Study of trm112, a unique methyltransferase activator at the interface between ribosome synthesis and function

Nhan Tran Van

université PARIS-SACLAY

NNT: 2017SACLX052

Thèse de doctorat

Study of Trm112, a unique methyltransferase activator, at the interface between ribosome synthesis and function

Thèse de doctorat de l'Université Paris-Saclay
préparée à l'Ecole Polytechnique

École doctorale n°573: Interfaces: approches interdisciplinaires, fondements, applications et innovation (Interfaces)
Spécialité de doctorat: BIOLOGIE

Thèse présentée et soutenue à Palaiseau, le 21 Septembre 2017, par

Nhan TRAN VAN

Composition du Jury:

M. **Hannu MYLLYKALLIO**	
Directeur de recherche, École Polytechnique –CNRS (UMR 7645)	Président
Mme. **Béatrice CLOUET-D'ORVAL**	
Directrice de recherche, CNRS - Université Paul Sabatier (UMR 5100)	Rapporteur
M. **Jean ARMENGAUD**	
Directeur de recherche, CEA-Marcoule	Rapporteur
Mme. **Tamara BASTA-LEBERRE**	
Maître de conférences, Université Paris-XI (UMR 9198)	Examinatrice
Mme. **Valérie DE CRÉCY-LAGARD**	
Professeur, University of Florida (Department of Microbiology)	Examinatrice
M. **Marc GRAILLE**	
Directeur de recherche, École Polytechnique – CNRS (UMR 7654)	Directeur de thèse

ACKNOWLEDGEMENT

First and foremost, I would like to express my sincere gratitude to my supervisor, Dr. Marc GRAILLE for accepting me as his PhD student. I am very much grateful for his wonderful guidance, continuous supports, helps and kindness in both scientific and non-scientific stuffs. I greatly appreciate all and words cannot express how thankful I am to him. Thank you so much, Marc.

Besides my supervisor, I would like to thank the rest of the jury members: Dr. Béatrice Clouet-D'orval, Dr. Jean Armengaud, Dr. Tamara Basta-Leberre, Prof. Valérie De Crécy-Lagard and Dr. Hannu Myllykallio for accepting to evaluate my thesis. Their insightful comments and detailed discussions helped me a lot to improve my thesis.

When I first came to the lab, Gabrielle Bourgeois taught me methods and showed me everything in the lab. She is a very good friend and teacher. I learnt a lot from her. Thank you so much Gaby. A very special gratitude is for Dr. Roxane Lestini (LOB, Ecole Polytechnique) for guiding me on archaeal genetics. Many thanks go to my teammates, both the former (Gabrielle Bourgeois, Clément Charenton, Régis Back, Juliette Letoquart) and the current (Amlan Roychowdhury, Ditipriya Hazra, Nathalie Ulryck) for their helps, nice talks and discussion. It was fantastic and honored to have the opportunity to work with them as a team.

I would like to thank all colleagues in BIOC for their hospitality, helps and nice talks. I actually enjoy the lab activities a lot and I will for sure miss them so much when I leave France.

Our research is carried out in a collaborating way; thank Dr. Sarah Cianferani and Leslie Muller for the MS experiments; thank Dr. Valérie Heurgué-Hamard and Prof. Denis Lafontaine for helpful discussion on the Trm112 project.

I want to acknowledge Synchrotron SOLEIL for facilitating our crystal structural studies; many thanks for Ecole Polytechnique for funding my PhD study.

Last but not the least, I would like to thank my parents for their unconditional love and cares; thank my parents-in-law for their encouragement and taking care of my daughter. Very special thank goes to the most beloved people in my life: my wife (Nguyễn Thị Ngọc Ái) and my daughter (Trần Nguyễn Khả Ân), who are the main motivation for my life and my study. This thesis is particularly a special gift for my daughter's 4^{th} birthday.

ACKNOWLEDGEMENT

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLE

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

I. Eukaryotic Translation

Protein synthesis is one of the most important processes in all living cells, leading to the synthesis of amino acid polymers based on messenger RNA (mRNA) templates. It is a tightly-regulated multi-step process occurring in the cytoplasm with the help of ribosome, a macromolecular machinery containing mainly ribosomal RNAs (rRNAs) and proteins (r-proteins), together with transfer RNAs (tRNAs) and other translational factors. The process begins when matured mRNAs synthesized in the nucleus by RNA polymerase II based on a DNA sequence are transported to the cytoplasm where the ribosome, made up of two subunits (60S and 40S), reads these mRNAs to produce corresponding proteins. By this way, the genetic information is systemically transferred from DNA to mRNA and finally proteins.

1. Translational factors and mechanism

1.1 The ribosome – the protein manufacturer

The ribosomes are known as ribozymes responsible for protein synthesis in all domains of life. They are all composed of two subunits, both built from RNAs and proteins (Figure 1A). Compared to bacterial ribosomes (sedimentation coefficient of 70S containing 30S and 50S subunits), the eukaryotic ribosomes are about 30 - 40% larger (sedimentation coefficient of 80S) and much more intricate (Ben-Shem et al., 2011; Jenner et al., 2012; Klinge et al., 2012). The small and large subunits are known as 40S and 60S subunits, respectively. Regarding its molecular weight, the eukaryotic 80S ribosome can range from 3.5MDa in lower eukaryotes to 4.5MDa in metazoa (Yusupova, & Yusupov, 2014).

1.1.1 Structure

To this day, several crystal and cryo-electron microscopy (cryo-EM) structures of eukaryotic ribosomes alone or in complex with tRNA, mRNA or proteins from different organisms have been determined, such as wheat 80S (Armache et al., 2010), yeast 80S (Ben-Shem et al., 2011), *Tetrahymena thermophila* 40S and 60S (Rabl et al., 2011; Klinge et al., 2012), human and *Drosophila melanogaster* 80S (Anger et al., 2013); *Trypanosoma brucei* 43S initiation complex (Hashem et al., 2013). These structures provide deep insights into structural landmarks as well as catalytic mechanism during protein synthesis by this huge machinery.

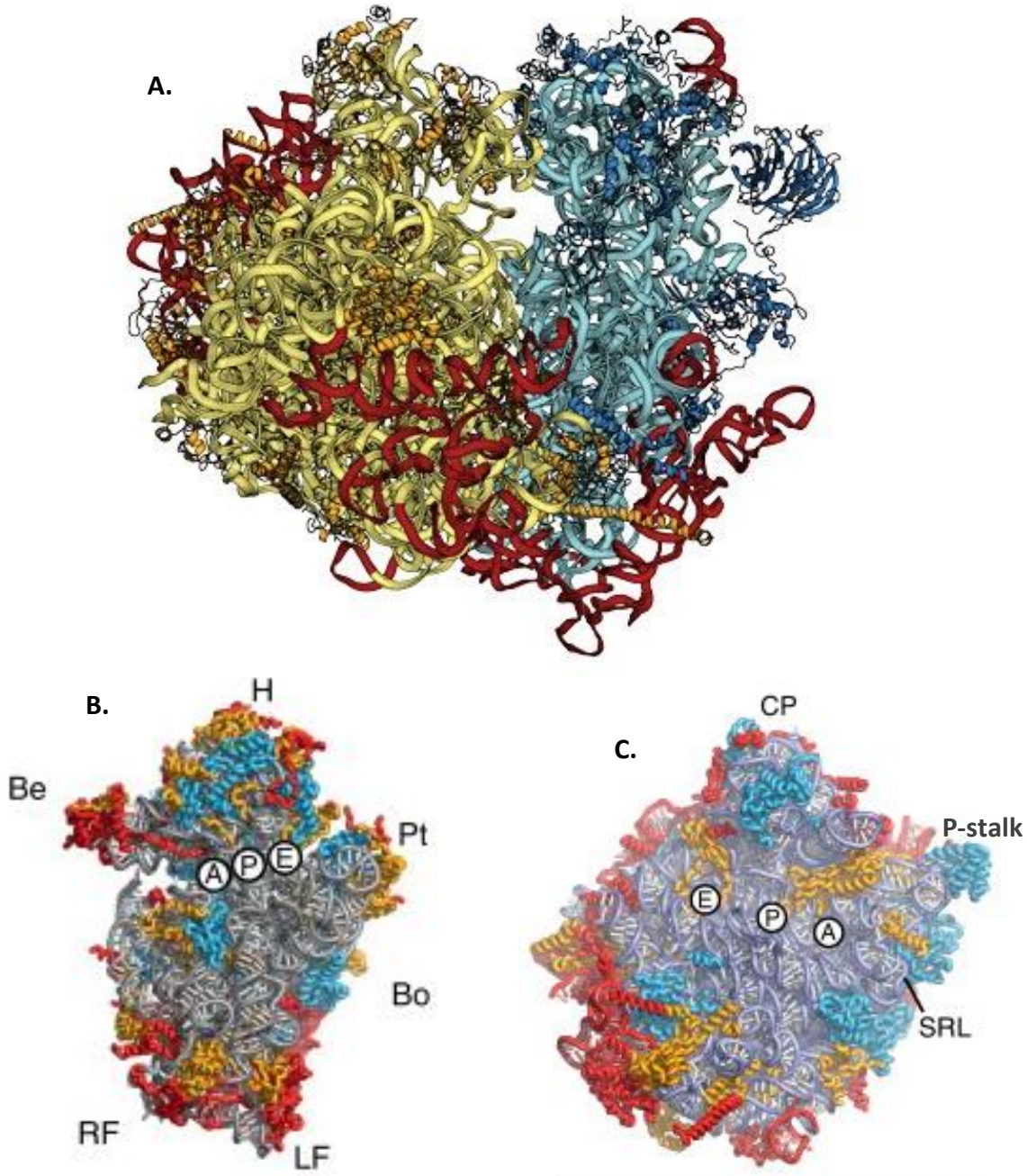


Figure 1. **The eukaryotic 80S ribosome**. (**A**) The 80S ribosome with 40S small subunit colored in blue, 60S large subunit in yellow. Eukaryotic expansion segments are shown in red (Ben-Shem et al., 2010). (**B**-**C**) Structures of 40S and 60S subunits respectively viewed at the subunit interface. The 40S subunit includes a head (H), beak (Be), platform (Pt), body (Bo), right foot (RF) and left foot (LF). The 60S subunit exhibits a central protuberance (CP) and P-stalk. A-P-E correspond to A-site, P-site and E-site (Adapted from Klinge *et al.,* 2012)

In yeast, the small 40S subunit consists of a 18S rRNA and 33 r-proteins, whereas the large 60S is formed by three rRNAs (5S, 5.8 S and 25S) and 46 r-proteins (Ben-Shem et al., 2011). Among those r-proteins, two thirds have homologs in bacteria and archaea, while the remaining are specific for eukaryotes (Ramakrishnan, 2011). Compared to their bacterial counterparts, the eukaryotic subunits have several additional expansion segments (ES) and variable regions (VR) in rRNAs.The r-proteins also have C-terminal insertions/extensions and there are eukaryotic-specific proteins, resulting in larger and more complex ribosomes in eukaryotes (Ben-Shem et al., 2011; Klinge et al., 2012; Wilson, & Doudna Cate, 2012; Anger et al., 2013).

Structurally, the eukaryotic 40S subunit similarly to prokaryotic 40S is divided into the head, body, platform, beak and shoulder, left foot, and right foot regions (Figure 1B) (Klinge et al., 2012; Yusupova, & Yusupov, 2014). This subunit contains an mRNA binding site and three binding sites for tRNAs at the subunit interface. The A-site binds the incoming aminoacyl-tRNA, the P-site holds the peptidyl-tRNA attached to the nascent polypeptide chain, and the E-site accommodates the deacylated P-site tRNA after peptide-bond formation before its release from the ribosome (Schmeing, & Ramakrishnan, 2009; Klinge et al., 2012). One very important feature of this subunit is the presence of the universally conserved decoding center located in the interface surface and created by the head, the shoulder and the penultimate stem. This is the place where the base-pairing interaction between the mRNA codon and the tRNA anticodon occurs (Melnikov et al., 2012; Yusupova, & Yusupov, 2014).

Like the prokaryotic 50S, the eukaryotic 60S subunit has a crown-like shape and is composed of the central protuberance, the L1 stalk and the P-stalk (Figure 1C) (Klinge et al., 2012; Wilson, & Doudna Cate, 2012; Yusupova, & Yusupov, 2014). This subunit contains the peptidyl transferase center (PTC) where the peptide bond formation takes place. It also harbors the ribosomal exit tunnel adjacent to the PTC, which allows the nascent polypeptide chain to thread through and access the solvent side in which it undergoes processing and folding (Tu, & Deutsch, 2010; Klinge et al., 2012; Yusupova, & Yusupov, 2014).

1.1.2 Biogenesis

Ribosome synthesis is one of the most complex and energetically consuming processes in all organisms (Henras et al., 2008; Kressler et al., 2010; Thomson et al., 2013). In eukaryotes, this intricate procedure is a multiple-step, error-prone process requiring at least 200 protein factors, numerous small nucleolar RNAs and non-ribosomal factors, such as AAA-ATPases, ATP-dependent RNA helicases and kinases. These are involved in the synthesis, the maturation and the transport of individual ribosomal components and their assembly into ribosomal subunits (Kressler et al., 2010) (Figure 2A).

The eukaryotic ribosome assembly starts in the nucleus where RNA polymerase I transcribes rDNA to produce a large polycistronic precursor rRNA (35S pre-rRNA) which then undergoes several chemical modifications on specific nucleotides and nucleolytic cleavages to obtain mature 5.8S, 18S and 25S rRNAs (Figure 2B). The fourth rRNA (5S) is transcribed by RNA polymerase III, whereas the RNA polymerase II synthesizes the pre-mRNAs encoding r-proteins and other ribosomal assembly-related protein factors (Grandi et al., 2002; Granneman, & Baserga, 2004; Henras et al., 2008; Woolford, & Baserga, 2013). During transcription, the long 35S pre-rRNA is assembled with r-proteins, assembly factors and small nucleolar RNAs to form a large 90S pre-ribosome or SSU processome. Cleavage events at sites A0, A1, and A2 in 35S pre-rRNA yield to 20S and 27S pre-rRNAs, which are further processed to generate pre-40S and pre-60S particles, respectively (Figure 2) (Henras et al., 2008; Kressler et al., 2010; Oeffinger, 2016). In the final step, these particles are transported through the nuclear pores to the cytoplasm, where numerous maturation steps are required to yield the mature, functional ribosomes (Henras et al., 2008; Kressler et al., 2010). It is particularly noteworthy that there are several O-methylated sugars and base methylation introduced in rRNAs at early and late stages of ribosome biogenesis (Sloan et al., 2016). Details regarding this point will be addressed in the next parts of the thesis.

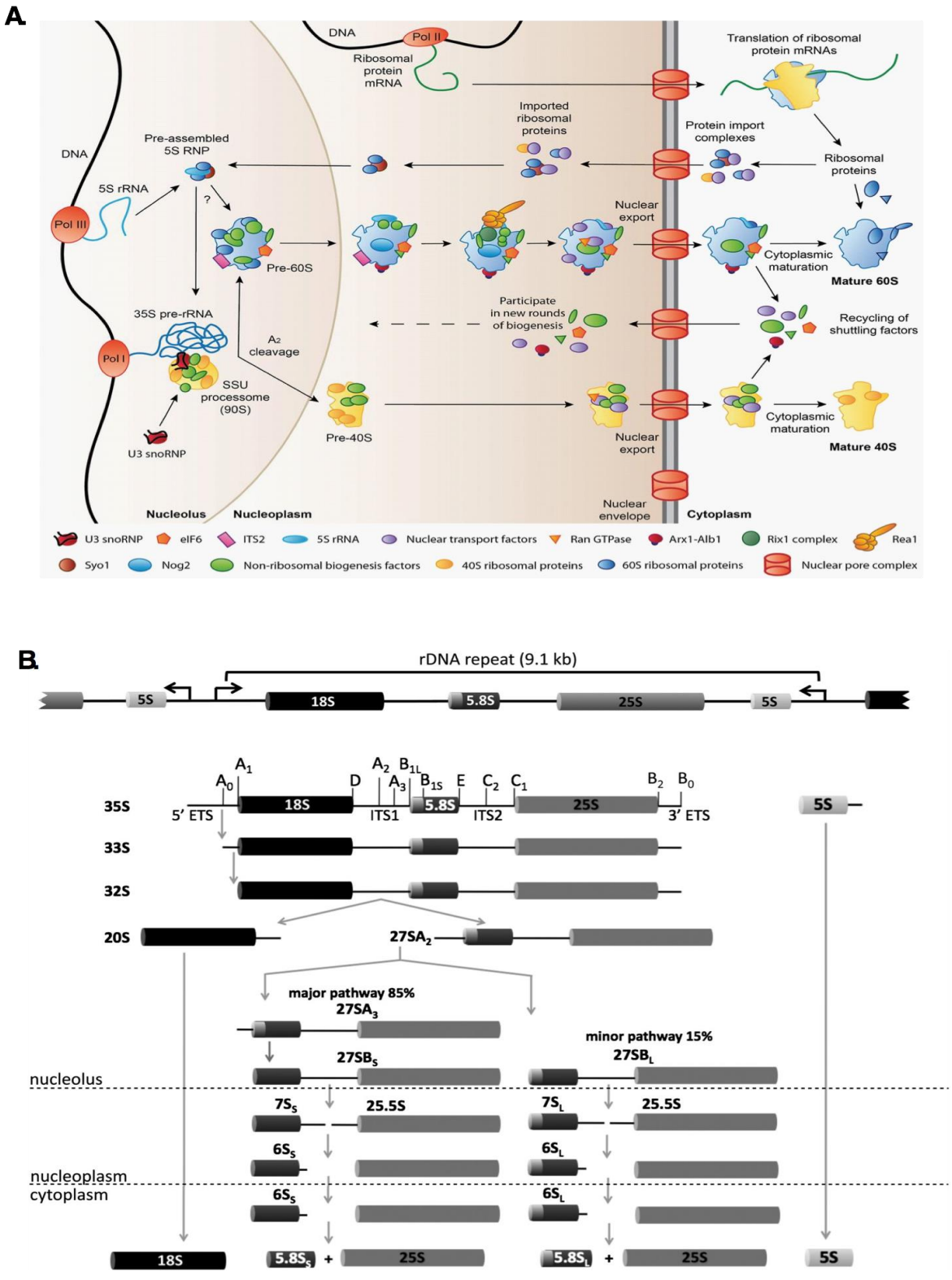


Figure 2. **Eukaryotic 80S ribosome biogenesis**. (**A**) Summary of 40S and 60S formation and maturation (Greber, 2016). (**B**) Pre-rRNA processing pathway (Woolford & Benserga, 2013).

1.2 mRNA – the translational template

Messenger RNAs (mRNA) carry genetic information transferred from DNA in the sequence of nucleotides, which are arranged into three-base codes called codons encoding corresponding amino acids. In eukaryotes, a primary precursor of mRNA (pre-mRNA) is synthesized by RNA polymerase II and requires a series of processing steps to produce the mature mRNA suitable for protein synthesis. These maturation steps include modifications of the 5 'and 3' ends as well as mRNA splicing.

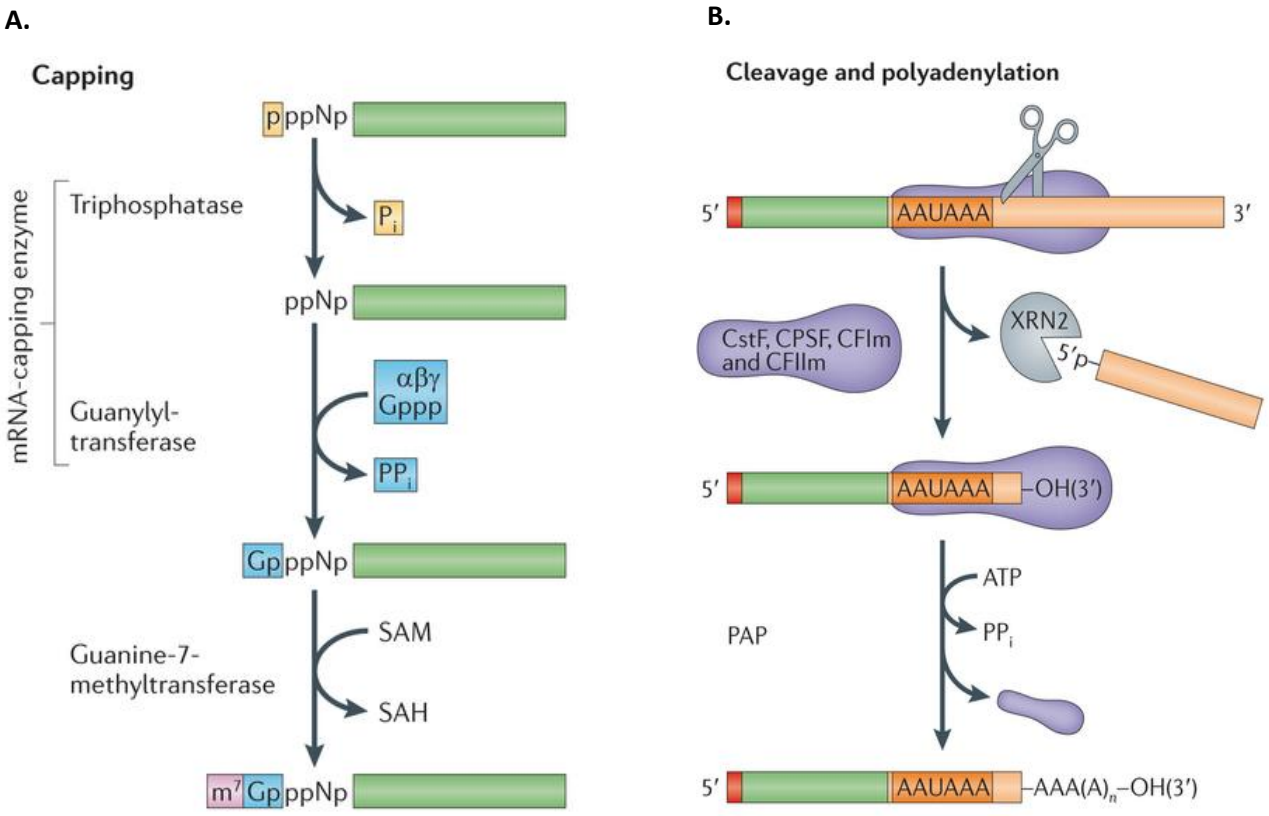


Figure 3. **mRNA processing**. (**A**) mRNA capping mechanism; (**B**) mRNA polyadenylation model (Bentley, 2014)

1.2.1 Capping at 5’ end

The nascent mRNA is co-transcriptionally modified by adding a 7-methylguanosine (m^7G) cap at the 5’ end through an unusual 5′ to 5′ triphosphate bond (Figure 3A). The m^7G cap

is a well-conserved modification feature in eukaryotic mRNAs. The mechanism responsible for cap addition is well-known and consists of several enzymes catalyzing different reactions (Banerjee, 1980; Ramanathan et al., 2016): (1) RNA triphosphatase cleaves the γ-phosphate at the 5′ triphosphate end of mRNA to generate 5′ diphosphate; (2) RNA guanylyltransferase (EC 2.7.7.50) adds a GMP group from GTP to the 5′ diphosphate via a lysine-GMP covalent intermediate, loosing a pyrophosphate and forming the 5′–5′ triphosphate linkage; (3) An mRNA (guanine-N7-)-methyltransferase (EC 2.1.1.56), a S-adenosyl-L-methionine (SAM)-dependent methyltransferase (MTase), then catalyzes methylation on the N7 atom of guanosine to form the m^7G cap (m^7GpppN) named cap 0; (4) Optionally, another SAM-dependent MTase modifies the nucleotide N on the 2' OH group of the ribose to generate the cap 1 ($m^7GpppNm$). With regard to its functions, the m^7G cap is known to protect mRNAs from the 5'-3' exonucleases (Hocine et al., 2010; Furuichi, 2015). It is also involved in splicing and in the regulation of mRNA export to the cytoplasm (Izaurralde et al., 1995). Another very important function of the mRNA cap is to promote the initiation step of the protein biosynthesis, acting like a signal for the ribosome recognition that then scans the mRNA to find the start codon (Hinnebusch, & Lorsch, 2012).

1.2.2 Polyadenylation at 3' end

Like the 5' end, the 3' end of the pre-mRNAs also undergoes modifications during the maturation process (Figure 3B). The polyadenylation mechanism includes two main steps (K. M. Brown, & Gilmartin, 2003; Mangus et al., 2003; Davila Lopez, & Samuelsson, 2008): (1) The pre-mRNA is cleaved at a site characterized by two signals, a highly conserved upstream AAUAAA sequence and a downstream G/U-rich sequence. This is carried out by a series of protein factors including cleavage/polyadenylation specificity factor (CPSF) and the cleavage stimulation factor (CstF)…; (2) A polyalanine (poly(A)) tail ranging from 50 to 250 nucleotides is added at the 3' end of the cleaved mRNA. Similarly to the 5' cap, the poly(A) tail is also known to protect mRNA but from the 3'-5' exonucleases. It also plays a role in mRNA transport to cytoplasm as well as in translation (Mangus et al., 2003; Hocine et al., 2010).

1.2.3 Splicing

Following transcription, eukaryotic pre-mRNAs contain an alternation of coding sequences (exons) and non-coding sequences (introns). In order to be ready for translation, the

pre-mRNAs are subjected to splicing steps to remove the introns and join the exons together. This is carried out by the spliceosome, a complex of small nuclear ribonucleoproteins (snRNPs), through a two-reaction process (Figure 4) (Fica et al., 2013; Scotti, & Swanson, 2016): (1) the formation of the intron lariat intermediate is catalyzed by a nucleophilic attack of the 2'OH of a specific branch-point nucleotide within the intron on the first nucleotide of the intron at the 5' splice site; (2) then the 3'OH of the released 5' exon performs a second nucleophilic attack on the nucleotide located just after the last nucleotide of the intron at the 3' splice site, thus connecting the exons and removing the intron lariat. Failure in RNA splicing as well as in its regulation can result in a growing number of human diseases ranging from retinal and developmental disorders to cancers, underlying the important roles of this mRNA maturation process (Scotti, & Swanson, 2016).

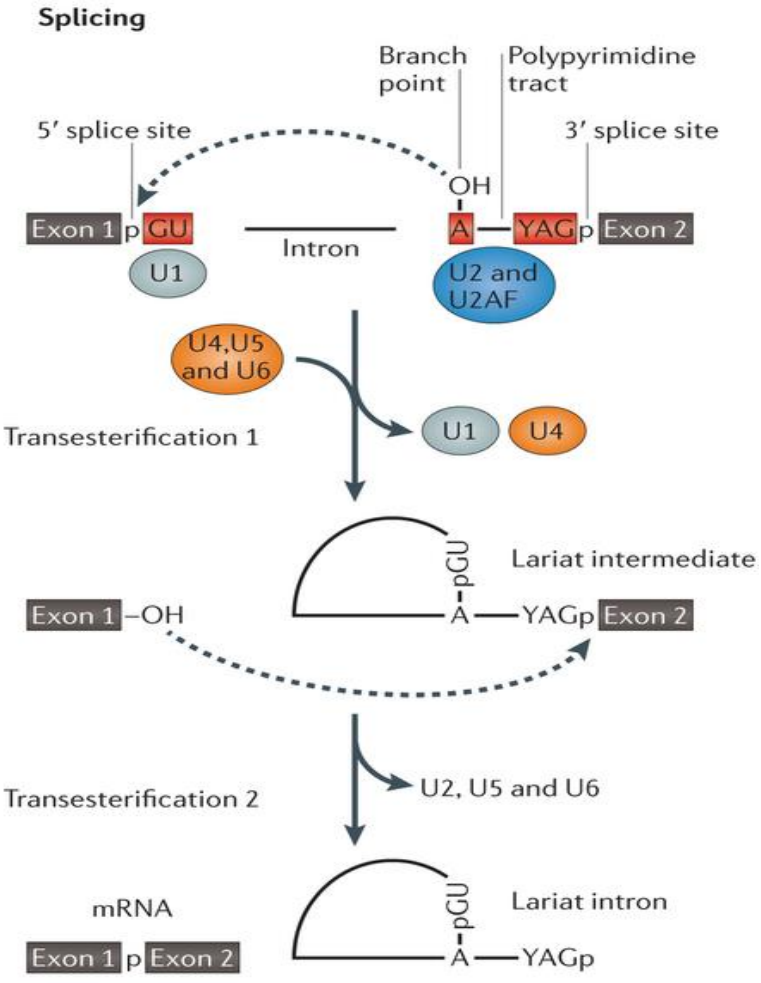


Figure 4. **mRNA splicing mechanism** (Bentley, 2014)

1.3 tRNA – an adaptor molecule

The tRNA plays an indispensable role as an adaptor molecule during the decoding process by physically linking the mRNA and the polypeptide chain. In eukaryotes, tRNAs are synthesized by RNA polymerase III as pre-tRNAs in the nucleus. In order to fulfill their complete functions, pre-tRNAs have to be processed by a variety of maturation steps: (1) removal of the 5’ leader sequence by RNase P; (2) removal of the 3’ trailer sequence by both endonuclease (tRNAse Z) and exonuclease (Rex1) enzymes; (3) addition of CCA; (4) removal of introns in some tRNAs by the combined action of an endonuclease and of a ligase; and (5) post-transcriptional modification of numerous tRNA nucleotides (Hopper, 2013; Wichtowska et al., 2013). Regarding those chemical modifications, it is known that tRNAs are the most heavily modified among the different RNAs and those modifications are required for tRNA folding, stability and function (Jackman, & Alfonzo, 2013; Vare et al., 2017). Details about this maturation step will be discussed later in the thesis.

1.3.1 tRNA structure

Similar to proteins, tRNAs also consist of primary, secondary and tertiary structures.

1.3.1.1 Primary structure

Eukaryotes can have tens to hundreds of different tRNAs depending on the organisms and some tRNAs can consist of the same anticodons (http://gtrnadb2009.ucsc.edu). In general, tRNAs have a sequence ranging from about 70 and up to 100 nucleotides (Figure 5A-B). As a conventional nomenclature (Sprinzl et al., 1998), tRNAs are numbered 1 for the nucleotide located at the 5' end and 76 for the nucleotide present at the 3' end. Similarly, the nucleotides forming the anticodon triplet are always numbered 34, 35 and 36.

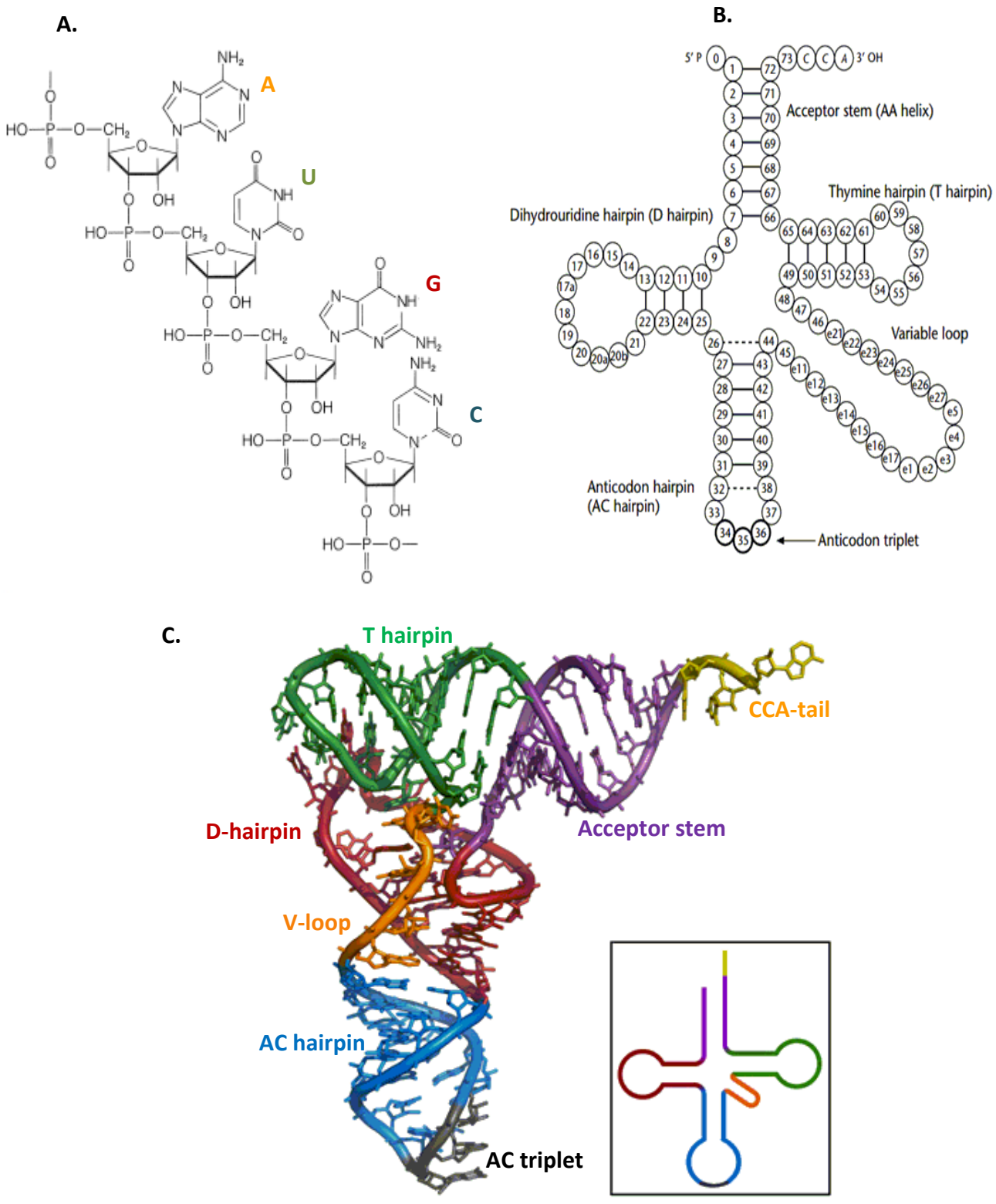


Figure 5. **tRNA structures**. (**A**) Primary structure of tRNA; (**B**) Cloverleaf-like secondary structure of tRNA; (**C**) L-shaped tertiary structure of tRNA

1.3.1.2 Secondary structure

tRNAs fold into a conserved secondary structure known as a cloverleaf (Figure 5B), which is formed by a set of canonical Watson-Crick base pairs as well as the wobble pairings (G-U) (Auffinger, & Westhof, 2001). The cloverleaf structure consists of five regions with four helices and 3-4 loops: (1) The acceptor arm helix is formed by the base pairing of the 5' and 3' ends, with the latter containing the CCA 3' end which is attached to the cognate amino acid by specific aminoacyl tRNA synthetases to form aa-tRNA. This region is normally formed by seven base pairs and possesses four unpaired nucleotides which sequence is –RCCA-3'OH at the 3' end. R is always a purine and this position acts as a discriminator for the selection of tRNAs by the cognate synthetases (Auffinger, & Westhof, 2001); (2) The dihydrouridine (D) hairpin or D-arm, which often contains the modified base dihydrouracil. This region generally consists of four base pairs and a 7-11 residues variable hairpin loop; (3) The anticodon (A) hairpin contains a 5-bp helix and a 7-residues loop with the anticodon triplet (residues 34-36). The first position of the anticodon is called the wobble base 34 while the position 33 of the loop is always a uridine (U33) and the residue 37 is highly modified (Auffinger, & Westhof, 2001; Grosjean et al., 2010); (4) The thymine (T) hairpin or T-arm, similarly to the anticodon arm, is composed of a 4-5 bp stem and loop with seven residues, which contains a pseudouridine (ψ), a modified nucleotide in which the link between the base and the sugar is C(1')-C5 instead of C(1')-N1; (5) The last region is a variable loop located between anticodon and T loops and present only in some tRNAs. The length of this loop can vary between 4 and 21 nucleotides and depending on its size, tRNAs are divided into class I (the variable loop has 4-5 residues) or class II (with longer variable loop) (Auffinger, & Westhof, 2001).

1.3.1.3 Tertiary structure

The three-dimensional structure of the tRNAs was a challenging problem until its crystal structure was solved for the first time in 1974 (Suddath et al., 1974; Auffinger, & Westhof, 2001). Up to now, several structures of free tRNAs and tRNAs in complex with their cognate aminoacyl tRNA synthetases and with amino acids exist. The cloverleaf tRNA is folded into a compact structure adopting an L shape (Figure 5C). This structure is created by four helical regions of the cloverleaf, from which pairs of helices stack on each other coaxially to form two

main arms or domains: the acceptor arm is made by the acceptor helix and the T-arm, whereas the anticodon arm is formed by the D-arm and the anticodon arm. In this L-shaped architecture, the two extremities are the anticodon triplet and the 3' CCA end, which are 75-80Å away (Auffinger, & Westhof, 2001). Moreover, this L-shaped structure is maintained by unusual hydrogen bonds occurring between the nucleotides located at invariant or semi-invariant positions (Auffinger, & Westhof, 2001; Oliva et al., 2006).

1.3.2 tRNA function

As discussed above, the main canonical function of the tRNAs is involved in the translation of genetic code as a linker between the mRNA and the amino acid sequence of the corresponding proteins by transferring amino acids, which will be inserted into the growing polypeptide chain. To perform this function, the 3'-CCA end of tRNAs are covalently attached with specific amino acids by their cognate aminoacyl-tRNA synthetases through a process called aminoacylation. This process is of particular importance since it affects the fidelity of the translation.

In addition to its well-known role in protein synthesis, tRNAs are known to possess additional functions. They are involved in the regulation of gene expression in both prokaryotes and eukaryotes (Raina, & Ibba, 2014). Moreover, aminoacylated tRNAs have been implicated as substrates for non-ribosomal peptide bond formation, post-translational protein labeling, bacterial cell envelope synthesis, and also antibiotic biosynthesis (Raina, & Ibba, 2014). tRNAs can act as precursors for the synthesis of other aa-tRNAs and could also have a potential catalytic role in certain reactions of peptidyl transfer, aminoacyl transfer and deacylation (Francklyn, & Minajigi, 2010). Furthermore, aa-tRNAs are known to be involved in the regulation of protein turnover via the N-terminal rule (Francklyn, & Minajigi, 2010).

1.4 Translation process

Translation consists of 4 different phases: initiation, elongation, termination and recycling of ribosome in all domains of life.

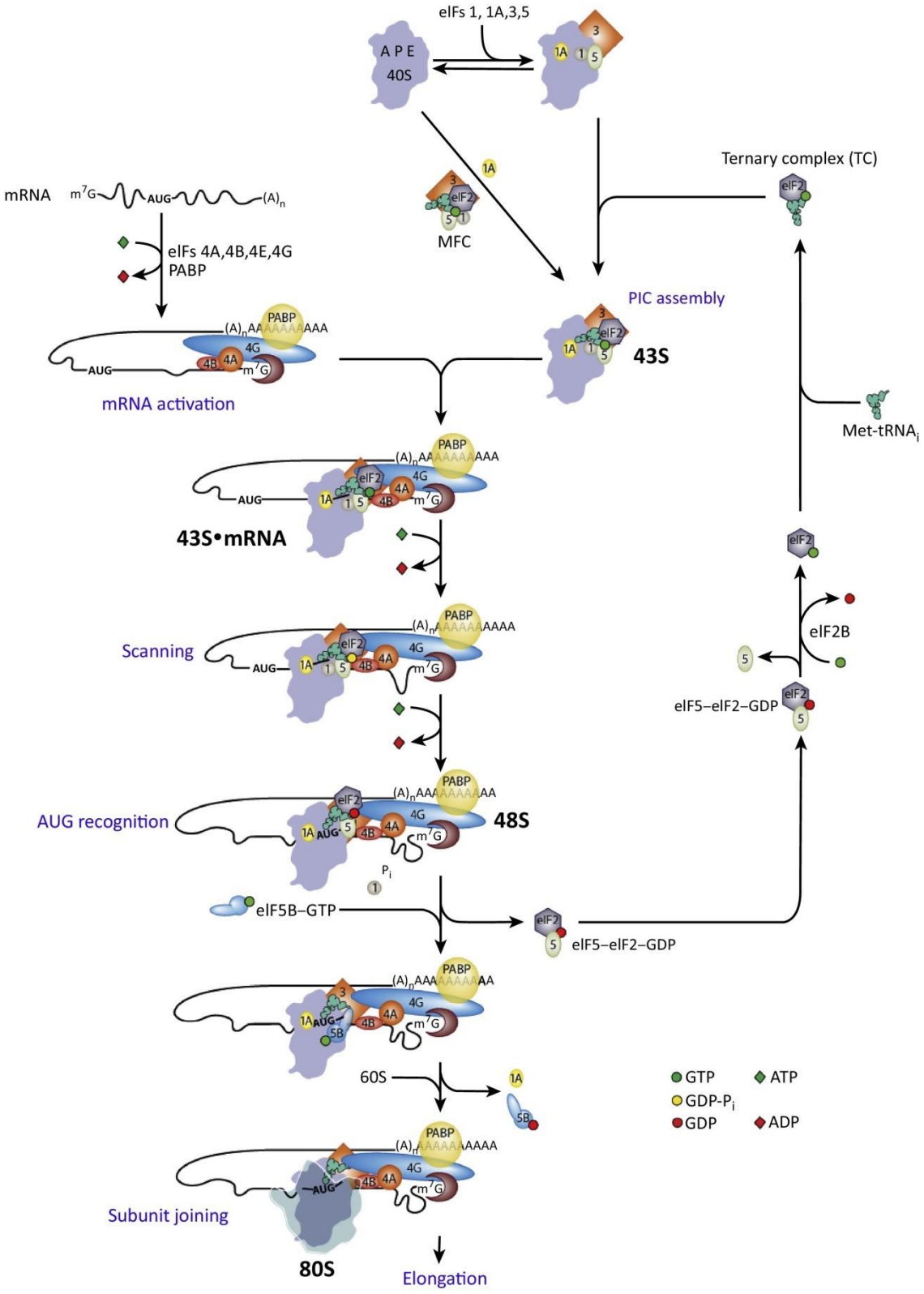


Figure 6. **Eukaryotic translation initiation model** (Adapted from Hinnebusch, 2017)

1.4.1 Translation initiation

In contrast to bacterial translation initiation, which needs the Shine–Dalgarno sequence of the mRNA to locate the start codon and require only three initiation factors (Voigts-Hoffmann et al., 2012), the eukaryotic translational initiation step is much more complex, requiring a scanning mechanism to identify the start codon and consists of at least 10 core initiation factors (eIFs) (Hussain et al., 2014; Aitken et al., 2016). The whole process can be divided into five different major steps (Jackson et al., 2010; Hinnebusch, & Lorsch, 2012) (Figure 6). In a first step, a methionylated initiator tRNA (Met-tRNA) forms a ternary complex (TC) with eIF2 (composed by α, β and γ subunits) and a GTP molecule before delivery to the ribosome. Next, this TC is associated with the 40S subunit linked to eIF1A, eIF1, eIF5 and eIF3s to create a pre-initiator complex 43S (PIC). Then, the PIC binds to mRNA via its interaction with mRNA cap-binding complex eIF4, thereby allowing the scanning of the mRNA in the 5' to 3' direction until the recognition of the start codon AUG in the P-site. The start codon recognition results in the formation of the 48S complex. This induces conformational changes leading to eIF1 displacement and to the hydrolysis of eIF2-bound GTP with the help of eIF5. Simultaneously, the eIF2-GDP, eIF1, eIF3 and eIF5 factors dissociate and the 60S subunit as well as eIF5B-GTP are recruited. Finally, the hydrolysis of the eIF5B-bound GTP is needed for dissociation of eIF1A and eIF5B-GDP, leaving the A-site empty to accommodate the first elongator tRNA. At the end of the initiation step, the initiator Met-tRNA occupies the P site of the functional 80S ribosome, which is ready for the elongation step.

1.4.2 Translation elongation

The translational elongation, a highly conserved step between eukaryotes and prokaryotes, is the stepwise polymerization of amino acids into the growing protein chain. In this step, the ribosome moves along the mRNA on a three-nucleotides basis called a codon in order to associate the corresponding aa-tRNA and thus allow the incorporation of the appropriate amino acid to the nascent polypeptide chain. This process can be divided into several steps and requires two eukaryotic elongation factors (eEFs): eEF1A/B and eEF2 (Figure 7) (Kapp, & Lorsch, 2004; Dever, & Green, 2012) (Figure 7).

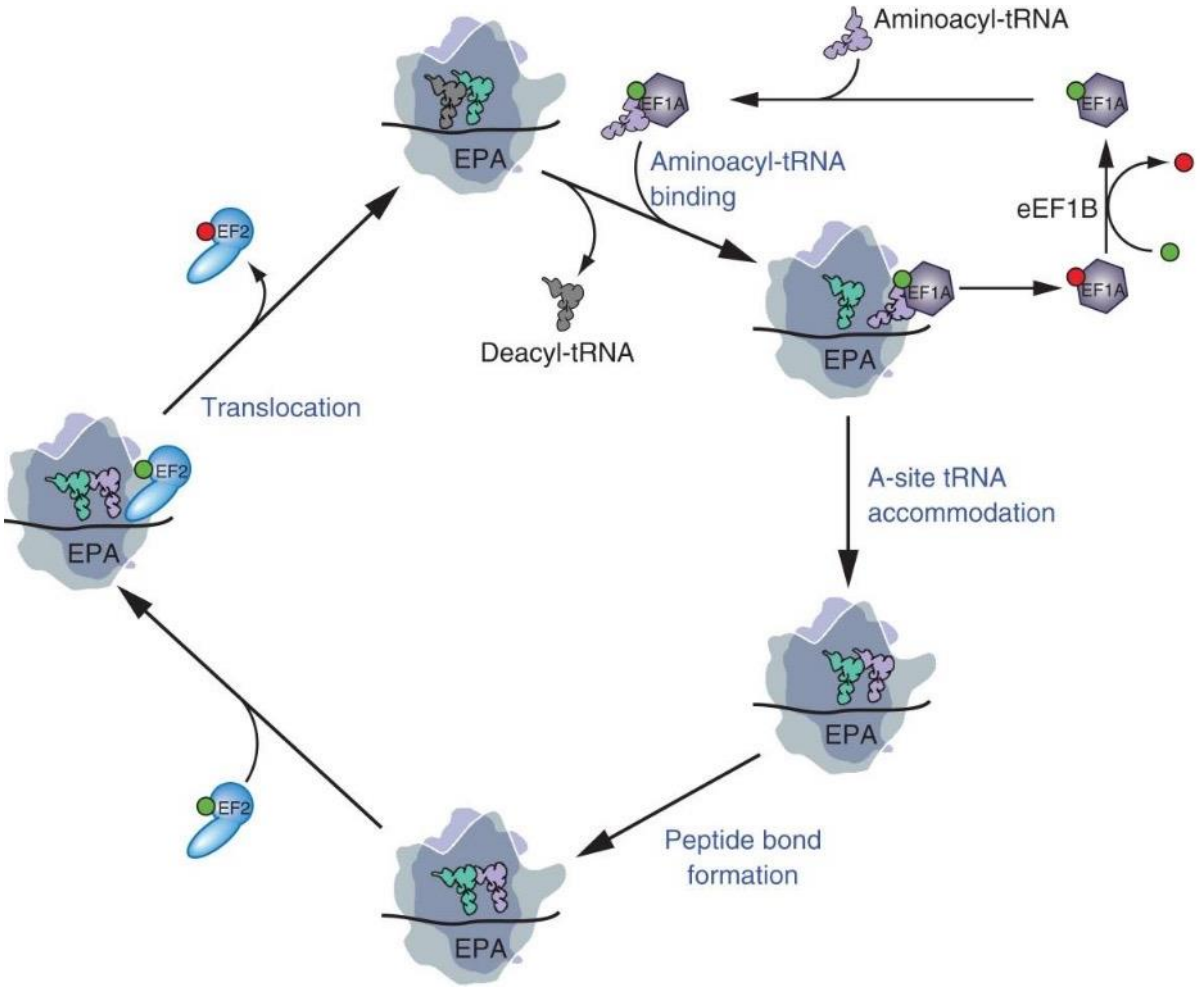


Figure 7. **Eukaryotic translation elongation model** (Adapted from Dever and Green, 2012)

The elongation step starts with the 80S ribosome poised on an mRNA with the P-site occupied by the initiator Met-tRNA containing the anticodon base-paired with the start codon. The A-site in which the second codon of the open reading frame (ORF) of mRNA is present is empty, awaiting for the cognate aa-tRNA (Kapp, & Lorsch, 2004). The first step of this process consists in the formation of a ternary complex between eEF1A and the aa-tRNA complementary to the second codon in the presence of GTP. This complex then enters the empty A-site on the ribosome, from which the anticodon of the aa-tRNA is matched against the codon positioned in the A-site. A correct codon-anticodon match results in the GTP hydrolysis of eEF1A-GTP and then dissociation of eEF1A-GDP, enabling the aa-tRNA to be accommodated into the A-site. Next, the aa-tRNA at the A-site re-orientates to bring its 3' end closer to that of the tRNA present

in the P-site while eEF1A-GDP is recycled to eEF1A-GTP by an exchange factor eEF1B. The formation of the peptide bond occurs rapidly upon the nucleophilic attack of the aminoacyl-tRNA in the A-site on the ester carbon of the P-site tRNA. This reaction is catalyzed by the peptidyl transferase center of the ribosome, which ideally positions the substrates for catalysis (Dever, & Green, 2012). It is then followed by the tRNAs movement to hybrid states with their acceptor arms in the P and E-sites while their anticodon loops still remain in the A and P-sites, respectively. The complete translocation of tRNAs from A and P sites to P and E sites, respectively requires the recruitment of eEF2-GTP and its GTP hydrolysis. eEF2-GDP is then released and the deacylated tRNA in the E-site dissociates from the ribosome (Kapp, & Lorsch, 2004; Rodnina, & Wintermeyer, 2011; Dever, & Green, 2012). The P-site then encompasses the tRNA bearing the elongated polypeptide chain whereas the A-site is empty, and ready for a new elongation cycle. This step is repeated until a stop codon (UAA, UAG or UGA) reaches the A-site, signaling the translation termination step.

1.4.3 Translation termination

The termination of protein synthesis takes place when one of the three stop codons (UAA, UAG and UGA) is detected in the ribosomal A-site. In eukaryotes, translation termination is catalyzed by two classes of release factors (RFs), class I eRF1 and class II eRF3 (Nakamura, & Ito, 2003) (Figure 8). The class I RF, eRF1 adopts a tRNA-like shape and recognizes all stop codons. It is responsible for the release of newly synthesized proteins from the tRNA at the P-site by triggering hydrolysis of the ester bond in peptidyl-tRNA. This results from the interaction between a universally conserved glycine-glycine-glutamine (GGQ) motif on eRF1 and the peptidyl transfer center (PTC) (Loh, & Song, 2010). Meanwhile, the class II RF, eRF3, an eEF1A-related translational GTPase, is associated to eRF1 and GTP and delivers eRF1 to the ribosome (Preis et al., 2014). eRF3 is required to complete the translation termination by enhancing the polypeptide release activity of eRF1 (Fan-Minogue et al., 2008; Eyler et al., 2013). It is known that eRF1 and the ribosome induce the hydrolysis of the GTP molecule bound to eRF3. This leads to dissociation of eRF3 from eRF1, further allowing conformational changes of eRF1 to project its GGQ motif into the PTC (Taylor et al., 2012; Preis et al., 2014).

Structurally, eRF1 consists of three well-defined domains: N, M and C (H. Song et al., 2000). The domain N (N-terminal) is involved in the recognition of the stop codons thanks to

several conserved motifs including GTx, NIKS and YxCxxxF (H. Song et al., 2000; Kryuchkova et al., 2013; Blanchet et al., 2015; A. Brown et al., 2015). The M domain (middle domain) containing the universally conserved GGQ motif is functionally analogous to the tRNA acceptor arm and this GGQ motif extends into the PTC to stimulate the polypeptide release (H. Song et al., 2000; Dever, & Green, 2012; Taylor et al., 2012; Wong et al., 2012). The C domain (C-terminal) is mainly responsible for the interaction with eRF3, mostly through hydrophobic interactions (Z. Cheng et al., 2009; Taylor et al., 2012). Moreover, it is particularly interesting to note that the glutamine side chain of GGQ motif is N5-methylated by a methyltransferase (Mtq2) in complex with a methyltransferase activator (Trm112). The methylation of the glutamine from the GGQ motif substantially increases the rate of peptide release on a subset of amino acids *in vitro* in bacteria (Pierson et al., 2016), raising the question of the impact of release factor methylation on the translation termination. The details about the methylation of the glutamine of GGQ motif by Mtq2-Trm112 complex will be discussed later in this thesis.

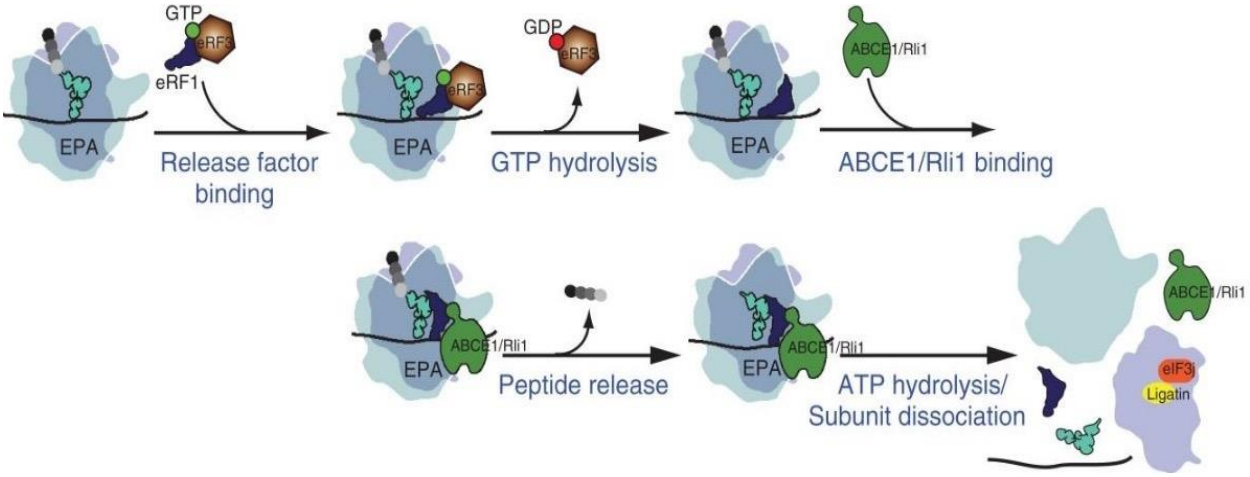


Figure 8. **Eukaryotic translation termination and recycling models** (Adapted from Dever and Green, 2012)

eRF3, for which its full-length structure is not yet available, has a variable N-terminal region known to be dispensable for translation termination (Kushnirov et al., 1988; Ter-Avanesyan et al., 1993) and a more conserved functional C-terminal region (Z. Cheng et al., 2009; Dever, & Green, 2012). The latter consists of a G-domain and of β-barrel domains 2 and 3, shares structural homology with elongation factors EF-Tu and eEF1A. It is directly interacting with domains M and C from eRF1 (Kong et al., 2004; Z. Cheng et al., 2009; Taylor et al., 2012). These interactions are important to stimulate eRF3 GTPase activity by eRF1 and by the ribosome, which is necessary for the translation termination (Z. Cheng et al., 2009; Taylor et al., 2012).

1.4.4 Recycling step

The recycling of the ribosome occurs once the complete polypeptide chain has been released. After the translation termination, the post-termination complex (post-TC) consisting of the 80S ribosome associated with an mRNA and to a deacylated tRNA and eRF1 in its P and A sites, respectively, needs to be dissociated (Pisarev et al., 2007; Jackson et al., 2012).

The exact mechanism of the recycling step is not fully understood to date, but there are two proposed mechanisms based on recent studies (Pisarev et al., 2007; Pisarev et al., 2010; Franckenberg et al., 2012; Jackson et al., 2012) (Figure 8). In a first model, the *in vitro* recycling of eukaryotic post-TCs is mediated by initiation factors eIF3, eIF1 and eIF1A through an energy-free mechanism functioning only in a narrow range of low Mg^{2+} concentrations (Pisarev et al., 2007; Pisarev et al., 2010). A second model has been recently proposed that requires ABCE1 or Rli1 in yeast, a highly conserved ATPase protein found in Eukaryotes and Archaea (Figure 8) (Pisarev et al., 2010; Kiosze-Becker et al., 2016). Following the dissociation of eRF3-GDP at termination step, eRF1 is available to interact with ABCE1, which is introduced between the two subunits of the ribosome. This leads to the hydrolysis of the ATP molecule bound to ABCE1 and the subsequent conformational changes of this later results in the dissociation of both ribosomal subunits (Becker et al., 2012; Kiosze-Becker et al., 2016). At this stage, the 60S subunit is free and ready for new cycles of translation while the small 40S subunit still holds the mRNA and the tRNA, which are subsequently released from 40S subunits either by eIF1/eIF1A/eIF3 or by ligatin (Jackson et al., 2012) (Figure 8).

2. Methylation as a major post-transcriptional and translational modification of translational machinery

It is known that eukaryotic protein synthesis is a tightly-regulated process requiring a variety of translational components. For their correct biosynthesis and functions, these factors are subjected to post-transcriptional and post-translational modifications (PTM). These PTM play key roles in enhancing and regulating the activity of the translational components and among different PTMs, methylation is the most prominent.

2.1 Methylation

Methylation is a chemical reaction, which catalyzes the transfer of a methyl group from a donor to an acceptor molecule in all living cells. Methylation plays an important role in several biological processes including biosynthesis, metabolism, detoxification, signal transduction, protein sorting and repair, and nucleic acid processing (Martin, & McMillan, 2002). This reaction is catalyzed by enzymes called methyltransferases (MTase) acting on a wide range of substrates such as nucleic acids, proteins, lipids, polysaccharides, and a range of small molecules (X. Cheng, & Roberts, 2001). Meanwhile, S-adenosyl-L-methionine (SAM) is considered as the most commonly used methyl donor for the methylation reaction (X. Cheng, & Roberts, 2001; Schubert et al., 2003; Fontecave et al., 2004).

2.1.1 S-Adenosyl-L-methionine (SAM or AdoMet)

SAM is a conjugate of a nucleotide adenosine and an amino acid methionine, known as a sulfonium compound (Figure 9A) (Fontecave et al., 2004; Kozbial, & Mushegian, 2005). SAM is known as the second most widely-used enzyme cofactor after ATP (Cantoni, 1975; Schubert et al., 2003). This is due to the fact that the strong electrophilic character of the SAM methyl group, which is brought into close proximity of a nucleophilic group from the substrate by SAM-dependent MTases, has highly favorable thermodynamics for SAM-dependent methyl-transfer reactions (Fontecave et al., 2004).

SAM is biosynthesized by SAM synthetase, which catalyzes the reaction between a methionine and an ATP. The newly synthesized SAM can further participate in a cycle, which is part of the general metabolism of sulfur-containing amino acid derivatives, by being converted to

S-adenosyl-homocysteine (SAH), which is subsequently hydrolyzed into adenosine and homocysteine by a SAH hydrolase. Homocysteine can be then either converted into glutathione or methylated to form methionine by methionine synthase. A new SAM cycle can start again with the resulting methionine (Figure 9B) (Fontecave et al., 2004). In addition to its role as the methyl donor for SAM-dependent MTases and its implication in sulfur-containing amino acid derivatives metabolism, SAM is known to act as precursor in the biosynthesis of the polyamines, nicotianamine and phytosiderophores, ethylene … (Kozbial, & Mushegian, 2005; Roje, 2006).

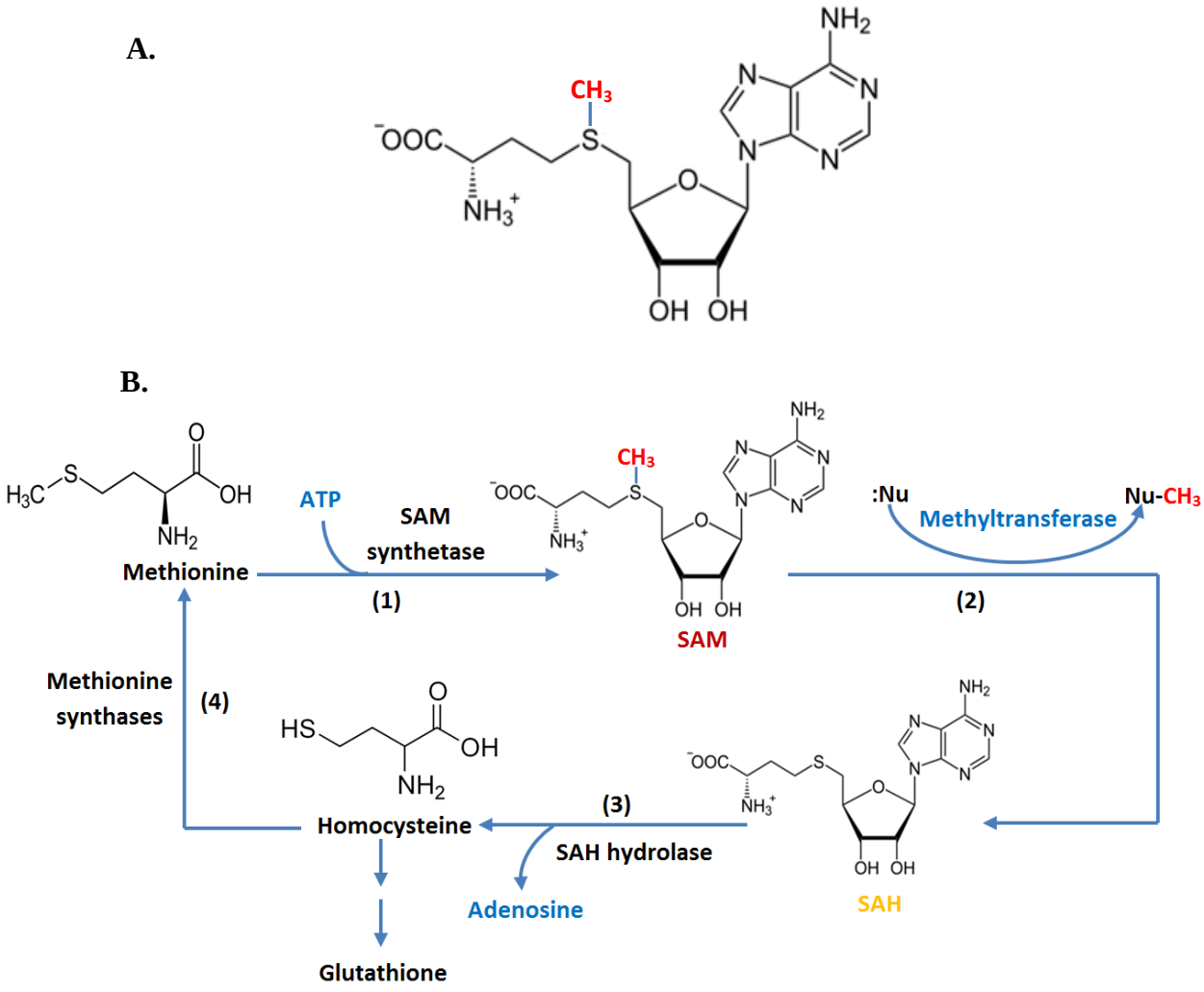


Figure 9. **S-Adenosyl-L-methionine (SAM).** (**A**) Chemical structure of SAM; (**B**) SAM cycle

2.1.2 SAM-dependent methyltransferases

In human, it is estimated that about 1% of all human genes encode for MTases (208 proteins) while this number is even higher in yeast, constituting around 1.2% of the genome. In addition, 30% of human MTases are linked to diseases such as cancers or mental disorders, stating the importance of this protein family (Petrossian, & Clarke, 2011).

2.1.2.1 General mode of action

The great majority of MTases (EC 2.1.1) are known to use SAM as a methyl donor to catalyze the methylation reactions (Martin, & McMillan, 2002; Loenen, 2006). Those MTases employ the general action mode for the methyl transfer reaction by promoting favorable orientation and bringing the SAM methyl group into close contact with nucleophile targets such as carbon, oxygen, nitrogen or sulfur, resulting in nucleophilic attack on the strong electrophilic methyl group of SAM and allowing the transfer of the methyl group to the various substrates (DNA, RNA, proteins, lipids, polysaccharides…) through an S_N2 reaction (O'Hagan, & Schmidberger, 2010).

2.1.2.2 Classification

Since the first crystal structure of the SAM-dependent C5-cytosine HhaI DNA MTase was solved in 1993 (X. Cheng et al., 1993; Schubert et al., 2003), there have been significant increases in the number of determined MTase crystal structures thanks to more advanced methods in structural biology. Although those SAM-dependent MTases can be defined by their action on a variety of substrates (DNA, RNA, proteins, lipids, polysaccharides, small molecules...) as well as on different atoms for methylation (nitrogen, oxygen, carbon, sulfur ...), they are structurally classified into at least five different classes (class I-V) based on their remarkably distinct structural features (Schubert et al., 2003). This characteristic of this protein family has been considered as an interesting example of the functional evolutionary convergence.

2.1.2.2.1 Class I SAM-dependent MTases

Class I MTases form the largest group with more than 60% of the total MTases in human and yeast. It consists almost all DNA MTases, some RNA as well as protein MTases but also enzymes acting on small molecules (Schubert et al., 2003; Petrossian, & Clarke, 2011; Struck et al., 2012). Despite little sequence identity, members of this class adopt a conserved Rossmann-like fold including a central seven-stranded β-sheet ending with a reversed β hairpin at the C-terminal extremity, which is surrounded by α-helices (Figure 10). The order of the β strands is ↑3↑2↑1↑4↑5↓7↑6, with the β strand 7 being anti-parallel to the other parallel strands. Helices Z, A and B are positioned on one side of the β-sheet whereas helices C, D and E are on the other side (Martin, & McMillan, 2002; Schubert et al., 2003).

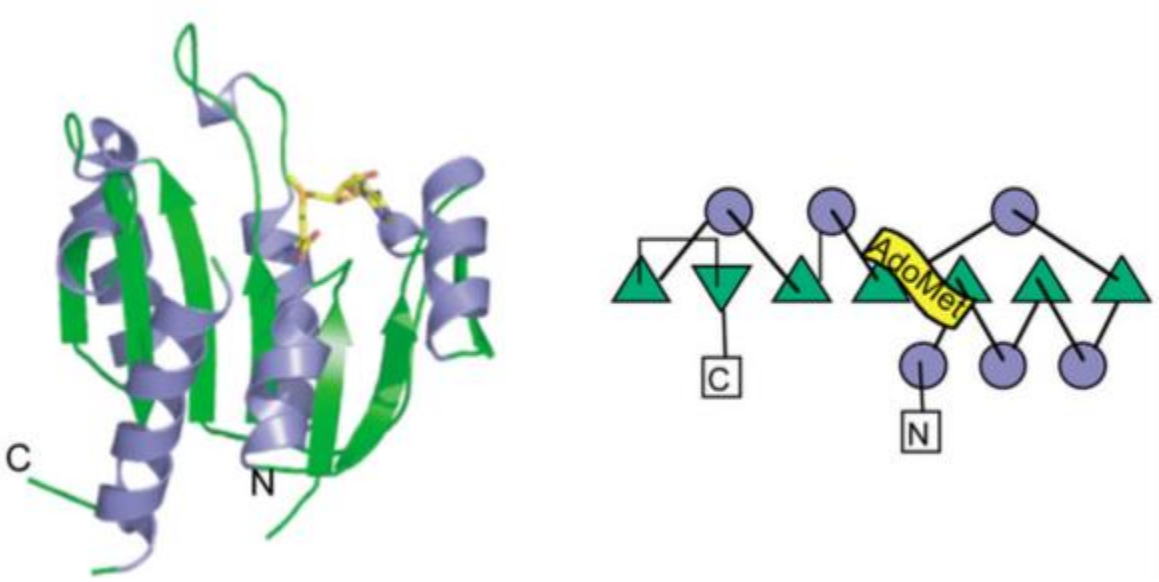


Figure 10. **Class I SAM-dependent MTases.** An example of class I MTase tertiary structure: M.HhaI (pdb: 6MHT) (left) and its topology diagram (right) (Schubert *et al.*, 2003)

The Rossmann-like fold is organized into two parts: the N-terminal part is responsible for SAM binding while the C-terminal part is mostly involved in substrate binding. The latter is known to be a tremendously variable region adapted to bind different kinds of substrates varying in shapes, sizes and chemistries (Martin, & McMillan, 2002). The SAM binding domain is is characterized by five highly conserved motifs (Kozbial, & Mushegian, 2005). Motif I, located in the loop between β-strand 1 and α-helix A, consists of the glycine-rich GxGxG sequence (or at least a GxG sequence), which directly interacts with the carboxypropyl moiety of SAM and is considered the hallmark of the SAM-binding site of the class I SAM-dependent MTases. Motif II

forms hydrogen bonds with the ribose hydroxyls of the SAM and is located in β-strand 2 and the adjoining turn while motif III is situated at the edge of β-strand 3 in the Rossmann-like fold and interacts with the SAM base. Motif IV interacts with the amino and sulfonium groups of the methionine moiety of SAM and encompasses β-strand 4 and the flanking loops followed by the helix corresponding to the motif V. Moreover, additional motifs such as IV, VI, VIII and/or X are known to be involved in substrate specificity whereas the motif V and VII play an important role mostly for the structural stability (Kozbial, & Mushegian, 2005).

2.1.2.2.2 Class II SAM-dependent MTases

Class II SAM-dependent MTases represent one of the smallest groups of MTases (Petrossian, & Clarke, 2011). These differ in their overall structural architecture (Figure 11) or their interaction with SAM compared to the class I MTases. This is the case of *E.coli* C-terminal MetH reactivation domain, which reactivates the cobalamin substrate by SAM-dependent methylation. This class II MTase has a three-dimensional structure dominated by a long central anti-parallel β-sheet flanked by groups of helices at both ends (Schubert et al., 2003). In addition, through an extended conformation, SAM binds into a shallow groove alongside the borders of the β-sheet and interacts with a conserved RxxxGY motif via hydrogen bonds (Schubert et al., 2003). In comparison to class I MTases, class II MTases undergo large conformational changes in order to position the substrate near the catalytic domain (Schubert et al., 2003).

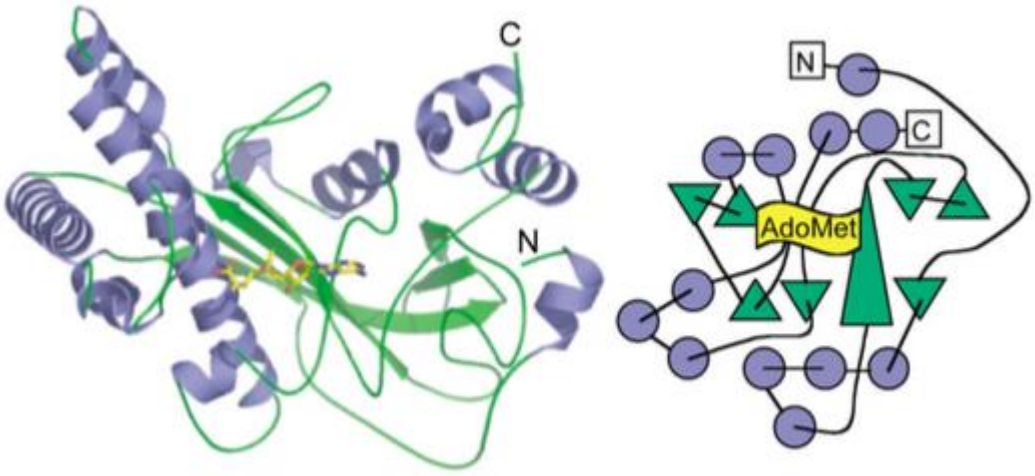


Figure 11. **Class II SAM-dependent MTases.** An example of class II MTase tertiary structure: MetH (pdb: 1MSK) (left) and its topology diagram (right) (Schubert *et al.*, 2003)

2.1.2.2.3 Class III SAM-dependent MTases

Like class II MTases, class III MTase family constitutes a very small group of MTases (Petrossian, & Clarke, 2011). This class of protein was first described in 1998 from the structure of CbiF, an MTase catalyzing SAM-dependent methylation on precorrin substrates through cobalamin biosynthesis (Schubert et al., 2003). Class III MTases contain two domains, each consisting of a five stranded β-sheet and 4 α-helices. The active site is located into a cleft at the interface between the two domains (Figure 12) (Schubert et al., 2003). Like class I MTases, members of this class exhibit a GxGxG motif at the C-terminal end of the first β-strand, but interestingly this motif is not involved in SAM binding. Meanwhile, SAM is bound to the protein between the two domains (Schubert et al., 1998; Schubert et al., 2003).

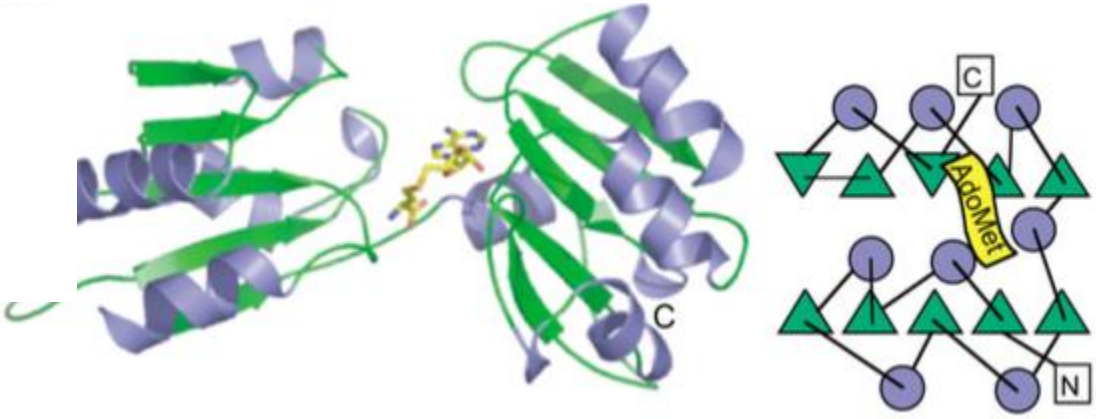


Figure 12. **Class III SAM-dependent MTases.** An example of class III MTase tertiary structure: CbiF (pdb: 1CBF) (left) and its topology diagram (right) (Schubert *et al.*, 2003)

2.1.2.2.4 Class IV SAM-dependent MTases

In human and yeast, the class IV MTase family is the third largest group (Petrossian, & Clarke, 2011), known as the SPOUT MTase superfamily (shortly for SpoU-TrmD). This goup mainly consists of tRNA and rRNA MTases (Anantharaman et al., 2002; Tkaczuk et al., 2007; R. J. Liu et al., 2013) but recently some protein MTases have been shown to belong to this group (Young et al., 2012; R. J. Liu et al., 2013). This class of MTases contains a unique SPOUT domain exhibiting an unusual α/β fold with a very deep topological knot (Figure 13) (Tkaczuk et al., 2007; R. J. Liu et al., 2013). This core domain including 5 or 6 parallel β-strands sandwiched by α-helices on both sides can be separated into 2 subdomains: (1) the N-terminal subdomain

forming a Rossmann-like fold; (2) the C-terminal subdomain with a deep conserved trefoil knot known to bind SAM (Tkaczuk et al., 2007; R. J. Liu et al., 2013). The class IV MTases can be divided into two subclasses: (1) the smallest SPOUT MTases such as TrmL and RlmH containing only the SPOUT domain in which the N-terminal subdomain is involved in substrate binding (R. J. Liu et al., 2013); (2) the larger SPOUT MTases like TrmH, RlmB and RsmE showing additional domains (for instance THUMP, PUA, OB fold, or L30e….) fused at the N- or C-termini or introduced into a linker between two subdomains of the SPOUT domain, responsible for the substrate binding (Tkaczuk et al., 2007; Petrossian, & Clarke, 2009; R. J. Liu et al., 2013). Another interesting point regarding this class IV MTase family is that almost all SPOUT members except the monomeric Trm10 protein reported so far were identified as dimeric proteins (Tkaczuk et al., 2007; Oerum et al., 2017). This characteristic is known to stabilize the SAM-binding loop in the knot of one monomer through interactions with the other monomer and to play an important role in the MTase activity as the active sites are created by residues from both monomers (Tkaczuk et al., 2007; Petrossian, & Clarke, 2009; Oerum et al., 2017).

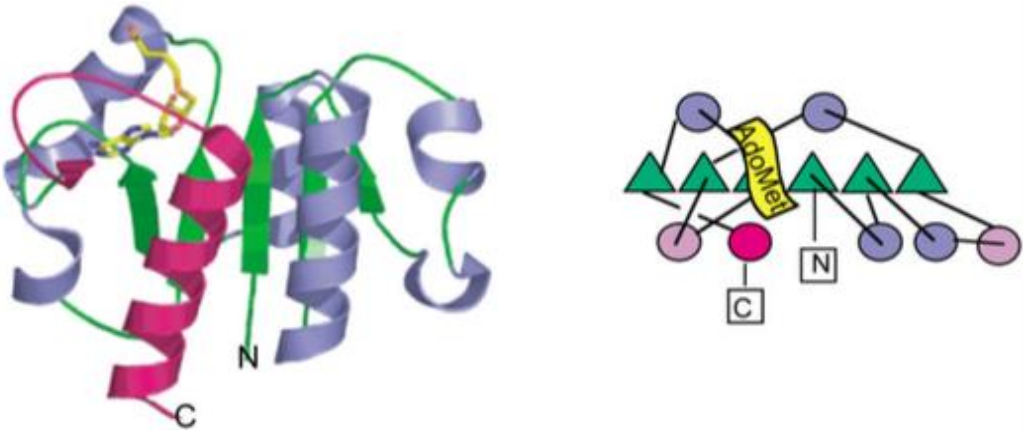


Figure 13. **Class IV SAM-dependent MTases.** An example of class IV MTase tertiary structure: YibK (pdb:1MXI) (left) and its topology diagram (right) (Schubert *et al.*, 2003)

2.1.2.2.5 Class V SAM-dependent MTases

The second largest family of SAM-dependent MTases is assigned to class V, with approximately 27% and 14% of the total MTases in human and yeast, respectively (Petrossian, & Clarke, 2011). The class V MTases contain a SET domain (Suppressor of variegation, Enhancer of zeste, Trithorax), and hence are also referred as SET-domain MTase superfamily. This family

encompasses almost all proteins known to catalyze histone lysine methylations crucial for the regulation of chromatin and gene expression. In addition to histones, the SET-domain MTases also methylate some other proteins like Rubisco (Dillon et al., 2005). The SET domain consists of eight curved β strands organized into three small β-sheets (Figure 14), including the C-terminus inserted below a surface loop to generate a knot-like structure as seen in the SPOUT MTases (Dillon et al., 2005; Petrossian, & Clarke, 2009). Class V MTases have sequence similarity in their N-terminal (N-SET) and C-terminal (C-SET) domains, which harbour conserved motifs I-II and motifs III-IV, respectively. Those are known to be responsible for catalysis, SAM-binding and substrate interaction (Petrossian, & Clarke, 2009). Similarly to class III MTases, the SAM bound to the SET domain binds to a shallow groove of the protein, formed by motif I, the N-terminal part of motif III and a tyrosine in motif IV (Schubert et al., 2003; Petrossian, & Clarke, 2009). One interesting point with regard to the SAM binding mode by proteins of this family is the presence of the GxG sequence in motif I similarly to the motif I glycine-rich GxGxG sequence found in class I MTases, despite no structural similarity between the two protein classes (Petrossian, & Clarke, 2009). Moreover, some MTases have the SET-domain flanked by diverse sequences named pre- and post-SET motifs. The pre-SET motif is known to stabilize the protein structure by interacting with different surfaces of the core SET domain while the post-SET motif constitutes part of the active site, vital for the MTase activity and maybe involved in the substrate recognition and specificity (Schubert et al., 2003; Qian, & Zhou, 2006).

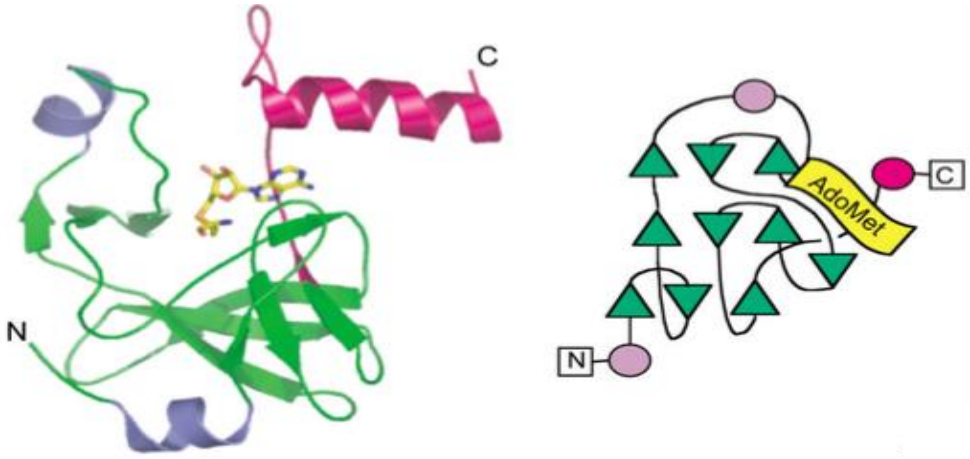


Figure 14. **Class V SAM-dependent MTases.** An example of class V MTase tertiary structure: Set7/9 (pdb:1O9S) (left) and its topology diagram (right) (Schubert *et al.*, 2003)

2.2 Methylation of translational machinery

Methylations are by far the most frequent and important PTMs for the control and optimal efficiency of mRNA translation. This is particularly the case for tRNAs, which are heavily methylated in order to enhance their stability as well as the efficiency and accuracy of translation. rRNAs are also targeted by various post-transcriptional modifications including 2'-OH methylation, base methylation, or more complex modifications for their maturation processes and functions (Sharma, & Lafontaine, 2015). Several recently and extensively studied modified nucleotides present in mRNAs, including methylations such as N6-methyladenosine (m^6A), N1-methyladenosine (m^1A), and 5-(hydroxyl)methylcytosine ((h)m^5C), have led to the emergence of the epitranscriptomics field (Dominissini et al., 2016; Gilbert et al., 2016). Finally, MTases also target ribosomal proteins and of translational factors (Polevoda, & Sherman, 2007).

2.2.1 tRNA methylation

According to the RNA modification database (http://mods.rna.albany.edu/mods/), 112 different modified nucleosides were found in all kinds of RNA in three domains of life to date. Among these, 93 modifications are present on tRNA molecules, with methylation being the major one. The four canonical nucleotides are the most regular substrates for tRNA MTases on either the base or the ribose moiety but modified nucleotides like pseudouridine (ψ), inosine (I) and more complex species are also subjected to methylation (Swinehart, & Jackman, 2015). Among tRNA MTases, almost all enzymes known to date belong to the class I Rossmann-like fold and class IV SPOUT-domain MTases. It is noteworthy that some recently discovered MTases could potentially form a novel class of MTase in the future (Swinehart, & Jackman, 2015).

2.2.1.1 Base methylation

The bases are the most common substrates of the tRNA MTases catalyzing modifications mainly at carbon, endocyclic nitrogen and exocyclic nitrogen atoms (Swinehart, & Jackman, 2015).

The methylation of the carbon atom at position 5 of pyrimidines (m^5C and m^5U) is present at multiple positions of tRNAs. These tRNA modifications are found throughout life,

with m^5U identified in all 3 domains of life while m^5C is present in Archaea and Eukarya (Hou, & Perona, 2010; Motorin, & Helm, 2011; Swinehart, & Jackman, 2015). Most tRNA MTases catalyzing m^5C and m^5U are class I MTases, with some representative members such as *E.coli* TrmA (yeast Trm2 and human TRMT2 homologs) catalyzing the conserved m^5U54 (also called ribothymidine) in the T-loop of tRNAs and yeast Trm4 (human NSUN2 homolog) generating m^5C at different positions (Hou, & Perona, 2010; Motorin, & Helm, 2011; Towns, & Begley, 2012; Swinehart, & Jackman, 2015). The catalytic mechanism of those MTases is probably as follows. First, a conserved cysteine acts as a nucleophile to attack C6 atom from the base, which then allows the C5 atom to attack the SAM methyl group. Second, the C5 proton is abstracted by a general base in order to generate the methylated product. For TrmA, the catalytic base has been identified as a glutamate while in Trm4, an aspartate fulfills this function (Hou, & Perona, 2010; Boschi-Muller, & Motorin, 2013; Swinehart, & Jackman, 2015). Moreover, methylation events can also occur on complex modifications such as cm^5U, $mcm^5(s^2)U$, ncm^5U… through more complex multi-step biosynthetic reactions. This will be discussed in detail in the next part of the thesis.

Methylations on the different endocyclic nitrogen atoms such as m^1A, m^1G, m^3G, m^3C, $m^1\psi$ and m^7G are found in almost all living cells except m^3C and $m^1\psi$ that are absent in Archaea and Bacteria, respectively (Swinehart, & Jackman, 2015). Some examples for this methylation group include the class I MTases like yeast Trm5 (human TRMT5) forming m^1G37 in Archaea and Eukarya; the yeast Trm6/Trm61 heterodimer (Trm61 is the catalytic subunit) (human TRM6) catalyzing m^1A58, a highly conserved modified A58 in the T-loop of tRNAs and also the SPOUT domain MTases such as bacterial TrmD forming m^1G37, yeast Trm10 (human TRMT10A) catalyzing formation of m^1G9 as well as archaeal TrmY for $m^1\psi 54$ formation (Motorin, & Helm, 2011; Towns, & Begley, 2012; Swinehart, & Jackman, 2015; Hori, 2017). In term of the potential methylation mechanism, some MTases (for instance Trm5 and Trm10) use general bases like glutamate and aspartate residues, respectively, which deprotonate N1 atom. This results in the nucleophilic attack of N1 on the SAM methyl group to form the methylated targets. In contrast, for m^7G class I MTases such as eukaryotic Trm8/Trm81 (human METTL1), bacterial TrmB or m^3C MTases like *Sc*Trm140, their catalytic mechanisms have not been clarified so far (Swinehart, & Jackman, 2015).

Methylation on the exocyclic nitrogen atoms like m^2G, $m^2{}_2G$, m^6A, and $m^6{}_2A$ is always catalyzed by class I MTases. The eukaryotic Trm112-Trm11 (human TRMT112-TRMT11) (where Trm11/TRMT11 are the catalytic subunit) catalyzes the formation of m^2G10 while archaeal TrmG10 is responsible for the formation of both m^2G10 and $m^2{}_2G10$ (Hirata et al., 2016). In addition, yeast Trm1 (human TRMT1) is known to form $m^2{}_2G26$ with m^2G26 as the intermediate while archaeal *Aquifex aeolicus* Trm1 modifies some tRNAs to generate m^2G26m^2G27 and $m^2{}_2G26m^2{}_2G27$ (Awai et al., 2009). The bacterial/archaeal Trm1 proteins contain a DPFG/DPPY conserved motif in their active site and the aspartate from this motif is likely acting as the general base to abstract a proton from N_2 atom, then promoting the nucleophilic attack on SAM methyl group for methylation to occur (Swinehart, & Jackman, 2015). Meanwhile, the m^6A modification also commonly seen in mRNA, rRNA and DNA is found in tRNAs at position 37 and is catalyzed by *E.coli* YfiC by a still unknown mechanism (Swinehart, & Jackman, 2015).

2.2.1.2 Ribose methylation

This methylation can be found on the 2' hydroxyl group of the ribose (*2'-O*-methylation) of any canonical nucleotides in all domains of life (Rana, & Ankri, 2016). A famous example for this group of MTases is bacterial TrmH (Trm3 in Eukarya), one of the founders for the SPOUT-domain MTases, generating a highly conserved *2'-O*-methylguanosine at position 18 (Gm18). This protein has been well-studied and its catalytic mode is proposed to act through a conserved arginine playing the role of the general base removing a proton from the 2'-hydroxyl group and then allowing the resulting oxygen to attack the SAM methyl group for the reaction to complete (Swinehart, & Jackman, 2015; Hori, 2017). Moreover, several important 2'-O-methylations are present at positions 32 and 34 in all domains of life. Some examples include SPOUT-domain MTases like *E.coli* TrmJ (Cm32 or Um32), TrmL (Cm34, $cmnm^5Um34$), aTrmJ (Cm32) but also class I MTases such as such as *Sc*Trm7/Trm732 (Cm32) and Trm7/Trm734 (Cm34) complexes or their human orthologues containing TRMT7 (also known as FTSJ1; Cm32, Um32, Cm34, Gm34). Meanwhile, some *2'-O*-methylations are specific to archaea (Cm56 catalyzed by SPOUT-domain aTrm56) or eukarya (Um44 formed by class-I *Sc*Trm44 or human METTL19) (Hori, 2017; Marchand et al., 2017).

2.2.1.3 Functions of tRNA methylations

Although the importance of tRNA modifications has often been underestimated as the loss of most single modifications usually results in only modest or no phenotypes, the PTMs in general and the methylations in particular are crucially involved in every aspects of the tRNA, namely its structure, function, and stability (Torres et al., 2014; Swinehart, & Jackman, 2015). The effects of these methylations on tRNAs can be classified based on the positions of the modifications.

The methylations in the anti-codon loop and at positions nearby are normally related to effects on the tRNA function, namely the fidelity of mRNA decoding. This is the case for mcm^5U, mcm^5s^2U, Cm and mnm^5Um at the wobble position 34, which are directly involved in decoding through the anti-codon:codon pairing and are necessary for the translation fidelity and efficiency (Letoquart et al., 2015b; Ranjan, & Rodnina, 2016; Tuorto, & Lyko, 2016; Hori, 2017). Meanwhile, the m^5C_{34} modification present in yeast $tRNA^{Leu(CAA)}$ is necessary for translation efficiency and its loss renders yeast hypersensitive to oxidative stress because of inefficient translation of UUG-containing stress response mRNAs (C. Gu et al., 2014; Swinehart, & Jackman, 2015; Ranjan, & Rodnina, 2016). Moreover, several modifications at position 37 such as a highly conserved m^1G37 are responsible for enhancing the translational accuracy, whereas chemically intricate modifications like wyosine (imG) and its derivative wybutosin (yW) in eukaryotic and archaeal phenylalanine-specific $tRNA^{Phe}$ are known to stabilize codon:anticodon interactions by providing base-stacking interactions between the anti-codon and the A-site codon. Such modifications have a key role in preventing translational frameshifting (Ranjan, & Rodnina, 2016; Tuorto, & Lyko, 2016).

The methylations of nucleotides located in the tRNA body usually affect the folding and stability of tRNAs. Some examples are the conserved modifications m^2G10 and $m^2{}_2G26$ at the stem of the D-loop, which both maintain the secondary and tertiary structure of tRNAs in the three domains of life (Lorenz et al., 2017; Vare et al., 2017). Meanwhile, two modifications (m^1G9 and m^1A9) have been identified to facilitate the correct folding of mitochondrial tRNAs (Vare et al., 2016; Oerum et al., 2017). It is also the case for m^5C48 and m^5C49, which are located at the junction between the variable loop and the T loop in archaeal and eukaryotic tRNAs. Furthermore, the m^5U_{54} (forming the conserved T54, also known as ribothymidine), and

the m^1A_{58} modifications, both located in the T loop of tRNAs, are known to stabilize a reverse Hoogsteen base-pairing interaction between those two bases, and thus necessary for the tertiary folding of the L-shaped tRNA (Hou, & Perona, 2010). Moreover, lack of m^1A58 catalyzed by Trm6/Trm61 (where Trm61 is the catalytic subunit) and m^7G46 modification generated by Trm8/Trm82 (where Trm8 is the catalytic subunit) results in rapid tRNA degradation (Alexandrov et al., 2006).

Apart from direct roles in tRNA decoding and structure, tRNA methylations also encompass other additional functions. For instance the 2´-*O*-methylation of the tRNA anticodon loop is known to ensure effective immune response against pathogens in plants (Ramirez et al., 2015). In addition, some modifications such as 2´-*O*-methylations alone or in combination with methylated bases have been found crucial for temperature adaptation in thermophilic and psychrophilic organisms (Lorenz et al., 2017; Vare et al., 2017). More importantly, growing evidence suggests that defects in tRNA methylations and MTases are involved in severe human disorders such as cancers, type II diabetes, and neurological diseases. Recent extensive efforts on understanding tRNA methylations and the enzymes responsible for these modifications could lead to future discovery of novel therapeutics (Torres et al., 2014; Grosjean, 2015).

2.2.2 rRNA methylation

Like tRNAs, rRNAs are frequently methylated in all living cells, however numbers and types of methylations vary from different organisms (Piekna-Przybylska et al., 2007; Piekna-Przybylska et al., 2008). *E. coli* has 19 base methylations and only four *2'-O*-methylations. The opposite pattern is observed in Eukaryotes. Indeed, budding yeast contains 10 base methylations but 55 ribose *2'-O*-methylations while in human, the numbers are 10 and around 100, respectively (Piekna-Przybylska et al., 2007; Piekna-Przybylska et al., 2008; Sharma, & Lafontaine, 2015; Krogh et al., 2016). Meanwhile, in archaea, rRNAs methylations like its eukaryotic counterparts are dominated by *2'-O*-methylations (Dennis et al., 2015). These differences between three domains of life seem to be due to the presence of box C/D snoRNPs in Archaea and Eukarya (Watkins, & Bohnsack, 2012; Krogh et al., 2016). Concerning the catalytic enzymes, most *2'-O*-methylations are catalyzed by small nucleolar RNPs (snoRNPs) guided by box C/D sno-RNAs, while base methylations are specifically added by conventional SAM-dependent MTases (Watkins, & Bohnsack, 2012; Sloan et al., 2016). It is also noteworthy that

2'-O-methylations are mostly added during early stages of ribosomal subunits maturation while the base modifications are generally believed to appear later despite no precise timing known for most of them till now (Sloan et al., 2016).

2.2.2.1 Base methylation

In yeast, 10 base methylations have been identified in rRNAs with six methylations located on the large subunit and four on the small one ((J. Yang et al., 2016); the enzymes responsible for these modifications are indicated in brackets). These SAM-dependent MTases catalyze mono- or di-methylation on specific atoms like N or C in both purine and pyrimidine rings (Sharma, & Lafontaine, 2015). The yeast 18S rRNAs contain $m^1acp^3\psi$1191 (ScEmg1) located at P-site, m^7G1575 (ScBud23-Trm112) at a P-site/E-site tRNA ridge and two $m^6{}_2$A1781/$m^6{}_2$A1782 (both generated by *Sc*Dim1 (Appel, & Maxwell, 2007; Motorin, & Helm, 2011; Gumienny et al., 2016)) situated close to the decoding center and platform (White et al., 2008; Sharma, & Lafontaine, 2015; Sharma et al., 2015; Sloan et al., 2016). In the 60S large subunit from *S. cerevisiae* yeast, the 25S rRNA includes two N1-adenosines m^1A645 (ScRrp8) located next to PTC and m^1A2142 (ScBmt2) at subunits interface, two C5-cytosines m^5C2278 (ScRcm1) also at subunit interface and m^5C2870 (Nop2) in the PTC, and two N3-uridines m^3U2634 (Bmt5) near PTC and m^3U2843 (Bmt6) near P-stalk (Peifer et al., 2013; Sharma et al., 2013; Bourgeois et al., 2015; Sharma, & Lafontaine, 2015; Sloan et al., 2016). Most of these MTases belong to class-I SAM-dependent MTases including Bud23, Dim1, Rrp8, Bmt2, Rmc1, Nop2, Bmt5 and Bmt6, whereas Emg1 is known as class-IV SPOUT-domain MTase (Sloan et al., 2016). These base modifications are conserved in human 18S rRNA. However, in human 28S rRNA, only 3 base methylations found in the yeast have been detected so far, including m^1A1322 (*Hs*NML), m^5C3782 (*Hs*NSUN5) and m^5C4447 (*Hs*NSUN1) (Sharma, & Lafontaine, 2015). Moreover, there is one 18S rRNA base methylation (m^6A) in human, at position 1832 (Maden, 1986), which is absent in yeast. However, no enzyme responsible for this methylation has been so far detected.

2.2.2.2 Ribose methylation

2'-O-methylation is one of the two most common modifications in eukaryotic rRNAs, together with pseudouridine. The *2'-O*-methylation is generally catalyzed by an addition of one

methyl group at the *2'-O*-position of the ribose on all canonical nucleosides. In Bacteria, this modification is known to be introduced by site specific or region specific MTases while in Eukarya and Archaea it is carried out by box C/D snoRNPs and box C/D sRNPs respectively, which are abundant in those organisms (Appel, & Maxwell, 2007; Dennis et al., 2015; Krogh et al., 2016; Taoka et al., 2016). Most box C/D snoRNAs consist of four conserved sequence motifs: boxes C, C' (5'-RUGAUGA-3') and boxes D, D' (5'-CUGA-3') in the order C-D'-C'-D from the 5' and 3' termini of the guide RNA, respectively (Appel, & Maxwell, 2007; Motorin, & Helm, 2011; Gumienny et al., 2016). Interestingly, this C/D box snoRNAs forms a classical K-turn motif facilitated by base pairing between the upstream regions of C and D boxes and then servs as a platform for C/D box snoRNP assembly (Appel, & Maxwell, 2007; Gumienny et al., 2016). In yeast and human, the C/D box snoRNPs contain a C/D box snoRNA associated with a core of four conserved proteins Nop1 (SAM-dependent MTase), Nop56, Nop58 and Snu13 while in archaea this core includes Nop1, Nop5 and L7Ae (Appel, & Maxwell, 2007; Motorin, & Helm, 2011). In this guide RNA complex, a short sequence complementary to the targeted position on rRNA is found upstream of boxes D, D' and the *2'-O*-methylation is catalyzed by the Nop1 protein (Galardi et al., 2002; Motorin, & Helm, 2011).

2.2.2.3 Functions of rRNA methylations

The roles of RNA methylations can be initially related to the stabilization of secondary and tertiary structures of rRNAs themselves. For example, *2'-O*-methylations are known to stabilize helices through enhancing base stacking as well as provide nucleotides with greater hydrophobicity potentially benefiting from inter- or intramolecular interactions while base methylations such as m^3U can stimulate hairpin formation and m^5C improves the stability of guanine base-pairing (Liang et al., 2007; Sloan et al., 2016).

The methylations are not found randomly in rRNAs but occur at specific, highly conserved residues of both ribosomal subunits that cluster in crucial regions such as peptidyl transferase center, the A-, P- and E-sites, the mRNA binding channel, the polypeptide exit tunnel, and the inter-subunit interface, indicating their important roles in the ribosome functions and biogenesis (Decatur, & Fournier, 2002; Piekna-Przybylska et al., 2008; Sharma, & Lafontaine, 2015; Sloan et al., 2016). In fact, combined loss of some or several *2'-O*-methylations in these functional sites has been reported to result in severe effects such as: (1)

reduced translation rates; (2) lowered translation efficiency through increasing stop codon read-through activity; (3) increased sensitivity to ribosome-based antibiotics; (4) strongly delayed pre-rRNA processing and reduced RNA levels (Liang et al., 2007, 2009).

In addition, all base methylations in 18S rRNA and some in 25S rRNA are known to be necessary for the ribosome biogenesis. In contrast with most MTases responsible for 25S base methylation are not essential, almost all MTases involved in 18S rRNA base methylations are essential (Emg1, Nep1 and Dim1) or crucial for growth (Bud23) (Sharma, & Lafontaine, 2015). It is interesting to note that these 18S rRNA MTases themselves but not their enzymatic activity are important for ribosome biosynthesis (Sharma, & Lafontaine, 2015; Sloan et al., 2016). Moreover, it is suggested that some of these MTases (Dim1 and Bud23) take part in robust quality control mechanisms during ribosome biogenesis that probe binding of those MTases to precursor rRNAs as requirement for processing (Zorbas et al., 2015). Moreover, defects in base methylations in conserved ribosomal regions increase translational infidelity as well as sensitivity to antibiotics and oxidative stress (Sharma, & Lafontaine, 2015).

Importantly there are growing evidences linking rRNA methylation deficiency and various human disorders such as cancers, Bowen-Conradi syndrome (neurodegenerative disease), obesity, Williams-Beuren syndrome…. This once again highlights the crucial roles these methylations play in the cells. However, the mechanisms responsible for these diseases and the potential role of these modifications still remain ambiguous for most cases (Sharma, & Lafontaine, 2015; Sloan et al., 2016).

2.2.3 mRNA methylation

Besides the modifications necessary for the maturation of mRNAs, namely 5’ end capping, 3’ end polyadenylation and splicing as discussed previously (section I.2), mRNAs like other RNA species are also targeted by methylations (Motorin, & Helm, 2011). However, in contrast to tRNA and rRNA methylations, which have been extensively studied for decades, the prominence of mRNA methylations has been neglected due to lack of efficient detection strategies. History of mRNA methylation started in 1974 when an internal m^6A was discovered in mRNA, followed by more than three-decades gap with less interest and efforts because it was thought to originate from contamination. It then made a remarkable comeback with great

achievements recently that contributed to the birth of the novel field – Epitranscriptomics (Meyer et al., 2012; Schwartz, 2016). So far, eukaryotic mRNAs are known to contain m^6A, m^1A and m^5C methylations (Hoernes, & Erlacher, 2017).

2.2.3.1 m^6A

m^6A is among the first and the most abundant PTM identified in mRNA. It is present in most eukaryotes and to some extent in prokaryotes (X. Deng et al., 2015). In eukaryotes, the m^6A constitutes around 1-3 residues per mRNA accounting for 0.1-0.5% of all A nucleotides (Gilbert et al., 2016; Hoernes, & Erlacher, 2017). On eukaryotic mRNAs, m^6A is enriched mostly in 3' UTR, near stop codon, and is also found in long exons as well as transcriptional start sites of mRNAs, generally belonging to a consensus motif A/G[G>A]m^6AC[U>A>C] (Dominissini et al., 2012; Meyer et al., 2012; Schwartz et al., 2014; Hoernes, & Erlacher, 2017). This m^6A methylation is reversible, catalyzed and recognized by a set of enzymes including writers, erasers and readers. The m^6A writers are a multi-protein complex acting like a MTase. This complex includes METTL3, METTL14 and WTAP proteins. METTL3 is the catalytic subunit while METTL14 is its activator and WTAP (Wilms' tumor 1-associating protein) plays a role in the complex localization (J. Liu et al., 2014; Sledz, & Jinek, 2016; P. Wang et al., 2016; X. Wang et al., 2016; Hoernes, & Erlacher, 2017). The m^6A erasers catalyze the demethylation of m^6A into A. So far, two enzymes have been described: FTO (fat mass and obesity-associated protein) and ALKBH5 in which the former catalyzes the reaction through two labile intermediates, N6-hydroxymethyladenosine (hm^6A) and N6-formyladenosine (f^6A) before formation of A while the latter directly converts m^6A to A (Hoernes, & Erlacher, 2017). Meanwhile, the roles of m^6A are to recruit specific binding proteins called m^6A readers, which can be direct or indirect m^6A-binding proteins. The direct readers have been characterized such as YTH domain-containing proteins YTHDF1-3 and YTHDC1, which precise biological roles are currently studied. The YTHDF2 is the best characterized one and it is known to target m^6A containing mRNAs to decay sites in mammalian cells by directly recognizing m^6A and in turn determining the half-life of the respective mRNAs. In addition, the YTHDF1 and YTHDF2 promote translation initiation via interacting possibly with the ribosome initiation factors within 3' end and 5'end regions of mRNAs, respectively (Hoernes, & Erlacher, 2017). Meanwhile, an indirect reader HNRNPC (heterogeneous nuclear ribonucleoprotein C) binds mRNAs via

changed local RNA structures triggered by the m^6A modification and is known to influence alternative splicing (Hoernes, & Erlacher, 2017).

2.2.3.2 m^1A

In addition to being a well-characterized modification in tRNAs and rRNAs, where it has crucial impacts on the RNA structure and function, m^1A is also found on thousands of different mRNA transcripts in eukaryotic cells ranging from yeast to mammals, constituting an average of 20% transcripts in human (Dominissini et al., 2016). However, this m^1A modification is present in mammalian mRNA at a low relative abundance with m^1A/A ratio of around 0.015%–0.054% in cell lines and up to 0.16% in tissues (Dominissini et al., 2016; Y. Wang, & Jia, 2016). In contrast to mRNA m^6A, which still forms base-pair with thymidine through the reverse transcription, m^1A is known to disturb the Watson-Crick base-pairs, leading to reverse transcription stops and read-throughs together with mismatches (Y. Wang, & Jia, 2016; J. Song, & Yi, 2017). There are two interesting features of m^1A employed for the identification of its location: (1) m^1A is able to rearrange to m^6A in mRNA under alkaline conditions (Dimroth rearrangement); (2) m^1A in mRNA is reversible and de-methylated by ALKBH3 (human ortholog of *E.coli* AlkB) (Dominissini et al., 2016; Li et al., 2016; Y. Wang, & Jia, 2016). Compared to m^6A, which is mostly found near stop codons and within 3' UTR, m^1A is enriched within the 5' UTR, around the start codon upstream of the first splicing site (Dominissini et al., 2016; Li et al., 2016). Moreover, while enzymes responsible for m^6A formation in mRNAs are well known, the MTase needed for m^1A formation is still unknown (J. Song, & Yi, 2017).

2.2.3.3 m^5C

Like m^1A, m^5C is found stable and mostly abundant in tRNAs and rRNAs but is also present in mRNAs (Squires et al., 2012; X. Yang et al., 2017). Quite differently from m^6A and m^1A sites, the m^5C sites are significantly positioned just downstream of the 5'UTR and distributed mostly in the coding sequence comprising 45% while 35% m^5C are sited at both 5' and 3' UTRs, and the remaining at intron sequences. Moreover, the consensus motifs for the mRNA sequence containing m^5C modification have been identified, including CG (55%), CHG (28%, H = A, C, U) and CHH (17%) (X. Yang et al., 2017). Pertaining to MTases for this mRNA modification, it is now known that NSUN2, a m^5C tRNA MTase, acts as a writer for

mRNA m^5C (Squires et al., 2012; X. Yang et al., 2017), while ALYREF/THOC4, a mammalian mRNA export adaptor, was recently identified as mRNA m^5C reader (X. Yang et al., 2017). Interestingly, the mRNA m^5C can be converted through oxidation to hm^5C (5-hydroxymethylcytosine) by Tet proteins (Ten–eleven translocation protein family) in a similar manner to that in DNA (Fu et al., 2014; Delatte et al., 2016).

2.2.3.4 Functions of mRNA methylations

As part of post-transcriptional modifications, mRNA methylations are involved in different steps of gene expression. In fact, m^6A is known to play crucial roles in several aspects of mRNAs: (1) mRNA processing, namely mRNA splicing through a regulatory mechanism of FTO-dependent m^6A demethylation (Zhao et al., 2014); (2) mRNA stability as inversely correlated with m^6A methylation levels since binding of the YTHDF2 reader to m^6A-containing mRNA directs the complex to cellular RNA decay sites (X. Wang et al., 2014) and as deletion of METTL3 and METTL14 in mouse leads to increased expression of mRNAs encoding many developmental regulators (Y. Wang et al., 2014); (3) mRNA export as knockout of the eraser ALKBH5 accelerates mRNA export while that of METTL3 results in delay of mRNA export to the cytoplasm (Fustin et al., 2013; Zheng et al., 2013); (4) translation efficiency through a reader YTHDF1-mediated mechanism in which YTHDF1 binds mRNAs at m^6A and then recruits translational initiation factor eIF3 for directly promoting the translation of m^6A-containing mRNAs (X. Wang et al., 2015), and also through a direct effect on elongation kinetics (Choi et al., 2016). Moreover, m^6A is also involved in circadian clock function by which the circadian period is prolonged in case of clock gene-specific METTL3 silencing (Fustin et al., 2013). Meanwhile, the biological function of m^1A and m^5C in mRNAs is not well understood although those modifications could have a putative role in mRNA translation (J. Song, & Yi, 2017). It is also important to mention that the proteins performing those methylations are linked to several human disorders such as cancers (Squires et al., 2012; Zhang et al., 2016), obesity (Dina et al., 2007), Alzheimer's disease (Keller et al., 2011), and mental disability (Abbasi-Moheb et al., 2012).

2.2.4 Translational protein methylation

Similarly to nucleic acids, translation-related proteins including translational factors and r-proteins are also highly methylated.

In *S. cerevisiae*, elongation factors such as eEF1A, eEF2 and the fungal-specific eEF3 are heavily methylated. The *Sc* eEF1A is subjected to several methylation events, *i.e.* mono-methylation on lysines 30 (K30) and 390 (K390), di-methylation on lysine 316 (K316) and tri-methylation on lysine 79 (K79) (Couttas et al., 2012; Dzialo et al., 2014), the recently identified tri-methylation on N-terminal glycine 2 (G2) and di-methylation on lysine 3 (K3) (Hamey et al., 2016). *Sc* eEF2 is di-methylated on K613 and tri-methylated on K509 while *Sc* eEF3 is also known to be the subject of tri-methylations on three positions, namely K187, K196 and K789 (Couttas et al., 2012; Dzialo et al., 2014). Until now, the MTases responsible for those methylations are well characterized for *Sc* eEF1A namely Efm1 (K30), Efm4 (K316), Efm5 (K79), Efm6 (K390) and Efm7 (G2 and K3) and for *Sc* eEF2 including Efm2 (K613) and Efm3 (K509); whereas little is known in the case of *Sc* eEF3 (Hamey et al., 2016). The biological functions of those methylations are still unclear for eEF1A and eEF3. However, knockout of Efm2 or Efm3 abolishing eEF2 methylations, renders yeast sensitive to a translation inhibitor like sordarin and also results in translational infidelity by increasing frameshifting during translation, supporting the importance of those methylations for eEF2 function during translation elongation (Davydova et al., 2014; Dzialo et al., 2014).

The class I release factors eRF1 in eukaryotes and RF1, RF2 in bacteria all contain a universally conserved GGQ motif, which is N5-methylated on the side chain of the Q residue. In bacteria, this methylation is catalyzed by the PrmC MTase (also known as HemK) and this modification is known to ensure normal translation termination *in vivo* as well as to increase peptide release activity *in vitro* (Dincbas-Renqvist et al., 2000; Graille et al., 2005; Mora et al., 2007; Liger et al., 2011). In yeast, the N5-methylated Q is generated by the Mtq2-Trm112 complex in which Mtq2 is a class I MTase acting as the catalytic subunit (Heurgue-Hamard et al., 2005; Heurgue-Hamard et al., 2006). However, the function of this eRF1 methylation is not known to date.

Furthermore, methylations are also found in ribosomal proteins from a wide range of organisms in both prokaryotes and eukaryotes. In bacteria, there are six methylated r-proteins including one from the small ribosomal subunit (S11) and five from the large subunit (L3, L11,

L7/L12, L16, and L33). Three MTases namely YfcB, PrmA and PrmB specific for L3, L11 and L3, respectively, have been characterized so far. Moreover, some of those methylations are known to participate in ribosome functions through interaction with translation factors as well as in ribosome biogenesis (Nesterchuk et al., 2011). In *S. cerevisiae*, four r-proteins from the 40S are methylated, including Rps2, Rps3, Rps25a/b and Rps27a/b, while six r-proteins (Rpl1a/b, Rpl3, Rpl12a/b, Rpl23a/b, Rpl42a/b and Rpl43) from the 60S contain methylations (Shirai et al., 2010; Al-Hadid et al., 2016a). Several MTases modifying those r-proteins have been studied including class I MTases such as Rkm5 (Rpl1a/b on K46), Hpm1 (Rpl3 on H243), Ntm1 (Rpl12a/b on N-terminal proline and Rps25a/b on N-terminal proline) and Rmt2 (Rpl12a/b on R66); SET-domain MTases such as Rkm1 (Rpl23a/b on specific K), Rkm2 (Rpl12a/b on K) and Rkm3 and Rkm4 (Rpl42a/b on K) (Young et al., 2012). The biological functions of those r-protein methylations are quite well characterized, and are known to be involved in ribosome biogenesis, translation regulation as well as translation fidelity (Shirai et al., 2010; Al-Hadid et al., 2014; Davydova et al., 2014; Al-Hadid et al., 2016a; Al-Hadid et al., 2016b).

II. Current knowledge about eukaryotic Trm112 network

In eukaryotes, Trm112 is a small zinc finger protein known to act as an activating platform for the functions as well as stability of at least four MTases modifying factors of the eukaryotic translational apparatus including rRNA (Bud23), tRNAs (Trm9 and Trm11) and class I release factor eRF1 (Mtq2), perfectly illustrating the important effects of methylation in translation. This chapter and the next one are adapted from our recent review published in Biomolecules, 2017 (*Trm112, a protein activator of methyltransferases modifying actors of the eukaryotic translational apparatus* (Gabrielle Bourgeois, Juliette Létoquart, Nhan van Tran and Marc Graille)).

1. Trm112

Trm112, a 15 kDa protein, is widely conserved in eukaryotic organisms. It has been mostly characterized by studies first conducted in yeast and later in human cells. In baker's yeast, the TRM112 gene was initially considered as an essential gene from a large-scale survey of the growth phenotype resulting from systematic deletion of individual genes (Giaever et al., 2002).

However, further studies revealed that *trm112*Δ yeast strain is very sick but still viable (Purushothaman et al., 2005; Mazauric et al., 2010; Chen et al., 2011; Figaro et al., 2012). Studies performed on SMO2, the *A. thaliana* Trm112 ortholog, have shown that as in *S. cerevisiae*, the inactivation of SMO2 gene leads to a defect in cell growth (Hu et al., 2010). SMO2 is also required for proper cell division and development, but the mechanisms underlying these phenotypes are still unknown. Finally, mouse Trm112 ortholog is strongly and ubiquitously expressed during mouse embryo development (T. Gu et al., 2012).

Sequence alignment of Trm112 orthologs from the three domains of life and crystal structures of eukaryotic Trm112 proteins either in an isolated form (Heurgue-Hamard et al., 2006) or in complex with MTase partners (see below, (Liger et al., 2011; Letoquart et al., 2014; Letoquart et al., 2015a)) have revealed an organization into two domains. The first domain, contributed by residues from the N- and C-terminal extremities of eukaryotic Trm112 proteins, is conserved within the three domains of life. It folds as a zinc-knuckle (Zn-knuckle) domain, composed of a short α-helix (α1) packed against the concave face of a curved anti-parallel β-sheet (Figure 15A). In the structure of isolated ScTrm112 (Heurgue-Hamard et al., 2006), this β-sheet is composed of three β-strands and Trm112 C-terminal extremity folds back onto a hydrophobic region of the Zn-knuckle domain. In the crystal structures of Trm112-MTase complexes (Liger et al., 2011; Letoquart et al., 2014; Letoquart et al., 2015a), this Trm112 C-terminal extremity adopts a radically different conformation and folds as a fourth β-strand (β4), which is engaged in the interaction with the MTase partners . The second domain is contributed by residues from the central region of Trm112 eukaryotic proteins and is absent in bacterial as well as in some archaeal orthologs. Depending on the solved structures, this later domain is formed by 3 or 4 α-helices.

All structures of eukaryotic Trm112 solved to date are from fungi (*S. cerevisiae* (Heurgue-Hamard et al., 2006; Letoquart et al., 2014), *Yarrowia lipolytica* (Letoquart et al., 2015a)) or from an intracellular parasite (*Encephalitozoon cuniculi* (Liger et al., 2011)) and they all exhibit one zinc atom coordinated by four cysteine residues in the so-called Zn-knuckle domain. These residues belong to two well-conserved motifs ($CX_{3-4}C$ and CX_2C from the N- and C-terminal parts, respectively; where X is for any amino acid; Figure 15B). However, these four cysteine residues are not conserved in metazoan Trm112 proteins, suggesting that Trm112 does not bind

zinc in these organisms. This is indeed the case for human TRMT112 (G. Bourgeois and M. Graille, unpublished results). A similar conservation scheme has been already observed for Ski2 helicase, a component of the SKI complex involved in 3' to 5' mRNA decay in eukaryotic organisms. Indeed, fungal Ski2 orthologs harbor a zinc binding site formed by four conserved cysteine residues while in metazoan Ski2 proteins, these residues are not conserved but residues present at the corresponding positions may play the same structural role (Halbach et al., 2012).

Moreover, Trm112 orthologs are also found in bacteria, which are typically quite short (usual length around 60 residues) and the representative member from this family is YcaR from *E. coli*. According to the few NMR (nuclear magnetic resonance) structures determined by a structural genomics consortium (PDB codes 2KPI and 2JS4; Figure 16A, these orthologs contain only a Zn-knuckle domain, which is highly similar to the corresponding domain in eukaryotic Trm112 proteins (rmsd value of 1.7-1.8Å; Figure 16B). In these proteins, one zinc atom is coordinated by conserved cysteine and/or aspartic acid residues. To our knowledge, nothing is known about the function of bacterial Trm112 orthologs. No genes with significant sequence similarities with eukaryotic MTases known to interact with Trm112 could be detected in bacterial genomes. Furthermore, G_{10} on bacterial tRNAs is not methylated (Cantara et al., 2011) and the wobble uridine (U_{34}) from tRNAs reading codons ending with a purine harbors a mnm^{5}U (5-methylaminomethyl-uridine) modification catalyzed by the MnmE-MnmG complex and the MnmC bifunctional enzyme (Armengod et al., 2012). Meanwhile, the bacterial class I translation termination factors RF1 and RF2 are methylated on the glutamine side chain of their GGQ motif by the PrmC MTase, which is active on its own (Heurgué-Hamard et al., 2002; Nakahigashi et al., 2002; Graille et al., 2005). Finally, the only known m^{7}G nucleotide found in 16S rRNA is located at position 527 in *E. coli* (Isaksson, & Phillips, 1968), a position differing radically from the one in eukaryotic 18S rRNA (G_{1575}), and this modification is catalyzed by the RsmG MTase (Nishimura et al., 2007). Hence, bacterial Trm112 are strongly conserved proteins with still unknown function. However, in some bacteria, Trm112 is fused to MTase domains, suggesting that similarly to eukaryotic Trm112, bacterial orthologs could interact with MTases. This is further supported by the strong structural similarity between bacterial and eukaryotic Trm112, in particular in the region involved in the interaction with MTases (Figure 16B). Future studies aimed at clarifying the role and the potential partners of these bacterial proteins are definitely needed.

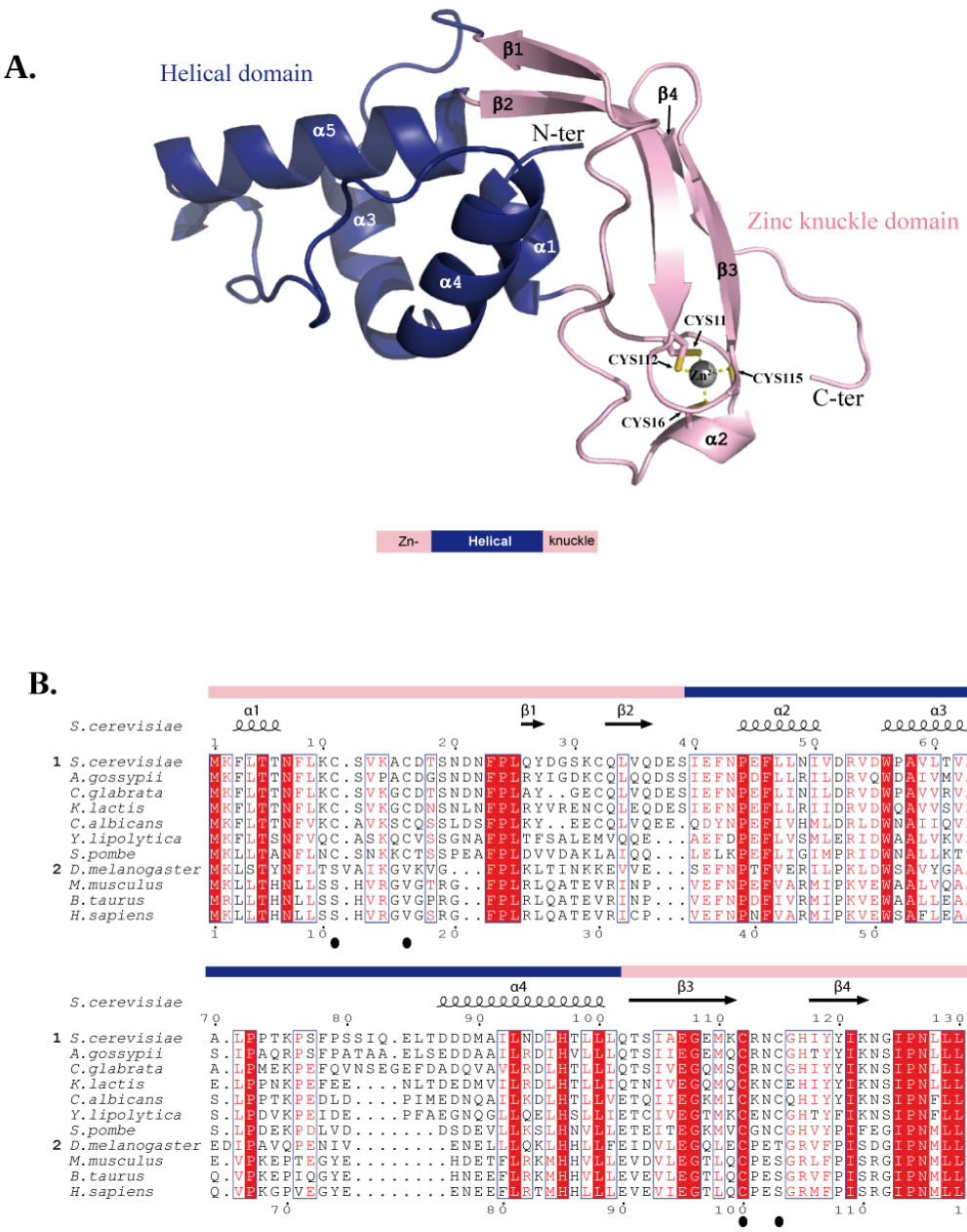


Figure 15. **Organization of eukaryotic Trm112 proteins.** (**A**) Ribbon representation of the crystal structure of isolated *S. cerevisiae* Trm112 protein with a schematic representation of eukaryotic Trm112 shown below with the domain's color code. (**B**) Sequence alignment of eukaryotic Trm112 protein sequences. Amino acids forming the Zn-knuckle and helical domains are identified by pink and blue bars, respectively, above the sequences. The positions of the four cysteine residues coordinating zinc atom in the structures of fungal and *E. cuniculi* Trm112 proteins are indicated by black spheres below the alignment. Secondary structure elements as observed in the structure of *S. cerevisiae* Bud23-Trm112 complex are indicated above the sequences (Letoquart et al., 2014). Sequences have been divided into two subgroups: fungal proteins (subgroup 1) and metazoans (subgroup 2). Strictly conserved residues are in white on a red background. Strongly conserved residues are in red. This figure was generated using the Espript server (Robert, & Gouet, 2014).

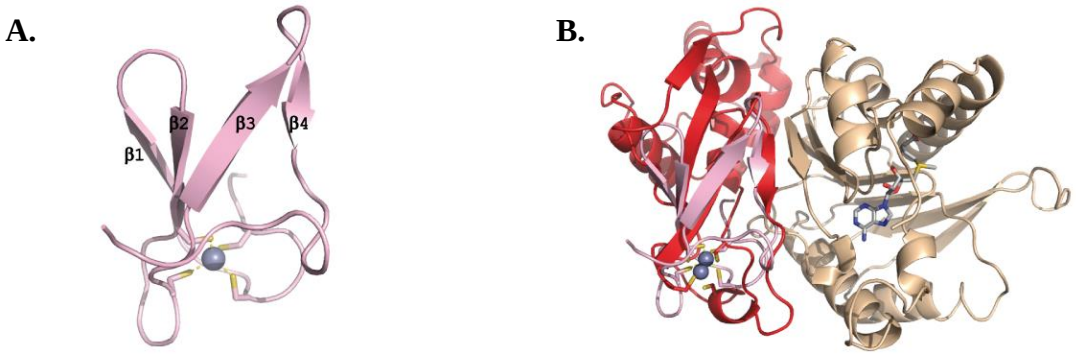


Figure 16. **Bacterial Trm112**. (**A**) Ribbon representation of *Streptomyces coelicolor* SCO3027 protein NMR structure (PDB code: 2KPI). Cysteine residues coordinating the zinc atom (grey sphere) bound to the protein are shown as sticks. (**B**) Superimposition of *Streptomyces coelicolor* SCO3027 structure (pink) onto Trm112 in the *Sc*Bud23-Trm112 complex (Trm112 and Bud23 are colored red and beige, respectively). The SAM molecule bound to Bud23 is shown as grey sticks.

2. Eukaryotic Trm112 interaction network

Eukaryotic Trm112 interaction network is defined as the network mainly studied in yeast in which a central conserved Trm112 protein interacts with and activates four MTases (Trm9, Trm11, Mtq2 and Bud23) (Figure 17), involved in different steps of protein synthesis.

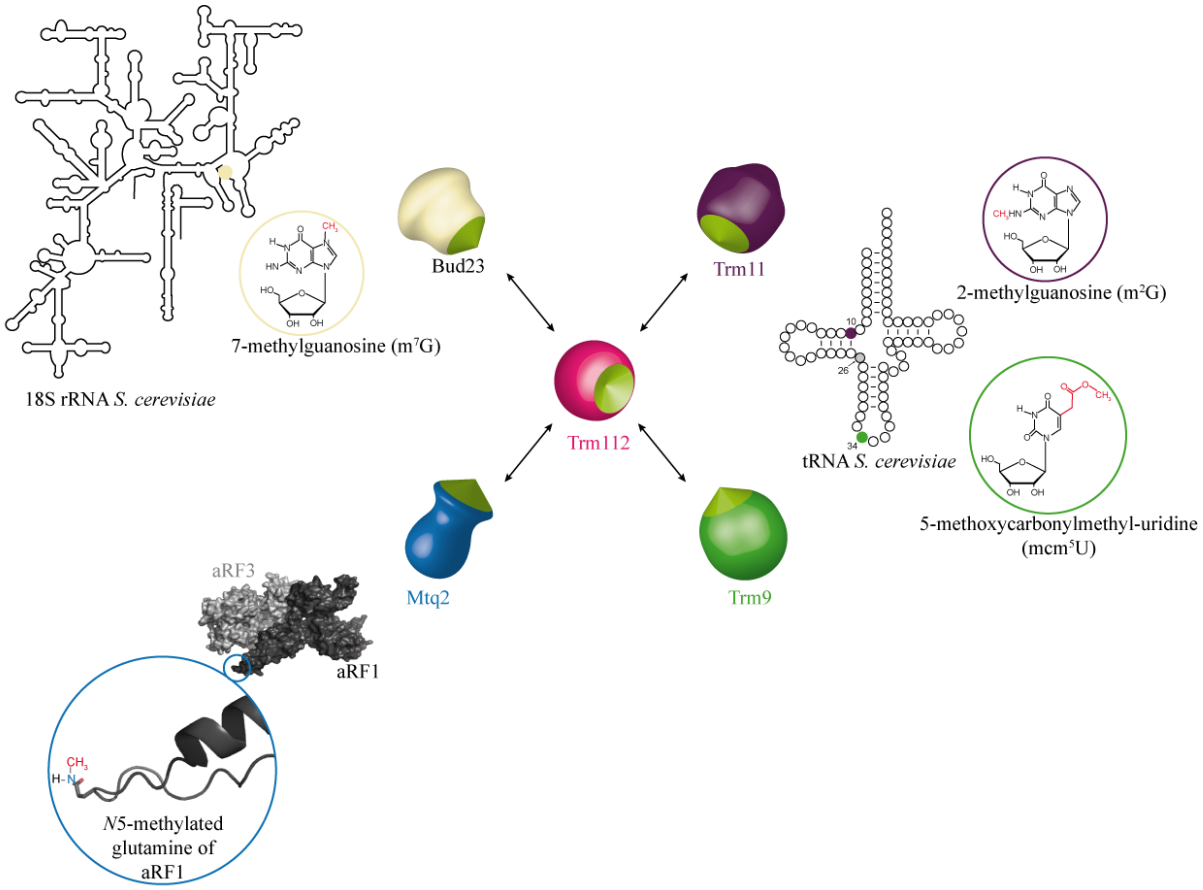


Figure 17. **Schematic representation of Trm112-MTase interaction network and of the substrates of these complexes.** The surface representation of the aRF1-aRF3 complex from *Aeropyrum pernix* archeon was generated using PDB code 3VMF (Kobayashi et al., 2012). Positions 10, 26 and 34 on a tRNA molecule are shown in purple, grey and green, respectively. Position of G_{1575} on 18S rRNA is shown as a beige sphere.

2.1. Trm9-Trm112

The large scale purification of budding yeast complexes using TAP-tag purification (tandem affinity purification) highlighted several partners for Trm112 including Trm9 (Gavin et al., 2002). The Trm9-Trm112 complex was further shown to be a tRNA MTase involved in the formation of $mcm^5(s^2)U$ (5-methoxycarbonylmethyl(2-thio)uridine) modifications at position 34 from the anticodon loop of some tRNAs (Figure 17; (Kalhor, & Clarke, 2003; Mazauric et al., 2010)).

In *S. cerevisiae*, the formation of $mcm^5(s^2)U$ involves at least 15 proteins. The first reaction, consisting in the addition of the carboxymethyl group at position 5 of the uracil to form 5-carboxymethyluridine (cm^5U), is catalyzed by the Elongator complex (Elp1-Elp6), which activity is regulated by seven additional proteins (Huang et al., 2008; S. Glatt et al., 2012; Sebastian Glatt et al., 2016). During the second step, the Trm9-Trm112 complex methylates cm^5U to yield mcm^5U. This modification is present at the wobble position of $tRNA^{Arg}_{(UCU)}$, $tRNA^{Gly}_{(UCC)}$, $tRNA^{Lys}_{(UUU)}$, $tRNA^{Gln}_{(UUG)}$ and $tRNA^{Glu}_{(UUC)}$. In the last three tRNAs, the oxygen atom attached to the C2 atom of the uracil ring is further substituted by a sulfur atom via the Ncs6/Urm1 synthesis pathway to form $mcm^5(s^2)U$ (Noma et al., 2009). The presence of the methyl group is important for an efficient thiolation as we detected tRNA mcm^5U_{34} product but not $mcm^5(s^2)U_{34}$ after *in vitro* enzymatic reaction of *Sc*Trm9-Trm112 on tRNAs purified from *trm9*Δ yeast strain. Furthermore, several groups observed a drop in $cm^5(s^2)U_{34}$ formation upon disruption of *TRM9* or *TRM112* genes (Mazauric et al., 2010; Chen et al., 2011; Letoquart et al., 2015a). In *S. cerevisiae*, this $mcm^5(s^2)U_{34}$ tRNA modification confers susceptibility to zymocin, a toxin secreted by the yeast *Kluyveromyces lactis,* which cleaves specifically the modified anticodon loop thereby inhibiting translation and leading to death (Lu et al., 2005).

The modifications in the anticodon loop of tRNAs are known to influence translation rate and fidelity of decoding. Indeed, the $mcm^5(s^2)U_{34}$ of tRNAs is involved in accurate and efficient reading of some codons in *S. cerevisiae* (Begley et al., 2007). Two studies based on integrated analysis of proteome, transcriptome and ribosome foot-printing, highlighted the link between these tRNA modifications and the regulation of global protein expression (Begley et al., 2007; W. Deng et al., 2015). The lack of $mcm^5(s^2)U_{34}$ modifications, upon deletion of *TRM9* gene, results in an increase of ribosomal pausing on mRNAs enriched with AGA and GAA codons. It is noteworthy that a significant portion of these mRNAs encodes proteins involved in protein synthesis, cell cycle control or DNA damage response and consequently, those proteins undergo a decreased expression in *trm9*Δ strain. These results rationalize the sensitive phenotype of the *trm9*Δ strain to methyl methanesulfonate (MMS) exposure. Indeed, MMS is a DNA damaging agent that methylates DNA mainly on N^7 and N^3 atoms of G and A bases, respectively, but also at other oxygen and nitrogen atoms of DNA bases, and thereby triggers DNA repair machineries. The *trm9*Δ strain also presents a delay in transition from G1 to S phase upon exposure to MMS

(Begley et al., 2007). Likewise, the absence of Trm9-catalyzed methylation causes translational infidelity and activation of protein stress response pathways (Patil et al., 2012).

To obtain information on Trm9 active site and Trm9-Trm112 complex organization, we solved the X-ray structure of this complex from the yeast *Yarrowia lipolytica* (*Yl*; (Letoquart et al., 2015a)). *Yl*Trm9 adopts the classical class I SAM-dependent MTase fold with a central seven-stranded β-sheet surrounded by two α-helices on each side. Moreover, a twisted two-stranded β-sheet forms a lid located on top of the C-terminal extremity of the central β-sheet and projects onto the active site (Figure 18A). Based on this structure, we mapped the active site of the *Sc*Trm9-Trm112 complex. This led to the identification of several mutants (R29A, H115A, R241A, Y243A and N271A) that strongly affect the affinity for tRNA (but not for SAM) as well as the enzymatic activity (k_{cat}). This supports a role of these conserved residues in tRNA binding, and particularly in the optimal orientation of the cm^5U_{34} nucleotide substrate in the active site, which is required for an efficient methyl transfer reaction by S_N2 mechanism (Figure 18B; (Letoquart et al., 2015a)).

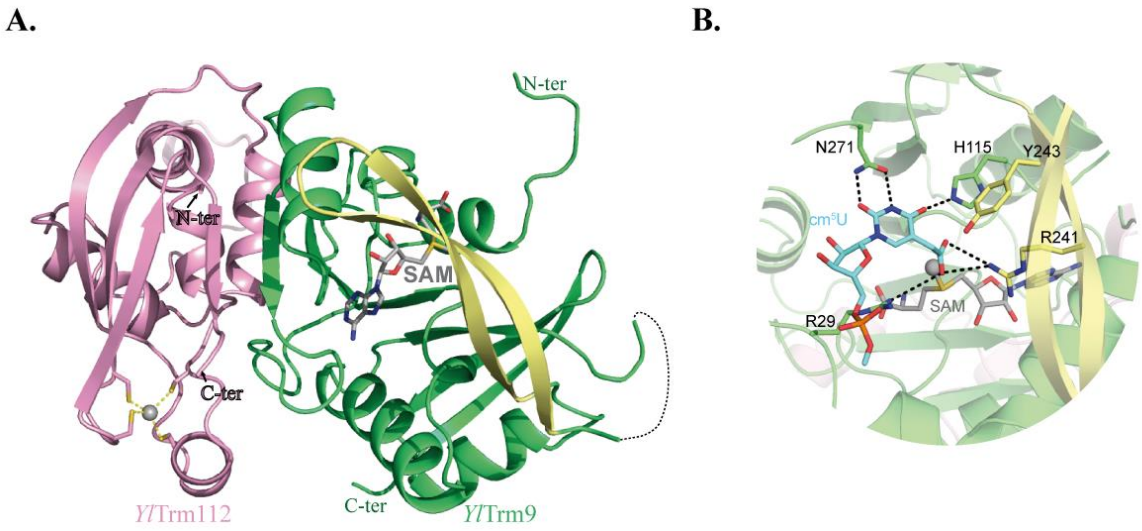


Figure 18. **Crystal structure of *Yl*Trm9-Trm112 complex.** (**A**) Ribbon representation of *Yl*Trm9-Trm112 complex. The SAM molecule (grey sticks), which was absent in the crystal structure, has been modeled by superimposing the SAM-bound structure of Bud23 onto *Yl*Trm9. Trm9 lid is colored yellow. (**B**) Model of cm^5U_{34} (blue sticks) docked into *Yl*Trm9 active site. The SAM methyl group to be transferred is depicted as a sphere. Residues are numbered according to *S. cerevisiae* protein.

Trm9 is largely conserved in eukaryotes. In human, two Trm9 orthologs are present: ABH8 and hTrm9L. ABH8 is a bifunctional enzyme encompassing a Trm9-like MTase domain converting cm^5U into mcm^5U, fused to an RRM domain (for RNA-recognition motif) and an AlkB-like domain responsible for the hydroxylation of mcm^5U into (S)-$mchm^5U$ ((S)-5-methoxycarbonylhydroxymethyluridine; (D. Fu et al., 2010; Y. Fu et al., 2010; Pastore et al., 2012)). Similarly to yeast, ABH8 needs to interact with TRMT112 to be active. The ABH8 protein is highly expressed in a variety of human cancer cells such as bladder cancer cells and its silencing suppresses tumor growth, angiogenesis and metastasis by inducing apoptosis of urothelial carcinoma cells (Shimada et al., 2009). ABH8 depletion also renders cells sensitive to DNA damaging agents (MMS) and to the bleomycin anti-cancer drug (D. Fu et al., 2010). Compared to ABH8, hTrm9L is only made of the MTase domain. The hTrm9L protein presents a Trm9-like tRNA MTase activity but to our knowledge, its interaction with TRMT112 has not been characterized (Begley et al., 2013). It acts as a negative regulator of tumor growth and the tumor cells deleted for the gene encoding hTrm9L are sensitive to paromomycin and gentamycin antibiotics (Begley et al., 2013). In *A. thaliana* (*At*), AT1G31600 (*At*TRM9), which is similar to Trm9 MTase, catalyzes the formation of mcm^5U_{34} and its activity is dependent on two Trm112 orthologs (*At*TRM112a and *At*TRM112b; (Leihne et al., 2011)). A second protein AT1G36310 (*At*ALKBH8), with similarity to ABH8 RRM and AlkB-like domains, has been shown to catalyze the hydroxylation of mcm^5U into (S)-$mchm^5U$ (Leihne et al., 2011).

2.2. Trm11-Trm112

The first Trm112 partner that has been described is the tRNA MTase Trm11, which catalyzes the formation of N^2-methylguanosine (m^2G) at position 10 of some tRNAs (Purushothaman et al., 2005). This modification is conserved in archaea and eukaryotes but absent in bacteria (Cantara et al., 2011). The m^2G_{10} is part of the body of the tRNA and is likely involved in tRNA folding and stability. The m^2G_{10} is stacked onto $m^2{}_2G_{26}$ nucleotide (Figure 17), which is methylated by Trm1. Interestingly, the simultaneous loss of both modifications induces strong growth defects (Purushothaman et al., 2005).

Bioinformatics analyses of eukaryotic Trm11 sequences suggested the presence of two domains. A N-terminal THUMP domain (for ThioUridine synthases, RNA Methyltransferases and pseudouridine synthetases; (Aravind, & Koonin, 2001)) formed by a NFLD (N-terminal

ferredoxin like domain) subdomain fused to a core-THUMP subdomain, and a C-terminal class I SAM-dependent MTase domain (Bujnicki et al., 2004; Purushothaman et al., 2005). Such modular organization has been confirmed by the recent crystal structure of the archaeal Trm11 ortholog from *Thermococcus kodakarensis* (Hirata et al., 2016) and is shared with archaeal Trm14 and bacterial TrmN, which are both responsible for m^2G formation at position 6 on tRNAs (Menezes et al., 2011; Fislage et al., 2012). In the 4-thiouridine synthase enzyme ThiI, the THUMP domain was shown to interact with the 3' CCA end (Neumann et al., 2014) and was then proposed to position the substrate nucleotide in the enzyme active site. It would then act as a molecular ruler that controls the distance between the tRNA CCA end and the nucleotide to be modified.

In *S. cerevisiae*, Trm112 is needed for the formation of m^2G_{10} modification by Trm11 (Purushothaman et al., 2005) whereas archaeal orthologs studied so far (PAB1283 from *P. abyssi* and aTrm11 from *T. kodakarensis*) are active on their own (Armengaud et al., 2004; Hirata et al., 2016). Initially, the m^2G_{10} modification could only be recapitulated *in vitro* using the *Sc*Trm11-Trm112 complex either purified directly from yeast cells (Purushothaman et al., 2005) or produced using wheat germ cell-free translation system (Okada et al., 2009), suggesting that post-translational modifications might be necessary for enzymatic activity. More recently, the *Sc*Trm11-Trm112 complex purified following co-expression of both subunits in *E. coli* turned out to be active on an *in vitro* synthesized tRNA. As mass spectrometry analyses showed that none of these proteins were post-translationally modified, post-translational/transcriptional modifications are then not mandatory for enzymatic activity (Bourgeois et al., 2017). Detailed analyses also indicate that Trm112 contributes to tRNA modification activity by influencing both SAM and tRNA binding either directly or indirectly (Bourgeois et al., 2017).

Although studies on Trm11 exclusively focused on *S. cerevisiae* and archaeal proteins so far, an orthologous gene was identified in human genome through BLAST searches. Indeed, the product of the C6orf75 gene is annotated as TRMT11, shares 34% identity and 54% similarity at the amino acid sequence level with Trm11, respectively, and has the same modular architecture as yeast Trm11. Two studies have linked defaults in human TRMT11 gene or transcript with advanced prostate cancer (Kohli et al., 2012; Yu et al., 2014) and this protein was shown to interact with at least three proteins from p53-family in fruit-fly (Lunardi et al., 2010). We have

successfully purified the human TRMT11-TRMT112 complex following co-expression in *E.coli* (unpublished data), indicating that similarly to yeast Trm11, its human counterpart interacts directly with TRMT112. Future studies are now needed to characterize the biochemical and biological functions of this human complex.

2.3. Mtq2-Trm112

Translation termination occurs when a stop codon is present in the ribosomal A-site. It is then not recognized by a cognate tRNA but by a protein factor known as class I release factor (RF1 or RF2 in bacteria, eRF1 in eukaryotes and aRF1 in archaea), which triggers the release of the newly synthesized proteins. Class I release factors are tRNA mimics as they recognize stop codon in the A-site through one domain and project a universally conserved GGQ motif (for Gly-Gly-Gln) from another domain into the ribosomal peptidyl transferase center (Klaholz, 2011; A. Brown et al., 2015). In bacteria, the side chain of the glutamine residue from this motif is N^5-methylated by the PrmC MTase (also known as HemK). This post-translational modification is important for normal translation termination *in vivo* and increases the affinity of the release factor for ribosomes (Pavlov et al., 1998; Dincbas-Renqvist et al., 2000; Heurgué-Hamard et al., 2002; Nakahigashi et al., 2002; Graille et al., 2005; Mora et al., 2007). Interestingly, the glutamine side chain of the GGQ motif from the eukaryotic class I release factor is also N^5-methylated (Figure 17; (Heurgue-Hamard et al., 2005; Polevoda et al., 2006)). The enzyme responsible for this modification in *S. cerevisiae* yeast is the Mtq2-Trm112 complex, where Mtq2 is the MTase catalytic subunit. Furthermore, this enzyme modifies eRF1 only when this later is associated with the GTP-bound form of class II translation termination factor eRF3 (Heurgue-Hamard et al., 2005; Heurgue-Hamard et al., 2006). Mtq2 orthologs have been described in human (HEMK2) and mouse (PRED28), where they also form a complex with the corresponding Trm112 orthologs and modify eRF1 translation termination factor (Figaro et al., 2008).

The crystal structure of Mtq2-Trm112 complex from *Encephalitozoon cuniculi* (*Ecu*) parasite obtained in the presence of SAM bound to the Mtq2 catalytic subunit, has confirmed the prediction that Mtq2 is a class I SAM-dependent MTase (Figure 19). It has also revealed the presence of a highly conserved surface surrounding the SAM methyl group (Liger et al., 2011). This region displays a negatively charged potential, which can ideally interact with the numerous

positively charged and conserved residues surrounding eRF1 GGQ motif. Furthermore, the crystal structure of the GTP-bound form of the archaeal aRF1-aRF3 complex reveals that aRF3 switches I and II regions, which are known to adopt different conformations between the GDP- or GTP-bound forms, are in close proximity of the GGQ motif (Kobayashi et al., 2012). Hence, these switches regions are very likely to interact directly with Mtq2-Trm112 complex, thereby explaining its specificity for the eRF1-eRF3-GTP form (Heurgue-Hamard et al., 2006). Comparison of the crystal structures of *Ecu*Mtq2-Trm112 and *E. coli* PrmC-RF1 complexes reveals that the NPPY active site signature from PrmC, which coordinates RF1 GGQ motif for proper methylation, structurally matches with Mtq2 NPPY signature, supporting a similar recognition mode of the GGQ motif across domains of life. Finally, it has also been shown that in the absence of TRMT112, HEMK2 does not exhibit enzymatic activity and cannot bind SAM while the purified HEMK2-TRMT112 complex is active and binds SAM. This indicates that TRMT112 activates HEMK2 by stimulating SAM binding (Figaro et al., 2008; Liger et al., 2011).

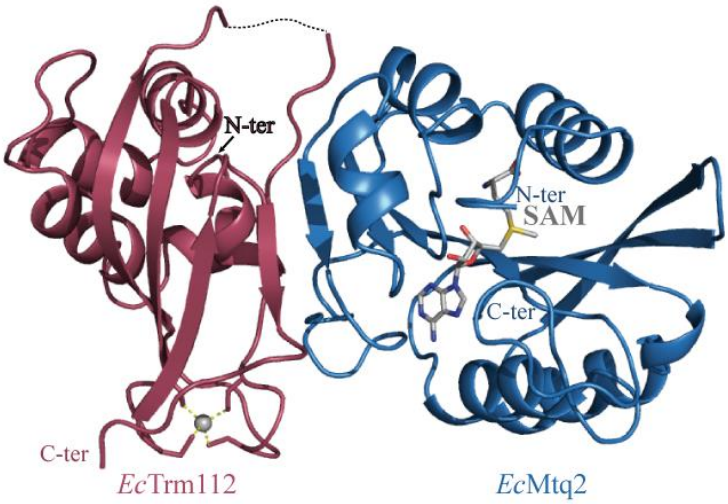


Figure 19. **Ribbon representation of the crystal structure of *Ecu*Mtq2-Trm112 complex bound to SAM** (grey sticks).

To date, the role of eRF1 methylation remains obscure but the conservation of this post-translational modification on the GGQ motif of at least bacterial and eukaryotic class I release factors, which adopt radically different three-dimensional structures, argues in favor of an important functional role. This is further supported by the growth defect phenotype of yeast cells lacking *MTQ2* gene (Polevoda et al., 2006; Mazauric et al., 2010; Chen et al., 2011), the cell proliferation defect with arrest in G1 phase of murine embryonic stem cells depleted of PRED28α isoform (Nie et al., 2009), the early mouse embryonic lethality upon disruption of PRED28α isoform (P. Liu et al., 2010) and the two-fold reduction in HEK293 human cells growth rate resulting from stable knock-down of *HEMK2* gene (P. Liu et al., 2010). Finally, murine PRED28α and human HEMK2 proteins appear to have a broad substrate specificity (Kusevic et al., 2016). Hence, future studies aimed at clarifying the role of the eukaryotic Mtq2-Trm112 complexes and of the methylation it is catalyzing, are needed.

2.4. Bud23-Trm112

The deletion of *TRM112* gene in *S. cerevisiae* results in a strong growth defect phenotype associated with strong defects in the synthesis of both ribosomal subunits and an increased sensitivity to paromomycin, a well-known inhibitor of protein synthesis (Figaro et al., 2012; Sardana, & Johnson, 2012). Recent studies have started to decipher Trm112 role in the very complex process of ribosome biogenesis.

2.4.1. The Bud23-Trm112 complex is involved in 40S maturation

To shed light on Trm112 role in 40S ribosomal subunit synthesis, a TAP-tag purification was conducted under milder conditions than in the initial large-scale study performed by Gavin and coworkers (Gavin et al., 2002) and new potential partners were identified by mass spectrometry analysis. Among these, Bud23, a SAM-dependent MTase involved in ribosome biogenesis and catalyzing the methylation of N^7 atom of G_{1575} in 18S rRNA, was an attractive candidate (Figure 17; (White et al., 2008; Figaro et al., 2012)). The interaction between Bud23 and Trm112 was shown to be direct by co-purification of both proteins following co-expression in *E. coli* (Figaro et al., 2012; Sardana, & Johnson, 2012). Furthermore, Trm112 is important for Bud23 cellular stability and for its activity (Figaro et al., 2012). In *S. cerevisiae*, the *bud23*Δ strain exhibits a strong growth defect, sensitivity to paromomycin as well as defects in the

synthesis and the nuclear export of the small ribosomal subunit 40S (White et al., 2008; Figaro et al., 2012). This *bud23*Δ mutant is affected in A2 cleavage, resulting in an accumulation of 35S and 20S rRNA intermediates, a depletion of 27SA2 rRNA intermediate and consequently a reduction of mature 18S rRNA (White et al., 2008; Figaro et al., 2012). Finally, Bud23 associates with the 90S particle at the intermediate stage before A2 cleavage (Sardana et al., 2013).

Crystal structures of *Sc*Bud23-Trm112 complex (lacking the Bud23 C-terminal extension rich in basic residues) in the presence or absence of SAM have brought useful information regarding the interaction mode between both proteins (see section 2.5) but also on G_{1575} binding (Figure 20; (Letoquart et al., 2014)). Indeed, comparison of the crystal structures of *Sc*Bud23 and *Coffea canephora* xanthine MTase bound to xanthosine has revealed striking similarities between enzyme active sites, suggesting that they both bind the purine ring of their substrates in a very similar manner (McCarthy, & McCarthy, 2007; Letoquart et al., 2014). This binding mode was validated experimentally by the characterization of Bud23 active site mutants. Based on these observations, the *Sc*Bud23-Trm112-GMP (guanosine monophosphate) model was generated (Figure 20B) and superimposed onto nucleotide G_{1575} in the structure of the mature *S. cerevisiae* 80S ribosome. Such superimposition reveals large steric clashes between Bud23-Trm112 and ribosomal components indicating that this Bud23-Trm112 complex cannot bind mature ribosomes and has to dissociate from the 40S subunit before completion of its biosynthesis. Additional experiments demonstrated that although Bud23 is recruited to pre-ribosomes at an early nucleolar stage, G_{1575} methylation, which is not essential for Bud23 cellular function, is a late event as it occurs on the 20S pre-rRNA (Letoquart et al., 2014). Finally, the Bud23-Trm112 complex physically interacts with the Dhr1 DEAH-helicase, which is involved in the dissociation of U3 small nucleolar RNA from the pre-40S prior to formation of the central pseudo-knot of the 40S subunit (Letoquart et al., 2014; Sardana et al., 2014; Sardana et al., 2015).

The Bud23-Trm112 complex is also found in human cells where TRMT112 interacts with the human Bud23 ortholog RNMT2 (also known as WBSCR22/Merm1; (Ounap et al., 2015; Zorbas et al., 2015)). Similarly to *Sc*Bud23, RNMT2 protein but not its MTase activity is required for ribosome biogenesis (Ounap et al., 2013; Zorbas et al., 2015). RNMT2 is associated

with several human diseases as it is one of the several genes deleted in the Williams-Beuren neurodevelopmental syndrome (Doll, & Grzeschik, 2001; Merla et al., 2002). It was also reported as a tumoral marker for invasive breast cancer, myeloma cells and hepatocarcinoma (Nakazawa et al., 2011; Tiedemann et al., 2012; Stefanska et al., 2014) and it might be involved in lung pathologies (Jangani et al., 2014) .

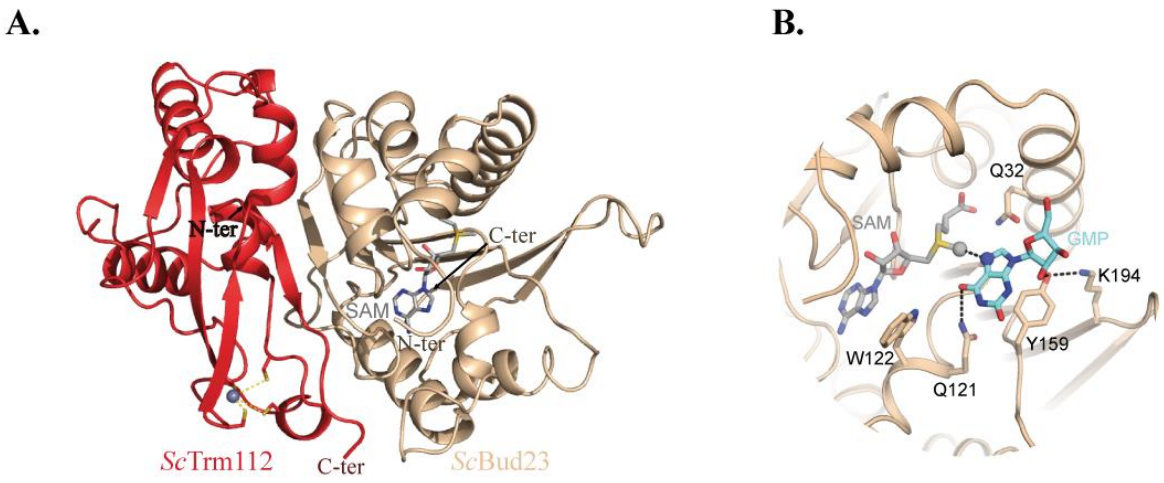


Figure 20. **Crystal structure of *Sc*Bud23-Trm112 complex.** (**A**) Ribbon representation of *Sc*Bud23-Trm112 complex bound to SAM. (**B**) Model of GMP (blue sticks) bound to *Sc*Bud23 active site.

2.4.2. Trm112 also influences 60S formation

The importance of Trm112 in ribosome biogenesis extends beyond the role of Bud23-Trm112 complex in 40S maturation as the disruption of *TRM112* gene in yeast also causes lower levels of 60S subunit (Figaro et al., 2012; Sardana, & Johnson, 2012). The effect of Trm112 depletion is less pronounced on 60S than on 40S levels but Trm112 is definitely important for the synthesis of both subunits. Trm112 role in 60S synthesis is supported by its co-immunoprecipitation with pre-60S and its association with Nop2 and Rcm1, two 25S rRNA-MTases involved in 60S biogenesis (Sardana, & Johnson, 2012; Sharma et al., 2013; Bourgeois et al., 2015). However, experimental evidences supporting a direct interaction between Trm112 and these two MTases have not been presented so far. Hence, Trm112 role in 60S synthesis is still unclear and future studies addressing this issue are needed.

2.5. Common themes in recognition and activation of these MTases partners by Trm112

Despite its small size, eukaryotic Trm112 is part of at least four heterodimeric MTase holoenzymes and acts as an activator of the MTase catalytic subunits. The description of the molecular mechanisms underlying Trm112 activation role was hindered by the difficulty to express and purify most isolated MTase subunits (Trm9, Mtq2 and Bud23) in sufficient amount for biochemical and biophysical studies (Heurgue-Hamard et al., 2006; Mazauric et al., 2010; Figaro et al., 2012; Sardana, & Johnson, 2012). Indeed, only the co-expression of Trm112 together with each of these three MTases allows the purification of the corresponding Trm112-MTase complexes (Heurgue-Hamard et al., 2006; Figaro et al., 2008; Mazauric et al., 2010; Chen et al., 2011; Liger et al., 2011; Figaro et al., 2012; Sardana, & Johnson, 2012; Letoquart et al., 2015a). This first led to the determination of the crystal structures of three Trm112-MTase complexes, namely *Ecu*Mtq2-Trm112 (Liger et al., 2011), *Sc*Bud23-Trm112 (Letoquart et al., 2014) and *Yl*Trm9-Trm112 (Letoquart et al., 2015a). These structures revealed that Trm112 interacts in a very similar way with these three MTase partners (rmsd values lower than 3Å when superimposing the structures of the complexes). To date, no crystal structure of a Trm11-Trm112 complex has been determined. However, information on this complex could be gleaned from the possibility to purify *S. cerevisiae* Trm11 alone in milligram amounts (Bourgeois et al., 2017). Hydrogen-deuterium exchange experiments coupled to mass spectrometry revealed that Trm11 regions involved in Trm112 binding match with the regions from Trm9, Mtq2 and Bud23 MTase domains that interact with Trm112. Reciprocally, Trm112 region contacted by Trm11 perfectly overlaps with the region involved in the interaction with the other MTases. Hence, these four MTases use the same surface of their MTase domain to bind to the same region of Trm112, and then compete directly to interact with Trm112. This is in agreement with previous reports showing that Trm11 over-expression in yeast reduces the amount of Trm112 co-immunoprecipitated with Trm9 and that Mtq2 over-expression in yeast results in decreased levels of Bud23 (Studte et al., 2008; Figaro et al., 2012).

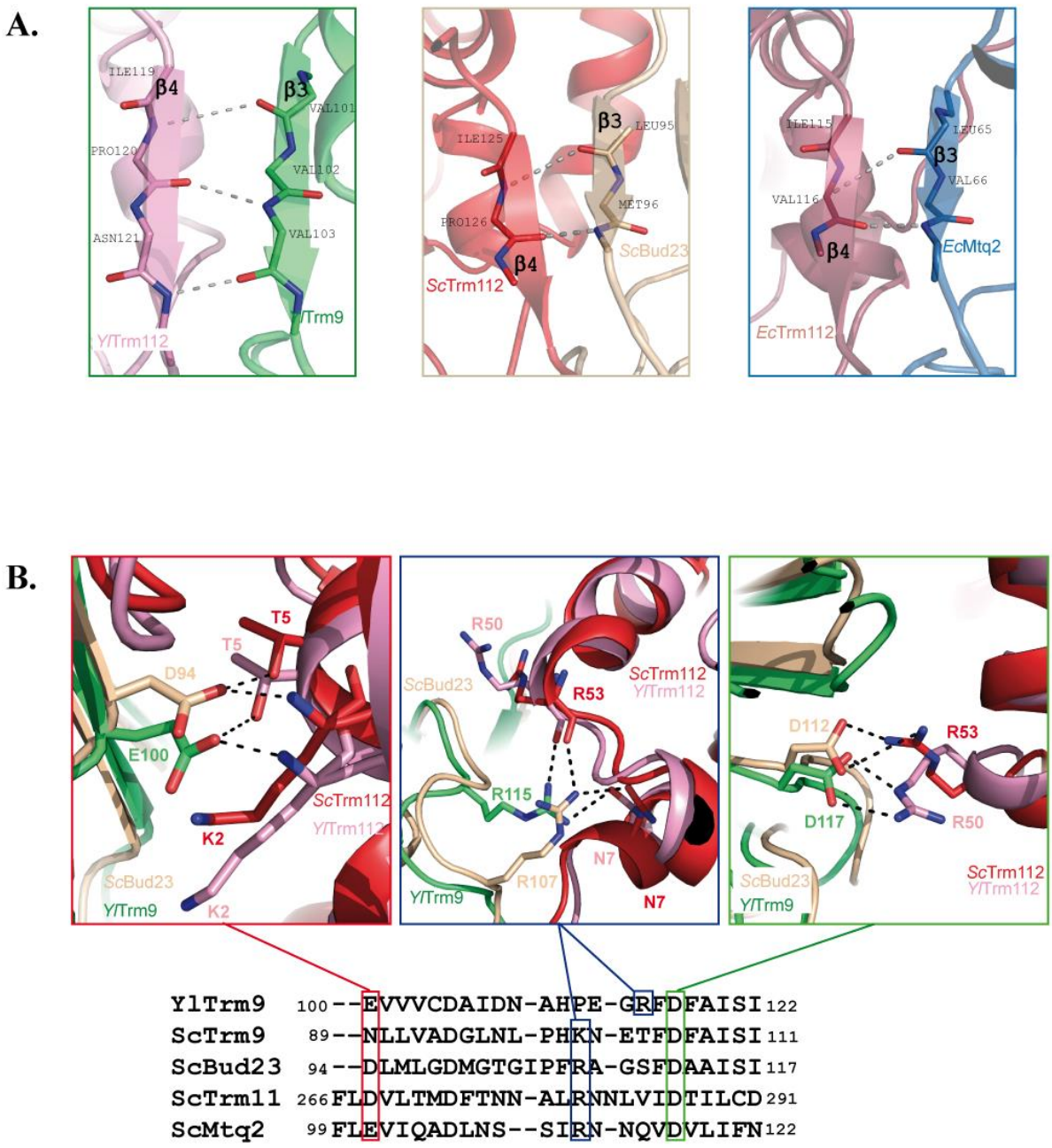


Figure 21. **Comparison of the Trm112-MTase interfaces.** (**A**) Ribbon representation of the β-zipper interaction between Trm112 and MTases. Hydrogen bonds formed between main chain atoms from both partners are depicted by grey dashed lines. (**B**) Comparison of *Sc*Bud23-Trm112 and *Yl*Trm9-Trm112 structures reveals conserved hotspots involved in complex formation. A structure-based sequence alignment of *Yl*Trm9 and of the four *S. cerevisiae* MTases interacting with Trm112 is shown in the lower panel. Only a small region of these MTases is shown for the sake of clarity.

These MTases interact mainly with the Zn-knuckle domain from Trm112 and complex formation buries a large hydrophobic region on the surface of both partners. This explains the requirement of Trm112 to express and purify three of these MTases in their soluble forms in *E. coli* and also to stabilize at least Bud23 in *S. cerevisiae* yeast cells (Heurgue-Hamard et al., 2006; Mazauric et al., 2010; Chen et al., 2011; Figaro et al., 2012; Sardana, & Johnson, 2012). Central to the Trm112-MTase interfaces is a β-zipper interaction formed between strand β3 from the MTase domain and strand β4 from Trm112 forming a continuous large eleven stranded β-sheet (Figure 21A). Such interaction relies on hydrogen bonds formed between main chain atoms from both partners and hence is much more dependent on the local three-dimensional structure than on conservation of amino acid residues at the positions involved in the formation of this β-zipper. Finally, three electrostatic interactions are also observed in both fungal Trm9-Trm112 and Bud23-Trm112 complexes. Structure-based sequence alignment between these four MTases from the same organism shows that the residues present at each of these three positions are strictly conserved or have the same propensity to form similar electrostatic interactions, *i.e.* basic residues (Lys or Arg) or similar hydrogen bonding properties (Asp, Glu or Asn; Figure 21B). This observation indicates that these interactions are most likely common to all Trm112-MTase complexes. Altogether, such binding mode features explain that Trm112 can interact with several MTases sharing less than 20% sequence identity. The ability of plant and metazoan Trm112 or MTases orthologs to complement at least partially for the deletion of the corresponding yeast gene further supports the plasticity in the interaction mode between these proteins (Figaro et al., 2008; Begley et al., 2013; Ounap et al., 2013; Letoquart et al., 2015a). Indeed, this indicates that chimeric Trm112-MTase complexes can be formed between a yeast protein and an ortholog from its binding partner.

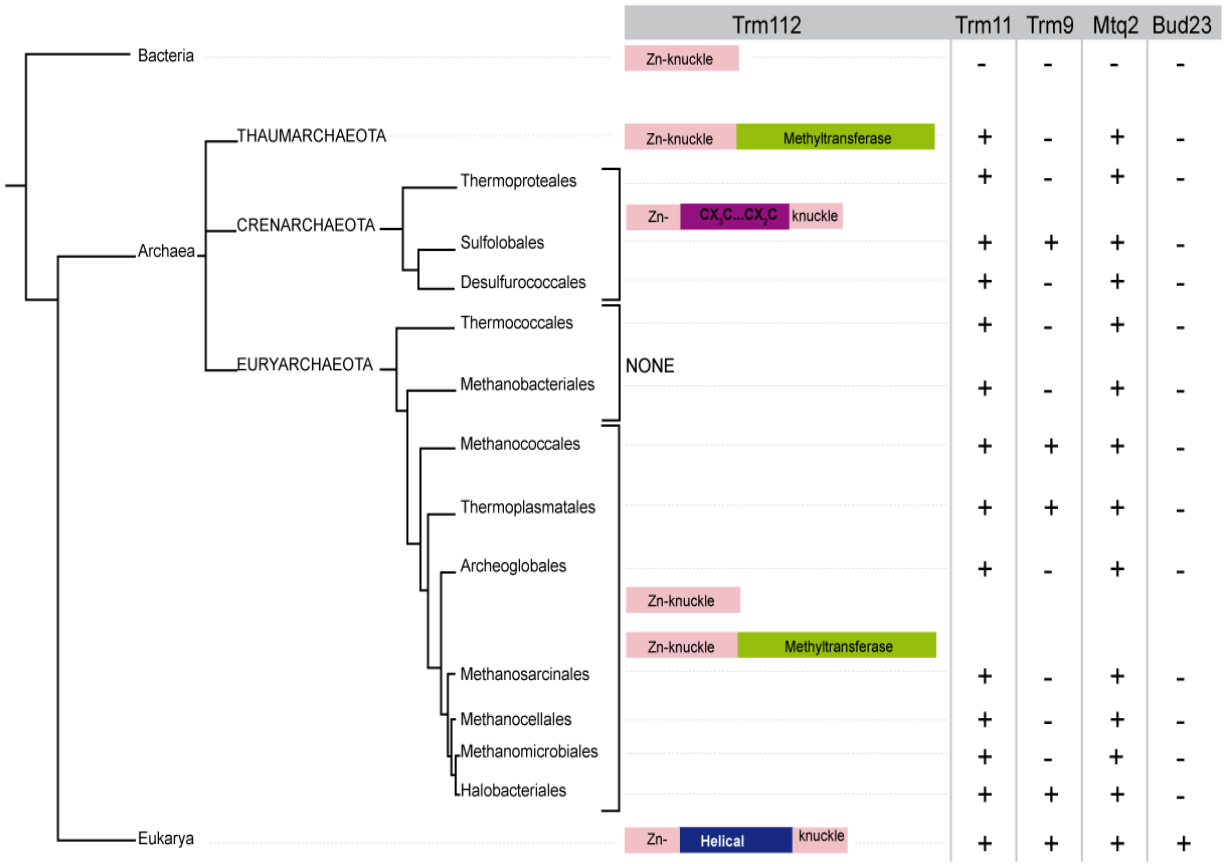


Figure 22. **Trm112 is present in the three domains of life.** Simplified phylogenetic tree of Trm112. Emphasis is given to archaeal phylogeny. The distribution of Trm11, Trm9, Mtq2 and Bud23 proteins within the three domains of life is indicated. The various Trm112 forms identified are schematically depicted.

III. Archaeal Trm112-related research

Since the initial analyses conducted on archaeal Trm112 sequences (Purushothaman et al., 2005; Heurgue-Hamard et al., 2006), many more archaeal genomes have been sequenced and the archaeal phylogeny has been revised (Brochier-Armanet et al., 2011). A new sequence analysis showed that Trm112 archaeal orthologs cluster into three subfamilies (Figure 22). The first one composed of proteins ranging in size from 60 to 80 residues harbors a [C/D]PX[C/D]$X_{19-36}CX_2C$ signature (where X is any residue). It is predicted to contain only the Zn-knuckle domain similarly to bacterial orthologs and is found almost exclusively in euryarchaeota. The second corresponds to proteins of about 130-140 amino acids found exclusively in crenarchaeota phylum. These proteins contain the Zn-knuckle domain as well as a central region. This later displays some similarity with eukaryotic Trm112 helical domain and a conserved putative $CX_3CX_{15-20}CX_2C$ Zn-binding signature. This observation is compatible with the eocyte phylogenetic tree proposed by Cox and coworkers (Cox et al., 2008), proposing that crenarchaeota are most closely related to eukaryotes. The third subfamily (composed of at least 29 members) is formed by proteins consisting of a N-terminal Trm112-like Zn-knuckle domain fused to a C-terminal SAM-dependent MTase domain. These are found only in some euryarchaeota or thaumarchaeota but not in crenarchaeota. Interestingly, bioinformatic analyses of the sequence of the MTase domains fused to Trm112 reveal that these MTases do not all belong to the same family and that some of these are putative archaeal orthologs of Trm9. Finally, no protein with significant sequence similarity with Trm112 could be identified in thermococcales and methanobacteriales from euryarchaeota phylum. This is noteworthy as Trm11 orthologs from two thermococcales archaea (*Pyrococcus abyssi* and *Thermococcus kodakarensis*) can form m^2G (and even $m^2{}_2G$) at position 10 of some tRNAs *in vitro* without requirement for a protein partner (Armengaud et al., 2004; Hirata et al., 2016). Bioinformatics analysis supports the existence of Trm11 orthologs in all archaea phyla and hence most of the time, these proteins co-occur with Trm112 suggesting that with the exception of thermococcales and methanobacteriales Trm11, the other archaeal Trm11 orthologs may exist as a complex with Trm112 and may require Trm112 to be active as observed in eukaryotes. Similarly, Mtq2 orthologs are present in all archaea phyla. The presence of Mtq2 orthologs in archaea is fully consistent with the conservation of the GGQ signature on eukaryotic (eRF1) and archaeal (aRF1)

class I translation termination factors and their strong structural similarity (H. Song et al., 2000; Saito et al., 2010; Kobayashi et al., 2012).

Modifications of archaeal tRNAs were investigated primarily in *Haloferax volcanii* and U at position 34 of some tRNAs were shown to harbor an unknown modification (Gupta, 1984). A more recent study identified 5-carbamoylmethyluridine (ncm^5U_{34}) at position 34 from $tRNA^{Leu}$(UAG) isolated from the *Thermoplasma acidophilum* (thermoplasmatales; euryarchaeota; (Tomikawa et al., 2013)). Several observations led Grosjean and coworkers to propose that $mcm^5(s^2)U$ could also be present in *Haloferax volcanii* (halobacteriales; euryarchaeota) and *Sulfolobus solfataricus* (sulfolobales; crenarchaeota) (Grosjean et al., 2008). Indeed, Elp3 (HVO_2888), Tuc1 (HVO_0580) and Trm9 (HVO_0574) orthologs were initially predicted in both organisms and genes encoding putative Elp3 and Trm9 orthologs are clustered in *Sulfolobus solfataricus.* More recently, HVO_1032 was proposed as a new *H. volcanii* Trm9 ortholog due to a better blast score (Phillips, & de Crécy-Lagard, 2011). In addition, recombinant Elp3 from *Methanocaldococcus infernus* archeon (methanococcales, euryarchaeota) was shown to catalyze cm^5U formation on tRNAs *in vitro* (Selvadurai et al., 2014). Blast analysis of *H. volcanii* HVO_1032 protein sequence against archaeal proteins identified Trm9 putative orthologs with E-values lower than $1e^{-48}$ only in sulfolobales (crenarchaeota) and in halobacteriales (euryarchaeota), suggesting a limited distribution of this modification. Finally, no protein with significant sequence homology with Bud23 could be identified by bioinformatics searches. This is in agreement with the fact that to our knowledge, the nucleotide that structurally matches with *S. cerevisiae* 18S rRNA G_{1575} has not been shown to be modified in archaeal 16S rRNAs analyzed so far.

Many interesting common features exist between archaeal and eukaryotic Trm112 proteins, opening a large field of investigation for future research aimed at understanding the functions of Trm112 in archaea. Those studies could also contribute to improving our understanding of the role of Trm112 in eukaryotes based on these similarities. Therefore, much more efforts on archaeal Trm112 research need to be made in order to expand our knowledge on this interesting topic.

OBJECTIVES OF THE PROJECT

The fine-tuning and tight control of protein synthesis process is ensured by PTMs of the translational components, with methylation being by far the most frequent. The significance of methylation on eukaryotic protein translation is perfectly illustrated by Trm112 protein, which is an activating platform, essential for the function and stability of four class-I SAM-dependent MTases including Bud23, Trm9, Trm11 and Mtq2. Until now, studies on those Trm112-MTase complexes have been extensively performed both functionally and structurally in eukaryotes (mostly *S. cerevisiae* and human). However, although Trm112 orthologs have been found in Archaea, little is known about its MTase partners as well as the roles of the methylation they catalyze in archaeal protein synthesis.

In addition, despite much effort made on eukaryotic Trm112-MTase complexes, it is still very difficult to study structure of those complexes bound to their substrates in order to get insight into the substrate recognition and modification mechanism. This struggle is due to the fact that protein complexes in eukaryotes are quite transient and eukaryotic proteins are normally larger and longer, containing flexible regions compared to bacterial and archaeal counterparts. These make it tough to study X-ray structure for eukaryotic protein complexes. Meanwhile, archaeal proteins are known to be more compact than their eukaryotic orthologs. Moreover, those archaeal proteins are also stable at extreme conditions like high salt concentration, high temperature as well as pH, leading to stronger protein complex formation in principle. In addition, the archaeal enzyme-substrate complex can be maintained much longer than the eukaryotic counterpart during crystallization process since the archaeal proteins are found less active at 20^0C and 4^0C, two commonly-used temperatures for crystallization. These characteristics can result in increasing chances for solving structure of archaeal protein complexes by X-ray crystallography. Therefore, it is worth finding out the Trm112 interacting network in Archaea.

For this project, the following specific aims have been conducted:

- **Characterization of the active site of *Sc*Trm9-Trm112**

Before I joined the lab, a PhD student had solved the crystal structure of Trm9-Trm112 from the *Yarrowia lypolytica* yeast and mapped the active site based on the structure, and zymocin killer assay. For this goal, some zymocin resistant mutants were selected to examine the effects of those mutations on the complex activity. Then kinetic values (Km, kcat) of some mutants have been determined to evaluate whether those residues were involved in substrate binding and catalysis.

- **Identification and characterization of Trm112-like network in Archaea**

Archaea were recognized as the third domain of life in 1977 by Carl Woese and co-worker (Woese, & Fox, 1977). Those organisms bear characteristics of both bacteria and eukaryotes, namely bacteria-similar morphology but eukaryotic-like genetic information processing pathways (Yutin et al., 2008). This provides an advantage of studying eukaryotic proteins and protein complexes based on their archaeal homologs by making use of their extremophilic properties. However, it is not surprising that research in Archaea still lags far behind that of in Bacteria and Eukarya. One of the reasons for this is the late availability of effective model organisms in Archaea like that in Bacteria (for example, *E. coli*) and Eukarya (*S. cerevisiae*). Fortunately, extensive efforts on archaeal genetic and biochemical research have so far produced some competent model organisms such as *Haloferax volcanii* (Allers et al., 2010).

Haloferax volcanii offers benefits as an archaeal model organism: (1) it is easy to grow under aerobic condition, optimum temperature of 45^{o}C and high salt content (NaCl) of 1.8-3.5M (Allers, 2010); (2) its whole genome has been sequenced (Hartman et al., 2010); (3) genetic tools have been developed, including effective transformation and gene knock-out systems based on selectable markers (Allers et al., 2010); (4) protein expression and purification system are developed (Allers, 2010; Allers et al., 2010); (5) there are ease and reduced cost during working with *H. volcanii* by means for example of lysing cells by the addition of water (Allers, 2010).

The Trm112-like interacting partners in Archaea have been identified using *H. volcanii* as the model organism. To do so, archaeal genetics (pop in/pop out, transformation) have first been applied to generate *H. volcanii* deleted for TRM112 gene for further complementation with a vector harboring aTrm112 fused to a FLAG tag to fish Trm112-like partners through co-immunoprecipitation (Co-IP) based on immuno-affinity between FLAG tag and Anti-FLAG

resin. Those partners were then detected by LC-MS/MS. After applying stringent filtering criteria, some putative MTase partners were finally confirmed by co-expression and co-purification in *E. coli* followed by other confirmation steps by MS and biophysical methods.

For Trm112-MTase complexes of predictable functions, their activity was determined through filter-binding enzymatic assays. Meanwhile, all complexes were subjected to crystallization trials resulting in the structure determination by X-ray crystallography for one complex.

- **Putative MTase partners of eukaryotic Trm112 based on HvoTrm112 network**

Newly identified MTase partners of *Hvo*Trm112 were analyzed with the hope to find out their orthologs in Eukaryotes, which could be eukaryotic Trm112 interacting partners.

MATERIALS AND METHODS

1. MATERIALS

1.1 Strains

Table **S1**. **Summary of different strains with their genotypes and purposes used in the thesis**

Strains	Genotype	Usage	Sources
A. E.coli			
XL1 - blue	*recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F´ proAB lacIq ZΔM15 Tn10 (Tet^R)]*	Cloning and plasmid preparation	Stratagene
BL21 (DE) Gold	*B F^- ompT hsdS(rB^- mB^-) dcm^+ Tet^R gal λ(DE3) endA Hte*	Protein expression	Agilent Technologies
BL21 (DE) Codon plus RIL	*B F^- ompT hsdS(rB^- mB^-) dcm^+ Tet^R gal λ(DE3) endA Hte [argU ileY leuW Cam^R]*	Protein expression	Agilent Technologies
B. S. cerevisiae			
YDL201	*MATa ura3-52 lys2-801_amber ade2-101_ochre trp1-Δ63 his1-Δ200 leu2-Δ1 TRM9::kanMX6*	tRNA purification	Létoquart et al; *NAR*; 2015
C. H. volcanii			
H98	*pyrE2*Δ *hdrB*Δ	Pop-in/pop out	Kind gifts from R. Lestini (LOB, Ecole Polytechnique)
H26	*pyrE2*Δ	Pop-in/pop-out, tRNA purification	
H26 *trm112*Δ	*pyrE2*Δ *trm112*Δ	tRNA purification	Kind gifts from J. Letoquart (Letoquart *et al.*, 2015)
H26 *trm9*Δ	*pyrE2*Δ *trm9*Δ		
H133 *elp3*Δ	*pyrE2*Δ *trpA*Δ *leuB*Δ *hdrB*Δ *elp3*Δ	tRNA purification	Kind gifts from Pr Mevarech/Altman (Altman-Price, & Mevarech, 2009)
H98 Trm112-FLAG	*pyrE2*Δ *hdrB*Δ *trm112*Δ	Co-IP	This study
H26 aRF1-FLAG	*pyrE2*Δ *aRF1-FLAG*	Co-IP	This study
H98 aRF1-FLAG *trm112*Δ	*pyrE2*Δ *hdrB*Δ *trm112*Δ *aRF1-FLAG*	Co-IP	This study
H26 aRF1-FLAG *mtq2*Δ	Δ*pyrE2 mtq2*Δ *aRF1-FLAG*	Co-IP	This study

1.2 Oligos and generated plasmids

Table **S2**. **Oligonucleotides and plasmids used for *in vivo* experiments in *H. volcanii***

***Hvo* genes**	**Names**	**Sequence**	**Enzyme**	**Plasmid generated**
Hvo_1131-US	oMG86	GGG**GGATCC**CGTCGTGACGGTG GATTGCG	**BamHI**	pMG613
	oMG87	CCC**TCTAGA**CGTGCTCGCCGGCC TCGACG	**XbaI**	
Hvo_1131-DS	oMG88	GGG**CTCGAG**CACGTCGACCGGC GGTTTGC	**XhoI**	
	oMG89	CCC**GGATCC**CACCTCATCGGAC CTGAACG	**BamHI**	
Hvo_Trm112 (Hvo_1131)	oMG320	GGGAA**CATATG**AAAGAATCCCT GATGGACATCCTCTGTGACCCC	**NdeI**	pMG772
	oMG321	TTTTT**GCGGCCGCtta**CTTGTCGT CATCGTCTTTGTAGTCGTCGTCTC GCATGTCCGGCGGTAGG	**NotI**	
Hvo_1131-US	oMG163	CATCGCCGTCAACCACCTCG	-	-
Hvo_1131-DS	oMG164	TCGGCTTCGACAGCGTCTCG	-	-
Hvo_1032-US	oMG72	GGG**GGATCC**GGACAAAGACCGC ACGGGTC	**BamHI**	pMG612
	oMG73	CCC**TCTAGA**GCGGAACCAGGAG ATGCTC	**XbaI**	
Hvo_1032-DS	oMG74	GGG**CTCGAG**GCGGAACCGGCCG TCGTCC	**XhoI**	

	oMG75	CCC**GGATCC**GCGTTGGTCGGGACTGGGGC	**BamHI**	
C-terminal Hvo_RF1-FLAG	oMG347	CCC**GGATCC**TTACTTGTCGTCATCGTCTTTGTAGTCG	**BamHI**	pMG798
	oMG348	CCCC**TCTAGA**CCGACCAAGAGGTCATCAAGG	**XbaI**	
Hvo_RF1-DS	oMG329	CCC**GAATTC**ATCCTATCGGCGCGTAGAGCG	**EcoRI**	
	oMG330	CCC**GGATCC**GCGACCTCCTCCTCGGCGTCCCCGGATCCGCGACCTCCTCCTCGGCGTC	**BamHI**	
Hvo_Mtq2-US	oMG90	GGGGGATCCTCGTCGCCTCCCTCGACGCC	**BamHI**	pMG614
	oMG91	CCCTCTAGAAGCCTCGTCAACCGGAGCCG	**XbaI**	
Hvo_Mtq2-DS	oMG92	GGGCTCGAGCCGGAGAAGGTGAGTTCGCC	**XhoI**	
	oMG93	CCCGGATCCCTAATGCGCATATGTTGATAC	**BamHI**	
Hvo_RF1-US	oMG332	GGGTCTAGATGACCGGTTCGAGCTCC	-	-

Restriction sites are highlighted in bold.

Stop codon is shown in bold lower cases and the DNA sequence encoding for the His_6-tag is underlined.

Table **S3**. **Oligonucleotides and plasmids used to over-express proteins in *E. coli***

***Hvo* genes (*Hvo* protein)**	**Name**	**Sequence**	**Enzyme**	**Plasmid generated**
Hvo_0773 (*Hvo*0773)	oMG401	CCCTTGCCATGGCTAAAGGAAAGGAGTGGTACCAGGCCGACG	**NcoI**	pMG814 (pET28b; Kan^R)
	oMG402	TTTTT**AAGCTTtta**ATGGTGATGGTGATGGTGCTCGGCGTCGCTCGCCCGCG	**HindIII**	
Hvo_0574 (*Hvo*0574)	oMG403	CCCTTGCCATGGCTCGACGTTTCTCCGAATCGTACCTCCG	**NcoI**	pMG813 (pET28b; Kan^R)
	oMG404	TTTTT**AAGCTTtta**ATGGTGATGGTGATGGTGCACTCGCCCCGTCGTCACGTC	**HindIII**	
Hvo_0019 (*Hvo*0019)	oMG405	CCCTTGCCATGGCTAGCGTCCGCGACGAGTTCGACGCCTG	**NcoI**	pMG812 (pET28b; Kan^R)
	oMG406	TTTTT**AAGCTTtta**ATGGTGATGGTGATGGTGGGGCGCGACGCCGACGGTCAG	**HindIII**	
Hvo_1715 (*Hvo*1715)	oMG407	CCCTTGCCATGGCTAGCGACGAGAAACGCCGAACCGCCG	**NcoI**	pMG811 (pET28b; Kan^R)
	oMG408	TTTTT**AAGCTTtta**ATGGTGATGGTGATGGTGCTCCGTTCGCGCTCGAACCAGCGC	**HindIII**	
Hvo_0475 (*Hvo*0475)	oMG409	CCCTTGCCATGGCTCCTACCCGCGACCGGTCGCAGTC	**NcoI**	pMG826 (pET28b; Kan^R)
	oMG410	TTTTT**AAGCTTtta**ATGGTGATGGTGATGGTGATCAACCGCACGCTCGCCAACGACCG	**HindIII**	
Hvo_2875 (*Hvo*2875)	oMG447	CCCTTGCCATGGCTCACGGCGCTGGCGACG	**NcoI**	pMG829 (pET28b; Kan^R)
	oMG448	TTTTT**AAGCTTtta**ATGGTGATGGTGATGGTGATGGCTCTCCCGCTTTTTTCCG	**HindIII**	

Hvo_1131 (*Hvo*Trm112)	oMG473	GGG**CATATG**AAAGAATCCCTGATGGA	**NdeI**	pMG564 (pET21a; Amp^R)
	oMG474	TTT**GCGGCCGC**tcaGTCGTCTCGCATGTCCGGCG	**NotI**	
Hvo_1032 (*Hvo*Trm9)	oMG469	GGG**CCATGG**ACGGAGACGAACCCCG	**NcoI**	pMG560 (pET28b; Kan^R)
	oMG470	TTT**GCGGCCGC**tcaATGGTGATGGTGATGGTGGCAGCCGGCGACGACGGCGT	**NotI**	
Hvo_2744 (*Hvo*Mtq2)	oMG327	CCCTTGCC**ATGGCT**ACCGACCTCGCCGAGCGCCGC	**NcoI**	pMG769 (pET28b; Kan^R)
	oMG328	TTTTT**GCGGCCGCtta**ATGGTGATGGTGATGGTGTTCCAACGCGAGCACCGTGAGCGTCTCG	**NotI**	
Hvo_0156 (*Hvo*Trm11)	oMG471	GGG**CCATGG**CATACGGATTGGAACTCGCGCCGG	**NcoI**	pMG562 (pET28b; Kan^R)
	oMG472	GGG**CTCGAG**tcaATGGTGATGGTGATGGTGCCGATGCAGCAGATGCACGT	**XhoI**	
Hvo_0321 (*Hvo*aRF1)	oMG322	GGGAA**CATATG**AGTAGCGACGCCGAGGAGGCGAGC	**NdeI**	pMG766 (pET21a; Amp^R)
	oMG324	TTTTT**GCGGCCGCtta**ATGGTGATGGTGATGGTGGACGCCGGTGGAGTACCGGAGGATGC	**NotI**	
Hvo_0359 (*Hvo*aRF3 or *Hvo*aEF1A)	oMG325	CCCTTGCC**ATGGCT**AGCGACAAACCCCACCAGAACCTGGCC	**NcoI**	pMG768 (pET28b; Kan^R)
	oMG326	TTTTT**GCGGCCGCtta**ATGGTGATGGTGATGGTGTCGCTCGTTGACTTCGAGCACTTTGCCGGC	**NotI**	

Restriction sites are highlighted in bold.

Stop codon is shown in bold lower cases and the his-tag DNA sequence in underlined.

Table **S4**. **Oligonucleotides used for site-directed mutagenesis and resulting plasmids**

***Hvo* genes (*Hvo* protein)**	**Names**	**Sequence**	**Template**	**Plasmid generated**
*Hvo*aRF1 Q187A	oMG387	CAGCGAAAGGGAGGTGCGTCCGC	pMG766	pMG797 (pET21a; Amp^R)
	oMG388	GAAACGCTGGGCGGACGCACCTC		
*Hvo*Mtq2 Y111A	oMG385	GTTCAACCCGCCGGCCCTCCCCG	pMG769	pMG796 (pET28b; Kan^R)
	oMG386	GAGGTCCTCGGGGAGGGCCGGC		

1.3 Media

1.3.1 Bacteria growth media

Luria-Bertani broth (**LB**): Tryptone (Bacto) 10g/l, yeast extract 5g/l and NaCl 5g/l, pH 7.5 (adjusted by NaOH).

2YT: Tryptone (Bacto) 16g/l, yeast extract 10g/l and NaCl 5g/l, pH 7.0 (adjusted by NaOH).

Terrific broth auto inducible (TBAI): Tryptone (Bacto) 12g/l, yeast extract 24g/l and MgSO4 0.15g/l, $(NH4)_2SO4$ 3.3g/l, KH_2PO_4 6.5g/l, Na_2HPO4 7.1g/l, glucose 0.5g/l, Alpha Lactose 2.0g/l, pH 7.0 (adjusted by NaOH).

1.3.2 Archaeal growth media

YPC: 18% salt water (SW: 2.46M NaCl, 90mM MgCl2, 84mM MgSO4, 54mM KCl and 12mM Tris HCl pH7.5), yeast extract (Difco) 5g/l, peptone (Oxoid) 1g/l, casamino acid 1g/l, 3mM $CaCl_2$.

1.4 Buffers

Table **S5. Different buffers used for different protein purifications in different organisms**

Steps	**Buffers**	**Components**		
		H. volcanii	*A. fulgidus*	*Human*
Cell lysis	Lysis buffer	2M NaCl, 50mM Tris HCl pH7.5, 5mM β-mercaptoethanol, 10µM $ZnCl_2$ and 10mM Imidazole	Similar as Hvo's but 0.5M NaCl	Similar as Hvo's but 0.2M NaCl
Ni-NTA	Washing buffer	Lysis buffer + 20mM Imidazole	Lysis buffer + 20mM Imidazole	Lysis buffer + 20mM Imidazole
	Elution buffer	Lysis buffer + 350mM Imidazole	Lysis buffer + 350mM Imidazole	Lysis buffer + 350mM Imidazole
Ion exchange chromatography *MonoQ or Heparin*	Buffer A	50mM NaCl, 50mM Tris HCl pH7.5, 5mM β-mercaptoethanol, 10µM $ZnCl_2$	Similar as Hvo's	Similar as Hvo's
	Buffer B	1M NaCl, 50mM Tris HCl pH7.5, 5mM β-mercaptoethanol, 10µM $ZnCl_2$	Similar as Hvo's	Similar as Hvo's
Gel filtration *S75 S200*	Buffer C	1M NaCl, 50mM Tris HCl pH7.5, 5mM β-mercaptoethanol, 10µM $ZnCl_2$	Similar as Hvo's but 0.5M NaCl	Similar as Hvo's but 0.2M NaCl

2. METHODS

2.1 Cloning

Several cloning experiments have been performed to generate plasmids containing target DNAs (from *H. volcanii*, *Homo sapiens* and *A. fulgidus*) used for protein expressions in *E. coli* or for archaeal genetics in this thesis. In general, DNA sequence encoding for the protein of interest was amplified using genomic DNA as templates by PCR with forward and reverse primers as listed in Table S2 and S3. All PCR steps were done with Phusion High-Fidelity DNA Polymerase (Thermo) or Q5 High-Fidelity DNA Polymerase (Biolabs) and PCR cycling steps were conducted according to the manufacturers' instructions. The PCR products were run on agarose gels (0.8-1.2%) and target DNA bands were excised from gels, followed by clean-up steps with commercial kits according to the manufacturers' instructions. Purified PCR fragments were digested by corresponding restriction enzymes designed in the oligos (Table S2 and S3) and then ligated using DNA T4 ligase into pET21a or pET28b vectors digested by the same restriction enzymes as the inserts. The plasmids harboring the target DNAs were transformed into chimio-competent *E. coli* XL1 Blue by the heat-shock transformation method. The transformants were then streaked on LB agar plates containing appropriate antibiotics (ampicillin (100µg/mL) for pET21a and kanamycin (50µg/mL) for pET28b), followed by o/n incubation at 37^0C. Some colonies for each clone were selected to inoculate 5 mL of LB medium supplemented with corresponding antibiotics and grown o/n at 37^0C. The plasmids were then extracted by MiniPrep kit (Qiagen) and finally sent for sequencing (Eurofins).

2.2 Site-directed mutagenesis

Plasmids encoding for *Hvo*aRF1 Q187A and *Hvo*Mtq2 Y111A mutants were generated by site-directed mutagenesis using partly overlapping oligos and templates as mentioned in Table S4. The PCR were performed using Phusion High-Fidelity DNA Polymerase (Thermo) (PCR conditions: one step of 180s at 98°C followed by 30 cycles (30s at 98°C; 30s at 60°C and 8min at 72°C) and a final step of 8 minutes at 72°C). The PCR products were subjected to DpnI restriction enzyme to remove the non-mutated plasmid templates. The resulting plasmids were purified by PCR clean up purification kit and then transformed into competent *E. coli* XL1 Blue

for plasmid production through Miniprep (Qiagen). The correctness of mutants was confirmed by sequencing (Eurofins).

2.3 Pop in/pop out

2.3.1 Deletion of *Hvo*Trm112 gene in *H. volcanii*.

For pop-in step, 1-4 colonies of *Haloferax volcanii* H98 strain were inoculated to 10 mL YPC media supplemented with uracil and thymidine (60 µg/mL) and cultured overnight (o/n) at 45°C. At OD_{650nm} around 0.8, the cells were harvested by centrifugation at 6000 rpm for 8 minutes at room temperature (RT). The pellet was then washed by gentle re-suspension in 2 mL buffered spheroblasting solution (1M NaCl, 27mM KCl, 50mM Tris HCl pH 8.5, 15% (w/v) sucrose) and centrifuged again in the same conditions. The pellet was next gently resuspended in 600 µL of the same buffer and 200 µL of the resuspended cells were used for one transformation. To generate spheroblasts, 20 µL of 0.5M EDTA pH 8.0 were added on the side of the tube followed by rapid mixing by inverting the tube and finally incubated 10 minutes at RT. In the next step, 30 µL of a mixture containing 15 µL of unbuffered spheroblasting solution (1M NaCl, 27mM KCl, 15% (w/v) sucrose, final pH 7.5), 5 µL of 0.5M EDTA pH 8.0 and 10 µL of *dam*$^-$ target DNA (pMG613) (about 1-2 µg) was added in the similar way as EDTA. After 5 minutes incubation at RT, 250 µL of a 60% PEG600 solution were added to the mixture and mixed gently before incubation at RT for 30 minutes. Next, 1.5 mL spheroblasting dilution solution (23% salt water (SW), 15% (w/v) sucrose, 3.75mM $CaCl_2$) was added, mixed by inverting the tube and incubated at RT for 2 minutes. The cells were pelleted at 6000 rpm for 8 minutes at RT. The supernatant was discarded and 1 mL of regeneration solution (23% SW, 15% (w/v) sucrose, 3 mM $CaCl_2$, 1xYPC (1.02g Yeast extract (Difco), 0.2 g Peptone (Oxoid), 0.2g Casamino Acids, 1.76 mM KOH in final 204 mL solution)) supplemented with thymidine (60 µg/mL) was added without disturbing the pellet. This mixture was then incubated at 45°C without rotation for 1.5 – 2 hours, followed by 3 – 4 hours incubation with rotation at 150 rpm. Then cells were harvested by centrifuging at 6000 rpm for 8 minutes at RT. The pellet was gently resuspended in 1 mL transformant dilution solution (23% SW, 15% (w/v) sucrose, 3mM $CaCl_2$) or SW solution 18%. Finally the suspension was plated on Casa-plates at different dilutions and then incubated at

45°C for at least 5 days. The grown colonies were then streaked on a new Casa-plate and grown at 45°C for at least 5 days.

For pop-out step, one isolated colony from the pop-in step was used to inoculate 3 mL of YPC (supplemented with thymidine 60 µg/mL), grown o/n at 45°C and 150 rpm. In the next day, 6 µL of o/n culture was inoculated to another 3 mL of YPC (supplemented with thymidine 60 µg/mL) and incubated o/n at 45°C and 150 rpm. This step was repeated once more. Then, the o/n culture was diluted to 10^{-2}-10^{-3} in 18% SW and plated on Casa-5-FOA plates and incubated at 45°C for at least 5 days.

Colonies were screened by colony lift to check for the deletion of the *HvoTRM112* sequence. All colonies from the pop-out step were transferred to YPC plates (+ thymidine 60 µg/mL) in a systematic way and then incubated at 45°C for 5 days. The colony lift started by transferring all colonies on a cellulose membrane followed by cell lysis by SDS 10%. DNA was denatured and neutralized by denaturation solution (1.5M NaCl, 0.5M NaOH) and neutralization solution (1.5M NaCl, 1mM EDTA, 0.5M Tris HCl pH7.5), respectively and then fixed on the membrane by UV cross-linking (using Biolink cross-linker with program for dosage of 0.120 J/cm^2). Next, membrane-bound DNA was hybridized with DIG-labeled *HvoTRM112* DNA oligo at 65°C o/n with agitation. After several washing steps to remove non-hybridized DNA, the DIG detection was done by using DIG Luminescent Detection kit (Roche) and a ChemiDoc MP (BioRad).

An additional validation step was performed by PCR on colony lift-derived colonies using the oMG163 and oMG164 oligos (Table S2) to amplify the whole upstream and downstream parts of *Hvotrm112* (Hvo_1131) gene. Parts of colonies were picked and lysed by mixing in 50 µL water, and then used as DNA templates for PCR reactions. The PCR was performed with Q5 High-Fidelity DNA Polymerase (Biolabs), according to manufacturer's instruction (PCR conditions: one step of 30s at 98°C followed by 30 cycles (10s at 98°C; 20s at 62°C and 30s at 72°C) and a final step of 2 minutes at 72°C).

2.3.2 Generation of *H. volcanii* strains containing aRF1-FLAG

The pop in/pop out was done as described previously for gene knockout with the aim of replacing the WT allele of *aRF1* gene by a FLAG-tagged allele on the chromosome. For that

purpose, we constructed plasmid pMG798 containing the C-terminal part of *aRF1* with the FLAG and the DS sequence of *aRF1*.

Pop-in/pop-out was performed with pMG798 using H26 strain to construct the *aRF1*-FLAG mutant or using *trm112*⊗ strain (in H98 background) to construct the *trm112*⊗ *aRF1*-FLAG mutant. In addition, a pop-in/pop-out was performed using plasmid pMG614 and *HvoaRF1*-FLAG mutant (in H26 background) to construct the *HvoaRF1*-FLAG *mtq2*⊗ mutant.

Colonies from pop-in/pop-out experiments were screened by PCR on colony. For that purpose colonies were picked and lysed by being just mixed in 50µL water, and then used as DNA templates. The PCR was performed with Q5 High-Fidelity DNA Polymerase (Biolabs), according to the manufacturers' instructions. For H26 *aRF1*-FLAG, PCR were done with oligos oMG329 and oMG332 (Table **S1**) designed to amplify the whole upstream and downstream parts of *Hvo* aRF1. PCR products were cut by BamHI to validate the presence of the FLAG tag (30% of colonies were *aRF1*-FLAG (3 out of 10 tested colonies)). In case of mutants, double checks namely presence of both mutants (*mtq2*Δ (by oligos oMG91-92) and *trm112*Δ (using oligos oMG163-164), Table **S1**) and aRF1-FLAG (using oligos oMG329 and oMG332) were performed by PCR. The aRF1 PCR products were subjected to digestion by BamHI (1.6% of colonies from the pop-in/pop-out in *Hvotrm112*⊗ strain were *aRF1*-FLAG (2 out of 120 tested colonies) and 2% of colonies from the pop-in/pop-out in Hvo*aRF1*-FLAG strain were *mtq2*⊗ (2 out of 104 tested colonies)).

2.4 Co-Immunoprecipitation (Co-IP)

To identify the partners of the *Hvo*Trm112 protein, we have performed co-immunoprecipitation under cross-linking conditions as described below.

For *Hvo*Trm112-FLAG expression, one colony of *Hvo* strain H98 *trm112*Δ - pTA962-Trm112-FLAG transformants was used to over-express *Hvo*Trm112-FLAG. First, the colonies were inoculated to 10 mL YPC (supplemented with thymidine 60 µg/mL), incubated o/n at 45°C and 150 rpm. In the next day, 2.5 mL o/n culture was applied to 1L YPC (+thymidine 60 µg/mL) and grown o/n at 45^0C and 150rpm. On the third day, 18% SW-dissolved tryptophan was added to the o/n culture to the final concentration of 5mM, followed by 6 hours incubating at 45°C and 150rpm for the protein over-expression. The cells were harvested by centrifuging at 4000rpm for

30 mins. For the control, the FLAG only, *Hvo* strain H98 *trm112*Δ – pTA927-FLAG transformants were used in a similar manner.

The co-immunoprecipitation (Co-IP) for *Hvo*Trm112-FLAG and the FLAG-only were carried out as described in Fischer *et al*., 2010, with some slight modifications. The pellets were washed with enriched PBS (2.5M NaCl, 150mM $MgCl_2$, 1×PBS (137mM NaCl, 2.7mM KCl, 10mM Na_2PO_4, 2mM $KHPO_4$, pH 7.4)). The cells were again pelleted and then re-suspended in enriched PBS containing 1% formaldehyde followed by incubation for 20 mins at 45 °C. To stop the cross-linking reaction, glycine was added to a final concentration of 0.25M and incubated for 5 mins at 45 °C. The cells were washed twice with enriched PBS at 4°C, and then lysis buffer (50mM Tris HCl pH 7.4, 1mM EDTA, 10mM $MgCl_2$, and 1mM $CaCl_2$) containing proteinase inhibitor was added. The solution was re-suspended and incubated for 1 hour at 4°C. Then RNase A was added to a final concentration of 200μg/mL, and the mixture was incubated for 30 min at 37°C followed by centrifugation at 20000cpm for 45 mins. Next, the supernatant was obtained and NaCl was added to a final concentration of 150mM, ready for the Co-IP.

The preparation of anti-FLAG M2 affinity gel (Sigma) was conducted as the manufacturer's protocol. For each Co-IP experiment, 1.6 mL of anti-FLAG M2 affinity gel was washed 10 times with 10 mL of ice-cold washing buffer (50mM Tris HCl, pH 7.4, 150mM NaCl) before the lysate was applied. After incubation overnight (14–16 hours) at 4°C, incubated anti-FLAG M2 affinity gel was washed eight times with 10 mL of cold washing buffer. The elution of the *Hvo*Trm112-FLAG was carried out by using 4 mL of washing buffer containing 3×FLAG peptide to the final concentration of 150ng/μL. The samples were incubated at 4°C for 1 hour with gentle rotation. For the second elution, the affinity gel was rinsed with 2 mL of washing buffer. For *Hvo*Trm112-FLAG and FLAG-only, the Co-IPs were performed four times and in triplicate, respectively.

In case of *Hvo*aRF1-FLAG in different *H. volcanii* strains, the Co-IP experiments were done in a similar manner as that for that *Hvo*Trm112-FLAG but with some modifications. First, *Hvo*aRF1-FLAG was expressed in native condition since aRF1-FLAG was integrated into *H. volcanii* genome. There was no tryptophan added for the protein induction. Second, no cross-linking was performed before Co-IP step for *Hvo*aRF1-FLAG purifications.

2.5 Mass spectrometry

The identification of *Hvo*Trm112 partners were identified by LC-MS/MS on Co-IP samples by our collaborators, Leslie Muller, PhD student in the group of Dr. Sarah Cianferani in Strasbourg University. The protocol is as the following:

After denaturation at 100°C for 10 mins in loading buffer (2% SDS, 0,1M DTT, 10% glycerol, 62.5 mM Tris pH 6.8), 20 µg of proteins of each sample were concentrated in one band with a 4% stacking SDS-PAGE. The gels were fixed with 45% methanol/3% acetic acid and stained with colloidal Silver Blue. Each band was excised and cut in four pieces prior to *in-gel* digestion. The gel pieces were washed four times with 100 µL of 75% acetonitrile (ACN) and 25% NH_4HCO_3 at 25 mM and dehydrated with 50 µL of ACN. Cysteine residues were reduced by adding 10 mM DTT for 30 min at 60°C and 30 min at room temperature, and alkylated by adding 55 mM iodo-acetamide for 20 min in the dark. The bands were then washed three times by adding 50 µL of 25 mM NH_4HCO_3 and 50 µL of ACN. After two dehydration steps with 50 µL of ACN, gel pieces were stored at -20 °C prior to enzymatic digestion. Proteins were cleaved in an adequate volume to cover all the gel pieces with a modified porcine trypsin (Promega) solution at a 1:100 w/w enzyme:protein ratio. Digestion was performed overnight at 37°C. Tryptic peptides were extracted twice under agitation, first with 40 µL of 60% ACN in 0.1% formic acid (FA) for 1 h and then with 30 µL of 100% ACN for 30 min. The collected extracts were pooled, the excess ACN was vacuum dried, and the samples were resolubilized with 10 µL of H_2O/ACN/FA (98/2/0.1 v/v/v).

The nanoLC-MS/MS analysis was performed on an Eksigent NanoLC 400 system coupled to a TripleTOF 6600 mass spectrometer (Sciex, Framingham, USA). Peptide separation was performed on a ZORBAX 300SB-C18 column (150 mm × 300 µm with 3.5 µm diameter particles – Agilent Technologies). The solvent system consisted of 0.1% FA in water (solvent A) and 0.1% FA in ACN (solvent B). The samples (4.6 µL) were loaded onto the column and the peptides were eluted at 5 µL/min with the following gradient of solvent B: from 2 to 35% over 95 min, and 35 to 80% over 1 min.

The Ion Spray Voltage Floating was set to 5.5 kV and the interface heater at 100 °C. The system was operated in Data Dependent Acquisition mode with automatic switching between MS and MS/MS modes. MS1 spectra were collected at 400-1250 m/z for 250 ms. The most

intense ions with charge states 2-4 were selected on each MS spectrum for further isolation and collision induced dissociation fragmentation. MS2 spectra were acquired in high sensitivity mode at 200-1600 m/z using dynamic accumulation, with an accumulation time for high intensity peaks of 25 ms and a total cycle time of 2.8 s. After fragmentation, the precursor ions were excluded for 18 s.

Raw data were converted into calibrated peak lists .mgf using ProteinPilot™ software (v. 5.0) before being subjected to a search against a concatenated target-decoy database including both forward (target) and reversed (decoy) UniProtKB *Haloferax volcanii* sequences (4 February 2016, 9858 total entries) and common contaminants using Mascot search algorithm (v.2.5.1). Searches were performed with a mass measurement tolerance of 15 ppm for precursor ions and 0.05 Da for fragment ions. Oxidation of methionine residues, carbamidomethylation as well as propionamidation of cysteine residues were searched as variable modifications. Full trypsic peptides only were searched with a maximum of one missed cleavage allowed. Proline software (http://proline.profiproteomics.fr; (Carapito et al., 2015)) was used to validate the identification results. For each sample, Peptide Spectrum Matches were filtered out if they did not meet following criteria: pretty rank ≤ 1, Mascot ion score ≥ 25, minimum peptide length of 7 residues and a maximum false discovery rate of 1 % on the adjusted e-value. Then, proteins were filtered out in order to obtain a final list with a maximum false discovery rate of 1 % based on the modified MudPit score. GO annotations were extracted for the identified proteins using an in-house developed software suite (Mass Spectrometry Data Analysis, https://msda.unistra.fr; (Carapito et al., 2014)). Only proteins identified with at least one unique peptide were considered.

2.6 Protein expression and purification

2.6.1 Expression test

Tests to analyse the expression of single proteins or co-expression of two proteins were performed in 5 mL media (TBAI or 2YT) using different *E.coli* strains (Gold or Codon Plus). Protein induction can be examined at different temperatures (18, 30, and 37^0C) and/or different IPTG concentrations (0.01-1mM in case of 2YT). The o/n pre-culture was carried out by inoculating the transformants into 5-10 mL LB containing appropriate antibiotics. In the next

day, 50 µL o/n pre-culture was introduced to 5 mL tested media containing appropriate antibiotics, first grown at 37^0C for 3 hours, then switching to different temperatures and/or adding IPTG at different tested concentrations for protein induction. The induction time was dependent on the temperature used to induce protein expression, *i.e.* 4-5 hours at 37^0C and o/n at 30^0C and lower temperatures. Cells were then harvested and lysed in 1 mL lysis buffers (Table S5) by sonication, followed by lysate clearance through centrifugation at 13000g at 4^0C for 30 min. The supernatant was then subjected to Ni-NTA in small-scale purification followed by washing and then eluting steps with corresponding buffers (Table S5). The protein expression was checked in SDS-PAGE 12%.

2.6.2 Large-scale purification

Large-scale (co-)expression and (co-)purification of proteins from different species were carried out in 1 L of appropriate media using the conditions determined from the expression tests. Cells were cultured and harvested as for the expression tests but in a large scale manner. Pellets were re-suspended in 30 mL lysis buffer (Table S5) and lysed by sonication on ice, followed by lysate clearance by centrifuging at 20000rpm for 45 mins. The supernatant was applied on Ni-NTA resin equilibrated with lysis buffer, incubated at 4^0C with gently rotation for 30 mins. After the flow-through, the protein-bound Ni-NTA resin was washed with 30 mL washing buffer (Table S5). Then 10 mL and 5 mL elution buffer (Table S5) were respectively used to elute the His-tagged proteins. The protein solutions were concentrated using a concentrator (10kDa cutoff) to 1 mL which was then 5 times diluted with buffer A (Table S5), ready for ion exchange chromatography. The next ion exchange and size-exclusion chromatography were done using Biologic DuoFlow (BioRad) or AKTA (GE Healthcare) systems. For ion exchange, Mono Q or Heparin 5 mL columns (GE Healthcare) were used and equilibrated with buffer A before 5 mL protein samples were injected and run at a flowrate of 2.5 mL/min. The proteins were eluted by NaCl concentration gradient from 50mM (buffer A) to 1M (buffer B, Table S5). The protein-containing fractions were analyzed by SDS-PAGE, collected and concentrated to 5 mL for the next chromatography step. The size-exclusion chromatography was performed by S75 16/60 column equilibrated with buffer C (Table S5). 5 mL protein samples were injected and the protein fractions were checked by SDS-PAGE, pooled and concentrated to 100-300 µL.

2.7 SEC-MALLS

To determine molecular weights of Trm112-MTase complexes in solution and then their oligomeric states, experiments of size exclusion chromatography coupled to multi-angle laser light scattering (SEC-MALLS) were performed. The advantage of this method over the size exclusion only is that by SEC-MALLS, we can directly estimate the molecular weight of protein complexes of any size and shape through scattered light at different angles without requirement for a calibration curve, which is known to be valid only for globular proteins.

For SEC-MALLS, each complex sample of 100 µL (1 mg/mL) was injected at a flow rate of 0.75 mL/min on a Superdex^{TM} 200 Increase 10/300 GL column (GE-Healthcare) in buffer C (Table S5). Elution was followed by a UV-Visible spectrophotometer, a RID-20A refractive index detector (Shimadzu), a MiniDawn TREOS detector (Wyatt Technology). The data were collected and processed with the program ASTRA 6.1 (Wyatt Technology). M_w was directly calculated from the absolute light scattering measurements using a dn/dc value of 0.183.

2.8 X-ray crystallography

2.8.1 Principle of X-ray crystallography

X-ray crystallography is a key X-ray based technique for determining the three-dimensional structure of molecules at the atomic level, which in turn provides crucial clues for understanding their biological functions. The principle behind this method is that when a crystal is subjected to an X-ray source, it will diffract X-rays into different directions, which can be recorded as a diffraction pattern containing many spots called reflections. Each reflection can be characterized by three parameters: the position (h,k,l), the intensity (I_{hkl}) and the phase σ_{hkl}. Knowing all parameters from each spots, we can obtain information about the position of each atom in the crystal. However, through diffraction pattern analysis, crystallographer can only determine the position and the intensity of the spots while the phase information is still missing. Hence, to solve the crystal structure the phase information has to be obtained based on a variety of methods.

2.8.2 Protein crystallization

2.8.2.1 Why do we need crystals for X-ray crystallography?

Diffraction signal of a single molecule will be too weak for measurement while that of different single molecules in solution can cause too many noises for detection. To avoid those problems, a crystal of molecules is needed. In fact, the crystal is a well-ordered array of molecules, which can contain billions of molecules, therefore acting as diffraction signal amplifiers.

2.8.2.2 Principle of protein crystallization

The first step of protein crystallization is to purify the protein of interest in sufficient amounts for crystallization. Keeping in mind that the more purified and homogenous protein sample, the more chance the protein crystal will grow. There are a number of commonly-used methods available for protein crystallization including liquid-liquid diffusion crystallization (batch, dialysis) and vapor diffusion crystallization (sitting drop, hanging drop). The crystal growth can be dependent on various conditions such as protein concentration, crystallization solutions and their concentration, pH, temperature, and additives.

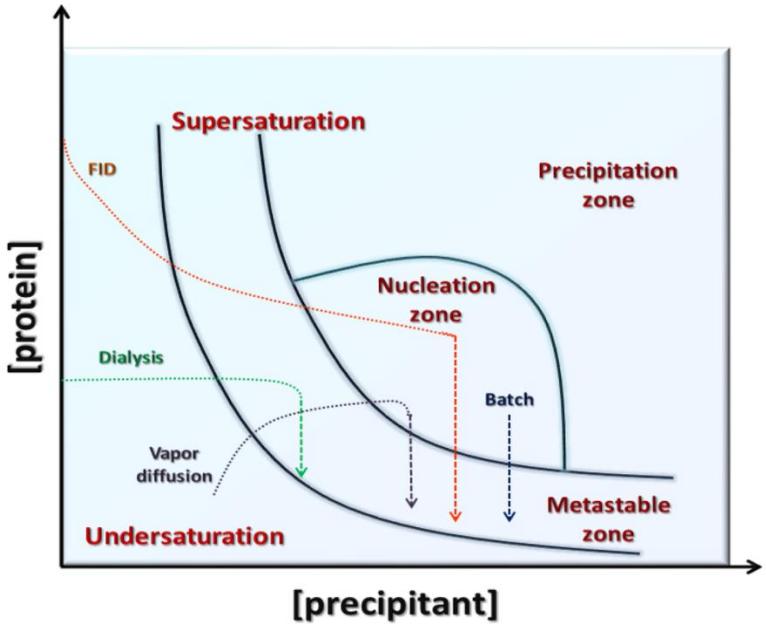


Figure S1. **Protein crystallization phase diagram from different crystallization methods** (FID: free interface diffusion, dialysis, batch and vapor diffusion)

In theory, the proteins can form crystals when the protein-dissolved solution is brought to a supersaturated state in a thermodynamic and kinetic favor, which is much dependent on both protein concentration and solution conditions. At the beginning of crystallization, the protein is dissolved in solution or in undersaturation state where no crystal can be formed. The protein concentration is then increased over its solubility curve by different crystallization methods as seen in the phase diagram during the crystallization process, reaching the supersaturated state (Figure S1). At this state, the system is not at equilibrium and thermodynamically driven to a new equilibrium situation with a new minimized free energy. Particular interactions can happen between individual molecules, leading to formation of aggregates, known as nucleation. Under suitable conditions, these aggregates can reach a critical size, forming stable nuclei. The next step is known as the crystal growth resulting from lowering protein concentration through nuclei formation in the nucleation to another phase shown in diagram as metastable zone. In this range, nucleation does not spontaneously occur and the stable nuclei will play a role as surface suitable for crystal growth, putatively leading to diffracting crystals in case suitable conditions are observed.

2.8.2.3 Crystallization techniques

Among different crystallization approaches available, in the thesis I have only applied the vapor diffusion method including the sitting drop and the hanging drop crystallization. The principle behind this method is that the protein sample is mixed with solution buffer normally in the ratio of 1:1 while the large solution buffer alone is kept in the reservoir (Figure S2). The system is then carefully sealed to make an enclosed environment. As a result of mixing, the concentration of precipitants in the sample drop is lower than that in the reservoir, resulting in a higher water concentration in the sample drop. In the equilibrium system, the water vapor will leave the sample drop and end up in the reservoir, leading to an increase in the concentration of both protein and precipitant in the sample drop to reach a level where the crystallization process can occur provided that optimal conditions are met.

For sitting drop crystallization, a Mosquito robot was used to screen hundreds of different commercial solution conditions in which 150 nL protein sample was mixed with 150 nL reservoir solution while 80 µL solution buffers was used in the reservoir.

Hanging drop set-up was used to optimize crystals obtained in the sitting drop by changing salt and precipitant concentration through applying hand-made gradients. To do it, 1 µL of protein sample was mixed with 1 µL of reservoir solution and 500-1000 µL was applied for the reservoir solution.

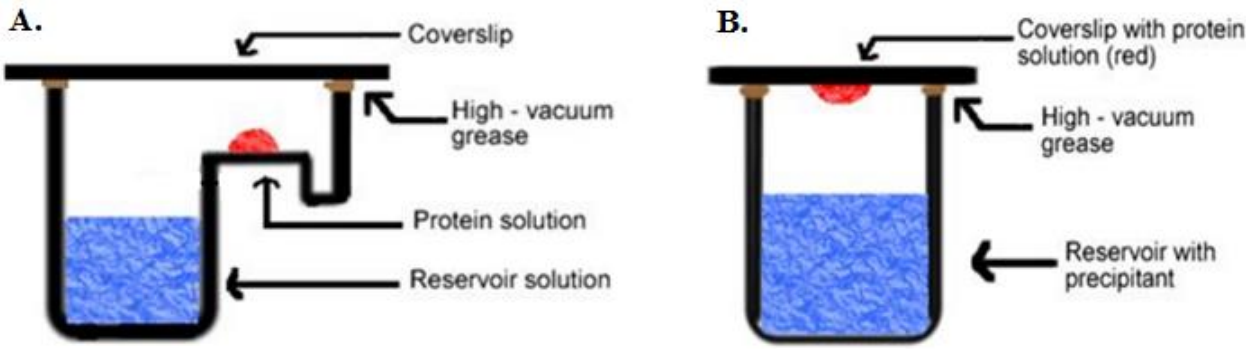


Figure S2. **Vapor diffusion crystallization**. **A.** Sitting drop method; **B.** Hanging drop method

2.9 *Hvo* tRNA purification

tRNAs from different *H. volcanii* strains (wild-type, *trm112*Δ, *trm9*Δ and *elp3*Δ) were purified based on phenol-chloroform extraction method. For each purification, 3 L of YPC (plus thymidine (60 µg/mL) in case of *elp3*Δ *Hvo*) were used to grow *H. volcanii*. For *Hvo* culture, 100mL YPC pre-culture was o/n performed at 45^0C. In the next day, 20mL pre-culture was used to inoculate into 1 L YPC. The culture was grown at 45^0C 150rpm for 28-30 hours or until at least mid-log phase. The cells were harvested by centrifugation at 4000rpm and 4^0C for 30 min. The cell pellet was re-suspended into sodium acetate 50mM pH 5.0 solution (3 mL for each mg of pellet). Then phenol saturated with Sodium acetate 50mM pH 5.0 was added to the suspension with the ratio of 1:1. The solution was well mixed, followed by gently shaking incubation at room temperature for overnight. In the next day, the mix was centrifuged at 5000rpm at RT for 20 min. The supernatant (top layer) was recovered followed by an addition of phenol in 1:1 ratio. The mix was strongly agitated in 2 min and then centrifuged at 5000rpm at RT for 20 min. The top layer supernatant was again collected and chloroform was added in 1:1 ratio, followed by 2-min vigorous agitation. The solution was centrifuged at 5000rpm at RT for 20 min. The

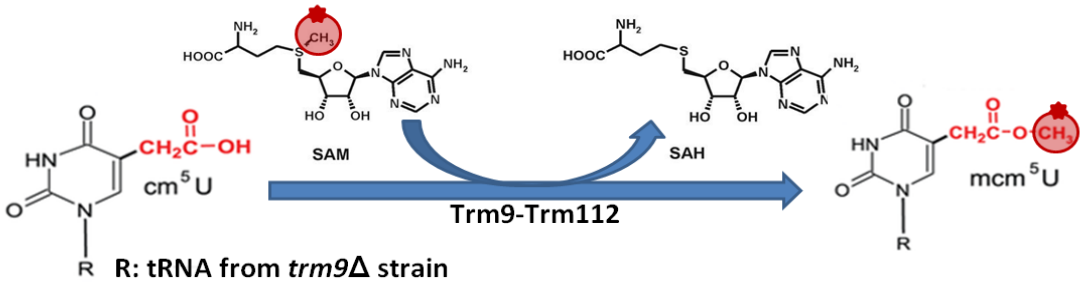
supernatant was recovered and 0.2 volume of isopropanol was added and well mixed, followed by 1 hour incubation at -20^{0}C. This step is to pellet genomic DNA and long RNAs. The solution was next centrifuged at 8000rpm at 4°C for 20 min. The supernatant was collected and 0.6 volume of isopropanol was added, mixed well and then incubated at -20^{0}C for at least 1 hour. The mix was centrifuged at 8000rpm at 4°C for 20 min. The supernatant was discarded and the pellet was washed by 80% ethanol. The mix was centrifuged at 8000rpm and 4°C for 20 min and the pellet was collected and dried. The pellet was then re-suspended in 0.1-1 mL sterile water or TE buffer. The total RNA was then run on polyacrylamide-urea gel (6-10%) in which bands corresponding to tRNAs (around 75 nucleotides) were cut into small pieces, covered with elution buffer (10mM Tris pH 7.5, 1mM EDTA, 0.1% SDS and 300mM NaCl) and then incubated at 37^{0}C with agitation for 30 min. The supernatant was recovered and this step was done in triplicate. Next, 2.5 volumes of pure ethanol were added to the supernatant and well mixed before incubation at -20^{0}C for overnight. The solution was pelleted at 12000rpm 4^{0}C for 30 min and the pellet was dried and finally re-suspended in 8mM $MgCl_2$.

10. Enzymatic assay

2.10.1 *Sc*Trm9-Trm112

The methyltransferase activity was determined by filter-binding assay based on the principle that the *Sc*Trm9-Trm112 enzyme complex transfers a tritium-labeled methyl group from [^{3}H]SAM to the tRNA substrates purified from a *S. cerevisiae trm9Δ* strain, i.e. containing cm^5U_{34} to generate tritium-labeled mcm^5U_{34}. The mcm^5U_{34}-harboring tRNAs were retained on glass filters (Whatman GF/C filters), which do not retain free radioactive SAM, and the radioactivity was detected using a scintillation counter. Details about these assays were described in the enclosed paper in chapter I and can be summarized as the following figure S3.

• Principle



• Procedure

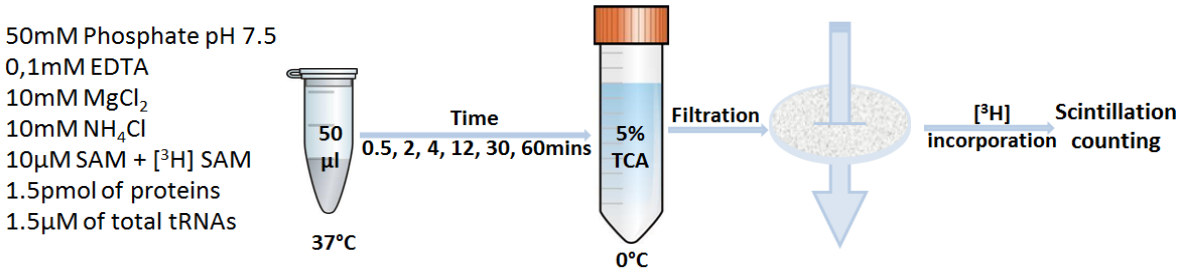


Figure S3. **Summary of filter-binding enzymatic assay method**. Indicated experimental conditions are those used for enzymatic assays on *Sc*Trm9-Trm112 (Adapted from Dr Juliette Létoquart)

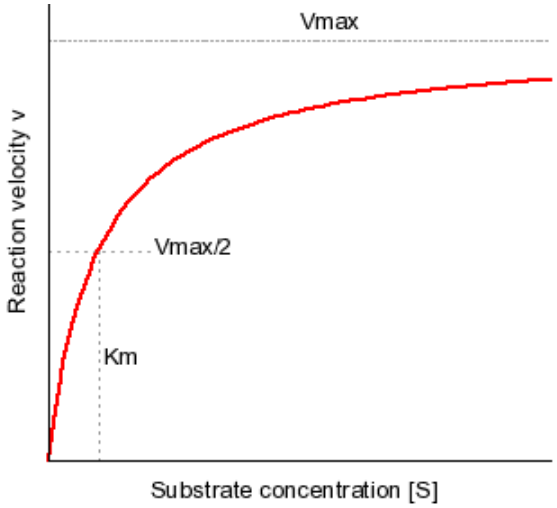
For calculation, the initial velocity (V_i) was derived from the formula:

$C_{tRNA} = V_i*(1 - \exp(-nt))/n$ where C_{tRNA} is the concentration of methylated tRNA, t the time in minutes, and n the relaxation rate constant of V_i by using the ORIGIN software according to Cao et al (Cao, & De La Cruz, 2013). The enzyme specific activity was then calculated by dividing V_i by enzyme quantity.

To determine the Km and Vmax values, the reaction velocity was plotted as a function of substrate concentration and the data were fitted with the ORIGIN software using the Michaelis-Menten equation as the following:

$$V = Vmax*[S]/(Km + [S])$$

where V is the reaction velocity, Vmax the maximal reaction velocity, [S] the substrate concentration and Km the Michaelis-Menten constant.



Then, k_{cat} was calculated by dividing Vmax by enzyme concentration.

2.10.2 *Hvo*Mtq2-Trm112

The methyltransferase assay was performed in a total volume of 10 µL containing 400mM phosphate buffer pH 7.5, 3M KCl, 2.5mM EDTA, 5mM $MgCl_2$, 5mM NH_4Cl, 0.25mg/mL Bovine Serum Albumin (BSA), 50µM SAM (containing 0.87 Ci/mmol of [3H]-SAM, Perkin Elmer) and 5 pmol of *Hvo*Mtq2-Trm112 complexes. The reaction was initiated by adding 100 pmol of *Hvo*aRF1 and *Hvo*aRF3 each to the mixture. The samples were incubated at 45°C for 2 hours. The reaction was stopped by precipitation with cold trichloroacetic acid (5%), followed by filtration on Whatman GF/C filters. The [^{3}H] incorporation was measured using a Beckman Coulter LS6500 scintillation counter. For buffer optimization, different phosphate concentrations (100-400mM) in combination of different concentrations of $MgCl_2$ (5 and 10mM) and NH_4Cl (5 and 10mM) were examined.

2.10.3 *Hvo*Trm9-Trm112

The methyltransferase assay was performed in a total volume of 10 µL containing 400mM phosphate buffer pH 7.5, 3M KCl, 2.5mM EDTA, 5mM $MgCl_2$, 5mM NH_4Cl, 0.25mg/mL Bovine Serum Albumin (BSA), 50µM SAM (containing 0.87 Ci/mmol of [^{3}H]-SAM, Perkin Elmer) and 2 pmol of *Hvo*Trm9-Trm112 complexes. The reaction was initiated by adding 100 pmol of total tRNAs purified from different strains to the enzyme mixture. The

samples were incubated at 45°C for 2 hours. The reaction was stopped by precipitation with cold trichloroacetic acid (5%), followed by filtration on Whatman GF/C filters. The [^{3}H] incorporation was measured using a Beckman Coulter LS6500 scintillation counter.

RESULTS

CHAPTER I

CHARACTERIZATION OF *Sc*TRM9-TRM112 ACTIVE SITE

When I joined the lab, the crystal structure of a truncated version of the Trm9-Trm112 complex from *Yarrowia lypolytica* (*Yl*Trm9N38) had been solved at 2.5Å resolution. Through structure comparison with *Sc*Bud23-Trm112 bound to SAM and *Rhodopseudomonas palustris* RPA2492 (highly structurally similar to Trm9) as well as by secondary structure prediction for the lacking N-terminal region of Trm9, the model of *Yl*Trm9-Trm112 in complex with SAM was generated. The sequence conservation was mapped at the molecular surface of this complex and several highly conserved residues surrounding the SAM methyl group have been identified. Corresponding residues from *Sc*Trm9 (including H24, R29, K31, H115, H116, W145, Q149, W168, R241, Y243, D270 and N271) were selected for functional analysis and were mutated to alanine. The first round of screening for the effect of these mutants was performed *in vivo* using the zymocin killer assay. This assay is based on the principle that the zymocin secreted from *K. lactis* cleaves *Sc*tRNAs containing mcm^5s^2 modification on U34, which is in part catalyzed by *Sc*Trm9-Trm112, finally leading to cell death. From this analysis, eight zymocin-resistant mutants (R29A, D72A, H116A, W145A, W168A, R241A, Y243A and N271A) and two zymocin-sensitive mutants (H115A and Q149A) as well as the WT protein as control were chosen for further functional investigation through enzyme kinetics.

The results that I have obtained on the enzymatic characterization of *Sc*Trm9-Trm112 complex will be presented in details below. Next, I will present the paper in which these results have been included and that has been published in Nucleic Acids Research.

RESULTS

1. Optimum pH for *Sc*Trm9-Trm112 enzymatic activity

To identify optimal pH for the enzymatic activity of *Sc*Trm9-Trm112, phosphate buffers with pH ranging from 5 to 8 have been tested. The enzyme specific activity at each pH has been calculated by using the initial velocity generated from enzymatic reaction of 1.5 pmol of enzyme *Sc*Trm9-Trm112 wild-type on 75 pmol of substrate (total tRNAs purified from *Sc trm9*Δ strain) at different time points. At phosphate pH 5.0, the enzyme was totally inactive (result not shown) whereas the enzymatic activity was somehow affected at other different pHs and followed a bell-shaped curve, with optimal initial velocity at pH 7.5 (Figure 23). Therefore, all the enzymatic assays aimed at characterizing the role of some specific residues on catalytic activity were performed in phosphate buffer and at pH 7.5.

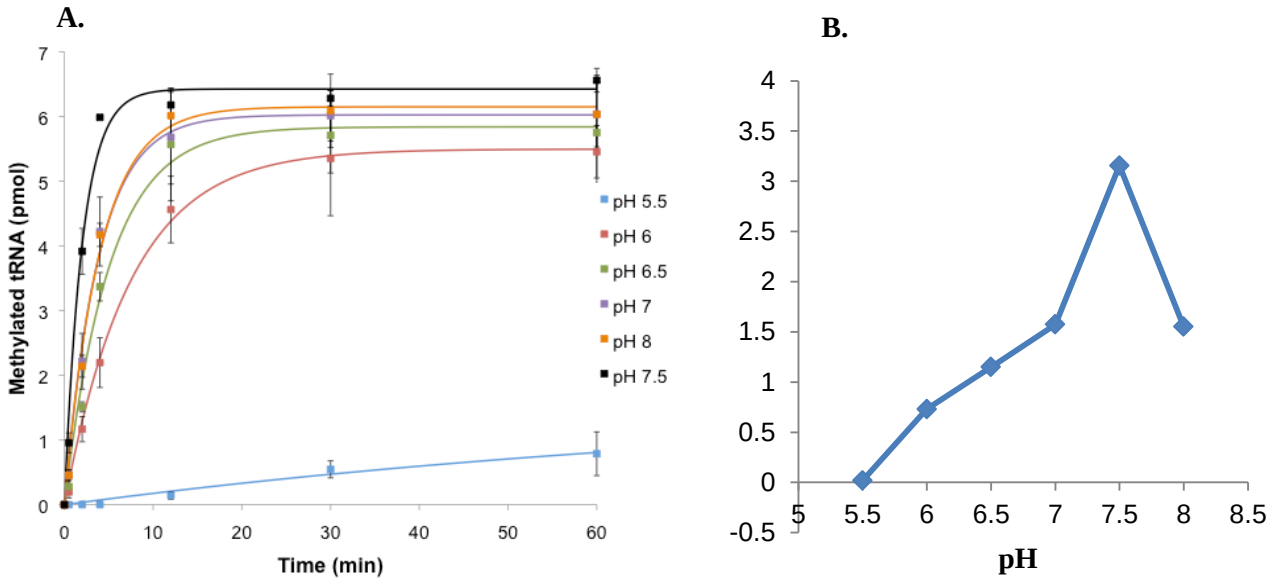


Figure 23. **pH dependence enzymatic activity of *Sc*Trm9-Trm112 complex.** (**A**). Effect of different pHs on enzyme activity. (**B**) Plot of initial velocity (Vi) as a function of pH.

Table 1. **Analysis of pH-dependence enzymatic activity of *Sc*Trm9-Trm112**

pH	**Initial velocity** (Vi)	**Standard error**	**Specific activity** (fmol of tRNA methylated /min/pmol of enzyme)
5.5	0.018	0.004	12
6	0.730	0.035	486
6.5	1.149	0.109	766
7	1.574	0.117	1049
7.5	3.157	0.331	2105
8	1.551	0.121	1034

2. Effect of *Sc*Trm9-Trm112 mutants on the methyltransferase activity

To evaluate whether analyzed residues were preliminarily involved in catalysis or tRNA substrate binding property of *Sc*Trm9-Trm112, the effect of the mutations on the methylation activity was determined by measuring apparent specific activity of the enzyme as for pH-dependence methylation activity. The results showed that zymocin resistant mutants have expectedly an impaired methyltransferase activity in which the D72A, H116A, W145A and W168A mutants were inactive whereas the R29A, R241A, Y243A and N271A mutants displayed severe effects on specific activity compared to wild-type enzyme (Figure 24). The zymocin-sensitive mutants (H115A and Q149A), as expected, were only slightly affected (H115A) or not at all (Q149A) in the enzymatic activity. This indicates a good correlation between zymocin phenotype and enzymatic activity.

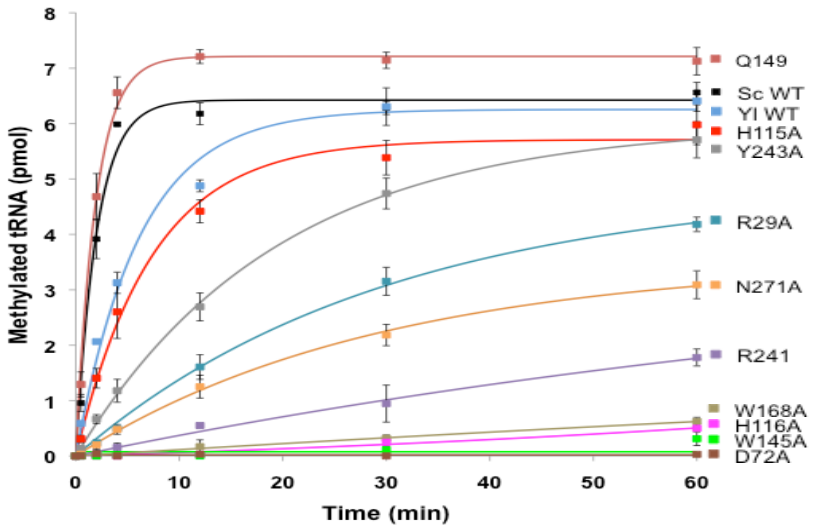


Figure 24. **Enzymatic activity of different *Sc*Trm9-Trm112 mutants**

Table 2. **Analysis of enzymatic activity of *Sc*Trm9-Trm112 mutants**

Proteins	**Initial velocity** (Vi)	**Standard error**	**Specific activity** (fmol of tRNA methylated/min/pmol of enzyme)
WT	3.157	0.331	2104.66 ± 220.66
R29A	0.162	0.007	108.00 ± 4.66
D72A	0	0	0
H115A	0.775	0.058	516.66 ± 38.66
H116A	0.005	0.001	<10
W145A	0	0	<10
Q149A	3.747	0.223	2498.00 ± 148.67
W168A	0.011	0.001	<10
R241A	0.039	0.004	26.00 ± 2.67
Y243A	0.309	0.007	206.00 ± 4.67
N271	0.120	0.005	80.00 ± 3.33

3. Kinetics analysis

To get insights into the enzyme catalytic mechanism, we have carried out steady-state kinetics analyses to determine apparent Km and k_{cat} for the WT, R29A, H115A, R241A, Y243A and N271A Trm9 mutants in the context of the *Sc*Trm9-Trm112 complex. The results revealed that all mutants are affected in Km and k_{cat} values. In fact, all tested mutants showed increase in Km and decrease in k_{cat} compared to the wild-type enzyme, therefore resulting in low catalytic efficiency k_{cat}/K_m. This indicates that these residues might directly participate in the binding of the tRNA substrate as well as in the catalysis. It is also important to mention that these conserved residues may play important role in orienting properly the substrate into the active site required for the methyl transfer to occur by means of S_N2 reaction, a common mechanism of class I SAM-dependent MTases. Therefore, mutants in those positions result in decrease in enzyme affinity for its substrate, leading to loose binding and coordination of the cm^5U substrate into the active site pocket, thereby decreasing the k_{cat}.

Table 3. **Analysis of kinetics of different *Sc*Trm112-Trm9 mutants**

***Sc*Trm9 mutants**	**Apparent specific activity** (fmol of tRNA methylated/min/pmol of enzyme)	**Apparent Km for tRNA** (μM)	**Apparent Kcat for tRNA** ($10^{-3}s^{-1}$)	**Apparent Kcat/Km for tRNA** ($M^{-1}. s^{-1}$)
Sc WT	2105 ± 221	0.08 ± 0.02	32 ± 2	$400x10^3$
R29A	108 ± 5	0.936 ± 0.149	1.9 ± 0.15	$2.03x10^3$
H115A	517 ± 39	0.351 ± 0.118	6 ± 0.7	$17.1x10^3$
R241A	26 ± 3	0.374 ± 0.008	0.8 ± 0.006	$2.14x10^3$
Y243A	206 ± 5	0.851 ± 0.222	4.5 ± 0.5	$5.29x10^3$
N271A	80 ± 3	0.337 ± 0.09	1 ± 0.078	$2.97x10^3$

Published online 4 October 2015 *Nucleic Acids Research, 2015, Vol. 43, No. 22* ***10989–11002***
doi: 10.1093/nar/gkv1009

Insights into molecular plasticity in protein complexes from Trm9-Trm112 tRNA modifying enzyme crystal structure

Juliette Létoquart[1,2], Nhan van Tran[1], Vonny Caroline[1], Alexey Aleksandrov[1], Noureddine Lazar[2], Herman van Tilbeurgh[2], Dominique Liger[2,*] and Marc Graille[1,2,*]

[1]Laboratoire de Biochimie, CNRS, UMR 7654, Ecole Polytechnique, F-91128 Palaiseau Cedex, France and [2]Fonction et Architecture des Assemblages Macromoléculaires, Département B3S, Institut de Biologie Intégrative de la Cellule (I2BC), CNRS, UMR 9198, CEA, Université Paris Sud, F-91405 Orsay Cedex, France

Received March 02, 2015; Revised September 17, 2015; Accepted September 24, 2015

ABSTRACT

Most of the factors involved in translation (tRNA, rRNA and proteins) are subject to post-transcriptional and post-translational modifications, which participate in the fine-tuning and tight control of ribosome and protein synthesis processes. In eukaryotes, Trm112 acts as an obligate activating platform for at least four methyltransferases (MTase) involved in the modification of 18S rRNA (Bud23), tRNA (Trm9 and Trm11) and translation termination factor eRF1 (Mtq2). Trm112 is then at a nexus between ribosome synthesis and function. Here, we present a structure-function analysis of the Trm9-Trm112 complex, which is involved in the 5-methoxycarbonylmethyluridine (mcm^{5}U) modification of the tRNA anticodon wobble position and hence promotes translational fidelity. We also compare the known crystal structures of various Trm112-MTase complexes, highlighting the structural plasticity allowing Trm112 to interact through a very similar mode with its MTase partners, although those share less than 20% sequence identity.

INTRODUCTION

The tRNAs play a central role in protein synthesis by bringing the amino acid corresponding to the mRNA codon present in the ribosomal A-site to the ribosomal peptidyl transferase center (PTC) during the elongation step of the translation process. Post-transcriptional maturation steps are essential for tRNA function. In particular, around 100 nucleoside modifications have been described for tRNAs (1) and were shown to mostly ensure either correct tRNA folding (2) or efficient and accurate decoding (3). Position 34 from the tRNA anticodon loop (also known as wobble) is frequently heavily modified, ensuring translational fidelity but also the recognition of several codons by a single tRNA molecule. In *Saccharomyces cerevisiae*, 13 out of 42 tRNAs have a uridine at position 34 (U34), which is modified into 5-carboxymethyluridine derivatives (xcm^{5}U) in 11 of those tRNAs (4). Among these modifications, 5-methoxycarbonylmethyl-(2-thio)uridine (mcm^{5}U and mcm^{5}s^{2}U) were shown to enhance accurate and efficient translation and codon pairing (5).

The synthesis of mcm^{5}U34 is very complex and requires at least 15 proteins in *S. cerevisiae*. Most of these proteins are involved in the first step of the reaction, i.e. addition of a carboxymethyl group at position 5 of the uracil to form cm^{5}U (6). The enzyme catalyzing this reaction is the Elongator complex composed by six subunits (Elp1–6) (7). This complex consists of two sub-complexes: a core complex Elp1–2–3 where Elp3 is the catalytic subunit endowed with acetyltransferase activity and the Elp4–5–6 complex, an hexameric ATPase regulating tRNA dissociation from the Elongator complex (8). The other factors (Kti11–14, Sit4, Sap185 and Sap190) seem to be involved in the regulation of the Elongator complex (9). The last step in the synthesis of this modification requires the methylation of cm^{5}U to form mcm^{5}U by the Trm9 methyltransferase (MTase) (10). This protein belongs to the class I S-adenosyl-L-methionine (SAM)-dependent MTase family and is active as a complex with Trm112 (11). The deletion of the gene encoding any of these 15 yeast proteins results in zymocin resistance phenotype (7). Indeed, zymocin, a toxin secreted by *Kluyveromyces lactis*, exclusively cleaves the tRNA anticodon loop containing mcm^{5}s^{2}U34, thereby inhibiting translation and leading to yeast death.

Trm112 is a small zinc binding protein that in addition to Trm9 interacts with and activates 3 other MTases: Trm11, Bud23 and Mtq2. These four Trm112-MTase com-

*To whom correspondence should be addressed. Tel: +33 1 69 33 48 90; Fax: +33 1 69 33 49 09; Email: marc.graille@polytechnique.edu
Correspondence may also be addressed to Dominique Liger. Tel: +33 1 69 15 79 68; Fax: +33 1 69 85 37 15; Email: dominique.liger@i2bc.paris-saclay.fr

plexes are involved in processes related to protein synthesis. Similarly to Trm9-Trm112, the Trm11-Trm112 complex methylates many tRNAs in yeast at guanosine 10 to form 2-methylguanosine (m^2G10) and hence is involved in translational elongation (12). The Mtq2-Trm112 complex methylates translation termination factor eRF1 on the amide group of the glutamine side chain from the universally conserved GGQ motif, which enters into the PTC and triggers the release of newly synthesized proteins (13). The Bud23-Trm112 complex is implicated in the synthesis of the small ribosomal subunit by catalyzing the N^7-methylation of guanosine 1575 of 18S rRNA (14,15). Finally, Trm112 is important for synthesis of the large ribosomal subunit by an unknown mechanism (16). The Trm9-Trm112 complex catalyzed tRNA modification enhances decoding of AGA, CAA, GAA and, to a lesser extent, AGG codons (17). Hence, it favors the translation of transcripts specifically enriched in these codons such as those coding for the DNA damage response key proteins Rnr1 and Rnr3 (18). In *S. cerevisiae*, the deletion of the *TRM9* gene results in increased sensitivity of the cell to DNA alkylating agent methyl methanesulfonate (MMS) and in delayed G1 to S phase transition after MMS treatment. Furthermore, tRNA hypomodification following Trm9 inactivation results in translational infidelity and triggers the activation of protein stress response pathways (19). Altogether, this indicates a connection between tRNA modification, regulation of translation in response to stress and DNA damage response.

The Trm9-Trm112 complex is highly conserved in eukaryotes and two Trm9 orthologues exist in human: ABH8 and hTrm9L (20). Similarly to yeast Trm9-Trm112, the human ABH8-TRM112 complex converts cm^5U into mcm^5U (20). In insects, worms and human, ABH8 proteins are bifunctional enzymes also encompassing an AlkB-like domain responsible for the hydroxylation of mcm^5U into (S)-5-methoxycarbonylhydroxymethyluridine ((S)-$mchm^5U$)(21). Compared to ABH8, hTrm9L is made of only the MTase domain (22). The ABH8 protein is highly expressed in a variety of human cancer cells. Furthermore, ABH8 silencing induces apoptosis of urothelial carcinoma cells thereby suppressing tumor growth, angiogenesis and metastasis (23). On the opposite, hTrm9L has been described as a negative regulator of tumor growth (24). Interestingly, ABH8 silencing renders cells sensitive to MMS and to the anti-cancer drug bleomycin, while the loss of the gene encoding hTrm9L renders tumor cells sensitive to paromomycin and gentamycin, two antibiotics known to induce translational errors (24). Hence, human hTrm9L and ABH8 proteins represent potent targets for the development of new anti-cancer drugs.

Here, we describe the crystal structure of the Trm9-Trm112 complex from *Yarrowia lipolytica* as well as *in vivo* and *in vitro* functional studies of *S. cerevisiae* enzyme with the aim of analyzing the mechanism of action of this protein. From the comparison of the various known crystal structures of Trm112-MTase complexes, we also unravel the molecular plasticity allowing Trm112 to interact with its various MTase partners, which share less than 20% sequence identity.

MATERIALS AND METHODS

Yeast strains, media and growth conditions

Saccharomyces cerevisiae strain YPH499 (Agilent technologies) was used as the wild-type host for all yeast gene manipulations. *Kluyveromyces lactis* AWJ137 and NK40 strains were used as source of zymocin and control strain, respectively, in zymocin killer eclipse assay and killer liquid assay (25).

Cultures were performed at 30°C in standard rich medium YEPD (1% yeast extract, 2% peptone, 2% dextrose) or selective minimal media (SD) with 2% dextrose or galactose. Yeast was transformed by the lithium acetate method as previously described (26). For selection, YEPD was supplemented with geneticin (200 μg/ml) and SD minus uracil or tryptophan was prepared.

Yeast strains construction

For clone selection, three steps were systematically considered: growth on a selective media, polymerase chain reaction (PCR) screening and DNA sequencing. For point mutation strain selection, the presence of the mutation was also investigated by restriction fragment length polymorphism prior to DNA sequencing.

Trm9 point mutant strains were generated using genomic DNA from strain YDL201 (YPH499 *TRM9::kanMX6*) as DNA template and the suitable Ftrm9/R1-trm9 set of primers for initial cassette amplification (27) and the technique already described (28) (Supplementary Tables S2 and S3).

Clones expressing wild-type or mutant Trm9–13Myc were obtained from transformation of the appropriate strain (encoding wild-type or mutant Trm9) with PCR product amplified from pFA6a-13Myc-TRP1 plasmid as a template using F2-Trm9/R1-trm9 primer set (27).

The *trm9Δ::URA3* strain was obtained by transforming wild-type strain with PCR product amplified from plasmid pESC-URA (Agilent Technologies) using F1-URA-trm9/R1-URA-trm9 primer set. Clones expressing *Yl*Trm9 or *Yl*Trm9N38 were obtained from a 5-FOA selection after transformation of *trm9Δ::URA3* strain with PCR product amplified from *Yarrowia lipolytica* genomic DNA as a template using respectively Ftrm9::YLTRM9 or Ftrm9::YLTRM9N38 / Rtrm9::YlTRM9 primer sets (Supplementary Table S3).

Zymocin killer eclipse assay

5 μl suspension of the yeast cells to test (OD_{600nm} = 0.5) were spotted on YEPD plate and air-dried. Next, using the tip of a toothpick, *K. lactis* AWJ137 and NK40 cells were placed onto the edge of the yeast spot. Plate was incubated at 30°C for 1–2 days. Sensitivity or resistance to zymocin was revealed by the presence or absence of a halo zone around AWJ137 colony.

Zymocin liquid killer assay

Filtered-sterile supernatants from overnight culture of *K. lactis* AWJ137 and NK40 were used as source of zymocin

and as a control, respectively. Serial dilutions of the supernatants (from 10^{-1} to 10^{-4}) were prepared in fresh selective SD media with galactose as carbon source. 5 ml sample of each dilution was inoculated at OD_{600nm} of 0.1 with yeast strain to be tested. Cultures were incubated for 20 h at 30°C. Growth was calculated as the ratio between the OD_{600nm} read for the 10^{-1} and 10^{-4} dilutions. The 100% sensitivity to zymocin was obtained from the wild-type strain.

Cloning expression and purification

A DNA sequence optimized for heterologous expression in *E. coli* was designed to encode *Yl*Trm112 (UniprotKB entry: Q6C4P5) and *Yl*Trm9 with a C-terminus His_6-tag (UniprotKB entry: Q6C999). This fragment was obtained by *de novo* gene synthesis (GenScript Corporation, Piscataway, NJ, USA) and was further subcloned into pET21-a between *Nde*I and *Xho*I sites. Two truncated forms of *Yl*Trm9 (*Yl*Trm9N19 and *Yl*Trm9N38 encompassing residues 19–324 and 38–324, respectively) were cloned in the same operon (Supplementary Tables S4 and S5). These complexes were expressed in *E. coli* BL21 (DE3) Gold (Novagen) in 800 ml autoinducing media (29) containing 10 μM $ZnCl_2$ and 100 μg/ml ampicillin, 5 h at 37°C and 15 h at 15°C.

Genes encoding wild-type *Sc*Trm9 and *Sc*Trm112 were amplified from *S. cerevisiae* S288C genomic DNA and cloned into a modified pET28a and a pACYCDUET-1 plasmids, respectively (Supplementary Tables S4 and S5). Genes encoding *Sc*Trm9 mutants were amplified from the genomic DNA isolated from YPH499 mutated strains and then cloned into pET21-a vector using *Nde*I and *Not*I restriction sites and introducing a C-terminal His_6-tag (Supplementary Tables S4 and S5). The expression of the *Sc*Trm9-Trm112 complexes was done in *E. coli* BL21(DE3)Gold strain (Novagen) co-transformed with the two vectors and cultured in 1 l of 2YT medium containing 10 μM $ZnCl_2$, 25 μg/ml chloramphenicol and 100 μg/ml ampicillin. Cultures were done at 37°C until OD_{600nm} reached 0.5–0.6, they were then shifted to 20°C and protein production was induced overnight with IPTG (final concentration of 0.5mM).

Bacteria were harvested and resuspended in 30 ml buffer A (20mM Tris-HCl pH7.5, 200mM NaCl, 5mM ß-mercaptoethanol, 10μM $ZnCl_2$) and stored at −20°C. Cells were lysed by sonication and the soluble fraction was cleared by centrifugation (15 000*g* for 30 min at 4°C). All His_6-tagged complexes were purified by Ni-NTA chromatography (QIAGEN Inc.), followed by an ion-exchange chromatography (5 ml Heparin column for *Sc* complexes and *Yl* full length complex and 5 ml MonoQ column for the *Yl* truncated forms (GE Healthcare)). The last purification step was performed on a Superdex 75 16/60 size-exclusion chromatography column (GE Healthcare) in buffer A.

Crystallization and structure solution

Crystals were grown after 3–4 days at 19°C from a mixture of *Yl*Trm9N38-Trm112 complex (10 mg/ml) with an equal volume of the crystallization solution containing 100mM tri-sodium citrate, 20% polyethylene glycol 4000 (PEG4K), 20% 2-propanol. Crystals were cryoprotected by transfer into the crystallization solution supplemented with 15 then 30% v/v ethylene glycol and then flash-frozen in liquid nitrogen. The data sets were collected at 100 K on Proxima-1 beam line (SOLEIL, St Aubin, France). The structure was solved by Zn-MAD (Multiple Anomalous Dispersion) using data sets collected at three wavelengths corresponding to inflexion, peak and remote of Zn-edge (see Table 1 for statistics). Data were processed with XDS (30) and scaled using XSCALE. The space group was I432 ($a = b = c =$ 176.2 Å) with one heterodimer per asymmetric unit, corresponding to a solvent content of 56.4%. As expected from the presence of one Trm112 protein in the asymmetric unit, one Zn atom could be located by the HYSS submodule of the PHENIX package in the 50–3 Å resolution range (31). Refinement of the Zn atom coordinates, phasing and density modification were performed with SHARP program using the 50–3 Å resolution range (32). The *Sc*Trm112 crystal structure (13) was positioned into the experimental density maps by molecular replacement using the MOLREP program (33) and then modified to match with the *Yl*Trm112 sequence. The Trm9 model was built into these maps using the COOT molecular modeling program (34). Iterative cycles of manual model rebuilding using COOT followed by refinement with PHENIX led to an almost complete model, which was then completed and refined using a higher resolution data set (2.5 Å) to yield the final model (final R and R_{free} values of 19.9% and 24.2%, respectively). The statistics for data collection and refinement are summarized in Table 1. The final model contains residues 1–124 from Trm112, residues 39–160 and 172–233 from Trm9, a zinc atom and 91 water molecules. According to PROCHECK (35), in this final model, 87% and 13% of the residues are in the most favored and in the allowed regions of the Ramachandran diagram, respectively. Due to the absence of electron density, the following regions are absent from the final model: residues 125–130 from Trm112 and 38, 161–171, 234 and the His_6 tag from Trm9.

Enzymatic assay

For enzymatic assays, total tRNAs were purified from *S. cerevisiae trm9Δ* and *elp1Δ* strains (*elp1Δ* strain is a kind gift from Dr B. Seraphin, IGBMC, France) using mostly the protocol described in Chen et al. (36). Yeast cells were grown in YEPD medium at 30°C until cell density reached 1 to 2×10^7 cells/ml. Cells were washed with water and harvested. The pellet was resuspended in one pellet volume of 0.9% NaCl then 2 volumes of phenol were added. The mixture was incubated on a rotating wheel for 30 min at room temperature. One volume of chloroform was added and the mixture was shaken for 15 min and then centrifuged for 20 min at 13000 rpm at 4°C. Aqueous phase was precipitated by 2.5 volumes of ethanol and 0.1 volume of 20% potassium acetate and washed with cold ethanol 95%. In order to remove bound amino acids, the precipitate was resuspended in 1.5 ml of 2 M Tris-HCl pH 8 (for 10 g of cells) and incubated for 90 min at 37°C. After ethanol precipitation, the pellet was resuspended in 3.4 ml of 2 M lithium acetate, 0.1 M potassium acetate pH 5 (for 10 g of cells) to remove ribosomal RNAs as described previously (37). The mix was shaken for 20 min at 4°C and centrifuged at 8000*g* for 20

Table 1. Data collection, phasing and refinement statistics

		Crystal 1		Crystal 2
Data collection				
Space group		I432		I432
Cell dimensions				
a, *b*, *c* (Å)		176.40		176.20
α, β, γ (°)		90.00		90.00
	Peak	*Inflection*	*Remote*	
Wavelength (Å)	1.2819	1.2826	1.2753	0.9801
Resolution (Å)	50–3 (3–3.17)	50–3.3 (3.3–3.5)	50–3.7 (3.7–3.97)	50–2.5 (2.5–2.65)
R_{sym} or $R_{merge\ (\%)}$	14.8 (67.7)	15.2 (63.6)	17.5(58.9)	6.9(50.1)
I / σI	20.3 (5.1)	22.1 (6.2)	10.9 (3.6)	14.2 (2.6)
Completeness (%)	99.9 (99.7)	99.9 (99.7)	99.9 (100)	97.9 (98.3)
Redundancy	20.9	23.5	6.9	4.8
Refinement				
Resolution (Å)				50–2.5
$R_{work}/R_{free\ (\%)}$				19.9/24.2
No. atoms				
Protein				1496 (Trm9) / 966 (Trm112)
Ligand/ion				1 (Zn)
Water				76
B-factors (Å2)				
Protein				49.5 (Trm9) / 64.7 (Trm112)
Ligand/ion				76.6
Water				50.9
R.m.s deviations				
Bond lengths (Å)				0.009
Bond angles (°)				1.193

min at 4°C. The pellet was extracted a second time with 1.7 ml of 2 M lithium acetate, 0.1 M potassium acetate pH 5. The pool of supernatants was dialyzed against 10 μM magnesium acetate during 2.5 h. The tRNAs were resuspended in 1 mM Tris pH 7.5, 10 mM magnesium acetate following ethanol precipitation.

The methyltransferase assay was performed in a final volume of 50 μl containing 50 mM phosphate buffer pH 7.5, 0.1 mM EDTA, 10 mM $MgCl_2$, 10 mM NH_4Cl, 0.1 mg/ml Bovine Serum Albumin, 10 μM SAM (containing 0.87 Ci/mmol of [^{3}H]-SAM, Perkin Elmer) and 1.5 pmol of Trm9-Trm112 complexes. The reaction was initiated by adding 1.5 μM of total tRNAs (75 pmol) purified from *trm9Δ* strain to the mixture. The samples were incubated at 37°C and aliquots were withdrawn at different time points. The reaction was stopped by precipitation with cold trichloroacetic acid (5%), followed by filtration on Whatman GF/C filters. The [^{3}H] incorporation was measured using a Beckman Coulter LS6500 scintillation counter. For the pH dependence enzymatic assays, phosphate buffer pH7.5 was replaced by other phosphate buffers ranging from pH 5–8. All reactions were performed in triplicates. The initial velocities (V_i) were calculated using the equation $C_{tRNA} = V_i{*}(1 - \exp(-nt))/n$ where C_{tRNA} is the concentration of methylated tRNA, t the time in minutes, V_i the initial enzyme cycling velocity and n the relaxation rate constant of V_i by fitting the experimental spots with the software ORIGIN according to Cao et al. (38).

Steady-state kinetics analyses were performed by using various tRNA (from 0.1 to 1.5 μM) or SAM (from 1 to 20 μM) concentrations. The initial rate calculated for each tRNA or SAM concentration was plotted as a function of the concentration. The data were fit to the Michaelis–Menten equation using the ORIGIN software and the K_m, k_{cat} and k_{cat}/K_m values were calculated from the curve fitting.

Equilibrium dialysis

The equilibrium dialysis set up consists of two chambers separated by a 3 kDa-cut off dialysis membrane. The experiment was performed in 100 mM Tris-HCl pH 7.5, 0.1 mM EDTA, 10 mM $MgCl_2$, 10 mM NH_4Cl. Initially, the ligand chamber contains 100 μM SAM (containing 0.87 Ci/mmol of [^{3}H]-SAM, Perkin Elmer) and the protein chamber 120 μM of Trm9-Trm112 complex. After 3 h rotation at 4°C, the SAM content of chambers is determined using a Beckman Coulter LS6500 scintillation counter.

Co-IP and western blot

The preparation of the soluble protein extracts and the co-immunoprecipitation assays were performed as previously described (26). Probing was performed using either mouse 9E10 anti-myc (Santa Cruz Biotechnology) or polyclonal rabbit anti-*Sc*Trm112 antibodies (Agrobio) as primary antibody (1/5000 dilution). Sheep anti-mouse or sheep anti-rabbit HRP-conjugated IgG were used as secondary antibody (1/5000, GE Healthcare).

Generation of *Yl*Trm9N20-Trm112 model

Residues 20–39 from *Yl*Trm9, which are missing in the crystal structure of *Yl*Trm9N38-Trm112 complex, were modeled by superimposing the crystal structure of RPA2492, the closest structural *Yl*Trm9 homologue (PDB code: 3E23), onto our structure. The SAM substrate and backbone atoms from residues 8–27 (corresponding to residues 20–39 from *Yl*Trm9) were retained from the crystal structure

3E23. Missing side chains were placed using the SCWRL4 software (39). The cm^5U residue was docked manually into the *Yl*Trm9 active site using constraints issued from the functional analysis of the various Trm9 mutants and by positioning the carboxyl group of cm^5U at a distance compatible with methyl transfer from SAM. Finally, as the loop containing R40 (corresponding to *Sc*Trm9 R39) is one residue shorter in *Yl*Trm9 than the corresponding region from RPA2492, which was used as a template to build this fragment, we applied a restraint with a force constant of 1000 kcal/mol/Å^2 between the carboxyl group of cm^5U and guanidinium group of R40 to position the side chain of this latter group toward the binding pocket. The atoms present in the crystal structure of *Yl*Trm9 were restrained to the experimental crystallographic position with a force constant of 5 kcal/mol/Å^2. The CHARMM27 force field was used (40). The complex was subsequently minimized in the CHARMM software (41) using the Powell algorithm and 10 000 steps of minimization.

2D thin layer chromatography for detection of modified nucleosides

The tRNAs obtained after *in vitro* methylation using [^{14}C]-SAM (containing 6.5 μl of [^{14}C]-SAM (58 mCi/mmol, Perkin Elmer)) were extracted with phenol/chloroform and ethanol precipitated. The tRNA pellet was dissolved in 10 μl of 50 mM ammonium acetate pH 5.3 supplemented with 1 μg of P1 nuclease from *Penicillium citrinum* (Sigma) and incubated overnight at 37°C. Digested tRNAs (2 μl) were mixed with 12μg of cold 5′ P-mononucleosides pA, pU, pG and pC and spotted on a CEL-300 cellulose plate (Merck). 5′ P-mononucleosides were separated using chromatographic solvents A, B or C as previously described (42). First dimension chromatography was performed in solvent A, and the second in either solvent B or C. The positions of the four major mononucleosides (pA, pG, pU and pC) were revealed by UV shadowing. The position of mcm^5U nucleoside carrying radiolabeled methyl group was revealed by phosphorimaging and compared with that of reference maps obtained under identical experimental conditions (43).

RESULTS AND DISCUSSION

Structure of the *Yl*Trm9-Trm112 complex

To determine the crystal structure of the Trm9-Trm112 complex, we expressed in *Escherichia coli* and purified this complex from different organisms (*Saccharomyces cerevisiae* (*Sc*), *Schizosaccharomyces pombe* (*Sp*), *Encephalitozoon cuniculi* (*Ec*) and *Yarrowia lipolytica* (*Yl*)). Unfortunately, no crystals could be obtained from these complexes formed by full-length proteins. Analyses of these purified complexes stored at 4°C for several weeks by SDS-PAGE followed by mass spectrometry peptide mass fingerprint revealed degradation of the N-terminal extremity from *Yl*Trm9. Therefore, two truncated *Yl*Trm9 forms lacking residues 1–18 or 1–37 (hereafter named *Yl*Trm9N19 and *Yl*Trm9N38, respectively) were cloned, co-expressed with *Yl*Trm112 and purified. Compared to full-length *Yl*Trm9-Trm112 complex, both truncated forms were strongly affected in their enzymatic activity of methylation toward a tRNA mixture isolated from a *S. cerevisiae trm9Δ* strain but displayed the same affinity for SAM, indicating that the N-terminal residues deleted in these constructs are required for optimal enzymatic activity (Supplementary Figure S1). Diffracting crystals were obtained for the *Yl*Trm9N38-Trm112 complex and its structure was solved by the MAD method using the anomalous signal of the zinc atom bound to Trm112. The final model was refined to 2.5 Å and yielded R and R_{free} values of 19.9% and 24.2%, respectively (Table 1, see Supplementary Figure S2 for electron density maps).

*Yl*Trm9N38 adopts the typical fold of the class I SAM-dependent MTases composed of a central seven-stranded ß-sheet surrounded by two α-helices on each side (Figure 1A). A twisted two-stranded antiparallel ß-sheet (strands ßA and ßB) is inserted between strand ß5 and helix αE and forms a lid positioned on top of the C-terminal extremity of the central ß-sheet. *Yl*Trm112 is formed by a zinc binding domain made of two α-helices (α1–2) and a four-stranded antiparallel ß-sheet as well as a helical domain of 3 α-helices (α3 to α5). The *Yl*Trm112 structure is very similar to the crystal structures of *Sc*Trm112 either in its free form (13) or bound to Bud23 (15) (rmsd of 1.21–1.38 Å; 49% sequence identity). The only major difference between *Yl*Trm112 and free *Sc*Trm112 results from a rearrangement of the C-terminal extremity from *Yl*Trm112 (corresponding to residues F122–L124). A similar rearrangement of this Trm112 region is also observed in the Mtq2-Trm112 and Bud23-Trm112 complexes and is required to avoid a steric clash between this Trm112 fragment and its MTase partner (26).

Trm9 binds mainly to Trm112 zinc-binding domain via a parallel ß-zipper interaction formed by Trm9 strand ß3 and Trm112 strand ß4, which results in the formation of an extended eleven-stranded ß-sheet (Figure 1A and B). Complex formation engages an interface area of 1160 Å^2 involving 20 and 21 residues from Trm112 and Trm9, respectively (Figure 1D and E). At the center of this interface, hydrophobic residues (M1, F8, V9, F43, M47, I119, P120 and F122) from Trm112 contact hydrophobic side chains from Trm9 (V59, F78, V102, A110, P112 and F116), thereby shielding this Trm9 region from exposure to solvent (Figure 1B and C). This explains the need to express Trm112 with Trm9 in *E. coli* to obtain soluble and active Trm9. This hydrophobic core is surrounded by several hydrogen bonds and salt bridges (see Supplementary Table S1 for details). Among these, the ß-zipper interaction is realized by two hydrogen bonds engaging main chain atoms from Trm9 V103 as well as P120 and F122 from Trm112 (Figure 1B). Other hydrogen bonds are formed by E100 from Trm9 with K2 and T5 from Trm112, by R115 from Trm9 with N7 and the carbonyl group of R50 from Trm112 and by A110 and E113 from Trm9 with Q10 from Trm112 (Figure 1B and C). Finally, a salt bridge is formed between Trm9 D117 and Trm112 R50. Among the Trm9 residues engaged in complex formation, E100, V102 and F116 from *Yl*Trm9 correspond to N89, L91 and F105 from *Sc*Trm9, which substitutions by Lys, Arg and Glu, respectively, were previously shown to disrupt and inactivate the *Sc*Trm9-Trm112 complex (26). It is noteworthy that the truncation of residues 263–279 from *Sc*Trm9 resulted in loss of interaction with Trm112 (11).

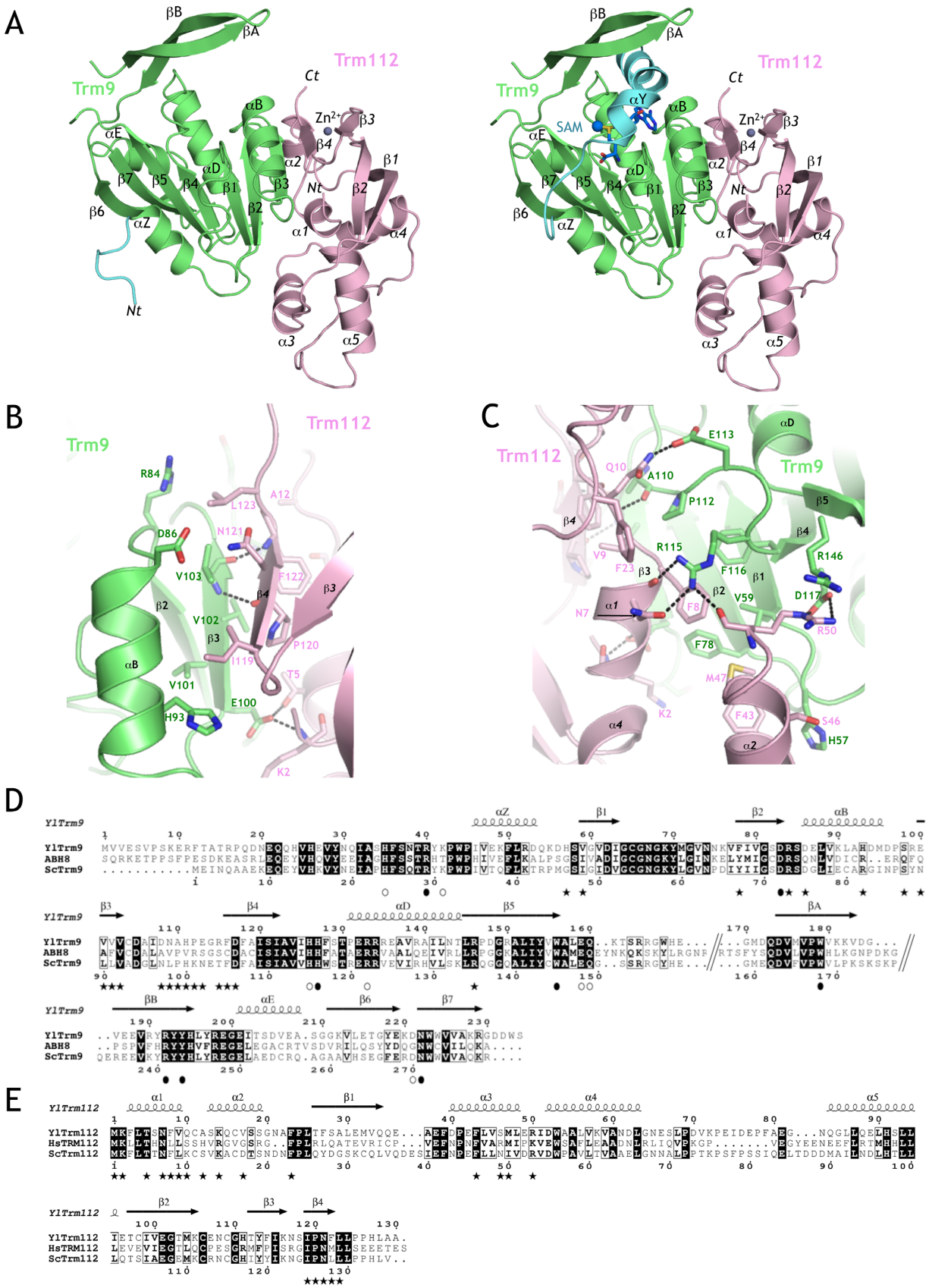


Figure 1. Structure of the *Yl*Trm9-Trm112 complex. (**A**) Ribbon representations of the crystal structure of the *Yl*Trm9N38-Trm112 complex (left) and of the model of the *Yl*Trm9N20-Trm112 complex (right). On the left panel, residues 39–44, which adopt different conformations between our crystal

The corresponding residues from *Yl*Trm9 (residues 214–234) constitute strands ß6 and ß7 and are not directly involved in Trm112 interaction. Hence, the loss of interaction with Trm112 most likely results from incorrect folding of this Trm9 truncated form.

Active site mapping

Despite extensive efforts, we could not obtain crystals of the *Yl*Trm9-Trm112 or *Yl*Trm9N38-Trm112 complexes bound to SAM or SAH. We have then modeled a SAM molecule by superimposing onto *Yl*Trm9 structure our SAM-bound structure of the *Sc*Bud23-Trm112 complex (15) (rmsd of 2 Å over 134 Cα atoms; 19% sequence identity). We have further observed that upon SAM binding, *Sc*Bud23 N-terminal region folds as an α-helix that lies onto SAM. In addition, a similar α-helix is present in the structure of RPA2492 from *Rhodopseudomonas palustris*, the protein sharing the highest structural similarity with Trm9 (Supplementary Figure S3A; rmsd of 1.9 Å over 200 Cα atoms; 17% sequence identity; PDB code: 3E23). Finally, secondary structure predictions for the N-terminal region from Trm9 proteins strongly suggest that this region has indeed a high propensity to fold as an α-helix. Based on these observations, in our model of the *Yl*Trm9-Trm112 complex bound to SAM, residues 20–45 from *Yl*Trm9 fold as two α-helices (Figure 1A) and the side chains from *Yl*Trm9 Y29 and F36 (*Sc*Trm9 Y18 and F25) match with RPA2492 Y17 and Y24, and are located onto SAM, further validating our model (Supplementary Figure S3B). Finally, in this model, the Trm9 loop connecting strands ß3 and ß4 is sandwiched between Trm112 on one side and SAM on the other. As shown for Mtq2-Trm112 (26), Trm112 could stabilize this loop and hence confer SAM binding activity to Trm9. This would then rationalize in part the role of Trm112 in Trm9 activation.

Mapping of the sequence conservation at the surface of this model reveals the presence of a patch formed by highly conserved residues centered on the SAM methyl group (Figure 2A). Several residues from this patch (H24, R29, K31, H115, H116, W145, Q149, W168, R241, Y243, D270 and N271; for clarity, *Sc*Trm9 numbering will be used in this paragraph as functional analysis were performed in *S. cerevisiae*, see Table 2 for correspondence between *S. cerevisiae* and *Y. lipolytica* numbering) or from other Trm9 regions (R122 and E148) were mutated into alanine to test their role in *S. cerevisiae* Trm9 activity. As a control, we have also considered the D72A catalytic mutant deficient in SAM binding (26). Mutations were introduced into the chromosomic copy of *S. cerevisiae TRM9* gene and the activity of these mutants was first tested *in vivo* using the zymocin killer eclipse assay. While seven mutants (H24A, K31A, H115A, R122A, E148A, Q149A and D270A) exhibited the same phenotype as WT Trm9, eight mutants (R29A, D72A, H116A, W145A, W168A, R241A, Y243A and N271A) were resistant to zymocin suggesting that their tRNA modification enzymatic activity is strongly impaired (Table 2 and Figure 2B). These eight mutants as well as two zymocin sensitive mutants (H115A and Q149A) affecting residues directly oriented toward the putative active site were selected for further functional analyses. Co-immunoprecipitation experiments demonstrated that the loss of *in vivo* activity of these mutants is not due to disruption of the Trm9-Trm112 complex (Figure 2C). This further indicated that all these mutants accumulate in the cells although to various extents depending on the mutants and that these mutations might directly impact the enzymatic activity of the complex.

All these complexes were expressed in *E. coli* and exhibited the same purification profile as the wild-type complex strongly suggesting that they are properly folded. The *S. cerevisiae* Trm9-Trm112 complex was previously demonstrated to convert cm^5U into mcm^5U using SAM as methyl donor (44) and we have confirmed this by 2D-TLC (Supplementary Figure S4A). We have further measured the *in vitro* enzymatic activity of Trm112-Trm9 mutants at pH 7.5, which corresponds to the optimum of activity (Supplementary Figure S4B and C), by using [^{3}H]-SAM and total tRNAs purified from *S. cerevisiae trm9Δ* strain as substrates. As this mixture of tRNAs contains both substrate tRNAs and non-substrate tRNAs, the latter potentially developing inhibitory effect on Trm9 activity, only apparent kinetic properties (initial velocity, specific activity, K_m and k_{cat}) could be determined. Among the mutants exhibiting zymocin resistance *in vivo*, the D72A, H116A, W145A and W168A mutants were inactive while the R29A, R241A, Y243A and N271A mutants exhibited apparent specific activities at least one order of magnitude lower than the *Sc*Trm9-Trm112 WT enzyme (Figure 2D, Table 2). The two mutants associated to the zymocin sensitivity phenotype proved to be only slightly affected (H115A) or as active (Q149A) as the wild-type enzyme, respectively. Hence, the effect of mutations on both zymocin phenotype and enzymatic activity are correlated. Interestingly, the H115A and Y243A mutants, which modify the same amount of tRNA after 1 h in our experimental conditions, exhibit sensitive and resistant zymocin phenotype, respectively. From detailed kinetics analysis, it appears that Y243A mutant ex-

structure and our model, are colored in blue. On the right panel, the modeled *Yl*Trm9 fragment encompassing residues 20–39 (helix αY) is shown in blue. The modeled SAM molecule is shown in sticks. *Yl*Trm112 secondary structure elements are labeled in italics. (**B** and **C**) Detailed views of the residues involved in *Yl*Trm9N38-Trm112 interface. Hydrogen bonds and salt bridges are show by black dotted lines. (**D**) Sequence alignments of Trm9 orthologs from *Y. lipolytica* (*Yl*Trm9), *Homo sapiens* (ABH8) and *S. cerevisiae* (*Sc*Trm9). Strictly conserved residues are in white on a black background. Partially conserved amino acids are boxed. Secondary structure elements assigned from the *Yl*Trm9N38 crystal structure are indicated above the alignment. Black stars indicate residues involved in complex formation. Filled or open circles below the alignment indicate residues, which mutation into Ala results in zymocin resistance or sensitivity, respectively. For clarity, both *Yl*Trm9 and *Sc*Trm9 numbering are indicated. Two regions corresponding to fragments present only in *Sc*Trm9 or ABH8 have been omitted. (**E**) Sequence alignments of Trm112 orthologs from *Y. lipolytica*, *Homo sapiens* and *S. cerevisiae*. Strictly conserved residues are in white on a black background. Partially conserved amino acids are boxed. Secondary structure elements assigned from the *Yl*Trm112 crystal structure are indicated above the alignment. Black stars indicate residues involved in complex formation. Panels D and E were generated using the ESPript server (49).

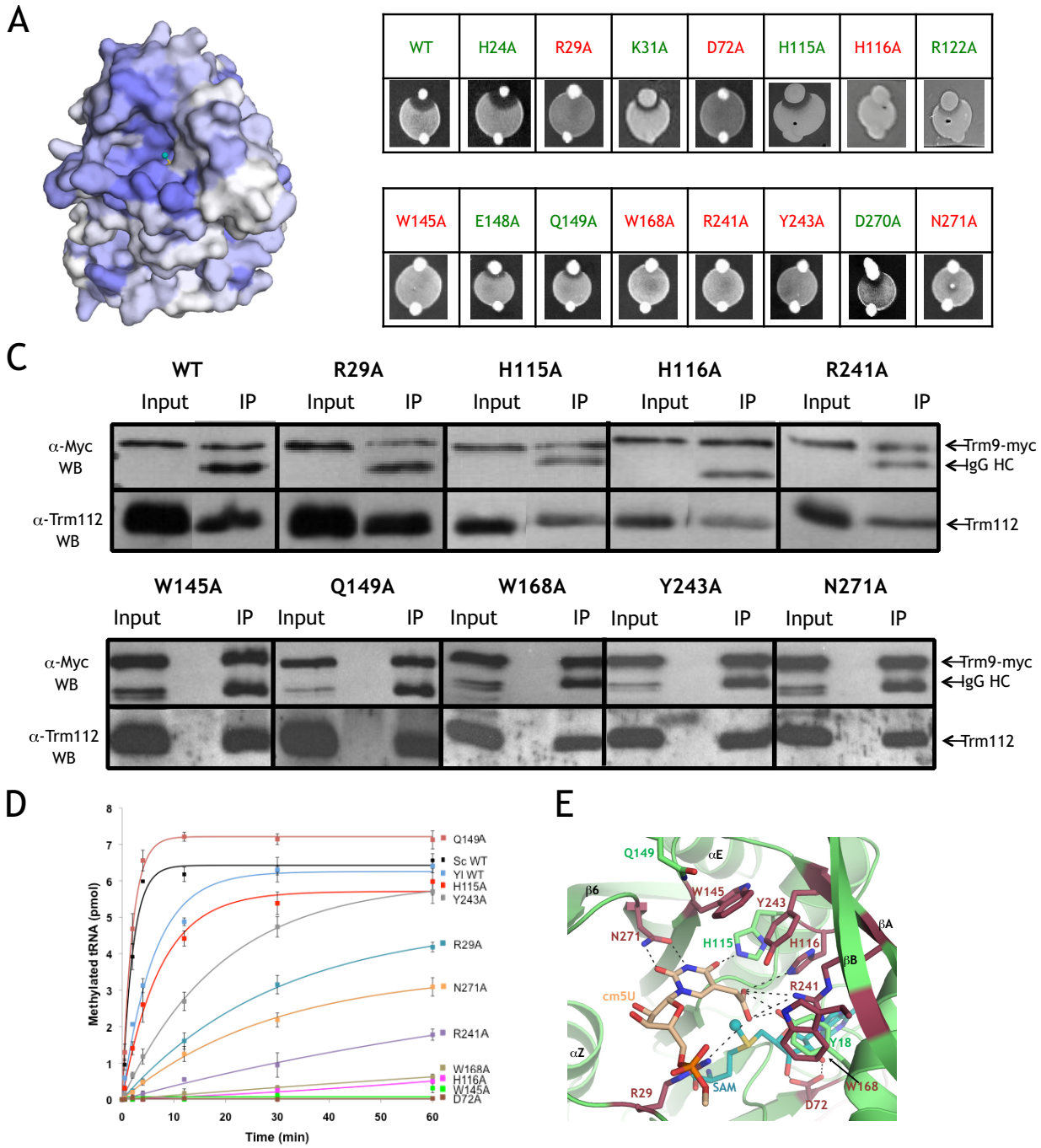


Figure 2. Trm9 active site mapping. (**A**) Mapping of the sequence conservation score at the surface of the *Yl*Trm9-Trm112 model. Coloring is from gray (low conservation) to blue (highly conserved). The conservation score was calculated using the CONSURF server (50). (**B**) Analyses of the zymocin phenotype of yeast mutant strains by eclipse assay. Each *S. cerevisiae* mutated strain was subjected to killer eclipse assay with the *Kluyveromyces lactis* AWJ137 killer strain (top) and the *Kluyveromyces lactis* NK40 non-killer strain (as a control, bottom). The presence or absence of an eclipse around the killer strain shows the sensitivity or resistance of the mutated strain to zymocin toxin, respectively. Resistant and sensitive strains are labeled in red and green, respectively. (**C**) Effect of Trm9 mutations on *Sc*Trm9/*Sc*Trm112 *in vivo* interaction. Soluble protein extracts (Input: $1/50^{th}$ of total proteins, i.e. 10 μg) and immunoprecipitates (IP: $1/10^{th}$ of immunoprecipitated material) were subjected to 15% SDS–PAGE analysis and immunoblotted using mouse anti-Myc (Trm9–13Myc) or rabbit anti-Trm112 as primary antibodies and sheep anti-mouse or sheep anti-rabbit HRP-conjugated IgG as secondary antibodies, respectively. Note that for the anti-myc probing, the lower bands observed in some of the input lanes result from Trm9-myc protein degradation whereas the band observed in the IP lanes comes from mouse primary antibody heavy chain cross-reacting with anti-mouse secondary antibody. IP were performed using 9E10 anti-myc monoclonal antibodies. (**D**) Enzymatic activity of Trm112-Trm9 mutants. The curves obtained after fitting of the experimental data with equation given in the Materials and Methods section are shown by lines using the same color code as for the symbols. (**E**) Detailed representation of the *Yl*Trm9 active site. Side chains from residues which substitution by Ala strongly (brown) or only weakly (green) affect enzymatic activity are shown as sticks. For clarity, *Sc*Trm9 numbering is used. A manually docked cm^5U nucleotide is shown as beige sticks and potential hydrogen bonds that it could form with Trm9 active site residues are depicted by dashed black lines. The modeled SAM molecule is shown as blue sticks and the methyl group to be transferred is shown as a sphere.

Table 2. Functional and enzymatic analysis of Trm9 mutants

*Sc*Trm9 mutants (*Yl*Trm9 numbering)	Zymocin sensitivity	Apparent specific activity[a]	SAM binding (%)[b]	Apparent K_m for tRNA (μM)[c]	Apparent k_{cat} for tRNA (10^{-3} s^{-1})[c]	Apparent k_{cat}/K_m for tRNA ($M^{-1}.s^{-1}$)	Apparent K_m for SAM (μM)[c]	Apparent k_{cat} for SAM (10^{-3} s^{-1})[c]	Apparent k_{cat}/K_m for SAM ($M^{-1}.s^{-1}$)
Sc WT	+	2105 ± 221	100	0.08 ± 0.02	32 ± 2	400x10^3	4.9 ± 1.1	23.7 ± 2.2	4837
H24A *(H35)*	+	ND	ND	ND	ND	ND	ND	ND	ND
R29A *(R40)*	-	108 ± 5	93 ± 21	0.936 ± 0.149	1.9 ± 0.15	2.03x10^3	5.4 ± 1.6	2.5 ±0.25	463
K31A *(K42)*	+	ND	ND	ND	ND	ND	ND	ND	ND
D72A *(D83)*	-	0	0	ND	ND	ND	ND	ND	ND
H115A *(H126)*	+	517 ± 39	143 ± 18	0.351 ± 0.118	6 ± 0.7	17.1x10^3	13.5 ± 2.8	11.7 ± 0.1	867
H116A *(H127)*	-	< 10	51 ± 5	ND	ND	ND	ND	ND	ND
R122A *(R133)*	+	ND	ND	ND	ND	ND	ND	ND	ND
W145A *(W156)*	-	< 10	85 ± 12	ND	ND	ND	ND	ND	ND
E148A *(E159)*	+	ND	ND	ND	ND	ND	ND	ND	ND
Q149A *(Q160)*	+	2498 ± 149	98 ± 18	ND	ND	ND	ND	ND	ND
W168A *(W179)*	-	< 10	46 ± 16	ND	ND	ND	ND	ND	ND
R241A *(R192)*	-	26 ± 3	173 ± 11	0.374 ± 0.008	0.8 ± 0.006	2.14x10^3	19 ± 16	0.85 ± 0.42	45
Y243A *(Y194)*	-	206 ± 5	126 ± 11	0.851 ± 0.222	4.5 ± 0.5	5.29x10^3	10 ± 3	2.9 ± 0.4	290
D270A *(D221)*	+	ND	ND	ND	ND	ND	ND	ND	ND
N271A *(N222)*	-	80 ± 3	67 ± 18	0.337 ± 0.09	1 ± 0.078	2.97x10^3	10 ± 6	0.25 ± 0.07	25
Yl WT	±	678 ± 65	ND	ND	ND	ND	ND	ND	ND

[a]Apparent specific activity (fmol of tRNA methylated/min/pmol of enzyme) calculated from apparent initial velocity. The kinetics were performed with 1.5 pmol of enzyme.
[b]These are relative values calculated fixing WT as 100%.
[c]These values were determined by fitting the data using the Michaelis–Menten equation.
ND: Not determined.

hibits reduced specific activity compared to the H115A mutant (Table 2). Therefore the threshold of Trm9 specific activity determining the cell phenotype toward zymocin lies somewhere in the window of specific activities defined by H115A and Y243A mutants. This indicates that the zymocin resistance phenotype *in vivo* is an effective tool for selecting mutants with significantly depressed Trm9 activity compared to the WT, phenotype toward zymocin switching from sensitive into resistant when *in vitro* Trm9 apparent specific activity is between 25% (H115A, sensitive) and 10% (Y243A, resistant) of the WT. We finally verified that the inactivation of these mutants did not result from defect in SAM binding using equilibrium dialysis (Table 2). As expected, the D72A mutant is inactive due to its complete loss of SAM binding capacity. The other mutants were still able to bind SAM although to different extent compared to WT complex. For these mutants, we do not observe any correlation between enzyme specific activity and SAM binding measurements, indicating that the decrease/loss of enzyme activity cannot be attributed to its reduced ability to bind SAM. Altogether, these results strongly suggest that the H116A, W145A, W168A, Y243A and to a lesser extent R29A, R241A, N271A and H115A *Sc*Trm9 mutants affect enzyme activity due to the direct involvement of these residues in catalysis or in binding of tRNA substrate.

Substrate binding and catalytic mechanism

Our site-directed mutagenesis strategy on strictly conserved residues surrounding the SAM methyl group in the *Yl*Trm9N38-Trm112 crystal structure has led to the identification of several Trm9 residues crucial for formation of mcm^5U34 in some tRNAs (Figure 2B–D). The Trm9-Trm112 complex catalyzes the O-methylation of the carboxylic function of the carboxymethyl group of cm^5U34. Due to the nucleophilic property of oxygen atoms, it is generally assumed that the methyl transfer reaction occurs through a direct SN_2 mechanism with the nucleophilic attack of the substrate (cm^5U) on the electrophilic SAM methyl group to form the methyl ester product (mcm^5U) and SAH (45). Such reaction requires the strict orientation of the incoming nucleophile to be optimal.

The comparison of previously described structures of MTases methylating carboxylic acids (i.e. TYW4 MTase, which is involved in the synthesis of wybutosine in Phe-tRNA (46), glutamate MTase CheR (47) and human LCMT-1 (48), which modifies the C-terminal leucine from PP2A) shows that two structurally conserved residues, an Arg and a Tyr (respectively R73 and Y203 in human LCMT-1), point toward the SAM methyl group and are ideally positioned to orient the substrate carboxylic group through electrostatic interactions (Supplementary Figure S5A). In our structure, the strictly conserved *Yl*Trm9 H127, corresponding to *Sc*Trm9 H116, which is crucial for enzymatic activity, structurally matches with Y203 from LCMT-1 (Supplementary Figure S5B). The H127 side chain forms a hydrogen bond with Y193 (Y242 in *Sc*Trm9) carbonyl group from strand ßB through its Nδ1 atom whereas the Nε2 atom forms a hydrogen bond with a water molecule. This water molecule, also coordinated by H126 (*Sc*Trm9 H115), is 2.75 Å away from the SAM methyl group and is very likely to occupy the position of one of the oxygen atoms from the cm^5U carboxylic group. In our structure, no basic residue exactly matches structurally with R73 from LCMT-1. However, our mutagenesis analysis has revealed that *Sc*Trm9 R29, which is located in the Trm9 region corresponding to LCMT-1 fragment encompassing R73, is important for activity (R29 mutant enzyme exhibits only 5% of the specific activity of the WT enzyme). We then suggest that *Sc*Trm9 R29 could play the same role as R73 from LCMT-1. In *Yl*Trm9N38, R40, corresponding to *Sc*Trm9 R29, is engaged in crystal packing and is oriented in an opposite direction relative to the active site. In our model of the *Yl*Trm9 structure bound to SAM, we have oriented R40 toward the SAM methyl group. This would allow its side chain to interact with the cm^5U carboxyl group and to contribute to its correct positioning for methyl transfer (Supplementary Figure S5B). Both R40 and H127 residues in *Yl*Trm9 would then be ideally positioned to participate

in the catalysis and might play the same role as the R/Y dyad in TYW4, CheR and LCMT-1 active sites, i.e. positioning the substrate carboxylate group (cm^5U in the case of Trm9-Trm112) for the nucleophilic attack onto SAM methyl group.

Furthermore, most of the strongly affected Trm9 mutants studied here are oriented toward the SAM methyl group to be transferred onto the cm^5U and are very likely to directly participate in the optimal orientation of the cm^5U34 nucleotide from the tRNA substrates into the active site. We have then manually docked a cm^5U nucleotide into the *Yl*Trm9-Trm112 active site (Figure 2E). In this model, N222 side chain (N271 in *Sc*Trm9) forms bidentate hydrogen bonds with the O2 and N3 atoms from U34 and then could be responsible for the selective recognition of pyrimidine ring at the wobble position. H126 (H115) forms a hydrogen bond with the O4 atom from U34 allowing Trm9 discriminating between U and C. In addition, a slight reorientation of W156 side chain (W145) into the Trm9 active site would allow it to stack onto the U34 pyrimidine ring. R192 (R241) and Y29 (Y18, not mutated in this study) side chains are also within hydrogen bond distances from the cm^5U carboxylic group. Finally, the W179 indole ring (W168 in *Sc*Trm9) is solvent-exposed and could stack against the ring of a nucleotide adjacent to U34. Last but not least, in this model, the U34 3′ and 5′ positions are oriented outward the active site and hence, this model is consistent with the binding of a tRNA molecule.

We have next performed kinetic analyses on the WT, R29A, H115A, R241A, Y243A and N271A Trm9 mutants to investigate the role of these residues in tRNA modification (Table 2). Overall, compared to WT Trm9-Trm112 complex, all the mutant complexes are strongly affected in their apparent K_m for tRNA (from 4- to 12-fold increase) and in their k_{cat} (up to 100-fold decrease) whereas their K_m for SAM is much less affected (up to 4-fold). The R29A, H115A and Y243A mutants exhibit a strict correlation between the increase of K_m for tRNA and the decrease of k_{cat} which indicates these residues play an equally important role in the binding of the tRNA and in the catalysis whereas their respective K_m for SAM are either unchanged (R29A) or only slightly increased (H115A and Y243A). This clearly demonstrates the involvement of these residues in the binding of the tRNA substrate rather than in the binding of SAM. For the R241A and N271A mutants, the overall effect on K_m is the same as described above (K_m for tRNA significantly more affected than K_m for SAM) with a more dramatic decrease of k_{cat}, which may also account for the catalytic instability of these complexes during the assay. Altogether, the joint effect of mutations observed on the K_m for tRNA and k_{cat} support that the conserved residues are involved in tRNA binding and participate directly in the proper orientation of the cm^5U34 substrate into the active site, which is required for optimal methyl transfer reaction by the SN_2 mechanism.

Comparison with other Trm112-MTase complexes

Trm112 interacts with and activates four MTases that adopt (Trm9, Mtq2 and Bud23) or are predicted to adopt (Trm11) the same fold. To date, we have determined the crystal structures of complexes between Trm112 and three MTase partners (Mtq2, Bud23 and Trm9, (15, 26) and this study). Interestingly, all these 3 MTases interact in a very similar manner with Trm112 as illustrated by the superimposition of the *Yl*Trm9-Trm112 complex onto *Sc*Bud23-Trm112 (rmsd value of 2.1 Å) and *Ec*Mtq2-Trm112 complexes (rmsd value of 3 Å; Figure 3A and B). As expected from phylogenetic distance between these organisms, the Mtq2-Trm112 complex from the *E. cuniculi* parasite is the most divergent among the three complexes, and the two complexes formed by fungal proteins share more similarities. This comparison clearly indicates that the previously proposed competition between these MTases to interact with Trm112 is not due to slightly overlapping binding sites but to binding of these MTases to exactly the same region from Trm112. Although these structures are derived from complexes originating from different organisms, this offers the unique opportunity to compare their binding modes and to understand the molecular mechanisms allowing Trm112 to interact with four structurally similar MTases sharing less than 20% sequence identity within the same organism. We therefore generated a structure-based sequence alignment of *Yl*Trm9, *Ec*Mtq2 and *Sc*Bud23 and extended this alignment by including *Sc*Mtq2, *Sc*Trm9 and *Sc*Trm11 sequences (Figure 3A). From this alignment, we assume that *Sc*Mtq2 and *Sc*Trm9 residues that align with *Ec*Mtq2 and *Yl*Trm9 residues involved in the interaction with *Ec*Trm112 and *Yl*Trm112, respectively, are also contacting *Sc*Trm112. As previous studies have shown that over-expression of *Sc*Trm9 decreases the amount of *Sc*Trm11 immunoprecipitated with *Sc*Trm112 (11), we also assume that Trm11 and Trm112 interact the same way, although no structure of this complex is available yet. Hence, we propose that *Sc*Trm11 residues aligning with interface residues from the other MTases are engaged in the interaction with *Sc*Trm112 (Figure 3A).

The structural comparison between these Trm112-MTase complexes as well as the structure-based alignment of the MTase proteins allow us to identify several common features, which could explain the ability of Trm112 to interact with its four MTase partners. First, a ß-zipper interaction is formed between strand ß4 from Trm112 and strand ß3 from the various MTases (Figure 3B). Such interaction mode implies formation of a hydrogen bond network between main chain atoms from both partners and hence is less altered by side chain variations at these positions of the four MTases. Second, Trm112 shields from the solvent a hydrophobic zone on these three MTases. Although the side chains corresponding to this hydrophobic core in the four MTases from *S. cerevisiae* exhibit some degree of variation (Figure 3C), the hydrophobic character of this core is conserved thereby explaining Trm112 solubilizing effect on most of these MTases. Finally, three electrostatic hot spots are conserved between fungal *Sc*Bud23-Trm112 and *Yl*Trm9-Trm112 crystal structures. The first one involves E100 from *Yl*Trm9 (N89 from *Sc*Trm9) that forms hydrogen bonds with K2 and T5 residues from *Yl*Trm112 (Supplementary Table S1; Figure 3D). *Yl*Trm9 E100 structurally matches with D94 from *Sc*Bud23, which is also engaged in hydrogen bonds with K2 and T5 residues from *Sc*Trm112. In *Sc*Mtq2 and *Sc*Trm11, the corresponding residues are

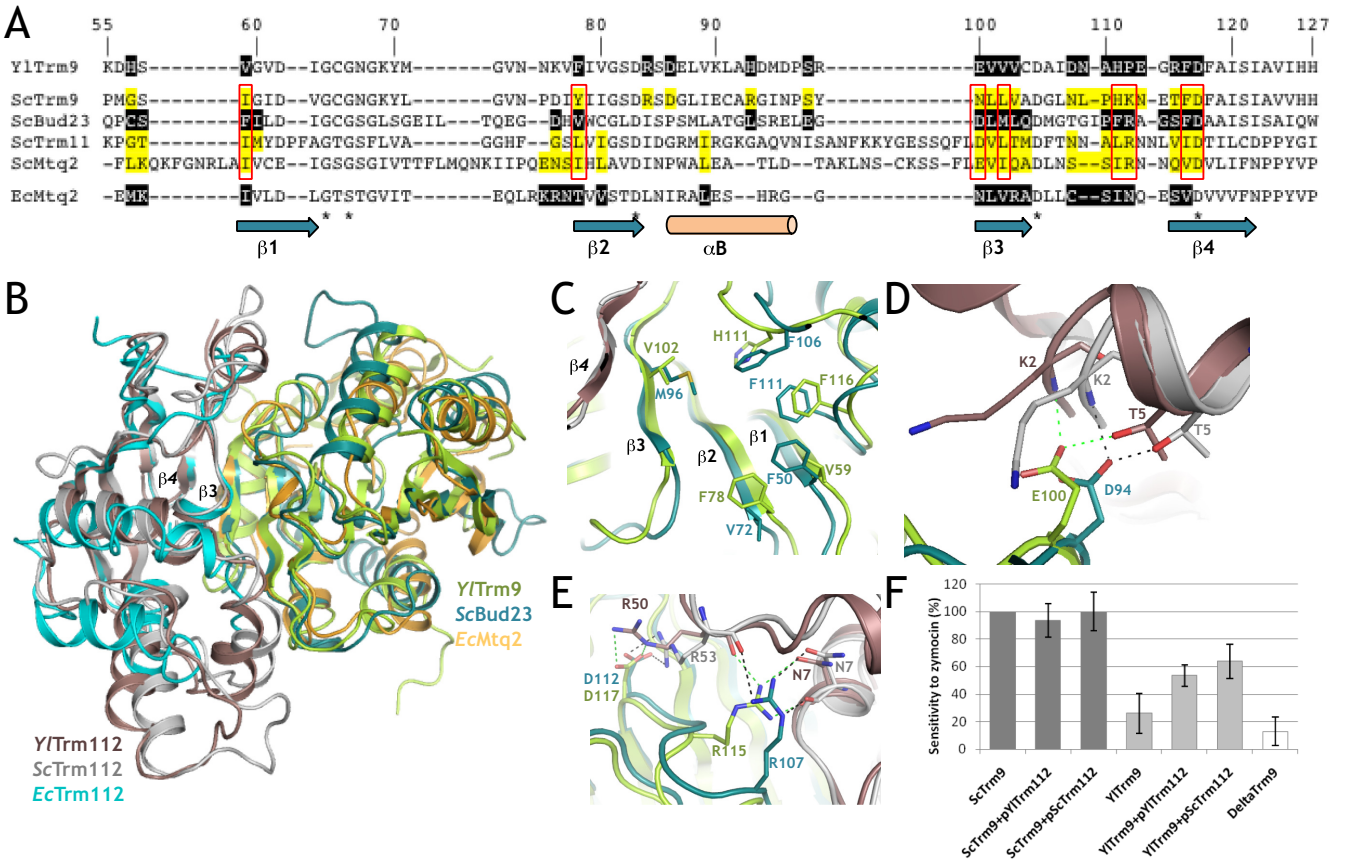


Figure 3. Structural comparison of Trm112-MTase complexes. (**A**) Structure-based sequence alignment of MTases interacting with Trm112. Only regions of these MTases that interact with Trm112 are shown. Residues directly contacting Trm112 in the X-ray structures of the complexes are in white on a black background. Residues from *Sc*Trm9, *Sc*Trm11 and *Sc*Mtq2, matching with interface residues from *Yl*Trm9 and *Ec*Mtq2, are shown in black on a yellow background. *Yl*Trm9 secondary structure elements are depicted below the alignment. Stars below the alignment indicate residues strictly conserved in all 4 *Sc*MTases considered. Residues highlighted in panels C to E are boxed in red. (**B**) Superimposition of the structures of the *Yl*Trm9-Trm112, *Sc*Bud23-Trm112 and *Ec*Mtq2-Trm112 complexes. (**C**–**E**) Detailed comparison of the hydrophobic core (C) and electrostatic hot spots (D and E) involved in *Yl*Trm9-Trm112 and *Sc*Bud23-Trm112 complexes. Same color code as panel B. (**F**) Effect of ectopic plasmid-driven expression of *Yl*Trm112 (+pYlTrm112) and *Sc*Trm112 (+pScTrm112) proteins on the susceptibility to zymocin of *S. cerevisiae* expressing either *Sc*Trm9 (dark gray) or *Yl*Trm9 (light gray) WT protein from the unique genomic copy of the corresponding gene under the control of *ScTRM9* natural promoter.

E101 and D268, respectively. Hence, in *S. cerevisiae*, the side chains found at this position of the four MTases have a similar size and possess a carbonyl group that can be engaged in hydrogen bonds with K2 and T5 from *Sc*Trm112. This is supported by our earlier observation that the substitution of N89 by Lys in *Sc*Trm9 (in combination with L91R mutation) prevents complex formation and therefore inactivates Trm9 (26). The second structurally conserved hot spot is a salt bridge formed between *Yl*Trm9 D117 with *Yl*Trm112 R50 and between *Sc*Bud23 D112 with *Sc*Trm112 R53 (Figure 3E). An Asp residue is conserved at this position in all MTases interacting with *Sc*Trm112 (Figure 3A) and we propose that a salt bridge between this Asp and *Sc*Trm112 R53 occurs in all these Trm112-Mtase complexes. We previously observed that although such interaction did not exist in the structure of the *Ec*Mtq2-Trm112 complex due to discrepancy in the Trm112 central domain containing this Arg residue, the *Sc*Trm112 R53E mutation strongly reduced the solubility and the stability of *Sc*Mtq2 and resulted in complete loss of enzyme activity (26). The third hot spot consists in three hydrogen bonds formed by *Yl*Trm9 R115 side chain with N7 main chain and side chain carbonyl groups and R50 carbonyl group from *Yl*Trm112 (Figure 3E). In our structural alignment, a basic residue is not conserved at the position corresponding to *Yl*Trm9 R115 (T104 in *Sc*Trm9, Q115 in *Sc*Mtq2, V284 in *Sc*Trm11 or S110 in *Sc*Bud23). However, the guanidinium group from *Sc*Bud23 R107 occupies exactly the same position as the guanidinium group from *Yl*Trm9 R115 and forms the same hydrogen bonding network with the N7 (also N7 in *Yl*Trm112) main chain and side chain carbonyl groups and the carbonyl group from *Sc*Trm112 R53 (R50 in *Yl*Trm9). In our alignment, the residues corresponding to *Sc*Bud23 R107 are K101 in *Sc*Trm9, R280 in *Sc*Trm11 and R112 in *Sc*Mtq2, indicating that this interaction network can exist in all these *Sc*Trm112-MTase complexes. Hence, these three crystal structures of Trm112-MTase complexes help us to understand in detail the structural plasticity that allows the small Trm112 protein to interact through the same mechanism with four MTases adopting the same fold but sharing less than 20% sequence identity.

To go deeper into the understanding of Trm112 ability to interact with various MTases, we have tested whether *Yl*Trm9, which shares 59% sequence identity with *Sc*Trm9, could functionally complement for the deletion of *TRM9* gene in *S. cerevisiae* and therefore could form a hybrid complex with *Sc*Trm112. First, we have performed enzymatic assays with the purified full-length *Yl*Trm9-Trm112 complex and tRNAs purified from *S. cerevisiae trm9Δ* strain as substrates. We observe that the *Yl* complex is functional on *S. cerevisiae* tRNAs and exhibits only a 3-fold decrease in apparent specific activity compared to *Sc* complex *in vitro* (Table 2). Furthermore, according to its apparent specific activity, the *Yl* complex should be associated with zymocin sensitivity phenotype *in vivo* (Table 2). Next, we have used an *in vivo* complementation approach in *S. cerevisiae* to characterize the biological activity of full length *Yl*Trm9. Following the replacement of genomic *ScTRM9* gene by *YlTRM9* gene (under the control of the *ScTRM9* natural promoter), the plasmid-driven expression of *Yl*Trm112 renders the strain sensitive to zymocin, indicating that an active *Yl*Trm9-*Yl*Trm112 complex can assemble and methylates *Sc* tRNAs *in vivo* (Figure 3F). In the absence of ectopic *Yl*Trm112 (i.e. in the presence of the sole endogenous *Sc*Trm112), the strain is more resistant to zymocin indicating a defect in *Yl*Trm9 activity. As *Yl*Trm9 is proved to be efficiently expressed and active in *S. cerevisiae*, the latter phenotype observed could account for the lack of interaction between *Yl*Trm9 and *Sc*Trm112. To test whether the likely competition between *Yl*Trm9 and *Sc* natural partners (Bud23, Mtq2 and Trm11) for *Sc*Trm112 could be responsible for this phenotype, we considered the potential effect of increasing *Sc*Trm112 expression level on cell phenotype toward zymocin. Plasmid-driven expression of *Sc*Trm112 resulted in zymocin sensitivity level comparable to the one previously obtained with *Yl*Trm112 (Figure 3F). This indicates that *Sc*Trm112 (50% sequence identity with *Yl*Trm112) is able to activate *Yl*Trm9 *in vivo* to a similar extent as *Yl*Trm112. This most likely occurs through a direct interaction between *Sc*Trm112 and *Yl*Trm9 (59% and 79% sequence identity and similarity with *Sc*Trm9, respectively). It should be noted that this chimeric *Yl*Trm9-*Sc*Trm112 complex is probably less stable than the corresponding complex formed by proteins from the same organism as it is necessary to increase expression level of *Sc*Trm112 to enhance zymocin sensitivity. This is further supported by preliminary co-IP experiments showing that no *Yl*Trm9-*Sc*Trm112 interaction could be characterized using a routine protocol suitable for measuring *Sc*Trm9-*Sc*Trm112 interaction.

Altogether, these results show that MTase-Trm112 binding mode is compatible with the formation of chimeric complexes between proteins from different organisms as illustrated by the ability of *Sc*Trm112 to activate *Yl*Trm9, which shares about 60% sequence identity with *Sc*Trm9. This is further supported by the ability of ABH8 (and hTrm9L) and WBSCR22 (Bud23 orthologue) human genes to partially complement the deletion of *TRM9* and *BUD23* genes in *S. cerevisiae*, respectively (24).

CONCLUSION

The structure of the Trm9-Trm112 tRNA MTase holoenzyme combined with functional studies has allowed the mapping of its active site and enabled us to propose a model of the catalytic mechanism and binding mode of the cm^5U moiety of the tRNA substrate. In addition, the detailed comparison of this structure to those of *Sc*Bud23-Trm112 and *Ec*Mtq2-Trm112 complexes highlights that during evolution, sequences of all these proteins within the same organism have evolved in a concerted manner so as to maintain the Trm112 ability to interact with its MTase partners that share less than 20% sequence identity.

ACCESSION NUMBERS

The atomic coordinates and structure factors have been deposited into the Brookhaven Protein Data Bank under accession numbers (5CM2).

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We are indebted to M. Argentini and D. Cornu for mass spectrometry analysis (SICaps, IMAGIF Platform, Gif/Yvette, France), to Dr V. Heurgué-Hamard for her help with MTase assay and for sharing with us plasmids and antibodies directed against *S. cerevisiae* Trm9-Trm112 complex. We thank Dr G. Bourgeois for helpful discussion and advices. We thank Dr B. Séraphin, Dr K. Breunig and Dr R. Schaffrath for sharing reagents with us. We acknowledge SOLEIL for provision of synchrotron radiation facilities.

FUNDING

This work was supported by the Centre National pour la Recherche Scientifique (CNRS) including a specific support by the ATIP-AVENIR program [to M.G.], the European Union Sixth Framework program '3D-Repertoire' [LSHG-CT-2005-512028], the Agence Nationale pour la Recherché (ANR; ANR14-CE09-0016-02), the University Paris-Sud and Ecole Polytechnique. J.L. and T.V.N. hold a PhD fellowship from the French Ministère de l'Enseignement Supérieur et de la Recherche (MESR). Funding for open access charge: This work was supported by the Centre National pour la Recherche Scientifique (CNRS) including a specific support by the ATIP-AVENIR program [to M.G.], the European Union Sixth Framework program '3D-Repertoire' [LSHG-CT-2005-512028], the Agence Nationale pour la Recherche (ANR, ANR14-CE09-0016-02), the University Paris-Sud and Ecole Polytechnique. J.L. and T.V.N. hold a PhD fellowship from the French Ministère de l'Enseignement Supérieur et de la Recherche (MESR).
Conflict of interest statement. None declared.

REFERENCES

1. Cantara,W.A., Crain,P.F., Rozenski,J., McCloskey,J.A., Harris,K.A., Zhang,X., Vendeix,F.A., Fabris,D. and Agris,P.F. (2011) The RNA Modification Database, RNAMDB: 2011 update. *Nucleic Acids Res.*, **39**, D195–D201.

2. Hopper,A.K. and Phizicky,E.M. (2003) tRNA transfers to the limelight. *Genes Dev.*, **17**, 162–180.
3. Johansson,M.J.O. and Bystrom,A.S. (2005) In: Grosjean,H (ed). *Fine-Tuning of RNA Functions by Modification and Editing*. Springer-Verlag, Heidelberg, pp. 87–119.
4. Johansson,M.J., Esberg,A., Huang,B., Bjork,G.R. and Bystrom,A.S. (2008) Eukaryotic wobble uridine modifications promote a functionally redundant decoding system. *Mol. Cell. Biol.*, **28**, 3301–3312.
5. Huang,B., Johansson,M.J. and Bystrom,A.S. (2005) An early step in wobble uridine tRNA modification requires the Elongator complex. *RNA*, **11**, 424–436.
6. Huang,B., Lu,J. and Bystrom,A.S. (2008) A genome-wide screen identifies genes required for formation of the wobble nucleoside 5-methoxycarbonylmethyl-2-thiouridine in Saccharomyces cerevisiae. *RNA*, **14**, 2183–2194.
7. Frohloff,F., Fichtner,L., Jablonowski,D., Breunig,K.D. and Schaffrath,R. (2001) Saccharomyces cerevisiae Elongator mutations confer resistance to the Kluyveromyces lactis zymocin. *EMBO J.*, **20**, 1993–2003.
8. Glatt,S., Letoquart,J., Faux,C., Taylor,N.M., Seraphin,B. and Muller,C.W. (2012) The Elongator subcomplex Elp456 is a hexameric RecA-like ATPase. *Nat. Struct. Mol. Biol.*, **19**, 314–320.
9. Fichtner,L. and Schaffrath,R. (2002) KTI11 and KTI13, Saccharomyces cerevisiae genes controlling sensitivity to G1 arrest induced by Kluyveromyces lactis zymocin. *Mol. Microbiol.*, **44**, 865–875.
10. Kalhor,H.R. and Clarke,S. (2003) Novel methyltransferase for modified uridine residues at the wobble position of tRNA. *Mol. Cell. Biol.*, **23**, 9283–9292.
11. Studte,P., Zink,S., Jablonowski,D., Bar,C., von der Haar,T., Tuite,M.F. and Schaffrath,R. (2008) tRNA and protein methylase complexes mediate zymocin toxicity in yeast. *Mol. Microbiol.*, **69**, 1266–1277.
12. Purushothaman,S.K., Bujnicki,J.M., Grosjean,H. and Lapeyre,B. (2005) Trm11p and Trm112p are both required for the formation of 2-methylguanosine at position 10 in yeast tRNA. *Mol. Cell. Biol.*, **25**, 4359–4370.
13. Heurgue-Hamard,V., Graille,M., Scrima,N., Ulryck,N., Champ,S., van Tilbeurgh,H. and Buckingham,R.H. (2006) The zinc finger protein Ynr046w is plurifunctional and a component of the eRF1 methyltransferase in yeast. *J. Biol. Chem.*, **281**, 36140–36148.
14. White,J., Li,Z., Sardana,R., Bujnicki,J.M., Marcotte,E.M. and Johnson,A.W. (2008) Bud23 methylates G1575 of 18S rRNA and is required for efficient nuclear export of pre-40S subunits. *Mol. Cell. Biol.*, **28**, 3151–3161.
15. Letoquart,J., Huvelle,E., Wacheul,L., Bourgeois,G., Zorbas,C., Graille,M., Heurgue-Hamard,V. and Lafontaine,D.L. (2014) Structural and functional studies of Bud23-Trm112 reveal 18S rRNA N7-G1575 methylation occurs on late 40S precursor ribosomes. *Proc. Natl. Acad. Sci. U.S.A.*, **111**, E5518–E5526.
16. Sardana,R. and Johnson,A.W. (2012) The methyltransferase adaptor protein Trm112 is involved in biogenesis of both ribosomal subunits. *Mol. Biol. Cell*, **23**, 4313–4322.
17. Patil,A., Dyavaiah,M., Joseph,F., Rooney,J.P., Chan,C.T., Dedon,P.C. and Begley,T.J. (2012) Increased tRNA modification and gene-specific codon usage regulate cell cycle progression during the DNA damage response. *Cell Cycle*, **11**, 3656–3665.
18. Begley,U., Dyavaiah,M., Patil,A., Rooney,J.P., DiRenzo,D., Young,C.M., Conklin,D.S., Zitomer,R.S. and Begley,T.J. (2007) Trm9-catalyzed tRNA modifications link translation to the DNA damage response. *Mol. Cell*, **28**, 860–870.
19. Patil,A., Chan,C.T., Dyavaiah,M., Rooney,J.P., Dedon,P.C. and Begley,T.J. (2012) Translational infidelity-induced protein stress results from a deficiency in Trm9-catalyzed tRNA modifications. *RNA Biol.*, **9**, 990–1001.
20. Fu,D., Brophy,J.A., Chan,C.T., Atmore,K.A., Begley,U., Paules,R.S., Dedon,P.C., Begley,T.J. and Samson,L.D. (2010) Human AlkB homolog ABH8 Is a tRNA methyltransferase required for wobble uridine modification and DNA damage survival. *Mol. Cell. Biol.*, **30**, 2449–2459.
21. Fu,Y., Dai,Q., Zhang,W., Ren,J., Pan,T. and He,C. (2010) The AlkB domain of mammalian ABH8 catalyzes hydroxylation of 5-methoxycarbonylmethyluridine at the wobble position of tRNA. *Angew. Chem. Int. Ed. Engl.*, **49**, 8885–8888.
22. Songe-Moller,L., van den Born,E., Leihne,V., Vagbo,C.B., Kristoffersen,T., Krokan,H.E., Kirpekar,F., Falnes,P.O. and Klungland,A. (2010) Mammalian ALKBH8 possesses tRNA methyltransferase activity required for the biogenesis of multiple wobble uridine modifications implicated in translational decoding. *Mol. Cell. Biol.*, **30**, 1814–1827.
23. Shimada,K., Nakamura,M., Anai,S., De Velasco,M., Tanaka,M., Tsujikawa,K., Ouji,Y. and Konishi,N. (2009) A novel human AlkB homologue, ALKBH8, contributes to human bladder cancer progression. *Cancer Res.*, **69**, 3157–3164.
24. Begley,U., Sosa,M.S., Avivar-Valderas,A., Patil,A., Endres,L., Estrada,Y., Chan,C.T., Su,D., Dedon,P.C., Aguirre-Ghiso,J.A. *et al.* (2013) A human tRNA methyltransferase 9-like protein prevents tumour growth by regulating LIN9 and HIF1-alpha. *EMBO Mol. Med.*, **5**, 366–383.
25. Wolf,K., Breunig,K. and Barth,G. (2003) *Non-conventional yeasts in genetics, biochemistry and biotechnology : practical protocols*. Springer, Berlin; NY.
26. Liger,D., Mora,L., Lazar,N., Figaro,S., Henri,J., Scrima,N., Buckingham,R.H., van Tilbeurgh,H., Heurgue-Hamard,V. and Graille,M. (2011) Mechanism of activation of methyltransferases involved in translation by the Trm112 'hub' protein. *Nucleic Acids Res.*, **39**, 6249–6259.
27. Longtine,M.S., McKenzie,A. 3rd, Demarini,D.J., Shah,N.G., Wach,A., Brachat,A., Philippsen,P. and Pringle,J.R. (1998) Additional modules for versatile and economical PCR-based gene deletion and modification in Saccharomyces cerevisiae. *Yeast*, **14**, 953–961.
28. Toulmay,A. and Schneiter,R. (2006) A two-step method for the introduction of single or multiple defined point mutations into the genome of Saccharomyces cerevisiae. *Yeast*, **23**, 825–831.
29. Studier,F.W. (2005) Protein production by auto-induction in high density shaking cultures. *Protein Expr. Purif.*, **41**, 207–234.
30. Kabsch,W. (1993) Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. Appl. Cryst.*, **26**, 795–800.
31. Adams,P.D., Grosse-Kunstleve,R.W., Hung,L.W., Ioerger,T.R., McCoy,A.J., Moriarty,N.W., Read,R.J., Sacchettini,J.C., Sauter,N.K. and Terwilliger,T.C. (2002) PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr. D. Biol. Crystallogr.*, **58**, 1948–1954.
32. Bricogne,G., Vonrhein,C., Flensburg,C., Schiltz,M. and Paciorek,W. (2003) Generation, representation and flow of phase information in structure determination: recent developments in and around SHARP 2.0. *Acta Crystallogr. D Biol. Crystallogr.*, **59**, 2023–2030.
33. Vagin,A. and Teplyakov,A. (1997) MOLREP: an automated program for molecular replacement. *J. Appl. Cryst.*, **30**, 1022–1025.
34. Emsley,P. and Cowtan,K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.*, **60**, 2126–2132.
35. Laskowski,R.A., MacArthur,M.W., Moss,D.S. and Thornton,J.M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.*, **26**, 283–291.
36. Chen,C., Huang,B., Anderson,J.T. and Bystrom,A.S. (2011) Unexpected accumulation of ncm(5)U and ncm(5)S(2) (U) in a trm9 mutant suggests an additional step in the synthesis of mcm(5)U and mcm(5)S(2)U. *PLoS One*, **6**, e20783.
37. Avital,S. and Elson,D. (1969) A convenient procedure for preparing transfer ribonucleic acid from Escherichia coli. *Biochim. Biophys. Acta*, **179**, 297–307.
38. Cao,W. and De La Cruz,E.M. (2013) Quantitative full time course analysis of nonlinear enzyme cycling kinetics. *Sci. Rep.*, **3**, 2658.
39. Krivov,G.G., Shapovalov,M.V. and Dunbrack,R.L. Jr (2009) Improved prediction of protein side-chain conformations with SCWRL4. *Proteins*, **77**, 778–795.
40. MacKerell,A.D., Bashford,D., Bellott, Dunbrack,R.L., Evanseck,J.D., Field,M.J., Fischer,S., Gao,J., Guo,H., Ha,S. *et al.* (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B*, **102**, 3586–3616.
41. Brooks,B.R., Brooks,C.L. 3rd, Mackerell,A.D. Jr, Nilsson,L., Petrella,R.J., Roux,B., Won,Y., Archontis,G., Bartels,C., Boresch,S.

et al. (2009) CHARMM: the biomolecular simulation program. *J. Comput. Chem.*, **30**, 1545–1614.
42. Grosjean,H., Droogmans,L., Roovers,M. and Keith,G. (2007) Detection of enzymatic activity of transfer RNA modification enzymes using radiolabeled tRNA substrates. *Methods Enzymol.*, **425**, 55–101.
43. Sturchler,C., Lescure,A., Keith,G., Carbon,P. and Krol,A. (1994) Base modification pattern at the wobble position of Xenopus selenocysteine tRNA(Sec). *Nucleic Acids Res.*, **22**, 1354–1358.
44. Mazauric,M.H., Dirick,L., Purushothaman,S.K., Bjork,G.R. and Lapeyre,B. (2010) Trm112p is a 15-kDa zinc finger protein essential for the activity of two tRNA and one protein methyltransferases in yeast. *J. Biol. Chem.*, **285**, 18505–18515.
45. Schubert,H.L., Blumenthal,R.M. and Cheng,X. (2003) Many paths to methyltransfer: a chronicle of convergence. *Trends Biochem. Sci.*, **28**, 329–335.
46. Suzuki,Y., Noma,A., Suzuki,T., Ishitani,R. and Nureki,O. (2009) Structural basis of tRNA modification with CO2 fixation and methylation by wybutosine synthesizing enzyme TYW4. *Nucleic Acids Res.*, **37**, 2910–2925.
47. Djordjevic,S. and Stock,A.M. (1997) Crystal structure of the chemotaxis receptor methyltransferase CheR suggests a conserved structural motif for binding S-adenosylmethionine. *Structure*, **5**, 545–558.
48. Stanevich,V., Jiang,L., Satyshur,K.A., Li,Y., Jeffrey,P.D., Li,Z., Menden,P., Semmelhack,M.F. and Xing,Y. (2011) The structural basis for tight control of PP2A methylation and function by LCMT-1. *Mol. Cell*, **41**, 331–342.
49. Robert,X. and Gouet,P. (2014) Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.*, **42**, W320–W324.
50. Ashkenazy,H., Erez,E., Martz,E., Pupko,T. and Ben-Tal,N. (2010) ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. *Nucleic Acids Res.*, **38**, W529–W533.

Supplementary files

Insights into molecular plasticity in protein complexes from Trm9-Trm112 tRNA modifying enzyme crystal structure

Juliette Létoquart[1,2], Nhan van Tran[1], Vonny Caroline[1], Alexey Aleksandrov[1], Noureddine Lazar[2], Herman van Tilbeurgh[2], Dominique Liger[2,*] and Marc Graille[1,2,*]

[1] Laboratoire de Biochimie, CNRS, UMR 7654, Ecole Polytechnique, F-91128 Palaiseau Cedex, France

[2] Fonction et Architecture des Assemblages Macromoléculaires, Département B3S, Institut de Biologie Intégrative de la Cellule (I2BC), CNRS,UMR 9198, CEA, Université Paris Sud, F-91405 Orsay Cedex, France Corresponding authors:
Dominique Liger. Tel: +33 (0)169157968. Email: dominique.liger@u-psud.fr
Marc Graille. Tel: +33 (0)169334890. Email: marc.graille@polytechnique.edu

Table S1: Details of the electrostatic interactions involved in *Yl*Trm9N38-Trm112 and *Sc*Bud23-Trm112 interactions.

***Yl*Trm9**	***Yl*Trm112**	***Sc*Bud23**	***Sc*Trm112**
Hydrogen bonds		***Hydrogen bonds***	
		Cys48 O	Asn49 Nd2
Glu100 Oe1	Lys2 N	Asp94 Od2	Lys2 N
Glu100 Oe1	Thr5 Og1	Asp94 Od2	Thr5 Og1
Val103 N	Pro120 O	Leu97 N	Pro126 O
Val103 0	Phe122 N	Leu97 O	Leu128 N
		Arg107 N	Phe8 O
Arg115 Nh2	Asn7 Od1	Arg107 Nh2	Asn7 Od1
Arg115 Nh1	Asn7 O	Arg107 Ne	Asn7 O
Arg115 Nh2	Arg50 O	Arg107 Nh2	Arg53 O
		Ser110 O	Arg53 Nh1
Ala110 O	Gln10 N	Pro105 O	Lys10 N
Glu113 Oe2	Gln10 Ne2		
		Glu92 O	Lys2 Nz
Salt bridges		***Salt bridges***	
		Arg139	Asp17
Asp117	Arg50	Asp112	Arg53

Interactions formed between structurally matching residues from both complexes are shown on the same line.

Table S2: Yeast strains

Strains	Genotype	Source
Kluyveromyces lactis		
AWJ137	*leu2 trp1*[k1+k2+]	*K. Breunig (University of Halle, Germany)*
NK40	*ade1 ade2 leu2* [k2+]	*R. Schaffrath (University of Leicester, UK)*
Saccharomyces cerevisiae		
YPH499	MAT**a** *ura3-52 lys2-801 amber ade2-101 ochre trp1-Δ63 his1-Δ200 leu2-Δ1*	*Agilent Technologies*
YDL201	YPH499 *TRM9::kanMX6*	*This study*
YDL202	YPH499 *TRM9(D72A)::kanMX6*	*This study*
YDL203	YPH499 *TRM9(N89K/L91R)::kanMX6*	*This study*
YDL204	YPH499 *TRM9(F105E)::kanMX6*	*This study*
yMG42	YPH499 *TRM9(H115A)::kanMX6*	*This study*
yMG43	YPH499 *TRM9(H116A)::kanMX6*	*This study*
yMG44	YPH499 *TRM9(R29A)::kanMX6*	*This study*
YDL205	YPH499 *TRM9(R241A)::kanMX6*	*This study*
YDL301	YPH499 *TRM9(W145A)13myc::kanMX6 TRM112-3HA::TRP1*	*This study*
YDL302	YPH499 *TRM9(Q149A)13myc::kanMX6 TRM112-3HA::TRP1*	*This study*
YDL303	YPH499 *TRM9(W168A)13myc::kanMX6 TRM112-3HA::TRP1*	*This study*
YDL304	YPH499 *TRM9(Y243A)13myc::kanMX6 TRM112-3HA::TRP1*	*This study*
YDL305	YPH499 *TRM9(N271A)13myc::kanMX6 TRM112-3HA::TRP1*	*This study*
yMG45	YPH499 *trm9Δ::URA3*	*This study*
yMG46	YPH499 *trm9Δ::kanMX6*	*This study*
yMG47	YPH499 *trm9Δ::YlTRM9*	*This study*
yMG48	YPH499 *trm9Δ::YlTRM9DN37*	*This study*

yMG5	YPH499 + empty pESC-URA	*This study*
yMG41	yMG46 + empty pESC-URA	*This study*
yMG6	yMG47 + empty pESC-URA	*This study*
yMG8	YPH499 + pMG571	*This study*
yMG40	yMG46 + pMG571	*This study*
yMG1	yMG47 + pMG571	*This study*
yMG37	YPH499 + pU-F*Sc*112	*This study*
yMG38	yMG46 + pU-F*Sc*112	*This study*
yMG39	yMG47 + pU-F*Sc*112	*This study*
yMG24	YPH499 *TRM9(H115A)-11Myc::TRP1*	*This study*
yMG25	YPH499 *TRM9(H116A) -13Myc::TRP1*	*This study*
yMG23	YPH499 *TRM9(R29A) -13Myc::TRP1*	*This study*
yMG29	YPH499 *TRM9(R241A) -11Myc::TRP1*	*This study*
BSY2534	BMA64 *(MATa ade 2-1 his3-11,15 leu2-3,112 trp1Δ ura3-1 can1-100) elp1Δ::kanMX6*	*From B. Séraphin*

Table S3: Primers for yeast constructions

Oligonucleotide	Sequence
F2 Trm9 (13myctagging)	5’CTTCGAGCGCGACAATTGGTGGGTGGTGGCCCAGAAGAGA CGGATCCCCGGGTTAATTAA 3’
F2-YlTrm9(13myctagging) oMG220	5’TAATTGGTGGGTGGTAGCCAAGAGAGGTGACGATTGGAGT CGGATCCCCGGGTTAATTAA 3’
F3 Trm9	5’CTTCGAGCGCGACAATTGGTGGGTGGTGGCCCAGAAGAGA TGAGGCGCGCCACTTCTAAA 3’
Ftrm9D72A	5'GGAGTGAACCCTGATATATATATTATCGGTTCA**GCG**CGCTCA GATGGTCTTATTGAGTGC 3'
Ftrm9N89K+L91R	5'GCCAGAGGAATAAACCCATCGTAT**AAA**TTA**CGG**GTGGCAGA CGGGCTGAACTTACCACAC 3'
Ftrm9F105E	5'CTGAACTTACCACACAAAAACGAAACA**GAA**GACTTCGCCA TCTCAATTGCTGTAGTGCAT3'
Ftrm9 R29A	5’GTGTATAATGAGATAGCTCCGCATTTCTCGCAAACT**GCA**TAT AAGCCATGGCCCATAGTG 3’
Ftrm9 R241A	5’GATCAAGAACAGGAAAGAGAAGAGGTAAAATAC**GCG**TACT ATCACTTATACCGAGAGGGC 3’
Ftrm9 H115A	5’AACATTTGACTTCGCCATCTCAATTGCTGTAGTG**GCT**CACT GGTCTACAAGGGAGAGACG 3’
Ftrm9 H116A	5’AACATTTGACTTCGCCATCTCAATTGCTGTAGTGCAT**GCC**TG GTCTACAAGGGAGAGACG 3’
oMG197 (Ftrm9 W145A)	5’AAGCTACGTCAGGGCGGACAAGCATTAATATATTGT**GCC**GC TCTAGAACA GGGCAGCTCC 3’
oMG199 (Ftrm9 Q149A)	5’GCGGACAAGCATTAATATATTGTTGGGCTCTAGAA**GCC**GGC AGCTCCCGT AGAGGTTACC 3’
oMG198 (Ftrm9 W168A)	5’TACCATGAAGGTATGGAGCAAGATGTCTTTGTCCCC**GCG**GT TCTTCCCAA GAGTAAATCC 3’
oMG200 (Trm9 N271A)	5’GGCGCTGCCGTTCATAGTGAGGGCTTCGAGCGCGAC**GCT**T GGTGGGTGGTGGCCCAGAAG 3’
oMG201 (Trm9 Y243A)	5’ATCAAGAACAGGAAAGAGAAGAGGTAAAATACCGgTAC**GC T**CACTTATACCGAGAGGGCG 3’
R1-trm9	5'CCTGCTGCTACAAAATACACTGTCTACCTATATATCACCTGAA TTCGAGCTCGTTTAAAC3'
Nucleotides corresponding to introduced mutations are underlined and in bold	

F1-URA-trm9 (deletion)	5’AGGTCTCGAAGAGCCAAGAAATAAAAGGTTAAGAACCAAC TAACTATGCGGCATCAGAGC 3’
R1-URA-trm9 (deletion)	5’CCTGCTGCTACAAAATACACTGTCTACCTATATATCACCTCCT GATGCGGTATTTTCTCC 3’
F1-URA-trm112 (deletion)	5’TCTCTTCGGCTCTACACATCATATTACTAGCCTAGTCAACTA ACTATGCGGCATCAGAGC 3’
R1-URA-trm112 (deletion)	5’TTTTCGTCTT GCGTGCCCACACACAGAGATCTCGCTTGAT CCTGATGCGGTATTTTCTCC 3’
Ftrm9::YLTRM9	5’ AGGTCTCGAA GAGCCAAGAAATAAAAGGTT AAGAACCAACATGGTGGTCGAATCGGTGCCC 3’
Ftrm9::YLTRM9DN38	5’ AGGTCTCGAA GAGCCAAGAA ATAAAAGGTT AAGAACCAACATG AACACGCGAT ACAAGCCGTG 3’
Rtrm9::YLTRM9	5’CCTGCTGCTACAAAATACACTGTCTACCTATATATCACCTTC AACTCCAATCGTCACCTC 3’
F1Trm9 (deletion)	5’AGGTCTCGAAGAGCCAAGAAATAAAAGGTTAAGAACCAAC CGGATCCCCGGGTTAATTAA 3’

Table S4: Primers for plasmid constructions

Oligonucleotide	Sequence
oMG5- F-BamH1-Trm112Yl (pESC-URA cloning)	5'GGGGGATCCATGAAATTCCTGACCTCAAACTTCG 3'
oMG6-R-NheI-Trm112Yl (pESC-URA cloning)	5' GGGGCTAGCTTAGGCAGCCAGGTGCGGCGGC 3'
F-NcoIYlTrm9N18	5'GGGCCATGGATAATGAACAGCAGCACGTCC 3'
F-NcoIYlTrm9N37	5'GGGCCATGGCAAATACCCGCTATAAACCGTGG 3'
R-XhoIYlTrm9	5'GGGCTCGAGTCAATGGTGATGGTGATGGTG 3'
F-NdeITrm9	5'GCGCATATGGAGATAAACCAAGCGGCTGA 3'
R-NotITrm9his	5'CTAGCGGCCGCTCAGTGATGGTGATGGTGATGTCTCTTCT GGGCC 3'
oMG196 (N271A Reverse)	5'ACCCACCAA**GC**GTCGCGCTCGAAGCCCTCA 3'
oMG195 (N271A Forward)	5' GAGCGCGAC**GC**TTGGTGGGTGGTGGCCCAG 3'
oMG194 (Y243A Reverse)	5' TATAAGTGA**GC**GTATCGGTATTTTACCTCT 3'
oMG193 (Y243A Forward)	5' TACCGATAC**GC**TCACTTATACCGAGAGGGC 3'
oMG192 (W168A Reverse)	5' GGAAGAACC**GC**GGGGACAAAGACATCTTGC 3'
oMG191 (W168A Forward)	5' TTTGTCCCC**GC**GGTTCTTCCCAAGAGTAAA 3'
oMG190 (Q149A Reverse)	5' GAGCTGCCC**GC**TTCTAGAGCCCAACAATAT 3'
oMG189 (Q149A Forward)	5' GCTCTAGAA**GC**GGGCAGCTCCCGTAGAGGT 3'
oMG188 (W145A Reverse)	5' TCTAGAGCC**GC**ACAATATATTAATGCTTGT 3'
oMG187 (W145A Forward)	5' ATATATTGT**GC**GGCTCTAGAACAGGGCAGC 3'

Nucleotides corresponding to introduced mutations are underlined and in bold

Table S5: plasmids

Plasmid name	Background	Gene cloned
pMG546	pET21a	*YlTRM112-TRM9-His$_6$*
pMG547	pET21a	*YlTRM112-TRM9N19-His$_6$*
pMG548	pET21a	*YlTRM112-TRM9N38-His$_6$*
pMG513	pACYCDUET-1	*ScTRM112*
pMG540	pET28a	*ScTRM9*
pMG576	pET21a	*ScTRM9-His$_6$* R241A mutant
pMG577	pET21a	*ScTRM9-His$_6$* R29A mutant
pMG579	pET21a	*ScTRM9-His$_6$* D72A mutant
pMG580	pET21a	*ScTRM9-His$_6$* H115A mutant
pMG581	pET21a	*ScTRM9-His$_6$* H116A mutant
pMG625	pET28a	*ScTRM9-His$_6$* W145A mutant
pMG626	pET28a	*ScTRM9-His$_6$* Q149A mutant
pMG627	pET28a	*ScTRM9-His$_6$* W168A mutant
pMG628	pET28a	*ScTRM9-His$_6$* Y243A mutant
pMG629	pET28a	*ScTRM9-His$_6$* N271A mutant
pMG571	pESC-URA	*YlTRM112* under P_{GAL1} control
pU-F*Sc*112	pESC-URA	*FLAG-ScTRM112* under P_{GAL10} control

Legends to supplementary figures.

Figure S1: The truncated forms of YlTrm9-Trm112 bind SAM but are inactive.

A. ITC measurements. The YlTrm9-Trm112 full-length (left), YlTrm9N19-Trm112 (middle) and YlTrm9N38-Trm112 (right) complexes at 100µM were titrated by 20 successive injections of 2µL SAM (1.2mM) at 20°C using a Microcal ITC200 machine. The stoechiometry (N), affinity constant (K), ΔH and ΔS values obtained with the ORIGIN program using a one binding site model are indicated in the inset. Kd values were calculated using the Kd=1/K relationship.

B. Activity measurements. The curves obtained after fitting of the experimental data with equation given in the Materials and Methods section are shown by lines using the same color code as for the symbols. The results obtained for *S. cerevisiae* wild-type and R241A mutant Trm9-Trm112 complexes are shown for comparison.

Figure S2: Electron density maps.

A. Experimental electron density map (contoured at 1 sigma) obtained after Zn-MAD phasing at 3 Å resolution. The final model is shown as sticks with Trm9 in green and Trm112 in pink.

B. Final 2Fo-Fc electron density map (contoured at 1 sigma) obtained at 2.5 Å resolution and covering the same region as panel A.

Figure S3: Superimposition of *Yl*Trm9-Trm112 and RPA2492 Xray structures to model Trm9 N-terminal helix.

A. Superimposition of *R. palustris* RPA2492 (deep salmon) onto *Yl*Trm9-Trm112 (in green and pink respectively). The modeled N-terminal helix of *Yl*Trm9 is colored in light green.

B. Superimposition of RPA2492 (deep salmon) onto our model of *Yl*Trm9-Trm112 (green). *Yl*Trm9 numbering is used. The SAM from RPA2492 structure is shown in blue sticks.

Figure S4: pH dependence of Sc Trm9-Trm112 enzymatic activity.

A. Detection of mcm^5U-modified nucleotide using 2D-TLC. Upper part shows chromatographic mobility in solvent system A+B, lower part in solvent system A+C. Left panels: reference maps (adapted from (1,2)) showing the theoretical position of mcm^5U nucleoside in both solvent systems. Central and right panels: ^{14}C-methyl group incorporated in modified nucleotides was detected by phosphor-imaging, and the positions of unmodified nucleotides were superimposed by aligning the spots visualized under UV. In the central panels, the negative control consists in the

analysis of nucleotides from tRNAs from Δ*elp1* strain while in the right panels, tRNAs used were purified from Δ*trm9* strain. Red arrows indicate the position of the labeled mcm^5U.

B.pH dependence of Trm9-Trm112 enzymatic activity. The curves obtained after fitting of the experimental data with equation given in the Materials and Methods section are shown by lines using the same color code as for the symbols.

C.Table summarizing the apparent specific activities derived from the fitting of the curves shown in panel B. Apparent specific activity (fmol of tRNA methylated / min / pmol of enzyme) calculated from apparent initial velocity. The kinetics were performed with 1.5 pmol of enzyme.

Figure S5: Comparison of structures of SAM-dependent MTases modifying carboxylic functions.

A. Superimposition of LCMT-1 (purple), CheR (grey) and TYW4 (pink). The SAM analog used to trap the C subunit of human PP2A (sand) in LCMT-1 active site is shown in cyan sticks. For clarity, only R73 and Y203 from LCMT-1 are labeled.

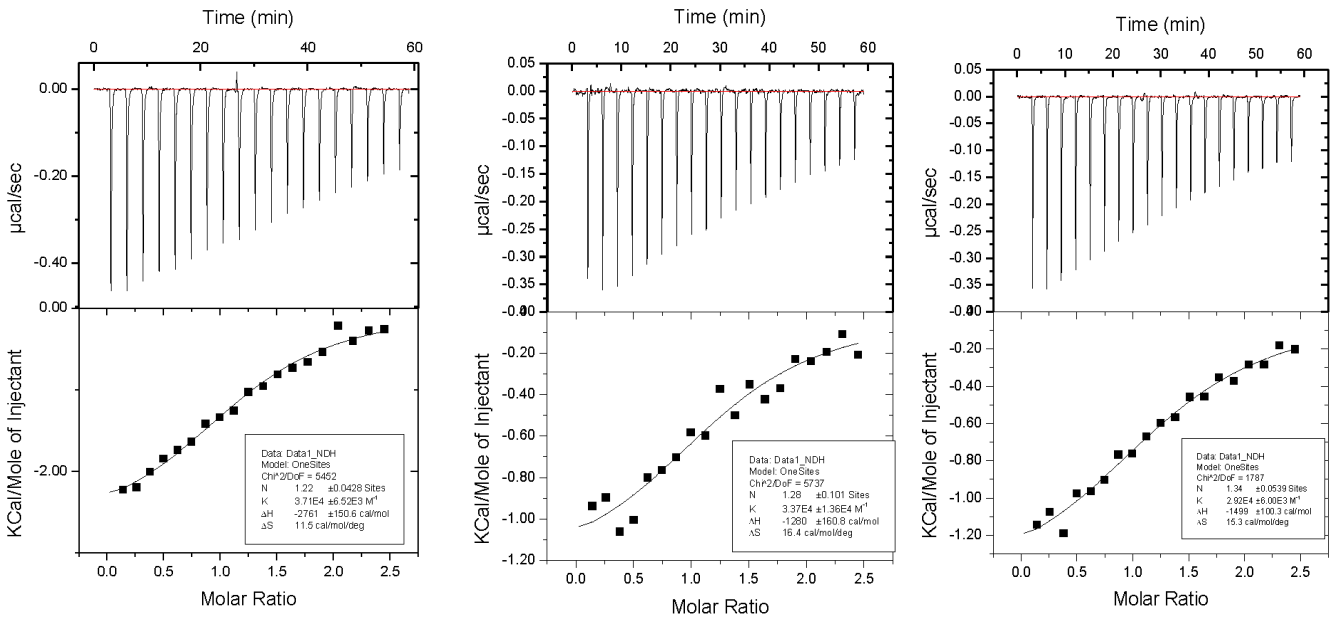
B. Superimposition of LCMT-1 (purple) bound to PP2A (sand) onto our model of *Yl*Trm9-Trm112 (green). *Yl*Trm9 numbering is used.

Bibliographie

1. Sturchler, C., Lescure, A., Keith, G., Carbon, P. and Krol, A. (1994) Base modification pattern at the wobble position of Xenopus selenocysteine tRNA(Sec). Nucleic acids research, 22, 1354-1358.

2. Grosjean, H., Droogmans, L., Roovers, M. and Keith, G. (2007) Detection of enzymatic activity of transfer RNA modification enzymes using radiolabeled tRNA substrates. Methods in enzymology, 425, 55-101.

A.


YlTrm9-Trm112
YlTrm9N19-Trm112
YlTrm9N38-Trm112
Time (min)
µcal/sec
KCal/Mole of Injectant
Molar Ratio
Kd ≈ 27µM
Kd ≈ 29µM
Kd ≈ 34µM

B.

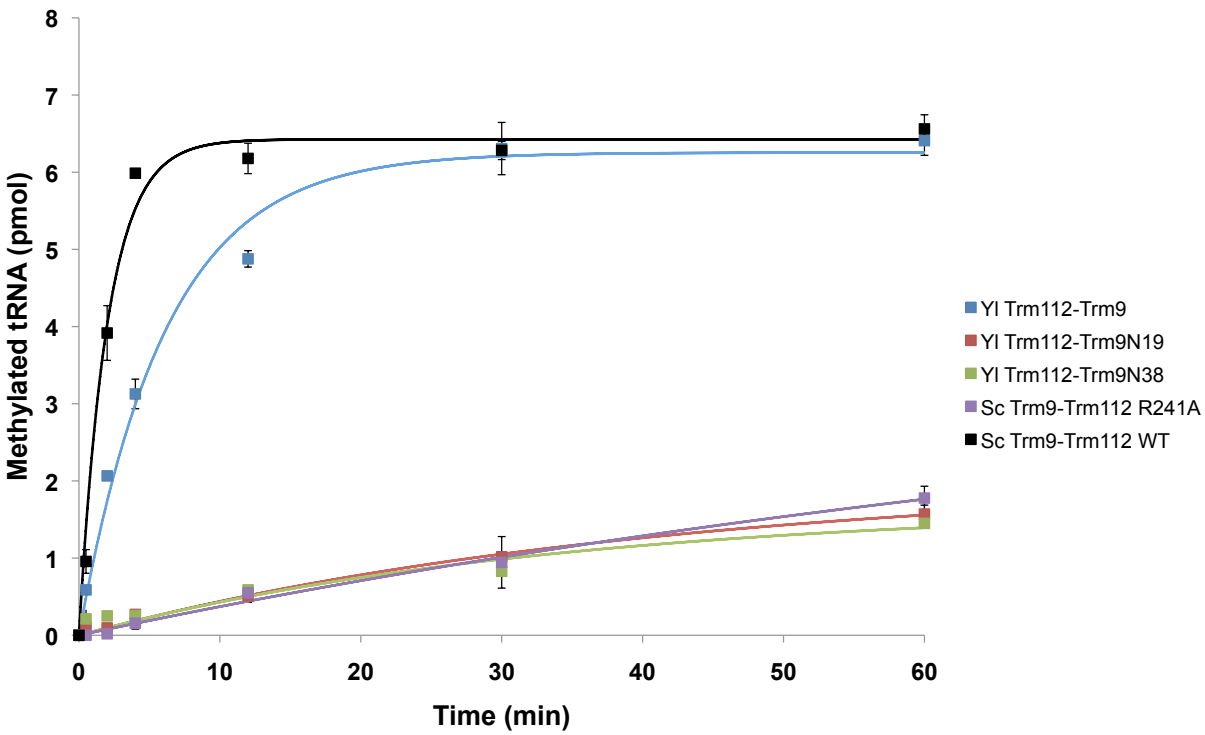
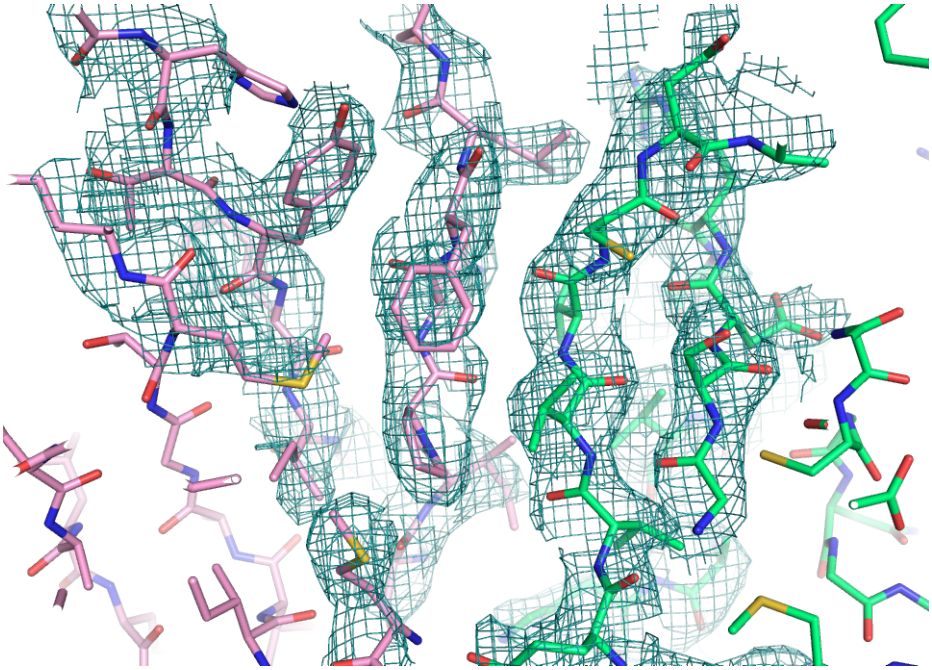

Methylated tRNA (pmol)
Time (min)
Yl Trm112-Trm9
Yl Trm112-Trm9N19
Yl Trm112-Trm9N38
Sc Trm9-Trm112 R241A
Sc Trm9-Trm112 WT

Figure S1

A.



B.

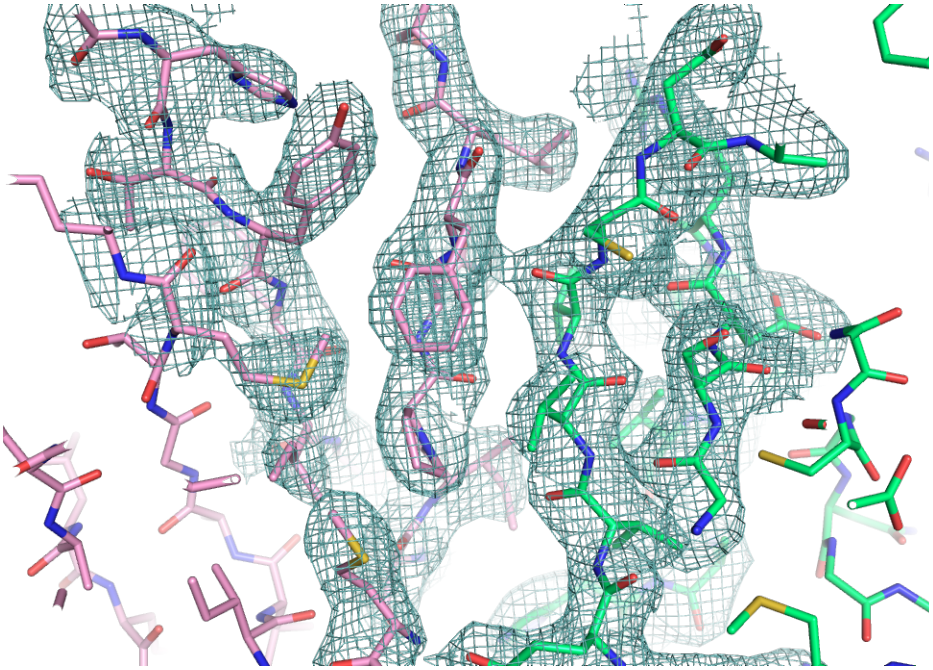
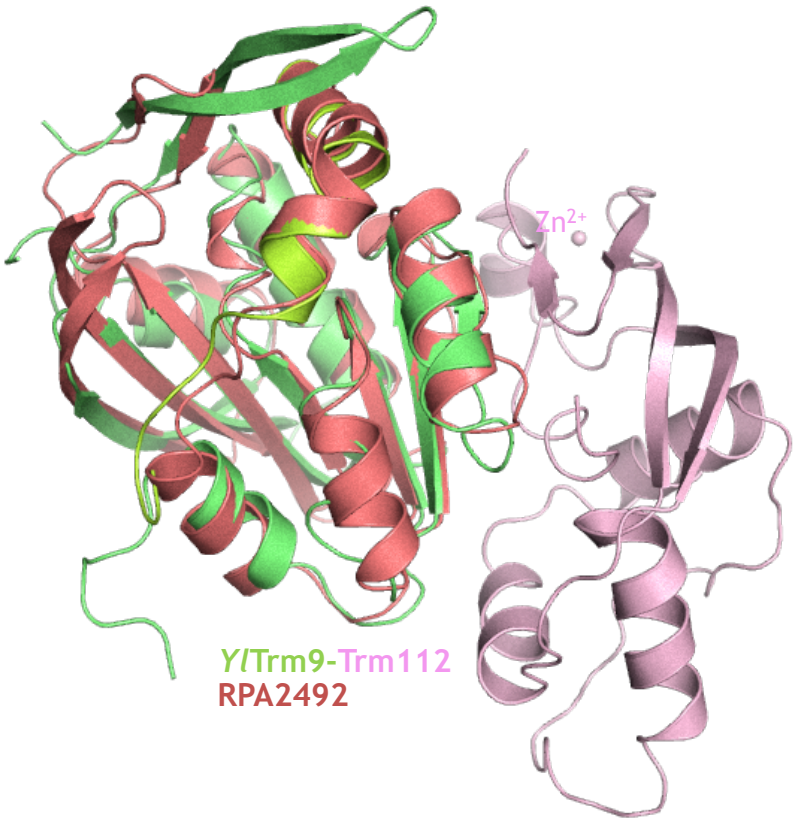


Figure S2

A.


Zn2+
YlTrm9-Trm112
RPA2492

B.

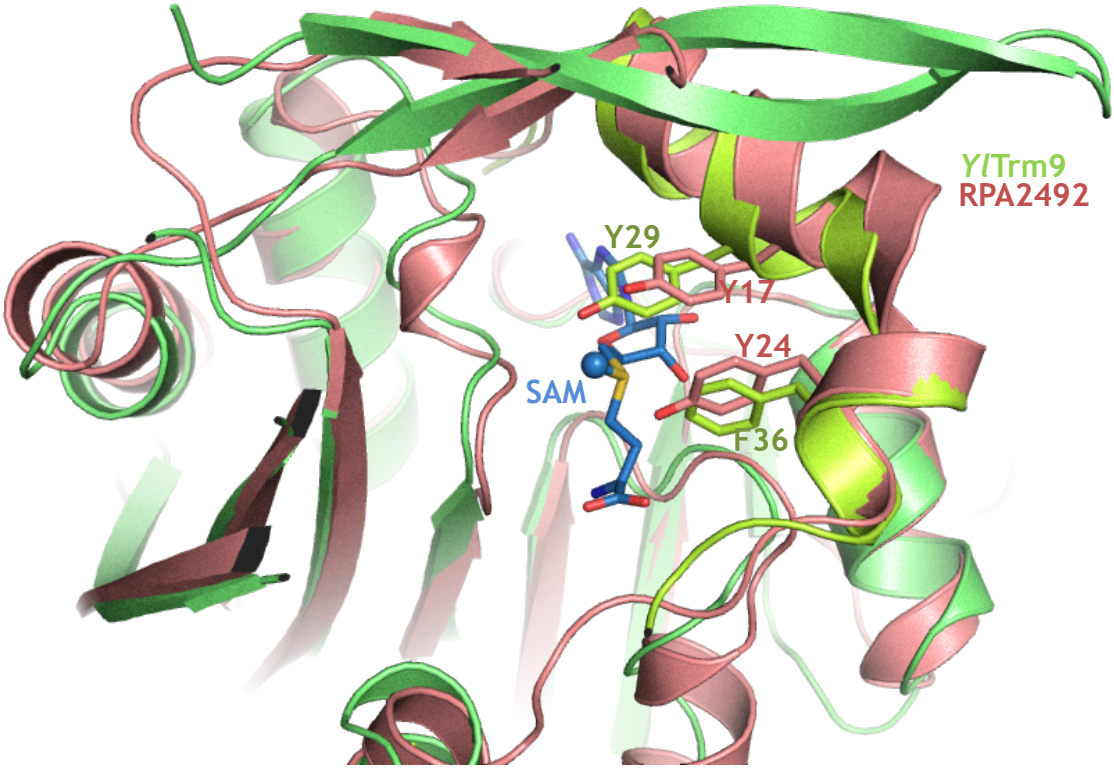
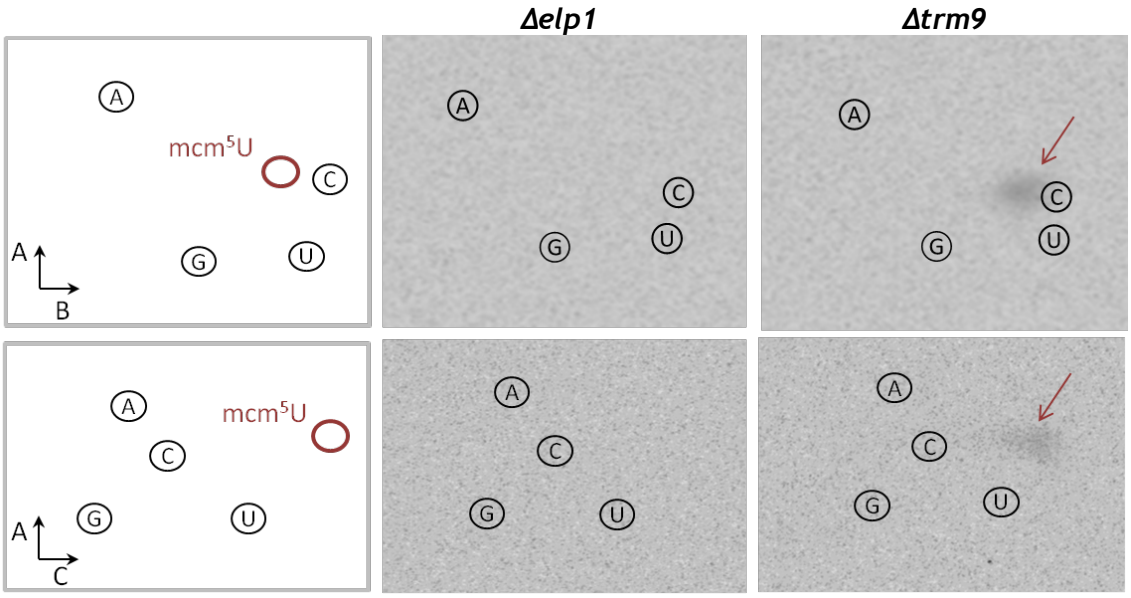
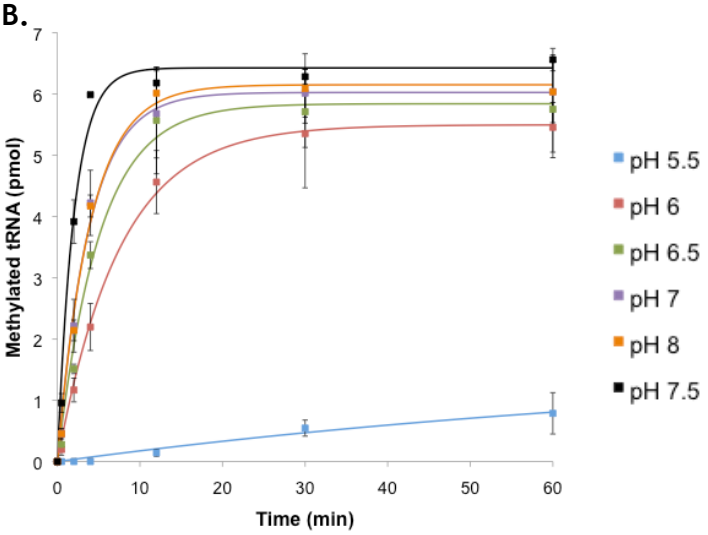

YlTrm9
RPA2492
Y29
Y17
Y24
SAM
F36

Figure S3

A.



B.

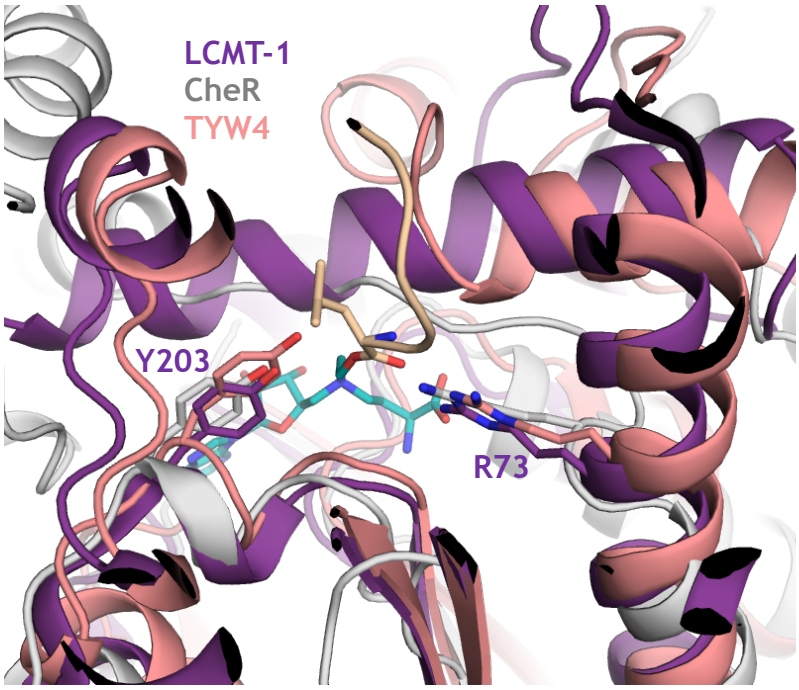


C.

pH	Apparent specific activity
5.5	12
6	486
6.5	766
7	1049
7.5	2105
8	1034

Figure S4

A.



B.

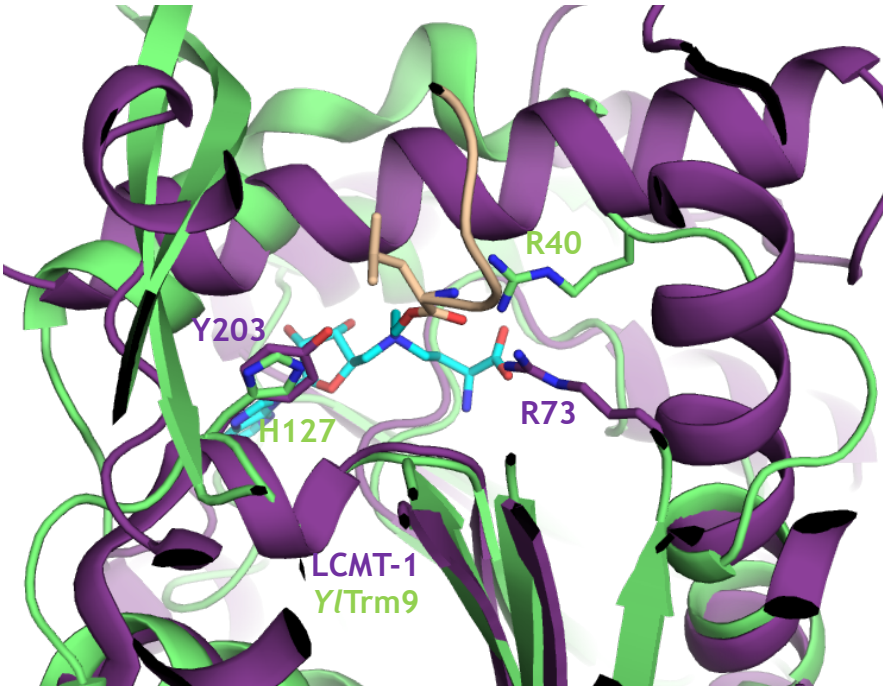


Figure S5

CHAPTER II

CHARACTERIZATION OF THE TRM112 INTERACTION NETWORK IN ARCHAEA

Bioinformatics analyses have revealed the presence of Trm112 orthologs in bacteria and archaea suggesting that its role might extend outside eukaryotic organisms (Purushothaman et al., 2005; Heurgue-Hamard et al., 2006; Bourgeois et al., 2017a). While nothing is known on bacterial Trm112 orthologs, the detection of Mtq2, Trm9 and Trm11 orthologs in archaeal genomes together with the strong similarity between eukaryotic and archaeal translation machineries suggest that archaeal Trm112 might play a similar role as eukaryotic Trm112 (Rivera et al., 1998; Yutin et al., 2008; Lyu, & Whitman, 2017). Indeed, Trm11 orthologs from *Pyrococcus abyssii* and *Thermococcus kodakarensis* have been biochemically characterized as enzymes methylating guanine nucleotide at position 10 of some tRNAs (Armengaud et al., 2004; Hirata et al., 2016). However, these enzymes not only catalyse the formation of N2-methylguanosine but also of N2,2-dimethylguanosine and are active on their own. This is of particular interest as the analysis of the distribution of Trm112 orthologs within archaeal genomes has revealed that Trm112 is absent in thermococcales and methanobacteriales, which encompass *Pyrococcus abyssii* and *Thermococcus kodakarensis* (Bourgeois et al., 2017a). Hence, this raises the question whether Trm112, when it is present in an archaeal organism, is important for Trm11 activity. Regarding Trm9, several observations argue in favor of its presence in some archaea. First, an initial survey of *Haloferax volcanii* genome suggested that the HVO_0574 gene encodes for a Trm9 ortholog (Grosjean et al., 2008) but a more recent analysis identified HVO_1032 as a better candidate (Gabriela Phillips, & de Crécy-Lagard, 2011). Second, genes encoding for proteins displaying some sequence similarity with the various enzymes (Elp3, Tuc1 and Trm9) involved in the formation of mcm^5s^2U modification at position 34 of tRNAs are present in *H. volcanii* and genes for Elp3 and Trm9 orthologs cluster in *Sulfolobus solfataricus* (Grosjean et al., 2008). Third, studies in the early 80's revealed the presence of unknown modifications at position U34 of some tRNAs from *H. volcanii* (Gupta, 1984). Regarding class I release factors, it is striking that despite radically different 3D-structures of the bacterial and eukaryotic factors, dedicated machineries have evolved to methylate the glutamine side chain of the universally conserved GGQ motif. Hence, one can imagine that such modification exists on archaeal aRF1. Considering the structural similarity between aRF1 and eukaryotic eRF1 factor (Song et al., 2000; Kobayashi et al., 2012), the enzyme responsible for this modification is very likely to be orthologous to Mtq2. This is further supported by the presence of a Mtq2 ortholog in all archaeal phyla (Heurgue-Hamard et

al., 2006; Bourgeois et al., 2017a). Finally, so far, no m^7G modification of the nucleotide corresponding to *S. cerevisiae* G_{1575} in archaeal 16S rRNAs has been found, in agreement with the absence of proteins with significant sequence homology with Bud23 in archaeal genomes (Bourgeois et al., 2017a).

To clarify the roles of archaeal Trm112, I have performed co-immunoprecipitation of Trm112 from the *H. volcanii* model organism to identify putative partners, validated some of these partners, characterized the enzymatic activity of two of these partners and solved the crystal structure of another one. Altogether, my results show that *H. volcanii* Trm112 (hereafter named *Hvo*Trm112) displays striking similarities with its eukaryotic orthologs but is also able to interact with a much larger number of MTases than yeast Trm112.

RESULTS

1. *Hvo*Trm112 interacting network in *H. volcanii*

1.1. Pop-in/Pop-out

In order to identify interacting partners of *H. volcanii* Trm112 protein (*Hvo*Trm112), I have employed an *in vivo* co-immunoprecipitation approach based on immuno-affinity of Anti-FLAG resin for FLAG tag fused to the C-terminal extremity of Trm112 from *H. volcanii (Hvo).* To do so, I first generated the target expression strain by deleting TRM112 gene in *Hvo* strain H98 using homologous recombination-based pop-in/pop-out method. The absence of TRM112 gene in colonies obtained after pop-in/pop-out was screened by colony lift experiment, which showed that the ratio between wild-type and deletion strains was nearly 1:1 as 51 colonies were *trm112* deleted out of 98 tested colonies (Figure 25A). The pop-in/pop-out result was furthermore confirmed by PCR on some colonies, *i.e.* those expected to be deleted for TRM112 gene (colonies 1-4, Figure 25B) and those expected to correspond to wild-type strain (colonies 5, 6 and negative control C, Figure 25B) according to the colony-lift results. Primers hybridizing in the upstream and downstream regions of the TRM112 gene in *H. volcanii* genome were used in the PCR reaction. This should result in PCR fragment of 800 bp if the TRM112 gene has been successfully deleted and in a PCR fragment of 983 bp in the case deletion failed. As shown in Figure 25B, the PCR and colony-lift results are fully consistent as strains 1 to 4, which were expected to be deleted for TRM112 gene from colony-lift assay, exhibit a PCR fragment of nearly 800 bp while strains 5 to 6 exhibit a PCR fragment of nearly 983 bp. This indicated that the *trm112*Δ *H. volcanii* strain is viable and hence that the TRM112 gene is not an essential gene.

1.2. Protein expression and Co-IP

For co-immunoprecipitation of flagged *Hvo*Trm112, the *Hvo* TRM112 gene fused with a FLAG sequence was cloned into tryptophan-induced vector pTA962 to yield pMG772 plasmid, which encodes for a C-terminally *Hvo*Trm112-FLAG. This plasmid was then successfully transformed into the *trm112*-deleted *Hvo* H98 strain. For the negative control, another tryptophan-inducible pTA927 plasmid harboring a sequence encoding the FLAG peptide only (kind gift from Pr. Anita Marchfelder, Ulm University, Germany) was also transformed into the same *H. volcanii* strain as the *Hvo*Trm112-FLAG protein. Protein expression was induced by

addition of 5mM tryptophan and cell lysis was done by osmotic shock (decreasing the salt concentration from 2.5M to around 100 mM).

As initial co-IP experiments resulted in a low number of specific peptide spectra in the mass spectrometry analysis (data not shown), I therefore introduced an *in vivo* cross-linking step with 1% formaldehyde before lysing the *Hvo* cells as previously described by Fischer and colleagues ((Fischer et al., 2010); Figure 26). Indeed, *Hvo*Trm112 could interact with some of its partners in a transient manner and such interactions might be maintained by high salt concentration while low salt concentration (as used during cell lysis or washing steps on the FLAG-resin) could induce dissociation and/or result in aggregation. This resulted in an increased number of proteins identified by mass spectrometry coupled to a higher number of specific peptide spectra.

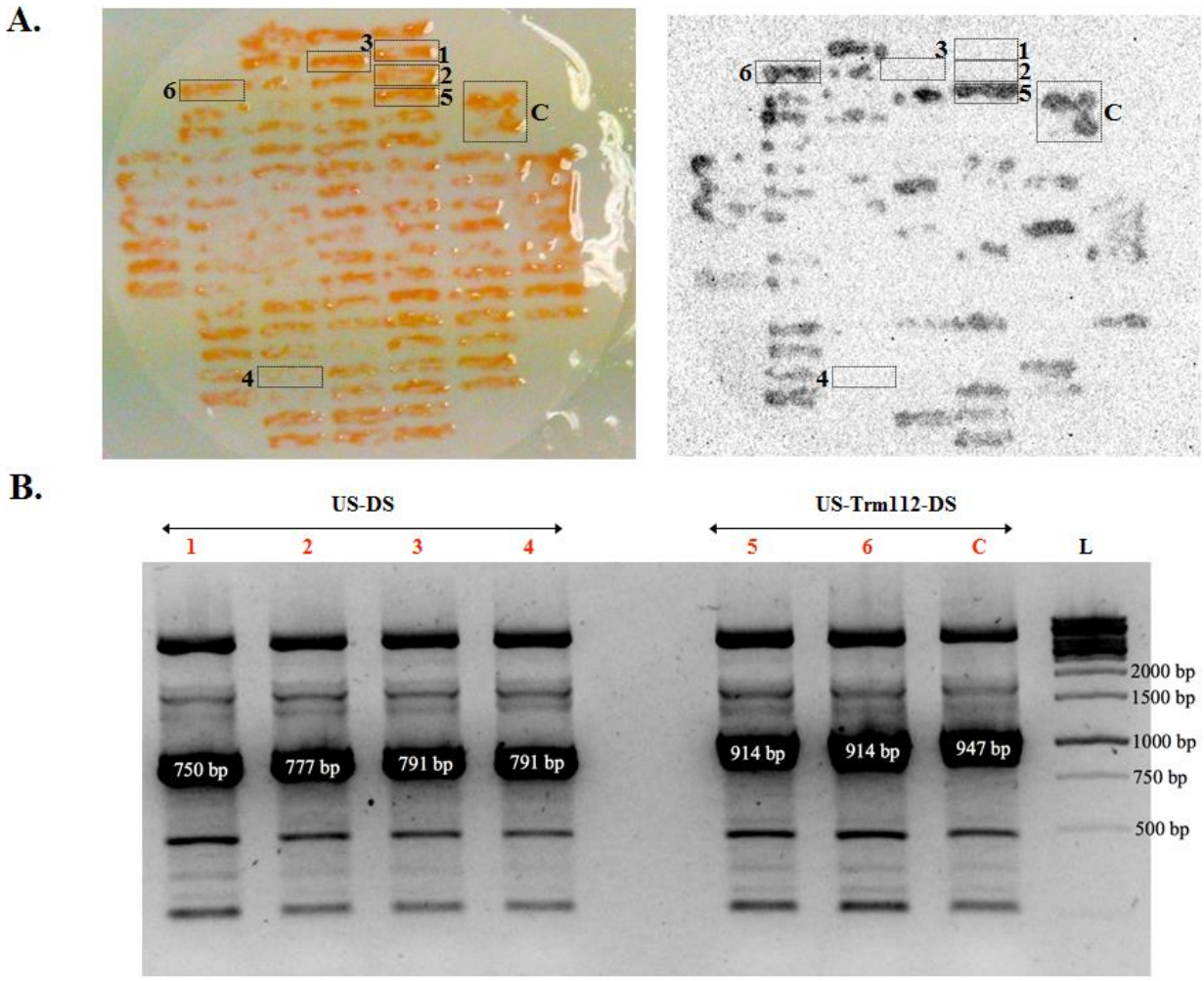


Figure 25. **Validation of the deletion of *H. volcanii* TRM112 gene by pop-in/pop-out.**

A. Colony lift assay. Out of the 98 colonies transferred onto a nitrocellulose membrane (left panel), 51 showed no signal on the membrane (right panel), indicating the deletion of *Hvo* TRM112 gene whereas 47 clones exhibited signal revealing the presence of *Hvo* TRM112 gene. The wild-type strain is shown as a negative control (C). Colonies labeled 1 to 6 were selected for a second validation step by PCR.

B. Validation by PCR. L – DNA ladder; 1-4: colonies without signal in the colony lift screen; 5-6 and C (control): colonies with signal in the colony lift screen. The lengths of US-DS and of US-Trm112-DS are 800bp and 983bp, respectively.

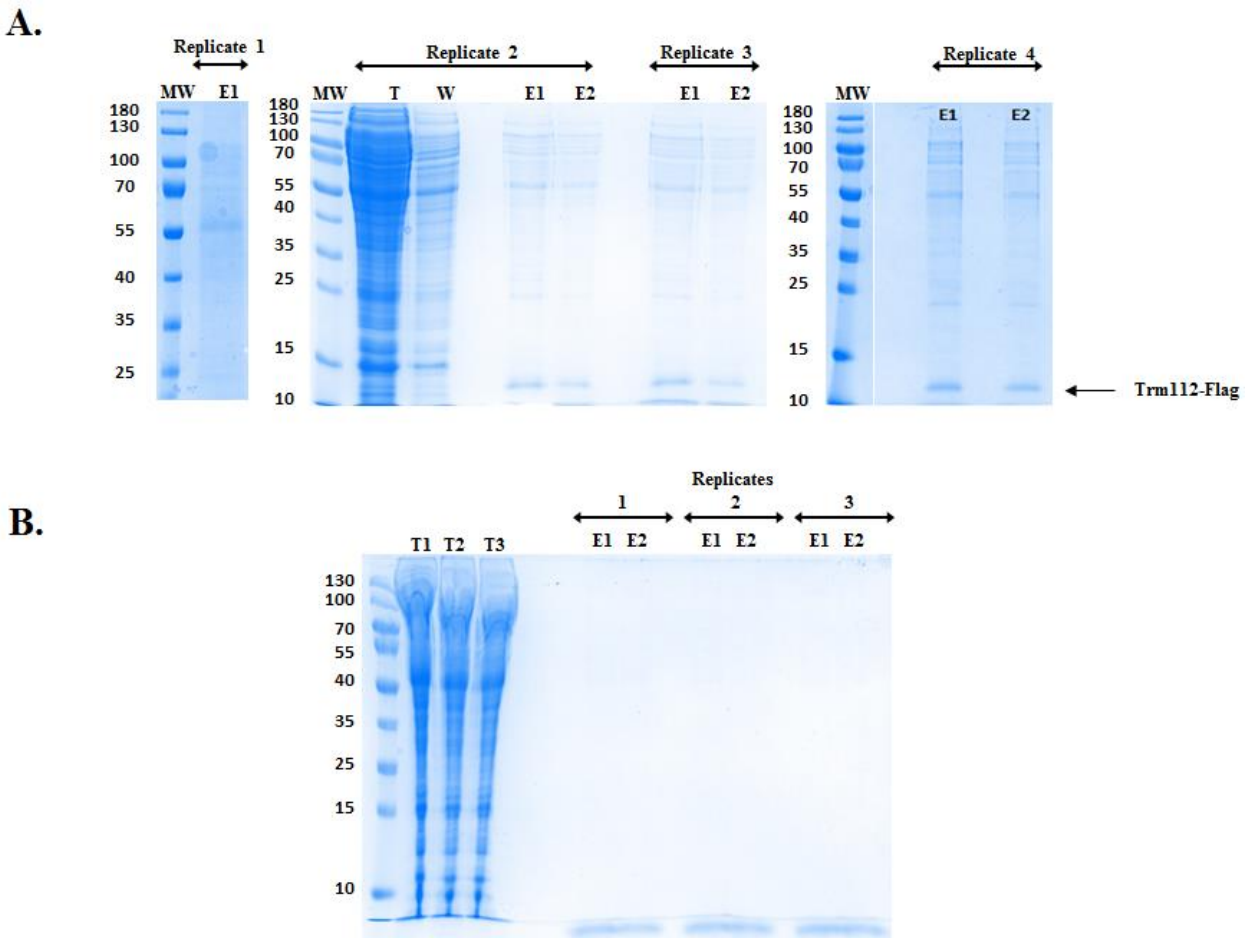


Figure 26. **Co-IP experiments.**

After Co-IPs, the cross-linked proteins were heated at 99^0C for 20mins to reverse cross-linking before running on 15% SDS-PAGE.

A. Co-IP performed with the *Hvo*Trm112-FLAG construct. This experiment was performed in quadruplicate. MW: Molecular weight marker. T: Total extract. W: Washing fraction. E1 and E2: Elution fractions. It is noteworthy that the SDS-PAGE for replicate 1 migrated too long and then the Trm112-FLAG protein ran out of the gel.

B. Co-IP performed with the FLAG-only construct. This experiment was performed in triplicate. Same legend as for panel A. MW: Molecular weight marker. Tn: Total extract for replicate n. E1 and E2: Elution fractions.

1.3. **Mass spectrometry and Trm112 partners identification**

As suggested by the SDS-PAGE analysis of the co-IP results (Figure 26), more than 1000 proteins were identified by LC-MS analyses of four independent co-immunoprecipitation experiments of *Hvo*Trm112-FLAG whereas only few hundreds of partners were identified in the negative control experiment with the FLAG-only construct. To filter out these results, we applied the following different filters: (1) proteins with mean number of specific spectra for the four *Hvo*Trm112-FLAG co-IP experiments lower or equal to 2 were removed; (2) proteins with ratio between mean number of specific spectra for the four *Hvo*Trm112-FLAG co-IP experiments and mean number of specific spectra for the three FLAG-only experiments lower or equal to 2 were removed; (3) proteins with maximum 2 specific spectra in two out of four *Hvo*Trm112-FLAG co-IP experiments were not considered; (4) proteins also identified by Fischer *et al (Fischer et al., 2010)* in their co-immuniprecipitation of *Hvo*Lsm-FLAG (performed in the same conditions as our analysis) were considered as non-specific interacting partners and were deleted. After strictly applying these filtering criteria, we ended up with a final list of 513 proteins.

We then focused on the 100 proteins that exhibit the higher mean values of specific spectra in the *Hvo*Trm112-FLAG co-IP experiments and sorted these in different families according to their molecular functions (*i.e.* nucleic acid binding, nuclease, ligase, MTase, kinase, synthase, peptidase, isomerase, …). As seen in Figure 27, four molecular functions (lyase, ligase, isomerase and MTase) are strongly enriched in the *Hvo*Trm112-FLAG experiment compared to the FLAG-only one while others (transporter, electron carrier, transcription) are completely absent. We also noticed a significant over-representation of proteins annotated to participate to translation or ribosomal function (translation factors, ribosomal proteins, …). The strong enrichment (more than 7-fold) in MTases (Figure 27) and to a lesser extent of proteins involved in translation is noteworthy as eukaryotic Trm112 is a well-known partner and activator of MTases modifying factors involved in translation (rRNAs, tRNAs, release factors). Further analysis of the final list of 513 proteins reveals that 26 MTases out of the 63 identified in *H. volcanii* proteome are detected in our experiment. Of particular interest is the presence in this list of proposed orthologs for eukaryotic Trm112 partners: Mtq2 (PrmC or HVO_2744, hereafter termed *Hvo*Mtq2), Trm9 (HVO_1032, named *Hvo*Trm9; (G. Phillips, & de Crecy-Lagard, 2011)) but also Trm11 (TrmG10, hereafter termed *Hvo*Trm11). This led us to propose that

*Hvo*Trm112 exhibits some similarity with its eukaryotic orthologs but may also have a larger number of MTase partners.

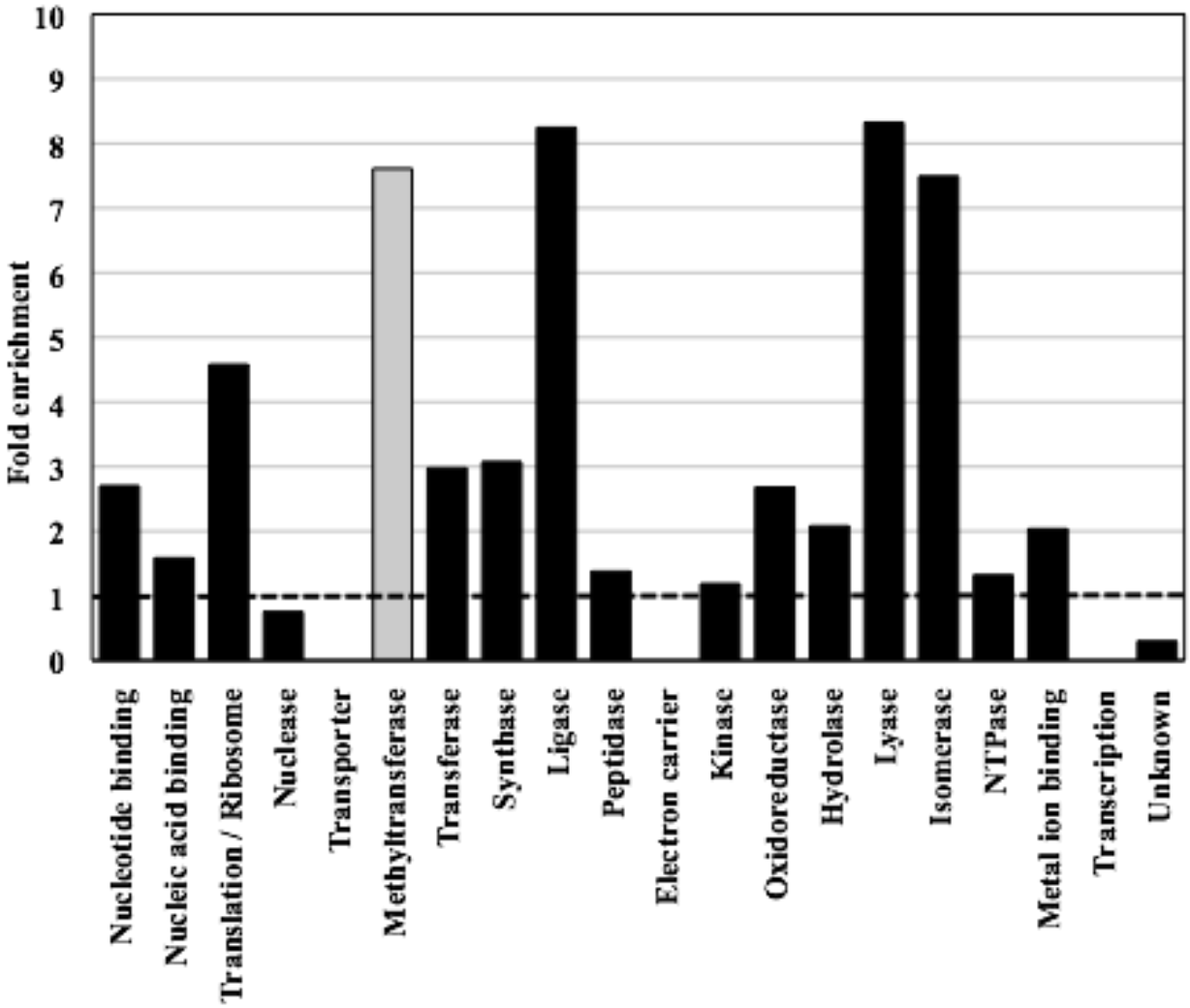


Figure 27**. Trm112 interacting network in *H. volcanii*.**

Enrichment of major "molecular function" GO terms in the 100 proteins exhibiting the higher mean value of specific spectra in the *Hvo*Trm112-FLAG co-IP experiments compared to the entire *H. volcanii* proteome. The dashed lines show an enrichment of one fold. Methyltransferase activity is shown in grey. Methyltransferases have not been considered as members of the transferase "molecular function" GO term in this analysis.

2. *Hvo*Trm112 partners characterization

To validate our co-IP results, we selected several MTases from this list and investigated whether these are indeed *bona fide* partners of *Hvo*Trm112. Besides putative orthologs of known eukaryotic Trm112 partners, namely *Hvo*Mtq2, *Hvo*Trm9 and *Hvo*Trm11, six other MTases were also chosen: HVO_0773 (exhibiting the largest number of specific spectra), HVO_0475 (third MTase with the largest number of specific spectra), HVO_0574 (initially proposed to be Trm9 archaeal ortholog (Grosjean et al., 2008)), HVO_0019, HVO_2875 and HVO_1715. These MTases were heterologously expressed in *E. coli* as fusion proteins with a N-terminal His-tag either alone or with untagged *Hvo*Trm112 and purified on Ni-NTA resin. Interestingly, similarly to *S. cerevisiae* Mtq2, Trm9 and Bud23 (Heurgue-Hamard et al., 2006; Mazauric et al., 2010; Figaro et al., 2012), some MTases (*Hvo*Mtq2, *Hvo*Trm9, HVO_0019, HVO_1715 and HVO_0773) can only be over-expressed as soluble proteins in the presence of *Hvo*Trm112 (Figure 28A) while others (HVO_0475 and HVO_0574) do not need *Hvo*Trm112 (Figure 28B). Furthermore, when co-expressed with these MTases, *Hvo*Trm112 co-purified with all of these as well as with HVO_2875 (expression assay was not performed for this protein in the absence of *Hvo*Trm112), supporting interaction. Unfortunately, despite extensive efforts (optimized gene, fusion with GST,…), it was not possible to express TrmG10/*Hvo*Trm11 as a soluble protein alone or in the presence of *Hvo*Trm112 (Figure 29).

To further confirm that these MTases interact directly with *Hvo*Trm112, a 3-steps (Ni-NTA, ion-exchange and size-exclusion chromatographies) purification protocol was used from *E. coli* cultures co-expressing each His-tagged MTase with *Hvo*Trm112 (illustrated for *Hvo*Trm112-HVO_0019 complex in Figure 30). For each MTase, we could observe a second band, which corresponds to *Hvo*Trm112 according to mass spectrometry analyses, on the Coomassie stained SDS-PAGE analyses of purified proteins, indicating that *Hvo*Trm112 interacts with all these MTases and that the resulting complexes are stable. SEC-MALLS analyses further revealed that these *Hvo*Trm112-MTase complexes adopt different oligomeric states (Table 4; Figure 31), *i.e.* heterodimers (*Hvo*Trm112-Trm9, *Hvo*Trm112-Mtq2, *Hvo*Trm112-HVO_0773, *Hvo*Trm112-HVO_0574 and *Hvo*Trm112-HVO_1715), heterotetramers (*Hvo*Trm112-HVO_0019) or heterohexamers (*Hvo*Trm112-HVO_0475).

We thus conclude that in the *H. volcanii* archeon, Trm112 interacts directly with at least seven different MTases (*Hvo*Mtq2, *Hvo*Trm9, HVO_0019, HVO_0773, HVO_0574, HVO_0475, HVO_2875 and HVO_1715) and hence its interaction network is more complex than for its eukaryotic orthologs studied so far (Bourgeois et al., 2017a).

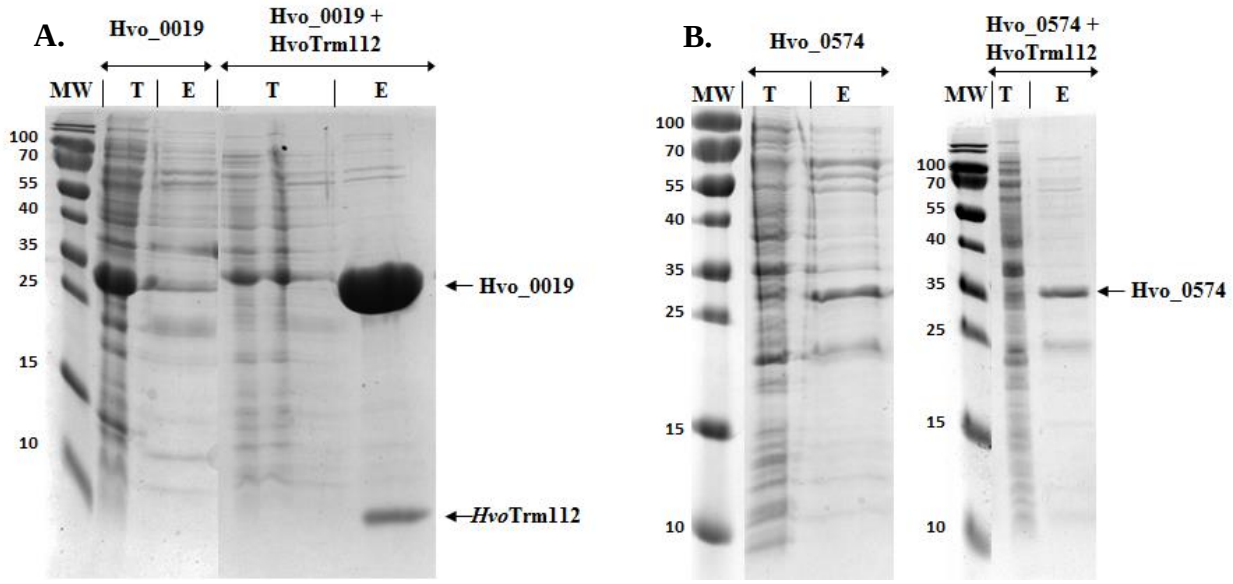


Figure 28**. *Hvo*Trm112 solubilizes most of its methyltransferase partners.**

A. *Hvo*Trm112 strongly enhances HVO_0019 solubility. SDS-PAGE analysis of total (T) and Ni-NTA elution (E) fractions obtained from expression of HVO_0019-His_6 either alone or together with *Hvo*Trm112 at 30°C using *E. coli* Codon+ strain in Terrific Broth Auto-inducible medium. MW: Molecular weight ladder (in kDa).

B. *Hvo*Trm112 has little effect on HVO_0574 solubility. SDS-PAGE analysis of total (T) and Ni-NTA elution (E) fractions obtained from expression of HVO_0574-His_6 either alone or together with *Hvo*Trm112 at 18°C using *E. coli* Codon+ strain in 2YT medium and upon induction by 0.05 mM IPTG. MW: Molecular weight ladder (in kDa).

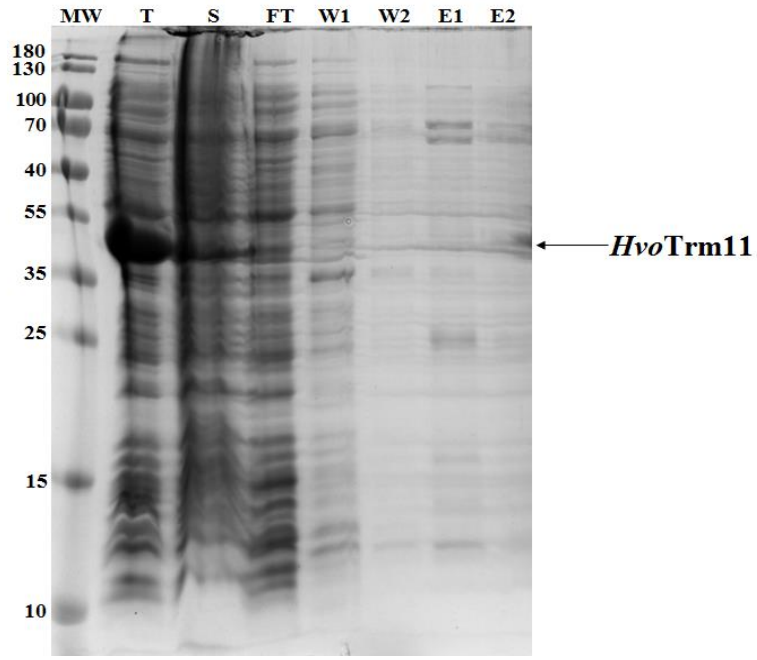
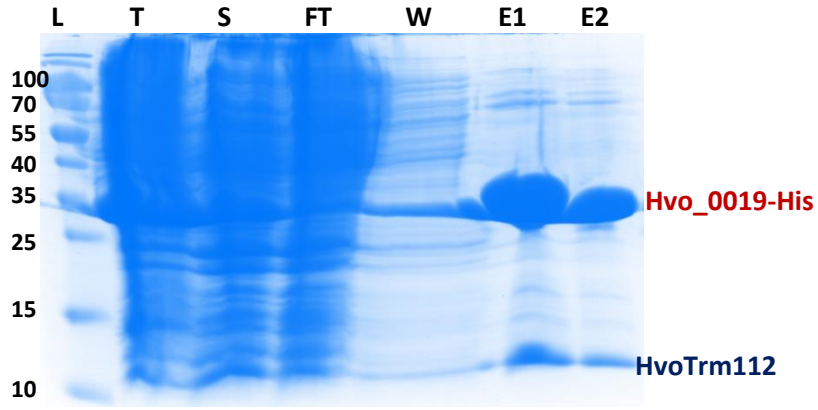


Figure 29. **Co-expression assay of *Hvo*Trm11-His$_6$ with *Hvo*Trm112.**

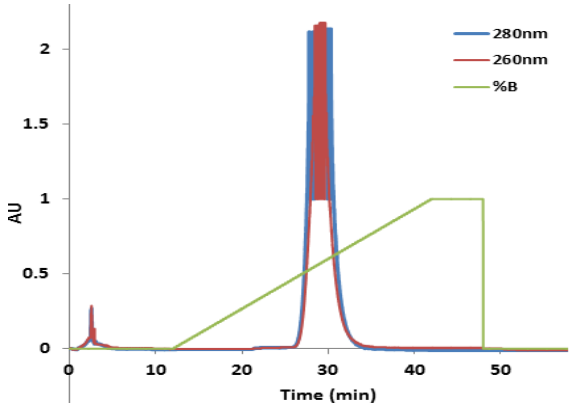
H. volcanii Trm11 and Trm112 proteins were co-expressed in *E. coli* Codon+ strain at 18°C using 2YT media. Protein expression was induced by adding IPTG (1mM final concentration) after a cold shock (15 minutes on ice). Similar results were obtained with either *E. coli* Gold strain, Terrific Broth Autoinducible or LB media, lower IPTG concentrations, optimized HVO_TRM11 gene or at 37°C. The theoretical molecular weight for *Hvo*Trm11-His$_6$ protein is 35 kDa. T: Total extract. S: Soluble extract. FT: Flow-through. Wn: Washing fraction n. E1 and E2: Elution fractions.

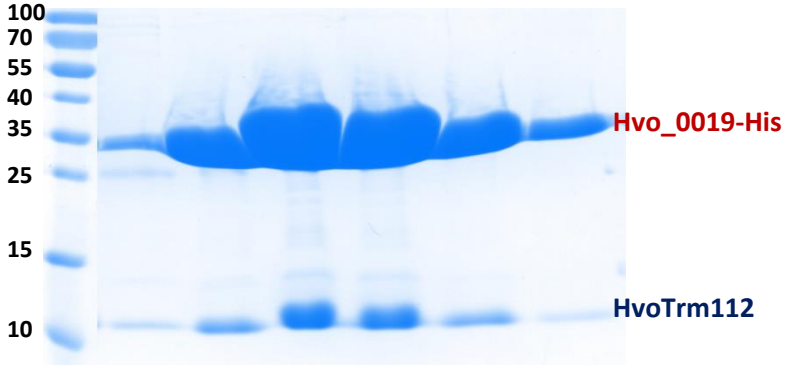
Ni-NTA chromatography



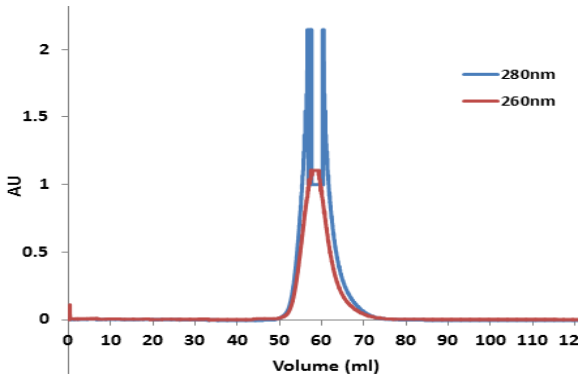
L: Protein ladder
T: Total extract
S: Soluble extract
FT: Flowthough
W: Wash
E1: Elution 1
E2: Elution 2

Ion-exchange chromatography: Mono Q





Size-exclusion chromatography: S75 16/60



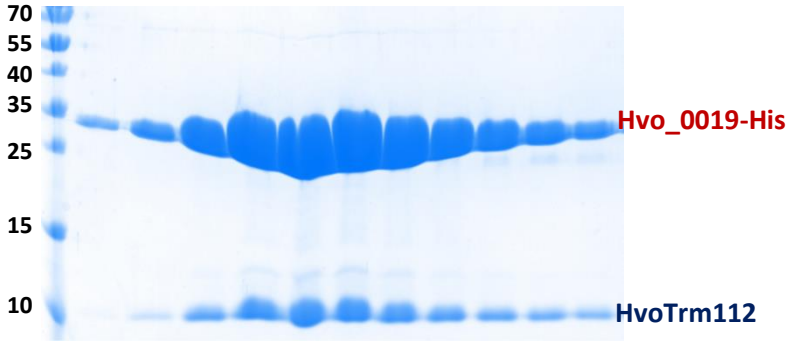


Figure 30. **An example of *Hvo*Trm112-*Hvo*MTase-His$_6$ purification through 3-step chromatographies**. The proteins were run on 15% SDS-PAGE gel

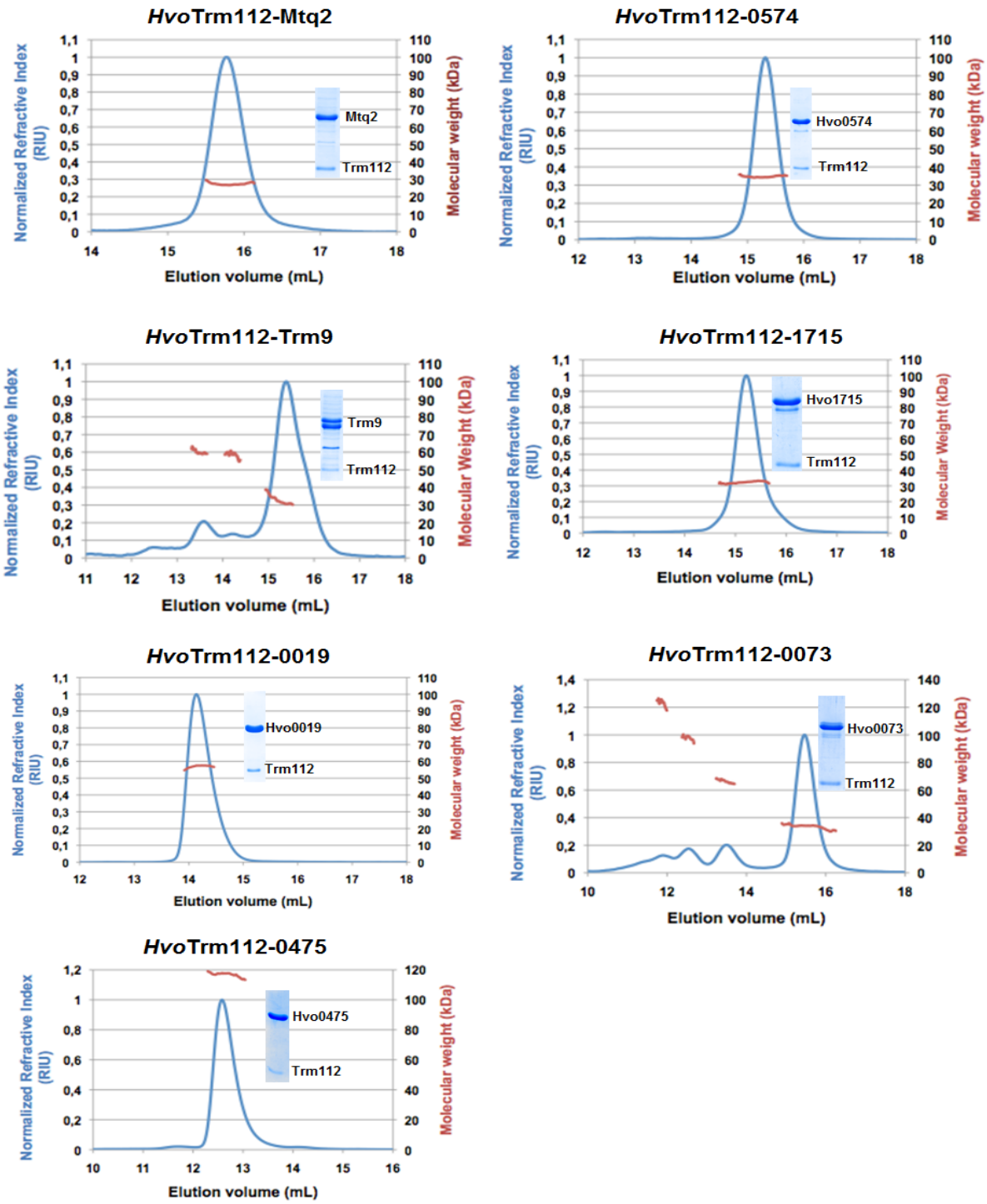


Figure 31. **Analyses of *Hvo*Trm112-MTase complexes by SEC-MALLS.**

A zoom centered on the main peak with the refractive index colored in blue (left y-axis) and the distribution of molecular mass calculated from light scattering along this peak colored in red (right y-axis) is shown for each *Hvo*Trm112-MTase complex studied. The Coomassie stained SDS-PAGE analyses of the proteins present in the main peak are shown.

Table 4. **Oligomeric states of *Hvo*Trm112-MTase complexes.**

***Hvo* Trm112-MTase complex**	**Theoretical MW of heterodimer (kDa)**	**Experimental MW determined by SEC-MALLS (kDa)**	**Oligomeric states**
*Hvo*Trm112-Trm9	31	29.5	Heterodimer
*Hvo*Trm112-Mtq2	28.9	27.3	Heterodimer
*Hvo*Trm112-HVO_0019	33.2	59.8	Heterotetramer
*Hvo*Trm112-HVO_0574	36.8	34.6	Heterodimer
*Hvo*Trm112-HVO_0475	40.1	117	Heterohexamer
*Hvo*Trm112-HVO_0773	35.4	33.9	Heterodimer
*Hvo*Trm112-HVO_1715	34.6	32.3	Heterodimer
*Hvo*Trm112-HVO_2875	28.6	ND	ND

3. Functional studies

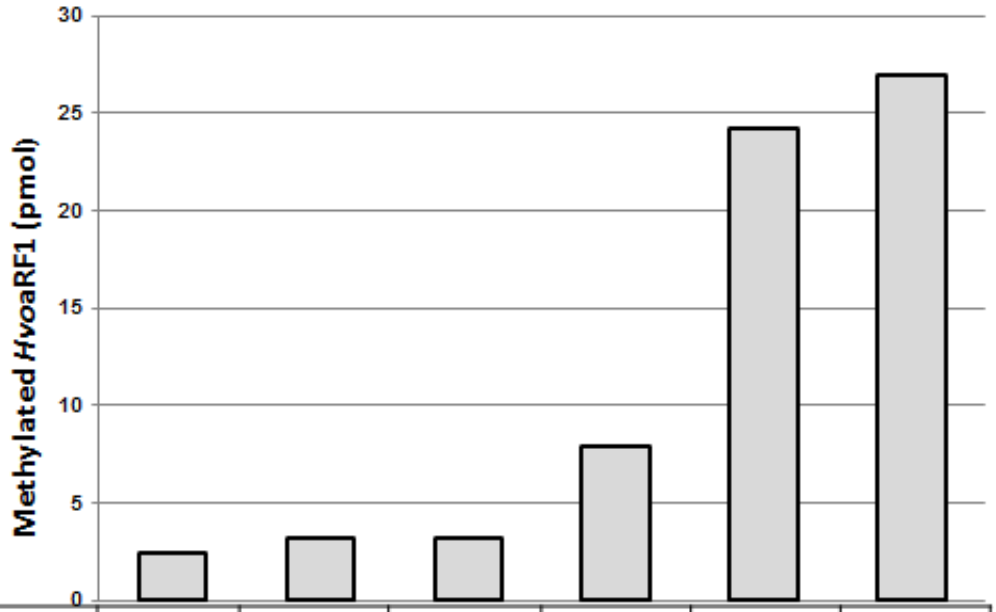
Among all newly-identified *Hvo*Trm112 partners, two proteins HVO_2744 and HVO_1032 are known as homologs of eukaryotic Trm112 partners, namely Mtq2 and Trm9 respectively, paving the road towards their functional characterization through enzymatic assay.

3.1. *Hvo*Mtq2-Trm112

3.1.1 *In vitro* enzymatic assay

In eukaryotes, the Trm112-Mtq2 complex has been shown to be enzymatically active on class I translation termination factor eRF1 but only when this later exists as a complex with the GTP-bound form of class II translation termination factor eRF3 (Heurgue-Hamard et al., 2006; Figaro et al., 2008). To characterize the enzymatic activity of the *Hvo*Trm112-Mtq2 complex *in vitro*, I have over-expressed in *E. coli* and purified *Hvo*aRF1 as well as *Hvo*aRF3 (also known as *Hvo*aEF1A). This later is known to contribute to several activities in archaea, namely translation elongation, translation termination and quality control pathways dedicated to the release of stalled ribosomes as ortholog of eEF1A, eRF3 and Hbs1, respectively (Saito et al., 2010). It is noteworthy that from the strong structural similarity between eukaryotic and archaeal class I translation termination factors and the striking differences between bacterial and eukaryotic enzymes modifying the GGQ motif of class I translation termination factors (Graille et al., 2012; Kobayashi et al., 2012), aRF1 is very unlikely to be methylated in *E. coli* and hence might be substrate for the *Hvo*Trm112-Mtq2 complex. First, I compared the enzymatic activity of *Hvo*Trm112-Mtq2 complex on *Hvo*aRF1 in presence of *Hvo*aRF3 and GTP, both in the absence of KCl or in the presence of 3M KCl, which corresponds to physiological conditions (Meury, & Kohiyama, 1989). In agreement with several reports on enzymes from *H. volcanii* archeon, a strong enzymatic activity could be detected only in the presence of 3M KCl (Figure 33). This activity also required high concentration (400mM) of phosphate buffer (Figure 32). In these conditions, *Hvo*Mtq2-Trm112 complex catalyzes the methylation of nearly 26 pmol of *Hvo*aRF1 out of 100 pmol total substrate in 2 hours. Next, we substituted Tyr111 from the NPPY signature in *Hvo*Mtq2 by Ala with the aim to inactivate *Hvo*Mtq2-Trm112 complex as the corresponding mutant of *S. cerevisiae* Mtq2-Trm112 complex resulted in complete loss of activity (Liger et al., 2011). The resulting mutant was indeed unable to methylate *Hvo*aRF1, indicating that as

suggested by the required high KCl concentration, the detected enzymatic activity does not originate from an *E. coli* contaminant (Figure 33). In addition, substitution of Gln187 from the *Hvo*aRF1 GGQ motif by Ala also resulted in undetectable activity. Finally, *Hvo*aRF3 and GTP are necessary for methylation of *Hvo*aRF1 by the *Hvo*Mtq2-Trm112 MTase. Altogether, this demonstrates that *Hvo*Mtq2-Trm112 holoenzyme catalyzes the methylation of *Hvo*aRF1 on the glutamine side chain from its GGQ motif in a *Hvo*aRF3- and GTP-dependent manner similarly to its yeast and human orthologs (Heurgue-Hamard et al., 2006; Figaro et al., 2008).



*Hvo*Trm112-Mtq2 WT (5 pmol)	+	+	+	+	+	+
*Hvo*aRF1 WT (100 pmol)	+	+	+	+	+	+
*Hvo*aRF3 (100 pmol)	+	+	+	+	+	+
GTP (60 nmol)	+	+	+	+	+	+
KCl (3 M)	+	+	+	+	+	+
EDTA (2.5mM)	+	+	+	+	+	+
Phosphate concentration (mM)	100	100	200	200	400	400
$MgCl_2$ (mM)	10	5	10	5	10	5
NH_4Cl (mM)	10	5	10	5	10	5

Figure 32. **Buffer optimization for enzymatic activity of *Hvo*Mtq2-Trm112**. Different buffer conditions tested were indicated in the table below the graph.

*Hvo*Trm112-Mtq2 WT (5 pmol)	+	+		+	+	+	+
*Hvo*Trm112-Mtq2 Y111A (5pmol)			+				
*Hvo*aRF1 WT (100 pmol)	+	+	+		+		+
*Hvo*aRF1 Q187A (100 pmol)				+			
*Hvo*aRF3 (100 pmol)	+	+	+	+		+	+
GTP (60 nmol)	+	+	+	+	+	+	
KCl (3 M)	+		+	+	+	+	+

Figure 33. **Enzymatic activities of *Hvo*Trm112-Mtq2.** The conditions (proteins, salt and ligand) used for each experiments are indicated in the table below the graph. The number of methylated substrate (in pmol) after a 2 h reaction is indicated for every condition. Errors bars have been calculated from the results of three independent experiments.

3.1.2 *In vivo* methylation identification

I have already confirmed the function of *Hvo*Mtq2-Trm112 on methylating *Hvo* aRF1 at glutamine side chain of GGQ motif by *in intro* enzymatic assay. However, the *in vivo* methylation status of this protein is not known. To examine that, I have already purified aRF1-FLAG from different *H. volcanii* strains, in particular the wild-type and two control mutants (*mtq2*Δ and *trm112*Δ) by Co-IP experiments (Figure 35). Before doing so, these different *H. volcanii* strains were successfully generated through pop in/pop out method (Figure 34). The aRF1-FLAG proteins were already sent for mass spectrometry. For aRF1 purified from wild-type strain, preliminary results suggest the presence of two methyl groups on the peptide containing the GGQ motif. Results are not clear for aRF1 purified from either *mtq2*Δ and *trm112*Δ strains. Further studies are needed.

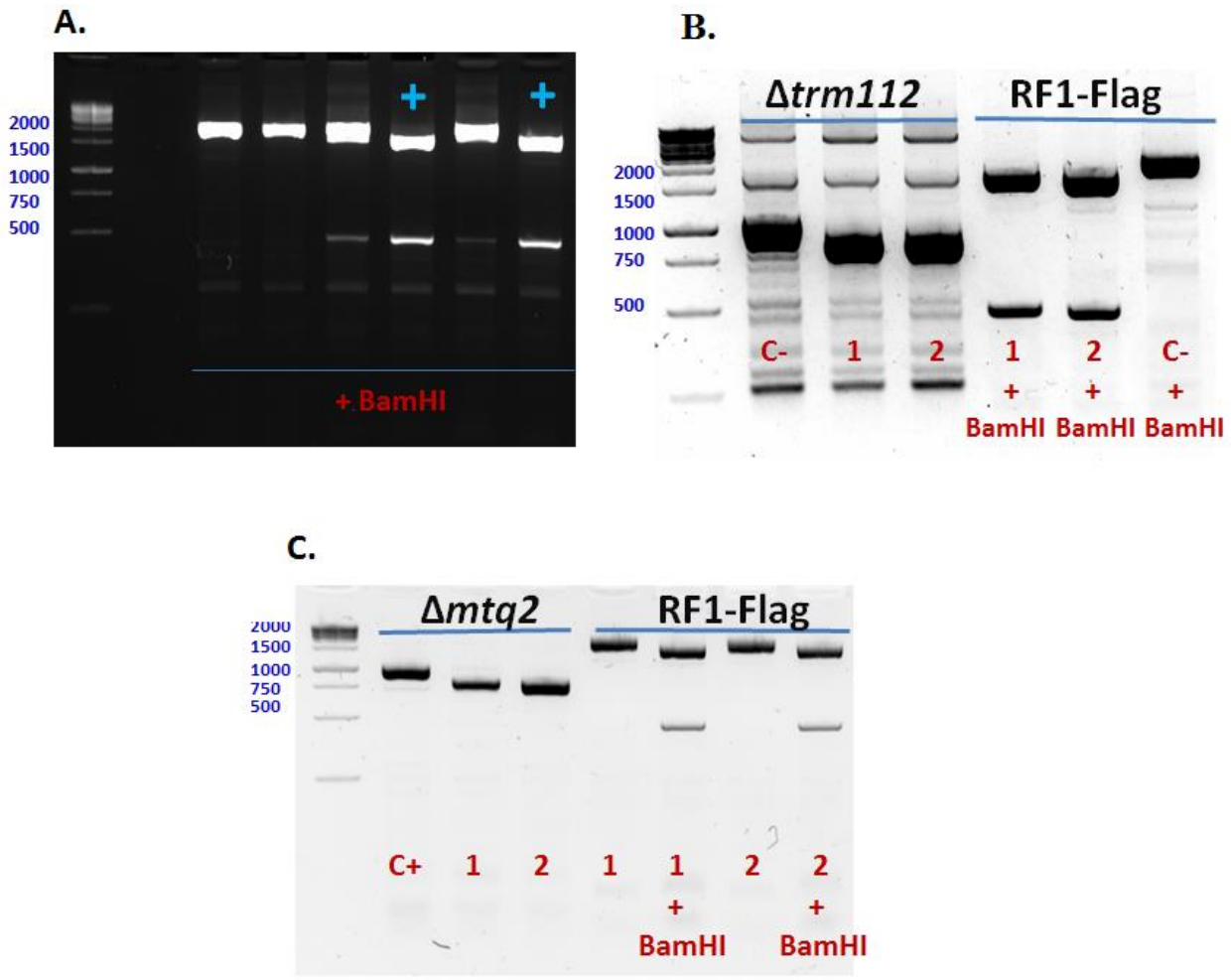


Figure 34**. Pop in/pop out result confirmed by PCR**. (**A)**. Presence of aRF1-FLAG in *Hvo* H26. PCR products in all tested clones were subjected to BamHI digestion (the BamHI site is present between the last codon of aRF1 gene and the first one of the FLAG tag sequence). Two clones with two fragments corresponding to approximately 1500 and 500bp are positive clones and are indicated by (+) sign. **(B).** Presence of both *Hvo* aRF1-FLAG and *trm112* Δ in *Hvo* H98. Two positive colonies (1-2) were confirmed. For *Hvo trm112*Δ check, colonies 1-2 contained only the US-DS (800bp) of a*trm112* in contrast to the US-a*trm112*-DS (around 980bp) in the negative control (C^-) while in case of aRF1-FLAG confirmation, PCR products amplified for US-aRF1-FLAG-DS from these two colonies cut into two fragments by BamHI indicating presence of aRF1-FLAG, which is not the case for the negative control (C^-) which PCR product (US-aRF1-DS) was not cut by BamHI. (**C).** Presence of both *Hvo* aRF1-FLAG and *Hvo mtq2*Δ in *Hvo* H26. For *Hvo mtq2*Δ check, positive colonies 1-2 contained only the US-DS (800bp) of *Hvo mtq2* while the positive control (C^+) was around 1000bp as it was a partly deleted a*mtq2* strain. For *Hvo* aRF1-FLAG check, PCR products amplifying US and DS of aRF1 were cut by BamHI in those two colonies, indicating the presence of aRF1-FLAG.

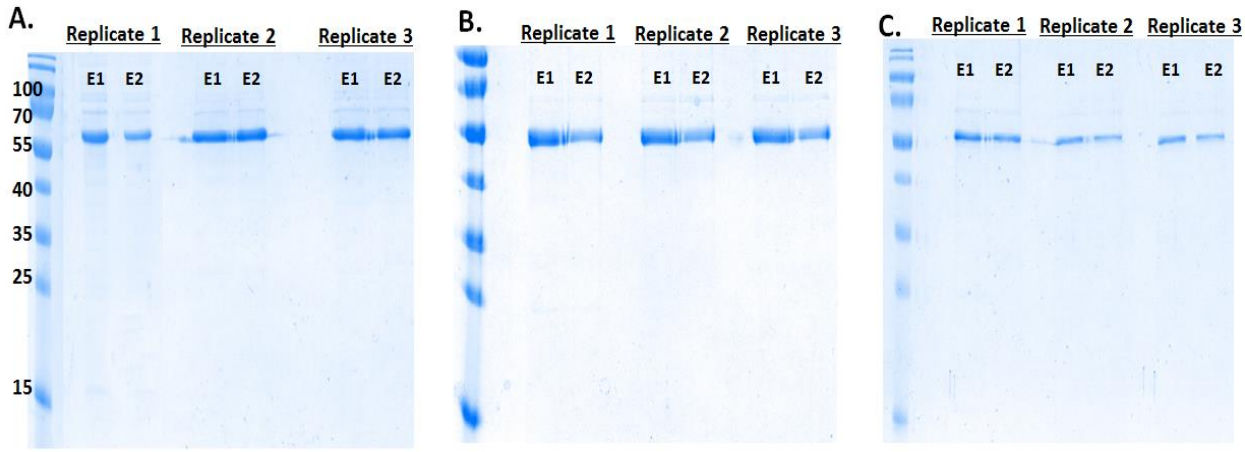


Figure 35**. Co-IP of aRF1-FLAG**. All experiments were done in three repeats which each contained elution 1 (E1) and elution 2 (E2). (**A).** WT *Hvo* H26 aRF1-FLAG. (**B).** *Hvo* H26 *trm112*Δ aRF1-FLAG. (**C).** *Hvo* H26 *mtq2*Δ aRF1-FLAG.

3.2. *Hvo*Trm9-Trm112

3.2.1 *In vitro* enzymatic activity

In parallel, we have investigated whether the *Hvo*Trm112-Trm9 acts as a tRNA modification enzyme similarly to its eukaryotic ortholog. To obtain putative tRNA substrates, we have generated an *H. volcanii* H26 strain deleted for HVO_1032 gene (hereafter termed *H. volcanii trm9*Δ), which encodes for *Hvo*Trm9, by the pop-in/pop-out method as described to build the *H. volcanii trm112*Δ (Figure 36). Our ability to obtain this strain indicates that the gene encoding for *Hvo*Trm9 is not essential. Total tRNAs purified from this strain as well as those purified from *S. cerevisiae trm9*Δ or *elp1*Δ strains (Letoquart et al., 2015a) were used for *in vitro* enzymatic assays. Surprisingly, the *Hvo*Trm112-Trm9 complex was active on total tRNAs from *S. cerevisiae trm9*Δ strain as substrates but not from *H. volcanii trm9*Δ strain (Figure 37A). *Hvo*Trm112-Trm9 complex (2 pmol) modifies up to 5.5 pmol of tRNAs from *S. cerevisiae trm9*Δ strain out of 100 pmol while *S. cerevisiae* Trm112-Trm9 complex (1.5 pmol) was modifying up to 7 pmol (out of 75 pmol) of the same tRNAs (Letoquart et al., 2015b). As observed for *Hvo*Trm112-Mtq2 complex, this complex is significantly more active in the presence of 3M KCl than in the absence of KCl, where only approximately 1.5 pmol of *Hvo*aRF1

are modified. Its activity depends on the presence of the cm^5U modification at position 34 of tRNA anticodon loop as tRNAs purified from *S. cerevisiae elp1*Δ strain are not substrates of this complex. Finally, as HVO_0475 gene product, which also interacts with *Hvo*Trm112, was initially predicted to be orthologous to eukaryotic Trm9 (Grosjean et al., 2008), we have included this purified complex in our enzymatic assay. As shown in Figure 37A, this complex does not exhibit enzymatic activity on total tRNAs extracted from both *S. cerevisiae trm9*Δ and *H. volcanii trm9*Δ strains. Altogether, these results demonstrate that the *Hvo*Trm112-Trm9 complex is indeed a tRNA methyltransferase, which, by analogy to its eukaryotic ortholog, most probably modifies cm^5U into mcm^5U at position 34 of the tRNAs as the cm^5U modification catalyzed by the Elp1-6 complex in eukaryotes (S. Glatt et al., 2012; Sebastian Glatt et al., 2016), is required for enzymatic activity.

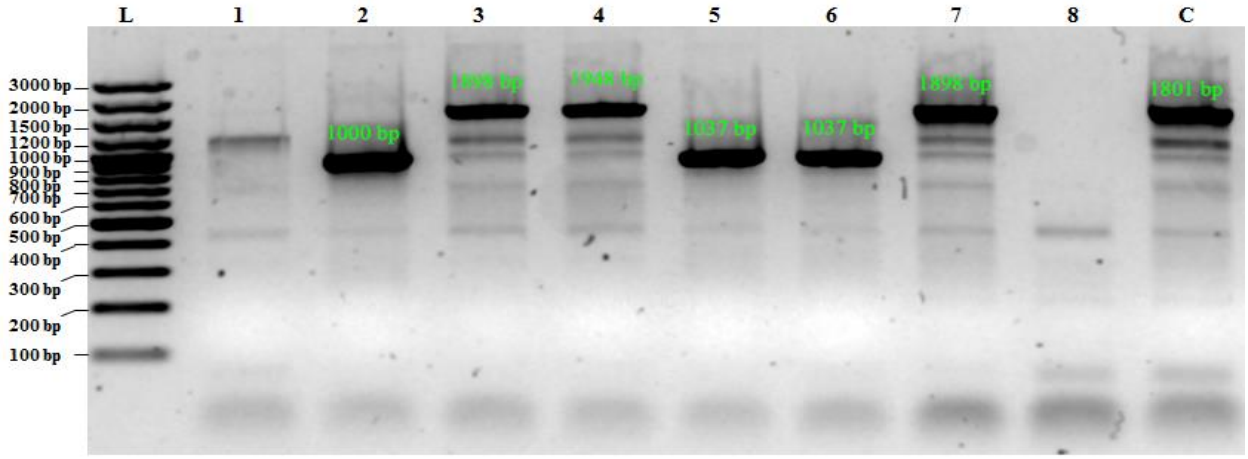
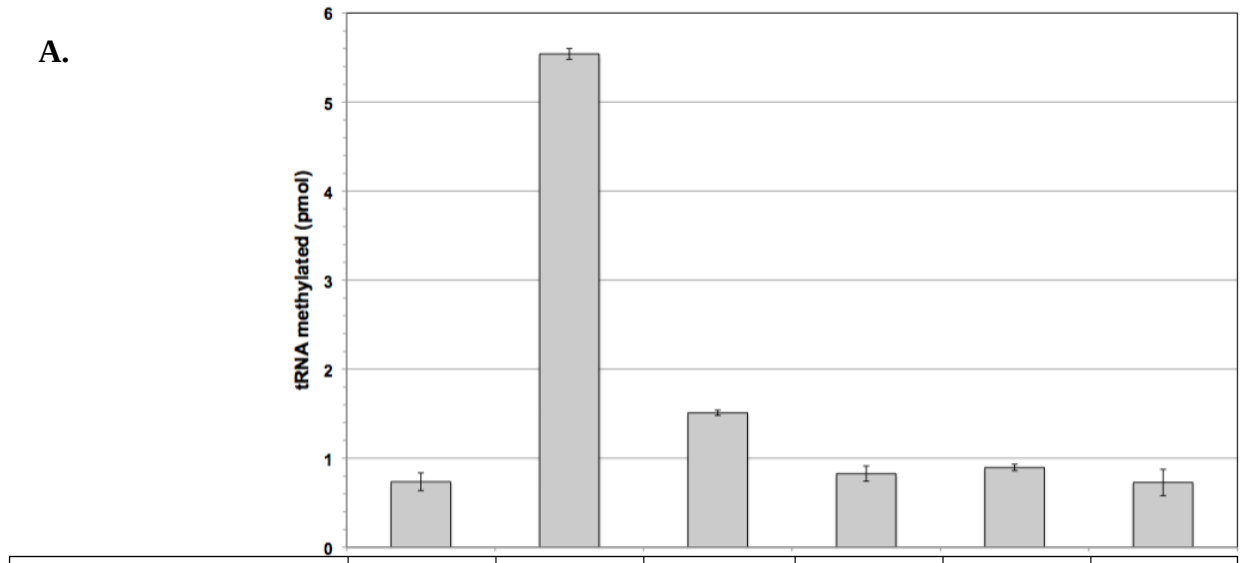


Figure 36. **Validation of the deletion of *H. volcanii* HVO_1032 (TRM9) gene by PCR.** PCR were performed on several randomly chosen colonies (lanes 1 to 8) as well as on wild-type *H. volcanii* H26 strain (C). L – DNA ladder. The theoretical lengths of US-DS and of US-HVO_1032-DS are 840 bp and 1494 bp, respectively. Colonies 2, 5 and 6 correspond to *H. volcanii trm9*Δ while colonies 3, 4 and 7 correspond to wild-type strain.

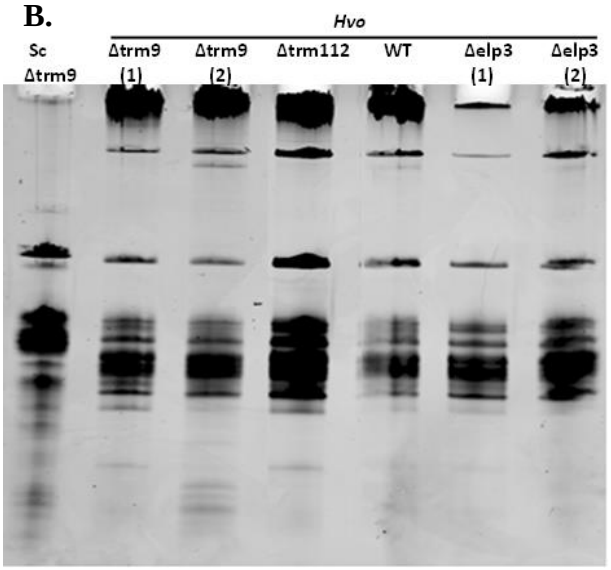
3.2.2 *In vivo* methylation identification

In case of *Hvo*Trm9-Trm112, I have detected enzymatic activity of the complex on total RNA from *trm9*Δ *S. cerevisiae*. Unfortunately no activity has been identified on total tRNAs from *H. volcanii trm9*Δ strain. Next, we plan to check whether tRNAs from different *H. volcanii* strains contain expected methylation modifications such as $cm^5(s^2)U$, $mcm^5(s^2)U$ and ncm^5U.

We will also focus on m^2G and $m^2{}_2G$ modifications, which are known to be catalyzed by archaeal Trm11. In particular, we will examine the possibility that *Hvo*Trm11 requires *Hvo*Trm112 for its activity. To achieve these, I purified several total RNAs from different *H. volcanii* strains including the wild-type and mutants (*trm9*Δ, *trm112*Δ and *elp3*Δ) (Figure 37B-C). These were sent to a collaborator in USA for identification of tRNA modifications by HPLC coupled to MS. This will also be of particular interest as so far only few information is available on tRNA modifications in archaea and we cannot exclude that new chemical structures will be identified.



*Hvo*Trm112-Trm9 WT (2 pmol)	+	+	+	+		
*Hvo*Trm112-Hvo_0574 (2 pmol)					+	+
*Sc*tRNA *trm9Δ* (100 pmol)		+	+		+	
*Sc*tRNA *elp1Δ* (100 pmol)				+		
*Hvo*tRNA *trm9Δ* (100 pmol)	+					+
KCl (3 M)	+	+		+	+	+



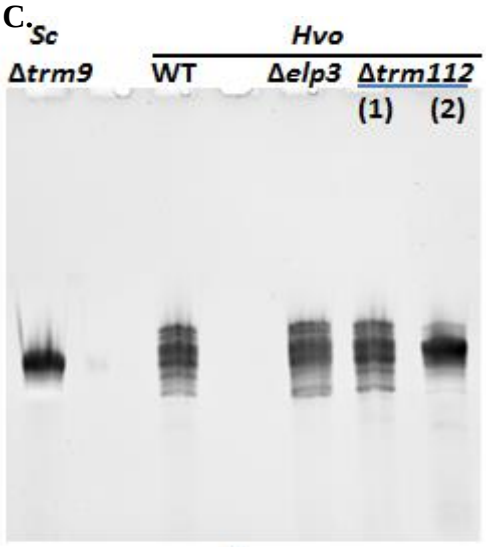


Figure 37. **Enzymatic assay for *Hvo*Trm112-Trm9 complex**. (**A)**. Enzymatic activities of *Hvo*Trm112-Trm9 complex. The conditions (proteins, salt and ligand) used for each experiments are indicated in the table below the graph. The number of methylated substrate (in pmol) after a 2 h reaction is indicated for every condition. Errors bars have been calculated from the results of three independent experiments. (**B)**. Total RNA purification from different *H. volcanii* strains by phenol-chloroform extraction; (**C)**. tRNA purification from different total RNAs by polyacrylamide-urea gel extraction.

4. Structural studies

4.1 Crystal structure of the *Hvo*Trm112-HVO_0019 complex

To gain insight into the molecular bases responsible for the interaction between *Hvo*Trm112 and its interacting MTase partners, different *Hvo*Trm112-MTase complexes were subjected to crystallization trials in the presence or in the absence of SAM. This was initially carried out by sitting drop vapor diffusion crystallization method in which 150 nL of the protein complexes (9-18mg/ml) with and without SAM was added to 150 nL of crystallization commercial crystallization screening kits (PEG Suit II, Crystal Screen 1 + 2, and JCSG) using 96-well TTP plate (Labtech) by Mosquito robot. The plates were carefully sealed and stored at 24^0C.

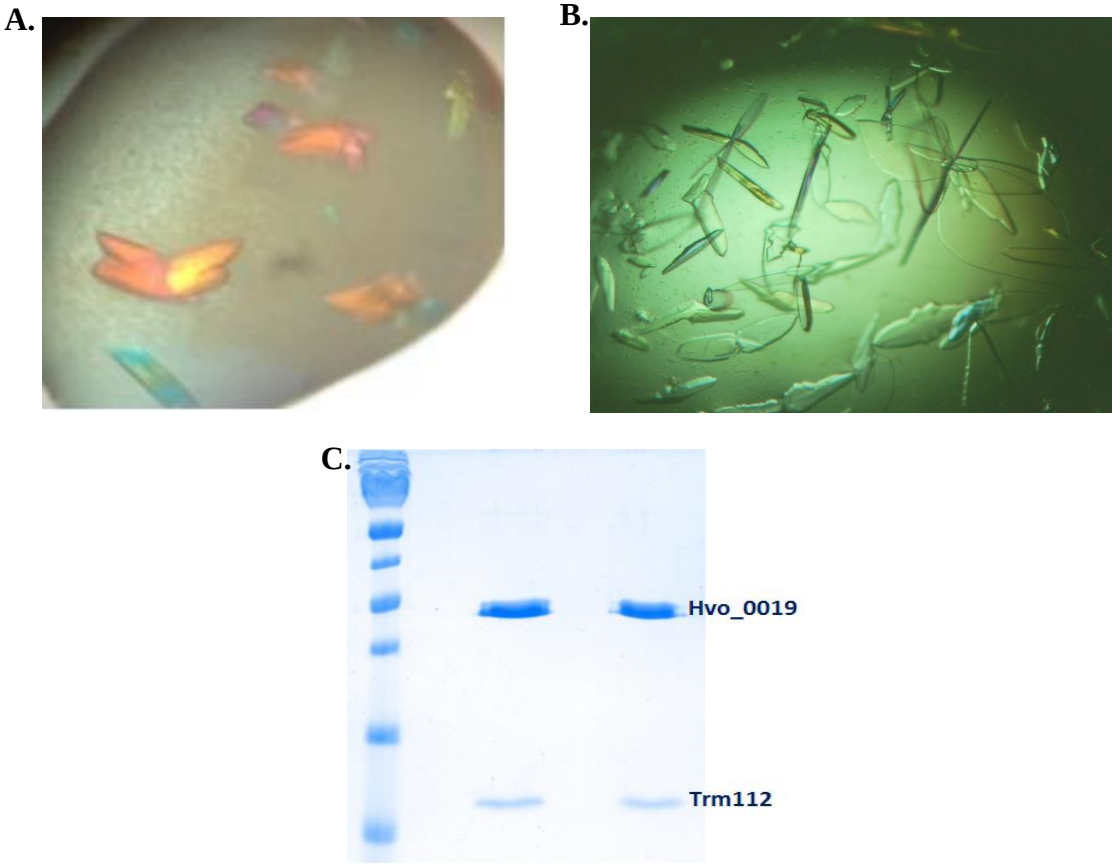


Figure 38. **HVO_0019-Trm112 crystallization**. (**A)**. Crystals of HVO_0019-Trm112 obtained from the sitting drop method (H10, JGSG: 0.2M Ammonium acetate, 0.1M BIS-TRIS pH5.5, 25% (w/v) PEG 3350) at 24^0C; (**B)**. Optimization of crystals of HVO_0019-Trm112 by hanging drop method at 4^0C (0.25M Ammonium acetate, 0.1M BIS-TRIS pH5.5, 25% (w/v) PEG 3350); (**C)**. Confirmation of the presence of HVO_0019-Trm112 complex in the crystals. Purified complex (middle lane) and dissolved crystals (right lane) were analyzed by SDS-PAGE.

In case of HVO_0019-Trm112 complex, small crystals appeared within few hours after setting up drops and then grew larger in the new few days in some conditions. To optimize the crystals, condition H10 (0.2M Ammonium acetate, 0.1M Bis-Tris pH 5,5, 25% w/v PEG3350) of the JCSG+ crystallization screen was selected for the hanging drop vapor diffusion crystallization method by applying hand-made gradients of ammonium acetate concentration (0.1 to 0.3M) and PEG3350 percentages (15 to 35%) in 24-well plates. Then 1 µL of the protein complex (with and without SAM) was mixed with 1 µL of the well solutions on a glass coverslip which was then inverted and sealed the well with the help of vacuum grease. The experiment was conducted both at 24^0C and 4^0C. The crystals were harvested, then cryo-protected by glycerol (15-30%) and finally flash-frozen in liquid nitrogen. The content of crystals was confirmed to contain HVO_0019 and *Hvo*Trm112 through collecting some crystals, followed by carefully washing them in the crystallization solution, then dissolving the crystals in water and finally running an SDS-PAGE 12% (Figure 38C).

All datasets were collected on beam-line Proxima-2A (Synchrotron SOLEIL, Saint-Aubin, France) at 100K. The structure was solved by Sulfur-SAD (Single Anomalous Dispersion) using a highly redundant dataset (Seven datasets of 1400° each collected on different regions of a single crystal) collected at 2.0664 Å to get the higher anomalous signal from sulfur atoms (see Table 5 for statistics). A 1.35 Å resolution dataset was obtained by merging two datasets collected at different crystal-detector distances on the same crystal (see Table S4 for statistics). Data were processed with XDS (Kabsch, 1993) and scaled using XSCALE.

- ***Crystal space group determination***

In crystals, a unit or motif repeated in three dimensions is known as the unit cell which can consist of multiple asymmetric units which are the minimal set of atoms (or ions or molecules) that create the whole content of the unit cell by applying the symmetry elements. The symmetry operators between the asymmetric units within unit cell are defined by the space group of the crystal. There are a possible total of 230 space groups, however only 65 ones are available for L-amino acid composed protein crystals, which can be divided in 14 Bravais lattices belonging to 7 crystal systems.

Data processing by XDS suggests different possible space groups based on parameters computed from the reflection data in the IDXREF step. Each possibility contains information about symmetry of the crystal and the penalty score (quality of fit). The most suitable space group is the case that is a combination of both the highest symmetry and the lowest penalty score. In case of HVO_0019-Trm112 crystal, possible space groups are shown in the following figure (Figure 39).

	LATTICE-CHARACTER	BRAVAIS-LATTICE	QUALITY OF FIT	UNIT CELL CONSTANTS (ANGSTROEM & DEGREES) a	b	c	alpha	beta	gamma
*	31	aP	0.0	80.6	82.3	88.0	90.0	89.9	90.0
*	44	aP	0.1	80.6	82.3	88.0	90.0	90.1	90.0
*	33	mP	0.6	80.6	82.3	88.0	90.0	90.1	90.0
*	35	mP	0.8	82.3	80.6	88.0	90.1	90.0	90.0
*	34	mP	1.1	80.6	88.0	82.3	90.0	90.0	90.1
*	32	oP	1.2	80.6	82.3	88.0	90.0	90.1	90.0
*	10	mC	21.9	115.3	115.2	88.0	90.0	90.1	91.2
*	14	mC	22.0	115.3	115.2	88.0	90.0	90.1	88.8
*	13	oC	22.8	115.3	115.2	88.0	90.0	90.1	88.8
*	11	tP	22.9	80.6	82.3	88.0	90.0	90.1	90.0
	20	mC	75.2	120.6	120.5	80.6	90.0	90.1	93.8
	25	mC	75.5	120.5	120.6	80.6	90.1	90.0	86.2
	23	oC	75.5	120.5	120.6	80.6	90.1	90.0	86.2
	21	tP	75.8	82.3	88.0	80.6	90.1	90.0	90.0
	2	hR	97.1	115.2	119.3	145.1	94.2	89.0	118.2

Figure 39. **Information on possible Bravais lattice of HVO_0019-Trm112 crystal derived from IDXREF step of the data processing**

From here, the Bravais lattice oP ($P2_n2_n2_n$, with n=1 for two-fold screw axis), which satisfies the criteria of both higher symmetry and lower penalty score is selected. In fact, when looking along each axis in reciprocal space namely (h,0,0), (0,k,0) and (0,0,l), we can see the signature of a two-fold screw axis for all axes that shows near zero intensity for the odd-numbered reflections (2n+1) while the even reflections (2n) have medium to strong intensities (Figure 40). Hence, the real space group of HVO_0019-Trm112 crystal is very likely to be $P2_12_12_1$.

H	K	L	RESOLUTION	INTENSITY	SIGMA	INTENSITY/SIGMA	#OBSERVED	
0	0	12	7.336	0.4829E+04	0.1051E+03	45.96	4	
0	0	13	6.772	-0.5141E+00	0.1519E+01	-0.34	4*	
0	0	14	6.288	0.3519E+03	0.1205E+02	29.21	2	⇨ 2
0	0	15	5.869	0.2253E+01	0.1781E+01	1.27	4*	
0	12	0	6.858	0.6155E+03	0.1379E+02	44.63	4	
0	13	0	6.330	0.1306E+01	0.1015E+01	1.29	4*	
0	14	0	5.878	0.2700E+04	0.5884E+02	45.89	4	⇨ 2
0	15	0	5.486	0.3443E+01	0.1354E+01	2.54	4*	
11	0	0	7.328	-0.2883E+01	0.1445E+01	-2.00	3*	
12	0	0	6.717	0.1690E+04	0.5251E+02	32.19	2	
13	0	0	6.201	-0.3775E+01	0.1587E+01	-2.38	3*	⇨ 2
14	0	0	5.758	0.9798E+03	0.2546E+02	38.49	3	

Figure 40. **Statistics on reflection intensity along each axis in reciprocal space (0,0,l), (0,k,0) and (h,0,0), taken from CORRECT.LP file.**

❖ ***Number of molecules in an asymmetric unit***

Next I estimated the number of molecules present in the asymmetric unit of the crystal. This can be done using the Matthews coefficient as the following formula:

where Vm is the Matthews coefficient (ranging between 1.66 to 4 corresponding to crystals with 30% to 75% solvent); Vcell, unit cell volume; Mw, molecular weight; Z, the number of asymmetric units in the unit cell; n, number of molecule in an asymmetric unit.

The number of copies of HVO_0019-Trm112 complex in the asymmetric unit can be manually calculated as the following:

Mw = *33244 Da;* Space group: $P2_12_12_1$ ➔ *Z = 4*

Cell parameters: *a = 80.6; b = 82.3; c = 88 and* $\alpha = \beta = \gamma = 90^0$

➔ Vcell = *80.6*82.3*88 = 583737* $Å^3$

So, Vm = *583737/(n*33244*4)*

If *n = 1*, Vm = *4.39* ➔ *too high*

If *n =2*, Vm = *2.19* ➔ *just right*

Hence, there are most likely two copies of HVO_0019-*Hvo*Trm112 complex in the asymmetric unit.

❖ ***Structure solving solution***

Crystals diffracting up to 1.35Å resolution were obtained for the *Hvo*Trm112-HVO_0019 complex. The structure of this complex was solved at 2.5Å resolution by the Sulfur-SAD method by taking advantage of the high number of sulfur atoms in this complex (11, excluding the initial methionine from HVO_0019, which is very likely to be excised in *E. coli*; (Hirel et al., 1989)). Twenty-four sulfur sites (eleven per HVO_0019-Trm112 complex and two from S-adenosyl-L-homocysteine (SAH) molecules bound to HVO_0019 methyltransferase) were successfully located using SHELXD (Schneider, & Sheldrick, 2002). Experimental phasing followed by density modification were performed with the PHASER_EP and RESOLVE programs implemented in the Phenix program (T. Terwilliger, 2004; McCoy et al., 2007; T. C. Terwilliger et al., 2008; Adams et al., 2010) (Figure 41). A first model was obtained by iterative cycles of building and refinement performed using COOT (Emsley et al., 2010) and BUSTER (Bricogne et al., 2016) programs, respectively (see Table 5). The structure was further refined at high resolution using the 1.35Å resolution native dataset. The final structure model was refined to 1.35Å with R and R_{free} values of 18.5% and 20.7%, respectively. The final model for HVO_0019-Trm112 complex contains HVO_0019 residues 2-227 (including the first histidine from the His_6-tag) and 2-12, 15-231 (including the five histidine residues from the His_6-tag) for protomers A and B, respectively as well as *Hvo*Trm112 residues 1-59 and 1-25, 29-58 for protomers C and D. In addition, 387 water molecules, two SAH molecules and two glycerol molecules from the cryoprotectant have also been modeled. There are two virtually identical copies of the *Hvo*Trm112-HVO_0019 complex in the asymmetric unit (Figure 42A; rmsd value of 0.27Å over 211 Cα atoms). Each complex is bound to a SAH molecule, most probably co-purified with *Hvo*Trm112-HVO_0019 complex.

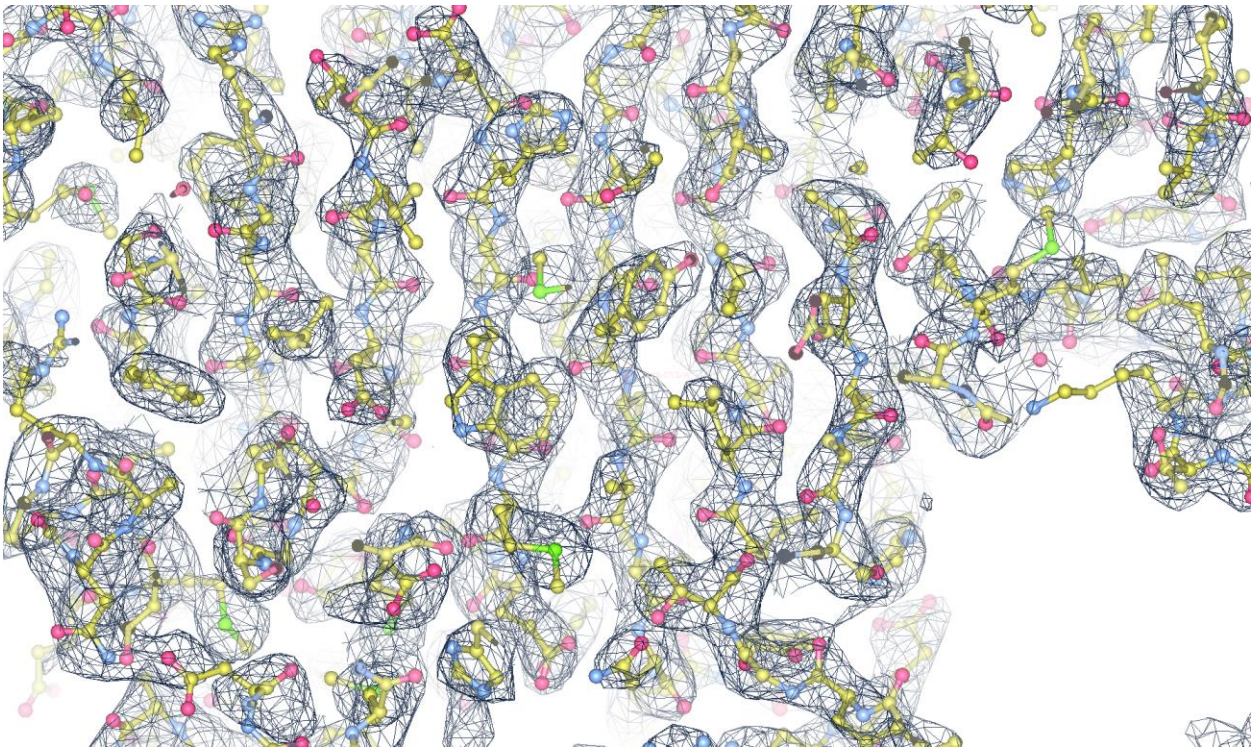


Figure 41. **Experimental electron density map calculated at 2.5Å resolution by Sulfur-SAD.**

The experimental electron density map is contoured at 1 sigma. The final model (refined at 1.35Å resolution) is shown as yellow sticks.

Table 5. **Data collection, phasing and refinement statistics**

Data collection	**S-SAD**	**High resolution**
Space group	$P2_12_12_1$	$P2_12_12_1$
Unit cell parameters		
a, *b*, *c* (Å)	77.35, 89.13, 89.64	80.6, 82.3, 88.0
α, ß, γ (°)	90.00	90.00
Wavelength	2.0664	0.98007
Resolution (Å)	50-2.5 (2.57-2.5)	50-1.35 (1.39-1.35)
R_{merge}	10.6 (66.7)	8.4 (172.8)
$I / \sigma I$	62.0 (5)	14.4 (1.3)
Completeness (%)	98.8 (86.8)	100 (100)
Redundancy	174	12.3
Observed reflections	7117251	1586836
Unique reflections	40910	128740
Refinement		
Resolution (Å)		48.19-1.35
No. reflections		128738
R / R_{free}		18.5/20.7
No. atoms		
Protein		4529
Ligands (Glycerol/SAH)		64
Water		387
B-factors (Å^2)		
Protein		25.6
SAH		17.2
Glycerol		59.4
Water		32.6
R.m.s deviations		
Bond lengths (Å)		0.01
Bond angles (°)		0.98

❖ ***Structure description***

Structurally, HVO_0019 adopts the classical class-I SAM-dependent MTase composed of a central seven-stranded β-sheet surrounded by three α-helices on one side and two on the other side (Figure 42B). On the top of the central β-sheet, four α-helices (αY, the N-terminal half of αZ, α1 and α2) contribute to the formation of a cavity centered on the sulfur atom from the SAH molecule bound to the HVO_0019 protein. The SAH binds to HVO_0019 in a canonical manner compared to class I SAM-dependent MTases. *Hvo*Trm112 structure consists of only the zinc finger-like domain previously observed in eukaryotic Trm112 proteins (rmsd values of 1.1-1.4Å over around 55 Cα atoms; (Heurgue-Hamard et al., 2006; Liger et al., 2011; Letoquart et al., 2014; Letoquart et al., 2015a)), including a small N-terminal α-helix (α1) and a four-stranded anti-parallel β-sheet (Figure 42B). Contrary to eukaryotic Trm112 proteins and some structurally similar bacterial proteins (PDB code: 2KPI) of known structures, there is no zinc bound to *Hvo*Trm112 in agreement with the lack of conservation of the cysteine residues involved in zinc coordination. The interface between *Hvo*Trm112 and HVO_0019 has an overall area of 975 $Å^2$ and is slightly smaller than the previously described interfaces between eukaryotic Trm112 and its MTase partners (Letoquart et al., 2014; Letoquart et al., 2015a). This most likely results from the absence of the helical domain, specific to eukaryotic Trm112 proteins, that was contributing to these interactions. Complex formation involves 20 and 21 residues from *Hvo*Trm112 and HVO_0019, respectively. The core of this interface is formed by hydrophobic residues (M1, L5, I8, L9, P12, I50, P51, L53, L54, P55 and M58 from *Hvo*Trm112 and F66, F89, L90, V91, L97, P98 and F99 from HVO_0019; Figure 43). Such a large hydrophobic interface rationalizes the strong solubilization effect that we observe for *Hvo*Trm112 upon co-expression with HVO_0019 in *E. coli* (Figure 28A). This hydrophobic core is surrounded by polar residues (K2, D7, C10, K15, E29, N52 and R59 from *Hvo*Trm112 and S2, R64, Q76, R79, D85, D86, S88, D93, D96, D100, S103, E125 and R128 from HVO_0019). Six hydrogen bonds and three salt bridges are also observed at the interface (Table 6). Two hydrogen bonds formed between V91 main chain atoms from HVO_0019 and P51 as well as L53 from *Hvo*Trm112, are responsible for the formation of a β-zipper interaction between *Hvo*Trm112 strand β4 and HVO_0019 strand β3 (Figure 42B), which is also observed in the eukaryotic MTase-Trm112 complexes. Other hydrogen bonds are formed between P98, L97, R64 and D100 from HVO_0019 and C10, K15,

S4 and I8 from *Hvo*Trm112, respectively. Finally, salt bridges are formed by K15 from *Hvo*Trm112 with D96 and E125 from HVO_0019, and by D7 from *Hvo*Trm112 with R64 from HVO_0019.

Table 6. **Details of hydrogen bonds and salt bridges involved on *Hvo*Trm112-HVO_0019 interaction**

Hydrogen bonds	
***Hvo*Trm112**	**HVO_0019**
Cys10 N	Pro 98 O
Lys15 NZ	Leu97 O
Leu53 N	Val91 O
Ser4 O	Arg64 NH1
Ile8 O	Asp100 N
Pro51 O	Val91 N
Salt bridges	
Lys15	Asp96
Lys15	Glu125
Asp7	Arg64

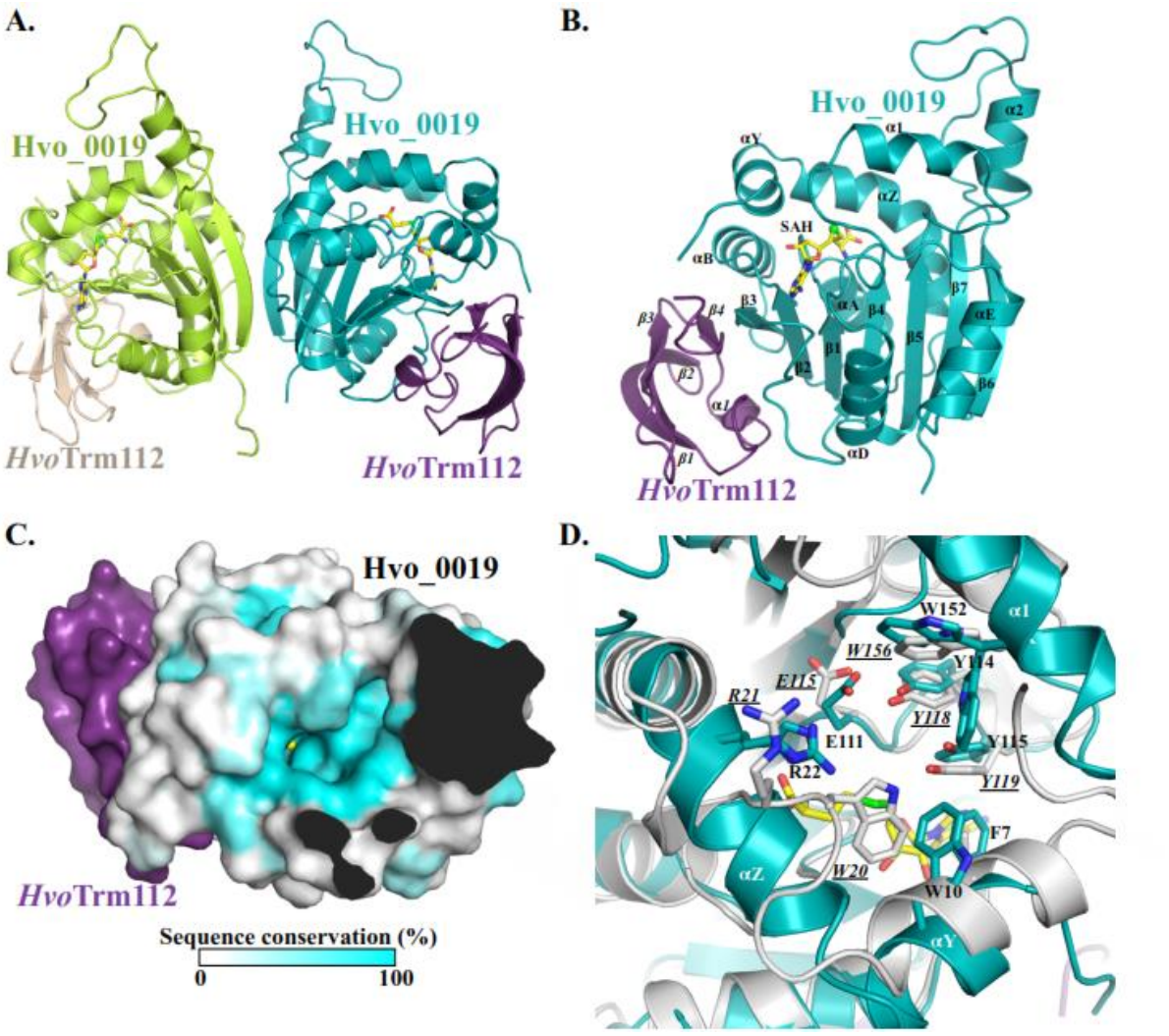


Figure 42. **Crystal structure of *Hvo*Trm112-HVO_0019 complex.**

A. Ribbon representation of the heterotetrameric *Hvo*Trm112-HVO_0019 complex. The SAH molecule bound to each HVO_0019 monomer is shown as yellow sticks.

B. Ribbon representation of the *Hvo*Trm112-HVO_0019 heterodimer.

C. Sequence conservation mapped at the surface of HVO_0019 protein structure. Only the heterodimer is shown for the sake of clarity. Conservation scores have been calculated from an alignment of 241 sequences using the Consurf server (Ashkenazy et al., 2010). The SAM methyl group modeled by superimposing the structure of *S. cerevisiae* Bud23 (Letoquart et al., 2014) onto the structure of HVO_0019 is shown as a yellow sphere.

D. Comparison of HVO_0019 (blue) and NodS (grey) active sites with conserved residues shown in sticks. Labels for NodS residues are underlined and in italics. SAH is shown as yellow sticks. It is noteworthy that W152 from HVO_0019 adopts an alternate conformation in the crystal structure.

As stated above, there are two copies of *Hvo*Trm112-HVO_0019 complex in the asymmetric unit and these are related by a two-fold symmetry axis (Figure 42A). As SEC-MALLS measurements on this complex have revealed that it forms heterotetramer in solution, this contact area most likely corresponds to the biological interface responsible for this oligomeric state. This homodimerization interface (area of 840 $Å^2$) is exclusively formed by residues from HVO_0019, more precisely from strands β6 and β7 and helix αZ. Hence, the MTase partner seems to rule the formation of oligomeric states, thereby explaining that depending on its MTase partner, *Hvo*Trm112 may exist either as heterodimers, heterotetramers or heterohexamers according to our SEC-MALLS analyses (Table 4).

- ***Structure comparison***

BLAST searches to identify proteins sharing sequence similarity with HVO_0019 identified orthologous proteins mostly from specific archaeal phyla such as halobacteriales and thaumarcheota but also from some bacteria (*Legionella shakespearei*, *Agrobacterium tumefaciens*, *rhodobacteriaceae*, *flavobacteriaceae*, *acidobacteria*). Multiple sequence alignments revealed the presence of few highly conserved residues (Figure 43B). Interestingly, in HVO_0019 structure, these residues cluster around the expected position of the SAM methyl group and form a strongly conserved pocket (Figure 42C). Comparison of HVO_0019 protein structure and active site with previously described structures of MTases revealed a significant degree of conservation with the MTase NodS from *Bradyrhizobium japonicum* (Cakici et al., 2010). NodS is involved in the biosynthesis of the Nod factor, a modified chitosaccharide acting as a signal molecule in rhizobia, by catalyzing the methylation of the NH_2 group of its glucosamine moiety. Indeed, residues R22, E111, Y114, Y115 and W152 from HVO_0019 structurally match with R21, E115, Y118, Y119 and W156 from NodS, respectively (Figure 42D). In addition, the side chains from HVO_0019 W10 and NodS W20 are also in close vicinity. Altogether, this suggests that HVO_0019 might bind a substrate with an hexose sugar ring and modify it. Further studies will be needed to characterize the enzymatic activity of the *Hvo*Trm112-HVO_0019 complex.

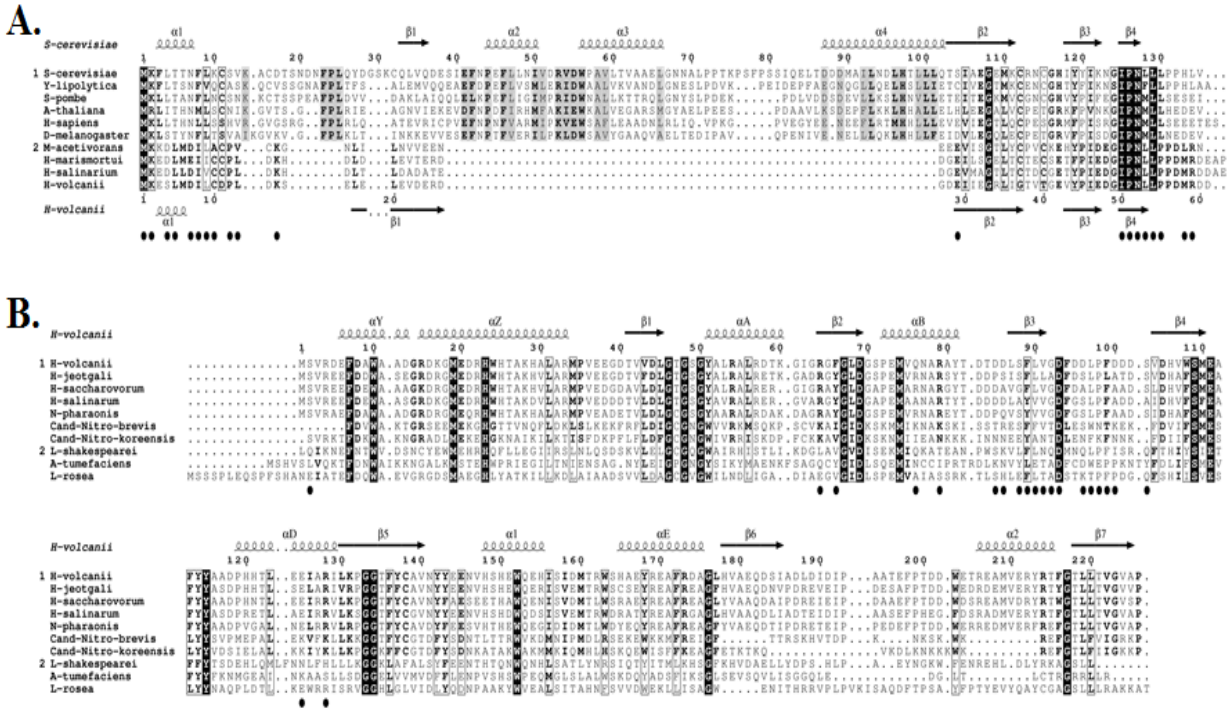


Figure 43. **Sequence alignment of *Hvo*Trm112 and HVO_0019 sequences.**

A. Sequence alignments of Trm112 orthologs from eukaryotes (group 1) and archaea (group 2). Strictly conserved residues are in white on a black background. Partially conserved amino acids are highlighted with a grey background. Secondary structure elements assigned from the *S. cerevisiae* Trm112 (Letoquart et al., 2014) and *Hvo*Trm112 crystal structures are indicated above and below the alignment, respectively. Black closed circles indicate residues involved in interaction with HVO_0019. Both panels were generated using the ESPript server (Robert, & Gouet, 2014).

B. Sequence alignments of HVO_0019 orthologs from archaea (group 1) and bacteria (group 2). Strictly conserved residues are in white on a black background. Partially conserved amino acids are boxed. Secondary structure elements assigned from the HVO_0019 crystal structure are indicated above the alignment. Black closed circles indicate residues involved in interaction with HVO_0019.

4.2 Other HvoMTase-Trm112 complexes

In addition to the determined crystal structure of *Hvo*_0019-Trm112, I have also obtained crystals for two other *Hvo* MTase-Trm112 complexes. Crystals of *Hvo*_0475-Trm112 diffracted up to 3 Å while in case of *Hvo*_0773-Trm112, tiny crystals diffracting to 7 Å were observed. Because the functions of these complexes are not known till now, efforts on solving their crystal structure are worth trying to get preliminary insights into their putative functions. In order to accomplish this, production and purification of selenomethionine-labeled *Hvo*_0475-Trm112 complex will be helpful for solving the complex structure. Meanwhile, for *Hvo*_0773-Trm112 complex, several optimization steps through hanging drop method are needed to improve the crystal size to get better diffraction.

4.3 MTase-Trm112 in complex with its substrate

Although several crystal structures of MTase-Trm112 complexes in different organisms have been solved, no crystal structures have been so far achieved for the protein complexes bound to their substrates. Knowing the fact that the enzyme-substrate complex structure is very useful in deciphering the substrate recognition and modification mechanism of MTase-Trm112 complexes, I have made several efforts on reconstituting and trying to crystallize the complex between Mtq2-Trm112 and its substrate aRF1-aRF3 in different organisms (*H. volcanii*, *Aeropyrum pernix* and *Archaeoglobus fulgidus*) during my thesis. However, until now I was only able to purify and reconstitute the complex of four proteins in presence of SAM and GTP from the latter organism (Figure 44). The crystallization trial has been performed but no crystals were obtained. Hence, this exciting structural study still remains to be solved.

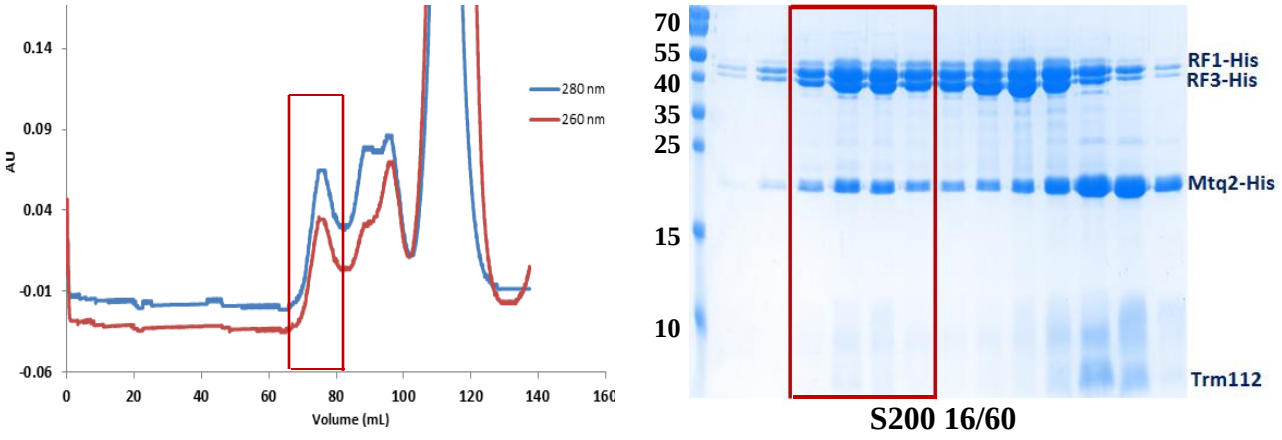


Figure 44. **Reconstitution of *A. fulgidus* Mtq2-Trm112-aRF1-aRF3 complex in presence of SAM and GTP**. *A. fulgidus* Mtq2-Trm112, aRF1 and aRF3 were first purified through 3-steps chromatography. The complex formation was performed by mixing 6 mg of each protein in presence of 5mM GTP + 5mM $MgCl_2$, followed by o/n incubation. The next day, SAM in excess was added to the mix, and then incubated at room temperature for 3hours. The mix was next run on S200 16/60 and different fractions were analyzed on SDS-PAGE 12%. The peak corresponding to quaternary complex was labelled by a box on both the chromatogram and the gel.

CHAPTER III

HUMAN METTL5-TRMT112 COMPLEX

INTRODUCTION

Sequence analysis through BLAST searches has been performed for putative newly identified *H. volcanii* Trm112 methyltransferase partners originated from our Co-IP experiments. For HVO_1475 MTase, we were able to identify a human ortholog – the METTL5 methyltransferase, which is surprisingly absent in *Saccharomyces cerevisiae*, the eukaryotic model organism.

This is of particular interest as it may lead to the description of a new Trm112 MTase partner in human. So far, our knowledge on human TRMT112 network was mainly based on the yeast counterpart. Indeed, TRMT112 has been shown to physically and directly interact with 4 human MTases, namely TRMT11 (GB and MG, unpublished results), ABH8 (D. Fu et al., 2010; Y. Fu et al., 2010), HEMK2 (Figaro et al., 2008) and WBSCR22 (Ounap et al., 2015; Zorbas et al., 2015), which are respectively orthologs of Trm11, Trm9, Mtq2 and Bud23 in yeast. As HVO_1475 is found in the list of putative *Hvo*Trm112 partners, we checked if these proteins indeed form a complex both in *H. volcanii* but also in human by co-expression and co-purification in *E. coli*.

RESULTS

1. HVO_1475 is another MTase partner of *Hvo*Trm112 in *H. volcanii*

As described for several *Hvo*Trm112 MTase partners in chapter 2, I characterized another putative *Hvo*Trm112 partner, HVO_1475 by means of co-expression and co-purification of these two proteins in *E.coli*. First of all, the HVO_1475 gene sequence was cloned similarly to other *Hvo*Trm112 MTase partners (as described in details in chapter 2) so as to be express a C-terminally His-tagged version of HVO_1475 together with untagged *Hvo*Trm112. First, expression of HVO_1475 alone in *E. coli* led to the protein in an insoluble form while upon co-expression with *Hvo*Trm112, the MTase was expressed as a soluble protein and with higher yield (Figure 45A). This is a common trend observed for several Trm112 MTase partners in yeast and *H. volcanii*, which strictly require Trm112 for their solubility. Next, the HVO_1475-*Hvo*Trm112 complex was expressed and purified through 3-steps chromatography purification (Ni-NTA, anion exchange and size-exclusion) (Figure 45B-D). Since the theoretical isoelectric point (pI) of

HVO_1475-Trm112 is around 4.4, it is supposed to be negatively charged at the purification buffer pH 7.5 and therefore it should bind to positively charged MonoQ column. The result of ion exchange chromatography then showed the presence of four main peaks (Figure 45C). The first large peak came before NaCl gradient and corresponds to the flow-through fraction, indicating no binding of a large proportion of the complex to the column. The fourth peak elutes at very high NaCl concentration and its UV signal is higher at 260 nm than at 280 nm, indicating that it corresponds most likely to nucleic acids. The second and third peaks were eluted at around 260mM and 450mM NaCl, respectively. Through SDS-PAGE these peaks were confirmed to contain HVO_1475 and Trm112 proteins. However, the asymmetric distribution of the proteins along peak 2 suggests that both proteins do not interact together in this peak (Figure 45C). The dissociation of the complex in this second peak could be due to the low salt concentration. Indeed, to perform the ion exchange chromatography, we had to lower the concentration of NaCl to nearly 100mM for the binding to the MonoQ column. We have observed such complex dissociation at low salt concentration for other *Hvo*Trm112-MTase complexes. Hence, HVO_1475-Trm112 was probably slowly dissociated at 100mM NaCl. Thanks to short-time protocol of ion exchange chromatography, some complexes were still there. As a result, three different species (the protein complex, HVO_1475 and *Hvo*Trm112 alone) were retained on the MonoQ column and eluted at different salt concentration, therefore leading to the formation of two peaks corresponding to proteins alone and the protein complex. The fractions corresponding to the protein complex, i.e. peak 3, were then collected, injected to gel filtration at 1M NaCl and finally eluted in a purer stage. The identities of HVO_1475 and *Hvo*Trm112 were then confirmed by MS. SEC-MALLS analysis yielded a measured molecular weight of around 29.9 kDa which is very similar to the theoretical molecular weight of the HVO_1475-HvoTrm112 complex (30.05kDa) (Figure 45E). Altogether, this indicates that these two proteins interact together and exist as an heterodimer in solution.

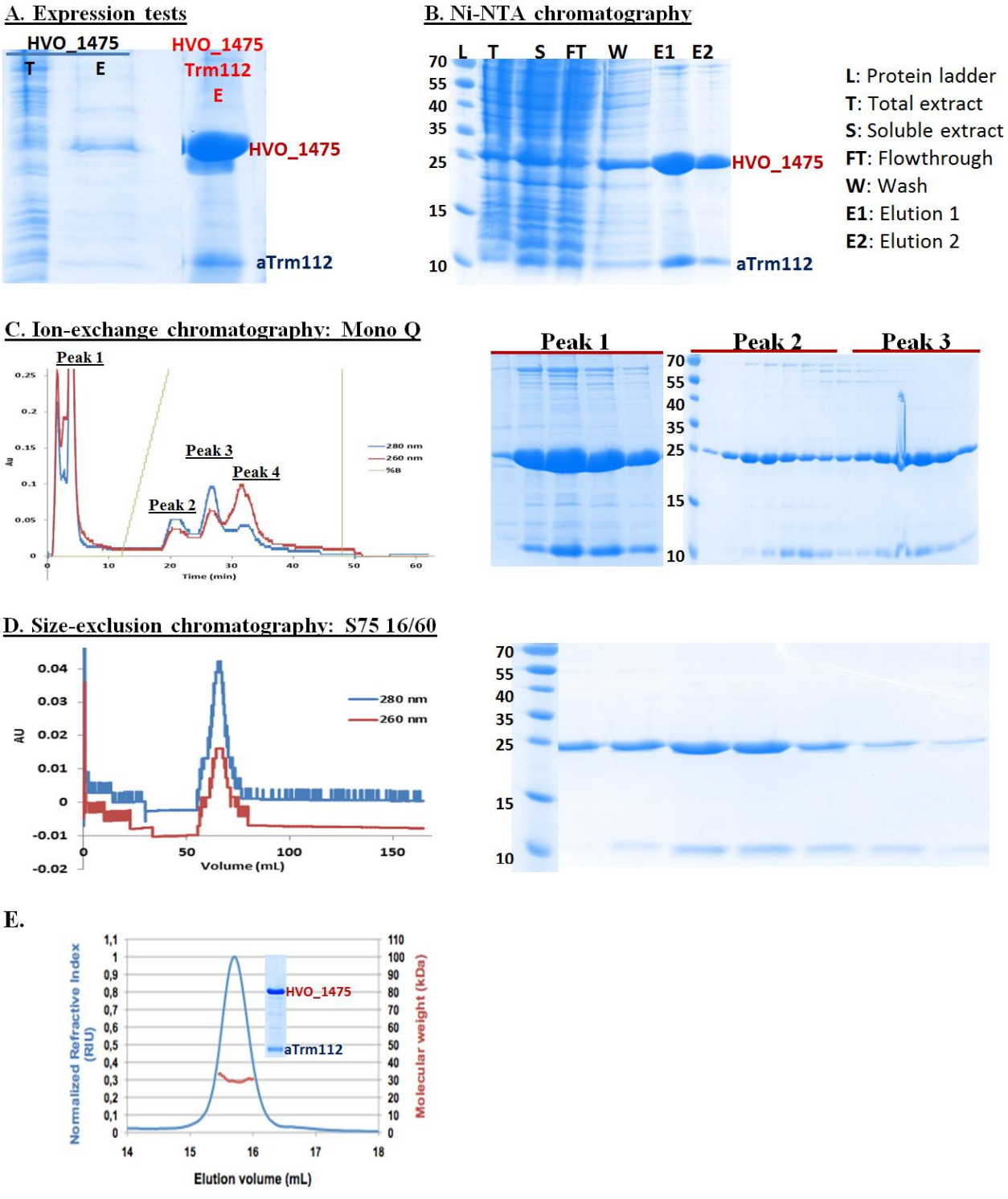


Figure 45. **HVO_1475-Trm112 complex**. (**A**). Expression tests of HVO_1475 in case of absence (band 1-2 from the left. T: total extract and E: elution) and presence (band 3. E: elution fraction) of *Hvo*Trm112.(**B-D**). 3-steps chromatography purification. (**E**). SEC-MALLS. The proteins were run on 12% SDS-PAGE gel.

2. Human METTL5 forms heterodimer with hTRMT112

Blast search analysis of the *H. volcanii* HVO_1475 protein sequence against all organism proteins led to the identification of several putative orthologs including the human putative methyltransferase METTL5 (E-value of $3e^{-30}$). Since I have demonstrated that HVO_1475 forms a complex with *Hvo*Trm112, I also tested whether its ortholog, the human METTL5, interacts with human TRMT112 or not. In doing so, a DNA sequence encoding human METTL5 fused to a His-tag at its C-terminal extremity was initially codon-optimized *in silico* and synthesized *de novo* (Integrated DNA Technologies, USA) for heterologous expression in *E. coli*. This synthetic DNA fragment was digested by NdeI and NotI and ligated into NdeI-NotI digested pET21a. The plasmids containing the target DNAs were transformed into competent *E. coli* XL1 Blue by the heat-shock transformation method. The transformants were then streaked on LB agar plate containing ampicillin (100µg/mL), followed by o/n incubation at 37^0C. Some colonies for each clone were selected to inoculate 5mL LB supplemented with ampicillin (100µg/mL) and then grown o/n at 37^0C. The plasmids were then extracted by MiniPrep kit and finally verified by sequencing. The plasmid encoding for untagged human TRMT112 is a kind gift from our collaborator Dr Valérie Heurgué-Hamard (IBPC, CNRS, Paris). I then co-expressed and co-purified these two human proteins in *E. coli*. First, small scale expression tests at different induction temperatures (37, 30 and 18^0C) indicated that the best expression yields are observed in *E. coli* Gold strain with auto inducible TBAI media at 18^0C overnight (Figure 46A). Similarly to several MTases in both yeast and *H. volcanii*, human METTL5 was also shown to strictly require TRMT112 for its solubility (Figure 46B).

The complex was then co-expressed in 1 L using the optimal conditions identified from the expression assay and co-purified through three steps chromatography purification as described in the method part. After the Ni-NTA affinity chromatography step, two main bands, migrating at the expected molecular weights for METTL5 and TRMT112, are present on the SDS-PAGE analysis of the elution fraction (Figure 46C). Next, as the theoretical pI of METTL5-TRMT112 complex is nearly 6, indicating that the complex should be negatively charged at pH 7.5, MonoQ anion exchange chromatography was therefore chosen. The main peak eluting at around 300mM NaCl also contains mostly those two proteins (Figure 46D) and it was injected on size exclusion chromatography (S75 Superdex 16/60). Analysis of the content of the single

peak from SEC also revealed the presence of two main bands (Figure 46E), which were shown to be METTL5 and TRMT112 by MS. SEC-MALLS analysis indicated a molecular weight of 41.6kDa, slightly different from the theoretical molecular weight of the complex (38.7kDa) (Figure 46F). Altogether, this indicates that the human METTL5 and TRMT112 proteins form a stable complex that exists as a heterodimer in solution. Hence, we have been able to identify a previously unanticipated partner of human TRMT112 from our initial study performed in the *H. volcanii* archeon.

3. Structural studies

In spite of being predicted as putative DNA/RNA methyltransferase in databases, the real functions of both HVO_1475 and its human ortholog METTL5 are still open questions. Taking advantage of possibility to obtain high amount of both purified protein complexes, we attempted to crystallize them with the hope to solve their crystal structures, which may give preliminary clues about their biological functions. Due to lack of time, I have so far tried only one crystallization trial for each complex. Unfortunately, no crystals have been so far obtained for both complexes.

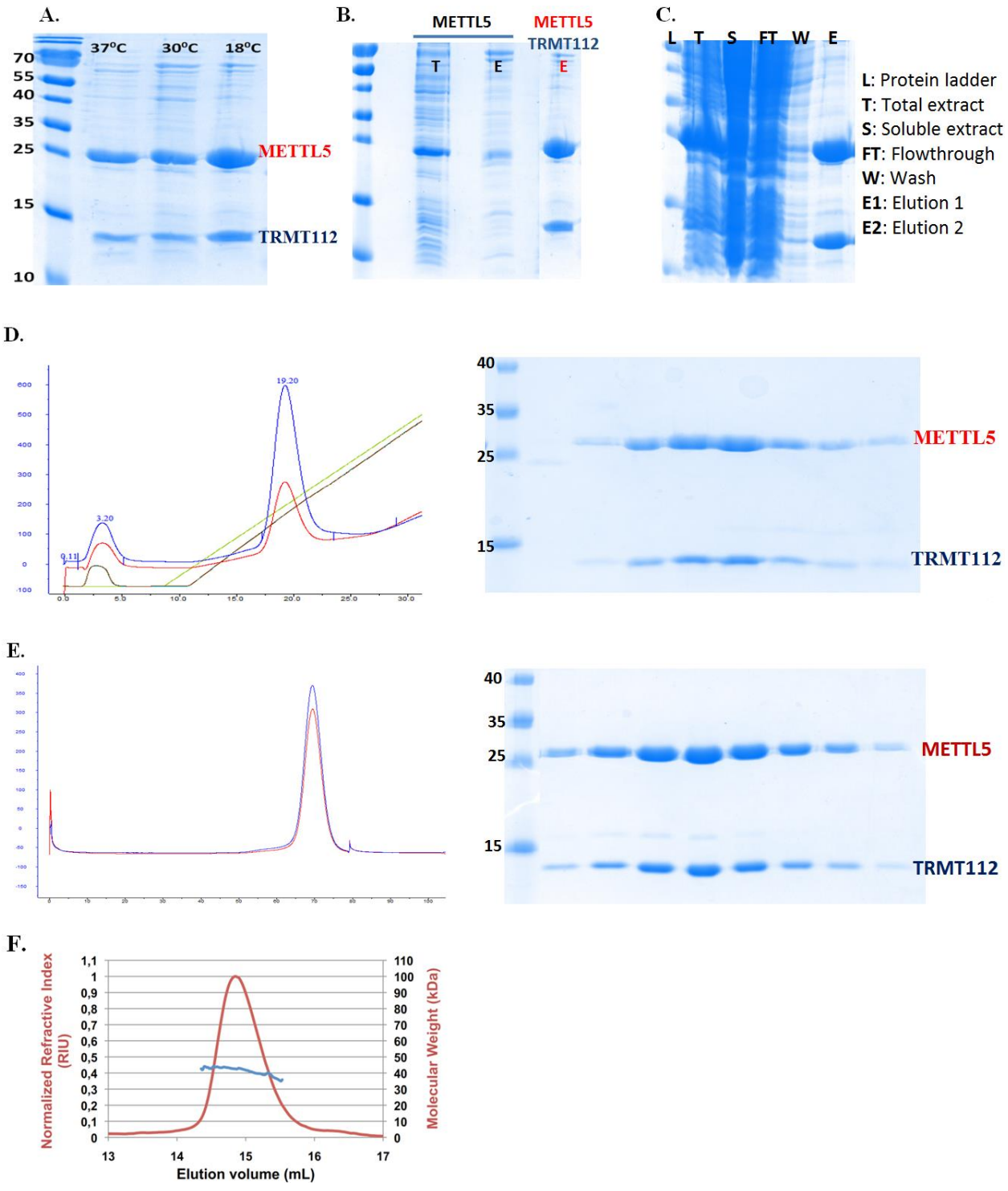


Figure 46. **Human METTL4-TRMT112 complex**. (**A).** Expression tests at different temperature (18, 30, and 37^{0}C). **(B)**. Expression tests of METTL5 in case of absence (band 2-3 from the left. T: total extract and E: elution) and presence (band 4. E: elution fraction) of TRMT112. (**C-E**) 3-step chromatography purification: Ni-NTA, Ion exchange (MonoQ) and gel filtration (S75 16/60), respectively. (**F**). SEC-MALLS. The proteins were run on 12% SDS-PAGE gel.

4. Putative biological function of human METTL5

Thanks to Trm112 interacting network in *H.volcanii*, we have so far extended our knowledge on the corresponding counterpart in human by characterizing one more MTase partner – METTL5, which ortholog is surprisingly not present in *S. cerevisiae*.

Grosjean et al, 2008 predicted that HVO_1475 could be the MTase responsible for m^6A modification at position 1432 in 16S rRNA. This modification is also found at the corresponding position from small ribosomal rRNAs in other archaea as well as in metazoan including human (Maden, 1986, Grosjean et al, 2008) but no m^6A has been reported in yeast 18S rRNA so far. In human, the MTase responsible for this modification has not been identified yet (Maden et al., 1986). This is of particular interest as bioinformatic analyses strongly suggest that the human METTL5 might be responsible for the methylation of ribosomal RNA. Hence, we postulate that METTL5, already annotated as RNA/DNA methyltransferase, is responsible for the m^6A formation in 18S rRNA in human and we were are currently testing this hypothesis in collaboration with Pr Denis Lafontaine (Belgium). If this hypothesis is true, this would enlarge the TRMT112 interaction network in human by including another MTase modifying an actor of translation, *i.e.* the 18S rRNA. Although no clear link has been established between defects in METTL5 and human disease, it is well-known that base methylations in rRNA affect the ribosome biogenesis, which defects normally link to human diseases (see Introduction I, 2.2.2.3). It is noteworthy that the four already-known MTase partners of human TRMT112 (TRM11, ABH8, HEMK2 and WBSCR22) have well established implications in human diseases (see Introduction II, 2.2).

DISCUSSION AND CONCLUSION

I. DISCUSSION

Protein translation is one of the most intricate processes in cell biology and requires highly numerous factors in a greatly-coordinated manner. Most of these translational factors such as tRNA, rRNA and proteins are frequently subjected to post transcriptional and post-translational modifications to perform correct functions. Methylation is so far known as the most prevalent modification. In yeast, effects of methylation on translation are perfectly illustrated by Trm112, an obligate activator for at least four methyltransferases, acting on tRNAs (Trm9 (Letoquart et al., 2015b) and Trm11 (Purushothaman et al., 2005; Bourgeois et al., 2017b)), 18S rRNA (Bud23, (Letoquart et al., 2014)) and translation termination factor eRF1 (Mtq2, (Liger et al., 2011)). Therefore, this Trm112 interacting network is at the heart of ribosome biogenesis and function. The importance of this interacting network is also reflected by its conservation in human where these MTases are associated with diseases (Bourgeois et al., 2017a). Through sequence analysis, Trm112 ortholog is also found in archaea, raising the question of the availability of the orthologous interacting network in this domain of life since it is known that most of the translational machineries are very similar between eukarya and archaea.

In the thesis, the main work was to identify the Trm112 interacting network in archaea using *H. volcanii* as a model organism, followed by its functional and structural characterizations. In addition, I have also had opportunities to perform studies on human and yeast MTase-Trm112 complexes. All of these will be discussed in the following parts.

1. Mechanism of *Sc*Trm9-Trm112

In yeast, the *Sc*Trm9-Trm112 complex is known to convert cm^5U into mcm^5U at position 34 in the anti-codon loop from different tRNAs, directly affecting the translation fidelity (Kalhor, & Clarke, 2003; Mazauric et al., 2010). In order to measure the enzymatic activity, the substrate tRNAs have to contain cm^5U34, which can be obtained through *in vivo* purification from *trm9*-deleted yeast based on phenol-chloroform extraction. Since the enzyme substrates used in this case are total RNAs, which in addition to tRNA substrates contain non-substrate RNAs (other tRNAs as well as rRNAs (Figure 37B, lane 1), and potentially tRNAs containing ncm^5U34, which is not the substrate for Trm9-Trm112 (Chen et al., 2011)), these contaminants can develop inhibiting effects on *Sc*Trm9-Trm112 activity, leading to the determination of only apparent kinetics parameters (initial velocity, specific activity, *Km* and *kcat*). Several conserved residues centered on the SAM methyl group were selected for mutations in order to examine their effects *in vivo* and on *in vitro* enzymatic activity. Results showed a good agreement with the zymocin-resistant phenotypes as some mutants (W145A, H116A and W168A) were nearly inactive while the remaining (R29A, H115A, R241A, Y243A and N271A) impaired the enzymatic activity to different extent. As a SAM-dependent methyltransferase modifying an oxygen atom, the catalytic mechanism of *Sc*Trm9-Trm112 is thought to be the conventional S_N2 reaction. For this reaction to occur, a linear orientation of the nucleophile (cm^5U), the methyl carbon atom (CH_3), and the thioester leaving group (C-S-C) in the transition state of the reaction is required. In fact, several mutants (R29A, H115A, R241A, Y243A and N271A) were kinetically shown to strongly affect the binding affinity for tRNA (*Km* higher) but not for SAM, as well as the enzymatic activity (*kcat* lower) compared to those for the *Sc*Trm9 wild type. This supports a model where those residues play a role in locking tRNA substrate in the active site pocket through an intricate hydrogen bonds network, so as to bring cm^5U34 into close proximity to the sulfonium group of the SAM for the methylation to be catalyzed (Figure 47).

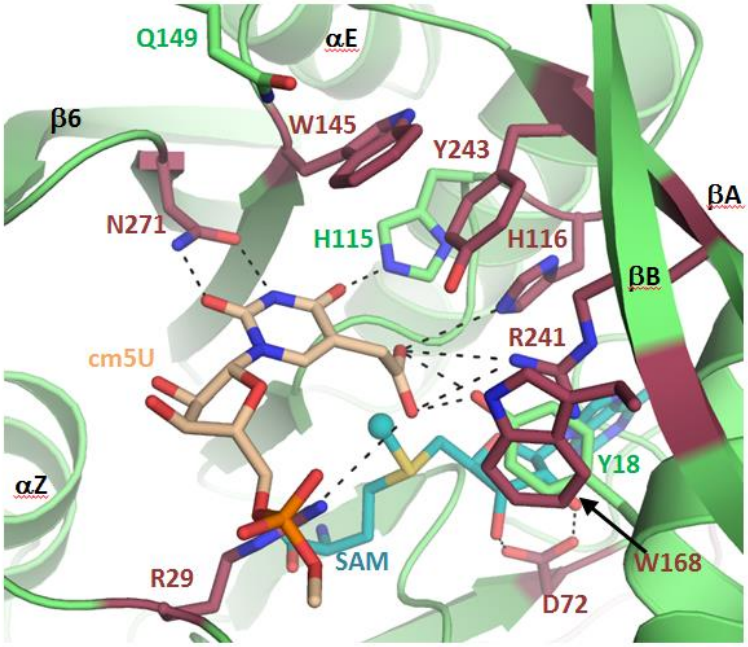


Figure 47. **Docking model of cm^{5}U into the active site of *Y. lipolytica* Trm9**. The residues are numbered based on the *S. cerevisiae* protein. A manually docked cm^{5}U nucleotide is shown as beige sticks and potential hydrogen bonds that it could form with Trm9 active site residues are depicted by dashed black lines. The modeled SAM molecule is shown as blue sticks and the methyl group to be transferred is shown as a sphere (Letoquart et al., 2015b).

2. *H. volcanii* Trm112 interacting network

2.1. More MTase partners of Trm112 are present in archaea than in yeast

The Trm112 interacting network was initially deciphered in yeast, where Trm112 acts as a hub protein interacting with and activating at least four different SAM-dependent methyltransferases involved in different facets of protein translation. With Trm9 or Trm11, the complex is responsible for tRNA methylation (Letoquart et al., 2015b; Bourgeois et al., 2017b) while when bound to Mtq2, the complex is involved in methylation of translation termination factor eRF1 (Liger et al., 2011). The last known partner is Bud23, which participates in 18S rRNA methylation (Figaro et al., 2012; Letoquart et al., 2014). Meanwhile, although the presence of Trm112, Trm9, Mtq2, and Trm11 orthologs in many archaea have been previously identified by bioinformatic analysis, no information is available whether the corresponding

interacting network exists in archaea. By taking advantage of recent advances in archaeal genetics and highly sensitive protein detection method based on liquid chromatography coupled to tandem mass spectroscopy (LC-MS/MS), we were able to detect several putative interacting partners of Trm112 in *H. volcanii* using co-immunoprecipitation coupled to *in vivo* cross-linking. So far and because Trm112 is a known interacting partner for several SAM-dependent MTases in eukaryotes, we only paid attention on putative or known MTases identified from our Co-IP experiments (Table 7). One crucial thing to remark is that comparing our Co-IP result with that of Lsm-FLAG (a protein involved in RNA metabolism) from Fischer et al (Fischer et al., 2010) reveals that there is no MTase detected in their experiments while in our case, as expected, MTases are highly enriched. By means of several rounds of *in vitro* complex characterizations (co-purification in *E. coli*, SEC-MALLS and MS), we have been able to validate at least 9 MTase partners for *Hvo*Trm112. In addition to proteins orthologous to known yeast Trm112 partners, namely Trm9 (HVO_1032), Mtq2 (HVO_2744), 7 other MTases were successfully shown to be real *Hvo*Trm112 partners, making the *Hvo*Trm112 interacting network unexpectedly much larger than that of the current yeast counterpart. Meanwhile, TrmG10/*Hvo*Trm11, orthologous to yeast Trm11, was found on the top of the Co-IP result list based on the high number of detected peptides. Unfortunately, we failed to prove the complex formation with *Hvo*Trm112 because TrmG10 was recovered in insoluble form despite attempting many efforts. It is also noteworthy that this *Hvo*Trm112 network could even be larger because there are still 15 other putative MTase partners that have not been tested so far (Table 7).

Table 7. **List of putative *Hvo*Trm112 MTase partners resulting from Co-IP experiments**

Rank	Symbol	UniProt accession number	Sequence coverage (%)	Specific spectra (Flag)	Specific spectra (Xlink)	Ratio (Xlink/ Flag)	Putative function	Validated interaction with HvoTrm112?
1	HVO_0773	D4GTS8	71.62	11.33	152.25	13.43	Unknown	YES
9	TrmG10 (aTrm11)	D4GZA0	57.17	4.67	88.75	19.02	tRNA (Guanine(10),N(2))-	NO
18	HVO_0475	D4GS15	62.10	1.00	63.25	63.25	Unknown	YES
20	SirC	D4GWE7	54.42	3.33	56.50	16.95	Uroporphyrin-III C-methyltransferase	Not tested
22	MenG	D4GZT8	44.69	7.33	54.75	7.47	Membrane protein	Not tested
25	HVO_A0501	D4GRG4	56.47	4.67	53.50	11.46	Pyridine nucleotide-disulfide oxidoreductase,	Not tested
31	cbiH2	D4GP59	44.12	10.00	49.00	4.90	Precorrin-3B C17-methyltransferase	Not tested
42	HVO_1475	D4GYB0	57.09	0.33	41.75	125.25	Putative DNA methylase	YES
55	HVO_0574	D4GSG9	41.17	1.00	36.00	36.00	Unknown	YES
77	HVO_1032 (aTrm9)	D4GVK8	31.57	0.33	30.00	90.00	tRNA (Uracil(34)) methyltransferase	YES
82	CbiT	D4GP64		12.67	29.25	2.31	Cobalt-precorrin-6B C(15)-methyltransferase	Not tested
102	HVO_0019	D4GYL4	23.79	0.00	25.25	-	24-sterol C-methyltransferase	YES
119	HVO_2875	D4GXJ1	43.78	0.00	22.75	-	Unknown	YES
161	CbiH1	D4GP60		7.67	18.00	2.35	Precorrin-3B C17-methyltransferase	Not tested
174	AglP	D4GYG5	29.81	0.00	17.00	-	Hexuronic acid methyltransferase	Not tested
228	MetE1	D4GW90	12.18	0.00	13.50	-	Methionine synthase	Not tested
229	CbiF	D4GP62	16.31	3.67	13.25	3.61	Cobalamin biosynthesis precorrin-3	Not tested
236	PrmC (aMtq2)	D4GW96	32.83	0.00	13.00	-	Class I translation termination factor	YES
295	HVO_1715	D4H060	22.50	0.00	10.00	-	Unknown	YES
328	Dph5	D4GUZ5	13.38	0.67	8.50	12.75	Diphtine synthase	Not tested
346	HVO_1534	D4GYH8	17.66	0.00	8.00	-	Unknown	Not tested
371	HVO_1093	D4GW13	19.95	0.00	7.25	-	Protein-L-isoaspartate O-methyltransferase	Not tested
385	HVO_2664	D4GV35	16.57	0.00	6.75	-	Unknown	Not tested
477	RlmE	D4GZC4	15.54	0.00	4.00	-	Ribosomal RNA large subunit	Not tested
488	Cna	D4GZ23	5.91	0.00	3.75	-	tRNA and rRNA cytosine-C5-methylase	Not tested
496	Nop5	D4GZN4	7.86	0.00	3.50	-	Archaeal nucleolar protein-like protein	Not tested

2.2. The role of salt in archaeal protein complex formation

As living optimally at high salt concentration (1.8-3.5M NaCl), *H. volcanii* cells have to adapt themselves by accumulating high ion concentration in their cytosol to avoid osmotic shock (Allers, 2010). As a result, this leads to salt-loving properties of *H. volcanii* proteins. In fact, in a low salt condition, some proteins were shown to be less active (*Hvo*Trm9-Trm112) or totally inactive (*Hvo*Mtq2-Trm112) (Chapter 2, part 3). Moreover, some protein complexes were observed to be salt-sensitive and dissociated under low salt conditions, while being stable under high salt conditions. This is for example the case of HVO_0475-Trm112 and to some extent HVO_1475-Trm112. As seen in the Figure 48A, at 500mM NaCl, the HVO_0475-Trm112 complex was nearly completely separated into two single proteins while at 1M NaCl, both proteins form a stable complex (Figure 48B). A similar behavior was observed for HVO_1475-Trm112 complex during ion exchange purification in which at 260mM NaCl, the complex was

also dissociated while the complex was stable at 450mM NaCl. This information is very useful for the crystallization of those protein complexes. For instance, it might be problematic to crystallize HVO_0475-Trm112 at 1M NaCl because through vapor diffusion crystallization methods we normally mix protein sample and reservoir solution in 1:1 ratio, resulting in the protein complex in a final solution with 500mM NaCl when the crystallization experiment is set-up. It is possible that this complex is gradually dissociating in this case, leading to heterogeneity in the protein sample, normally not good for the crystallization. Hence it is better and necessary to increase salt concentration for this complex when performing crystallization. The same is also true for HVO_1475-Trm112 complex.

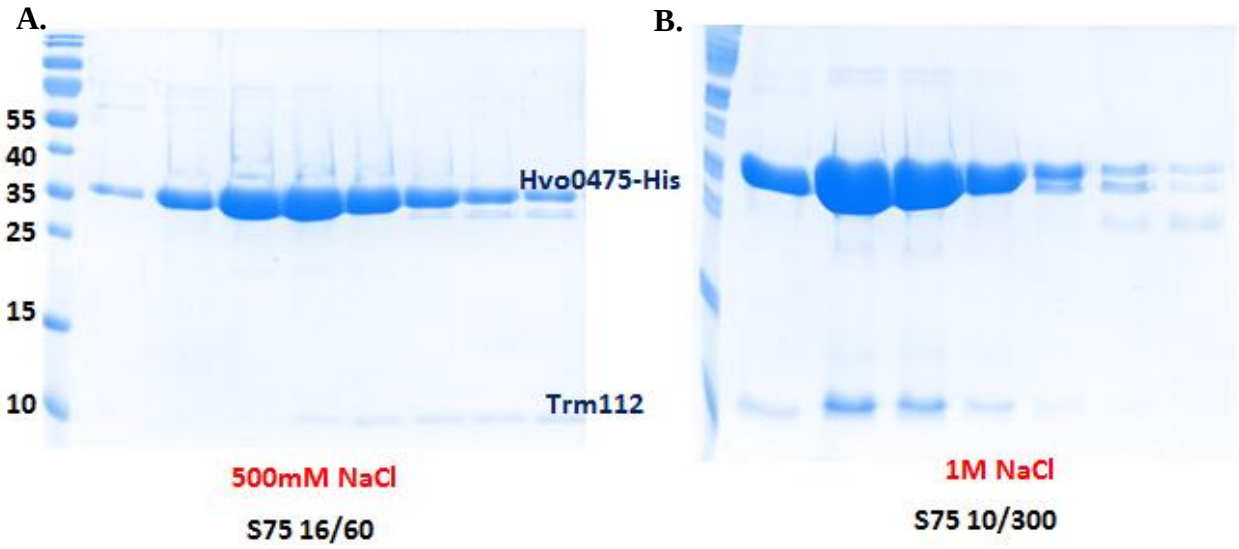


Figure 48. **HVO_0475-Trm112 complex purification**. (**A)**. Purification of the complex by gel filtration at 500mM NaCl. The complex was completely dissociated; (**B)**. All fractions from A were collected and adjusted to 1M NaCl and incubated for 1hour before loading on another gel filtration. The complex was able to form again in this condition.

2.3. Interacting mode between MTases and Trm112

In yeast, although Trm9, Mtq2 and Bud23 share very low sequence identity (less than 20%), the way they interact with Trm112 is very similar (Letoquart et al., 2015b). The same seems to be true for Trm11. Indeed, although no crystal structure for Trm11-Trm112 has been so

far obtained, hydrogen-deuterium exchange experiments coupled to mass spectrometry have shown that the peptides from Trm112 and from Trm11 that are affected upon complex formation correspond to the regions involved at the interface between Trm112 and its other partners in the crystal structures of Mtq2-Trm112, Trm9-Trm112 and Bud23-Trm112 (Bourgeois et al., 2017b). From the crystal structure of HVO_0019-Trm112 complex solved during my thesis – the first archaeal Trm112-MTase complex structure, this commonly interacting pattern is once again observed in *H. volcanii* MTase-Trm112 complex, including a number of characteristics. First, comparing the crystal structures of HVO_0019-Trm112 and *Sc*Bud23-Trm112 complexes, we observe the way both MTases interact with corresponding Trm112 proteins is very similar (rmsd value is 1.33Å) (Figure 49). HVO_0019 adopts the same Rossmann fold-like domain as all four yeast MTases that all interact with *Sc*Trm112 in the same region of Zn-binding domain. This later domain is conserved in *H. volcanii* Trm112 protein, but does not bind Zn^{2+} ion as the Zn-binding motif ($CX_{3-4}C$ and CX_2C from the N- and C-terminal parts in yeasts and some archaea, respectively; where X is for any amino acid) is not conserved. Second, the interaction mode between MTase and Trm112 is totally similar between yeast and *H. volcanii* complexes with the presence of a β-zipper interaction formed between Trm112 strand β4 and MTase strand β3 implying formation of hydrogen bonds between main chain atoms. This explains why different proteins can maintain a similar interaction mode that is in fact much more dependent on the local three-dimensional structure rather than on conservation of amino acid residues at the positions involved in the formation of this β-zipper. Third, Trm112-MTase complex formation encompasses a large hydrophophic region on the surface of both proteins. This explains the prerequisite of Trm112 to stabilize and purify several MTases (Trm9, Mtq2 and Bud23 in *S. cerevisiae* and HVO_0019 and most probably HVO_0773, HVO_1475 as well as human METTL5) in their soluble forms in *E. coli*. In addition, Trm112 is recognized to induce SAM binding capability as well as to contribute to substrate binding by Trm11, Mtq2 and probably Trm9 and Bud23 in yeast (Liger et al., 2011; Bourgeois et al., 2017b). We cannot exclude these characteristics are also present in archaea but more studies are needed.

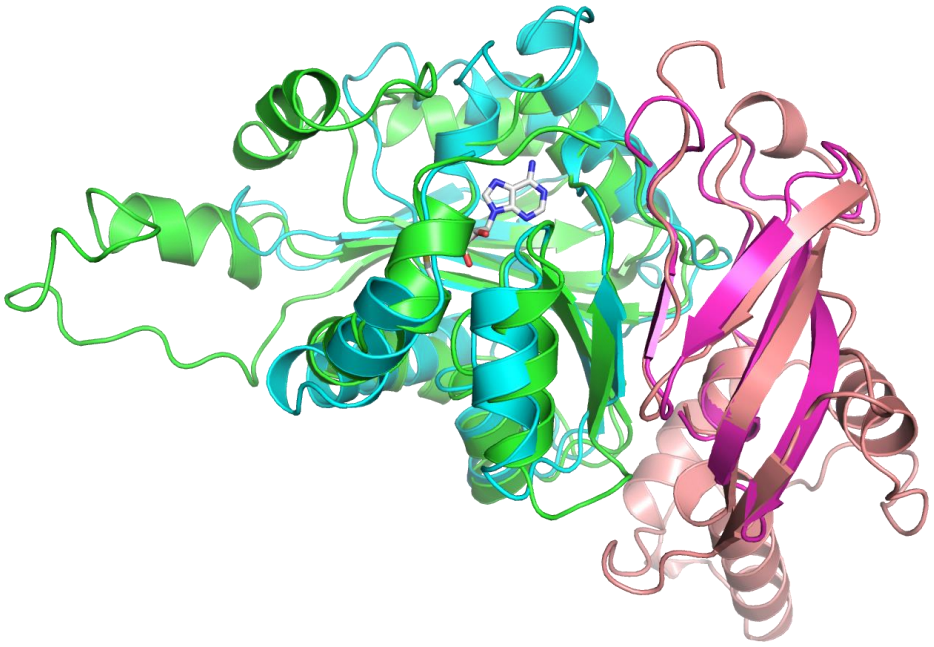


Figure 49. **The superimposition of *Sc*Bud23-Trm112** (blue and light pink, respectively) **onto HVO_0019-*Hvo*Trm112** (green and pink, respectively)

2.4. Functions of different *Hvo*Trm112 MTase partners

2.4.1. *Hvo*Mtq2

In spite of having distinct evolutionary origins, both eukaryotic and bacterial class I release factor contain a universally conserved GGQ motif which is known to be N5-methylated at glutamine side chain, underlying the importance of this modification (Dincbas-Renqvist et al., 2000; Heurgue-Hamard et al., 2005). In fact, the methylation of Q residue in bacterial RF2 GGQ motif is established to affect the efficiency of the polypeptide release (Dincbas-Renqvist et al., 2000; Pierson et al., 2016). In bacteria, PrmC (or HemK) is the MTase responsible for the RF1 and RF2 QQG methylation (Heurgue-Hamard et al., 2002; Graille et al., 2005) while in yeast, a catalytic subunit Mtq2 in complex with Trm112 is responsible for this methylation on eRF1 in an eRF3-dependent manner (Heurgue-Hamard et al., 2005; Polevoda et al., 2006). In archaea, the translation termination process is quite comparable to the case of eukaryotes and the archaeal class I release factor aRF1 shares similar properties as its eukaryotic counterpart, both differing from those in bacteria (Dontsova et al., 2000; Kobayashi et al., 2012). This led to an assumption

that aRF1 is also methylated. Initially, as no ortholog of eRF3 was detected in archaeal genomes, this process was supposed to be in a class-II release factor (aRF3)-independent manner (Heurgue-Hamard et al., 2005). However, since then, it has been shown that in archaea, a single protein (aEF1α), in addition to act as elongation factor carrying tRNAs during translation elongation also acts in translation termination as eRF3 ortholog as well as in mRNA surveillance by interacting with aPelota (Saito et al., 2010). In the thesis, I demonstrated that in archaea, aRF1 is methylated in a similar manner to the eukaryotic counterpart by a eukaryotic orthologous complex between HVO_2744 (*Hvo*Mtq2) and HVO_1131 (*Hvo*Trm112) (Figure 33). It is remarkable to mention that this methylation is also dependent on salt concentration, a commonly-observed property of proteins from halophilic organisms. The role of high salt in this case is possibly involved in enhancing the enzyme and substrate stability as well as their conformations, which in turn benefit the enzymatic activity by means of conventional S_N2 mechanism of SAM-dependent MTases. The effect of aRF1 GGQ methylation on archaeal translation termination as well as the substrate recognition mechanism by archaeal Mtq2-Trm112, remains to be resolved as it is the case for their eukaryotic counterparts.

Regarding enzymatic activity assays, it is normally not an easy task for *H. volcanii* proteins in particular and for archaea in general since the optimal conditions for these proteins are difficult to be uncovered (high salt, pH, temperature, …). Initially, almost no activity could be detected for *Hvo*Mtq2-Trm112 enzyme by use of optimal conditions from the yeast enzymes (50mM phosphate buffer pH7.5, 0.1mM EDTA, 10mM $MgCl_2$ and 10mM NH_4Cl) and even in case of high salt concentration (2M KCl). Fortunately, through keeping incubating the reaction overnight at 45^0C, I was able to detect enzymatic activity. This led me to optimize the experimental buffer conditions by increasing the concentrations of buffer contents. It is due to the fact that incubation at 45^0C for a long time resulted in water evaporation, therefore gradually increasing the reaction solution concentrations which turned out to be the optimal conditions for the *Hvo*Mtq2-Trm112 activity. At the end of buffer optimization process (Figure 32), I ended up with 400mM phosphate as the best condition for enzyme activity among those tested. It is also important to note that the high concentration of $MgCl_2$ (5mM instead of 2.5mM) was not a good option for the activity, possibly resulting from precipitation trouble as a result of high phosphate and high $MgCl_2$ concentration combination. Moreover, to control changes in reaction contents via evaporation, all experiments were then carried out in a closed incubator.

*2.4.2. Hvo*Trm9

As the Trm9-Trm112 complex in *S. cerevisiae*, the complex formed between orthologous proteins from *H. volcanii* (HVO_1032 (Trm9) and HVO_1131 (Trm112)) was shown to be active on total tRNAs, thereby most likely catalyzing the conversion of cm^5U to mcm^5U. However, this activity is only detected for the chimeric enzymatic reaction between the *H. volcanii* Trm9-Trm112 enzyme complex on the yeast substrate. In case of *H. volcanii* total RNA substrates, no activity at all was detected. There are several reasons that could explain this result. The most probable is that the tRNA substrates purified from the *H. volcanii trm9*Δ strain do not contain cm^5U34. Indeed, in *S. cerevisiae*, deletion of *TRM9* gene can result in an accumulation of ncm^5U34 (5-carbamoylmethyl uridine) which is not the right substrate for *Sc*Trm9-Trm112 (Chen et al., 2011). We cannot exclude this could also occur in *H. volcanii* where mutant *trm9*Δ could initially lead to accumulation of cm^5U34, which could then potentially be converted to ncm^5U (or another modification) by an unknown enzyme or an unknown mechanism. The quality of tRNA substrates could also be another issue for the absence of enzymatic activity. In fact, total RNA purified from *H. volcanii trm9*Δ contained lots of contaminants such as genomic DNA, rRNAs compared to the *Sc* counterpart (Figure 37B). This could as a result develop inhibitory effects on the activity of *Hvo*Trm9-Trm112. We also could not exclude that this buffer condition is not optimal for activity of *Hvo*Trm9-Trm112 on *Hvo* total RNAs, despite the fact that it was compatible with the detection of enzymatic activity on *Sc* total RNAs. Moreover, similar to the case of *Hvo*Mtq2-Trm112 complex, *Hvo*Trm9-Trm112 requires high salt concentration for optimal enzymatic activity. Meanwhile, HVO_0574 was initially proposed as Trm9 ortholog (Grosjean et al., 2008; G. Phillips, & de Crecy-Lagard, 2011). In the thesis, I have performed enzymatic activity on the HVO_0574-Trm112 complex and could not detect activity on tRNAs. Hence, the biochemical function of HVO_0574-Trm112 complex remains to be unraveled.

*2.4.3. Hvo*TrmG10

Different from other known or putative Trm112 MTase partners, TrmG10 (Trm11 ortholog) is the only protein already well studied in archaea. Functional and structural characterization of TrmG10 has been so far carried out in *Pyrococcus abyssi* and *Thermococcus*

kodakarensis in which TrmG10 is responsible for the formation of m^2G10 and $m^2{}_2G10$ (Armengaud et al., 2004; Urbonavicius et al., 2006; Hirata et al., 2016), quite similar to the role of Trm11 in yeast, which catalyzes the former modification in some tRNAs (Purushothaman et al., 2005; Bourgeois et al., 2017b). However, Trm11 is known to require its Trm112 partner to be active in yeast (Purushothaman et al., 2005; Bourgeois et al., 2017b) while its archaeal orthologs studied so far were shown to be active on their own (Urbonavicius et al., 2006; Hirata et al., 2016). This occurrence initially led to the hypothesis that Trm112 was not mandatory for Trm11 activity in archaea. However, bioinformatics analysis revealed that although Trm11 orthologs are present in all archaeal phyla, archaeal Trm112 orthologs are absent in Thermococcales and Methanobacteriales from the Euryarchaeota phylum (Figure 22). This then could explain why Trm11 orthologs from two Thermococcales archaea (*Pyrococcus abyssi* and *Thermococcus kodakarensis*) are able to generate m^2G and/or $m^2{}_2G$ at position 10 of some tRNAs *in vitro* without the help of Trm112 partner. Meanwhile, for most cases, Trm11 orthologs co-occur with Trm112, raising the possibility that with the exception of Thermococcales and Methanobacteriales TrmG10, the other archaeal Trm11 orthologs may be in complex with Trm112 and also entail Trm112 to be active as demonstrated in eukaryotes. In the thesis, *Hvo*Trm11 (*Hvo*_0156) was found on the top of our Co-IP results and predicted to be an *Hvo*Trm112 partner with high confidence. However, although several efforts have been made so far to analyse the formation of *Hvo*TrmG10-Trm112 complex, *i.e.* using different culture media, expression strains, inducer concentrations, temperatures and codon-optimized synthetic DNA sequence, *H. volcanii* TrmG10 was found expressed but only in an insoluble form in inclusion bodies. Hence, this leaves the question related to *Hvo*TrmG10-Trm112 complex formation still opened. One reason for the protein insolubility could be that *Hvo*TrmG10 is a salt-loving protein, probably resulting in mis-folding and aggregates under low ionic conditions in *E. coli*. The next attempt to study the interaction of these two proteins should be co-purification of these two proteins in their native organism or expression of orthologous proteins from another archaea. This would ideally help to characterize their interaction through biochemical and biophysical approaches. This task is experimentally feasible since methods to over-express and purify proteins directly in *H. volcanii* have been developed and already performed by Allers and colleagues (Allers et al., 2010).

2.4.4. HVO_0019

Thanks to the crystal structure of HVO_0019-Trm112, we were able to take a first glance at the putative HVO_0019 function. Using DALI server (Holm, & Rosenstrom, 2010) to search Protein Data Bank (PDB) for proteins sharing strong structural similarity with HVO_0019 MTase domain, we identified of hundreds of crystal structures of MTases with Z-scores ranging from 21 to 10. Among these structures, based on structural analysis, we found out one candidate (NodS from *Bradyrhizobium japonicum* (Cakici et al., 2010)) sharing lots of features in common with the HVO_0019 MTase active site. For example, residues R22, E111, Y114, Y115 and W152 from HVO_0019 structurally match with R21, E115, Y118, Y119 and W156 from NodS, respectively (Figure 42D). In addition, the side chains from HVO_0019 W10 and NodS W20 are also seen in close proximity. NodS is known to methylate the NH_2 group of its glucosamine moiety, then participating in the biosynthesis of the Nod factor, a modified chitosaccharide acting as a signal molecule in rhizobia (Cakici et al., 2010). The similarity in the active site of both proteins suggests that HVO_0019 might also methylate a substrate with an hexose sugar ring. Further biochemical experimental studies will be needed to characterize the enzymatic activity of the *Hvo*Trm112-HVO_0019 complex on putative sugar substrates.

2.4.5. HVO_1475

Among those newly identified Trm112 MTase partners, HVO_1475, which belongs to the Clusters of Orthologous Groups 2263 (COG2263), was proposed to catalyze the formation of m^6A at position 1432 on 16S rRNA (Grosjean et al., 2008). This m^6A modification is found in small ribosomal subunit rRNA from some organisms, in particular archaea *H. volcanii* and *S. Solfataricus* archeon (Noon et al., 1998) and eukaryotes such as *Homo sapiens (Maden, 1986),* plants (Cecchini, & Miassod, 1979) and other metazoans (*Xenopus laevis* and *Rattus norvegicus*) but is absent from *S. cerevisiae* (Grosjean et al., 2008). The COG2263 family annotated as RNA MTase or N6-DNA-methylase distributed in most archaea and many eukaryotes was anticipated to be responsible for m^6A modification in those cases (Grosjean et al., 2008). This is quite coincident because there is no HVO_1475 ortholog identified in *S. cerevisiae* whereas this ortholog is found in *Homo sapiens* known as METTL5, which I was able to identify as a new MTase partner of human TRMT112 and in our prediction probably to be responsible for the m^6A 18S rRNA modification. However, more functional and structural studies on *H. volcanii* HVO_1475-Trm112 and human METTL5-TRMT112 are required to decipher this hypothesis.

2.4.6. Other MTase partners

In addition to *Hvo*Trm112 MTase partners with known or putative functions, there are still five other confirmed partners without any information about their functions, including HVO_0773, HVO_0574, HVO_0475, HVO_1715 and HVO_2875. Through sequence analysis, those MTases are recognized to belong to the class I SAM-dependent methyltransferases. However, these proteins have very loose annotations as SAM-dependent MTases in public protein databases such as UniProt and NCBI. This is not a surprising issue because these online servers are normally suffering from severe annotation errors as a result of automatic annotation procedures by computational programs (Pfeiffer, & Oesterhelt, 2015). Thus, experimental validations based on structural and biochemical studies for those complexes remain to be made in order to provide insight into their specific biological functions.

3. Human METTL5

3.1. Human TRMT112 network

The current human TRMT112 network is characterized based on the yeast counterpart, which contain four complexes like those in yeast. Except TRMT11 (Trm11 ortholog) whose biochemical and biological functions have not been characterized so far, ABH8 (Trm9 ortholog), HEMK2 (Mtq2 ortholog) and WBSCR22 (Bud23 ortholog) are well characterized, requiring TRMT112 for enzymatic activity and are involved in human diseases (Bourgeois et al., 2017a). One exciting aspect of the results obtained during my thesis is that thanks to Trm112 interacting network in *H. volcanii*, we have so far extended our knowledge on the corresponding counterpart in human by characterizing one more MTase partner – METTL5, which ortholog is not present in *S. cerevisiae*.

3.2. Putative biological function of human METTL5

The real function of METTL5 is not exactly known so far. In human and some metazoans, there is presence of m^6A in 18S rRNA but no enzyme is currently determined to be responsible for this modification (Maden, 1986). Meanwhile, an ortholog of METTL5 in *H. volcanii*, HVO_1475 is predicted to methylate m^6A at position 1432 of 16S rRNA (Grosjean et al., 2008). Coincidently, both the enzyme ortholog and m^6A modification are not observed in *S.*

cerevisiae. The human METTL5 belonging to COG2263, annotated as RNA/DNA methyltransferase, is in our hypothesis linked to the m^6A formation in 18S rRNA in human. However, more experimental attempts are needed to characterize this hypothesis. This would be an exciting discovery if it is the case because base methylations in rRNAs are well acknowledged to be important for ribosome biogenesis, which defects normally link to human diseases (see Introduction I, 2.2.2.3). However, in case this hypothesis is true, it is a very crucial but not surprising result because four previously-known MTase partners of human TRMT112 including TRMT11, ABH8, HEMK2 and WBSCR22 are known to modify translation actors and it is well established that they have implications in human diseases (see Introduction II, 2.2).

II. CONCLUSION

Trm112 is a perfect example for investigating the importance of methylations involved in translation where a small protein can interact with and activate many MTase partners, interestingly in the same way. Knowing the fact that the translation process is very compatible between eukaryotes and archaea, we paid our attention on resolving the Trm112 interacting network in the latter organism on the purpose of making use of advantages of studying archaeal proteins for the dissection of more complex eukaryotic counterparts. The study surprisingly led us to a much larger Trm112 interacting network in archaea compared to well-known equivalent ones in yeast and human. Moreover, this is not only just findings for archaea; it also in turn benefits the human knowledge of TRMT112 network by identifying one more MTase partner. Studies of archaeal Trm112 interacting network are still at their infancies and further studies are necessary to clarify the role of archaeal Trm112 and its interacting partners in archaeal biological functions.

REFERENCES

Abbasi-Moheb, L., Mertel, S., Gonsior, M., Nouri-Vahid, L., Kahrizi, K., Cirak, S., . . . Kuss, A. W. (2012). Mutations in NSUN2 cause autosomal-recessive intellectual disability. *Am J Hum Genet, 90*(5), 847-855. doi: 10.1016/j.ajhg.2012.03.021

Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., . . . Zwart, P. H. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr, 66*(Pt 2), 213-221. doi: 10.1107/S0907444909052925

Aitken, C. E., Beznoskova, P., Vlckova, V., Chiu, W. L., Zhou, F., Valasek, L. S., . . . Lorsch, J. R. (2016). Eukaryotic translation initiation factor 3 plays distinct roles at the mRNA entry and exit channels of the ribosomal preinitiation complex. *Elife, 5*. doi: 10.7554/eLife.20934

Al-Hadid, Q., Roy, K., Chanfreau, G., & Clarke, S. G. (2016a). Methylation of yeast ribosomal protein Rpl3 promotes translational elongation fidelity. *RNA, 22*(4), 489-498. doi: 10.1261/rna.054569.115

Al-Hadid, Q., Roy, K., Munroe, W., Dzialo, M. C., Chanfreau, G. F., & Clarke, S. G. (2014). Histidine methylation of yeast ribosomal protein Rpl3p is required for proper 60S subunit assembly. *Mol Cell Biol, 34*(15), 2903-2916. doi: 10.1128/MCB.01634-13

Al-Hadid, Q., White, J., & Clarke, S. (2016b). Ribosomal protein methyltransferases in the yeast Saccharomyces cerevisiae: Roles in ribosome biogenesis and translation. *Biochem Biophys Res Commun, 470*(3), 552-557. doi: 10.1016/j.bbrc.2016.01.107

Alexandrov, A., Chernyakov, I., Gu, W., Hiley, S. L., Hughes, T. R., Grayhack, E. J., & Phizicky, E. M. (2006). Rapid tRNA decay can result from lack of nonessential modifications. *Mol Cell, 21*(1), 87-96. doi: 10.1016/j.molcel.2005.10.036

Allers, T. (2010). Overexpression and purification of halophilic proteins in Haloferax volcanii. *Bioeng Bugs, 1*(4), 288-290. doi: 10.4161/bbug.1.4.11794

Allers, T., Barak, S., Liddell, S., Wardell, K., & Mevarech, M. (2010). Improved strains and plasmid vectors for conditional overexpression of His-tagged proteins in Haloferax volcanii. *Appl Environ Microbiol, 76*(6), 1759-1769. doi: 10.1128/AEM.02670-09

Altman-Price, N., & Mevarech, M. (2009). Genetic evidence for the importance of protein acetylation and protein deacetylation in the halophilic archaeon Haloferax volcanii. *J Bacteriol, 191*(5), 1610-1617. doi: 10.1128/JB.01252-08

Anantharaman, V., Koonin, E. V., & Aravind, L. (2002). SPOUT: a class of methyltransferases that includes spoU and trmD RNA methylase superfamilies, and novel superfamilies of predicted prokaryotic RNA methylases. *J Mol Microbiol Biotechnol, 4*(1), 71-75.

Anger, A. M., Armache, J. P., Berninghausen, O., Habeck, M., Subklewe, M., Wilson, D. N., & Beckmann, R. (2013). Structures of the human and Drosophila 80S ribosome. *Nature, 497*(7447), 80-85. doi: 10.1038/nature12104

Appel, C. D., & Maxwell, E. S. (2007). Structural features of the guide:target RNA duplex required for archaeal box C/D sRNA-guided nucleotide 2'-O-methylation. *RNA, 13*(6), 899-911. doi: 10.1261/rna.517307

Aravind, L., & Koonin, E. V. (2001). THUMP--a predicted RNA-binding domain shared by 4-thiouridine, pseudouridine synthases and RNA methylases. *Trends Biochem Sci, 26*(4), 215-217.

Armache, J. P., Jarasch, A., Anger, A. M., Villa, E., Becker, T., Bhushan, S., . . . Beckmann, R. (2010). Cryo-EM structure and rRNA model of a translating eukaryotic 80S ribosome at 5.5-A resolution. *Proc Natl Acad Sci U S A, 107*(46), 19748-19753. doi: 10.1073/pnas.1009999107

Armengaud, J., Urbonavicius, J., Fernandez, B., Chaussinand, G., Bujnicki, J. M., & Grosjean, H. (2004). N2-methylation of guanosine at position 10 in tRNA is catalyzed by a THUMP domain-containing, S-adenosylmethionine-dependent methyltransferase, conserved in Archaea and Eukaryota. *J Biol Chem, 279*(35), 37142-37152. doi: 10.1074/jbc.M403845200

Armengod, M. E., Moukadiri, I., Prado, S., Ruiz-Partida, R., Benitez-Paez, A., Villarroya, M., . . . Navarro-Gonzalez, C. (2012). Enzymology of tRNA modification in the bacterial MnmEG pathway. *Biochimie, 94*(7), 1510-1520. doi: 10.1016/j.biochi.2012.02.019

Ashkenazy, H., Erez, E., Martz, E., Pupko, T., & Ben-Tal, N. (2010). ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. *Nucleic Acids Res, 38*(Web Server issue), W529-533. doi: gkq399 [pii] 10.1093/nar/gkq399

Auffinger, P., & Westhof, E. (2001). An extended structural signature for the tRNA anticodon loop. *RNA, 7*(3), 334-341.

Awai, T., Kimura, S., Tomikawa, C., Ochi, A., Ihsanawati, Bessho, Y., . . . Hori, H. (2009). Aquifex aeolicus tRNA (N2,N2-guanine)-dimethyltransferase (Trm1) catalyzes transfer of methyl groups not only to guanine 26 but also to guanine 27 in tRNA. *J Biol Chem, 284*(31), 20467-20478. doi: 10.1074/jbc.M109.020024

Banerjee, A. K. (1980). 5'-terminal cap structure in eucaryotic messenger ribonucleic acids. *Microbiol Rev, 44*(2), 175-205.

Becker, T., Franckenberg, S., Wickles, S., Shoemaker, C. J., Anger, A. M., Armache, J. P., . . . Beckmann, R. (2012). Structural basis of highly conserved ribosome recycling in eukaryotes and archaea. *Nature, 482*(7386), 501-506. doi: 10.1038/nature10829

Begley, U., Dyavaiah, M., Patil, A., Rooney, J. P., DiRenzo, D., Young, C. M., . . . Begley, T. J. (2007). Trm9-catalyzed tRNA modifications link translation to the DNA damage response. *Mol Cell, 28*(5), 860-870. doi: S1097-2765(07)00633-8 [pii] 10.1016/j.molcel.2007.09.021

Begley, U., Sosa, M. S., Avivar-Valderas, A., Patil, A., Endres, L., Estrada, Y., . . . Begley, T. (2013). A human tRNA methyltransferase 9-like protein prevents tumour growth by regulating LIN9 and HIF1-alpha. *EMBO Mol Med, 5*(3), 366-383. doi: 10.1002/emmm.201201161

Ben-Shem, A., Garreau de Loubresse, N., Melnikov, S., Jenner, L., Yusupova, G., & Yusupov, M. (2011). The structure of the eukaryotic ribosome at 3.0 A resolution. *Science, 334*(6062), 1524-1529. doi: 10.1126/science.1212642

Blanchet, S., Rowe, M., Von der Haar, T., Fabret, C., Demais, S., Howard, M. J., & Namy, O. (2015). New insights into stop codon recognition by eRF1. *Nucleic Acids Res, 43*(6), 3298-3308. doi: 10.1093/nar/gkv154

Boschi-Muller, S., & Motorin, Y. (2013). Chemistry enters nucleic acids biology: enzymatic mechanisms of RNA modification. *Biochemistry (Mosc), 78*(13), 1392-1404. doi: 10.1134/S0006297913130026

Bourgeois, G., Letoquart, J., van Tran, N., & Graille, M. (2017a). Trm112, a Protein Activator of Methyltransferases Modifying Actors of the Eukaryotic Translational Apparatus. *Biomolecules, 7*(1). doi: 10.3390/biom7010007

Bourgeois, G., Marcoux, J., Saliou, J. M., Cianferani, S., & Graille, M. (2017b). Activation mode of the eukaryotic m2G10 tRNA methyltransferase Trm11 by its partner protein Trm112. *Nucleic Acids Res, 45*(4), 1971-1982. doi: 10.1093/nar/gkw1271

Bourgeois, G., Ney, M., Gaspar, I., Aigueperse, C., Schaefer, M., Kellner, S., . . . Motorin, Y. (2015). Eukaryotic rRNA Modification by Yeast 5-Methylcytosine-Methyltransferases and Human Proliferation-Associated Antigen p120. *PLoS One, 10*(7), e0133321. doi: 10.1371/journal.pone.0133321

Bricogne, G., Blanc, E., Brandl, M., Flensburg, C., Keller, P., Paciorek, W., . . . Womack, T. O. (2016). BUSTER version 2.10.2. Cambridge, United Kingdom: Global Phasing Ltd.

Brochier-Armanet, C., Forterre, P., & Gribaldo, S. (2011). Phylogeny and evolution of the Archaea: one hundred genomes later. *Curr Opin Microbiol, 14*(3), 274-281. doi: 10.1016/j.mib.2011.04.015

Brown, A., Shao, S., Murray, J., Hegde, R. S., & Ramakrishnan, V. (2015). Structural basis for stop codon recognition in eukaryotes. *Nature, 524*(7566), 493-496. doi: 10.1038/nature14896

Brown, K. M., & Gilmartin, G. M. (2003). A mechanism for the regulation of pre-mRNA 3' processing by human cleavage factor Im. *Mol Cell, 12*(6), 1467-1476.

Bujnicki, J. M., Droogmans, L., Grosjean, H., Purushothaman, S. K., & Lapeyre, B. (2004). Bioinformatics-guided identification and experimental char- acterization of novel RNA methyltransferases. In J. M. Bujnicki (Ed.), *Practical Bioinformatics* (Vol. 15, pp. 139-168). Heidelberg, Germany: Springer-Verlag.

Cakici, O., Sikorski, M., Stepkowski, T., Bujacz, G., & Jaskolski, M. (2010). Crystal structures of NodS N-methyltransferase from Bradyrhizobium japonicum in ligand-free form and as SAH complex. *J Mol Biol, 404*(5), 874-889. doi: 10.1016/j.jmb.2010.10.016

Cantara, W. A., Crain, P. F., Rozenski, J., McCloskey, J. A., Harris, K. A., Zhang, X., . . . Agris, P. F. (2011). The RNA Modification Database, RNAMDB: 2011 update. *Nucleic Acids Res, 39*(Database issue), D195-201. doi: gkq1028 [pii]10.1093/nar/gkq1028

Cantoni, G. L. (1975). Biological methylation: selected aspects. *Annu Rev Biochem, 44*, 435-451. doi: 10.1146/annurev.bi.44.070175.002251

Cao, W., & De La Cruz, E. M. (2013). Quantitative full time course analysis of nonlinear enzyme cycling kinetics. *Sci Rep, 3*, 2658. doi: 10.1038/srep02658

Carapito, C., Burel, A., Guterl, P., Walter, A., Varrier, F., Bertile, F., & Van Dorsselaer, A. (2014). MSDA, a proteomics software suite for in-depth Mass Spectrometry Data Analysis using grid computing. *Proteomics, 14*(9), 1014-1019. doi: 10.1002/pmic.201300415

Carapito, C., Lane, L., Benama, M., Opsomer, A., Mouton-Barbosa, E., Garrigues, L., . . . Vandenbrouck, Y. (2015). Computational and Mass-Spectrometry-Based Workflow for the Discovery and Validation of Missing Human Proteins: Application to Chromosomes 2 and 14. *J Proteome Res, 14*(9), 3621-3634. doi: 10.1021/pr5010345

Cecchini, J. P., & Miassod, R. (1979). Studies on the methylation of cytoplasmic ribosomal RNA from cultured higher plant cells. *Eur J Biochem, 98*(1), 203-214.

Chen, C., Huang, B., Anderson, J. T., & Bystrom, A. S. (2011). Unexpected accumulation of ncm(5)U and ncm(5)S(2) (U) in a trm9 mutant suggests an additional step in the synthesis of mcm(5)U and mcm(5)S(2)U. *PLoS One, 6*(6), e20783. doi: 10.1371/journal.pone.0020783 PONE-D-11-05734 [pii]

Cheng, X., Kumar, S., Posfai, J., Pflugrath, J. W., & Roberts, R. J. (1993). Crystal structure of the HhaI DNA methyltransferase complexed with S-adenosyl-L-methionine. *Cell, 74*(2), 299-307.

Cheng, X., & Roberts, R. J. (2001). AdoMet-dependent methylation, DNA methyltransferases and base flipping. *Nucleic Acids Res, 29*(18), 3784-3795.

Cheng, Z., Saito, K., Pisarev, A. V., Wada, M., Pisareva, V. P., Pestova, T. V., . . . Song, H. (2009). Structural insights into eRF3 and stop codon recognition by eRF1. *Genes Dev, 23*(9), 1106-1118. doi: 10.1101/gad.1770109

Choi, J., Ieong, K. W., Demirci, H., Chen, J., Petrov, A., Prabhakar, A., . . . Puglisi, J. D. (2016). N(6)-methyladenosine in mRNA disrupts tRNA selection and translation-elongation dynamics. *Nat Struct Mol Biol, 23*(2), 110-115. doi: 10.1038/nsmb.3148

Couttas, T. A., Raftery, M. J., Padula, M. P., Herbert, B. R., & Wilkins, M. R. (2012). Methylation of translation-associated proteins in Saccharomyces cerevisiae: Identification of methylated lysines and their methyltransferases. *Proteomics, 12*(7), 960-972. doi: 10.1002/pmic.201100570

Cox, C. J., Foster, P. G., Hirt, R. P., Harris, S. R., & Embley, T. M. (2008). The archaebacterial origin of eukaryotes. *Proc Natl Acad Sci U S A, 105*(51), 20356-20361. doi: 10.1073/pnas.0810647105

Davila Lopez, M., & Samuelsson, T. (2008). Early evolution of histone mRNA 3' end processing. *RNA, 14*(1), 1-10. doi: 10.1261/rna.782308

Davydova, E., Ho, A. Y., Malecki, J., Moen, A., Enserink, J. M., Jakobsson, M. E., . . . Falnes, P. O. (2014). Identification and characterization of a novel evolutionarily conserved lysine-specific methyltransferase targeting eukaryotic translation elongation factor 2 (eEF2). *J Biol Chem, 289*(44), 30499-30510. doi: 10.1074/jbc.M114.601658

Decatur, W. A., & Fournier, M. J. (2002). rRNA modifications and ribosome function. *Trends Biochem Sci, 27*(7), 344-351.

Delatte, B., Wang, F., Ngoc, L. V., Collignon, E., Bonvin, E., Deplus, R., . . . Fuks, F. (2016). RNA biochemistry. Transcriptome-wide distribution and function of RNA hydroxymethylcytosine. *Science, 351*(6270), 282-285. doi: 10.1126/science.aac5253

Deng, W., Babu, I. R., Su, D., Yin, S., Begley, T. J., & Dedon, P. C. (2015). Trm9-Catalyzed tRNA Modifications Regulate Global Protein Expression by Codon-Biased Translation. *PLoS Genet, 11*(12), e1005706. doi: 10.1371/journal.pgen.1005706

Deng, X., Chen, K., Luo, G. Z., Weng, X., Ji, Q., Zhou, T., & He, C. (2015). Widespread occurrence of N6-methyladenosine in bacterial mRNA. *Nucleic Acids Res, 43*(13), 6557-6567. doi: 10.1093/nar/gkv596

Dennis, P. P., Tripp, V., Lui, L., Lowe, T., & Randau, L. (2015). C/D box sRNA-guided 2'-O-methylation patterns of archaeal rRNA molecules. *BMC Genomics, 16*, 632. doi: 10.1186/s12864-015-1839-z

Dever, T. E., & Green, R. (2012). The elongation, termination, and recycling phases of translation in eukaryotes. *Cold Spring Harb Perspect Biol, 4*(7), a013706. doi: 10.1101/cshperspect.a013706

Dillon, S. C., Zhang, X., Trievel, R. C., & Cheng, X. (2005). The SET-domain protein superfamily: protein lysine methyltransferases. *Genome Biol, 6*(8), 227. doi: 10.1186/gb-2005-6-8-227

Dina, C., Meyre, D., Gallina, S., Durand, E., Korner, A., Jacobson, P., . . . Froguel, P. (2007). Variation in FTO contributes to childhood obesity and severe adult obesity. *Nat Genet, 39*(6), 724-726. doi: 10.1038/ng2048

Dincbas-Renqvist, V., Engstrom, A., Mora, L., Heurgue-Hamard, V., Buckingham, R., & Ehrenberg, M. (2000). A post-translational modification in the GGQ motif of RF2 from Escherichia coli stimulates termination of translation. *EMBO J, 19*(24), 6900-6907.

Doll, A., & Grzeschik, K. H. (2001). Characterization of two novel genes, WBSCR20 and WBSCR22, deleted in Williams-Beuren syndrome. *Cytogenet Cell Genet, 95*(1-2), 20-27. doi: 57012

Dominissini, D., Moshitch-Moshkovitz, S., Schwartz, S., Salmon-Divon, M., Ungar, L., Osenberg, S., . . . Rechavi, G. (2012). Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. *Nature, 485*(7397), 201-206. doi: 10.1038/nature11112

Dominissini, D., Nachtergaele, S., Moshitch-Moshkovitz, S., Peer, E., Kol, N., Ben-Haim, M. S., . . . He, C. (2016). The dynamic N(1)-methyladenosine methylome in eukaryotic messenger RNA. *Nature, 530*(7591), 441-446. doi: 10.1038/nature16998

Dontsova, M., Frolova, L., Vassilieva, J., Piendl, W., Kisselev, L., & Garber, M. (2000). Translation termination factor aRF1 from the archaeon Methanococcus jannaschii is active with eukaryotic ribosomes. *FEBS Lett, 472*(2-3), 213-216.

Dzialo, M. C., Travaglini, K. J., Shen, S., Roy, K., Chanfreau, G. F., Loo, J. A., & Clarke, S. G. (2014). Translational roles of elongation factor 2 protein lysine methylation. *J Biol Chem, 289*(44), 30511-30524. doi: 10.1074/jbc.M114.605527

Emsley, P., Lohkamp, B., Scott, W. G., & Cowtan, K. (2010). Features and development of Coot. *Acta Crystallogr D Biol Crystallogr, 66*(Pt 4), 486-501. doi: 10.1107/S0907444910007493

Eyler, D. E., Wehner, K. A., & Green, R. (2013). Eukaryotic release factor 3 is required for multiple turnovers of peptide release catalysis by eukaryotic release factor 1. *J Biol Chem, 288*(41), 29530-29538. doi: 10.1074/jbc.M113.487090

Fan-Minogue, H., Du, M., Pisarev, A. V., Kallmeyer, A. K., Salas-Marco, J., Keeling, K. M., . . . Bedwell, D. M. (2008). Distinct eRF3 requirements suggest alternate eRF1 conformations mediate peptide release during eukaryotic translation termination. *Mol Cell, 30*(5), 599-609. doi: 10.1016/j.molcel.2008.03.020

Fica, S. M., Tuttle, N., Novak, T., Li, N. S., Lu, J., Koodathingal, P., . . . Piccirilli, J. A. (2013). RNA catalyses nuclear pre-mRNA splicing. *Nature, 503*(7475), 229-234. doi: 10.1038/nature12734

Figaro, S., Scrima, N., Buckingham, R. H., & Heurgue-Hamard, V. (2008). HemK2 protein, encoded on human chromosome 21, methylates translation termination factor eRF1. *FEBS Lett, 582*(16), 2352-2356. doi: 10.1016/j.febslet.2008.05.045

Figaro, S., Wacheul, L., Schillewaert, S., Graille, M., Huvelle, E., Mongeard, R., . . . Heurgue-Hamard, V. (2012). Trm112 is required for Bud23-mediated methylation of the 18S rRNA at position G1575. *Mol Cell Biol, 32*(12), 2254-2267. doi: MCB.06623-11 [pii] 10.1128/MCB.06623-11

Fischer, S., Benz, J., Spath, B., Maier, L. K., Straub, J., Granzow, M., . . . Marchfelder, A. (2010). The archaeal Lsm protein binds to small RNAs. *J Biol Chem, 285*(45), 34429-34438. doi: 10.1074/jbc.M110.118950

Fislage, M., Roovers, M., Tuszynska, I., Bujnicki, J. M., Droogmans, L., & Versees, W. (2012). Crystal structures of the tRNA:m2G6 methyltransferase Trm14/TrmN from two domains of life. *Nucleic Acids Res, 40*(11), 5149-5161. doi: 10.1093/nar/gks163

Fontecave, M., Atta, M., & Mulliez, E. (2004). S-adenosylmethionine: nothing goes to waste. *Trends Biochem Sci, 29*(5), 243-249. doi: 10.1016/j.tibs.2004.03.007

Franckenberg, S., Becker, T., & Beckmann, R. (2012). Structural view on recycling of archaeal and eukaryotic ribosomes after canonical termination and ribosome rescue. *Curr Opin Struct Biol, 22*(6), 786-796. doi: 10.1016/j.sbi.2012.08.002

Francklyn, C. S., & Minajigi, A. (2010). tRNA as an active chemical scaffold for diverse chemical transformations. *FEBS Lett, 584*(2), 366-375. doi: 10.1016/j.febslet.2009.11.045

Fu, D., Brophy, J. A., Chan, C. T., Atmore, K. A., Begley, U., Paules, R. S., . . . Samson, L. D. (2010). Human AlkB homolog ABH8 Is a tRNA methyltransferase required for wobble uridine modification and DNA damage survival. *Mol Cell Biol, 30*(10), 2449-2459. doi: MCB.01604-09 [pii] 10.1128/MCB.01604-09

Fu, L., Guerrero, C. R., Zhong, N., Amato, N. J., Liu, Y., Liu, S., . . . Wang, Y. (2014). Tet-mediated formation of 5-hydroxymethylcytosine in RNA. *J Am Chem Soc, 136*(33), 11582-11585. doi: 10.1021/ja505305z

Fu, Y., Dai, Q., Zhang, W., Ren, J., Pan, T., & He, C. (2010). The AlkB domain of mammalian ABH8 catalyzes hydroxylation of 5-methoxycarbonylmethyluridine at the wobble position of tRNA. *Angew Chem Int Ed Engl, 49*(47), 8885-8888. doi: 10.1002/anie.201001242

Furuichi, Y. (2015). Discovery of m(7)G-cap in eukaryotic mRNAs. *Proc Jpn Acad Ser B Phys Biol Sci, 91*(8), 394-409. doi: 10.2183/pjab.91.394

Fustin, J. M., Doi, M., Yamaguchi, Y., Hida, H., Nishimura, S., Yoshida, M., . . . Okamura, H. (2013). RNA-methylation-dependent RNA processing controls the speed of the circadian clock. *Cell, 155*(4), 793-806. doi: 10.1016/j.cell.2013.10.026

Galardi, S., Fatica, A., Bachi, A., Scaloni, A., Presutti, C., & Bozzoni, I. (2002). Purified box C/D snoRNPs are able to reproduce site-specific 2'-O-methylation of target RNA in vitro. *Mol Cell Biol, 22*(19), 6663-6668.

Gavin, A. C., Bosche, M., Krause, R., Grandi, P., Marzioch, M., Bauer, A., . . . Superti-Furga, G. (2002). Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature, 415*(6868), 141-147. doi: 10.1038/415141a

415141a [pii]

Giaever, G., Chu, A. M., Ni, L., Connelly, C., Riles, L., Veronneau, S., . . . Johnston, M. (2002). Functional profiling of the Saccharomyces cerevisiae genome. *Nature, 418*(6896), 387-391. doi: 10.1038/nature00935

Gilbert, W. V., Bell, T. A., & Schaening, C. (2016). Messenger RNA modifications: Form, distribution, and function. *Science, 352*(6292), 1408-1412. doi: 10.1126/science.aad8711

Glatt, S., Letoquart, J., Faux, C., Taylor, N. M., Seraphin, B., & Muller, C. W. (2012). The Elongator subcomplex Elp456 is a hexameric RecA-like ATPase. *Nat Struct Mol Biol, 19*(3), 314-320. doi: nsmb.2234 [pii] 10.1038/nsmb.2234

Glatt, S., Zabel, R., Kolaj-Robin, O., Onuma, O. F., Baudin, F., Graziadei, A., . . . Muller, C. W. (2016). Structural basis for tRNA modification by Elp3 from Dehalococcoides mccartyi. *Nat Struct Mol Biol, 23*(9), 794-802. doi: 10.1038/nsmb.3265 http://www.nature.com/nsmb/journal/v23/n9/abs/nsmb.3265.html#supplementary-information

Graille, M., Figaro, S., Kervestin, S., Buckingham, R. H., Liger, D., & Heurgue-Hamard, V. (2012). Methylation of class I translation termination factors: structural and functional aspects. *Biochimie, 94*(7), 1533-1543. doi: 10.1016/j.biochi.2012.01.005

Graille, M., Heurgue-Hamard, V., Champ, S., Mora, L., Scrima, N., Ulryck, N., . . . Buckingham, R. H. (2005). Molecular basis for bacterial class I release factor methylation by PrmC. *Mol Cell, 20*(6), 917-927.

Grandi, P., Rybin, V., Bassler, J., Petfalski, E., Strauss, D., Marzioch, M., . . . Hurt, E. (2002). 90S pre-ribosomes include the 35S pre-rRNA, the U3 snoRNP, and 40S subunit processing factors but predominantly lack 60S synthesis factors. *Mol Cell, 10*(1), 105-115.

Granneman, S., & Baserga, S. J. (2004). Ribosome biogenesis: of knobs and RNA processing. *Exp Cell Res, 296*(1), 43-50. doi: 10.1016/j.yexcr.2004.03.016

Grosjean, H. (2015). RNA modification: the Golden Period 1995-2015. *RNA, 21*(4), 625-626. doi: 10.1261/rna.049866.115

Grosjean, H., de Crecy-Lagard, V., & Marck, C. (2010). Deciphering synonymous codons in the three domains of life: co-evolution with specific tRNA modification enzymes. *FEBS Lett, 584*(2), 252-264. doi: 10.1016/j.febslet.2009.11.052

Grosjean, H., Gaspin, C., Marck, C., Decatur, W. A., & de Crecy-Lagard, V. (2008). RNomics and Modomics in the halophilic archaea Haloferax volcanii: identification of RNA modification genes. *BMC Genomics, 9*, 470. doi: 10.1186/1471-2164-9-470

Gu, C., Begley, T. J., & Dedon, P. C. (2014). tRNA modifications regulate translation during cellular stress. *FEBS Lett, 588*(23), 4287-4296. doi: 10.1016/j.febslet.2014.09.038

Gu, T., He, H., Zhang, Y., Han, Z., Hou, G., Zeng, T., . . . Wu, Q. (2012). Trmt112 gene expression in mouse embryonic development. *Acta Histochem Cytochem, 45*(2), 113-119. doi: 10.1267/ahc.11047

Gumienny, R., Jedlinski, D. J., Schmidt, A., Gypas, F., Martin, G., Vina-Vilaseca, A., & Zavolan, M. (2016). High-throughput identification of C/D box snoRNA targets with CLIP and RiboMeth-seq. *Nucleic Acids Res*. doi: 10.1093/nar/gkw1321

Gupta, R. (1984). Halobacterium volcanii tRNAs. Identification of 41 tRNAs covering all amino acids, and the sequences of 33 class I tRNAs. *Journal of Biological Chemistry, 259*(15), 9461-9471.

Halbach, F., Rode, M., & Conti, E. (2012). The crystal structure of S. cerevisiae Ski2, a DExH helicase associated with the cytoplasmic functions of the exosome. *RNA, 18*(1), 124-134. doi: 10.1261/rna.029553.111

Hamey, J. J., Winter, D. L., Yagoub, D., Overall, C. M., Hart-Smith, G., & Wilkins, M. R. (2016). Novel N-terminal and Lysine Methyltransferases That Target Translation Elongation Factor 1A in Yeast and Human. *Mol Cell Proteomics, 15*(1), 164-176. doi: 10.1074/mcp.M115.052449

Hartman, A. L., Norais, C., Badger, J. H., Delmas, S., Haldenby, S., Madupu, R., . . . Eisen, J. A. (2010). The complete genome sequence of Haloferax volcanii DS2, a model archaeon. *PLoS One, 5*(3), e9605. doi: 10.1371/journal.pone.0009605

Hashem, Y., des Georges, A., Fu, J., Buss, S. N., Jossinet, F., Jobe, A., . . . Frank, J. (2013). High-resolution cryo-electron microscopy structure of the Trypanosoma brucei ribosome. *Nature, 494*(7437), 385-389. doi: 10.1038/nature11872

Henras, A. K., Soudet, J., Gerus, M., Lebaron, S., Caizergues-Ferrer, M., Mougin, A., & Henry, Y. (2008). The post-transcriptional steps of eukaryotic ribosome biogenesis. *Cell Mol Life Sci, 65*(15), 2334-2359. doi: 10.1007/s00018-008-8027-0

Heurgue-Hamard, V., Champ, S., Engstrom, A., Ehrenberg, M., & Buckingham, R. H. (2002). The hemK gene in Escherichia coli encodes the N(5)-glutamine methyltransferase that modifies peptide release factors. *EMBO J, 21*(4), 769-778. doi: 10.1093/emboj/21.4.769

Heurgue-Hamard, V., Champ, S., Mora, L., Merkulova-Rainon, T., Kisselev, L. L., & Buckingham, R. H. (2005). The glutamine residue of the conserved GGQ motif in Saccharomyces cerevisiae release factor eRF1 is methylated by the product of the YDR140w gene. *J Biol Chem, 280*(4), 2439-2445. doi: M407252200 [pii]10.1074/jbc.M407252200

Heurgue-Hamard, V., Graille, M., Scrima, N., Ulryck, N., Champ, S., van Tilbeurgh, H., & Buckingham, R. H. (2006). The zinc finger protein Ynr046w is plurifunctional and a component of the eRF1 methyltransferase in yeast. *J Biol Chem, 281*(47), 36140-36148.

Hinnebusch, A. G., & Lorsch, J. R. (2012). The mechanism of eukaryotic translation initiation: new insights and challenges. *Cold Spring Harb Perspect Biol, 4*(10). doi: 10.1101/cshperspect.a011544

Hirata, A., Nishiyama, S., Tamura, T., Yamauchi, A., & Hori, H. (2016). Structural and functional analyses of the archaeal tRNA m2G/m22G10 methyltransferase aTrm11 provide mechanistic insights into site specificity of a tRNA methyltransferase that contains common RNA-binding modules. *Nucleic Acids Res, 44*(13), 6377-6390. doi: 10.1093/nar/gkw561

Hirel, P. H., Schmitter, M. J., Dessen, P., Fayat, G., & Blanquet, S. (1989). Extent of N-terminal methionine excision from Escherichia coli proteins is governed by the side-chain length of the penultimate amino acid. *Proc Natl Acad Sci U S A, 86*(21), 8247-8251.

Hocine, S., Singer, R. H., & Grunwald, D. (2010). RNA processing and export. *Cold Spring Harb Perspect Biol, 2*(12), a000752. doi: 10.1101/cshperspect.a000752

Hoernes, T. P., & Erlacher, M. D. (2017). Translating the epitranscriptome. *Wiley Interdiscip Rev RNA, 8*(1). doi: 10.1002/wrna.1375

Holm, L., & Rosenstrom, P. (2010). Dali server: conservation mapping in 3D. *Nucleic Acids Res, 38*(Web Server issue), W545-549. doi: gkq366 [pii] 10.1093/nar/gkq366

Hopper, A. K. (2013). Transfer RNA post-transcriptional processing, turnover, and subcellular dynamics in the yeast Saccharomyces cerevisiae. *Genetics, 194*(1), 43-67. doi: 10.1534/genetics.112.147470

Hori, H. (2017). Transfer RNA methyltransferases with a SpoU-TrmD (SPOUT) fold and their modified nucleosides in tRNA. *Biomolecules, 7*(1). doi: 10.3390/biom7010023

Hou, Y. M., & Perona, J. J. (2010). Stereochemical mechanisms of tRNA methyltransferases. *FEBS Lett, 584*(2), 278-286. doi: 10.1016/j.febslet.2009.11.075

Hu, Z., Qin, Z., Wang, M., Xu, C., Feng, G., Liu, J., . . . Hu, Y. (2010). The Arabidopsis SMO2, a homologue of yeast TRM112, modulates progression of cell division during organ growth. *Plant J, 61*(4), 600-610. doi: TPJ4085 [pii] 10.1111/j.1365-313X.2009.04085.x

Huang, B., Lu, J., & Bystrom, A. S. (2008). A genome-wide screen identifies genes required for formation of the wobble nucleoside 5-methoxycarbonylmethyl-2-thiouridine in Saccharomyces cerevisiae. *RNA, 14*(10), 2183-2194. doi: rna.1184108 [pii] 10.1261/rna.1184108

Hussain, T., Llacer, J. L., Fernandez, I. S., Munoz, A., Martin-Marcos, P., Savva, C. G., . . . Ramakrishnan, V. (2014). Structural changes enable start codon recognition by the eukaryotic translation initiation complex. *Cell, 159*(3), 597-607. doi: 10.1016/j.cell.2014.10.001

Isaksson, L. A., & Phillips, J. H. (1968). Studies on microbial RNA. V. A comparison of the in vivo methylated components of ribosomal RNA from Escherichia coli and Saccharomyces cerevisiae. *Biochim Biophys Acta, 155*(1), 63-71.

Izaurralde, E., McGuigan, C., & Mattaj, I. W. (1995). Nuclear localization of a cap-binding protein complex. *Cold Spring Harb Symp Quant Biol, 60*, 669-675.

Jackman, J. E., & Alfonzo, J. D. (2013). Transfer RNA modifications: nature's combinatorial chemistry playground. *Wiley Interdiscip Rev RNA, 4*(1), 35-48. doi: 10.1002/wrna.1144

Jackson, R. J., Hellen, C. U., & Pestova, T. V. (2010). The mechanism of eukaryotic translation initiation and principles of its regulation. *Nat Rev Mol Cell Biol, 11*(2), 113-127. doi: 10.1038/nrm2838

Jackson, R. J., Hellen, C. U., & Pestova, T. V. (2012). Termination and post-termination events in eukaryotic translation. *Adv Protein Chem Struct Biol, 86*, 45-93. doi: 10.1016/B978-0-12-386497-0.00002-5

Jangani, M., Poolman, T. M., Matthews, L., Yang, N., Farrow, S. N., Berry, A., . . . Ray, D. W. (2014). The methyltransferase WBSCR22/Merm1 enhances glucocorticoid receptor function and is regulated in lung inflammation and cancer. *J Biol Chem, 289*(13), 8931-8946. doi: 10.1074/jbc.M113.540906

Jenner, L., Melnikov, S., Garreau de Loubresse, N., Ben-Shem, A., Iskakova, M., Urzhumtsev, A., . . . Yusupov, M. (2012). Crystal structure of the 80S yeast ribosome. *Curr Opin Struct Biol, 22*(6), 759-767. doi: 10.1016/j.sbi.2012.07.013

Kabsch, W. (1993). Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. Appl. Cryst., 26*, 795-800.

Kalhor, H. R., & Clarke, S. (2003). Novel methyltransferase for modified uridine residues at the wobble position of tRNA. *Mol Cell Biol, 23*(24), 9283-9292.

Kapp, L. D., & Lorsch, J. R. (2004). The molecular mechanics of eukaryotic translation. *Annu Rev Biochem, 73*, 657-704. doi: 10.1146/annurev.biochem.73.030403.080419

Keller, L., Xu, W., Wang, H. X., Winblad, B., Fratiglioni, L., & Graff, C. (2011). The obesity related gene, FTO, interacts with APOE, and is associated with Alzheimer's disease risk: a prospective cohort study. *J Alzheimers Dis, 23*(3), 461-469. doi: 10.3233/JAD-2010-101068

Kiosze-Becker, K., Ori, A., Gerovac, M., Heuer, A., Nurenberg-Goloub, E., Rashid, U. J., . . . Tampe, R. (2016). Structure of the ribosome post-recycling complex probed by chemical cross-linking and mass spectrometry. *Nat Commun, 7*, 13248. doi: 10.1038/ncomms13248

Klaholz, B. P. (2011). Molecular recognition and catalysis in translation termination complexes. *Trends Biochem Sci, 36*(5), 282-292. doi: 10.1016/j.tibs.2011.02.001

Klinge, S., Voigts-Hoffmann, F., Leibundgut, M., & Ban, N. (2012). Atomic structures of the eukaryotic ribosome. *Trends Biochem Sci, 37*(5), 189-198. doi: 10.1016/j.tibs.2012.02.007

Kobayashi, K., Saito, K., Ishitani, R., Ito, K., & Nureki, O. (2012). Structural basis for translation termination by archaeal RF1 and GTP-bound EF1alpha complex. *Nucleic Acids Res, 40*(18), 9319-9328. doi: 10.1093/nar/gks660

Kohli, M., Riska, S. M., Mahoney, D. W., Chai, H. S., Hillman, D. W., Rider, D. N., . . . Cerhan, J. R. (2012). Germline predictors of androgen deprivation therapy response in advanced prostate cancer. *Mayo Clin Proc, 87*(3), 240-246. doi: 10.1016/j.mayocp.2011.09.009

Kong, C., Ito, K., Walsh, M. A., Wada, M., Liu, Y., Kumar, S., . . . Song, H. (2004). Crystal structure and functional analysis of the eukaryotic class II release factor eRF3 from S. pombe. *Mol Cell, 14*(2), 233-245.

Kozbial, P. Z., & Mushegian, A. R. (2005). Natural history of S-adenosylmethionine-binding proteins. *BMC Struct Biol, 5*, 19. doi: 10.1186/1472-6807-5-19

Kressler, D., Hurt, E., & Bassler, J. (2010). Driving ribosome assembly. *Biochim Biophys Acta, 1803*(6), 673-683. doi: 10.1016/j.bbamcr.2009.10.009

Krogh, N., Jansson, M. D., Hafner, S. J., Tehler, D., Birkedal, U., Christensen-Dalsgaard, M., . . . Nielsen, H. (2016). Profiling of 2'-O-Me in human rRNA reveals a subset of fractionally modified positions and provides evidence for ribosome heterogeneity. *Nucleic Acids Res, 44*(16), 7884-7895. doi: 10.1093/nar/gkw482

Kryuchkova, P., Grishin, A., Eliseev, B., Karyagina, A., Frolova, L., & Alkalaeva, E. (2013). Two-step model of stop codon recognition by eukaryotic release factor eRF1. *Nucleic Acids Res, 41*(8), 4573-4586. doi: 10.1093/nar/gkt113

Kusevic, D., Kudithipudi, S., & Jeltsch, A. (2016). Substrate Specificity of the HEMK2 Protein Glutamine Methyltransferase and Identification of Novel Substrates. *J Biol Chem, 291*(12), 6124-6133. doi: 10.1074/jbc.M115.711952

Kushnirov, V. V., Ter-Avanesyan, M. D., Telckov, M. V., Surguchov, A. P., Smirnov, V. N., & Inge-Vechtomov, S. G. (1988). Nucleotide sequence of the SUP2 (SUP35) gene of Saccharomyces cerevisiae. *Gene, 66*(1), 45-54.

Leihne, V., Kirpekar, F., Vagbo, C. B., van den Born, E., Krokan, H. E., Grini, P. E., . . . Falnes, P. O. (2011). Roles of Trm9- and ALKBH8-like proteins in the formation of modified wobble uridines in Arabidopsis tRNA. *Nucleic Acids Res, 39*(17), 7688-7701. doi: 10.1093/nar/gkr406

Letoquart, J., Huvelle, E., Wacheul, L., Bourgeois, G., Zorbas, C., Graille, M., . . . Lafontaine, D. L. (2014). Structural and functional studies of Bud23-Trm112 reveal 18S rRNA N7-G1575 methylation occurs on late 40S precursor ribosomes. *Proc Natl Acad Sci U S A, 111*(51), E5518-5526. doi: 10.1073/pnas.1413089111

Letoquart, J., Tran, N. V., Caroline, V., Aleksandrov, A., Lazar, N., van Tilbeurgh, H., . . . Graille, M. (2015a). Insights into molecular plasticity in protein complexes from Trm9-Trm112 tRNA modifying enzyme crystal structure. *Nucleic Acids Res, 43*(22), 10989-11002.

Li, X., Xiong, X., Wang, K., Wang, L., Shu, X., Ma, S., & Yi, C. (2016). Transcriptome-wide mapping reveals reversible and dynamic N(1)-methyladenosine methylome. *Nat Chem Biol, 12*(5), 311-316. doi: 10.1038/nchembio.2040

Liang, X. H., Liu, Q., & Fournier, M. J. (2007). rRNA modifications in an intersubunit bridge of the ribosome strongly affect both ribosome biogenesis and activity. *Mol Cell, 28*(6), 965-977. doi: 10.1016/j.molcel.2007.10.012

Liang, X. H., Liu, Q., & Fournier, M. J. (2009). Loss of rRNA modifications in the decoding center of the ribosome impairs translation and strongly delays pre-rRNA processing. *RNA, 15*(9), 1716-1728. doi: 10.1261/rna.1724409

Liger, D., Mora, L., Lazar, N., Figaro, S., Henri, J., Scrima, N., . . . Graille, M. (2011). Mechanism of activation of methyltransferases involved in translation by the Trm112 'hub' protein. *Nucleic Acids Res, 39*(14), 6249-6259. doi: gkr176 [pii] 10.1093/nar/gkr176

Liu, J., Yue, Y., Han, D., Wang, X., Fu, Y., Zhang, L., . . . He, C. (2014). A METTL3-METTL14 complex mediates mammalian nuclear RNA N6-adenosine methylation. *Nat Chem Biol, 10*(2), 93-95. doi: 10.1038/nchembio.1432

Liu, P., Nie, S., Li, B., Yang, Z. Q., Xu, Z. M., Fei, J., . . . Xu, G. L. (2010). Deficiency in a glutamine-specific methyltransferase for the release factor causes mouse embryonic lethality. *Mol Cell Biol, 30*(17), 4245-4253. doi: MCB.00218-10 [pii] 10.1128/MCB.00218-10

Liu, R. J., Zhou, M., Fang, Z. P., Wang, M., Zhou, X. L., & Wang, E. D. (2013). The tRNA recognition mechanism of the minimalist SPOUT methyltransferase, TrmL. *Nucleic Acids Res, 41*(16), 7828-7842. doi: 10.1093/nar/gkt568

Loenen, W. A. (2006). S-adenosylmethionine: jack of all trades and master of everything? *Biochem Soc Trans, 34*(Pt 2), 330-333. doi: 10.1042/BST20060330

Loh, P. G., & Song, H. (2010). Structural and mechanistic insights into translation termination. *Curr Opin Struct Biol, 20*(1), 98-103. doi: 10.1016/j.sbi.2009.12.005

Lorenz, C., Lunse, C. E., & Morl, M. (2017). tRNA Modifications: Impact on Structure and Thermal Adaptation. *Biomolecules, 7*(2). doi: 10.3390/biom7020035

Lu, J., Huang, B., Esberg, A., Johansson, M. J., & Bystrom, A. S. (2005). The Kluyveromyces lactis gamma-toxin targets tRNA anticodons. *RNA, 11*(11), 1648-1654. doi: 11/11/1648 [pii] 10.1261/rna.2172105

Lunardi, A., Di Minin, G., Provero, P., Dal Ferro, M., Carotti, M., Del Sal, G., & Collavin, L. (2010). A genome-scale protein interaction profile of Drosophila p53 uncovers additional nodes of the human p53 network. *Proc Natl Acad Sci U S A, 107*(14), 6322-6327. doi: 10.1073/pnas.1002447107

Lyu, Z., & Whitman, W. B. (2017). Evolution of the archaeal and mammalian information processing systems: towards an archaeal model for human disease. *Cell Mol Life Sci, 74*(2), 183-212. doi: 10.1007/s00018-016-2286-y

Maden, B. E. (1986). Identification of the locations of the methyl groups in 18 S ribosomal RNA from Xenopus laevis and man. *J Mol Biol, 189*(4), 681-699.

Mangus, D. A., Evans, M. C., & Jacobson, A. (2003). Poly(A)-binding proteins: multifunctional scaffolds for the post-transcriptional control of gene expression. *Genome Biol, 4*(7), 223. doi: 10.1186/gb-2003-4-7-223

Marchand, V., Pichot, F., Thuring, K., Ayadi, L., Freund, I., Dalpke, A., . . . Motorin, Y. (2017). Next-Generation Sequencing-Based RiboMethSeq Protocol for Analysis of tRNA 2'-O-Methylation. *Biomolecules, 7*(1). doi: 10.3390/biom7010013

Martin, J. L., & McMillan, F. M. (2002). SAM (dependent) I AM: the S-adenosylmethionine-dependent methyltransferase fold. *Curr Opin Struct Biol, 12*(6), 783-793.

Mazauric, M. H., Dirick, L., Purushothaman, S. K., Bjork, G. R., & Lapeyre, B. (2010). Trm112p is a 15-kDa zinc finger protein essential for the activity of two tRNA and one protein methyltransferases in yeast. *J Biol Chem, 285*(24), 18505-18515. doi: 10.1074/jbc.M110.113100

McCarthy, A. A., & McCarthy, J. G. (2007). The structure of two N-methyltransferases from the caffeine biosynthetic pathway. *Plant Physiol, 144*(2), 879-889. doi: pp.106.094854 [pii] 10.1104/pp.106.094854

McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., & Read, R. J. (2007). Phaser crystallographic software. *Journal of Applied Crystallography, 40*(4), 658-674. doi: doi:10.1107/S0021889807021206

Melnikov, S., Ben-Shem, A., Garreau de Loubresse, N., Jenner, L., Yusupova, G., & Yusupov, M. (2012). One core, two shells: bacterial and eukaryotic ribosomes. *Nat Struct Mol Biol, 19*(6), 560-567. doi: 10.1038/nsmb.2313

Menezes, S., Gaston, K. W., Krivos, K. L., Apolinario, E. E., Reich, N. O., Sowers, K. R., . . . Perona, J. J. (2011). Formation of m2G6 in Methanocaldococcus jannaschii tRNA catalyzed by the novel methyltransferase Trm14. *Nucleic Acids Res, 39*(17), 7641-7655. doi: 10.1093/nar/gkr475

Merla, G., Ucla, C., Guipponi, M., & Reymond, A. (2002). Identification of additional transcripts in the Williams-Beuren syndrome critical region. *Hum Genet, 110*(5), 429-438. doi: 10.1007/s00439-002-0710-x

Meury, J., & Kohiyama, M. (1989). ATP is required for K+ active transport in the archaebacterium Haloferax volcanii. *Archives of Microbiology, 151*(6), 530-536. doi: 10.1007/BF00454870

Meyer, K. D., Saletore, Y., Zumbo, P., Elemento, O., Mason, C. E., & Jaffrey, S. R. (2012). Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. *Cell, 149*(7), 1635-1646. doi: 10.1016/j.cell.2012.05.003

Mora, L., Heurgue-Hamard, V., de Zamaroczy, M., Kervestin, S., & Buckingham, R. H. (2007). Methylation of bacterial release factors RF1 and RF2 is required for normal translation termination in vivo. *J Biol Chem, 282*(49), 35638-35645. doi: M706076200 [pii] 10.1074/jbc.M706076200

Motorin, Y., & Helm, M. (2011). RNA nucleotide methylation. *Wiley Interdiscip Rev RNA, 2*(5), 611-631. doi: 10.1002/wrna.79

Nakahigashi, K., Kubo, N., Narita, S., Shimaoka, T., Goto, S., Oshima, T., . . . Inokuchi, H. (2002). HemK, a class of protein methyl transferase with similarity to DNA methyl transferases, methylates polypeptide chain release factors, and hemK knockout induces defects in translational termination. *Proc Natl Acad Sci U S A, 99*(3), 1473-1478.

Nakamura, Y., & Ito, K. (2003). Making sense of mimic in translation termination. *Trends Biochem Sci, 28*(2), 99-105. doi: 10.1016/S0968-0004(03)00006-9

Nakazawa, Y., Arai, H., & Fujita, N. (2011). The novel metastasis promoter Merm1/Wbscr22 enhances tumor cell survival in the vasculature by suppressing Zac1/p53-dependent apoptosis. *Cancer Res, 71*(3), 1146-1155. doi: 10.1158/0008-5472.CAN-10-2695

Nesterchuk, M. V., Sergiev, P. V., & Dontsova, O. A. (2011). Posttranslational Modifications of Ribosomal Proteins in Escherichia coli. *Acta Naturae, 3*(2), 22-33.

Neumann, P., Lakomek, K., Naumann, P. T., Erwin, W. M., Lauhon, C. T., & Ficner, R. (2014). Crystal structure of a 4-thiouridine synthetase-RNA complex reveals specificity of tRNA U8 modification. *Nucleic Acids Res, 42*(10), 6673-6685. doi: 10.1093/nar/gku249

Nie, D. S., Liu, Y. B., & Lu, G. X. (2009). Cloning and primarily function study of two novel putative N5-glutamine methyltransferase (Hemk) splice variants from mouse stem cells. *Mol Biol Rep, 36*(8), 2221-2228. doi: 10.1007/s11033-008-9437-7

Nishimura, K., Hosaka, T., Tokuyama, S., Okamoto, S., & Ochi, K. (2007). Mutations in rsmG, encoding a 16S rRNA methyltransferase, result in low-level streptomycin resistance and antibiotic overproduction in Streptomyces coelicolor A3(2). *J Bacteriol, 189*(10), 3876-3883. doi: 10.1128/JB.01776-06

Noma, A., Sakaguchi, Y., & Suzuki, T. (2009). Mechanistic characterization of the sulfur-relay system for eukaryotic 2-thiouridine biogenesis at tRNA wobble positions. *Nucleic Acids Research, 37*(4), 1335-1352. doi: 10.1093/nar/gkn1023

Noon, K. R., Bruenger, E., & McCloskey, J. A. (1998). Posttranscriptional modifications in 16S and 23S rRNAs of the archaeal hyperthermophile Sulfolobus solfataricus. *J Bacteriol, 180*(11), 2883-2888.

O'Hagan, D., & Schmidberger, J. W. (2010). Enzymes that catalyse SN2 reaction mechanisms. *Nat Prod Rep, 27*(6), 900-918. doi: 10.1039/b919371p

Oeffinger, M. (2016). Structural biology: Moulding the ribosome. *Nature, 537*(7618), 38-40. doi: 10.1038/537038a

Oerum, S., Degut, C., Barraud, P., & Tisne, C. (2017). m1A Post-Transcriptional Modification in tRNAs. *Biomolecules, 7*(1). doi: 10.3390/biom7010020

Okada, K., Muneyoshi, Y., Endo, Y., & Hori, H. (2009). Production of yeast (m2G10) methyltransferase (Trm11 and Trm112 complex) in a wheat germ cell-free translation system. *Nucleic Acids Symp Ser (Oxf)*(53), 303-304. doi: nrp152 [pii] 10.1093/nass/nrp152

Oliva, R., Cavallo, L., & Tramontano, A. (2006). Accurate energies of hydrogen bonded nucleic acid base pairs and triplets in tRNA tertiary interactions. *Nucleic Acids Res, 34*(3), 865-879. doi: 10.1093/nar/gkj491

Ounap, K., Kasper, L., Kurg, A., & Kurg, R. (2013). The human WBSCR22 protein is involved in the biogenesis of the 40S ribosomal subunits in mammalian cells. *PLoS One, 8*(9), e75686. doi: 10.1371/journal.pone.0075686 PONE-D-13-07916 [pii]

Ounap, K., Leetsi, L., Matsoo, M., & Kurg, R. (2015). The Stability of Ribosome Biogenesis Factor WBSCR22 Is Regulated by Interaction with TRMT112 via Ubiquitin-Proteasome Pathway. *PLoS One, 10*(7), e0133841. doi: 10.1371/journal.pone.0133841

Pastore, C., Topalidou, I., Forouhar, F., Yan, A. C., Levy, M., & Hunt, J. F. (2012). Crystal structure and RNA binding properties of the RNA recognition motif (RRM) and AlkB domains in human AlkB homolog 8 (ABH8), an enzyme catalyzing tRNA hypermodification. *J Biol Chem, 287*(3), 2130-2143. doi: M111.286187 [pii] 10.1074/jbc.M111.286187

Patil, A., Chan, C. T., Dyavaiah, M., Rooney, J. P., Dedon, P. C., & Begley, T. J. (2012). Translational infidelity-induced protein stress results from a deficiency in Trm9-catalyzed tRNA modifications. *RNA Biol, 9*(7), 990-1001. doi: 10.4161/rna.20531

Pavlov, M. Y., Freistroffer, D. V., Dincbas, V., MacDougall, J., Buckingham, R. H., & Ehrenberg, M. (1998). A direct estimation of the context effect on the efficiency of termination. *J Mol Biol, 284*(3), 579-590. doi: S0022-2836(98)92220-3 [pii] 10.1006/jmbi.1998.2220

Peifer, C., Sharma, S., Watzinger, P., Lamberth, S., Kotter, P., & Entian, K. D. (2013). Yeast Rrp8p, a novel methyltransferase responsible for m1A 645 base modification of 25S rRNA. *Nucleic Acids Res, 41*(2), 1151-1163. doi: 10.1093/nar/gks1102

Petrossian, T. C., & Clarke, S. G. (2009). Multiple Motif Scanning to identify methyltransferases from the yeast proteome. *Mol Cell Proteomics, 8*(7), 1516-1526. doi: 10.1074/mcp.M900025-MCP200

Petrossian, T. C., & Clarke, S. G. (2011). Uncovering the human methyltransferasome. *Mol Cell Proteomics, 10*(1), M110 000976. doi: 10.1074/mcp.M110.000976

Pfeiffer, F., & Oesterhelt, D. (2015). A manual curation strategy to improve genome annotation: application to a set of haloarchael genomes. *Life (Basel), 5*(2), 1427-1444. doi: 10.3390/life5021427

Phillips, G., & de Crecy-Lagard, V. (2011). Biosynthesis and function of tRNA modifications in Archaea. *Curr Opin Microbiol, 14*(3), 335-341. doi: 10.1016/j.mib.2011.03.001

Piekna-Przybylska, D., Decatur, W. A., & Fournier, M. J. (2007). New bioinformatic tools for analysis of nucleotide modifications in eukaryotic rRNA. *RNA, 13*(3), 305-312. doi: 10.1261/rna.373107

Piekna-Przybylska, D., Przybylski, P., Baudin-Baillieu, A., Rousset, J. P., & Fournier, M. J. (2008). Ribosome performance is enhanced by a rich cluster of pseudouridines in the A-site finger region of the large subunit. *J Biol Chem, 283*(38), 26026-26036. doi: 10.1074/jbc.M803049200

Pierson, W. E., Hoffer, E. D., Keedy, H. E., Simms, C. L., Dunham, C. M., & Zaher, H. S. (2016). Uniformity of Peptide Release Is Maintained by Methylation of Release Factors. *Cell Rep, 17*(1), 11-18. doi: 10.1016/j.celrep.2016.08.085

Pisarev, A. V., Hellen, C. U., & Pestova, T. V. (2007). Recycling of eukaryotic posttermination ribosomal complexes. *Cell, 131*(2), 286-299. doi: 10.1016/j.cell.2007.08.041

Pisarev, A. V., Skabkin, M. A., Pisareva, V. P., Skabkina, O. V., Rakotondrafara, A. M., Hentze, M. W., . . . Pestova, T. V. (2010). The role of ABCE1 in eukaryotic posttermination ribosomal recycling. *Mol Cell, 37*(2), 196-210. doi: 10.1016/j.molcel.2009.12.034

Polevoda, B., & Sherman, F. (2007). Methylation of proteins involved in translation. *Mol Microbiol, 65*(3), 590-606. doi: 10.1111/j.1365-2958.2007.05831.x

Polevoda, B., Span, L., & Sherman, F. (2006). The yeast translation release factors Mrf1p and Sup45p (eRF1) are methylated, respectively, by the methyltransferases Mtq1p and Mtq2p. *J Biol Chem, 281*(5), 2562-2571. doi: M507651200 [pii] 10.1074/jbc.M507651200

Preis, A., Heuer, A., Barrio-Garcia, C., Hauser, A., Eyler, D. E., Berninghausen, O., . . . Beckmann, R. (2014). Cryoelectron microscopic structures of eukaryotic translation termination complexes containing eRF1-eRF3 or eRF1-ABCE1. *Cell Rep, 8*(1), 59-65. doi: 10.1016/j.celrep.2014.04.058

Purushothaman, S. K., Bujnicki, J. M., Grosjean, H., & Lapeyre, B. (2005). Trm11p and Trm112p are both required for the formation of 2-methylguanosine at position 10 in yeast tRNA. *Mol Cell Biol, 25*(11), 4359-4370. doi: 10.1128/MCB.25.11.4359-4370.2005

Qian, C., & Zhou, M. M. (2006). SET domain protein lysine methyltransferases: Structure, specificity and catalysis. *Cell Mol Life Sci, 63*(23), 2755-2763. doi: 10.1007/s00018-006-6274-5

Rabl, J., Leibundgut, M., Ataide, S. F., Haag, A., & Ban, N. (2011). Crystal structure of the eukaryotic 40S ribosomal subunit in complex with initiation factor 1. *Science, 331*(6018), 730-736. doi: 10.1126/science.1198308

Raina, M., & Ibba, M. (2014). tRNAs as regulators of biological processes. *Front Genet, 5*, 171. doi: 10.3389/fgene.2014.00171

Ramakrishnan, V. (2011). Molecular biology. The eukaryotic ribosome. *Science, 331*(6018), 681-682. doi: 10.1126/science.1202093

Ramanathan, A., Robb, G. B., & Chan, S. H. (2016). mRNA capping: biological functions and applications. *Nucleic Acids Res, 44*(16), 7511-7526. doi: 10.1093/nar/gkw551

Ramirez, V., Gonzalez, B., Lopez, A., Castello, M. J., Gil, M. J., Etherington, G. J., . . . Vera, P. (2015). Loss of a Conserved tRNA Anticodon Modification Perturbs Plant Immunity. *PLoS Genet, 11*(10), e1005586. doi: 10.1371/journal.pgen.1005586

Rana, A. K., & Ankri, S. (2016). Reviving the RNA World: An Insight into the Appearance of RNA Methyltransferases. *Front Genet, 7*, 99. doi: 10.3389/fgene.2016.00099

Ranjan, N., & Rodnina, M. V. (2016). tRNA wobble modifications and protein homeostasis. *Translation (Austin), 4*(1), e1143076. doi: 10.1080/21690731.2016.1143076

Rivera, M. C., Jain, R., Moore, J. E., & Lake, J. A. (1998). Genomic evidence for two functionally distinct gene classes. *Proc Natl Acad Sci U S A, 95*(11), 6239-6244.

Robert, X., & Gouet, P. (2014). Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res, 42*(Web Server issue), W320-324. doi: 10.1093/nar/gku316

Rodnina, M. V., & Wintermeyer, W. (2011). The ribosome as a molecular machine: the mechanism of tRNA-mRNA movement in translocation. *Biochem Soc Trans, 39*(2), 658-662. doi: 10.1042/BST0390658

Roje, S. (2006). S-Adenosyl-L-methionine: beyond the universal methyl group donor. *Phytochemistry, 67*(15), 1686-1698. doi: 10.1016/j.phytochem.2006.04.019

Saito, K., Kobayashi, K., Wada, M., Kikuno, I., Takusagawa, A., Mochizuki, M., . . . Ito, K. (2010). Omnipotent role of archaeal elongation factor 1 alpha (EF1alpha in translational elongation and termination, and quality control of protein synthesis. *Proc Natl Acad Sci U S A, 107*(45), 19242-19247. doi: 10.1073/pnas.1009599107

Sardana, R., & Johnson, A. W. (2012). The methyltransferase adaptor protein Trm112 is involved in biogenesis of both ribosomal subunits. *Mol Biol Cell, 23*(21), 4313-4322. doi: mbc.E12-05-0370 [pii] 10.1091/mbc.E12-05-0370

Sardana, R., Liu, X., Granneman, S., Zhu, J., Gill, M., Papoulas, O., . . . Johnson, A. W. (2015). The DEAH-box helicase Dhr1 dissociates U3 from the pre-rRNA to promote formation of the central pseudoknot. *PLoS Biol, 13*(2), e1002083. doi: 10.1371/journal.pbio.1002083

Sardana, R., White, J. P., & Johnson, A. W. (2013). The rRNA methyltransferase Bud23 shows functional interaction with components of the SSU processome and RNase MRP. *RNA, 19*(6), 828-840. doi: 10.1261/rna.037671.112

Sardana, R., Zhu, J., Gill, M., & Johnson, A. W. (2014). Physical and functional interaction between the methyltransferase Bud23 and the essential DEAH-box RNA helicase Ecm16. *Mol Cell Biol, 34*(12), 2208-2220. doi: 10.1128/MCB.01656-13

Schmeing, T. M., & Ramakrishnan, V. (2009). What recent ribosome structures have revealed about the mechanism of translation. *Nature, 461*(7268), 1234-1242. doi: 10.1038/nature08403

Schneider, T. R., & Sheldrick, G. M. (2002). Substructure solution with SHELXD. *Acta Crystallogr D Biol Crystallogr, 58*(Pt 10 Pt 2), 1772-1779.

Schubert, H. L., Blumenthal, R. M., & Cheng, X. (2003). Many paths to methyltransfer: a chronicle of convergence. *Trends Biochem Sci, 28*(6), 329-335. doi: 10.1016/S0968-0004(03)00090-2

Schubert, H. L., Wilson, K. S., Raux, E., Woodcock, S. C., & Warren, M. J. (1998). The X-ray structure of a cobalamin biosynthetic enzyme, cobalt-precorrin-4 methyltransferase. *Nat Struct Biol, 5*(7), 585-592. doi: 10.1038/846

Schwartz, S. (2016). Cracking the epitranscriptome. *RNA, 22*(2), 169-174. doi: 10.1261/rna.054502.115

Schwartz, S., Mumbach, M. R., Jovanovic, M., Wang, T., Maciag, K., Bushkin, G. G., . . . Regev, A. (2014). Perturbation of m6A writers reveals two distinct classes of mRNA methylation at internal and 5' sites. *Cell Rep, 8*(1), 284-296. doi: 10.1016/j.celrep.2014.05.048

Scotti, M. M., & Swanson, M. S. (2016). RNA mis-splicing in disease. *Nat Rev Genet, 17*(1), 19-32. doi: 10.1038/nrg.2015.3

Selvadurai, K., Wang, P., Seimetz, J., & Huang, R. H. (2014). Archaeal Elp3 catalyzes tRNA wobble uridine modification at C5 via a radical mechanism. *Nat Chem Biol, 10*(10), 810-812. doi: 10.1038/nchembio.1610http://www.nature.com/nchembio/journal/v10/n10/abs/nchembio.1610.html#supplementary-information

Sharma, S., & Lafontaine, D. L. (2015). 'View From A Bridge': A New Perspective on Eukaryotic rRNA Base Modification. *Trends Biochem Sci, 40*(10), 560-575. doi: 10.1016/j.tibs.2015.07.008

Sharma, S., Langhendries, J. L., Watzinger, P., Kotter, P., Entian, K. D., & Lafontaine, D. L. (2015). Yeast Kre33 and human NAT10 are conserved 18S rRNA cytosine acetyltransferases that modify tRNAs assisted by the adaptor Tan1/THUMPD1. *Nucleic Acids Res, 43*(4), 2242-2258. doi: 10.1093/nar/gkv075

Sharma, S., Yang, J., Watzinger, P., Kotter, P., & Entian, K. D. (2013). Yeast Nop2 and Rcm1 methylate C2870 and C2278 of the 25S rRNA, respectively. *Nucleic Acids Res, 41*(19), 9062-9076. doi: 10.1093/nar/gkt679

Shimada, K., Nakamura, M., Anai, S., De Velasco, M., Tanaka, M., Tsujikawa, K., . . . Konishi, N. (2009). A novel human AlkB homologue, ALKBH8, contributes to human bladder cancer progression. *Cancer Res, 69*(7), 3157-3164. doi: 10.1158/0008-5472.CAN-08-3530

Shirai, A., Sadaie, M., Shinmyozu, K., & Nakayama, J. (2010). Methylation of ribosomal protein L42 regulates ribosomal function and stress-adapted cell growth. *J Biol Chem, 285*(29), 22448-22460. doi: 10.1074/jbc.M110.132274

Sledz, P., & Jinek, M. (2016). Structural insights into the molecular mechanism of the m(6)A writer complex. *Elife, 5*. doi: 10.7554/eLife.18434

Sloan, K. E., Warda, A. S., Sharma, S., Entian, K. D., Lafontaine, D. L., & Bohnsack, M. T. (2016). Tuning the ribosome: The influence of rRNA modification on eukaryotic ribosome biogenesis and function. *RNA Biol*, 1-16. doi: 10.1080/15476286.2016.1259781

Song, H., Mugnier, P., Das, A. K., Webb, H. M., Evans, D. R., Tuite, M. F., . . . Barford, D. (2000). The crystal structure of human eukaryotic release factor eRF1--mechanism of stop codon recognition and peptidyl-tRNA hydrolysis. *Cell, 100*(3), 311-321.

Song, J., & Yi, C. (2017). Chemical Modifications to RNA: A New Layer of Gene Expression Regulation. *ACS Chem Biol, 12*(2), 316-325. doi: 10.1021/acschembio.6b00960

Sprinzl, M., Horn, C., Brown, M., Ioudovitch, A., & Steinberg, S. (1998). Compilation of tRNA sequences and sequences of tRNA genes. *Nucleic Acids Res, 26*(1), 148-153.

Squires, J. E., Patel, H. R., Nousch, M., Sibbritt, T., Humphreys, D. T., Parker, B. J., . . . Preiss, T. (2012). Widespread occurrence of 5-methylcytosine in human coding and non-coding RNA. *Nucleic Acids Res, 40*(11), 5023-5033. doi: 10.1093/nar/gks144

Stefanska, B., Cheishvili, D., Suderman, M., Arakelian, A., Huang, J., Hallett, M., . . . Szyf, M. (2014). Genome-wide study of hypomethylated and induced genes in patients with liver cancer unravels novel anticancer targets. *Clin Cancer Res, 20*(12), 3118-3132. doi: 10.1158/1078-0432.CCR-13-0283

Struck, A. W., Thompson, M. L., Wong, L. S., & Micklefield, J. (2012). S-adenosyl-methionine-dependent methyltransferases: highly versatile enzymes in biocatalysis, biosynthesis and other biotechnological applications. *Chembiochem, 13*(18), 2642-2655. doi: 10.1002/cbic.201200556

Studte, P., Zink, S., Jablonowski, D., Bar, C., von der Haar, T., Tuite, M. F., & Schaffrath, R. (2008). tRNA and protein methylase complexes mediate zymocin toxicity in yeast. *Mol Microbiol, 69*(5), 1266-1277. doi: MMI6358 [pii] 10.1111/j.1365-2958.2008.06358.x

Suddath, F. L., Quigley, G. J., McPherson, A., Sneden, D., Kim, J. J., Kim, S. H., & Rich, A. (1974). Three-dimensional structure of yeast phenylalanine transfer RNA at 3.0angstroms resolution. *Nature, 248*(5443), 20-24.

Swinehart, W. E., & Jackman, J. E. (2015). Diversity in mechanism and function of tRNA methyltransferases. *RNA Biol, 12*(4), 398-411. doi: 10.1080/15476286.2015.1008358

Taoka, M., Nobe, Y., Yamaki, Y., Yamauchi, Y., Ishikawa, H., Takahashi, N., . . . Isobe, T. (2016). The complete chemical structure of Saccharomyces cerevisiae rRNA: partial pseudouridylation of U2345 in 25S rRNA by snoRNA snR9. *Nucleic Acids Res, 44*(18), 8951-8961. doi: 10.1093/nar/gkw564

Taylor, D., Unbehaun, A., Li, W., Das, S., Lei, J., Liao, H. Y., . . . Frank, J. (2012). Cryo-EM structure of the mammalian eukaryotic release factor eRF1-eRF3-associated termination complex. *Proc Natl Acad Sci U S A, 109*(45), 18413-18418. doi: 10.1073/pnas.1216730109

Ter-Avanesyan, M. D., Kushnirov, V. V., Dagkesamanskaya, A. R., Didichenko, S. A., Chernoff, Y. O., Inge-Vechtomov, S. G., & Smirnov, V. N. (1993). Deletion analysis of the SUP35 gene of the yeast Saccharomyces cerevisiae reveals two non-overlapping functional regions in the encoded protein. *Mol Microbiol, 7*(5), 683-692.

Terwilliger, T. (2004). SOLVE and RESOLVE: automated structure solution, density modification and model building. *J Synchrotron Radiat, 11*(Pt 1), 49-52.

Terwilliger, T. C., Grosse-Kunstleve, R. W., Afonine, P. V., Moriarty, N. W., Zwart, P. H., Hung, L. W., . . . Adams, P. D. (2008). Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. *Acta Crystallogr D Biol Crystallogr, 64*(Pt 1), 61-69. doi: 10.1107/S090744490705024X

Thomson, E., Ferreira-Cerca, S., & Hurt, E. (2013). Eukaryotic ribosome biogenesis at a glance. *J Cell Sci, 126*(Pt 21), 4815-4821. doi: 10.1242/jcs.111948

Tiedemann, R. E., Zhu, Y. X., Schmidt, J., Shi, C. X., Sereduk, C., Yin, H., . . . Stewart, A. K. (2012). Identification of molecular vulnerabilities in human multiple myeloma cells by RNA interference lethality screening of the druggable genome. *Cancer Res, 72*(3), 757-768. doi: 0008-5472.CAN-11-2781 [pii] 10.1158/0008-5472.CAN-11-2781

Tkaczuk, K. L., Dunin-Horkawicz, S., Purta, E., & Bujnicki, J. M. (2007). Structural and evolutionary bioinformatics of the SPOUT superfamily of methyltransferases. *BMC Bioinformatics, 8*, 73. doi: 10.1186/1471-2105-8-73

Tomikawa, C., Ohira, T., Inoue, Y., Kawamura, T., Yamagishi, A., Suzuki, T., & Hori, H. (2013). Distinct tRNA modifications in the thermo-acidophilic archaeon, Thermoplasma acidophilum. *FEBS Lett, 587*(21), 3575-3580. doi: 10.1016/j.febslet.2013.09.021

Torres, A. G., Batlle, E., & Ribas de Pouplana, L. (2014). Role of tRNA modifications in human diseases. *Trends Mol Med, 20*(6), 306-314. doi: 10.1016/j.molmed.2014.01.008

Towns, W. L., & Begley, T. J. (2012). Transfer RNA methytransferases and their corresponding modifications in budding yeast and humans: activities, predications, and potential roles in human health. *DNA Cell Biol, 31*(4), 434-454. doi: 10.1089/dna.2011.1437

Tu, L. W., & Deutsch, C. (2010). A folding zone in the ribosomal exit tunnel for Kv1.3 helix formation. *J Mol Biol, 396*(5), 1346-1360. doi: 10.1016/j.jmb.2009.12.059

Tuorto, F., & Lyko, F. (2016). Genome recoding by tRNA modifications. *Open Biology, 6*(12). doi: Artn 160287 10.1098/Rsob.160287

Urbonavicius, J., Armengaud, J., & Grosjean, H. (2006). Identity elements required for enzymatic formation of N2,N2-dimethylguanosine from N2-monomethylated derivative and its possible role in avoiding alternative conformations in archaeal tRNA. *J Mol Biol, 357*(2), 387-399. doi: 10.1016/j.jmb.2005.12.087

Vare, V. Y., Eruysal, E. R., Narendran, A., Sarachan, K. L., & Agris, P. F. (2017). Chemical and Conformational Diversity of Modified Nucleosides Affects tRNA Structure and Function. *Biomolecules, 7*(1). doi: 10.3390/biom7010029

Voigts-Hoffmann, F., Klinge, S., & Ban, N. (2012). Structural insights into eukaryotic ribosomes and the initiation of translation. *Curr Opin Struct Biol, 22*(6), 768-777. doi: 10.1016/j.sbi.2012.07.010

Wang, P., Doxtader, K. A., & Nam, Y. (2016). Structural Basis for Cooperative Function of Mettl3 and Mettl14 Methyltransferases. *Mol Cell, 63*(2), 306-317. doi: 10.1016/j.molcel.2016.05.041

Wang, X., Feng, J., Xue, Y., Guan, Z., Zhang, D., Liu, Z., . . . Yin, P. (2016). Structural basis of N(6)-adenosine methylation by the METTL3-METTL14 complex. *Nature, 534*(7608), 575-578. doi: 10.1038/nature18298

Wang, X., Lu, Z., Gomez, A., Hon, G. C., Yue, Y., Han, D., . . . He, C. (2014). N6-methyladenosine-dependent regulation of messenger RNA stability. *Nature, 505*(7481), 117-120. doi: 10.1038/nature12730

Wang, X., Zhao, B. S., Roundtree, I. A., Lu, Z., Han, D., Ma, H., . . . He, C. (2015). N(6)-methyladenosine Modulates Messenger RNA Translation Efficiency. *Cell, 161*(6), 1388-1399. doi: 10.1016/j.cell.2015.05.014

Wang, Y., & Jia, G. (2016). New Edges of RNA Adenosine Methylation Modifications. *Genomics Proteomics Bioinformatics, 14*(3), 172-175. doi: 10.1016/j.gpb.2016.05.003

Wang, Y., Li, Y., Toth, J. I., Petroski, M. D., Zhang, Z., & Zhao, J. C. (2014). N6-methyladenosine modification destabilizes developmental regulators in embryonic stem cells. *Nat Cell Biol, 16*(2), 191-198. doi: 10.1038/ncb2902

Watkins, N. J., & Bohnsack, M. T. (2012). The box C/D and H/ACA snoRNPs: key players in the modification, processing and the dynamic folding of ribosomal RNA. *Wiley Interdiscip Rev RNA, 3*(3), 397-414. doi: 10.1002/wrna.117

White, J., Li, Z., Sardana, R., Bujnicki, J. M., Marcotte, E. M., & Johnson, A. W. (2008). Bud23 methylates G1575 of 18S rRNA and is required for efficient nuclear export of pre-40S subunits. *Mol Cell Biol, 28*(10), 3151-3161. doi: MCB.01674-07 [pii] 10.1128/MCB.01674-07

Wichtowska, D., Turowski, T. W., & Boguta, M. (2013). An interplay between transcription, processing, and degradation determines tRNA levels in yeast. *Wiley Interdiscip Rev RNA, 4*(6), 709-722. doi: 10.1002/wrna.1190

Wilson, D. N., & Doudna Cate, J. H. (2012). The structure and function of the eukaryotic ribosome. *Cold Spring Harb Perspect Biol, 4*(5). doi: 10.1101/cshperspect.a011536

Woese, C. R., & Fox, G. E. (1977). Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proc Natl Acad Sci U S A, 74*(11), 5088-5090.

Wong, L. E., Li, Y., Pillay, S., Frolova, L., & Pervushin, K. (2012). Selectivity of stop codon recognition in translation termination is modulated by multiple conformations of GTS loop in eRF1. *Nucleic Acids Res, 40*(12), 5751-5765. doi: 10.1093/nar/gks192

Woolford, J. L., Jr., & Baserga, S. J. (2013). Ribosome biogenesis in the yeast Saccharomyces cerevisiae. *Genetics, 195*(3), 643-681. doi: 10.1534/genetics.113.153197

Yang, J., Sharma, S., Watzinger, P., Hartmann, J. D., Kotter, P., & Entian, K. D. (2016). Mapping of Complete Set of Ribose and Base Modifications of Yeast rRNA by RP-HPLC and Mung Bean Nuclease Assay. *PLoS One, 11*(12), e0168873. doi: 10.1371/journal.pone.0168873

Yang, X., Yang, Y., Sun, B. F., Chen, Y. S., Xu, J. W., Lai, W. Y., . . . Yang, Y. G. (2017). 5-methylcytosine promotes mRNA export - NSUN2 as the methyltransferase and ALYREF as an m5C reader. *Cell Res*. doi: 10.1038/cr.2017.55

Young, B. D., Weiss, D. I., Zurita-Lopez, C. I., Webb, K. J., Clarke, S. G., & McBride, A. E. (2012). Identification of methylated proteins in the yeast small ribosomal subunit: a role for SPOUT methyltransferases in protein arginine methylation. *Biochemistry, 51*(25), 5091-5104. doi: 10.1021/bi300186g

Yu, Y. P., Ding, Y., Chen, Z., Liu, S., Michalopoulos, A., Chen, R., . . . Luo, J. H. (2014). Novel fusion transcripts associate with progressive prostate cancer. *Am J Pathol, 184*(10), 2840-2849. doi: 10.1016/j.ajpath.2014.06.025

Yusupova, G., & Yusupov, M. (2014). High-resolution structure of the eukaryotic 80S ribosome. *Annu Rev Biochem, 83*, 467-486. doi: 10.1146/annurev-biochem-060713-035445

Yutin, N., Makarova, K. S., Mekhedov, S. L., Wolf, Y. I., & Koonin, E. V. (2008). The deep archaeal roots of eukaryotes. *Mol Biol Evol, 25*(8), 1619-1630. doi: 10.1093/molbev/msn108

Zhang, C., Samanta, D., Lu, H., Bullen, J. W., Zhang, H., Chen, I., . . . Semenza, G. L. (2016). Hypoxia induces the breast cancer stem cell phenotype by HIF-dependent and ALKBH5-mediated m(6)A-demethylation of NANOG mRNA. *Proc Natl Acad Sci U S A, 113*(14), E2047-2056. doi: 10.1073/pnas.1602883113

Zhao, X., Yang, Y., Sun, B. F., Shi, Y., Yang, X., Xiao, W., . . . Yang, Y. G. (2014). FTO-dependent demethylation of N6-methyladenosine regulates mRNA splicing and is required for adipogenesis. *Cell Res, 24*(12), 1403-1419. doi: 10.1038/cr.2014.151

Zheng, G., Dahl, J. A., Niu, Y., Fedorcsak, P., Huang, C. M., Li, C. J., . . . He, C. (2013). ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. *Mol Cell, 49*(1), 18-29. doi: 10.1016/j.molcel.2012.10.015

Zorbas, C., Nicolas, E., Wacheul, L., Huvelle, E., Heurgue-Hamard, V., & Lafontaine, D. L. (2015). The human 18S rRNA base methyltransferases DIMT1L and WBSCR22-TRMT112 but not rRNA modification are required for ribosome biogenesis. *Mol Biol Cell, 26*(11), 2080-2095. doi: 10.1091/mbc.E15-02-0073

ANNEX

RÉSUMÉ DE THÈSE

Lors du processus de traduction, le ribosome, une machinerie très sophistiquée et formée des sous-unités 40S et 60S chez les eucaryotes, recrute des ARN de transfert (ARNt) ainsi que des facteurs de traduction pour décoder chaque ARN messager (ARNm) et synthétiser les protéines correspondantes selon les règles du code génétique. La traduction repose sur une notion de haute fidélité dont le but est de réduire au maximum les erreurs lors de la synthèse des protéines. Elle est également finement régulée de façon à produire la quantité adéquate de chaque protéine et de permettre aux organismes et aux cellules de s'adapter aux stimuli environnementaux. Pour ce faire, la grande majorité des facteurs impliqués dans la synthèse des protéines subissent des étapes de maturation qui vont optimiser leur mode d'action. Parmi ces mécanismes de maturation, les plus fréquents sont les modifications post-transcriptionnelles et post-traductionnelles et plus particulièrement les méthylations des divers ARN mais aussi des protéines.

Mon laboratoire d'accueil s'intéresse à la protéine Trm112 qui chez la levure *Saccharomyces cerevisiae*, interagit avec et active 4 méthyltransférases de classe I ou MTases (Mtq2, Trm9, Trm11 et Bud23) qui modifient des acteurs de la traduction. Ces enzymes dépendent du donneur de groupement méthyl S-adenosyl-L-méthionine (SAM). Le complexe Mtq2-Trm112 méthyle le facteur de terminaison eRF1 au niveau de la chaîne latérale du résidu glutamine (Q) du motif GGQ universellement conservé et qui est impliqué directement dans la libération des protéines nouvellement synthétisées. Le complexe Trm11-Trm112 catalyse la formation de 2-méthylguanosine en position 10 des ARNt, une modification dont le rôle serait de stabiliser la structure de ces ARNt. Le complexe Trm9-Trm112 catalyse l'addition d'un groupement méthyl sur l'uridine présente en position 34 de la boucle anticodon de certains ARNt. Cette modification est très vraisemblablement impliquée dans la fidélité du décodage et des expériences suggèrent que cette modification serait impliquée dans la réponse cellulaire aux stress génotoxiques. Le complexe Bud23-Trm112 modifie spécifiquement une base de l'ARNr 18S et intervient dans le mécanisme de biosynthèse de la sous-unité 40S du ribosome. Ces complexes sont également présents dans les cellules humaines et leurs défauts de fonctionnement ont été associés à des maladies neuro-dégénératives et à des cancers.

Lorsque j'ai débuté ma thèse, j'avais deux objectifs:

1. Réaliser l'étude enzymatique du complexe Trm9-Trm112 de la levure *S. cerevisiae*.

Avant mon arrivée au laboratoire, la structure cristallographique du complexe Trm9-Trm112 avait été déterminée et des études *in vivo* et *in vitro* avaient été initiées pour cartographier le site actif de cette enzyme. Pour ma part, j'ai déterminé les constantes cinétiques de mutants alanine de certains résidus très fortement conservés au niveau du site actif de Trm9 pour les comparer à celles de l'enzyme sauvage. Ces études m'ont permis de montrer que les résidus conservés et situés autour du groupement méthyl du SAM sont importants pour la fixation du substrat ARNt (augmentation du Km) mais aussi pour l'activité enzymatique (diminution du k_{cat}). Ces observations sont en accord avec un mécanisme réactionnel de type S_N2, qui est fréquemment décrit pour les MTases de classe I qui sont dépendantes du SAM. En effet, le mécanisme S_N2 consiste en un transfert spontané du groupement méthyl vers son substrat à la condition que ce dernier soit parfaitement positionné par rapport au groupement méthyl du SAM. Ainsi, les résidus mutés apparaissent comme très importants pour orienter idéalement le nucléotide à modifier dans le site actif de l'enzyme pour que le transfert du groupement méthyl puisse avoir lieu. A partir de ces expériences, il a été possible de modéliser le nucléotide substrat dans le site actif de l'holoenzyme Trm9-Trm112. Ceci pourrait avoir des implications futures dans le développement de molécules visant à inhiber l'activité de la protéine ABH8 (ou AlkBH8, orthologue humaine de Trm9) dans la mesure où des études sur des cellules humaines suggèrent que la déplétion de la protéine ABH8 rend ces cellules plus sensibles à des molécules anti-cancéreuses.

2. Etudier le réseau d'interaction de la protéine Trm112 chez les archées.

Des orthologues de la protéine Trm112 de *S. cerevisiae* sont présents dans les trois domaines du vivant (eucaryotes, procaryotes et archées). Au début de ma thèse, les études s'étaient uniquement concentrées sur les complexes formés entre la protéine Trm112 et des MTases chez les eucaryotes. Afin de comprendre d'un point de vue évolutif le ou les rôles des orthologues de Trm112, j'ai caractérisé le réseau d'interaction de la protéine Trm112 de l'archée *Haloferax volcanii*, un organisme modèle d'études pour lequel des outils génétiques ont été développés. De plus, compte-tenu des fortes similarités entre les machineries de traduction

eucaryotes et d'archées, les études menées chez les archées pourraient se révéler informatives pour une meilleure compréhension des mécanismes de la traduction chez les eucaryotes.

J'ai donc généré une souche d'*H. volcanii* exprimant la protéine Trm112 fusionnée à une étiquette "Flag" à son extrémité C-terminale pour réaliser des expériences de co-immunoprécipitation et identifier les partenaires potentiels par spectrométrie de masse. A partir de ces expériences, une liste finale d'environ 500 protéines a été obtenue. De façon très intéressante, nous avons noté un enrichissement significatif de protéines annotées comme MTases (26 au total). Parmi celles-ci, figuraient les orthologues des protéines Trm11, Trm9 et Mtq2 de *S. cerevisiae*. En co-exprimant Trm112 d'*H. volcanii* avec plusieurs MTases différentes chez *E. coli*, j'ai pu montrer que chez cette archée, Trm112 interagit avec au moins 9 MTases différentes dont les orthologues de Trm9 et de Mtq2. Je n'ai malheureusement pas pu valider la formation du complexe Trm11-Trm112 dans la mesure où Trm11 est exprimée uniquement sous forme insoluble. Si pour beaucoup de ces MTases, il est compliqué de prédire une fonction biochimique, j'ai toutefois pu démontrer par des tests enzymatiques que les complexes Mtq2-Trm112 et Trm9-Trm112 d'*H. volcanii* possèdent des activités enzymatiques similaires à leurs orthologues eucaryotes, à savoir la méthylation du facteur de terminaison de la traduction aRF1 et des ARNt, respectivement. J'ai également déterminé la structure cristallographique à 1.35Å de résolution d'un complexe d'*H. volcanii* formé entre Trm112 et une MTase de fonction inconnue. Cette structure montre que les bases moléculaires de l'interaction entre Trm112 et cette MTase (et donc très vraisemblablement des autres MTases validées comme étant des partenaires de Trm112) sont comparables à ce qui a été décrit chez les eucaryotes. Cela permet de rationaliser le fait que Trm112 d'*H. volcanii* puisse interagir avec un si grand nombre de MTases. Des cristaux ont été obtenus pour d'autres complexes et diffractent à 3Å de résolution dans un cas et 7Å de résolution dans un autre mais par manque de temps, je n'ai pas encore pu résoudre leur structure.

Enfin, parmi les MTases partenaires de Trm112 chez *H. volcanii*, une est particulièrement intéressante dans la mesure où elle possède un orthologue proche chez les métazoaires dont l'homme (la protéine METTL5) mais pas chez la levure *S. cerevisiae*. J'ai donc cloné la protéine METTL5 et l'ai co-exprimée avec la protéine humaine TRMT112 chez *E. coli*. Cela m'a permis de mettre en évidence l'interaction directe entre ces deux protéines humaines et donc d'identifier un nouveau partenaire de TRMT112. Celui-ci n'avait pas pu être mis en

évidence jusqu'à présent puisque les études sur les complexes entre TRMT112 humaine et ses MTases partenaires ont toutes été réalisées par analogie à l'état de nos connaissances chez la levure *S. cerevisiae*. A l'heure actuelle, les fonctions biochimiques de METTL5 ne sont pas connues mais des analyses bioinformatiques et bibliographiques suggèrent que cette protéine soit responsable de la formation de m6A (méthylation de la base sur l'azote en position 6 d'une adenine) en position 1832 de l'ARNr 18S. En effet, une co-occurence très claire existe entre la présence de la protéine METTL5 (et de ses orthologues) et de la modification m6A en position 1832 (ou équivalent suivant les organismes) chez les métazoaires et chez les archées. De plus, il est intéressant de noter que cette modification est absente chez la levure *S. cerevisiae* qui ne possède pas d'orthologue de METTL5. Des études seront poursuivies en collaboration avec le Pr. Denis L.J. Lafontaine (Université Libre de Bruxelles, FNRS, Belgique) afin de caractériser la fonction du complexe TRMT112-METTL5.

En conclusion, les travaux menés pendant ma thèse m'ont permis d'approfondir l'état de nos connaissances du réseau d'interaction de la protéine Trm112 qui joue un rôle très important dans le mécanisme global de la traduction des ARNm en protéines comme l'attestent les nombreuses maladies décrites pour être associées à des dysfonctionnements des complexes entre TRMT112 et des MTases humaines. L'étude du réseau d'interaction de Trm112 chez les archées a abouti à plusieurs découvertes surprenantes telles qu'un nombre beaucoup plus important de partenaires chez les archées que chez les eucaryotes mais aussi la découverte d'un nouveau partenaire de TRMT112 humaine. Ceci suggère que le réseau d'interactions de TRMT112 avec des MTases humaines soit plus complexe que précédemment anticipé sur la base des études réalisées initialement chez l'organisme modèle eucaryote *S. cerevisiae*. Cette étude souligne également l'intérêt d'utiliser les archées pour obtenir des informations sur des mécanismes cellulaires essentiels et conservés entre les archées et les eucaryotes.

Titre: Etude de Trm112, un activateur unique de méthyltransférases à l'interface entre la synthèse du ribosome et sa fonction

Mots clés: *méthyltransférases, synthèse des protéines, ribosome, biochimie, cristallographie aux rayons X*

La traduction des ARNm est un processus très complexe qui en plus des nombreux facteurs impliqués, nécessite également des étapes de maturation des protéines et ARN pour la production fidèle des protéines. Parmi ces évènements, des modifications post-transcriptionnelles et post-traductionnelles, dont la méthylation est la plus fréquente, sont trouvées dans tous les composants et principalement chez les eucaryotes. Le rôle des méthylations dans la traduction est parfaitement illustré par la protéine Trm112, qui est un activateur essentiel pour la fonction de 4 méthyltransférases (MTase) (Trm9, Trm11, Bud23 et Mtq2) qui modifient des facteurs impliqués dans la synthèse des protéines. Chez la levure, les complexes Trm9-Trm112 et Trm11-Trm112 catalysent la formation de mcm^5U34 et m^2G10, respectivement sur certains ARNts. Le complexe Bud23-Trm112 modifie l'ARNr 18S pour former la m^7G1575 tandis que le complexe Mtq2-Trm112 modifie le facteur de terminaison de classe I eRF1sur la chaine latérale de la glutamine du motif GGQ. Jusqu'à présent, des études structurales et fonctionnelles du réseau d'interaction de la protéine Trm112 se sont uniquement focalisées chez les eucaryotes alors que cette protéine est trouvée dans les 3 domaines du vivant. Dans cette étude, des expériences de co-immunoprécipitations couplées à de la LC-MS/MS ont permis d'étudier le réseau d'interaction de la protéine Trm112 chez l'archée *H. volcanii*. Celui-ci s'avère être composé de plus de MTase que chez les eucaryotes. Pour la première fois, la structure cristallographique d'un complexe Trm112-MTase d'archée a été déterminée, révélant un mode d'interaction conservé par rapport aux complexes eucaryotes malgré une très faible identité de séquence. De façon très intéressante, un des partenaires de Trm112 chez *H. volcanii* est orthologue d'une protéine humaine dont nous avons pu démontré qu'elle est une nouveau partenaire de la protéine TRMT112 humaine.

Title: Study of Trm112, a unique methyltransferase activator, at the interface between ribosome synthesis and function

Keywords: *methyltransferase, protein synthesis, ribosome, biochemistry, X-ray crystallography*

mRNA translation is a highly complex process, which in addition to the numerous factors involved also requires RNA and protein maturation events for faithful and timely protein productions. Among those events, post-transcriptional and post-translation modifications, with methylation being the most prominent, are found on all components and largely in eukaryotes. The effects of methylation on translation are perfectly illustrated by the Trm112 protein, which is an activating platform, essential for the function of four SAM-dependent methyltransferases (MTase) (Trm9, Trm11, Bud23 and Mtq2) modifying factors involved in protein synthesis. In yeast, Trm9-Trm112 and Trm11-Trm112 complexes catalyze the formation of mcm^5U34 and m^2G10, respectively in some tRNAs. The Bud23-Trm112 complex modifies 18S rRNA to form m^7G1575 while the Mtq2-Trm112 complex methylates class I translation termination factor eRF1 at glutamine side chain of GGQ motif. Until now, the study of Trm112 network in Eukaryotes has been quite clear structurally and functionally, however, little is known for corresponding proteins in Archaea. In this study, using co-immunoprecipitation coupled with LC-MS/MS, we characterized the Trm112 interacting network in the *H. volcanii* archeon, which is surprisingly composed of more MTase partners than that known in Eukaryotes. For the first time, an archaeal Trm112-MTase complex crystal structure has been solved, which shows a conserved interaction mode between Trm112 and MTase partner compared to eukaryotic counterparts despite very less sequence identity shared between those MTases. More importantly, a study on one new archaeal Trm112 partner has helped us extending the TRMT112 network in human by identifying a new MTase Partner.

Université Paris-Saclay
Espace Technologique / Immeuble Discovery
Route de l'Orme aux Merisiers RD 128 / 91190 Saint-Aubin, France