these options is accurate could not be confirmed in this thesis.

The standalone 3'-constructs (*H3-FRET* and *H3d-FRET*) showed a different behavior in the smFRET histograms in comparison to the 5'-hairpins. Addition of WNCG (as well as scNCG) led to the formation of a new high-FRET population (see Figure 5.28, page 167), termed packed state (E-FRET 0.82), most likely representing compaction of pseudouridylation pocket between upper and lower stem of the guide RNA. This contrary observation for the 3'-constructs might be

an explanation for the inactivity of the standalone hairpins in the single turnover assays. Presumably, the packed state represents a misalignment of the partial 3'-RNPs, leading to a loss of essential functionalities such as substrate recruitment of the Ψase. This further means that presence of the 5'-hairpin promotes correct 3'-RNP assembly, since effective pseudouridylation of 3'-substrate was observed for the complete snR81 RNP.

Substrate Turnover

Substrate recruitment capability of Cy3-labeled guide RNAs (*H5-Cy3* and *H5d-Cy3*) was tested with 5fU-scSub-Cy5 (see Figure 5.24, page 154). Due to the fact that this target nucleotide yielded a significantly lower reaction rate in the activity experiments, the observation of substrate-enzyme complexes should be accessible for a longer period of time. The RNAs were annealed in presence of scNCG or WNCG to investigate potential impacts of Nhp2-presence for complete or shortened 5'-hairpin. Overall, substrate-bound complexes were observed showing that the labeled immobilized RNPs were capable of target RNA binding. Removal of Nhp2 led to a notable loss of detected substrate-enzyme complexes, giving further evidence for the assumption that the presence of the proteins stabilizes the H/ACA-RNP and therefore promotes substrate recruitment.

In order to investigate structural rearrangements within the snoRNA during pseudouridylation, unlabeled target RNA (U-scSub) was added to the immobilized guide RNAs (*H5-FRET* and *H5d-FRET*) in presence of WNCG. No hints on structural dynamics were found; neither in the histograms (similar FRET-distribution as in absence of U-scSub), nor in in single molecule fluorescence trajectories (see Figure 5.23, page 153). Once the guide RNA formed a conformation the particle remained in this state, indicating that no major rearrangements between lower and upper stem of the snoRNA occur during substrate binding, turnover and product release. Comparing these results to the various observed states (free, docked and compact) for the archaeal H/ACA-RNA, the passiveness of the snoRNA during pseudouridylation was not anticipated. A possible explanation for these dissimilar observations might be within the different labeling sites. The archaeal sRNA was labeled directly next to the pseudouridylation pocket (and at either 5'-end or within the loop). *H5-* and *H5d-FRET* however, were labeled within the upper stem four base pairs above the substrate binding site (and at the 5'-end). It is likely that after complex assembly most of the structural dynamics happen directly at the pseudouridylation

pocket and cannot be observed by the chosen labeling site within the upper stem. Another reasonable explanation might be that the snoRNA is overall less dynamic than the thermophile sRNA of Archaea and the eukaryotic H/ACA-RNAs remain passive during Ψ-formation. Which of these two hypotheses is correct cannot be validated from the obtained data, different constructs and further investigations would be required.

Finally, substrate turnover was investigated for *H5-Cy3* via usage of labeled target RNAs carrying U-, Ψ- and 5fU as target nucleotides (X-scSub-Cy5). These substrate analogues were used to mimic different states during pseudouridylation, as was done for the archaeal H/ACA-RNP. Furthermore, absence and presence of scGar1 were tested to see potential impacts of this protein on target RNA binding and release. In total, omission of scGar1 led to a notable decrease in detected target RNA-bound RNPs (see Figure 5.25, page 155), indicating that besides product release this protein also promotes substrate recruitment in a direct or indirect manner. Investigation of fluorescence trajectories yielded the observation of spontaneous FRET-events, representing target RNA binding to the RNP. These binding events were categorized in different dwell times and are summarized in Figure 5.29 on page 170. For the full RNP, U- and Ψ-scSub-Cy5 showed a very similar frequency-pattern in dwell times, with close to equal distribution of the categorized binding events (from ultra-short to very long). Usage of 5fU-scSub-Cy5 as well as omission of scGar1 led to a significant increase in the relative frequency of very long (> 10 s) dwell times. This behavior is in overall agreement with the fact that 5fU slows down catalytic turnover and that scGar1 promotes RNA release, since in both cases longer dwell times of substrate-enzyme complex would be expected. Due to the fact that for each sample all events were observed, a clear statement on which of the categorized binding events represents an actual complete pseudouridylation process cannot be made. Nevertheless, the results allowed educated guesses for individual kinetic steps in RNA guided pseudouridylation. To begin with, target RNA binding seems to be independent of nucleotide identity at the modification site (U or Ψ), since no significant alterations between U- and Ψ-scSub-Cy5 in dwell times for the complete RNP were observed. For the same reason, it can be assumed that the protein catalyzed isomerization is a fast process – on a millisecond to sub millisecond time scale. Furthermore, absence of scGar1 significantly reduces the target RNA off-rate, yielding longer dwell times for the substrate analogues. Considering these assumptions, the rate limiting step in RNA guided

pseudouridylation should be the formation of an active enzyme-substrate complex, for which proper assembly of a functional H/ACA-RNP is required.

Chapter VI

Summary and Outlook

In this thesis the aim was to get new insights into the structural dynamics of H/ACA-ribonucleoproteins, an enzyme class whose main function is to generate Ψ in rRNAs and snRNAs. The major knowledge about these complexes has been derived from single hairpin crystal structures of archaeal sRNPs[52-56], while structure and function of the bipartite eukaryotic snoRNPs are less well understood. For investigation of archaeal and eukaryotic RNA guided pseudouridylation, single molecule FRET spectroscopy was the method of choice, since this technique allows direct observation of structural changes under physiological conditions. This thesis was divided into two projects: 1) an archaeal H/ACA-RNP from the thermophile organism *Pyrococcus furiosus*, and 2) a eukaryotic H/ACA-RNP from *Saccharomyces cerevisiae*, which is active at ambient temperatures. The focus was on the structural dynamics of the H/ACA-RNAs. To gain new insights into structure and function of this enzyme class, complex assembly, substrate recruitment and individual reactive states were investigated by smFRET spectroscopy. Furthermore, biochemical binding and activity assays with radioactively labeled target RNAs were performed to validate RNP formation and to study potential impacts of individual proteins on RNA guided pseudouridylation. For these purposes, archaeal and eukaryotic H/ACA-proteins as well as labeled RNAs were prepared, purified and tested for biological function. The RNA binding capability of the proteins was verified in EMSAs. For the archaeal constructs the fluorescence anisotropies of individual dyes and labeling sites were determined, showing that RNP formation had no major influence on fluorophore rotation. Catalytic activity of labeled guide RNAs was verified in single and multiple turnover experiments, demonstrating that insertion of the dyes had no significant impact on enzymatic function. Hence, preparation, modification and purification of the H/ACA-RNP components were overall successful and yielded highly active Ψases for smFRET microscopy.

H/ACA-RNP Assembly and Protein Function

In this study it was shown how individual proteins contribute to RNP formation and catalytic function. Archaeal and eukaryotic H/ACA-RNP assembly seem to occur in a similar manner. Upon protein binding the pseudouridylation pocket of the guide RNA undergoes a structural rearrangement – representing an opening of the substrate binding site – which appears to be necessary for substrate recruitment in archaeal sRNAs. Efficient opening of the H/ACA-RNA was only observed in the presence of all proteins, suggesting that even the presence of Gar1 – the only protein that is not in direct contact with the guide RNA[52,54]– promotes complex formation.

However, in recruitment assays Gar1 was not found to be essential for substrate recruitment, unlike L7Ae and Nop10; providing experimental evidence for the assumption that L7Ae promotes correct substrate positioning in the RNP[53]. Furthermore, this thesis gives new insights into the role and function of the smallest of the four proteins, Nop10. Anchored and bound in between L7Ae (binding the kink-loop), Cbf5 (the active enzyme) and the guide RNA, this protein seems to stabilize the overall RNP structure, which is required for effective target RNA binding.

The role of eukaryotic H/ACA-proteins in pseudouridylation was investigated in multiple turnover assays using different guide RNA constructs (containing either the canonical full-length loop or an artificially introduced tetra-loop in the upper stem of the 5'-hairpin). Unlike scGar1, Nhp2 was not found to be essential for snoRNP function. Apparently, the latter protein promotes catalytic activity of the Ψase via stabilization of the large loop structure in the upper stem of the snoRNA. This Nhp2-dependent effect appears to be not as necessary for the tetra-loop construct as for the full-length 5'-hairpin, most likely due to the fact that the intramolecular tetra-loop stabilizes the hairpin structure and compensates for the lack of Nhp2 in the naturally occurring snoRNA. Removal of scGar1 however led to a complete inactivity of all RNP constructs, probably due to the loss of the presumed substrate positioning and product releasing functions of this protein[63,64]. Furthermore, it was observed that individual hairpins in the bipartite snoRNP depend on presence of the other hairpin for accurate complex assembly and function. The standalone 3'-hairpin displayed no catalytic activity (in contrast to the bipartite snoRNP) in single turnover experiments, which can probably be attributed to the observed packing instead of opening for the pseudouridylation pocket after RNP formation. Regarding future assembly experiments, it would be interesting to see how presence of one of the hairpins influences assembly of the other hairpin. This could be achieved by enzymatic ligation of a FRET-labeled construct to an unlabeled guide RNA and repetition of the step-by-step assembly investigation.

Catalytic States during Pseudouridylation

Reactive states during substrate binding, uridine isomerization and product release were investigated in a different manner for archaeal and eukaryotic pseudouridylation. Since the required temperatures needed to simulate native conditions for the Ψase from *P. furiosus* were not accessible by smFRET, individual states during catalysis (reactive, release and uncatalytic state) were mimicked by atomic mutagenesis at the target nucleotide and attempted to be trapped by rapid cool down during incubation. In these experiments, evidence for a presumed release

state in the sRNP was found. Apparently, the to-be-released target RNA experiences a conformational rearrangement before it is removed from the RNP. Furthermore, the upper stem of the archaeal H/ACA-RNP showed the ability to undergo a dense compaction when a substrate analogue carrying C at the target site (mimicking a release state) was bound. Whether this compaction represents the presumed releasing function of the sRNP or a possible discrimination ability of the Ψase to mismatched RNAs cannot be stated in this thesis. A possible way to target this question would be to perform smFRET investigations at higher temperatures (e.g. 37 °C, where the prepared complex was found to be active), which is however demanding considering the surface immobilization in these experiments. An increase to such temperatures might also lead to other challenges, such as fluorophore stability and proper function of the oxygen scavenging system. Furthermore, the mismatched target RNA proves to be an interesting target for crystallization studies of the archaeal H/ACA-RNP, since the dense compaction in the upper stem has not been reported yet.

For the eukaryotic H/ACA-RNP it was attempted to observe actual pseudouridylation events under the microscope, since the prepared complexes were active at modest conditions (20 °C). In the guide RNA no rearrangements were observed upon substrate addition, indicating that the snoRNAs remains in one conformation after RNP assembly during the reaction process. Furthermore, usage of labeled substrate analogues carrying different target nucleotides allowed the observation of various binding events (ranging from dwell times of a few hundred milliseconds to over a minute). A definite classification of these binding events to individual processes in Ψ-formation was not possible. Still, the observed trends indicate that functional H/ACA-RNPs and actual uridine isomerization were observed, giving confidence for future experiments to obtain kinetical data and dissect the individual processes of RNA guided pseudouridylation by smFRET spectroscopy. Before this can be achieved, several optimizations have to be made; probably the most important point being the preparation of stable scGar1. Recombinantly expressed scGar1 is known to be prone to degradation[59], as was also observed for the used scNCG complex in this thesis. This protein was shown to be essential for activity, and even partial degradation might lead to inactive RNPs which cannot be distinguished from active complexes by smFRET. Thus, attempts were made to produce more stable scGar1 constructs via variation of the N- or C-terminus containing the HisTag. Analytical gels of these constructs after affinity chromatography showed significantly less protein degradation. Hence, in future

experiments these constructs have to be tested for stability and activity and are promising candidates for further smFRET experiments. Furthermore, with the active biomolecules at hand, a stable scGar1 construct and the full-length transcribed snoRNA cryo-EM studies of the snoRNP are possible. These could be used to obtain the first high-resolution structure of a functional and complete eukaryotic snoRNP.

Conclusively it can be stated that this thesis gave new insights into structure, assembly and dynamics of H/ACA-RNPs. For the better understood Ψase from *P. furiosus* results were put into literature context, and new conclusions regarding RNA and protein functions were made. The acquired expertise in preparation of archaeal H/ACA-RNPs was successfully transferred to the snoRNPs from yeast and allowed a comparison of archaeal and eukaryotic RNA guided pseudouridylation. Furthermore, the accomplishments in the second project lay the foundation for promising experiments regarding the still mysterious process of eukaryotic uridine isomerization and contribute to an overall understanding of the formation of the most important RNA modification.

List of References

1 Busch, H., Reddy, R., Rothblum, L. & Choi, Y. C. Snrnas, Snrnps, and Rna Processing. *Annual review of biochemistry* **51**, 617-654, doi:10.1146/annurev.bi.51.070182.003153 (1982).

2 Karijolich, J. & Yu, Y. T. Spliceosomal Snrna Modifications and Their Function. *RNA biology* **7**, 192-204, doi:10.4161/rna.7.2.11207 (2010).

3 Valadkhan, S. & Gunawardane, L. S. Role of Small Nuclear Rnas in Eukaryotic Gene Expression. *Essays in biochemistry* **54**, 79-90, doi:10.1042/bse0540079 (2013).

4 Lee, R. C., Feinbaum, R. L. & Ambros, V. The C. Elegans Heterochronic Gene Lin-4 Encodes Small Rnas with Antisense Complementarity to Lin-14. *Cell* **75**, 843-854, doi:10.1016/0092-8674(93)90529-y (1993).

5 He, L. & Hannon, G. J. Micrornas: Small Rnas with a Big Role in Gene Regulation. *Nature reviews. Genetics* **5**, 522-531, doi:10.1038/nrg1379 (2004).

6 Bartel, D. P. Micrornas: Target Recognition and Regulatory Functions. *Cell* **136**, 215-233, doi:10.1016/j.cell.2009.01.002 (2009).

7 Fabian, M. R., Sonenberg, N. & Filipowicz, W. Regulation of Mrna Translation and Stability by Micrornas. *Annual review of biochemistry* **79**, 351-379, doi:10.1146/annurev-biochem-060308-103103 (2010).

8 Ambros, V. & Ruvkun, G. Recent Molecular Genetic Explorations of Caenorhabditis Elegans Micrornas. *Genetics* **209**, 651-673, doi:10.1534/genetics.118.300291 (2018).

9 Tucker, B. J. & Breaker, R. R. Riboswitches as Versatile Gene Control Elements. *Current opinion in structural biology* **15**, 342-348, doi:10.1016/j.sbi.2005.05.003 (2005).

10 Mironov, A. S. *et al.* Sensing Small Molecules by Nascent Rna: A Mechanism to Control Transcription in Bacteria. *Cell* **111**, 747-756, doi:10.1016/s0092-8674(02)01134-0 (2002).

11 Winkler, W., Nahvi, A. & Breaker, R. R. Thiamine Derivatives Bind Messenger Rnas Directly to Regulate Bacterial Gene Expression. *Nature* **419**, 952-956, doi:10.1038/nature01145 (2002).

12 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, doi:10.1093/nar/gkx110 (2017).

13 Reining, A. *et al.* Three-State Mechanism Couples Ligand and Temperature Sensing in Riboswitches. *Nature* **499**, 355-359, doi:10.1038/nature12378 (2013).

14 Dupuis-Sandoval, F., Poirier, M. & Scott, M. S. The Emerging Landscape of Small Nucleolar Rnas in Cell Biology. *Wiley interdisciplinary reviews. RNA* **6**, 381-397, doi:10.1002/wrna.1284 (2015).

15 Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N. & Altman, S. The Rna Moiety of Ribonuclease P Is the Catalytic Subunit of the Enzyme. *Cell* **35**, 849-857, doi:10.1016/0092-8674(83)90117-4 (1983).

16 Kuo, M. Y., Sharmeen, L., Dinter-Gottlieb, G. & Taylor, J. Characterization of Self-Cleaving Rna Sequences on the Genome and Antigenome of Human Hepatitis Delta Virus. *Journal of virology* **62**, 4439-4444 (1988).

17 Prody, G. A., Bakos, J. T., Buzayan, J. M., Schneider, I. R. & Bruening, G. Autolytic Processing of Dimeric Plant Virus Satellite Rna. *Science* **231**, 1577-1580, doi:10.1126/science.231.4745.1577 (1986).

18 Kim, S. H. *et al.* Three-Dimensional Structure of Yeast Phenylalanine Transfer Rna: Folding of the Polynucleotide Chain. *Science* **179**, 285-288, doi:10.1126/science.179.4070.285 (1973).

19 Boccaletto, P. *et al.* Modomics: A Database of Rna Modification Pathways. 2017 Update. *Nucleic acids research* **46**, D303-D307, doi:10.1093/nar/gkx1030 (2018).

20 Lyons, S. M., Fay, M. M. & Ivanov, P. The Role of Rna Modifications in the Regulation of Trna Cleavage. *FEBS letters* **592**, 2828-2844, doi:10.1002/1873-3468.13205 (2018).

21 Lewis, C. J., Pan, T. & Kalsotra, A. Rna Modifications and Structures Cooperate to Guide Rna-Protein Interactions. *Nature reviews. Molecular cell biology* **18**, 202-210, doi:10.1038/nrm.2016.163 (2017).

22 Chow, C. S., Lamichhane, T. N. & Mahto, S. K. Expanding the Nucleotide Repertoire of the Ribosome with Post-Transcriptional Modifications. *ACS chemical biology* **2**, 610-619, doi:10.1021/cb7001494 (2007).

23 Klassen, R. *et al.* Loss of Anticodon Wobble Uridine Modifications Affects Trna(Lys) Function and Protein Levels in Saccharomyces Cerevisiae. *PloS one* **10**, e0119261, doi:10.1371/journal.pone.0119261 (2015).

24 Konevega, A. L. *et al.* Purine Bases at Position 37 of Trna Stabilize Codon-Anticodon Interaction in the Ribosomal a Site by Stacking and Mg2+-Dependent Interactions. *Rna* **10**, 90-101, doi:10.1261/rna.5142404 (2004).

25 Zhao, X. *et al.* Fto-Dependent Demethylation of N6-Methyladenosine Regulates Mrna Splicing and Is Required for Adipogenesis. *Cell research* **24**, 1403-1419, doi:10.1038/cr.2014.151 (2014).

26 Dominissini, D. *et al.* Topology of the Human and Mouse M6a Rna Methylomes Revealed by M6a-Seq. *Nature* **485**, 201-206, doi:10.1038/nature11112 (2012).

27 Polikanov, Y. S., Melnikov, S. V., Soll, D. & Steitz, T. A. Structural Insights into the Role of Rrna Modifications in Protein Synthesis and Ribosome Assembly. *Nature structural & molecular biology* **22**, 342-344, doi:10.1038/nsmb.2992 (2015).

28 Decatur, W. A. & Fournier, M. J. Rrna Modifications and Ribosome Function. *Trends in biochemical sciences* **27**, 344-351, doi:10.1016/s0968-0004(02)02109-6 (2002).

29 Zebarjadian, Y., King, T., Fournier, M. J., Clarke, L. & Carbon, J. Point Mutations in Yeast Cbf5 Can Abolish in Vivo Pseudouridylation of Rrna. *Molecular and cellular biology* **19**, 7461-7472, doi:10.1128/mcb.19.11.7461 (1999).

30 Tollervey, D., Lehtonen, H., Jansen, R., Kern, H. & Hurt, E. C. Temperature-Sensitive Mutations Demonstrate Roles for Yeast Fibrillarin in Pre-Rrna Processing, Pre-Rrna

Methylation, and Ribosome Assembly. *Cell* **72**, 443-457, doi:10.1016/0092-8674(93)90120-f (1993).

31 Sloan, K. E. *et al.* Tuning the Ribosome: The Influence of Rrna Modification on Eukaryotic Ribosome Biogenesis and Function. *RNA biology* **14**, 1138-1152, doi:10.1080/15476286.2016.1259781 (2017).

32 Keith, G. Mobilities of Modified Ribonucleotides on Two-Dimensional Cellulose Thin-Layer Chromatography. *Biochimie* **77**, 142-144, doi:10.1016/0300-9084(96)88118-1 (1995).

33 Ofengand, J. *et al.* Pseudouridines and Pseudouridine Synthases of the Ribosome. *Cold Spring Harbor symposia on quantitative biology* **66**, 147-159, doi:10.1101/sqb.2001.66.147 (2001).

34 Carlile, T. M. *et al.* Pseudouridine Profiling Reveals Regulated Mrna Pseudouridylation in Yeast and Human Cells. *Nature* **515**, 143-146, doi:10.1038/nature13802 (2014).

35 Davis, F. F. & Allen, F. W. Ribonucleic Acids from Yeast Which Contain a Fifth Nucleotide. *The Journal of biological chemistry* **227**, 907-915 (1957).

36 Davis, D. R. Stabilization of Rna Stacking by Pseudouridine. *Nucleic acids research* **23**, 5020-5026, doi:10.1093/nar/23.24.5020 (1995).

37 Kierzek, E. *et al.* The Contribution of Pseudouridine to Stabilities and Structure of Rnas. *Nucleic acids research* **42**, 3492-3501, doi:10.1093/nar/gkt1330 (2014).

38 Hudson, G. A., Bloomingdale, R. J. & Znosko, B. M. Thermodynamic Contribution and Nearest-Neighbor Parameters of Pseudouridine-Adenosine Base Pairs in Oligoribonucleotides. *Rna* **19**, 1474-1482, doi:10.1261/rna.039610.113 (2013).

39 Schroeder, K. T., Skalicky, J. J. & Greenbaum, N. L. Nmr Spectroscopy of Rna Duplexes Containing Pseudouridine in Supercooled Water. *Rna* **11**, 1012-1016, doi:10.1261/rna.2270205 (2005).

40 Spenkuch, F., Motorin, Y. & Helm, M. Pseudouridine: Still Mysterious, but Never a Fake (Uridine)! *RNA biology* **11**, 1540-1554, doi:10.4161/15476286.2014.992278 (2014).

41 Penzo, M. *et al.* Human Ribosomes from Cells with Reduced Dyskerin Levels Are Intrinsically Altered in Translation. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* **29**, 3472-3482, doi:10.1096/fj.15-270991 (2015).

42 Yang, C., McPheeters, D. S. & Yu, Y. T. Psi35 in the Branch Site Recognition Region of U2 Small Nuclear Rna Is Important for Pre-Mrna Splicing in Saccharomyces Cerevisiae. *The Journal of biological chemistry* **280**, 6655-6662, doi:10.1074/jbc.M413288200 (2005).

43 Ge, J. & Yu, Y. T. Rna Pseudouridylation: New Insights into an Old Modification. *Trends in biochemical sciences* **38**, 210-218, doi:10.1016/j.tibs.2013.01.002 (2013).

44 Rintala-Dempsey, A. C. & Kothe, U. Eukaryotic Stand-Alone Pseudouridine Synthases - Rna Modifying Enzymes and Emerging Regulators of Gene Expression? *RNA biology* **14**, 1185-1196, doi:10.1080/15476286.2016.1276150 (2017).

45 Aittaleb, M. *et al.* Structure and Function of Archaeal Box C/D Srnp Core Proteins. *Nature structural biology* **10**, 256-263, doi:10.1038/nsb905 (2003).

46 Nolivos, S., Carpousis, A. J. & Clouet-d'Orval, B. The K-Loop, a General Feature of the Pyrococcus C/D Guide Rnas, Is an Rna Structural Motif Related to the K-Turn. *Nucleic acids research* **33**, 6507-6514, doi:10.1093/nar/gki962 (2005).

47 Yip, W. S., Shigematsu, H., Taylor, D. W. & Baserga, S. J. Box C/D Srna Stem Ends Act as Stabilizing Anchors for Box C/D Di-Srnps. *Nucleic acids research* **44**, 8976-8989, doi:10.1093/nar/gkw576 (2016).

48 Lapinaite, A. *et al.* The Structure of the Box C/D Enzyme Reveals Regulation of Rna Methylation. *Nature* **502**, 519-523, doi:10.1038/nature12581 (2013).

49 Lin, J. *et al.* Structural Basis for Site-Specific Ribose Methylation by Box C/D Rna Protein Complexes. *Nature* **469**, 559-563, doi:10.1038/nature09688 (2011).

50 Xue, S. *et al.* Structural Basis for Substrate Placement by an Archaeal Box C/D Ribonucleoprotein Particle. *Molecular cell* **39**, 939-949, doi:10.1016/j.molcel.2010.08.022 (2010).

51 Yu, G., Zhao, Y. & Li, H. The Multistructural Forms of Box C/D Ribonucleoprotein Particles. *Rna* **24**, 1625-1633, doi:10.1261/rna.068312.118 (2018).

52 Li, L. & Ye, K. Crystal Structure of an H/Aca Box Ribonucleoprotein Particle. *Nature* **443**, 302-307, doi:10.1038/nature05151 (2006).

53 Liang, B., Xue, S., Terns, R. M., Terns, M. P. & Li, H. Substrate Rna Positioning in the Archaeal H/Aca Ribonucleoprotein Complex. *Nature structural & molecular biology* **14**, 1189-1195, doi:10.1038/nsmb1336 (2007).

54 Duan, J., Li, L., Lu, J., Wang, W. & Ye, K. Structural Mechanism of Substrate Rna Recruitment in H/Aca Rna-Guided Pseudouridine Synthase. *Molecular cell* **34**, 427-439, doi:10.1016/j.molcel.2009.05.005 (2009).

55 Liang, B. *et al.* Structure of a Functional Ribonucleoprotein Pseudouridine Synthase Bound to a Substrate Rna. *Nature structural & molecular biology* **16**, 740-746, doi:10.1038/nsmb.1624 (2009).

56 Zhou, J., Liang, B. & Li, H. Functional and Structural Impact of Target Uridine Substitutions on the H/Aca Ribonucleoprotein Particle Pseudouridine Synthase. *Biochemistry* **49**, 6276-6281, doi:10.1021/bi1006699 (2010).

57 Rozhdestvensky, T. S. *et al.* Binding of L7ae Protein to the K-Turn of Archaeal Snornas: A Shared Rna Binding Motif for C/D and H/Aca Box Snornas in Archaea. *Nucleic acids research* **31**, 869-877, doi:10.1093/nar/gkg175 (2003).

58 Rashid, R. *et al.* Crystal Structure of a Cbf5-Nop10-Gar1 Complex and Implications in Rna-Guided Pseudouridylation and Dyskeratosis Congenita. *Molecular cell* **21**, 249-260, doi:10.1016/j.molcel.2005.11.017 (2006).

59 Li, S. *et al.* Reconstitution and Structural Analysis of the Yeast Box H/Aca Rna-Guided Pseudouridine Synthase. *Genes & development* **25**, 2409-2421, doi:10.1101/gad.175299.111 (2011).

60 Majumder, M., Bosmeny, M. S. & Gupta, R. Structure-Function Relationships of Archaeal Cbf5 During in Vivo Rna-Guided Pseudouridylation. *Rna* **22**, 1604-1619, doi:10.1261/rna.057547.116 (2016).

61 Czudnochowski, N. *et al.* The Mechanism of Pseudouridine Synthases from a Covalent Complex with Rna, and Alternate Specificity for U2605 Versus U2604 between Close Homologs. *Nucleic acids research* **42**, 2037-2048, doi:10.1093/nar/gkt1050 (2014).

62 McDonald, M. K., Miracco, E. J., Chen, J., Xie, Y. & Mueller, E. G. The Handling of the Mechanistic Probe 5-Fluorouridine by the Pseudouridine Synthase Trua and Its Consistency with the Handling of the Same Probe by the Pseudouridine Synthases Trub and Rlua. *Biochemistry* **50**, 426-436, doi:10.1021/bi101737z (2011).

63 Wang, P., Yang, L., Gao, Y. Q. & Zhao, X. S. Accurate Placement of Substrate Rna by Gar1 in H/Aca Rna-Guided Pseudouridylation. *Nucleic acids research* **43**, 7207-7216, doi:10.1093/nar/gkv757 (2015).

64 Yang, X. *et al.* Kinetic and Thermodynamic Characterization of the Reaction Pathway of Box H/Aca Rna-Guided Pseudouridine Formation. *Nucleic acids research* **40**, 10925-10936, doi:10.1093/nar/gks882 (2012).

65 Koo, B. K. *et al.* Structure of H/Aca Rnp Protein Nhp2p Reveals Cis/Trans Isomerization of a Conserved Proline at the Rna and Nop10 Binding Interface. *Journal of molecular biology* **411**, 927-942, doi:10.1016/j.jmb.2011.06.022 (2011).

66 Hamma, T. & Ferre-D'Amare, A. R. Structure of Protein L7ae Bound to a K-Turn Derived from an Archaeal Box H/Aca Srna at 1.8 a Resolution. *Structure* **12**, 893-903, doi:10.1016/j.str.2004.03.015 (2004).

67 Li, H. Unveiling Substrate Rna Binding to H/Aca Rnps: One Side Fits All. *Current opinion in structural biology* **18**, 78-85, doi:10.1016/j.sbi.2007.11.004 (2008).

68 Richard, P. *et al.* A Common Sequence Motif Determines the Cajal Body-Specific Localization of Box H/Aca Scarnas. *The EMBO journal* **22**, 4283-4293, doi:10.1093/emboj/cdg394 (2003).

69 Darzacq, X. *et al.* Cajal Body-Specific Small Nuclear Rnas: A Novel Class of 2'-O-Methylation and Pseudouridylation Guide Rnas. *The EMBO journal* **21**, 2746-2756, doi:10.1093/emboj/21.11.2746 (2002).

70 Jady, B. E. *et al.* Modification of Sm Small Nuclear Rnas Occurs in the Nucleoplasmic Cajal Body Following Import from the Cytoplasm. *The EMBO journal* **22**, 1878-1888, doi:10.1093/emboj/cdg187 (2003).

71 Wang, C. & Meier, U. T. Architecture and Assembly of Mammalian H/Aca Small Nucleolar and Telomerase Ribonucleoproteins. *The EMBO journal* **23**, 1857-1867, doi:10.1038/sj.emboj.7600181 (2004).

72 Caton, E. A., Kelly, E. K., Kamalampeta, R. & Kothe, U. Efficient Rna Pseudouridylation by Eukaryotic H/Aca Ribonucleoproteins Requires High Affinity Binding and Correct Positioning of Guide Rna. *Nucleic acids research* **46**, 905-916, doi:10.1093/nar/gkx1167 (2018).

73 Khanna, M., Wu, H., Johansson, C., Caizergues-Ferrer, M. & Feigon, J. Structural Study of the H/Aca Snornp Components Nop10p and the 3' Hairpin of U65 Snorna. *Rna* **12**, 40-52, doi:10.1261/rna.2221606 (2006).

74 Nguyen, T. H. D. *et al.* Cryo-Em Structure of Substrate-Bound Human Telomerase Holoenzyme. *Nature* **557**, 190-195, doi:10.1038/s41586-018-0062-x (2018).

75 Fatica, A., Dlakic, M. & Tollervey, D. Naf1 P Is a Box H/Aca Snornp Assembly Factor. *Rna* **8**, 1502-1514 (2002).

76 Darzacq, X. *et al.* Stepwise Rnp Assembly at the Site of H/Aca Rna Transcription in Human Cells. *The Journal of cell biology* **173**, 207-218, doi:10.1083/jcb.200601105 (2006).

77 McMahon, M., Contreras, A. & Ruggero, D. Small Rnas with Big Implications: New Insights into H/Aca Snorna Function and Their Role in Human Disease. *Wiley interdisciplinary reviews. RNA* **6**, 173-189, doi:10.1002/wrna.1266 (2015).

78 Meier, U. T. The Many Facets of H/Aca Ribonucleoproteins. *Chromosoma* **114**, 1-14, doi:10.1007/s00412-005-0333-9 (2005).

79 He, J. *et al.* Variable Expression of Dkc1 Mutations in Mice. *Genesis* **47**, 366-373, doi:10.1002/dvg.20509 (2009).

80 Sarek, G., Marzec, P., Margalef, P. & Boulton, S. J. Molecular Basis of Telomere Dysfunction in Human Genetic Diseases. *Nature structural & molecular biology* **22**, 867-874, doi:10.1038/nsmb.3093 (2015).

81 Le Guen, T. *et al.* Human Rtel1 Deficiency Causes Hoyeraal-Hreidarsson Syndrome with Short Telomeres and Genome Instability. *Human molecular genetics* **22**, 3239-3249, doi:10.1093/hmg/ddt178 (2013).

82 Stepanov, G. A. *et al.* Regulatory Role of Small Nucleolar Rnas in Human Diseases. *BioMed research international* **2015**, 206849, doi:10.1155/2015/206849 (2015).

83 Xia, T., Li, N. & Fang, X. Single-Molecule Fluorescence Imaging in Living Cells. *Annual review of physical chemistry* **64**, 459-480, doi:10.1146/annurev-physchem-040412-110127 (2013).

84 Murata, K. & Wolf, M. Cryo-Electron Microscopy for Structural Analysis of Dynamic Biological Macromolecules. *Biochimica et biophysica acta. General subjects* **1862**, 324-334, doi:10.1016/j.bbagen.2017.07.020 (2018).

85 Erni, R., Rossell, M. D., Kisielowski, C. & Dahmen, U. Atomic-Resolution Imaging with a Sub-50-Pm Electron Probe. *Physical review letters* **102**, 096101, doi:10.1103/PhysRevLett.102.096101 (2009).

86 Merk, A. *et al.* Breaking Cryo-Em Resolution Barriers to Facilitate Drug Discovery. *Cell* **165**, 1698-1707, doi:10.1016/j.cell.2016.05.040 (2016).

87 Hell, S. W. & Wichmann, J. Breaking the Diffraction Resolution Limit by Stimulated Emission: Stimulated-Emission-Depletion Fluorescence Microscopy. *Optics letters* **19**, 780-782, doi:10.1364/ol.19.000780 (1994).

88 Rust, M. J., Bates, M. & Zhuang, X. Sub-Diffraction-Limit Imaging by Stochastic Optical Reconstruction Microscopy (Storm). *Nature methods* **3**, 793-795, doi:10.1038/nmeth929 (2006).

89 Betzig, E. *et al.* Imaging Intracellular Fluorescent Proteins at Nanometer Resolution. *Science* **313**, 1642-1645, doi:10.1126/science.1127344 (2006).

90 Sydor, A. M., Czymmek, K. J., Puchner, E. M. & Mennella, V. Super-Resolution Microscopy: From Single Molecules to Supramolecular Assemblies. *Trends in cell biology* **25**, 730-748, doi:10.1016/j.tcb.2015.10.004 (2015).

91 Hell, S. W. Nanoscopy with Focused Light (Nobel Lecture). *Angewandte Chemie* **54**, 8054-8066, doi:10.1002/anie.201504181 (2015).

92 Moerner, W. E. Single-Molecule Spectroscopy, Imaging, and Photocontrol: Foundations for Super-Resolution Microscopy (Nobel Lecture). *Angewandte Chemie* **54**, 8067-8093, doi:10.1002/anie.201501949 (2015).

93 Betzig, E. Single Molecules, Cells, and Super-Resolution Optics (Nobel Lecture). *Angewandte Chemie* **54**, 8034-8053, doi:10.1002/anie.201501003 (2015).

94 Spahn, C. K. *et al.* A Toolbox for Multiplexed Super-Resolution Imaging of the E. Coli Nucleoid and Membrane Using Novel Paint Labels. *Scientific reports* **8**, 14768, doi:10.1038/s41598-018-33052-3 (2018).

95 Kim, J. Y., Kim, C. & Lee, N. K. Real-Time Submillisecond Single-Molecule Fret Dynamics of Freely Diffusing Molecules with Liposome Tethering. *Nature communications* **6**, 6992, doi:10.1038/ncomms7992 (2015).

96 Hengesbach, M., Kim, N. K., Feigon, J. & Stone, M. D. Single-Molecule Fret Reveals the Folding Dynamics of the Human Telomerase Rna Pseudoknot Domain. *Angewandte Chemie* **51**, 5876-5879, doi:10.1002/anie.201200526 (2012).

97 Roy, R., Hohng, S. & Ha, T. A Practical Guide to Single-Molecule Fret. *Nature methods* **5**, 507-516, doi:10.1038/nmeth.1208 (2008).

98 Iino, R., Koyama, I. & Kusumi, A. Single Molecule Imaging of Green Fluorescent Proteins in Living Cells: E-Cadherin Forms Oligomers on the Free Cell Surface. *Biophysical journal* **80**, 2667-2677, doi:10.1016/S0006-3495(01)76236-4 (2001).

99 Hanspach, G., Trucks, S. & Hengesbach, M. Strategic Labelling Approaches for Rna Single-Molecule Spectroscopy. *RNA biology* **16**, 1119-1132, doi:10.1080/15476286.2019.1593093 (2019).

100 Kolb, H. C., Finn, M. G. & Sharpless, K. B. Click Chemistry: Diverse Chemical Function from a Few Good Reactions. *Angewandte Chemie* **40**, 2004-2021, doi:10.1002/1521-3773(20010601)40:11<2004::aid-anie2004>3.3.co;2-x (2001).

101 Tornoe, C. W., Christensen, C. & Meldal, M. Peptidotriazoles on Solid Phase: [1,2,3]-Triazoles by Regiospecific Copper(I)-Catalyzed 1,3-Dipolar Cycloadditions of Terminal Alkynes to Azides. *The Journal of organic chemistry* **67**, 3057-3064, doi:10.1021/jo011148j (2002).

102 Agard, N. J., Prescher, J. A. & Bertozzi, C. R. A Strain-Promoted [3 + 2] Azide-Alkyne Cycloaddition for Covalent Modification of Biomolecules in Living Systems. *Journal of the American Chemical Society* **126**, 15046-15047, doi:10.1021/ja044996f (2004).

103 Laughlin, S. T., Baskin, J. M., Amacher, S. L. & Bertozzi, C. R. In Vivo Imaging of Membrane-Associated Glycans in Developing Zebrafish. *Science* **320**, 664-667, doi:10.1126/science.1155106 (2008).

104 Ning, X., Guo, J., Wolfert, M. A. & Boons, G. J. Visualizing Metabolically Labeled Glycoconjugates of Living Cells by Copper-Free and Fast Huisgen Cycloadditions. *Angewandte Chemie* **47**, 2253-2255, doi:10.1002/anie.200705456 (2008).

105 Horner, K. A., Valette, N. M. & Webb, M. E. Strain-Promoted Reaction of 1,2,4-Triazines with Bicyclononynes. *Chemistry* **21**, 14376-14381, doi:10.1002/chem.201502397 (2015).

106 Blackman, M. L., Royzen, M. & Fox, J. M. Tetrazine Ligation: Fast Bioconjugation Based on Inverse-Electron-Demand Diels-Alder Reactivity. *Journal of the American Chemical Society* **130**, 13518-13519, doi:10.1021/ja8053805 (2008).

107 Kurra, Y. *et al.* Two Rapid Catalyst-Free Click Reactions for in Vivo Protein Labeling of Genetically Encoded Strained Alkene/Alkyne Functionalities. *Bioconjugate chemistry* **25**, 1730-1738, doi:10.1021/bc500361d (2014).

108 Xiang, C. C. *et al.* Amine-Modified Random Primers to Label Probes for DNA Microarrays. *Nature biotechnology* **20**, 738-742, doi:10.1038/nb0702-738 (2002).

109 Scaringe, S. A., Wincott, F. E. & Caruthers, M. H. Novel Rna Synthesis Method Using 5‘-O-Silyl-2‘-O-Orthoester Protecting Groups. *Journal of the American Chemical Society* **120**, 11820-11821, doi:10.1021/ja980730v (1998).

110 Grosjean, H. *DNA and Rna Modification Enzymes*. 1st Edition edn, 550-559 (CRC Press, 2009).

111 Akiyama, B. M. & Stone, M. D. Assembly of Complex Rnas by Splinted Ligation. *Methods in enzymology* **469**, 27-46, doi:10.1016/S0076-6879(09)69002-9 (2009).

112 Ho, C. K., Wang, L. K., Lima, C. D. & Shuman, S. Structure and Mechanism of Rna Ligase. *Structure* **12**, 327-339, doi:10.1016/j.str.2004.01.011 (2004).

113 Nakajima, N. & Ikada, Y. Mechanism of Amide Formation by Carbodiimide for Bioconjugation in Aqueous Media. *Bioconjugate chemistry* **6**, 123-130, doi:10.1021/bc00031a015 (1995).

114 England, T. E. & Uhlenbeck, O. C. 3'-Terminal Labelling of Rna with T4 Rna Ligase. *Nature* **275**, 560-561, doi:10.1038/275560a0 (1978).

115 Tessier, D. C., Brousseau, R. & Vernet, T. Ligation of Single-Stranded Oligodeoxyribonucleotides by T4 Rna Ligase. *Analytical biochemistry* **158**, 171-178, doi:10.1016/0003-2697(86)90606-8 (1986).

116 Nudler, E., Gusarov, I. & Bar-Nahum, G. Methods of Walking with the Rna Polymerase. *Methods in enzymology* **371**, 160-169, doi:10.1016/S0076-6879(03)71011-8 (2003).

117 Liu, Y. *et al.* Synthesis and Applications of Rnas with Position-Selective Labelling and Mosaic Composition. *Nature* **522**, 368-372, doi:10.1038/nature14352 (2015).

118 Liu, Y. *et al.* Incorporation of Isotopic, Fluorescent, and Heavy-Atom-Modified Nucleotides into Rnas by Position-Selective Labeling of Rna. *Nature protocols* **13**, 987-1005, doi:10.1038/nprot.2018.002 (2018).

119 Hanspach, G. *Einzelmolekülspektroskopische Untersuchungen Eines H/Aca-Ribonucleoprotein-Komplexes* Master thesis, J. W. Goethe University, (2015).

120 Ma, X. *et al.* Pseudouridylation of Yeast U2 Snrna Is Catalyzed by Either an Rna-Guided or Rna-Independent Mechanism. *The EMBO journal* **24**, 2403-2413, doi:10.1038/sj.emboj.7600718 (2005).

121 Wu, G., Xiao, M., Yang, C. & Yu, Y. T. U2 Snrna Is Inducibly Pseudouridylated at Novel Sites by Pus7p and Snr81 Rnp. *The EMBO journal* **30**, 79-89, doi:10.1038/emboj.2010.316 (2011).

122 Rueden, C. T. *et al.* Imagej2: Imagej for the Next Generation of Scientific Image Data. *BMC bioinformatics* **18**, 529, doi:10.1186/s12859-017-1934-z (2017).

123 Bokinsky, G. *et al.* Single-Molecule Transition-State Analysis of Rna Folding. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 9302-9307, doi:10.1073/pnas.1133280100 (2003).

124 Kibbe, W. A. Oligocalc: An Online Oligonucleotide Properties Calculator. *Nucleic acids research* **35**, W43-46, doi:10.1093/nar/gkm234 (2007).

125 Lakowicz, J. R. *Principles of Fluorescence Spectroscopy*. (Springer, Boston, MA, 2006).

126 Dale, R. E., Eisinger, J. & Blumberg, W. E. The Orientational Freedom of Molecular Probes. The Orientation Factor in Intramolecular Energy Transfer. *Biophysical journal* **26**, 161-193, doi:10.1016/S0006-3495(79)85243-1 (1979).

127 Gholamalipour, Y., Karunanayake Mudiyanselage, A. & Martin, C. T. 3' End Additions by T7 Rna Polymerase Are Rna Self-Templated, Distributive and Diverse in Character-Rna-Seq Analyses. *Nucleic acids research* **46**, 9253-9263, doi:10.1093/nar/gky796 (2018).

128 Schmidt, A. *Investigation of Structural Dynamics of H/Aca Rnps by Smfret Spectroscopy & Design and Applications of Versatile Small Mole-Cule Dependent Rna of Interest Release Platform* Doctorate thesis, J. W. Goethe University, (2019).

129 Green, N. M. Avidin. *Advances in protein chemistry* **29**, 85-133, doi:10.1016/s0065-3233(08)60411-8 (1975).

130 Rasnik, I., McKinney, S. A. & Ha, T. Nonblinking and Long-Lasting Single-Molecule Fluorescence Imaging. *Nature methods* **3**, 891-893, doi:10.1038/nmeth934 (2006).

131 Samuelsson, T. Interactions of Transfer Rna Pseudouridine Synthases with Rnas Substituted with Fluorouracil. *Nucleic acids research* **19**, 6139-6144, doi:10.1093/nar/19.22.6139 (1991).

132 Patton, J. R. Ribonucleoprotein Particle Assembly and Modification of U2 Small Nuclear Rna Containing 5-Fluorouridine. *Biochemistry* **32**, 8939-8944, doi:10.1021/bi00085a027 (1993).

133 Huang, L., Pookanjanatavip, M., Gu, X. & Santi, D. V. A Conserved Aspartate of Trna Pseudouridine Synthase Is Essential for Activity and a Probable Nucleophilic Catalyst. *Biochemistry* **37**, 344-351, doi:10.1021/bi971874+ (1998).

134 Zhao, X. & Yu, Y. T. Incorporation of 5-Fluorouracil into U2 Snrna Blocks Pseudouridylation and Pre-Mrna Splicing in Vivo. *Nucleic acids research* **35**, 550-558, doi:10.1093/nar/gkl1084 (2007).

135 Hohng, S., Joo, C. & Ha, T. Single-Molecule Three-Color Fret. *Biophysical journal* **87**, 1328-1337, doi:10.1529/biophysj.104.043935 (2004).

136 Sun, Y., Shopova, S. I., Wu, C. S., Arnold, S. & Fan, X. Bioinspired Optofluidic Fret Lasers Via DNA Scaffolds. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 16039-16042, doi:10.1073/pnas.1003581107 (2010).

137 Fourmann, J. B. *et al.* Comparative Study of Two Box H/Aca Ribonucleoprotein Pseudouridine-Synthases: Relation between Conformational Dynamics of the Guide Rna, Enzyme Assembly and Activity. *PloS one* **8**, e70313, doi:10.1371/journal.pone.0070313 (2013).

138 Mörl, M. & Hartmann, R. in *Handbook of Rna Biochemistry* Vol. 2 (eds RK. Hartmann, A. Bindereif, A. Schön, & E. Westhof) Ch. 2, 29-44 (WILEY-VCH, 2014).

139 Helmling, C. *et al.* Rapid Nmr Screening of Rna Secondary Structure and Binding. *Journal of biomolecular NMR* **63**, 67-76, doi:10.1007/s10858-015-9967-y (2015).

140 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, doi:10.1261/rna.058743.116 (2017).

List of Abbreviations

5fU	5-fluoro-uracil
A	Adenine
A	Fluorescence anisotropy
ACE	Bis(2-Acetoxyethoxy)methyl
Amp	Ampicillin
bp	Base pairs
C	Cytidine
c_{final}	Final concentration
CG	Cbf5-Gar1
DNA	Deoxyribonucleic acid
DnL	T4-DNA-ligase
DMSO	Dimethyl sulfoxide
E. coli	*Escherichia coli*
E-FRET	FRET-efficiency
e.g.	Exempli gratia “for example”
EDC	1-Ethyl-3(3-dimethylaminopropyl)carbodiimide
EDTA	Ethylenediaminetetraacetic acid
EM	Electron microscopy
EMSA	Electrophoretic mobility shift assay
EPR	Electron paramagnetic resonance
et al.	*Et alia* “and others”
etc.	Et cetera “and so forth”
FRET	Förster Resonance Energy Transfer
G	Guanine
GFP	Green fluorescent protein
GTP	Guanosine triphosphate
HIV	Hepatitis delta virus
HPLC	High-performance liquid chromatography
IEC	Ion-exchange chromatography
IPTG	Isopropyl β-D-1-thiogalactopyranoside
Kan	Kanamycin
K_d	Dissociation constant
LNCG	L7ae-Nop10-Cbf5-Gar1
m^6A	N^6-methyladenosine
mRNA	Messenger RNA
miRNA	Micro RNA

m/z	Mass-to-charge ratio
NC	Nop10-Cbf5
NCG	Nop10-Cbf5-Gar1
NEB	New England Biolabs
NHS	*N*-Hydroxysuccinimide
Ni-NTA	Nickel(II) nitrilotriacetic acid
NMR	Nuclear magnetic resonance
nt	Nucleotides
NTP	Nucleotide triphosphates
Ψ	Pseudouridine
Ψases	Pseudouridylase
PAA	Polyacrylamide
PAGE	Polyacrylamide gel electrophoresis
PALM	Photo-Activated Localization Microscopy
PCR	Polymerase chain reaction
PDB	Protein database
P. furiosus (or *pf*)	*Pyrococcus furiosus*
PLOR	Position-Selective Labeling of RNA
PUA	Pseudouridine synthase archaeosine transglycosylase
rab	Ribose-abasic
RNA	Ribonucleic acid
RNase	Ribonuclease
RnL	T4-RNA-ligase 2
rRNA	Ribosomal RNA
scNCG	scNop10-scCbf5-scGar1
SDS	Sodium dodecyl sulfate
SEC	Size exclusion chromatography
sRNA	Small RNA
sRNP	Small ribonucleoprotein
snRNA	Small nuclear RNAs
snRNP	Small nuclear ribonucleoprotein
snoRNA	Small nucleolar RNA
snoRNP	Small nucleolar ribonucleoprotein
smFRET	Single molecule FRET
SRM	Super resolution microscopy
STED	Stimulated Emission Depletion Microscopy
STORM	Stochastic Optical Reconstruction Microscopy
$Sub_{1/2}$	Substrate half-life

tRNA	Transfer RNA
TIRF	Total internal reflection fluorescence
TLC	Thin layer chromatography
U	Uridine
v/v	Volume per volume
Vol. eq.	Volume equivalent
v_{start}	Initial velocity
w	Full width at half maximum
WNCG	Nhp2-scNop10-scCbf5-scGar1
WNC	Nhp2-scNop10-scCbf5
wt	Wildtype
w/v	Weight per volume
x_c	Peak center

Appendix

A1 – Genes

Codon Optimized Genes of Archaeal H/ACA-Proteins

Nop10
CAT**ATG**GCGGACTTATTCCGGATCCGCAAATGCCCGAAATGCGGTCGTTACACGCTGAAAG
AGGTGTGTCCGGTATGTGGGGAAAAGACCAAAGTCGCCCATCCACCTCGCTTTAGCCCGG
AAGATCCCTATGGCGAATATCGCCGTCGTTGGAAACGCGAAGTTCTGGGCATTGGTCGCAA
AGAGAAG**TAA**GCGGCCGC

Cbf5
CC**ATG**GCTCGCGATGAAGTACGGCGCATTTTACCCGCCGACATTAAACGCGAAGTTCTCAT
CAAGGACGAAAATGCGGAAACAAATCCGGACTGGGGTTTTCCGCCGGAAAAGCGGCCAAT
CGAGATGCACATTCAGTTTGGGGTCATCAACCTGGATAAACCGCCTGGTCCAACCTCTCAC
GAAGTGGTGGCCTGGATTAAGAAAATCCTCAACCTGGAGAAAGCGGGACATGGCGGAACG
TTGGATCCGAAAGTAAGCGGTGTCCTTCCAGTAGCCCTTGAAAAGGCGACTCGTGTTGTGC
AAGCGTTACTGCCTGCTGGCAAAGAATATGTCGCTCTTATGCATCTGCATGGCGATGTTCC
GGAAGATAAAATCATCCAGGTCATGAAGGAGTTTGAGGGCGAAATCATCCAACGCCCACCC
TTGCGCTCAGCGGTCAAACGTCGTCTGCGCACCCGCAAAGTTTACTACATTGAAGTGCTGG
AAATTGAAGGTCGCGATGTGCTGTTCCGTGTGGGCGTAGAAGCCGGAACCTATATTCGCTC
GCTGATTCATCACATCGGTTTAGCGCTCGGCGTAGGTGCACACATGAGCGAGTTGCGTCGT
ACTCGCTCCGGTCCGTTTAAAGAGGATGAAACCCTGATTACGCTGCATGACCTCGTCGACT
ACTATTATTTCTGGAAAGAAGATGGCATTGAGGAATACTTCCGCAAAGCCATCCAGCCTATG
GAGAAAGCCGTTGAGCATCTGCCGAAAGTGTGGATCAAAGACAGCGCGGTTGCAGCAGTT
ACACATGGGGCGGATCTGGCTGTTCCGGGCATTGCGAAATTACATGCGGGCATTAAACGT
GGGGATCTGGTGGCAATTATGACGTTGAAAGACGAACTGGTCGCACTGGGCAAAGCCATG
ATGACCAGTCAGGAAATGCTGGAAAAGACCAAAGGTATTGCAGTGGATGTGGAAAAAGTGT
TTATGCCGCGTGATTGGTATCCGAAGCTGTGGGAGAAACGTGATCGCTCGCATCACCACCA
CCATCAC**TAA**GCGGCCGC

Gar1
CC**ATG**GAACGCTTAGGGAAAGTTCTCCATTATGCCAAACAAGGCTTCCTGATTGTTCGCAC
CAATTGGGTACCTAGTCTGAACGATCGTGTGGTGGATAAACGGTTGCAGTTTGTGGGCATT
GTCAAAGACGTGTTTGGTCCAGTCAAAATGCCGTATGTTGCGATCAAACCGAAAGTGAGCA
ATCCCGAAATCTACGTAGGTGAAGTCCTGTACGTTGATGAGCGTAAACGCAAAGAATCGCC
GAAAAAGAACAAAGAGAAACGCATGAAGAAGAAGAAACGTCTGAACCGCGCGGCCGCACT
CGAG

Legend:
Restriction sites
Start and stop codons

A2 – Mass Spectrometry Data

Archaeal H/ACA-RNAs

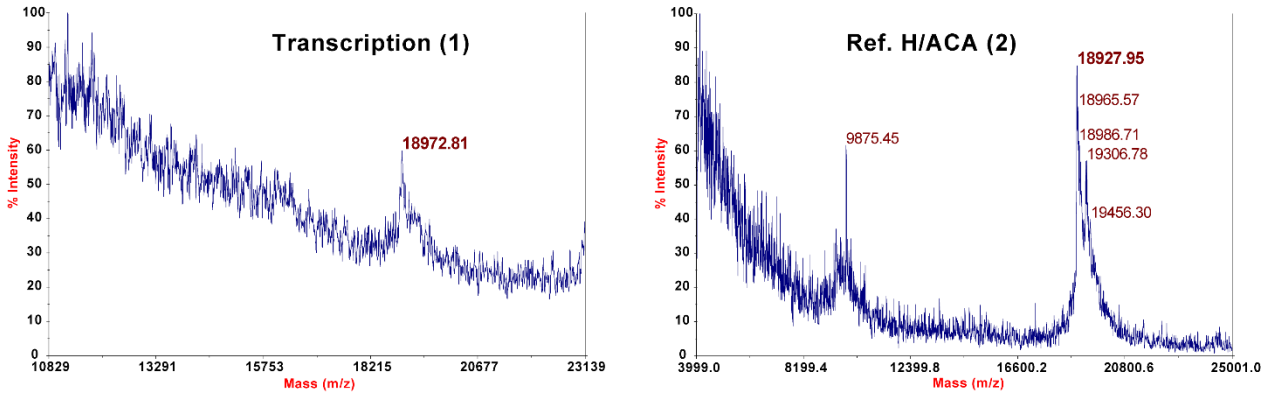


Figure A2.1: MALDI-spectrograms of archaeal H/ACA-RNAs.

X-Sub-bio RNAs

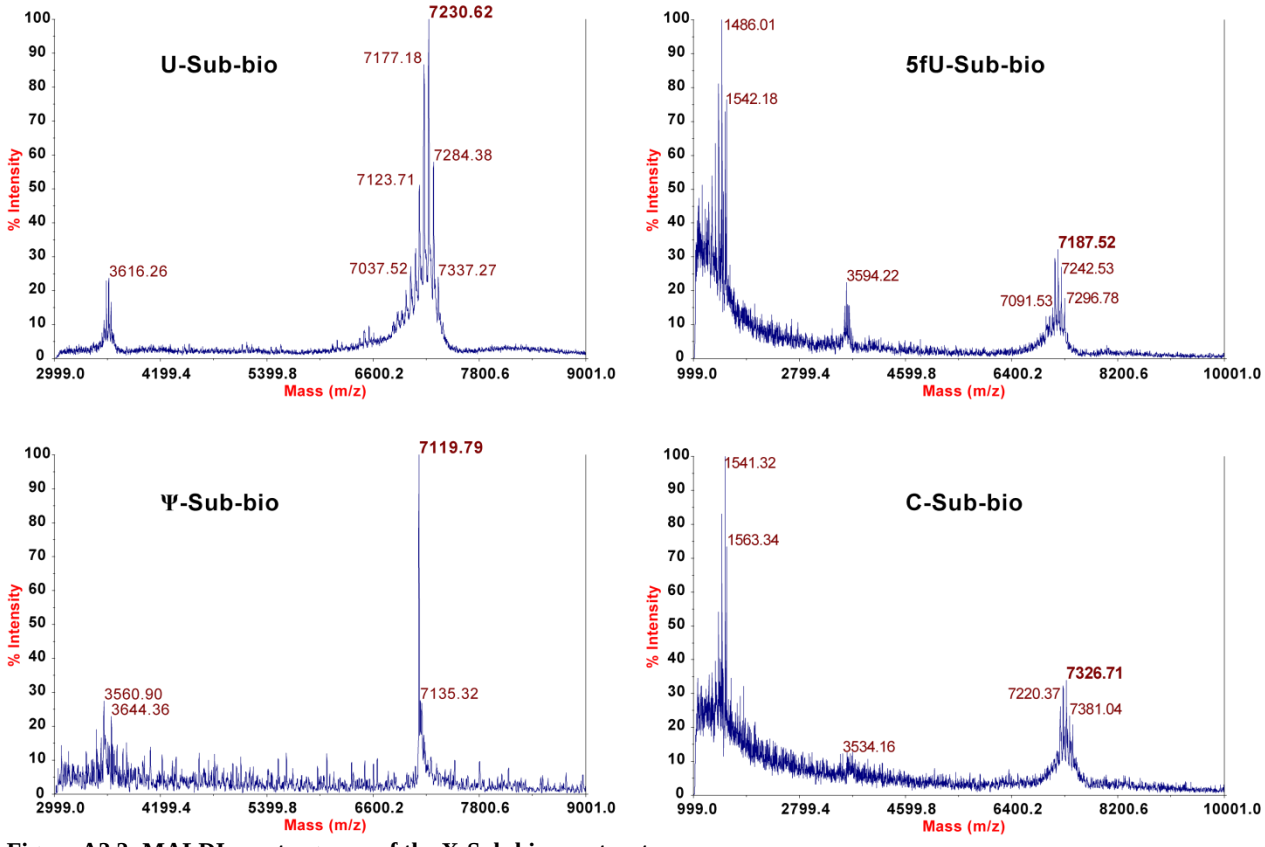


Figure A2.2: MALDI-spectrograms of the X-Sub-bio constructs.

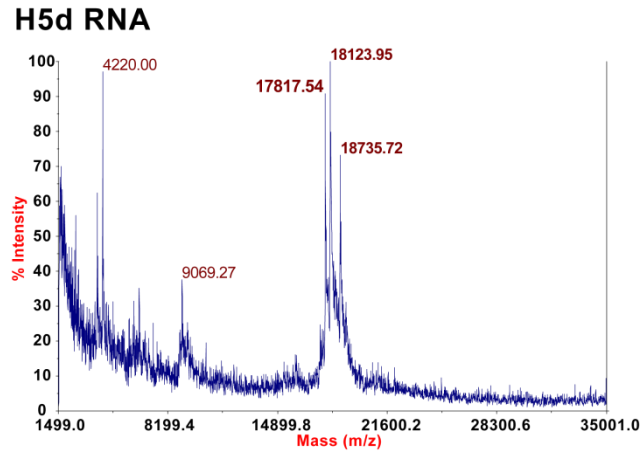


Figure A2.3: MALDI-spectrogram of the H5d construct.

A3 – Supplementary Figures

Project 1 – Archaeal H/ACA-RNP

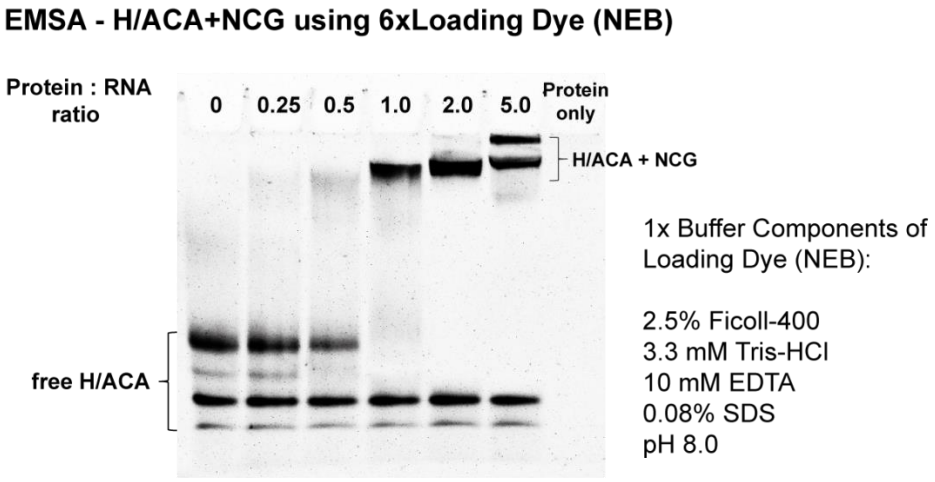


Figure A3.1: Native 8% PAGEl of archaeal wt-H/ACA-RNA with NCG in 0.8x nSEC-buffer. The RNA concentration was adjusted to 1 µM in each sample and NCG concentration ranged from 0 to 5 µM in a 10 µL scale per sample. For annealing samples were incubated at 80 °C for 2 min and subsequently at 50 °C for 60 min. Afterwards 6xLoading Dye (NEB) was added and adjusted to 1x final concentration for sample loading. The EMSA was performed at 60 V for 105 min. RNAs were stained with GelRed.

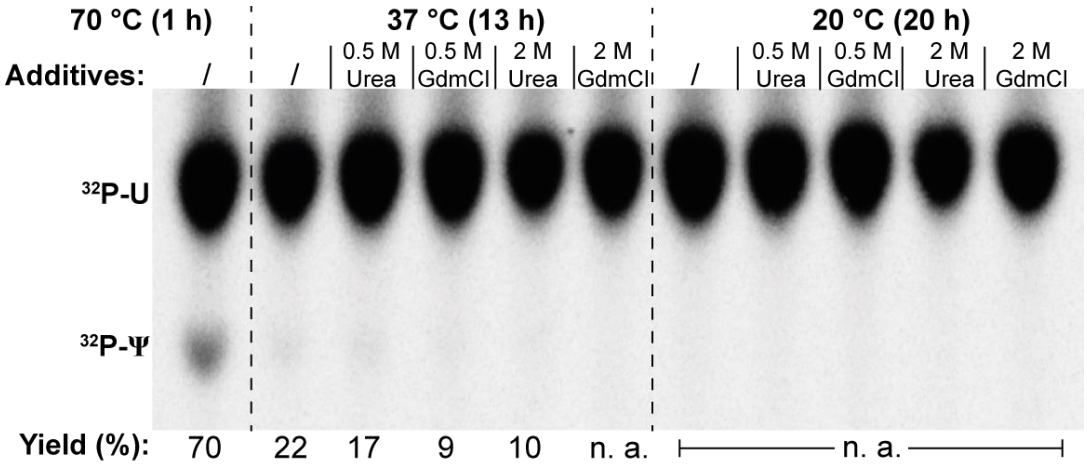


Figure A3.2: TLC of single turnover activity assay of the full archaeal H/ACA-RNPs (Typhoon settings: Phosphor screen). Activity was tested at various temperatures and in absence or presence (0.5 M or 2 M) of different additives (urea or guanidine hydrochloride). This was done to imitate the denaturing conditions at higher temperatures. Addition of these denaturing agents had no positive impact on catalytic activity of the Ψase. At 20 °C no apparent Ψ-formation was detected (n. a.), even in absence of urea or guanidine hydrochloride.

Time/min	Normalized Yield		
	wt	5'-FRET	Loop-FRET
1	0.04280	0.00886	0.00179
5	0.27382	0.10301	0.07859
15	0.50877	0.24015	0.19503
30	0.64453	0.37704	0.30378
60	0.71030	0.48960	0.41646
180	0.74852	0.62928	0.58220

Fitting equation:

$$y = \frac{y_{max} \cdot x}{Sub_{1/2} + x}$$

Initial velocity calculation:

$$v_{start} = \frac{y_{(fit)}}{x_{(fit)}} \cdot 20$$

Figure A3.3: The table (left) summarizes the obtained results from densitometric analysis (ImageJ) of the separated and visualized radioactive nucleotides (by TLC and autoradiography) for the multiple turnover assays of the designed archaeal H/ACA-constructs. The individual time points were plotted with the given fitting equation (Michealis Menten kinetics) in Origin. The fitting point closest to 1 min was taken for calculation of the initial velocity.

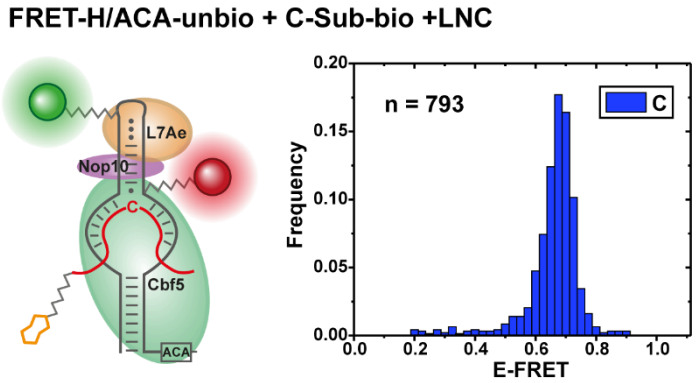


Figure A3.4: smFRET histogram of FRET-H/ACA-unbio plus C-Sub-bio in presence of LNC (without Gar1) for comparison to the full RNP (compare Figure 4.28). The absolute number of observed molecules with E-FRET > 0.2 is represented by 'n'. Annealing and sample treatment were performed as for the full RNP (with Gar1). For this sample only one FRET-population at E-FRET = 0.68 was observed. The compact state (E-FRET = 0.9) did not appear in absence of Gar1.

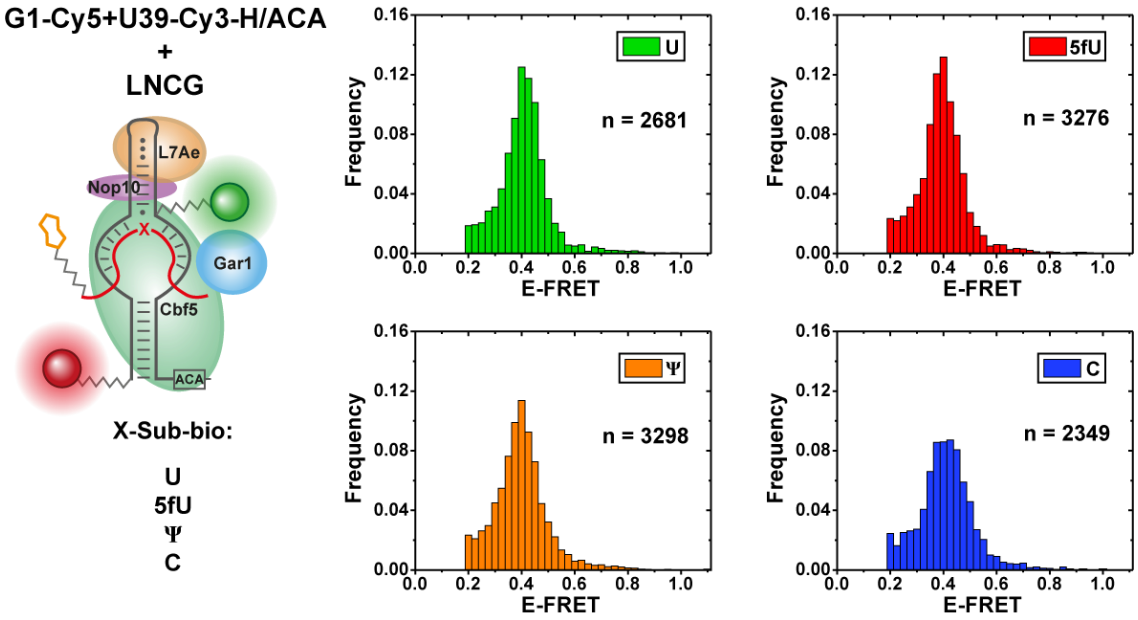


Figure A3.5: smFRET histograms of unbiotinylated G1-Cy5+U39-Cy3-H/ACA plus X-Sub-bio in presence of LNCG. Different species of X-Sub-Cy5 are color coded. The absolute number of observed molecules with E-FRET > 0.2 is represented by 'n'. Annealing and sample treatment were performed as for FRET-H/ACA-unbio (labeled at U26 instead of G1) samples. For all constructs only one major FRET-population at E-FRET = 0.40 was observed.

Project 2 – Eukaryotic H/ACA-RNP

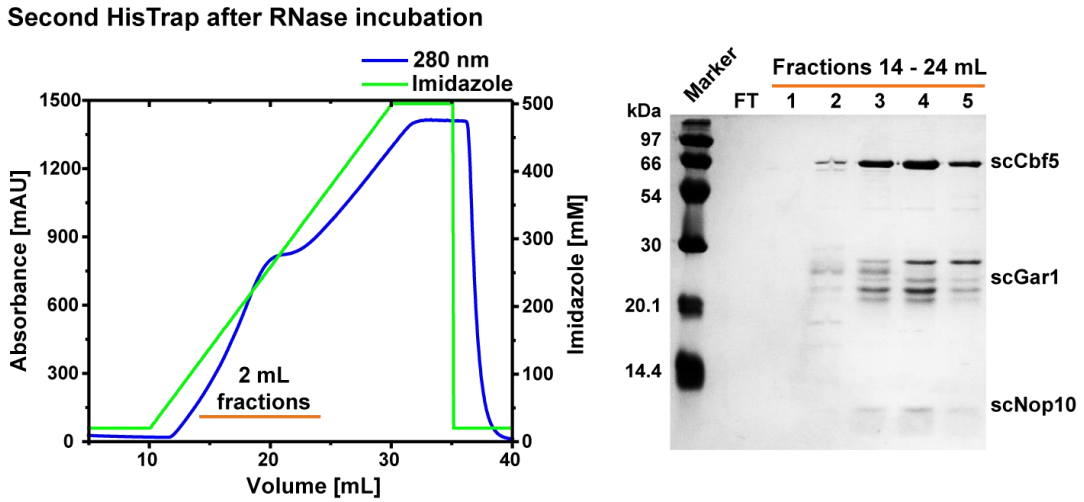


Figure A3.6: Ni-NTA-chromatogram (left) of scNCG after RNase treatment on 1 mL HisTrap and the resulting 15% SDS-PAGE (right) of fractions showing high UV-absorption. Chromatogram and gel were performed by Jason Nigel Martins within the scope of a bachelor thesis. Fractions 2 to 5 were used for further purification.

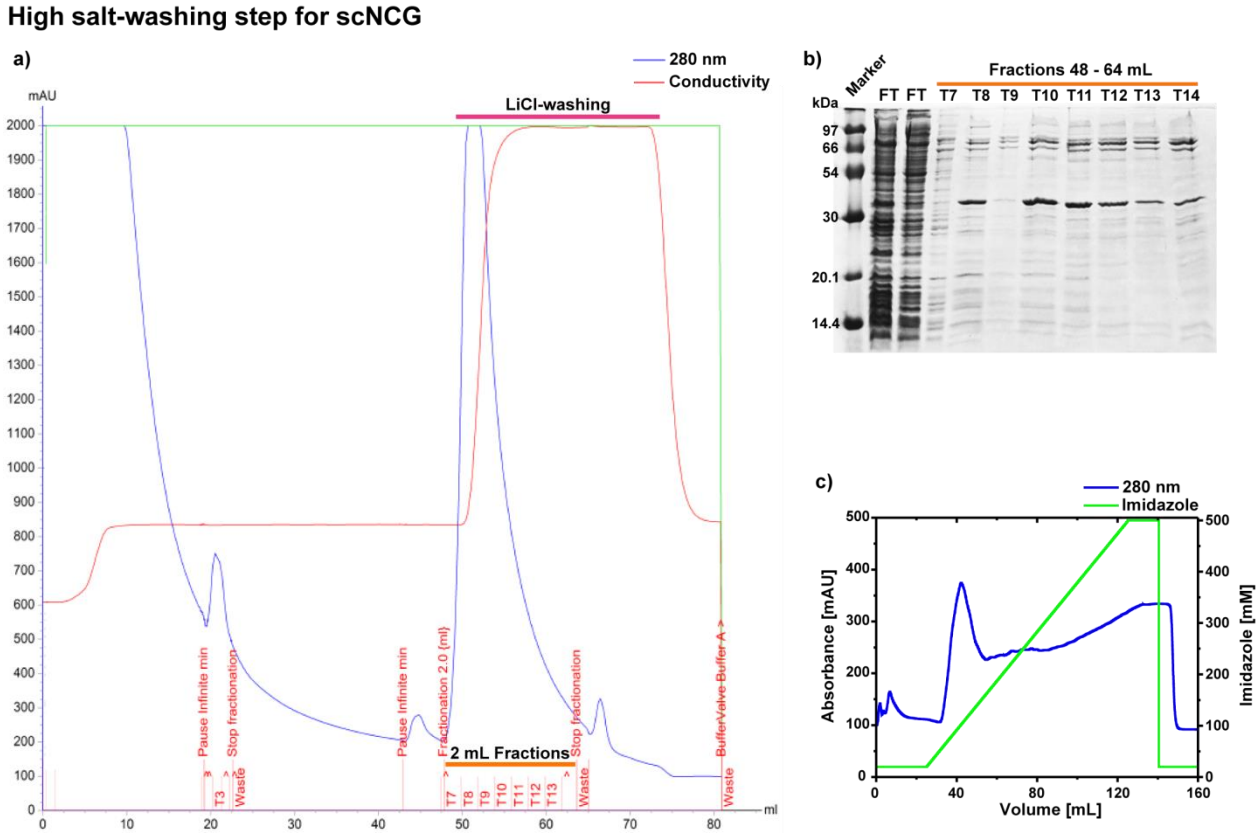


Figure A3.7: *a)* Run-log from the LiCl-washing step of scNCG-Ctag on a 5 mL Ni-NTA-column. The blue line represents the detected absorption at 280 nm, and the red line the measured conductivity. LiCl-washing (pink) and fractionation (orange) are highlighted in the log. The rapid increase in conductivity (at ca. 50 mL) is caused by the high salt buffer (2 M LiCl) and leads to an increase in absorbance. *b)* Resulting 15% SDS-PAGE of the washing step. No proteins of the scNCG complex were found. *c)* Subsequent Ni-NTA-affinity chromatogram of scNCG-Ctag after LiCl-washing.

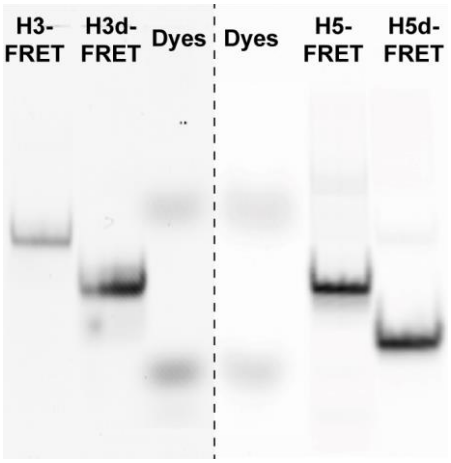


Figure A3.8: Native 8% PAGE of H3- and H3d-FRET (Typhoon settings: FRET-scan) as well as H5- and H5d-FRET (Typhoon settings: Cy5-scan). Separately run gels for the individual hairpin constructs were aligned (border marked by the dotted line). The running dyes (xylene cyanol and bromophenol blue) were added for comparison. Each constructs shows a homogeneous fold in the native PAGE. The full-length constructs (H3- and H5-FRET) run higher as the truncated constructs (H3d- and H5d-FRET), as was expected.

Data for unlabeled guide RNAs

RNA	H5						H5d				
Proteins	WNCG						WNCG				
Temperature	30 °C				20 °C	(on ice)	30 °C				20 °C
Time/min	Data set 1	Data set 2	Ø	Error	Data set 1	Data set 1	Data set 1	Data set 2	Ø	Error	Data set 1
5	0.44455	0.36263	0.40359	0.05793	0.08137	0.00194	0.33701	0.30456	0.32079	0.02294	0.07733
20	0.79766	0.69986	0.74876	0.06916	0.25936	0.00764	0.70133	0.63518	0.66826	0.04678	0.20795
60	0.86104	0.85035	0.85569	0.00756	0.49942	0.01187	0.84713	0.82640	0.83676	0.01466	0.42114
240	0.94289	0.92023	0.93156	0.01602	0.79878	0.00061	0.89367	0.88620	0.88994	0.00528	0.71734

RNA	H5					(+5fU-scSub)	H5d			
Proteins	scNCG				WNC	WNCG	scNCG			
Temperature	30 °C						30 °C			
Time/min	Data set 1	Data set 2	Ø	Error	Data set 1	Data set 1	Data set 1	Data set 2	Ø	Error
5	0.02578	0.04765	0.03671	0.01547	-	0.04679	0.11332	0.08866	0.10099	0.01744
20	0.10521	0.13846	0.12184	0.02352	-	0.06568	0.26693	0.24707	0.25700	0.01404
60	0.22742	0.27224	0.24983	0.03170	0	0.10534	0.45625	0.43736	0.44680	0.01336
240	0.44611	0.57099	0.50855	0.08830	-	0.18227	0.74067	0.72747	0.73407	0.00933

Data for labeled guide RNAs

RNA	H5-FRET		H5d-FRET	
Proteins	WNCG		WNCG	
Temperature	30 °C	20 °C	30 °C	20 °C
Time/min	Data set 1	Data set 1	Data set 1	Data set 1
5	0.30462	0.10534	0.23208	0.08532
20	0.56320	0.23755	0.52383	0.16448
60	0.77785	0.42162	0.77451	0.29805
240	0.87907	0.65189	0.88262	0.57940

Fitting equation:

$$y = \frac{y_{max} \cdot x}{Sub_{1/2} + x}$$

Initial velocity calculation:

$$v_{start} = \frac{y_{(fit)}}{x_{(fit)}} \cdot 40$$

Figure A3.9: The tables summarize the obtained results from densitometric analysis (ImageJ) of the separated and visualized radioactive nucleotides (by TLC and autoradiography) for the multiple turnover assays of the designed eukaryotic H/ACA-constructs. The individual time points were plotted with the given fitting equation (Michealis Menten kinetics) in Origin. The fitting point closest to 5 min was taken for calculation of the initial velocity.

A4 – Scripts

Matlab-Script – 'readhisttracesandor3'

```
%Program to read binary traces file output by Hazens IDL program written by
%Michael Stone 01/17/2004
clear all;
total_frets = [];
n_molecules = [];


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 = 5000; % 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 = 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:n_fr);   %% this gets rid of the frame number col, which is
the first colum in a traces file

%RemoveBleachedMolecules


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
%using 12% for the 630 nm dichroics

total_acceptor_corrected(1:n_peaks-1,1:n_fr)= total_acceptor(1:n_peaks-1,1:n_fr)-
0.2*total_donor(1:n_peaks-1,1:n_fr);
%total_acceptor_corrected = total_acceptor;

for i = 1:n_peaks-1
    i
    dyesum = total_donor(i,:)+total_acceptor_corrected(i,:);
```

```
    FRET = total_acceptor_corrected(i,:)./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

%this creates the vector for binning into 0.025 FRET steps
bins = [-0.1    -0.075  -0.05   -0.025  0   0.025   0.05    0.075   0.1 0.125   0.15    0.175
0.2 0.225   0.25    0.275   0.3 0.325   0.35    0.375   0.4 0.425   0.45    0.475   0.5 0.525
0.55    0.575   0.6 0.625   0.65    0.675   0.7 0.725   0.75    0.775   0.8 0.825   0.85    0.875
0.9 0.925   0.95    0.975   1   1.025   1.05    1.075   1.1];
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)









%
% h = figure('Position', [7 449 1652 504]);
%
% i = input('What molecule would you like start on?');
%
%
%
% %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([-.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');
%            %A = 2;
%        if A == 1;
```

```
% %             q = helpdlg('Click on start and end of bleach event for normalization', 'Point Selection');
% %             uiwait(q)
%             [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([-.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([-.5 1.2])
%             C = menu('Accept this trace', 'yes', 'no');
%             if C == 1;
%                 accepted_traces = [accepted_traces donor_norm' acceptor_norm'];
%             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
%         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);
% %save the data to ascii text file for later analysis...this is time
% %intensive so might not be the way to go.
% % file2=strcat('accepted',fileName);
% % dlmwrite(file2,accepted_traces,'\t');
%
%
% r = 2;%= menu('Would you like to print data now?', 'yes', 'no');
%
% if r == 1;
% % % %now that we have saved the data, lets generate a series of figures to be
% % % %printed and saved in .ai format.
% % % set up the figure printing format
% set(gcf, 'PaperPositionMode', 'manual');
% set(gcf, 'PaperOrientation', 'landscape');
% set(gcf, 'PaperUnits', 'inches');
% set(gcf, 'PaperPosition', [0 0 10 7.5]);
% %
% % %print five traces at a time
% data = accepted_traces; %rename matirx as 'data'
% [rows,columns] = size(data);
% m = 1; % molecule counter
%
% for i = 1:8:columns;
%     %make and setup figure properites
%     h = figure;
%     set(h, 'Position', [1 35 1680 929]);
%     set(gcf, 'PaperPositionMode', 'manual');
%     set(gcf, 'PaperOrientation', 'landscape');
%     set(gcf, 'PaperUnits', 'inches');
%     set(gcf, 'PaperPosition', [0.5 0.5 10 7.5]);
%     time = (1/10:1/10:120)';
%     for j = 1:2:8;
%       donor = data(:,i);
%       accep = data(:,i+1);
%       subplot(4,2,j)
%       plot(time,donor,'g');
%       title(strcat(fileName, '  molecule #', num2str(m)), 'FontSize', 8);
%       grid on
%       xlim([0 120])
%       ylim([-100000000 500000000])
%       hold on
%       ylabel('Intensity')
%       plot(time,accep,'r')
%       hold off
%       subplot(4,2,j+1)
%       dyesum = accep+donor
%       [rows] = find(dyesum <fretthreshold);
%       FRET = accep./dyesum;
%       FRET(rows) = 0;
%       plot(time,FRET);
%       grid on
%       xlim([0 120])
%       ylim([-.2 1.2])
%       ylabel('FRET')
%       i = i+2;
%       m= m+1;
%     end
%     print -dwinc; %print document to current printer in color
%     close(h);
% end
% end
%
%
%
```

Matlab-Script - '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);


%% 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);    %% this gets rid of the frame number col, which is
the first colum in a traces file


%RemoveBleachedMolecules
%FindBindingEvents1


total_donor = [];
total_acceptor = [];

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
%using 12% for the 630 nm dichroics

total_acceptor_corrected(1:n_peaks-1,1:n_fr)=                 total_acceptor(1:n_peaks-1,1:n_fr)-
0.2*total_donor(1:n_peaks-1,1:n_fr);

%total_acceptor_corrected(1:n_peaks-1,1:n_fr)=                total_acceptor(1:n_peaks-1,1:n_fr)-
0.11*total_donor(1:n_peaks-1,1:n_fr);
%total_acceptor_corrected = total_acceptor;
h = figure('Position', [219 28 1015 689]);
%h = figure('Position', [7 449 1652 504]);




i = input('What molecule would you like start on?');


%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([-.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');
           %A = 2;
        if A == 1;
%               q = helpdlg('Click on start and end of bleach event for normalization', 'Point
Selection');
%            uiwait(q)
            [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([-.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([-.5 1.2])
            C = menu('Accept this trace', 'yes', 'no');
            if C == 1;
                accepted_traces = [accepted_traces donor_norm' acceptor_norm'];
            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
        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);
%save the data to ascii text file for later analysis...this is time
%intensive so might not be the way to go.
% file2=strcat('accepted',fileName);
% dlmwrite(file2,accepted_traces,'\t');


r = 2;%= menu('Would you like to print data now?', 'yes', 'no');

if r == 1;
% % %now that we have saved the data, lets generate a series of figures to be
% % %printed and saved in .ai format.
% % set up the figure printing format
set(gcf, 'PaperPositionMode', 'manual');
set(gcf, 'PaperOrientation', 'landscape');
set(gcf, 'PaperUnits', 'inches');
set(gcf, 'PaperPosition', [0 0 10 7.5]);
%
% %print five traces at a time
data = accepted_traces; %rename matirx as 'data'
[rows,columns] = size(data);
m = 1; % molecule counter

for i = 1:8:columns;
    %make and setup figure properites
    h = figure;
    set(h, 'Position', [1 35 1680 929]);
    set(gcf, 'PaperPositionMode', 'manual');
    set(gcf, 'PaperOrientation', 'landscape');
    set(gcf, 'PaperUnits', 'inches');
    set(gcf, 'PaperPosition', [0.5 0.5 10 7.5]);
    time = (1/10:1/10:120)';
    for j = 1:2:8;
      donor = data(:,i);
      accep = data(:,i+1);
      subplot(4,2,j)
      plot(time,donor,'g');
      title(strcat(fileName, '  molecule #', num2str(m)), 'FontSize', 8);
      grid on
      xlim([0 120])
      ylim([-100000000 500000000])
      hold on
```

```
      ylabel('Intensity')
      plot(time,accep,'r')
      hold off
      subplot(4,2,j+1)
      dyesum = accep+donor
      [rows] = find(dyesum <fretthreshold);
      FRET = accep./dyesum;
      FRET(rows) = 0;
      plot(time,FRET);
      grid on
      xlim([0 120])
      ylim([-.2 1.2])
      ylabel('FRET')
      i = i+2;
      m= m+1;
    end
    print -dwinc; %print document to current printer in color
    close(h);
end
end
```

Deutsche Zusammenfassung

In den letzten Jahrzehnten hat sich die allgemeine Auffassung zur Rolle und Funktion von RNAs in der Zelle grundlegend verändert. Ausgehend vom ursprünglichen Bild des passiven Mediators in der Proteinbiosynthese ist mittlerweile bekannt, dass die RNA weitaus mehr enzymatische und regulatorische Funktionen übernimmt. Ein Grund für die Vielseitigkeit dieser Molekülklasse ist die Anzahl an Nukleotid-Modifikationen die in RNAs vorkommen können. Zurzeit sind über 170 spezifische solcher RNA-Modifikationen bekannt und es werden stetig neue gefunden[19]. Aus all diesen Modifikationen ist Pseudouridin (Ψ) die bei weitem am häufigsten vorkommende Variante. Ψ kann entweder durch eigenständige Proteine oder auf einem RNA basierten Mechanismus eingeführt werden. RNA-Klassen die diese Modifikation dirigieren, werden als H/ACA-RNAs bezeichnet und können unterteilt werden in snoRNAs (Eukaryoten) und sRNAs (Archaeen). Diese bilden eine bis drei Haarnadelstrukturen aus (in Eukaryoten sind exklusiv nur snoRNAs mit zwei Haarnadelstrukturen bekannt). Jede dieser Haarnadeln besteht aus einem unteren Stamm und einem oberen Stamm (enthält einen Loop von variabler Länge) zwischen denen sich die sogenannte Pseudouridylierungstasche (hier bindet die Substrat-RNA) befindet. Für enzymatische Aktivität benötigen die H/ACA-RNAs einen Satz an Proteinen: Nhp2 (L7Ae in Archaeen), Nop10, Cbf5 (Dyskerin oder NAP57 in Wirbeltieren), und Gar1. Hierbei ist Cbf5 die eigentlich reaktive Pseudouridylase (Ψase). Zusammen formen diese den sogenannten H/ACA-Ribonukleoprotein Komplex (H/ACA-RNP) und sind verantwortlich für die Ψ-Synthese in spliceosomalen und ribosomalen RNAs. Fehlfunktionen dieser H/ACA-RNPs wurden mit letalen Knochenmarkserkrankungen assoziiert[74,79,80]. Darüber hinaus wurden snoRNAs und der Prozess der Pseudouridylierung mit verschiedenen Krebsarten und Viruserkrankungen in Verbindung gebracht[82], was die Bedeutung in Bezug auf ein umfassendes Verständnis dieser RNA-Klasse weiter hervorhebt.

Das Hauptwissen über diese Enzymklasse stammt aus Kristallstrukturen von sRNPs aus Archaeen[52-56], während Struktur und Funktion von den eukaryotischen snoRNPs weniger bekannt sind. Ziel dieser Thesis war es neue Erkenntnisse zu Struktur und Dynamiken von H/ACA-RNPs zu gewinnen. Die Untersuchung von archaealer und eukaryotischer RNA-gesteuerter Pseudouridylierung sollte mittels Einzelmolekül-FRET-(smFRET)-Spektroskopie erfolgen. FRET ist eine Technik die eine zeitaufgelöste und direkte Beobachtung struktureller

Veränderungen (ca. 1 – 10 nm) von Fluorophor-markierten Domänen unter physiologischen Bedingungen ermöglicht. Diese Arbeit wurde in zwei Projekte unterteilt: 1) Untersuchung eines thermophilen H/ACA-RNP aus *Pyrococcus furiosus* (Archaea) und 2) eines eukaryotischen H/ACA-RNP aus *Saccharomyces cerevisiae* (Hefe), welcher bei moderaten Temperaturen aktiv ist. smFRET-Spektroskopie wurde angewandt um Komplexbildung, Substrat-Rekrutierung und individuell reaktive Zustände zu beobachten. Darüber hinaus wurden biochemische Bindungs- und Aktivitätstests mit fluorophor- und radioaktivmarkierten RNAs durchgeführt, um die RNP-Bildung zu validieren und mögliche Einflüsse einzelner Proteine in der RNA-gesteuerten Pseudouridylierung zu untersuchen. Für diese Zwecke wurden archaeale und eukaryotische H/ACA-Proteine sowie Fluorophor-markierte RNAs im Rahmen dieser Thesis hergestellt, gereinigt und auf ihre biochemische Funktion getestet. Dabei lag in dieser Dissertation der Fokus auf den strukturellen Dynamiken der H/ACA-RNAs.

Ausgehend von vorangegangen Arbeiten[119] musste die Herstellung der archaealen Ψase optimiert werden. Sobald Präparation und smFRET Experimente für diesen Komplex etabliert waren sollte die gewonnene Expertise auf das zweite Projekt, mit dem weniger erforschten eukaryotischen H/ACA-RNP, angewandt werden, um schlussendlich Vergleiche zwischen archaealer und eukaryotischer Pseudouridylierung zu ziehen. Die Herstellung einzelner Proteine sowie Proteinkomplexe (zum Beispiel des ternären Nop10-Cbf5-Gar1 Komplexes ‚NCG') erfolgte mittels rekombinanter Expression von Codon-optimierten Plasmiden (aus Verständnisgründen wird eukaryotischen Versionen der Proteine und Komplexe mit gleichem Namen das Präfix ‚sc' vorangestellt). Zur Aufreinigung dieser Proteine wurde unter anderem Affinitäts- und Größenausschlusschromatographie angewandt. Außerdem musste auf eine Nukleinsäure-freie Präparation der Proteine geachtet werden, da es sich um RNA-bindende Makromoleküle handelte. Um smFRET-Spektroskopie mit den H/ACA-RNPs betreiben zu können wurden die H/ACA-RNAs positionsspezifisch mit Fluorophoren (FRET-Donor und -Akzeptor) markiert. Die Herstellung dieser modifizierten RNAs erfolgte über enzymatische Ligation individuell markierbarer Oligomere zu den Volllängekonstrukten. Diese Teilfragmente waren kommerziell erhältlich und wurden während der Synthese positionsspezifisch mit funktionellen Gruppen zur Fluorophorkopplung oder mit Biotin-Modifikationen für die Komplex-Immobilisierung zwecks smFRET-Mikroskopie versehen. Um die modifizierten volllänge H/ACA-RNAs zu erhalten, wurden die einzelne Teilfragmente individuell an Fluorophore gekoppelt, mittels HPLC

gereinigt, alle nötigen Oligomere enzymatisch ligiert und eine finale gelelektrophoretische Aufreinigung durchgeführt. Auf diese Weise wurden diverse Fluorophor-markierte H/ACA-RNA Konstrukte hergestellt, die die Untersuchung globaler Strukturdynamiken in sRNA und snoRNAs ermöglichen.

Die hergestellten Protein- und RNA-Konstrukte wurden mittels auf ihre Funktionalität überprüft. Dazu wurde in sogenannten gelelektrophoretischen Tests die RNA-Bindefähigkeit der hergestellten Proteine kontrolliert. In allen Tests zeigte sich die Bildung eines RNPs, auch wenn aufgrund von Aggregation an der Geloberfläche keine Aussagen über Spezifität der Bindung getroffen werden konnten. Außerdem wurden für den H/ACA-RNP aus Archaea die Fluoreszenzanisotropien einzelner Fluorophore und Markierungsstellen bestimmt. Die erhaltenen Werte zeigten, dass die Ausbildung des RNPs keinen signifikanten Einfluss auf die Rotationsfreiheit der Fluorophore hatte, was eine Grundvoraussetzung ist um zuverlässige Abstandsinformation via FRET zu erhalten. Die Funktionalität von markierten H/ACA-RNAs wurde in Aktivitätsassays mit einem Vielfachen an Substrat RNAs überprüft. Es zeigte sich, dass die Insertion von Fluorophoren keinen wesentlichen Einfluss auf die enzymatische Funktion der Ψase hatte. Somit konnte nachgewiesen werden, dass Herstellung, Modifizierung und Reinigung der H/ACA-RNPs erfolgreich war und valide Konstrukte für die Untersuchungen von Strukturdynamiken mittels smFRET-Spektroskopie lieferten.

In dieser Arbeit wurde gezeigt, welchen Beitrag einzelne Proteine zur Ψase-Formation und ihrer biologischen Funktion leisten. Die Assemblierung archaealer und eukaryotischer H/ACA-RNPs scheint auf ähnliche Weise zu erfolgen. Die Ausbildung des RNA-Protein Komplexes führt eine strukturelle Umlagerung in der H/ACA-RNA herbei. Dabei öffnet sich die Pseudo-uridylierungstasche, was für die Substratbindung in archaealen sRNAs notwendig zu sein scheint. Eine effiziente Öffnung der reaktionsdirigierenden RNA wurde nur in Gegenwart aller Proteine beobachtet, was darauf hindeutet, dass auch die Anwesenheit von Gar1 – das einzige Protein, das nicht in direktem Kontakt mit der H/ACA-RNA steht[52,54] – die Komplexbildung fördert. In Rekrutierungstests wurde jedoch gezeigt, dass Gar1 im Gegensatz zu L7Ae und Nop10 nicht essentiell für eine effiziente Substratbindung der Ψase ist. Dies ist ein experimentelles Indiz für die von Liang *et al*. erstellte Hypothese, dass L7Ae eine korrekte Substratpositionierung im RNP fördert[53]. Des Weiteren konnten in dieser Arbeit neue Einblicke in die Rolle und Funktion des

kleinsten der vier Proteine (Nop10) erworben werden. Verankert zwischen L7Ae (gebunden am sogenannten Kink-Loop im oberen Stamm), Cbf5 (dem aktiven Enzym) und der H/ACA-RNA, scheint dieses Protein die gesamte RNP-Struktur zu stabilisieren, die für eine effektive Substratbindung erforderlich ist.

Um die Rolle der eukaryotischen H/ACA-Proteine im Pseudouridylierungsprozess zu untersuchen, wurden mehrere partielle snoRNA-Konstrukte designt. Der zweiteilige Komplex wurde in 5‘- und 3‘-Haarnadelstrukturen unterteilt, die jeweils ein Set an Proteinen (Nhp2, scNop10, scCbf5 und scGar1) binden. In vorangegangen Studien konnte gezeigt werden, das auch diese partiellen snoRNPs katalytisch aktiv sein können[59,72]. Darüber hinaus wurden von den 5‘- und 3‘-snoRNAs Volllängekonstrukte (H5 und H3), die die kanonische Loopsequenz enthalten, und verkürzte Konstrukte (H5d und H3d), bei denen ein Teil des oberen Stamms fehlt und die große Loopstruktur durch einen Tetraloop ersetzt wurde, erstellt. Es konnte gezeigt werden, dass Nhp2 – im Gegensatz zu scGar1 – nicht essentiell für die Funktion von snoRNPs ist. Für Nhp2 konnte eine Stabilisierungsfunktion im snoRNP identifiziert werden. Die Abwesenheit dieses Proteins hatte eine merklich geringere Reduzierung der Reaktivität zur Folge für das verkürzte Konstrukt H5d im Vergleich zu H5. Dies ist höchstwahrscheinlich auf die Tatsache zurückzuführen, dass der intramolekulare Tetraloop von H5d die verbleibende Struktur der H/ACA-RNA stabilisiert und das Fehlen von Nhp2 in der natürlichen snoRNA kompensiert. Die Entfernung von scGar1 führte jedoch zur vollständigen Inaktivität der Ψase, was wahrscheinlich auf den Verlust der angenommen Substratpositionierungs- und Produkt-freisetzungsfunktion zurückzuführen ist[63,64]. Interessanterweise wurde für die archaealen Äquivalente von Nhp2 und scGar1 – L7Ae und Gar1 – eine konträre Signifikanz für die Aktivität der H/ACA-RNPs gefunden[128]. Anscheinend ist der Stabilisierungseffekt von L7Ae wichtiger für die Funktion des archaealen sRNPs, als die Substratpositionierungs- und Produktfreisetzungsfunktionen von Gar1. Diese gegensätzliche Bedeutsamkeit der Proteine zur Ψase-Funktion in Eukaryoten und Archaeen lässt sich mit den natürlich vorkommenden Reaktionstemperaturen der Organismen erklären. Als thermophiler Organismus ist *Pyrococcus furiosus* normalerweise sehr hohen Temperaturen (ca. 100 °C) ausgesetzt, während Hefe optimale Bedingungen zwischen 20 °C und 30 °C hat. Und bei diesen hohen Temperaturen scheint die Stabilität des RNA-Protein Komplexes wichtiger für die Funktion des H/ACA-RNPs zu sein, als die Positionierung von Substrat und Produkt, da Letzteres aufgrund der intrinsisch

vorhandenen Bewegungsenergie auch ohne Gar1 stattfinden kann. Darüber hinaus konnte gezeigt werden, dass die Aktivität einzelner Haarnadeln des zweigeteilten snoRNPs von der Anwesenheit der jeweils anderen Domäne abhängig sein können. Die eigenständigen 3‘-Konstrukte zeigten allesamt keine katalytische Aktivität in Abwesenheit der 5‘-Domäne. Vermutlich ist dies auf eine Fehlfaltung des partiellen H/ACA-RNPs zurückzuführen. Nur für die 3‘-Konstrukte wurde eine Kompaktion der Pseudouridylierungs-tasche in den smFRET-Experimenten beobachtet. In Bezug auf zukünftige Experimente wäre es interessant zu sehen wie die Anwesenheit einer Domäne die Strukturen der jeweils anderen beeinflusst. Dazu könnte man die FRET-markierten 3‘-Konstrukte an unmarkierte 5‘-Haarnadel-RNAs ligieren und Assemblierungsexperimente für den vollständigen snoRNP wiederholen.

Reaktive Zustände während der RNA dirigierten Pseudouridylierung wurden auf unterschiedliche Weise für archaela und eukaryotische H/ACA-RNPs untersucht. Da für die themophile Ψase von *Pyrococcus furiosus* keine nativen Temperaturen mittels smFRET-Spektroskopie erreicht werden konnten, wurden einzelne Katalysierungszustände (reaktiv, produktähnlich und unkatalytisch) mittels Verwendung diverser Zielnukleotide in der Substrat-RNA imitiert und es wurde versucht die Zustände durch schnelle Abkühlung während der Inkubation festzuhalten. So konnten Hinweise auf eine Produktfreisetzungsfunktion des sRNPs gefunden werden. Anscheinend erfährt die Produkt-RNA eine strukturelle Umlagerung, bevor sie aus von der Ψase befreit wird. Des Weiteren konnte eine dichte Komprimierung im oberen Stamm des H/ACA-RNPs festgestellt werden, wenn eine Ziel-RNA mit einem Cytidin (produktähnlicher Zustand) an der zu modifizierenden Stelle an den Komplex gebunden war. Ob es sich bei der beobachteten Komprimierung um die vermutete Produktfreisetzungsfunktion oder eine potentielle Diskrimierungsfunktion der Ψase gegenüber fehlgepaarten RNAs handelt, konnte in dieser Arbeit nicht final geklärt werden. Die modifizierte Ziel-RNA – mit Cytidin an der Zielposition – ein stellt interessantes Konstrukt für zukünftige Kristallisationsstudien dar, da bisher noch nichts über die kompakte Verdichtung im oberen Stamm des archaealen H/ACA-RNPs berichtet wurde.

Für den eukaryotischen H/ACA-RNP wurde versucht, Pseudouridylierungsereignisse direkt unter dem Mikroskop zu beobachten, da dieser Komplex auch bei Raumtemperatur aktiv ist. In 5‘-Konstrukten wurden keine Strukturänderung mittels smFRET-Spektroskopie nach Substratzugabe beobachtet, was darauf hindeutet, dass die snoRNAs nach Komplexbildung in

einer Konformation während des Reaktionsprozesses verweilen. Bindungsereignisse zwischen Substrat RNA und dem H/ACA-RNP konnten mittels Fluorophor-markierter Substratanaloga, die verschiedene Zielnukleotide besaßen, beobachtet werden. Die Verweilzeiten dieser Substrat-Enzym-Komplexe reichte von einigen hundert Millisekunden bis zu mehr als einer Minute an. Eine eindeutige Zuordnung dieser Bindungsereignisse zu einzelnen Prozessen in der Pseudouridylierung (Substrat-Rekrutierung, Isomerisierung oder Produktabgabe) war nicht möglich. Jedoch deuteten die beobachteten Trends darauf hin, dass funktionelle H/ACA-RNPs und tatsächliche Isomerisierungen beobachtet wurden. Damit wurde die Grundlage für zukünftige Experimente geschaffen, um kinetische Daten zu den verschiedenen auftretenden Prozessen während der RNA-gesteuerten Pseudouridylierung mittels smFRET zu erhalten. Bevor dies erreicht werden kann, müssen jedoch mehrere Optimierungen vorgenommen werden. Dabei ist wahrscheinlich die Präparation von stabilem scGar1 der wichtigste Punkt. Es ist bekannt, dass rekombinantes scGar1 anfällig für Degradation ist[59], so wie es in den verwendeten scNCG-Konstrukten auch in dieser Thesis teilweise der Fall war. Wie gezeigt wurde ist scGar1 essentiell für die Aktivität der Ψase, sodass nicht auszuschließen ist, dass einige der hergestellten H/ACA-RNPs in ihrer enzymatischen Funktion massiv beeinträchtigt oder komplett inaktiv sind. Deshalb wurden gegen Ende der Thesis Versuche unternommen stabile scGar1-Varianten herzustellen. Dabei zeigte sich, dass Veränderungen des N- oder C-Terminus durch eine Umplatzierung oder Löschung des Histidin-Tags vielversprechende Ansätze sind. Diese Konstrukte könnten ein Zuordnung der verschieden Bindungsereignisse maßgebend erleichtern. Des Weiteren könnten mit den vorliegenden Biomolekülen und dem stabilen scGar1 Strukturstudien mittels Kryo-Elektronenmikroskopie unternommen werden. So könnten erste hochauflösende Strukturen von funktionellen und vollständigen snoRNPs erhalten werden.

Zusammenfassend lässt sich festhalten, dass diese Arbeit neue Einblicke in Aufbau, Struktur und Dynamiken von H/ACA-RNPs gewährt hat und Vergleiche zwischen eukaryotischer und archaealer RNA-dirigierender Pseudouridylierung gezogen werden konnten. Des Weiteren hat die erfolgreiche Etablierung der Präparation von Ψasen aus Hefe die Grundlage für Folgeexperimente gelegt, die tiefere Einblicke in den essentiellen Prozess der Pseudouridylierung ermöglichen. Somit trägt diese Thesis einen bedeutenden Anteil zum allgemeinen Verständnis dieser wichtigen Enzymklasse bei.

Publications

Peer-reviewed

Schmidt A. & Hanspach G., Hengesbach M., 'Structural Dynamics govern Substrate Recruitment and Catalytic Turnover in H/ACA-RNP Pseudouridylation', *RNA Biology* 2020; preprint.

Alriquet M., Calloni G., Martínez-Limon A., Hanspach G., Hengesbach M., Tartaglia G. G., Vabulas R. M., 'The protective role of m1A during stress-induced granulation', *Journal of Molecular Cell Biology* 2020; preprint.

Hanspach G. & Trucks S., Hengesbach M., 'Strategic labelling approaches for RNA single molecule spectroscopy', *RNA Biology* 2019; 16:1119-32.

Alriquet M., Martínez-Limon A., Hanspach G., Hengesbach M., Tartaglia G. G., Calloni G., Vabulas R. M., 'Assembly of Proteins by Free RNA during the Early Phase of Proteostasis Stress', *Journal of Proteome Research* 2019; 18: 2835-2847.

In revision

Trucks S. & Hanspach G., Hengesbach M., 'Eukaryotic specific RNA and protein features facilitate assembly and catalysis of H/ACA snoRNPs'.

Conference Contributions

Oral presentations

Hanspach, G. 'smFRET-Analysis of Structural Dynamics in Eukaryotic H/ACA-Complexes', *CRC 902 - IRTG Symposium*: '*Understanding RNA-based Regulation in Cells*', Frankfurt am Main, Germany, August 2018

Hanspach, G. 'RNA Structural Dynamics govern H/ACA-Ribonucleoprotein Function', *CRC 902 - IRTG Symposium*: '*RNA Biology*', Friedrichsdorf, Germany, November 2017

Poster presentations

Hanspach, G., Trucks, S., Hengesbach, M. 'smFRET-Analysis of Structural Dynamics in Eukaryotic H/ACA-Complexes', *10th RNA Biochemistry Meeting*, Bonn, Germany, October 2018

Hanspach, G., Schmidt, A., Hengesbach, M. 'RNA Structural Dynamics govern H/ACA-Ribonucleoprotein Function', *22nd Annual Meeting of the RNA Society*, Prague, Czech Republic, May 2017

Hanspach, G., Schmidt, A., Hengesbach, M. 'Structure, Assembly and Dynamics of H/ACA-Ribonucleoproteins', *9th RNA Biochemistry Meeting*, Bonn, Germany, October 2016

Hanspach, G., Schmidt, A., Hengesbach, M. 'Structural Dynamics of H/ACA-RNA Modification Enzyme Complexes', *VII. Nucleinsäurechemie-Treffen*, Berlin, Germany, September 2015

Hanspach, G., Schmidt, A., Wilhelm, J., Hengesbach, M. 'Structural Dynamics of H/ACA-mediated RNA Modification', *8th RNA Biochemistry Meeting*, Bonn, Germany, October 2014

List of Academic Teachers

Dr. Markus Braun - Dr. Georg Bruls - Prof. Dr. Bernd Brutschy - Prof. Dr. Irene Burghardt - Dr. Stephanie Cronje - Prof. Dr. Andreas Dreuw - Prof. Dr. Ernst Egert - Prof. Dr. Joachim Engels - Prof. Dr. Michael Göbel - Dr. Kenta Goto - Prof. Dr. Alexander Heckel - Dr. Rainer Hegger - Dr. Martin Hengesbach - Prof. Dr. Max Holthausen - Dr. Eltahmash Israr - Dr. Martin Kind - Prof. Dr. Bernd Kolbesen - Prof. Dr. Arnim Lühken - Prof. Dr. Thomas Prisner - Prof. Dr. Matthias Rehahn - Prof. Dr. Martin Schmidt - Prof. Dr. Harald Schwalbe - Prof. Dr. Teruo Shinmyozu - Prof. Dr. Andreas Terfort - Prof. Dr. Josef Wachtveitl - Prof. Dr. Matthias Wagner