VILNIUS UNIVERSITY
LIFE SCIENCES CENTER

Yaniv Ali

Recombinant Expression and Purification of Human PCIF1, a Cap-Specific Adenosine-N6 Methyltransferase

Master's Thesis

Molecular Biotechnology study program

Thesis supervisor
Dr. Albin Widmark.

Consultant
Dr. Darius Kavaliauskas

Thesis completed at
Thermo Fisher Scientific Baltics, UAB

Vilnius 2025

Contents

Literature review

1.1. mRNA Modifications

1.1.1. mRNA Modifications Overview

The modification of mRNA subsequent to transcription has become a central regulatory dimension of the control of gene expression, influencing the destiny and function of transcripts and the genetic code (Hoernes et al., 2016; Delaunay et al., 2024; Gilbert & Nachtergaele, 2023). These chemical modifications, which together form the epitranscriptome, enable cells to respond dynamically to developmental cues and environmental stresses by imposing reversible changes on mRNA metabolism, frequently functioning as subtle regulators of translational control, RNA stability, and localization (Gilbert et al., 2016; Boo & Kim, 2020; Shi et al., 2020). Unlike DNA chemical modifications, RNA modifications are highly reversible, enabling tight temporal and spatial regulation by dedicated 'writer' proteins that install the modifications, 'reader' proteins that recognize and interpret these modifications to influence downstream processes, and 'eraser' proteins that remove them. (Boo & Kim, 2020; Delaunay et al., 2024; Shi et al., 2020). Notably, these changes are not stochastic but rather occur at defined nucleotide motifs or structural features of the mRNA and thereby affect translation control, splicing, nuclear export, and degradation in transcript-specific manner (Gilbert et al., 2016; Gilbert & Nachtergaele, 2023).

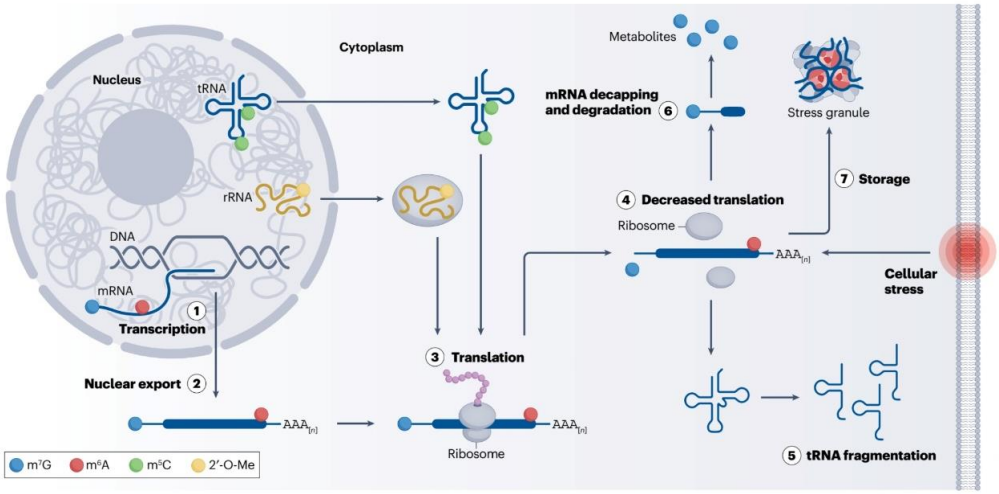

Figure 1.1. Functional roles of mRNA modifications throughout the transcript life cycle.
This figure illustrates how different chemical modifications of mRNA (such as m^6A, m^5C, pseudouridine, etc.) regulate various stages of the mRNA life cycle. Key processes include alternative splicing, mRNA nuclear export, translation enhancement or repression, and transcript degradation. These regulatory effects highlight the dynamic influence of epitranscriptomic marks on gene expression. Adapted from Delaunay et al. (2024).

Although the field of epitranscriptomics is developing very rapidly, it has become clear modifications of mRNA play a central role in diverse biological processes, including but not limited to cell fate determination, stress response. (Hoernes et al., 2016; Shi et al., 2020).

1.1.2 mRNA Modification Types and Dispersion

More than 170 different RNA modifications have been discovered on various RNA species; however, one group comprising N^6-methyladenosine (m^6A), 5-methylcytidine (m^5C), pseudouridine (Ψ), N1-methyladenosine (m^1A), and 2′-O-methylation is particularly prevalent in mRNA and has been well investigated functionally (Gilbert et al., 2016; Delaunay et al., 2024; Shi et al., 2020). Among these changes, m^6A is the most common internal change, occurring in consensus motifs (RRACH) within coding sequences and untranslated regions. The modification plays a significant role in regulating mRNA stability, splicing, translation efficiency, and decay (Boo & Kim, 2020; Gilbert & Nachtergaele, 2023; Shi et al., 2020). m^5C and Ψ, which are less prevalent, although commonly studied in rRNA and tRNA, these modifications in mRNA have also been found to contribute to transcript stabilization their in transcript stabilization and translation regulation under stress in mRNA (Gilbert et al., 2016; Delaunay et al., 2024; Shi et al., 2020). The modifications generally exhibit tissue-specific, cell type–specific, or developmental stage dependent patterns, which testify to their regulatory specificity and adaptive roles in gene expression (Gilbert & Nachtergaele, 2023; Boo & Kim, 2020). Each type of mRNA modification is associated with specific locations within the transcript and is regulated by distinct sets of writer, reader, and eraser proteins.

each type of modification is composed by specific methyltransferases or isomerases (writers) and interpreted by (reader) such as YTH domain proteins, which determine RNA destiny by selective recognition and binding (Boo & Kim, 2020; Shi et al., 2020; Delaunay et al., 2024). Removal is governed by demethylases or editing enzymes (erasers) to allow for reversibility and regulatory plasticity e.g., the FTO and ALKBH5 proteins demethylate m^6A in context-dependent manner (Gilbert & Nachtergaele, 2023; Delaunay et al., 2024). Figure 1.2. illustrates that mRNA modifications tend to occur at specific transcript regions, and these positional patterns significantly influence the regulatory outcomes of post-transcriptional processes. (Gilbert et al., 2016; Shi et al., 2020; Boo & Kim, 2020).

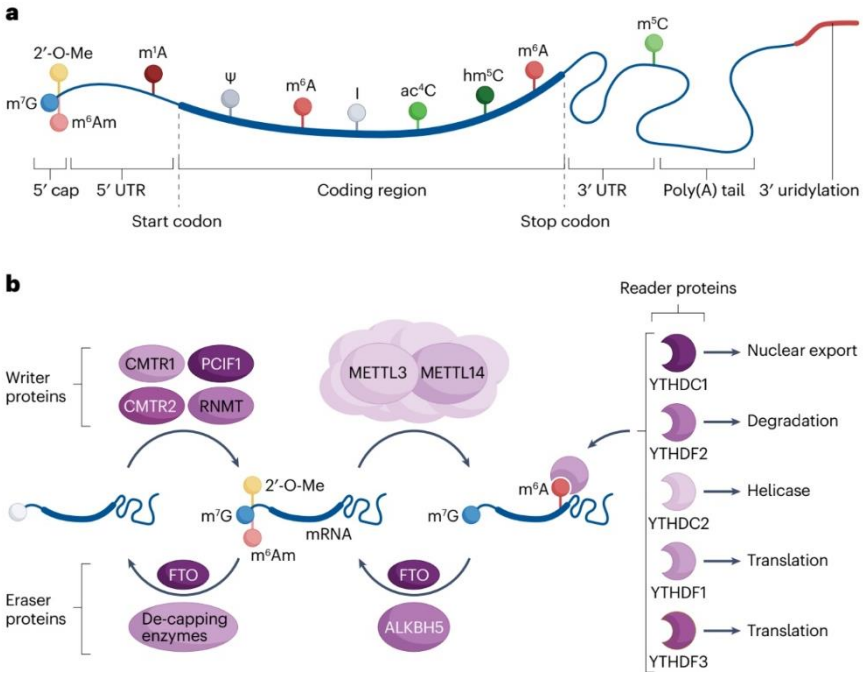

Figure 1.2. Types, locations, and regulatory proteins of mRNA modifications

(A) shows the chemical structures of major mRNA modifications, including m^6A, m^5C, m^1A, and Ψ. Panel B maps where these modifications are typically located within the mRNA molecule (5′ UTR, coding sequence, 3′ UTR). **(B)** list the associated "writer," "reader," and "eraser" proteins for each modification, emphasizing the complexity of the regulatory machinery. Adapted from Delaunay et al. (2024).

1.1.3 N^6,2′-O-Dimethyladenosine m^6Am Modification

Among the diverse array of chemical modifications found in eukaryotic mRNAs, N^6,2′-O-dimethyladenosine (m^6Am) has a singularly defined position, located immediately adjacent to the +1 nucleotide directly adjacent to the 7-methylguanosine (m^7G) cap. The modification plays a role in the generation of an extended cap structure with distinctive biochemical and regulatory features (Jin et al., 2024; Sun et al., 2021; Cesaro et al., 2023). Compared to the internal N^6-methyladenosine (m^6A), which is mostly within coding regions and 3′ untranslated regions, m^6Am is exclusively in the very first transcribed nucleotide of mRNAs (mRNAs) and only if this nucleotide is adenosine (Oerum et al., 2021; Cesaro et al., 2023). Structurally, the modification is a two-step enzymatic process in which, in the first step, 2′-O-methylation of the ribose yields Am, followed by N^6-methylation of the adenine base to yield m^6Am creating a dinucleotide cap structure m^7Gpppm^6Am (Mauer et al., 2017; Cesaro et al., 2023).

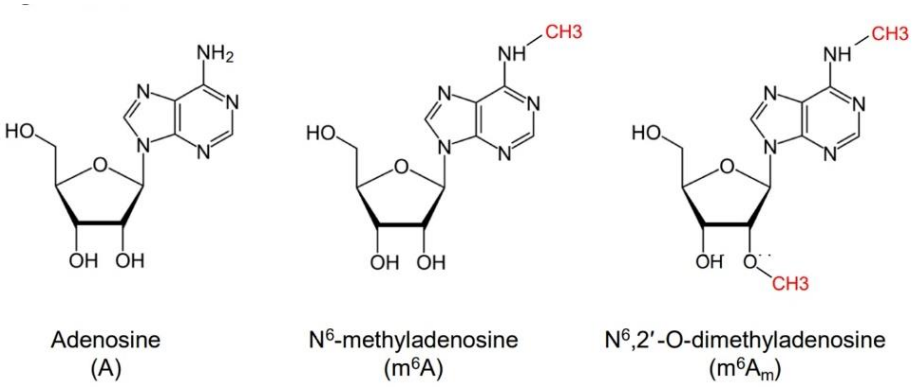

Figure 1.3. Chemical structures of adenosine (A), m⁶A, and m⁶Am

This figure compares the chemical differences between adenosine (A), N⁶-methyladenosine (m⁶A), and N⁶,2′-O-dimethyladenosine (m⁶Am). It highlights the added methyl group at the N⁶ position of the adenine base in both m⁶A and m⁶Am, and the additional 2′-O-methyl group on the ribose in m⁶Am, which distinguishes it structurally and functionally. Adapted from Cesaro et al. (2023).

This methylation at the levels of both ribose and base introduces hydrophobicity and alters base-pairing potential, which has been hypothesized to influence cap-binding protein recruitment and downstream processing of the transcript (Sun et al., 2021; Jin et al., 2024). Recent research demonstrates that the concurrent methylation of m⁶Am generates a distinct biochemical signature that is different from that of m⁶A or Am alone. This specific identity is sensed and regulated by specific molecular mechanisms such as PCIF1, which catalyzes the addition of the N⁶-methyl modification specifically with regard to a pre-existing Am at the cap (Oerum et al., 2021; Sun et al., 2021). The structural specificity of m⁶Am being located exclusively at the first transcribed nucleotide adjacent to the cap—determines not only its placement but also its biological function, including enhanced mRNA stability and longer transcript half-life as illustrated in Figure 1.4. (Mauer et al., 2017; Cesaro et al., 2023; Jin et al., 2024). Besides its function in the structural recognition and stabilization of transcripts, m⁶Am is also implicated in selective regulation of transcripts associated with cellular proliferation, metabolic regulation, and stress response (Sun et al., 2021; Cesaro et al., 2023). Transcriptome-scale studies have found m⁶Am to be particularly enriched in transient, highly expressed mRNAs, suggesting that m⁶Am plays an active role in preventing rapid degradation of transcripts, thereby contributing to their increased stability and expression levels. (Sun et al., 2021; Cesaro et al., 2023). Preferential enrichment and stabilization of rapidly degraded mRNAs point toward m⁶Am serving as a regulatory model for fine-tuning adaptive gene expression programs (Sun et al., 2021; Cesaro et al., 2023).

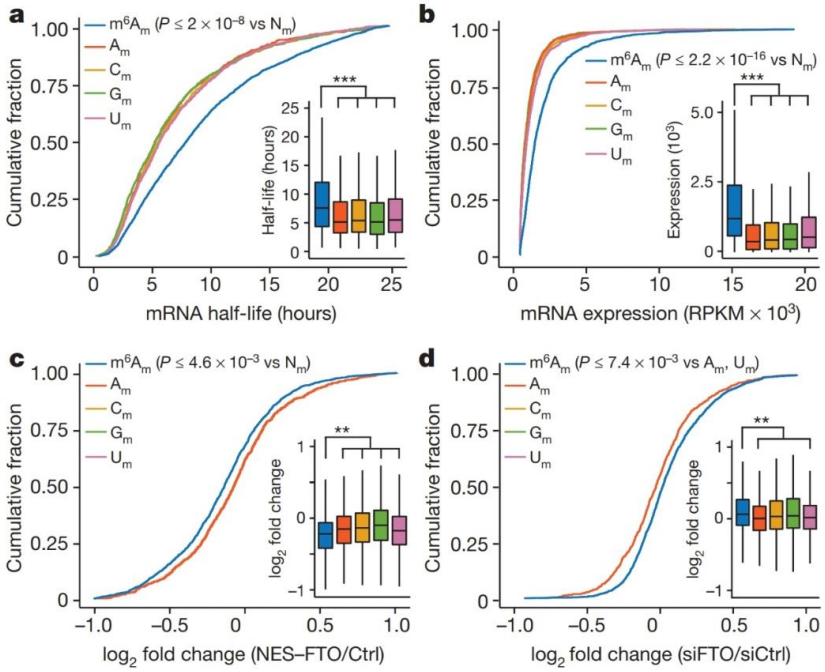

Figure 1.4. m⁶Am Enhances mRNA Stability and Half-Life

(A) Cumulative distribution analysis showing that transcripts beginning with m⁶Am have significantly longer half-lives compared to those starting with Am, Cm, Gm, or Um. **(B)** Box plots of mRNA expression levels further support that m⁶Am-containing transcripts are more abundant, reflecting increased stability. **(C)** Expression of cytosolic NES–FTO leads to selective degradation of m⁶Am-marked RNAs, highlighting its role in dynamic transcript regulation. **(D)** Knock-down of FTO results in elevated levels of m⁶Am transcripts, confirming its stabilizing impact. Adapted from Mauer et al. (2017).

Furthermore, the global transcriptome profiling through m⁶Am-seq has revealed that m⁶Am is exclusively positioned at transcription start sites (TSSs) with an observation of strong enrichment in transcripts implicated in cellular growth and stress response pathways. This observation points to a regulatory role in gene expression based on cellular state and environmental demand (Sun et al., 2021). In contrast to m⁶A, which is distributed throughout coding and non-coding sequences, m⁶Am generates a very localized signal near the cap, thereby establishing itself as an instructional determinant of mRNA destiny during the initial phase of transcript life cycle (Mauer et al., 2017; Sun et al., 2021). The presence of m⁶Am has been tightly linked to enhanced transcript stability and degradation resistance, thereby characterizing it as a stabilizing epitranscriptomic mark that provides a selective advantage by prolonging mRNA half-life and supporting the persistence of transcripts critical for stress response or proliferation. Cesaro et al., 2023; Jin et al., 2024). A particularly interesting example of this role is the capacity to suppress decapping by the exonuclease DCP2, an influence characterizing transcript with m⁶Am from transcripts modified with Am or possessing an unmodified A at the cap (Mauer et al., 2017). This decapping resistance is reflective of a more general regulatory

mechanism whereby m⁶Am extends the mRNA half-life, thereby increasing the stability and translational potential of individual transcripts most conspicuously those engaged in rapid adaptation or cell proliferation (Mauer et al., 2017; Cesaro et al., 2023).

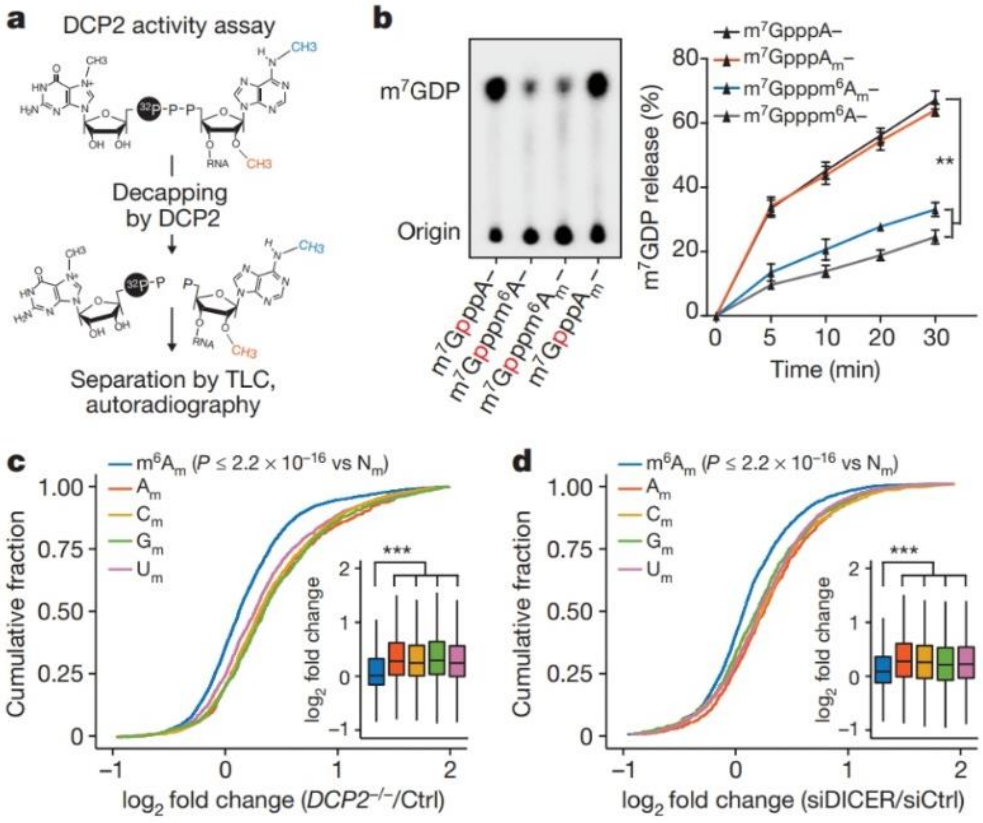

Figure 1.5. m⁶Am Protects mRNA from DCP2-and degradation

(A) Schematic of the decapping assay showing how DCP2 releases m⁷GDP from capped RNA, with m⁶Am-modified RNAs resisting cleavage. **(B)** Experimental TLC assay and kinetic analysis confirm reduced m⁷GDP release from m⁶Am-bearing transcripts. **(C)** Transcriptomic data from DCP2-deficient cells show stabilization of Am-, Cm-, Gm-, and Um-capped RNAs, while m⁶Am-capped RNAs remain largely unchanged, indicating innate resistance. **(D)** m⁶Am-marked RNAs show reduced upregulation following DICER/AGO2 knockdown, suggesting resistance to microRNA-mediated degradation. Adapted from Mauer et al. (2017).

To conclude, m⁶Am enhances transcript stability by protecting it from decapping and degradation. (Cesaro et al., 2023). This function relies on the precision methylation of cap-proximal adenosine by the PCIF1 enzyme (Oerum et al., 2021).

1.2. Cap-Specific Adenosine Methyltransferase

1.2.1. Introduction

Phosphorylated CTD-interacting factor 1 (PCIF1), also known cap-specific adenosine methyltransferase (CAPAM), is the only enzyme found in mammals that catalyzes the N^6-methylation of 2′-O-methyladenosine (Am) adjacent to the 7-methylguanosine (m^7G) cap of mRNA, thus forming the N^6,2′-O-dimethyladenosine (m^6Am) cap modification (Akichika et al., 2019; Sendinc et al., 2019). The human PCIF1 gene is mapped to chromosome 20q13.12 and encodes a protein of 704 amino acids with a predicted molecular weight of around 80.7 kDa (Fan et al., 2003; Akichika et al., 2019). PCIF1 is expressed ubiquitously in all tissues and primarily localized to the nucleus, in line with its function as a transcription-coupled modifier (Fan et al., 2003; Sendinc et al., 2019; Hirose et al., 2008). The enzyme was first characterized by its association with the RNA polymerase II (RNAPII) complex and has since been functionally characterized as a key regulator of m^6Am installation (Fan et al., 2003; Akichika et al., 2019). *PCIF1* is the last enzyme of a sequential capping pathway. The process begins with RNGTT-catalyzed addition of the 5′ guanosine cap, RNMT-catalyzed N^7-methylation of guanosine, and CMTR1-catalyzed 2′-O-methylation of the first transcribed nucleotide (Cowling, 2019; Akichika et al., 2019). *PCIF1* specifically recognizes transcripts where the starting nucleotide is adenosine and methylates the N^6 position, yielding the extended cap structure m^7Gpppm^6Am (Akichika et al., 2019; Sendinc et al., 2019). Notably, the enzyme is recruited to the transcription complex through binding to the Ser5-phosphorylated C-terminal domain (CTD) of RNA polymerase II (RNAPII), thereby enabling co-transcriptional methylation and forming a tight coupling between RNA synthesis and modification (Akichika et al., 2019; Hirose et al., 2008). A step-by-step illustration of this capping process is provided in Figure 1.6.

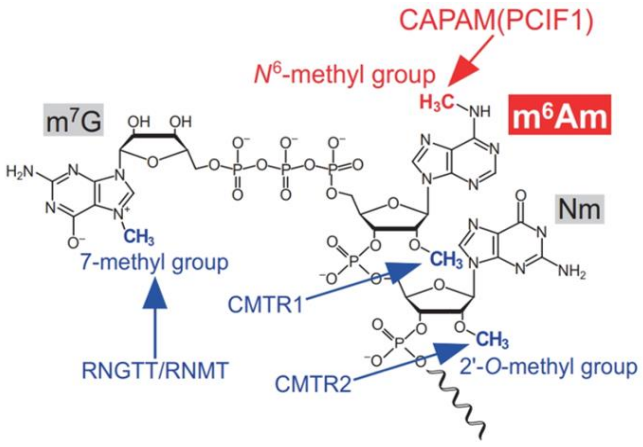

Figure 1.6. Stepwise modification of the 5′ cap structure in vertebrate mRNAs

The figure outlines the sequential conversion of m^7GpppA to m^7Gpppm^6Am through the coordinated actions of RNGTT, RNMT, CMTR1, and *PCIF1* during early transcription.

Adapted from: Akichika et al., 2019.

1.2.2. Structure and Catalytic Domains

PCIF1 contains three structurally distinct domains: an N-terminal WW domain (residues 47–77), a central helical domain, and a C-terminal methyltransferase (MTase) domain (residues 446–616) (Fan et al., 2003; Akichika et al., 2019). The WW domain plays a critical role in binding the phosphorylated carboxy-terminal domain (CTD) of RNA polymerase II (RNAPII) and is therefore necessary for the recruitment of *PCIF1* to the early elongation complex (Fan et al., 2003; Sugita et al., 2021). The middle domain, which is predicted to form a positively charged groove, is involved in RNA substrate binding and domain stability (Akichika et al., 2019). The MTase domain contains a conserved NPPF catalytic motif that is important in the transfer of a methyl group from S-adenosylmethionine (SAM) to the N^6 of adenosine (Akichika et al., 2019; Sendinc et al., 2019). This domain configuration supports *PCIF1*'s dual role in cap recognition and enzymatic methylation during transcription. The structural details and recruitment mechanism are depicted in Figure 1.7.

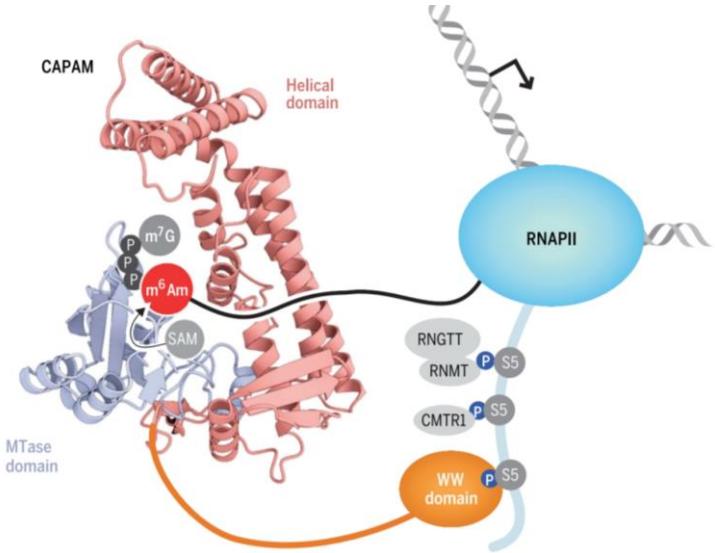

Figure 1.7. Structure and recruitment of *PCIF1*

The figure illustrates the crystal structure of the *PCIF1* catalytic core, including the MTase and helical domains, and its recruitment to RNAPII via WW domain interaction with the Ser5-phosphorylated CTD. Adapted from Akichika et al. (2019) and Wu et al. (2023).

Crystal structures have confirmed *PCIF1* 's high specificity for cap linked Am with minimal activity for unmethylated adenosine, corroborating its role in hierarchical cap modification (Akichika et al., 2019). Enzyme kinetics and mutational studies show that *PCIF1* methylates Am over A preferentially, with significantly higher catalytic efficiency (lower Km for Am substrates), and is inactive on internal adenosines, unequivocally demarcating its activity from the METTL3-METTL14 complex (Sendinc et al., 2019; Wu et al., 2023). Loss-of-function models all show a complete loss of m^6Am with intact

m⁶A levels, attesting to the nonredundant and cap-specific activity of *PCIF1* (Sendinc et al., 2019; Sugita et al., 2021).

1.2.3. Biological Function and Expression Regulation

PCIF1 activity has downstream effects on the transcript fate. Initial studies reported m⁶Am with elevated mRNA stability; however, more recent investigations indicate that *PCIF1* -catalyzed methylation more accurately controls translation efficiency as opposed to transcript longevity (Sendinc et al., 2019; Wu et al., 2023). Numerous transcripts displaying m⁶Am modification exhibit changed ribosome occupancy and reduced translation when *PCIF1* is knocked out, without a corresponding alteration in the overall mRNA quantities (Sendinc et al., 2019). This suggests that m⁶Am is a post-transcriptional control regulatory mechanism that influences protein synthesis without involving the degradation of transcripts. Expression analyses have shown that PCIF1 is not simply nuclear but under dynamic control based on cell type and stress status, with some findings suggesting that *PCIF1* activity is altered in proliferative or metabolically active states (Hirose et al., 2008; Sugita et al., 2021). Enzyme regulation therefore integrates transcriptional cues with post-transcriptional gene expression control. This coordinated recruitment of capping enzymes, including *PCIF1*, during transcription is illustrated in Figure 1.8.

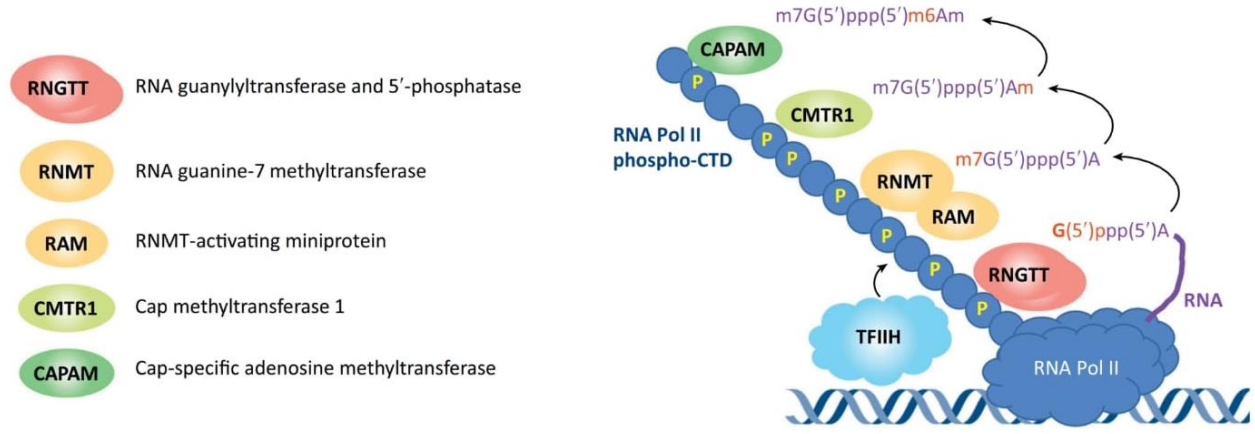

Figure 1.8. Recruitment of mRNA capping enzymes to RNA Polymerase II during transcription.
The figure illustrates the stepwise association of capping enzymes—including RNGTT, RNMT–RAM, CMTR1, and *PCIF1* —with the Ser5-phosphorylated C-terminal domain (CTD) of RNA polymerase II (RNAPII). Each enzyme modifies the growing 5′ cap structure in sequence, culminating in *PCIF1* -mediated N⁶-methylation to form the m⁷Gpppm⁶Am cap. Adapted from Cowling (2019).

Together, these results characterize *PCIF1* as a unique, transcription-coupled methyltransferase with specificity for mRNA cap modification, thereby bridging nuclear RNA processing and translation control. Several studies have expressed recombinant human *PCIF1* in *E. coli*, enabling structural and enzymatic characterization of its methyltransferase domain (Akichika et al., 2019)."
(Akichika et al., 2019).

1.3. Recombinant Protein Expression

1.3.1 Overview

Recombinant DNA technology is a fundamental molecular biology technique that creates new gene combinations by inserting DNA sequences into vectors to be replicated and expressed in host systems (Clark & Pazdernik, 2016). A significant application of this technology is the generation of recombinant proteins, like insulin, through the process of moving target protein-coding genes to suitable host organisms capable of expressing the proteins in functional and large-scale terms (Clark & Pazdernik, 2016; Rosano & Ceccarelli, 2014). Selection of Expression System The choice of host system is very important to protein yield efficiency and is controlled by a number of biological along with operational parameters like expression speed, scalability, cost, post-translational modifications (PTMs) capacity, and genetic tractability (Choi et al., 2006). Of the bacterial hosts, E. coli remains the favorite expression system due to the fact that its genetics are well known, it grows rapidly, has low nutrient demands, and can be fermented to high density (Baneyx, 1999). To better aid decision-making for recombinant protein production, certain performance attributes must be considered in the selection of an expression system. These include scalability, cost, genetic manipulability, capacity for protein folding, and potential for post-translational modification. Figure 1.9. presents a radar chart comparing various expression platforms based on these qualities. This comparative illustration was prepared on the basis of a survey carried out among 27 European academic and industrial organizations within the EP4EU network project. The results provide a helpful guide framework, aiding researchers in matching their choices of expression systems to specific production goals and protein features (Schütz et al., 2023).

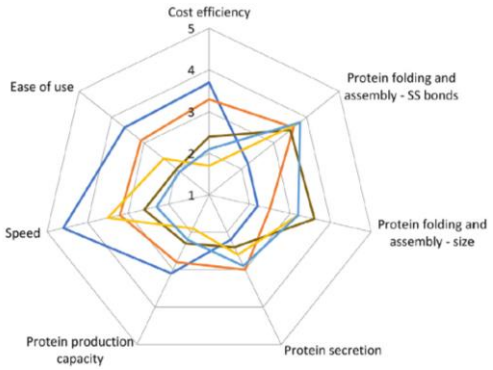

Figure 1.9. Comparative radar chart of expression systems

This radar compares various expression systems based on critical performance criteria, which are scalability, speed of expression, cost effectiveness, genetic manipulability, and capacity for post-translational modification. Higher scores denote superior performance in specific categories. The chart serves to graphically emphasize the comparative strengths and weaknesses of E. coli, yeast, insect, and mammalian systems, thereby facilitating informed decision making in selection.Following Schütz et al. (2023).

1.3.2 Recombinant Protein Expression in Escherichia coli

The process of recombinant protein production in Escherichia coli starts with the cloning of a gene of interest into a plasmid expression vector, which is specifically tailored to enable controlled transcription and translation in bacterial cells (Rosano & Ceccarelli, 2014). In order to be functional, these vectors have several key elements: strong promoters (e.g., T7, araBAD) to enhance transcription, a multiple cloning site (MCS) to insert the gene, an origin of replication (ori) for plasmid replication, selection markers (e.g., antibiotic resistance genes) for plasmid maintenance, and optional affinity tags for purification (Rosano & Ceccarelli, 2014; Hayat et al., 2018). Figure 1.10. provides a schematic representation of these features, emphasizing the structural arrangement of an expression plasmid.

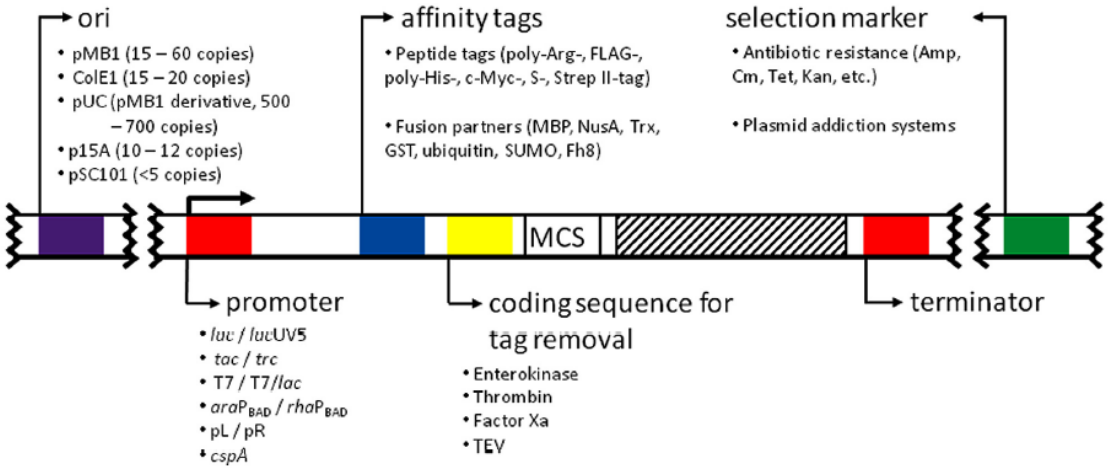

Figure 1.10. Anatomy of an expression vector.

The illustration shows basic features such as promoters, ori, MCS, tag removal sites, and selection markers required for recombinant protein expression in E. coli. Adapted from Rosano & Ceccarelli (2014).

Once the vector construction is done, it is then inserted into competent E. coli cell transformation is a process that facilitates entry and multiplication of the recombinant plasmid into the host bacterium (Rosano & Ceccarelli, 2014). Following the transformation process, expression is triggered by the introduction of chemical inducers specific for the selected promoter system employed (Makrides,

1996). For example, IPTG induces the T7 promoter in the pET system, and arabinose induces the araBAD promoter in the pBAD system, allowing inducible transcription of the inserted gene (Guzman et al., 1995). The ensuing expression is contingent upon the specific pair of the promoter and inducer, which regulates both the level and duration of protein expression. In order to comprehend this regulation process, Figure 1.11. illustrates how the pET system works under the presence or absence of IPTG.

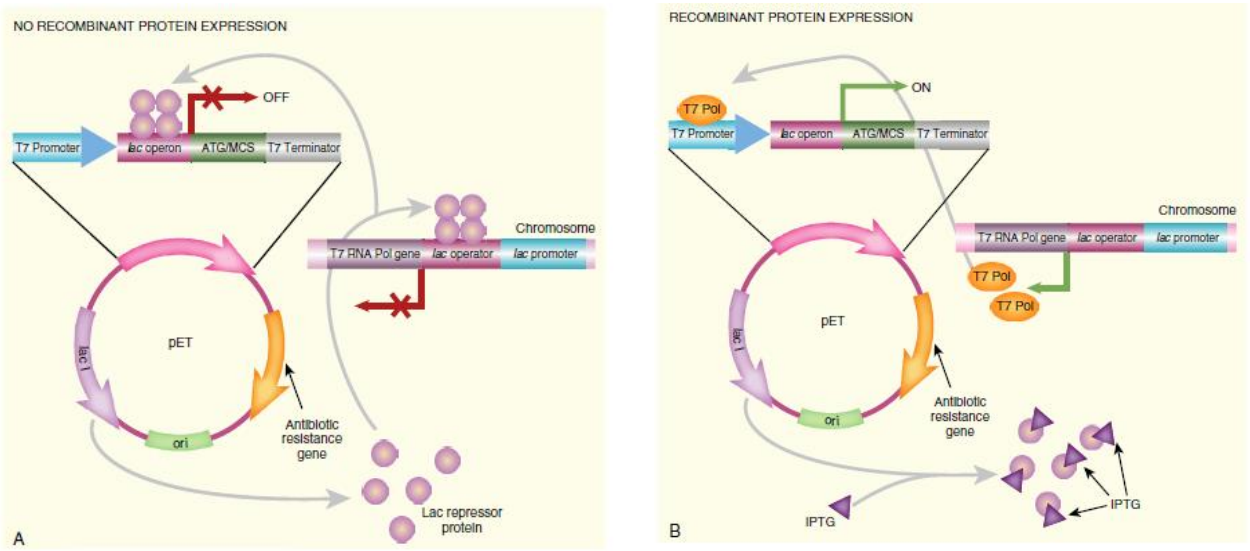

Figure 1.11. Regulation of recombinant protein expression in the pET system

Panel A illustrates the OFF state of the system, where the LacI repressor inhibits transcription by blocking T7 RNA polymerase expression. In **Panel B**, upon induction with IPTG, the repressor is released, enabling the T7 RNA polymerase to initiate transcription of the recombinant gene inserted under the T7 promoter. This visualizes the principle of controlled expression via chemical induction. Adapted from Clark & Pazdernik (2016).

1.3.3 Recombinant Protein Expression in Escherichia coli

In spite of its extensive application, recombinant protein production in Escherichia coli encounters various limitations such as misfolding, inclusion body formation, low solubility, and absence of post-translational modifications (Rosano & Ceccarelli, 2014). These limitations usually result in low quantities of active protein or biologically inactive aggregates. Moreover, E. Escherichia coli is unable to carry out intricate post-translational modifications, including glycosylation and the formation of disulfide bonds, that are critical for the functional activity of most eukaryotic proteins (Jia et al., 2021). Furthermore, toxicity of proteins is a significant limitation, particularly in the production of proteins that disrupt host metabolic processes or affect cell viability (Hayat et al., 2018). Also, hydrophobic characteristics or misfolding of recombinant proteins may cause aggregation and inactivation, whereas proteolytic degradation may result in truncated or inactive products (Singh et

al., 2015). To try to surmount these difficulties, different optimization strategies have been developed, where focus has been placed on critical parameters like temperature, induction time, inducer concentration, and expression strain or vector choice (Mairhofer et al., 2015). Decreasing the culture temperature can increase protein solubility and reduce the formation of inclusion bodies by slowing down the translation rate, thereby allowing more time for proper folding (Rosano & Ceccarelli, 2014). In the same manner, optimization of induction time and IPTG or arabinose concentration prevents overloading of host machinery and reduces toxicity. Selection of vectors containing regulated promoters, fusion partners, or signal peptides has the potential to optimize solubility and direct expression to desired cellular compartments (Hayat et al., 2018). For instance, fusion of the target protein with solubility-promoting tags like GST or MBP frequently results in improved folding and less aggregation. Along with vector and induction tuning, selecting a suitable host strain is another important element. Strains genetically modified to provide rare codon tRNAs or minimize protease activity have been shown to work well for enhancing yield and protein integrity (Singh et al., 2015). In addition, co-expression of molecular chaperones proteins that facilitate proper folding of newly synthesized polypeptides has been shown to be a valuable approach to avert aggregation and maintain proper conformation of complicated proteins. They keep unfolded or partially folded intermediates together during translation to allow the polypeptide to fold into its native structure. (Clark & Pazdernik, 2016).

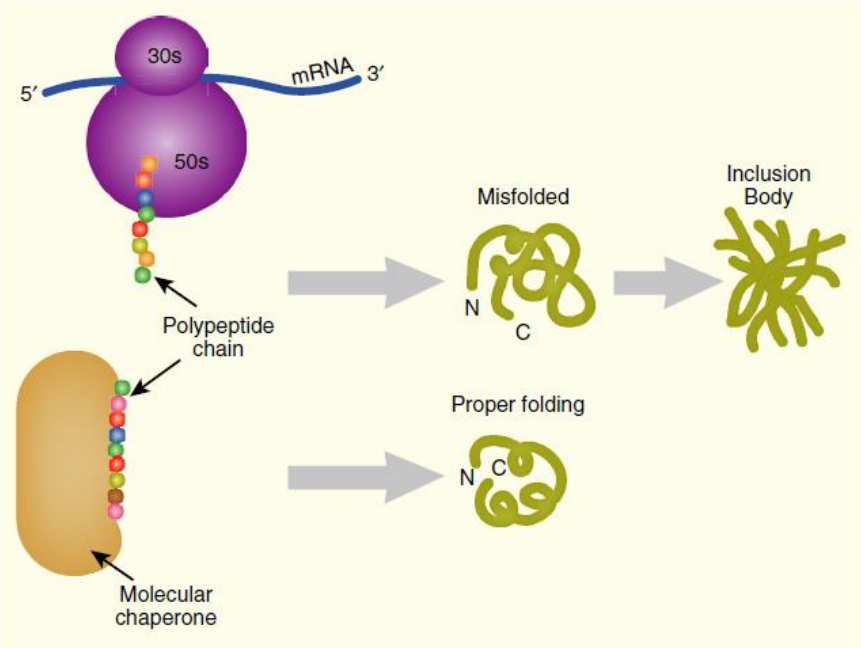

Figure 1.12. Chaperone-assisted protein folding. The illustration shows the destiny of a nascent polypeptide when synthesized on the ribosome. Misfolding in the lack of molecular chaperones results in the formation of inclusion bodies. With chaperone help, the polypeptide gets correct folding and becomes completely functional. Adapted from Clark & Pazdernik (2016).

VILNIUS UNIVERSITY
LIFE SCIENCES CENTER

Yaniv Ali

Master's thesis

Recombinant Expression and Purification of Human PCIF1, a Cap-Specific Adenosine-N6 Methyltransferase.

ABSTRACT

RNA modifications have gained increasing attention due to their critical roles in gene regulation and expression. One such modification, N^6,2'-O-dimethyladenosine (m^6Am), significantly enhances mRNA stability and resistance to degradation. This modification is catalyzed by the cap-specific methyltransferase **PCIF1**. However, the recombinant production of human PCIF1 is challenging because of its structural complexity and tendency to aggregate when expressed in bacterial systems. Therefore, this study aimed to express human PCIF1 recombinantly in *Escherichia coli* and to develop strategies to achieve soluble production and efficient isolation.

The PCIF1 gene was cloned into suitable expression vectors and transformed into *E. coli*. Initial attempts to express His-tagged PCIF1 (without additional fusion proteins) resulted in poor solubility, with the protein accumulating predominantly in inclusion bodies. To improve protein solubility, fusion tags were employed, leading to enhanced folding and significantly increased soluble protein expression; however, the activity of these tagged constructs was not evaluated. Additionally, co-expression with molecular chaperones (GroEL/GroES) notably improved the solubility of His-tagged PCIF1 (no fusion protein), clearly demonstrating their beneficial role in protein folding. Both tagged and His-tagged versions of PCIF1 were successfully isolated by affinity chromatography; however, additional purification steps are necessary to achieve high purity.

In conclusion, this study demonstrates that soluble expression of human PCIF1 in *E. coli* can be effectively achieved either by using fusion tags or through chaperone-assisted folding. These findings offer promising strategies for future structural and functional studies, as well as biotechnological applications involving this important enzyme.

References

1. *Akichika, S., Hirano, S., Shichino, Y., et al. (2019). Cap-specific terminal N6-methylation of RNA by an RNA polymerase II–associated methyltransferase. Science, 363(6423), eaav0080.*
2. *Asahreen, S., Sharif, S., Ahmad, H., & Mukhtar, H. (2023). Comparison of host expression systems used for efficient recombinant protein production. Proceedings of the Pakistan Academy of Sciences: B, 60(1), 33–46.*
3. *Baneyx, F. (1999). Recombinant protein expression in Escherichia coli. Current Opinion in Biotechnology, 10(5), 411–421.*
4. *Boo, S. H., & Kim, Y. K. (2020). The emerging role of RNA modifications in the regulation of mRNA stability. Experimental & Molecular Medicine, 52(3), 400–408.*
5. *Cesaro, E., Zampieri, M., & Fazi, F. (2023). The multifaceted roles of m6Am in mRNA regulation. RNA Biology, 20(1), 203–212.*
6. *Choi, J. H., & Lee, S. Y. (2006). Secretory and extracellular production of recombinant proteins using Escherichia coli. Applied Microbiology and Biotechnology, 64(5), 625–635.*
7. *Clark, D. P., & Pazdernik, N. J. (2016). Biotechnology: Applying the Genetic Revolution (2nd ed.). Academic Cell.*
8. *Cowling, V. H. (2019). CAPAM: the mRNA cap adenosine N6-methyltransferase. Trends in Biochemical Sciences, 44(3), 183–185.*
9. *Delaunay, S., Friedländer, M. R., & Alló, M. (2024). Epitranscriptomic control of gene expression via cap modifications. Trends in Biochemical Sciences, 49(1), 34–48.*
10. *Fan, J., Wang, L., & Zhao, Y. (2003). WW domains and their binding partners in cell signaling and development. Cellular Signalling, 15(2), 179–189.*
11. *Gilbert, W. V., Bell, T. A., & Schaening, C. (2016). Messenger RNA modifications: Form, distribution, and function. Cell, 165(5), 1215–1228.*
12. *Gilbert, W. V., & Nachtergaele, S. (2023). Regulation of translation by epitranscriptomic RNA modifications. Annual Review of Biochemistry, 92, 159–186.*
13. *Guzman, L. M., Belin, D., Carson, M. J., & Beckwith, J. (1995). Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. Journal of Bacteriology, 177(14), 4121–4130.*
14. *Hayat, S. M. G., Golichenari, B., & Sahebkar, A. (2018). Recombinant protein expression in Escherichia coli: Advances and challenges. Current Pharmaceutical Design, 24(40), 4700–4713.*
15. *Hirose, Y., & Ohkuma, Y. (2008). Phosphorylated CTD-interacting factor 1 forms a stable complex with TFIIH and inhibits transcriptional elongation. Journal of Biological Chemistry, 283(44), 29950–29960.*
16. *Hoernes, T. P., Faserl, K., Jantsch, M. F., & Hüttenhofer, A. (2016). Modification of the m6A epitranscriptome: Insights from mass spectrometry. Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms, 1859(6), 316–324.*
17. *Jia, B., Wu, Y., Li, B., et al. (2021). Enhancing disulfide bond formation for improved production of complex proteins in E. coli. Biotechnology Advances, 49, 107751.*
18. *Jin, Y., Liu, H., & Luo, L. (2024). Functional implications of cap-proximal m6Am modification in gene regulation. Frontiers in Genetics, 15, 1200090.*

19. *Makrides, S. C. (1996). Strategies for achieving high-level expression of genes in Escherichia coli. Microbiological Reviews, 60(3), 512–538.*

20. *Manole, C., Gourru, R., Deruelle, M., et al. (2012). Actin- and microtubule-dependent transport of vaccinia virus inside host cells. PLoS Pathogens, 8(2), e1002726.*

21. *Mauer, J., Luo, X., Blanjoie, A., et al. (2017). Reversible methylation of m6Am in the 5′ cap controls mRNA stability. Nature, 541(7637), 371–375.*

22. *Miroux, B., & Walker, J. E. (1996). Over-production of proteins in Escherichia coli: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. Journal of Molecular Biology, 260(3), 289–298.*

23. *Nason, E. L., Rothagel, R., Mukherjee, S. K., et al. (2004). Interactions between the inner and outer capsids of Bluetongue virus. Journal of Virology, 78(15), 8059–8067.*

24. *Rosano, G. L., & Ceccarelli, E. A. (2014). Recombinant protein expression in Escherichia coli: Advances and challenges. Frontiers in Microbiology, 5, 172.*

25. *Sahreen, S., et al. (2023). Host expression systems for recombinant protein production: a comparative overview. Applied Microbiology and Biotechnology, 107(5), 2003–2021.*

26. *Schütz, M., et al. (2023). EP4EU report: Survey of expression systems across EU biotech research centers. European Protein Production Network.*

27. *Sendinc, E., Valle-Garcia, D., Dhall, A., et al. (2019). PCIF1 catalyzes m6Am mRNA methylation to regulate gene expression. Molecular Cell, 75(3), 620–630.*

28. *Shi, H., Wei, J., & He, C. (2020). Where, when, and how: Context-dependent functions of RNA methylation writers, readers, and erasers. Molecular Cell, 74(4), 640–650.*

29. *Singh, A., et al. (2015). Strategies to overcome bottlenecks in E. coli protein production. Biotechnology Journal, 10(1), 56–64.*

30. *Stewart, M., & Roy, P. (2015). Bluetongue virus entry: disruption of endosomal membranes by VP5. Journal of Virology, 89(8), 4319–4331.*

31. *Sugita, M., Tanaka, S., & Watanabe, Y. (2021). Recruitment of CAPAM during transcription. Nature Communications, 12(1), 1657.*

32. *Sun, H., Li, J., He, S., et al. (2021). Cap m6Am methylation regulates transcript stability and proliferation pathways. Nature Cell Biology, 23(6), 620–630.*

33. *Sutton, G., Walton, J., Liebenberg, J., et al. (2007). The crystal structure of Bluetongue virus VP4, a multifunctional capping enzyme. Journal of Virology, 81(12), 6841–6847.*

34. *Wu, R., Liu, Y., Yao, Y., & He, C. (2023). Structural insights into CAPAM activity. Cell Reports, 42(5), 112313.*

35. *Zheng, S., & Shuman, S. (2008). The vaccinia virus D12 subunit is a cofactor for the D1 polyribonucleotide kinase/guanylyltransferase. Journal of Biological Chemistry, 283(14), 8883–8891.*

36. *Oerum, S., Meynier, V., Catala, M., & Tisné, C. (2021). A comprehensive review of m6A/m6Am RNA methyltransferase structures. Nucleic Acids Research, 49(13), 7239–7255. https://doi.org/10.1093/nar/gkab378*