

---

# DNA and RNA Pyrimidine Nucleobase Alkylation at the Carbon-5 Position

Yuri Motorin, Salifu Seidu-Larry, and Mark Helm

---

## Abstract

The carbon 5 of pyrimidine nucleobases is a privileged position in terms of nucleoside modification in both DNA and RNA. The simplest modification of uridine at this position is methylation leading to thymine. Thymine is an integral part of the standard nucleobase repertoire of DNA that is synthesized at the nucleotide level. However, it also occurs in RNA, where it is synthesized posttranscriptionally at the polynucleotide level. The cytidine analogue 5-methylcytidine also occurs in both DNA and RNA, but is introduced at the polynucleotide level in both cases. The same applies to a plethora of additional derivatives found in nature, resulting either from a direct modification of the 5-position by electrophiles or by further derivatization of the 5-methylpyrimidines. Here, we review the structural diversity of these modified bases, the variety of cofactors that serve as carbon donors, and the common principles shared by enzymatic mechanisms generating them.

---

Y. Motorin (✉)

IMoPA UMR7365 CNRS-UL, BioPole de l'Université de Lorraine,  
9 avenue de la Forêt de Haye, 54505 Vandoeuvre-les-Nancy, France  
e-mail: [motorine5@univ-lorraine.fr](mailto:motorine5@univ-lorraine.fr)

S. Seidu-Larry

Department of Biochemistry, University of Cape Coast, College of Agriculture and Natural Sciences, School of Biological Sciences, Cape Coast, Ghana

Institute of Pharmacy and Biochemistry, Johannes Gutenberg University Mainz,  
Staudingerweg 5, 55128 Mainz, Germany

M. Helm (✉)

Institute of Pharmacy and Biochemistry, Johannes Gutenberg University Mainz,  
Staudingerweg 5, 55128 Mainz, Germany  
e-mail: [mhelm@uni-mainz.de](mailto:mhelm@uni-mainz.de)

## 1 Introduction

One of only two chemical differences between RNA and DNA is the presence of a methyl group in deoxythymidine (T, also abbreviated dT, dm<sup>5</sup>U). This substituent at the carbon 5 is thus part of one out of four integral building blocks of DNA, while no methyl group is present in its counterpart ribouridine (U, also abbreviated rU). The fact that both are metabolically derived from uridine monophosphate (containing ribose) is one of several arguments often used to support the claim that DNA has evolved from RNA (Kun et al. 2015; Muller 2006). It is also a clear indication of the importance of this methyl group, which, from this perspective, constitutes a nucleoside modification. Remarkably, and in contrast to most other nucleoside modifications, the obligatory presence of that methyl group in DNA is cemented by its introduction prior to nucleotide polymerization. Interestingly, thymidine also occurs in RNA (ribothymidine, also rT, m<sup>5</sup>U, rm<sup>5</sup>U), most prominently as the namesake of the T(T $\Psi$ )-loop in transfer RNA, where it is introduced posttranscriptionally. Methylated versions of the sister nucleobase cytidine are also found in both DNA (5mC) and RNA (m<sup>5</sup>C); however, all are introduced at the polynucleotide level.

Beyond simple methyl groups, numerous chemically more complex modifications of pyrimidines are known that contain a carbon modification at the 5-position. This group may easily represent the largest group of all known modifications of nucleobases, because it includes a large part of the modifications of the U34-position in the anticodon of transfer RNAs (tRNAs) with its several dozen species (Machnicka et al. 2013). Again, the vast majority of these modifications are introduced at the polynucleotide level, the only exceptions being 5-hydroxymethylpyrimidines (5hmC, 5hmU) and their glycosylated derivatives found in phage DNA, which are generated at the mononucleotide level (Gommers-Ampt and Borst 1995). Interestingly, the resulting triphosphate nucleotides are then incorporated by phage DNA polymerase despite being sterically encumbered to a large degree. Even more interestingly, this tolerance for modifications at the 5-position appears to be a general feature of nucleotide polymerases on both the DNA and RNA level. Most prominently, T7-RNA polymerase, arguably the most commonly used enzyme for RNA synthesis, efficiently incorporates triphosphate ribonucleotides of rT and rm<sup>5</sup>C, and even more sterically, demanding carbon-5 modifications have been incorporated into RNA this way (Vaught et al. 2004). Similarly, synthetic modifications are available for incorporation into DNA in PCR reactions by non-phage polymerases (Vaught et al. 2010).

The enzymatic mechanisms of pyrimidine C-5 modification involved here are of particular interest, since the central step involves a C–C bond formation. This reaction type is of increased interest to organic chemists and in natural product metabolism. In a large number of cases, proper understanding of the mechanism of bond-forming reactions requires the identification of the nucleophilic partner on one hand and the electrophilic partner on the other hand. This is typically easy for the formation of C–N or C–O bonds but more sophisticated for C–C bonds. In the case at hand, the carbon-5 position in pyrimidines is catalytically activated to form an intermittent carbon nucleophile, while the carbon side chains result from an

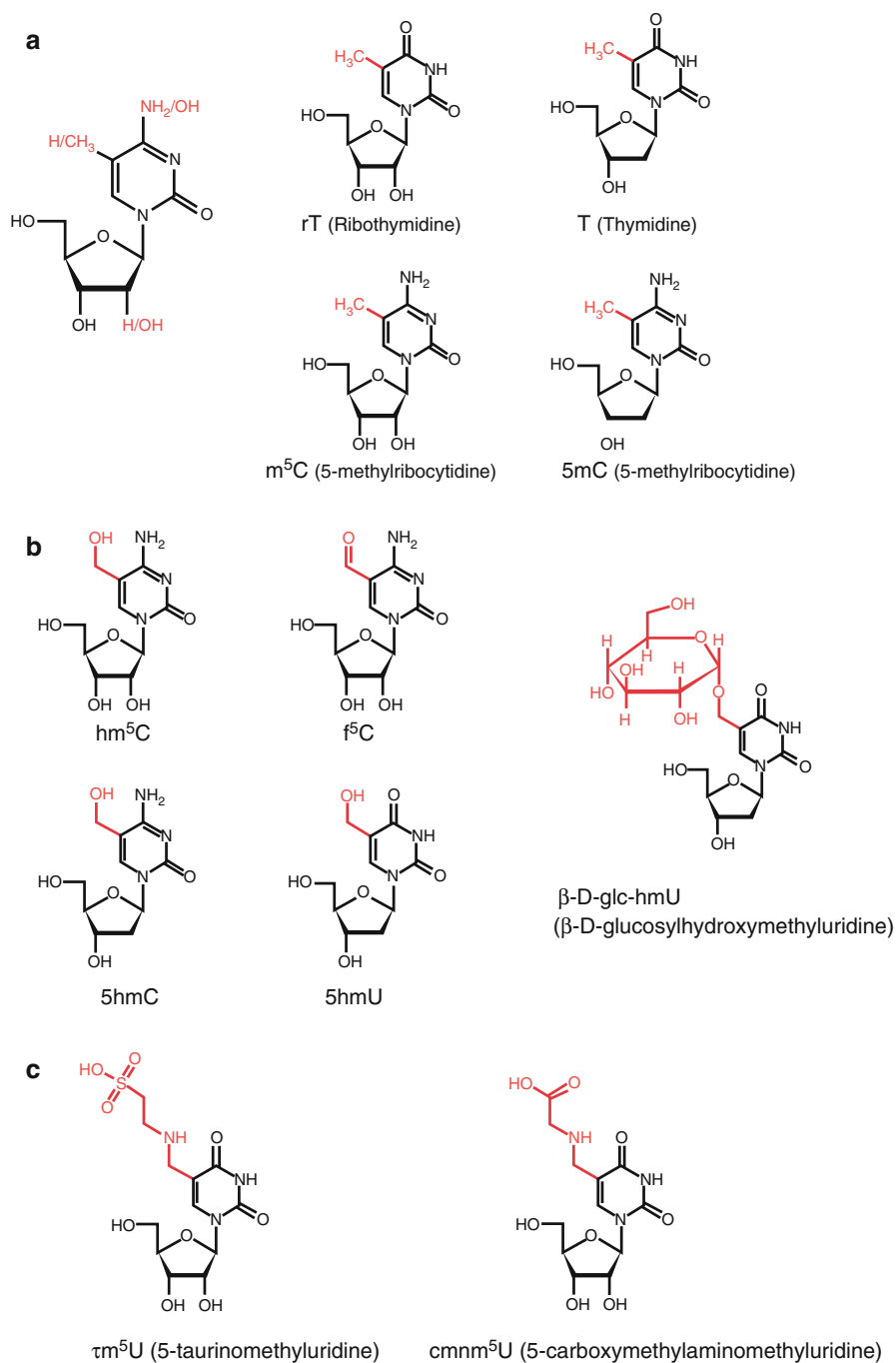
electrophilic metabolite, such as S-adenosyl-L-methionine (AdoMet) or  $N^5,N^{10}$ -methylenetetrahydrofolate ( $\text{CH}_2$ -THF). Hence, before addressing the biocatalysis of pyrimidine alkylation, we will discuss the reactivity of the 5-position from the perspective of the organic chemist and then do the same with the various electrophilic carbon scaffolds supplied as cofactors by the modification enzymes. Only then will we discuss a series of enzymatic reactions, of 5-pyrimidine alkylation and related relevant processes. Of note, the mechanism of the methyl group oxidations by TET enzymes leading, e.g., to 5hmC and 5hmU (Fu et al. 2014; Pfaffeneder et al. 2014) will not be discussed here. Instead, we will turn to equally fascinating modifications typically found in transfer RNA (tRNA) featuring a uridine at 34-position of the anticodon loop. Although this group of modifications includes a bewildering variety of sophisticated chemical structures, the initial modification reaction bears similarities with the aforementioned relatively simple modifications.

## 1.1 Chemical Structure and Occurrence of Pyrimidine C-5 Modifications

A surprising variety of pyrimidine modifications at the 5-position are known as of today and have been known for some time. Permutation of functional groups at three positions on the pyrimidine nucleoside, namely, H vs. OH at the 2'-position,  $\text{NH}_2$  vs. OH at the 4-position, and H vs.  $\text{CH}_3$  at the 5-position, results in eight species of pyrimidine nucleosides. Ribothymidine (rT, Fig. 1a) is ubiquitous in tRNA and very frequent in ribosomal RNA (rRNA), but not known elsewhere (Motorin and Helm 2011). 5mC is present in bacterial DNA as a guard against restriction nucleases (Roberts et al. 2005), while its presence in promoter and coding regions of eukaryotic genes participates in the regulation of transcription (Bogdanovic and Gomez-Skarmeta 2014). The distribution of  $\text{m}^5\text{C}$  in RNA, which is yet more complicated, was recently reviewed by us (Motorin et al. 2010) and has since encountered renewed interest through transcriptome-wide studies (Burgess et al. 2015; Militello et al. 2014; Squires et al. 2012).

The number of chemical species dramatically increases if 5-modifications other than methyl groups are admitted into this perusal. For example, recent research has discovered, or rediscovered, oxidation products of the 5-methyl group, including 5-hydroxymethyl and 5-formyl derivatives (Fig. 1b) (Kriaucionis and Heintz 2009; Tahiliani et al. 2009; Pfaffeneder et al. 2011). While it is common knowledge that DNA obligatorily contains T as a 5-methylated pyrimidine nucleobase, a less well-known exception is that the abovementioned glycosylated derivatives of 5hmC and 5hmU (termed “J-base”) are not just spurious modifications in the DNA of certain phages, but exist as near quantitative surrogates of the conventional C and T nucleosides (Fig. 1b) (Gommers-Ampt and Borst 1995). The unglycosylated precursor 5hmC was discovered in phage DNA as early as 1953 (Wyatt and Cohen 1953).

In eukaryotes, the existence of 5-hydroxymethylcytidine in DNA is a more recent discovery (Kriaucionis and Heintz 2009; Tahiliani et al. 2009) with a strong impact in fields such as epigenetics and developmental biology, while the corresponding



**Fig. 1** DNA and RNA modifications arising from alkylation of the pyrimidine 5-position. (a) Thymidine, ribothymidine, m<sup>5</sup>C, and 5mC. (b) Modified nucleosides resulting from 5-mPy oxidation. (c) Modified uridines involved in codon recognition at 34-position of the tRNA anticodon

modification in RNA has also been reported decades ago, although incompletely characterized (Racz et al. 1978). Similarly, continued investigations have revealed 5-formylcytosine in DNA (fC, 5fC) (Pfaffeneder et al. 2011), while the corresponding ribonucleotide (f<sup>r</sup>C) had been described in tRNA as early as 1994 (Moriya et al. 1994). However, further oxidation of 5fC leads to 5-carboxydeoxycytidine (caC, 5caC) (He et al. 2011), of which the ribonucleoside has yet to be discovered. 5hmU as a constituent of mammalian DNA has been discovered in traces and demonstrated to be a consequence of thymidine oxidation by TET enzymes (Pfaffeneder et al. 2014). Finally, the largest structural variety is found in aminomethyluridines, which are ribothymidine derivatives at the oxidation step of 5-hydroxymethyluridine and which predominate at 34-position in the anticodon of tRNAs (Machnicka et al. 2013). In contrast to 5hmU, these are not biochemically formed by oxidation of thymidine, but their biogenesis involves the use of a single carbon building block at the oxidation state of formaldehyde, namely, CH<sub>2</sub>-THF, as will be detailed below.

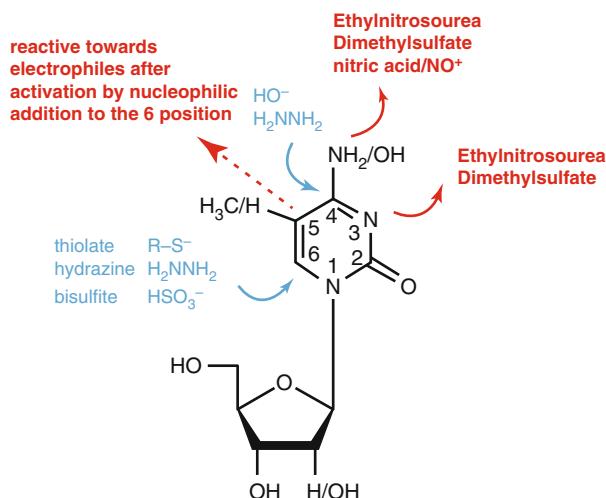
## 1.2 Reactivity of Pyrimidines

A closer look at the catalytic strategies employed by modification enzymes acting on the 5-position of pyrimidines reveals that these exploit the intrinsic chemical reactivity of the pyrimidine ring. While this is not a surprising finding in general, the situation of pyrimidines is counterintuitive to the untrained biochemist, and a brief look at pyrimidine reactivity is conducive to a more intuitive mechanistic understanding of the involved enzyme.

Both nitrogen atoms within pyrimidines exert an electron withdrawing effect, resulting in an electron poor aromatic ring that is susceptible to nucleophiles. A nucleophilic attack, e.g., by bisulphite, at 6-position can be viewed as a Michael addition, while 4-position corresponds directly to the electrophilic center of a carbonyl functionality. Certain reactions with nucleophiles, such as hydrazine treatment, or the deamination reaction used in the so-called bisulphite sequencing (Schaefer et al. 2009; Frommer et al. 1992), sequentially exploit the electrophilic nature of both positions (Fig. 2).

In contrast, their electron poverty leaves pyrimidines relatively inert toward electrophilic reagents such as alkylating reagents, with the 3- and 5-positions being the exceptions (Motorin et al. 2010). The N3-position reacts with electrophilic reagents such as kethoxal or CMC, which is exploited in structural probing experiments (Giege et al. 1999). Concerning carbon 5, uridine reacts with the electrophilic formaldehyde under relatively mild acidic conditions to form 5-hydroxymethyluridine. Uracil (Kong et al. 2009), deoxyuridine (Conte et al. 1992), as well as cytidines (Khursid et al. 1982) are reported to also yield 5-hydroxymethylpyrimidines under alkaline conditions. The mechanism under acidic conditions can be understood in analogy to a Friedel–Crafts alkylation/acylation, which involves a stabilization by the lone electron pair of nitrogen 1 of the positive charge introduced by the alkylating agent.

Under alkaline conditions, an intermittent Michael addition of hydroxide at the 6-position would plausibly generate an enolate-type carbon nucleophile, which



**Fig. 2** Chemical reactivity of the pyrimidine C-5 position. Nucleophilic reagents and attack vectors are depicted in blue, electrophilic reagents and vectors are in red

then reacts with the electrophilic formaldehyde, followed by elimination of the hydroxide to restore the aromatic ring. Note that indeed, the enzymatic mechanisms discussed below for alkylation, acylation, or hydroxymethylation all involve such a Michael attack by a nucleophile, typically a cysteine thiolate (Jurkowski et al. 2008; Motorin et al. 2010). Interestingly, mechanisms discussed for the enzymatic decarboxylation of 5cC and 5fC employ the same path in reverse. Carell et al. described a nonenzymatic *in vitro* decarboxylation proceeding in the presence of high concentrations of thiol but at low pH (Schiesser et al. 2013; Schiesser et al. 2012). Under the same conditions, removal of formaldehyde from 5hmC was inefficient. Reaction with formaldehyde may be conducted in the presence of amine, resulting in aminomethylation, thus leading to modified pyrimidines that closely resemble native counterparts typically found at 34-position of tRNA. Here again, the catalytic mechanism in the biosynthesis of these modified bases bears similarities (Helm and Alfonzo 2014; El Yacoubi et al. 2012) with that of other modifications using CH<sub>2</sub>-THF (*vide infra*).

## 2 Enzymatic Mechanisms of Pyrimidine Alkylation

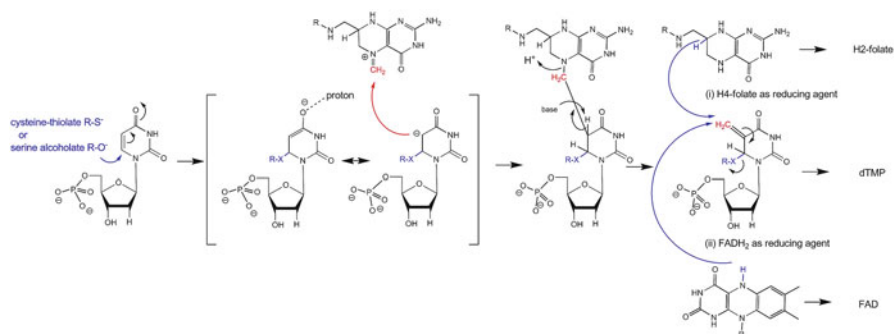
Attachment of the alkyl (most frequently -CH<sub>3</sub>) group to carbon 5 of pyrimidines U and C can be catalyzed by a variety of enzymes which differ in their origin, sequence, and structure yet employ some common principles of catalysis. At the nucleotide level, this reaction is catalyzed by the extensively studied thymidylate synthase (TS), which is a key target enzyme in certain anticancer and immunosuppressive treatments. TS catalyzes the conversion of dUMP into dTMP, an essential reaction

for the synthesis of DNA nucleotides. At the polynucleotide level, the methylation of C5 in U and C is insured either by specific DNA-MTases (for 5mC formation) or by RNA-specific m<sup>5</sup>U-methyltransferases as well as m<sup>5</sup>C-methyltransferases.

## 2.1 The Thymidylate Synthase Family

Thymidylate synthase (TS, EC 2.1.1.45) catalyzes the synthesis of dTMP via a reductive methylation of dUMP. This important enzyme family has been extensively studied for almost 40 years, starting in the late 1970s (Santi 1986; Carreras and Santi 1995). The first characterized enzymes used CH<sub>2</sub>-THF as co-substrate, yielding dihydrofolate, from which CH<sub>2</sub>-THF was regenerated from serine and FADH<sub>2</sub>. More recent studies (Koehn and Kohen 2010; Graziani et al. 2006; Agrawal et al. 2004) revealed the existence of a second unusual class of TS, which also act on dUMP and use CH<sub>2</sub>-THF, but require FADH<sub>2</sub> as a direct reaction cofactor. This family is now called FDTs for flavin-dependent TS. The catalytic mechanism is now established for both enzyme families (Hong et al. 2007; Koehn et al. 2009; Mishanina et al. 2012, 2014).

In the first “classical” family of TS enzymes, the initial step of catalysis relies on a highly conserved Cys residue, which is positioned in the active site of the enzyme. This residue is responsible for the activation of the C5 via addition to the C5=C6 double bond in the pyrimidine ring, resulting in an enolate intermediate. The enolate’s nucleophilic C5 attacks the methylene CH<sub>2</sub> group of the folate co-substrate, forming a covalent ternary complex between the enzyme, dUMP, and the folate. The next step of this reaction is a hydride transfer, which allows the formation of the methylated pyrimidine ring, and is followed by the release of the enzyme via a concerted reaction mechanism corresponding to an elimination that reconstitutes the 5–6 carbon double bond (Islam et al. 2014) (Fig. 3). In the Flavin-dependent TS family, the initial step of the reaction may depend on the enzyme nucleophile (generally an OH-group) or on a direct attachment of FADH<sub>2</sub> at the 6-position of the pyrimidine base. In the case of an enzyme nucleophile, the major reaction steps are



**Fig. 3** Enzymatic mechanisms for 2'-deoxythymidine formation in DNA building blocks

rather similar to the “classical” TS, except the last step of hydride transfer, where FADH<sub>2</sub> serves as a hydride donor rather than THF. In the case of the direct activation of dUTP by FADH<sub>2</sub>, after hydride transfer, the FADH<sub>2</sub> is replaced by CH<sub>2</sub>-THF, and the reaction proceeds by the “classical” way, but without covalent intermediate at the TS active site. The final hydride transfer thus proceeds by an intermolecular rather than an intramolecular reaction (not shown), and THF is the cofactor product as opposed to dihydrofolate in the case of the classical enzymes.

## 2.2 Enzymes Performing 5-Pyrimidine Methylation in Nucleic Acids

The rT (m<sup>5</sup>U) was among the first modified nucleotides discovered in tRNAs, and the respective bacterial enzyme (TrmA or RUMT) catalyzing its biosynthesis was characterized in *E. coli* in the early 1980s (Greenberg and Dudock 1980; Ny and Bjork 1980; Lindstrom et al. 1985). Studies of its enzymatic mechanism identified AdoMet as its CH<sub>3</sub>-group donor and Cys324 as a catalytic nucleophile and suggested a simple displacement mechanism of the methylation reaction (Kealey and Santi 1991; Kealey et al. 1991). RUMT catalyzes the modification of U54 in tRNAs, and in addition it is capable of modifying synthetic 16S rRNA *in vitro* (Gu et al. 1994). The yeast homologue of RUMT was also characterized and its tRNA recognition properties studied using synthetic tRNA transcripts (Nordlund et al. 2000; Becker et al. 1997). rT was also found in bacterial rRNA and a different MTase (ygcA, renamed to RumA/RlmD) was found to be responsible for its formation. Mutagenesis of RUMT and structural studies of RumA identified the residues involved in catalysis (Santi and Hardy 1987; Kealey et al. 1994). Bacteria also have an additional enzyme of the same family (RlmC/RumB), catalyzing m<sup>5</sup>U747 formation in 23S rRNA. Activity of m<sup>5</sup>U-MTases was also detected in Archaea (Constantinesco et al. 1999); however their presence is restricted to the Thermococcales and Nanoarchaeota groups. In *Pyrococcus abyssi*, two close homologues of RlmD fulfill the cellular functions of TrmA (m<sup>5</sup>U54 in tRNA) and RlmC (equivalent of m<sup>5</sup>U747 in 23S rRNA) (Auxilien et al. 2011). The analysis of m<sup>5</sup>U54 formation in *B. subtilis* revealed an unexpected m<sup>5</sup>U54-MTase in these gram-positive bacteria. The flavoprotein TrmFO enzyme from *B. subtilis* uses CH<sub>2</sub>-THF as a carbon donor, akin to ThyA and ThyX thymidylate synthases (Urbonavicius et al. 2005; Hamdane et al. 2012, 2013). In addition, TrmFO uses the same flavin FADH<sub>2</sub> cofactor as the TSFD family, as the reducing agent in the CH<sub>3</sub>-group transfer. A similar enzyme was found to catalyze the formation of m<sup>5</sup>U1939 in *M. capricolum* 23S rRNA (Lartigues et al. 2014).

The distribution of m<sup>5</sup>C in cellular RNAs from different life domains is complex. In bacteria, this modified residue is present in rRNA, but not in other RNA species; in eukaryotes it is found in tRNA, rRNA, and mRNA (Squires et al. 2012; Hussain et al. 2013), while in Archaea its presence seems to be restricted to tRNAs and some sites in mRNAs (Edelheit et al. 2013). Three m<sup>5</sup>C residues in *E. coli* rRNA are formed by three specific enzymes, while in yeast three homologues modify both tRNAs and rRNAs. In higher eukaryotes, at least seven or eight specific proteins are



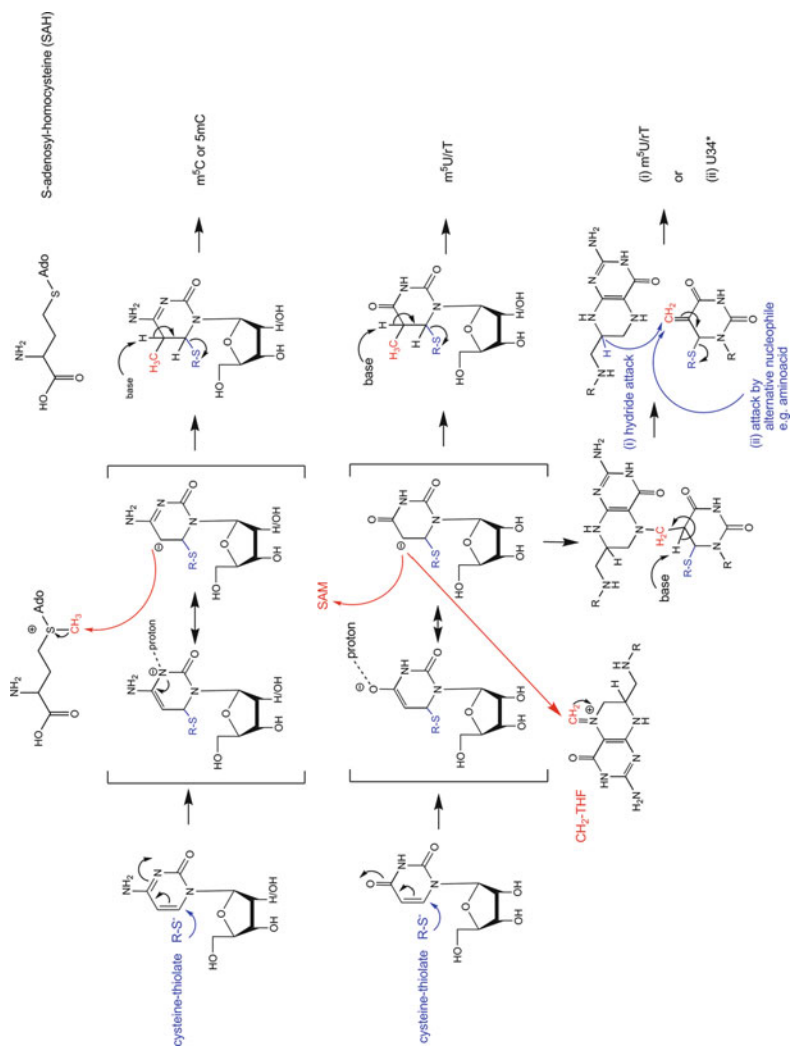
required for the modification of tRNA and cytoplasmic and mitochondrial rRNAs and mRNAs (see Motorin et al. (2010) for further information).

The known enzymes transferring methyl groups from AdoMet to nucleic acids belong to the SPOUT and MTase superfamilies, the latter containing a Rossmann fold for the accommodation of the cofactor. Structure–function relationships in the m<sup>5</sup>C-MTase family were combined with bioinformatic analyses (Bujnicki et al. 2004), resulting in a subdivision of the known m<sup>5</sup>C-MTases into four major sub-families: two groups related to Nop2/Nol1 and YebU/Trm4, a large group related to RsmB or Ynl022c, and a small group represented by *P. horikoshi* PH1991 and human NSUN6. Further inspection of homologues in higher eukaryotes (Pavlopoulou and Kossida 2009) suggested the existence of a new subgroup of m<sup>5</sup>C-MTase-related proteins, termed RCMT9, with members distantly related to Trm4 and a distribution restricted to four taxons. A detailed discussion of the distribution of m<sup>5</sup>C-forming enzymes in the different kingdoms is given in Motorin et al. (2010).

Higher eukaryotes also have another distinct family of m<sup>5</sup>C-RNA-MTases derived from former m<sup>5</sup>C:DNA-MTases (DNMT2-related family). These enzymes have different catalytic mechanisms but evolved to modify tRNAs at 38-position (Goll et al. 2006). For information on m<sup>5</sup>C-MTases acting on DNA, i.e., enzymes of the DNMT family, see elsewhere in this book.

### 2.3 Catalytic Mechanisms in the Formation of rT, m<sup>5</sup>C, and 5mC

The catalytic strategies employed for the alkylation of the carbon 5 in pyrimidines share some common elements, which derive from the heterocycle reactivity, as outlined above. Some basic elements already appeared in the discussion of the thymidylate synthase in above. In all cases, the Michael addition of an anionic nucleophile to the 6-position of the pyrimidine ring produces a nucleophilic carbon with partial carbanion character at 5-position (Fig. 4). In uridines, the Michael addition produces an intermediate, in which the negative charge is delocalized in an enolate structure. Arguably, this intermediate might be stabilized by a hydrogen bond of the enolate oxygen before reacting as a nucleophile with the carbon electrophile provided as a cofactor in the form of AdoMet or CH<sub>2</sub>-THF. In cytidine substrates, the mechanisms comprise an enamine intermediate instead of an enolate, and the mechanisms discussed in literature typically include an acidic residue in the catalytic site, which may intermittently protonate nitrogen 3 to stabilize this enamine intermediate. So far, the known enzymes acting on cytidines in both RNA and DNA exclusively use AdoMet as an electrophilic carbon source, while rT can be formed from either AdoMet or CH<sub>2</sub>-THF. This has interesting implications, namely, (i) that a cytidine methyltransferase using CH<sub>2</sub>-THF might so far have eluded detection and (ii) that the formation of rT has been invented multiple times with different cofactors in the course of evolution, in particular at 54-position of tRNA (Hamdane et al. 2012; Nordlund et al. 2000; Ny and Bjork 1980). Furthermore, with an eye to the more sophisticated U34 modifications occurring in tRNA, we note that a covalent



**Fig. 4** Posttranscriptional ribouridine and cytidine methylation with the AdoMet cofactor. The first step uses a catalytic cysteine in the enzyme active site for activation of the C5=C6 bond via formation of a covalent enzyme-RNA enolate intermediate. This allows the transfer of the CH<sub>3</sub>-group from AdoMet to the 5-position of the base, followed by deprotonation via Glu358 and the enzyme release from the covalent enzyme-RNA complex (Reviewed in Kealey et al. 1994)

enzyme-thiol-pyrimidine-methylene-folate intermediate (as reported/postulated for TS (Fig. 3) and TrmFO) can not only be resolved by a reduction with a hydride equivalent from the folate (Fig. 4(i)) but also with other nucleophiles (Helm and Alfonzo 2014) (Fig. 4 (ii)). The aromaticity of the pyrimidine base is restored in an elimination step featuring an abstraction of a proton from the C5, which regenerates the thiolate used in the initial activation step via Michael addition. In a number of RNA m<sup>5</sup>C:MTases, a second cysteine was reported to be crucial to this regeneration (reviewed in Motorin et al. (2010)).

In the framework of the above-described common elements, the various enzymes differ from one another by the amino acids that embody the different roles outlined above, such as activating nucleophile, general acid, general base, etc. In addition, the position of such residues, while relatively conserved in the spatial arrangement of the active site, may vary within the polypeptide sequence of the different enzymes. This has been especially well studied in m<sup>5</sup>C:MTases, where these residues are located within several conserved motives numbered I through X, which, in their order of appearance in the primary sequence, undergo permutation among the different enzymes of the bacterial and plant DNA-MTase family and some of the RNA-m<sup>5</sup>C:MTase (reviewed in Motorin et al. (2010)).

## 2.4 Catalytic Mechanisms in the Formation of Exotic U34 Modifications

As already mentioned several times, the most bewildering variety of 5-pyrimidine modifications are found at uridine 34 in the anticodon of tRNAs, where they play a crucial role in mRNA decoding. Alkylations predominate among these modifications, and some of the enzymatic mechanisms bear strong similarity with those applied in simple methylations. For example, certain enzymes use CH<sub>2</sub>-THF to transfer a formaldehyde equivalent to the C5, and instead of reducing it to the methyl group using a hydride donor, such as tetrahydrofolate (H4-folate) or FADH<sub>2</sub> (Fig. 4), the amino group of certain amino acids such as taurine or glycine serves as the attacking nucleophile, leading to the structures displayed in Fig. 1c, which correspond to the overall product of a Mannich reaction (Helm and Alfonzo 2014). Recently, two groups discovered new types of reactive intermediates formed and employed in catalytic mechanisms of U34 modification. The Almo group reported the conversion of the conventional AdoMet into a novel derivative, carboxy-S-adenosyl-L-methionine, which is used by the bacterial CmoB enzyme to introduce the carboxymethyl into 5-hydroxyuridine (ho5U), yielding a 5-oxyacetyluridine (cmo<sup>5</sup>U) (Kim et al. 2013, 2015). Most interestingly, Huang's group reported the use of AdoMet by a radical AdoMet enzyme from the elongator complex to generate a radical from the methyl group of the acetic acid moiety in acetyl CoA, which would then add to the C5–C6 double bond of uridine34 in Archaea and Eukarya (Selvadurai et al. 2014).

### 3 Functions of Alkylated Pyrimidine Nucleosides

In view of the plethora of different structures of alkylated pyrimidines already discovered, it is clear that there cannot be one function common to all of them. Indeed, new facets of functions are being continuously discovered in very diverse areas of molecular life sciences, and since this is not the focus of this chapter, we will only provide references to few known functions. One common biophysical property of 5-methylpyrimidines is that they enhance stacking in A- and B-helices of nucleic acids, leading to a structural stabilization that is typically reflected in an increased thermal stability detected in melting experiments. This applies to thymidine and 5-methylcytidine in DNA and RNA alike. The role of 5-methylcytidine in mammalian epigenetics, as well as in the restriction/methylation systems of bacteria, is detailed elsewhere in this book. Of interest, certain bacteriophages use particular 5-pyrimidine modifications to escape bacterial restriction (Gommers-Ampt and Borst 1995). Curiously, the roles of ribothymidine and 5-methylribocytidine in RNA have remained little understood despite their long-standing tenure in the zoo of known RNA modifications (Motorin and Helm 2011). This is likely due to the fact that their principle occurrences in tRNA and rRNA concern heavily modified RNAs, where a plethora of modifications cooperate in a network fashion to modulate RNA activity (Motorin and Helm 2010). The role of U34 modifications in tRNA has already been alluded to, although the generic explanation of mRNA decoding on the ribosome does not do justice to the plethora of structures found here. Apparently, there is no universally perfect modification at this site that suits all organisms, and the variety of conditions under which protein synthesis must take place has led to the emergence of numerous chemical solutions in different species. Along this line, the recent findings that tRNA anticodon modifications are dynamically responding to stress conditions point to an especially sensitive environment that is subject to constant tuning and further evolution.

---

### References

- Agrawal N, Lesley SA, Kuhn P, Kohen A. Mechanistic studies of a flavin-dependent thymidylate synthase. *Biochemistry*. 2004;43(32):10295–301. doi:[10.1021/bi0490439](https://doi.org/10.1021/bi0490439).
- Auxilien S, Rasmussen A, Rose S, Brochier-Armanet C, Husson C, Fourmy D, et al. Specificity shifts in the rRNA and tRNA nucleotide targets of archaeal and bacterial m5U methyltransferases. *RNA*. 2011;17(1):45–53. doi:[10.1261/rna.2323411](https://doi.org/10.1261/rna.2323411).
- Becker HF, Motorin Y, Sissler M, Florentz C, Grosjean H. Major identity determinants for enzymatic formation of ribothymidine and pseudouridine in the T psi-loop of yeast tRNAs. *J Mol Biol*. 1997;274(4):505–18. doi:[10.1006/jmbi.1997.1417](https://doi.org/10.1006/jmbi.1997.1417).
- Bogdanovic O, Gomez-Skarmeta JL. Embryonic DNA methylation: insights from the genomics era. *Brief Funct Genomics*. 2014;13(2):121–30. doi:[10.1093/bfpg/elt039](https://doi.org/10.1093/bfpg/elt039).
- Bujnicki JM, Feder M, Ayres CL, Redman KL. Sequence-structure-function studies of tRNA:m5C methyltransferase Trm4p and its relationship to DNA:m5C and RNA:m5U methyltransferases. *Nucleic Acids Res*. 2004;32(8):2453–63. doi:[10.1093/nar/gkh564](https://doi.org/10.1093/nar/gkh564).
- Burgess AL, David R, Searle IR. Conservation of tRNA and rRNA 5-methylcytosine in the kingdom Plantae. *BMC Plant Biol*. 2015;15:199. doi:[10.1186/s12870-015-0580-8](https://doi.org/10.1186/s12870-015-0580-8).

- Carreras CW, Santi DV. The catalytic mechanism and structure of thymidylate synthase. *Annu Rev Biochem.* 1995;64:721–62. doi:[10.1146/annurev.bi.64.070195.003445](https://doi.org/10.1146/annurev.bi.64.070195.003445).
- Constantinesco F, Motorin Y, Grosjean H. Transfer RNA modification enzymes from *Pyrococcus furiosus*: detection of the enzymatic activities in vitro. *Nucleic Acids Res.* 1999;27(5):1308–15.
- Conte MR, Galeone A, Avizonis D, Hsu VL, Mayol L, Kearns DR. Solid phase synthesis of 5-hydroxymethyluracil containing DNA. *Bioorg Med Chem Lett.* 1992;2:79–81.
- Edelheit S, Schwartz S, Mumbach MR, Wurtzel O, Sorek R. Transcriptome-wide mapping of 5-methylcytidine RNA modifications in bacteria, archaea, and yeast reveals m5C within archaeal mRNAs. *PLoS Genet.* 2013;9(6):e1003602. doi:[10.1371/journal.pgen.1003602](https://doi.org/10.1371/journal.pgen.1003602).
- El Yacoubi B, Bailly M, de Crecy-Lagard V. Biosynthesis and function of posttranscriptional modifications of transfer RNAs. *Annu Rev Genet.* 2012;46:69–95. doi:[10.1146/annurev-genet-110711-155641](https://doi.org/10.1146/annurev-genet-110711-155641).
- Frommer M, McDonald LE, Millar DS, Collis CM, Watt F, Grigg GW, et al. A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. *Proc Natl Acad Sci U S A.* 1992;89(5):1827–31.
- Fu L, Guerrero CR, Zhong N, Amato NJ, Liu Y, Liu S, et al. Tet-mediated formation of 5-hydroxymethylcytosine in RNA. *J Am Chem Soc.* 2014;136(33):11582–5. doi:[10.1021/ja505305z](https://doi.org/10.1021/ja505305z).
- Giege R, Helm M, Florentz C. Chemical and enzymatic Probing of RNA Structure. *Comprehensive Natural Product Chemistry.* In: Söll D, Nishimura S (editors). Oxford: Pergamon Press; 1999;6:63–77.
- Goll MG, Kirpekar F, Maggert KA, Yoder JA, Hsieh CL, Zhang X, et al. Methylation of tRNA<sup>Asp</sup> by the DNA methyltransferase homolog Dnmt2. *Science.* 2006;311(5759):395–8. doi:[10.1126/science.1120976](https://doi.org/10.1126/science.1120976).
- Gommers-Ampt JH, Borst P. Hypermodified bases in DNA. *FASEB J Off Publ Fed Am Soc Exp Biol.* 1995;9(11):1034–42.
- Graziani S, Bernauer J, Skouloubris S, Graille M, Zhou CZ, Marchand C, et al. Catalytic mechanism and structure of viral flavin-dependent thymidylate synthase ThyX. *J Biol Chem.* 2006;281(33):24048–57. doi:[10.1074/jbc.M600745200](https://doi.org/10.1074/jbc.M600745200).
- Greenberg R, Dudock B. Isolation and characterization of m5U-methyltransferase from *Escherichia coli*. *J Biol Chem.* 1980;255(17):8296–302.
- Gu X, Ofengand J, Santi DV. In vitro methylation of *Escherichia coli* 16S rRNA by tRNA(m5U54)-methyltransferase. *Biochemistry.* 1994;33(8):2255–61.
- Hamdane D, Argentini M, Cornu D, Golinelli-Pimpaneau B, Fontecave M. FAD/folate-dependent tRNA methyltransferase: flavin as a new methyl-transfer agent. *J Am Chem Soc.* 2012;134(48):19739–45. doi:[10.1021/ja308145p](https://doi.org/10.1021/ja308145p).
- Hamdane D, Bruch E, Un S, Field M, Fontecave M. Activation of a unique flavin-dependent tRNA-methylating agent. *Biochemistry.* 2013;52(49):8949–56. doi:[10.1021/bi4013879](https://doi.org/10.1021/bi4013879).
- He YF, Li BZ, Li Z, Liu P, Wang Y, Tang Q, et al. Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science.* 2011;333(6047):1303–7. doi:[10.1126/science.1210944](https://doi.org/10.1126/science.1210944).
- Helm M, Alfonzo JD. Posttranscriptional RNA Modifications: playing metabolic games in a cell's chemical Legoland. *Chem Biol.* 2014;21(2):174–85. doi:[10.1016/j.chembiol.2013.10.015](https://doi.org/10.1016/j.chembiol.2013.10.015).
- Hong B, Maley F, Kohen A. Role of Y94 in proton and hydride transfers catalyzed by thymidylate synthase. *Biochemistry.* 2007;46(49):14188–97. doi:[10.1021/bi701363s](https://doi.org/10.1021/bi701363s).
- Hussain S, Aleksic J, Blanco S, Dietmann S, Frye M. Characterizing 5-methylcytosine in the mammalian epitranscriptome. *Genome Biol.* 2013;14(11):215. doi:[10.1186/gb4143](https://doi.org/10.1186/gb4143).
- Islam Z, Strutzenberg TS, Gurevic I, Kohen A. Concerted versus stepwise mechanism in thymidylate synthase. *J Am Chem Soc.* 2014;136(28):9850–3. doi:[10.1021/ja504341g](https://doi.org/10.1021/ja504341g).
- Jurkowski TP, Meusbürger M, Phalke S, Helm M, Nellen W, Reuter G, et al. Human DNMT2 methylates tRNA(Asp) molecules using a DNA methyltransferase-like catalytic mechanism. *RNA.* 2008;14(8):1663–70. doi:[10.1261/ma.970408](https://doi.org/10.1261/ma.970408).

- Kealey JT, Santi DV. Identification of the catalytic nucleophile of tRNA (m5U54)methyltransferase. *Biochemistry*. 1991;30(40):9724–8.
- Kealey JT, Lee S, Floss HG, Santi DV. Stereochemistry of methyl transfer catalyzed by tRNA (m5U54)-methyltransferase--evidence for a single displacement mechanism. *Nucleic Acids Res*. 1991;19(23):6465–8.
- Kealey JT, Gu X, Santi DV. Enzymatic mechanism of tRNA (m5U54)methyltransferase. *Biochimie*. 1994;76(12):1133–42.
- Khursid M, Khan A, Qudrat-E-Khuda M. Synthesis of 5-Hydroxymethylcytidine. *J Chem Soc Pak*. 1982;4(3):167–8.
- Kim J, Xiao H, Bonanno JB, Kalyanaraman C, Brown S, Tang X, et al. Structure-guided discovery of the metabolite carboxy-SAM that modulates tRNA function. *Nature*. 2013;498(7452):123–6. doi:[10.1038/nature12180](https://doi.org/10.1038/nature12180).
- Kim J, Xiao H, Koh J, Wang Y, Bonanno JB, Thomas K, et al. Determinants of the CmoB carboxy-methyl transferase utilized for selective tRNA wobble modification. *Nucleic Acids Res*. 2015;43(9):4602–13. doi:[10.1093/nar/gkv206](https://doi.org/10.1093/nar/gkv206).
- Koehn EM, Kohen A. Flavin-dependent thymidylate synthase: a novel pathway towards thymine. *Arch Biochem Biophys*. 2010;493(1):96–102. doi:[10.1016/j.abb.2009.07.016](https://doi.org/10.1016/j.abb.2009.07.016).
- Koehn EM, Fleischmann T, Conrad JA, Palfey BA, Lesley SA, Mathews II, et al. An unusual mechanism of thymidylate biosynthesis in organisms containing the thyX gene. *Nature*. 2009;458(7240):919–23. doi:[10.1038/nature07973](https://doi.org/10.1038/nature07973).
- Kong LQ, Zhao J, Fan L, Yang DC. The facile synthetic method of 5-hydroxymethyluracil. *Chin Chem Lett*. 2009;20(3):314–6.
- Kriaucionis S, Heintz N. The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. *Science*. 2009;324(5929):929–30. doi:[10.1126/science.1169786](https://doi.org/10.1126/science.1169786).
- Kun A, Szilagy A, Konnyu B, Boza G, Zachar I, Szathmary E. The dynamics of the RNA world: insights and challenges. *Ann NY Acad Sci*. 2015;1341:75–95. doi:[10.1111/nyas.12700](https://doi.org/10.1111/nyas.12700).
- Lartigue C, Lebaudy A, Blanchard A, El Yacoubi B, Rose S, Grosjean H, et al. The flavoprotein Mcap0476 (RlmFO) catalyzes m5U1939 modification in *Mycoplasma capricolum* 23S rRNA. *Nucleic Acids Res*. 2014;42(12):8073–82. doi:[10.1093/nar/gku518](https://doi.org/10.1093/nar/gku518).
- Lindstrom PH, Stuber D, Bjork GR. Genetic organization and transcription from the gene (trmA) responsible for synthesis of tRNA (uracil-5)-methyltransferase by *Escherichia coli*. *J Bacteriol*. 1985;164(3):1117–23.
- Machnicka MA, Milanowska K, Osman Oglou O, Purta E, Kurkowska M, Olchowik A, et al. MODOMICS: a database of RNA modification pathways--2013 update. *Nucleic Acids Res*. 2013;41(Database issue):D262–7. doi:[10.1093/nar/gks1007](https://doi.org/10.1093/nar/gks1007).
- Militello KT, Chen LM, Ackerman SE, Mandarano AH, Valentine EL. A map of 5-methylcytosine residues in *Trypanosoma brucei* tRNA revealed by sodium bisulfite sequencing. *Mol Biochem Parasitol*. 2014;193(2):122–6. doi:[10.1016/j.molbiopara.2013.12.003](https://doi.org/10.1016/j.molbiopara.2013.12.003).
- Mishanina TV, Koehn EM, Conrad JA, Palfey BA, Lesley SA, Kohen A. Trapping of an intermediate in the reaction catalyzed by flavin-dependent thymidylate synthase. *J Am Chem Soc*. 2012;134(9):4442–8. doi:[10.1021/ja212082z](https://doi.org/10.1021/ja212082z).
- Mishanina TV, Corcoran JM, Kohen A. Substrate activation in flavin-dependent thymidylate synthase. *J Am Chem Soc*. 2014;136(30):10597–600. doi:[10.1021/ja506108b](https://doi.org/10.1021/ja506108b).
- Moriya J, Yokogawa T, Wakita K, Ueda T, Nishikawa K, Crain PF, et al. A novel modified nucleoside found at the first position of the anticodon of methionine tRNA from bovine liver mitochondria. *Biochemistry*. 1994;33(8):2234–9.
- Motorin Y, Helm M. tRNA stabilization by modified nucleotides. *Biochemistry*. 2010;49(24):4934–44. doi:[10.1021/bi100408z](https://doi.org/10.1021/bi100408z).
- Motorin Y, Helm M. RNA nucleotide methylation. *Wiley Interdiscip Rev RNA*. 2011;2(5):611–31. doi:[10.1002/wrna.79](https://doi.org/10.1002/wrna.79).
- Motorin Y, Lyko F, Helm M. 5-methylcytosine in RNA: detection, enzymatic formation and biological functions. *Nucleic Acids Res*. 2010;38(5):1415–30. doi:[10.1093/nar/gkp1117](https://doi.org/10.1093/nar/gkp1117).
- Muller UF. Re-creating an RNA world. *Cell Mol Life Sci*. 2006;63(11):1278–93. doi:[10.1007/s00018-006-6047-1](https://doi.org/10.1007/s00018-006-6047-1).

- Nordlund ME, Johansson JO, von Pawel-Rammingen U, Bystrom AS. Identification of the TRM2 gene encoding the tRNA(m5U54)methyltransferase of *Saccharomyces cerevisiae*. *RNA*. 2000;6(6):844–60.
- Ny T, Bjork GR. Cloning and restriction mapping of the *trmA* gene coding for transfer ribonucleic acid (5-methyluridine)-methyltransferase in *Escherichia coli* K-12. *J Bacteriol*. 1980;142(2):371–9.
- Pavlopoulou A, Kossida S. Phylogenetic analysis of the eukaryotic RNA (cytosine-5)-methyltransferases. *Genomics*. 2009;93(4):350–7. doi:[10.1016/j.ygeno.2008.12.004](https://doi.org/10.1016/j.ygeno.2008.12.004).
- Pfaffeneder T, Hackner B, Truss M, Munzel M, Muller M, Deiml CA, et al. The discovery of 5-formylcytosine in embryonic stem cell DNA. *Angewandte Chemie*. 2011;50(31):7008–12. doi:[10.1002/anie.201103899](https://doi.org/10.1002/anie.201103899).
- Pfaffeneder T, Spada F, Wagner M, Brandmayr C, Laube SK, Eisen D, et al. Tet oxidizes thymine to 5-hydroxymethyluracil in mouse embryonic stem cell DNA. *Nat Chem Biol*. 2014;10(7):574–81. doi:[10.1038/nchembio.1532](https://doi.org/10.1038/nchembio.1532).
- Racz I, Kiraly I, Lasztily D. Effect of light on the nucleotide composition of rRNA of wheat seedlings. *Planta*. 1978;142(3):263–7. doi:[10.1007/BF00385075](https://doi.org/10.1007/BF00385075).
- Roberts RJ, Vincze T, Posfai J, Macelis D. REBASE--restriction enzymes and DNA methyltransferases. *Nucleic Acids Res*. 2005;33(Database issue):D230–2. doi:[10.1093/nar/gki029](https://doi.org/10.1093/nar/gki029).
- Santi DV. The mechanism and structure of thymidylate synthetase. *Nucleic Acids Symp Ser*. 1986;17:125–6.
- Santi DV, Hardy LW. Catalytic mechanism and inhibition of tRNA (uracil-5-)methyltransferase: evidence for covalent catalysis. *Biochemistry*. 1987;26(26):8599–606.
- Schaefer M, Pollex T, Hanna K, Lyko F. RNA cytosine methylation analysis by bisulfite sequencing. *Nucleic Acids Res*. 2009;37(2):e12. doi:[10.1093/nar/gkn954](https://doi.org/10.1093/nar/gkn954).
- Schiesser S, Hackner B, Pfaffeneder T, Muller M, Hagemeyer C, Truss M, et al. Mechanism and stem-cell activity of 5-carboxycytosine decarboxylation determined by isotope tracing. *Angewandte Chemie*. 2012;51(26):6516–20. doi:[10.1002/anie.201202583](https://doi.org/10.1002/anie.201202583).
- Schiesser S, Pfaffeneder T, Sadeghian K, Hackner B, Steigenberger B, Schroder AS, et al. Deamination, oxidation, and C-C bond cleavage reactivity of 5-hydroxymethylcytosine, 5-formylcytosine, and 5-carboxycytosine. *J Am Chem Soc*. 2013;135(39):14593–9. doi:[10.1021/ja403229y](https://doi.org/10.1021/ja403229y).
- Selvadurai K, Wang P, Seimetz J, Huang RH. Archaeal Etp3 catalyzes tRNA wobble uridine modification at C5 via a radical mechanism. *Nat Chem Biol*. 2014;10(10):810–2. doi:[10.1038/nchembio.1610](https://doi.org/10.1038/nchembio.1610).
- Squires JE, Patel HR, Nusch M, Sibbritt T, Humphreys DT, Parker BJ, et al. Widespread occurrence of 5-methylcytosine in human coding and non-coding RNA. *Nucleic Acids Res*. 2012;40(11):5023–33. doi:[10.1093/nar/gks144](https://doi.org/10.1093/nar/gks144).
- Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, et al. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science*. 2009;324(5929):930–5. doi:[10.1126/science.1170116](https://doi.org/10.1126/science.1170116).
- Urbonavicius J, Skouloubris S, Myllykallio H, Grosjean H. Identification of a novel gene encoding a flavin-dependent tRNA:m5U methyltransferase in bacteria--evolutionary implications. *Nucleic Acids Res*. 2005;33(13):3955–64. doi:[10.1093/nar/gki703](https://doi.org/10.1093/nar/gki703).
- Vaught JD, Dewey T, Eaton BE. T7 RNA polymerase transcription with 5-position modified UTP derivatives. *J Am Chem Soc*. 2004;126(36):11231–7. doi:[10.1021/ja049009h](https://doi.org/10.1021/ja049009h).
- Vaught JD, Bock C, Carter J, Fitzwater T, Otis M, Schneider D, et al. Expanding the chemistry of DNA for in vitro selection. *J Am Chem Soc*. 2010;132(12):4141–51. doi:[10.1021/ja908035g](https://doi.org/10.1021/ja908035g).
- Wyatt GR, Cohen SS. The bases of the nucleic acids of some bacterial and animal viruses: the occurrence of 5-hydroxymethylcytosine. *Biochem J*. 1953;55(5):774–82.