



HAL
open science

Dihydrouridine in the transcriptome: New life for this ancient RNA chemical modification

Damien Brégeon, Ludovic Pecqueur, Sabrine Toubdji, Claudia Sudol, Murielle Lombard, Marc Fontecave, Valérie de Crécy-Lagard, Yuri Motorin, Mark Helm, Djemel Hamdane

► To cite this version:

Damien Brégeon, Ludovic Pecqueur, Sabrine Toubdji, Claudia Sudol, Murielle Lombard, et al.. Dihydrouridine in the transcriptome: New life for this ancient RNA chemical modification. *ACS Chemical Biology*, 2022, 17 (7), pp.1638-1657. 10.1021/acscchembio.2c00307 . hal-03836746

HAL Id: hal-03836746

<https://hal.science/hal-03836746v1>

Submitted on 2 Nov 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Dihydrouridine in the transcriptome: New life for this ancient RNA chemical modification

Damien Brégeon¹, Ludovic Pecqueur⁷, Sabine Toubdji^{1,7}, Claudia Sudol^{1,7}, Murielle Lombard⁷, Marc Fontecave⁷, Valerie De Crecy-Lagard^{2,3}, Yuri Motorin^{4,5}, Mark Helm⁶, Djemel Hamdane*⁷

¹IBPS, Biology of Aging and Adaptation, Sorbonne Université, Paris, 75252, France.

²Department of Microbiology and Cell Science, University of Florida, Gainesville, Florida, 32611, USA.

³University of Florida, Genetics Institute, Gainesville, Florida, 32610, USA.

⁴Université de Lorraine, CNRS, INSERM, UMS2008/US40 IBSLor, EpiRNA-Seq Core Facility, Nancy, F-54000, France.

⁵Université de Lorraine, CNRS, UMR7365 IMoPA, Nancy, F-54000, France.

⁶Institut für pharmazeutische und biomedizinische Wissenschaften (IPBW), Johannes Gutenberg-Universität, Mainz 55128, Germany.

⁷Laboratoire de Chimie des Processus Biologiques, CNRS-UMR 8229, Collège De France, Université Pierre et Marie Curie, 11 place Marcelin Berthelot, 75231 Paris Cedex 05, France.

To whom correspondence should be addressed:

Djemel Hamdane, Laboratoire de Chimie des Processus Biologiques, CNRS-UMR 8229, Collège de France, 11 place Marcelin Berthelot, 75231 Paris Cedex 05, France, Tel : +33-(0)1-44271645, Email :

djemel.hamdane@college-de-france.fr

Abstract

Until recently, post-transcriptional modifications of RNA were largely restricted to non-coding RNA species. However, this belief seems to have quickly dissipated with the growing number of new modifications found in mRNA that were originally thought to be primarily tRNA-specific, such as dihydrouridine. Recently, transcriptomic profiling, metabolic labeling, and proteomics have identified unexpected dihydrouridylation of mRNAs, greatly expanding the catalog of novel mRNA modifications. These data also implicated dihydrouridylation in meiotic chromosome segregation, protein translation rates and cell proliferation. Dihydrouridylation of tRNAs and mRNAs are introduced by flavin-dependent dihydrouridine synthases. In this review, we will briefly outline the current knowledge on the distribution of dihydrouridines in the transcriptome, their chemical labeling, and highlight structural and mechanistic aspects regarding the dihydrouridine synthases enzyme family. A special emphasis on important research directions to be addressed will also be discussed. This new entry of dihydrouridine into mRNA modifications has definitely added a new layer of information that controls protein synthesis.

Keywords

tRNA: transfer RNA is an adaptor molecule typically composed of around 76 nucleotides that serves as a physical link between mRNA and the amino acid sequence of proteins. The tRNA performs this function by transporting an amino acid to the ribosome where it pairs via its anticodon to the codon of an mRNA.

mRNA: messenger ribonucleic acid is a single-stranded molecule of RNA that corresponds to the genetic sequence of a gene, and is read by the ribosome machinery during the translation process.

rRNA: ribosomal ribonucleic acid is non-coding RNA found as the primary component of ribosomes. rRNA allows tRNA and mRNA to process and translate the latter into proteins.

RNA modifications: chemically modified nucleotides found in mature RNA species. The modifications are specifically catalyzed by enzymes after transcription and are one of the maturation steps leading to functional RNAs.

Epitranscriptome: all functionally relevant biochemical modifications of the RNAs (the transcriptome) within a cell.

Dihydrouridine: One of the most abundant modifications of the transcriptome. This base is formed by the reduction of uridine and is therefore non-aromatic.

Dihydrouridine synthases: A family of flavoenzymes using flavin as a coenzyme and NADPH as a reductant to catalyze the dihydrouridylation (dihydrouridine synthesis) of tRNAs and mRNAs.

Flavin: flavin is a redox-active organic coenzyme associated with various enzymes (flavoenzyme).

Introduction

Many canonical nucleotide residues found in RNA polymers undergo extensive chemical modifications after the transcriptional step catalyzed by a wide variety of enzymes called "RNA modifying enzymes" (1). These post-transcriptional modifications are part of the complex maturation processes that eventually generate functional RNA molecules. Seventy years ago, pseudouridine (Ψ), also called the "fifth nucleotide" of RNA, was discovered (2). This modified base has since proven to be one of the most abundant modifications in the transcriptome. Today, more than 170 distinct chemical modifications have been identified and their number is still steadily climbing. Modifications are widely distributed among different types of RNA, including transfer RNA (tRNA), ribosomal RNA (rRNA), long non-coding RNA (lncRNA) and messenger RNA (mRNA) (1). After having been the realm of specialists for many years, enthusiasm for this field of research has been rekindled as more RNA modifications are shown to have essential biological roles, particularly in regulating gene expression (3). The understanding that some of RNA modifications (at least m^6A (4), and, probably m^1A , m^3C , m^1G and $m^{22}G$ ((5), (6), (7), (8))) may be removed by carbon oxidation and thus reversible and respond to environmental changes have revived a renewed interest in the discovery of their biological functions (9). Most of the modified bases identified to date have been in tRNA and rRNA, however, rapid advances in the field notably with new developments in the accurate detection and quantification of these epitranscriptomic marks, combined with the ability to detect low-content RNA modifications, have expanded the list of such modifications in mRNA (10). Although, detecting RNA modifications remains a challenging task, these technical advances, particularly with the availability of much more specific labeling, have allowed to establish the patterns of distribution of these modifications at the level of the whole transcriptome. On the other hand, many aspects of the enzymology and structural biology of RNA-modifying enzymes, which provide the molecular basis underlying the biogenesis of these modifications, are still lagging far behind.

Because tRNAs are the most heavily modified RNA molecules, their study has provided the core of our understanding of the RNA modification machinery and progressively revealed the functions of many RNA modifications in recent years (11-16). Most RNA modifications have a fairly simple chemical nature and often involve methylation, isomerization, reduction, or deamination reactions. For example, 5-methyluridine (m^5U), pseudouridine (Ψ), 1-methyladenosine (m^1A), dihydrouridine (D), or 2'-O-methylguanosine (G_m) (**Figure 1**) are often observed in loop regions of tRNAs and rRNAs. Evolutionarily conserved enzymes typically catalyze the biosynthesis of these classes of modified nucleosides, whose position and identity are conserved in the majority of these RNA species. Others may have a very complex chemical nature but they are usually found in the anticodon loop of tRNAs. They facilitate or prevent interactions with amino-acyl-tRNA synthetases, translation factors and they allow a precise decoding of mRNAs on the ribosome via accurate codon/anticodon interactions (17). This second category represents hypermodified bases and their biosynthesis often involves many enzymatic steps, sometimes within protein complexes. In mRNA, only simple modifications have been identified to date (**Figure 1**) and their functions are slowly being deciphered. Modifications in mRNA (N6-methyladenosine (m^6A), 5-methylcytidine (m^5C), pseudouridine (Ψ), inosine (I), N1-methyladenosine (m^1A) and 7-methylguanosine (m^7G) N4-acetylcytidine (ac^4C)), play important structural roles and are involved in RNA stability (directly or indirectly), and also found to regulate several mRNA cell cycle processes like mRNA export, splicing and translation (4, 16, 18-21). The role of m6A on RNA stability depends on the m6A binding proteins. The general belief is that m6A on mRNA promotes RNA degradation. However, when m6A is recognized by IGF2BP proteins, the mRNA is stabilized (22). Recently, D, one of the most abundant modified bases in tRNAs, that gave its name to tRNA's D-loop structure, has just entered the world of mRNA modifications with important physiological roles in cell growth (23, 24). In this review, we attempt to summarize the current knowledge about the distribution/frequencies of D modification in the transcriptome, the chemical labeling tools used to detect D, the enzymology of D formation as well as the relevance of these modifications in translation and cancer biology. In addition to the

already existing structural information on Dus proteins, we also present the first accurate structural models of human enzymes catalyzing RNA dihydrouridylation obtained by AlphaFold. The accuracy of the 3D models can be found at <https://alphafold.ebi.ac.uk/>. All these structural elements serve as a solid basis to discuss some important aspects related to the functional involvement of these proteins in epitranscriptomic under normal and pathological cellular states.

Structural versus chemical properties of dihydrouridine and quantitative approaches

D is one of the most conserved and abundant modified bases in the transcriptome, second only after pseudouridine, primarily due to its high presence in tRNAs. The D content in tRNAs seems to correlate with the growth temperature, since at high growth temperatures D ring undergoes hydrolysis (25). Thus, the highest tRNA D-content is observed in psychrophilic organisms, where 40-70% more of D was found, compared to mesophilic bacteria, while much lower D-content was found in hyperthermophiles (26-28). D is formed by the reduction of the C5-C6 double bond of the pyrimidine ring of uridine leading to a saturated base (**Figure 1**), a unique feature found to date in nucleic acids. We will see below that this property, which distinguishes it from other modifications, gives it unique physicochemical and structural properties, which can be judiciously exploited for analytical purposes. Initially discovered in tRNAs by Holley's group in the mid-1960s (29), this modified base was later observed in much lower abundance in rRNAs and some long non-coding RNAs. Early studies by Bonner had shown that D can also be found in histone-associated RNAs in plants and mammals (30), albeit without localization data. The very recent and independent studies of Dai *et al.* (23) and Finet *et al.* (24) using two different approaches came to the same conclusion, namely the presence of D in mRNAs. We now know that it is more than likely that the distribution of D in the transcriptome is broader than expected.

Structural properties of dihydrouridine

The peculiar structural properties of D-nucleoside were defined early on from the crystallographic structure of free D hemihydrate (31) and the first structures of mature tRNAs isolated from *S.*

cerevisiae (32-35) (**Figure 2**), while NMR studies on D-containing oligonucleotides have provided important information regarding the dynamical attributes that this modified base plays on RNA (36-38). These structures reveal the lack of planarity of the D heterocycle and a shift of carbons C5 and C6 to opposite sides of the plane formed by positions N1, C2, N3 and C4, resulting in a half-chair conformation (**Figure 2**, zoom on D sites). This has notable structural consequences, among them D cannot participate in stacking interactions with neighboring aromatic bases. Moreover, the presence of a CH₂ in place of a CH increases the electronic volume of the base thus promoting the inherently flexible C2'-endo conformation of its ribose. To date, this modification is the only one known to favor such a conformation whereas others RNA modifications such as 2'-O-ribose methylation (39), 2-thiolation of ribothymidine(40), and pseudouridylation(41) rather favor the C3'-endo conformation observed in A-type RNA helix, conferring stability to the RNA by reinforcing base stacking (14). Thus, these structural properties of D, and in particular the C2'-endo conformation of the sugar, locally affect the RNA structure by introducing functionally important local flexibility. This dynamic property will have different consequences depending on whether the RNA has a particular structure or not. Indeed, this increased local flexibility mediated by D may facilitate formation of interactions between neighboring tertiary bases in the critical tRNA elbow region. In fact, this region, formed by the kissing between the D and TΨC loops, involves several critical interactions that are highly conserved in cytosolic tRNAs, including the Hoogsteen-reverse base pair T54-A58, the interloop base pairs G18-Ψ55, G19-C56, and the stack of four mutually interspersed purine bases, A58-G18-A/G57-G19 (**Figure 2**). The decrease in melting temperature of *E. coli* tRNAs observed in the absence of dihydrouridine (42) could well be explained by the loss of flexibility of the D-loop required to accommodate these tertiary interactions essential for the maintenance of the tRNA 3D structure. In the course of their investigation on siRNA, Sipa *et al.* evaluated the thermodynamic stability and gene silencing activity of a series of nucleobase-modified RNA duplexes containing modified bases nucleosides, including D (43). Unlike tRNA, RNA duplexes lack tertiary structures and rather show a decrease in melting temperature when a D unit is present in their central part. Again, this effect can obviously be

attributed to the destabilizing effect of a D nucleobase on the C3'-endo sugar conformation and its nonplanar nature that disrupts stacking interactions with neighboring nucleobases. These negative effects of D on duplex formation are likely the driving force for the modified D arm to exclusively adopt the hairpin conformation in solution. Moving in that direction, Dyubankova *et al.* performed an NMR study to understand the role of dihydrouridine modification on a 15-nucleotides long D-arm of a *S. pombe* tRNA_i^{Met} mimetic. While the unmodified oligonucleotide adopts several undefined conformations that interchange rapidly in solution, the presence of D triggers a hairpin folding with a stable stem and a flexible loop region (37). Although the structural role of D in mRNAs has not yet been investigated, based on the studies we have cited above, it is likely that this base may also serve some regions to adopt a hairpin structure or perhaps to prevent alternative RNA structures in the cell.

Dihydrouridine labelling and detection

The study of modifications and their biosynthesis requires the development of tools allowing their accurate detection and quantification. Initially, modified nucleosides including D were identified solely on the basis of their chromatographic mobility involving ³²P and/or ¹⁴C labeling, 2D-electrophoresis combined with thin layer chromatography or anion exchange chromatography and HPLC analysis (44). However, these methods suffer from low specificity and reproducibility and identification becomes problematic as the number of modifications or the length of the RNA chain increases. In contrast, analysis of oligonucleoside fragments by mass spectrometry generated by treatment of the RNA polymer with RNase has proven to be a much better technique for analysis of post-transcriptional modifications, as almost all modifications produce a change in the mass of canonical nucleosides (10, 45-48). In the case of dihydrouridine an $m/z + 2$ is expected as compared to an unmodified uridine. In that respect, we have successfully applied such a methodology to determine the specificity of Dus enzymes from *Escherichia coli* and that of *Mycoplasma capricolum* (see below), but also to determine the specificity of other RNA-modifying enzymes including m⁵U

methyltransferases (49, 50) and TrmI from *Pyrococcus abyssi*, which catalyzes the sequential double methylation of A57-A58 to m¹A57-m¹A58 (51-53). However, although this technique allows accurate identification of modifications, it does not have the ability to map all D sites in the transcriptome of a given organism on a large scale and in a single shot. Newer detection and quantitative assessment methods that examine RNA modifications on a larger scale, such as in tissues or whole cells, are sequencing methods that target a subset of RNA modifications open to detection by reverse transcription (generally coupled with selective chemical treatment), yet with many limitations.

Beyond its structural properties, the saturation of the C5-C6 bond of D offers interesting consequences on its chemical reactivity, which have been judiciously exploited for RNA labeling and D-site mapping. Earlier works showed that under moderate alkaline hydrolysis, the dihydrouracil undergoes a ring opening via hydrolysis of the N3-C4 bond (54, 55). This reaction produces a β -ureidopropionic acid adduct, which is thought to lose base-pairing ability and generates primer extension arrest at D sites. Xing *et al.* successfully applied this procedure to map dihydrouridine modification sites of several cytoplasmic tRNAs from yeast with the 5' end ³²P radio-labeled primer extension technique using a primer complementary to the 3'-end region of tRNAs(56). Alternatively, the heterocyclic ring of D can also be subjected to reductive cleavage by sodium borohydride (NaBH₄) under alkaline conditions, yielding in this case a 3-ureidopropanol bound to the ribose C-1' position (57) (**Figure 3**). Interestingly, Zachau and his coworkers and others demonstrated that several R-NH₂ compounds (e.g., amines, hydrazines, hydrazides) could, in principle, be used to replace 3-ureidopropanol within tRNA (58-60), which opened up an efficient pathway for tRNA labeling as well as applications for sequencing purposes (**Figure 3**). For instance, in studies of tRNA-ribosome interactions, this approach has been used to label tRNA at the D-site by fluorophores bearing a primary amino group such as proflavin or rhodamine 110 (61). Fluorescent-labeled tRNAs have been used extensively to provide mechanistic insights into in vitro protein synthesis and protein synthesis within intact cells.

A conceptually similar approach consisting in producing an abasic site but this time for broad sequencing purposes and detection of certain modifications including D in RNA fragments from tRNAs or rRNAs has been recently developed (62). This new approach, termed AlkAniline-Seq, can map abasic sites and modified nucleotides in successive treatments that combine: i) alkaline hydrolysis of RNA at high temperatures, ii) extensive 5'- and 3'-dephosphorylation, and iii) aniline-dependent cleavage of the sugar moiety and subsequent formation of RNA abasic sites. This methodology enables simultaneous detection of 7-methylguanosine (m^7G), 3-methylcytidine (m^3C), and D in RNA at single nucleotide resolution. It should be noted, however, that the AlkAniline-Seq signal intensity is considerably lower for D than for m^7G and m^3C , which is most likely due to the fact that reduction of uridine at many potential D-sites is often partial.

In their quest to understand the mechanisms of fluorescent labeling of D in tRNA, Kaur *et al.* showed that D is converted to tetrahydrouridine in the presence of a large excess of $NaBH_4$ under acidic conditions and in a second step, its C4 hydroxyl group is replaced by benzohydrazide via nucleophilic substitution ultimately producing a covalent adduct with the base (63). Finet *et al.* have recently developed a Rho-seq integrated pipeline, based on a concept similar to that described by Kaur, but with a variation involving the replacement of the hydrazide nucleophile with rhodamine (Rho), for the purpose of transcription-wide mapping of D (24). This approach is expected to produce an irreversible RNA-base-Rho adduct primarily at D sites, providing a specific labeling for this modification (**Figure 3**). The presence of a bulky rhodamine moiety produces a specific and clear stop during reverse transcription at the predicted D-sites in tRNAs validating the specificity of this experimental procedure. Application of this Rho-seq to the *E. coli*, *Schizosaccharomyces pombe* and human transcriptome showed (i) the absence of D in *E. coli* mRNAs while the expected D2449 in 23S rRNA as well as D in tRNAs were detected, (ii) the absence of D in *S. pombe* yeast rRNA as expected, (iii) established of D distribution in tRNAs and, (iv) led to the discovery of D in mRNAs (see below).

Distribution of dihydrouridine in the transcriptome

Location and frequencies of dihydrouridine residues in non-coding RNAs

The study of tRNAs has revealed some interesting paradigm by which D influences the RNA structure/function relationship, although this modification was long neglected compared to other modifications such as pseudouridine for example. D is often found in multiple locations in bacterial and eukaryotic tRNAs and its abundance varies with both the organism and the type of tRNA (1). For example, there are up to five positions where D can be found in prokaryotes (**Figure 4A**), most frequently at positions 16, 17, 20, and 20a, all of which are located in the "D loop", and at position 47 in the variable loop (V loop), however, the latter has so far been observed in only one *Bacillus subtilis* tRNA, tRNA^{Met}_{CAU}. In eukaryotes, D is in general more abundant than in prokaryotes, since D is observed in no less than six different sites, five of which are in the D-loop (D16, D17, D20, D20a, and D20b) and one is in the variable loop (D47) (**Figure 4A**). This persistent presence of D at these positions in so many different organisms reveals the evolutionary importance of the modification and alteration sites. Beyond these "D-canonical positions", D can also locate at other positions such as 14, 17a, and 21 in D loops and at position 48 in the variable arm, but all of these cases are exceptional (see below). An accurate picture of D-distribution in tRNAs can be obtained from *Saccharomyces cerevisiae* since they have all been sequenced. These can be viewed in MODOMICS (<http://genesilico.pl/modomics>), a database of RNA modifications that provides comprehensive information concerning the location of modified residues in RNA sequences (1). Interestingly, all uridine residues in yeast tRNAs at positions 16, 17, 20, 20a and 20b in the D-loop can be converted to dihydrouridine, which may be a consequence of the high solvent accessibility observed for all dihydrouridine-modified positions as evidenced from the crystal structures of three well known yeast tRNAs, namely tRNA^{Phe}, tRNA^{Asp}, and tRNA^{Met}_i (32-35)(see **Figure 2**). To take this analysis a step further, we examined the modification pattern of D sites in all cytosolic tRNA sequences from single cells of fungi and metazoa that are available in RNA databases i.e. 173 sequences from 22 species. The observation is again consistent with the fact that uridines at positions 16, 17, 20, 20a, 20b, and 47 are predominantly modified to D (**Table 1**). Dihydrouridine is also present in mitochondrial and

plastid tRNAs, but is less frequent there, leaving many uridines unmodified (124 mt-tRNA sequences from 18 species, see **Table 1**). A comprehensive analysis performed on the bovine mt-tRNAs identified 15 types of modified nucleosides distributed over 7.5% of all mt-tRNA bases (64). Although post-transcriptional modifications in mt-tRNAs are less abundant than in cytosolic tRNAs, their biogenesis requires a large panel of specialized enzymes, some exclusively dedicated to function in mitochondria (65). Among the D residues, those at positions 16 and 20 seem to be the most widespread modifications in eukaryotic tRNAs, whether cytosolic or mitochondrial. Indeed, a recent transcriptome-wide D mapping using the Rho-seq method was able to confirm this in *S. pombe* (24). Among the 228 modified positions identified on 141 *S. pombe* tRNAs, 98.7% of the identified D-sites were located within the D-loop, with the most prevalent positions being the expected D16 and D20.

tRNAs can often harbor doubly modified bases such as m^2_2G26 , $cmnm^5s^2U54$, m^1I or even ms^2i^6A , for example, and **Table 1** shows that D is no exception to this rule. Starting with a simple modification, additional modifications can quickly lead to a hypermodified base. In search of new modifications, Krog *et al.* identified, by MS analysis of *Trypanosoma brucei* tRNA^{Lys}_{UUU} digestion fragments, a new D-derivative, namely 3-(3-amino-3-carboxypropyl)-5,6-dihydrouridine (acp^3D) at position 47 (66)(**Figure 1**). Interestingly, a pattern, with dihydrouridylation and/or acp^3 modification of various tRNA species seems to exist; however, how these modifications are introduced remains unclear. The fact that we find uridines bearing only an acp^3 group such as acp^3U47 may suggest that the biosynthetic reaction-giving rise to acp^3D takes place in a particular order where reduction would occur as the last step even if we cannot rule out that the reduction occurs first. In fact, both types of scenarios can be encountered in the biosynthesis of modified bases. For instance, a sequential synthesis of ms^2i^6A37 is observed where the initial grafting of an isopentenyl by MiaA is essential to allow the second enzyme MiaB to introduce a thiomethyl group on the C2 of A37 (67, 68). Stepwise modification also extends to the case of m^1I biosynthesis. After the conversion of A37 to I37 by ADAT1, SAM-dependent Trm5, also known for its ability to synthesize m^1G37 , methylates directly the inosine N1 nitrogen (69-72). In archaea, a SAM-dependent TrmI enzyme first methylates A57 before

it becomes a substrate for deamination to inosine (73) by a yet unidentified deaminase. The C32 in the anticodon loop of *T. brucei* tRNA^{Thr} is methylated to 3-methylcytosine (m³C) by Trm140 as a prerequisite for C-to-U deamination by the deaminase ADAT2/3 (74). The introduction of stepwise modifications at positions 34 and 37 in the anti-codon loop is frequently observed and is attributed to the fact that the first modification acts as an additional recognition determinant for subsequent modifying enzymes (75). Cases where the order does not matter are also found, such as in cmnm⁵s²U54 biosynthesis, where the C2 thiolation can occur before or after C5-U54 carboxyamino-methylation (76).

Table 1. Dihydrouridine modification profile for tRNA sequences from cytosol and mitochondria. x before U means that the uridine is modified by an unidentified chemical group.

Cytosolic tRNAs		
<i>Positions</i>	<i>modifications</i>	<i>Frequency of occurrence</i>
16	D/U	123/25
17	D/U	39/0
20	acp ³ U/D/U	4/118/11
20a	acp ³ U/D/ψ/xU/U	6/64/2/2/2
20b	acp ³ U/D/ψ/U	8/6/2
47	D/xU/U	83/1/16
Mitochondrial tRNAs		
14	D/U	1/3
16	D/U	27/21
17	D/U	8/19
17a	D/U	2/5
20	D/U	49/9
20a	D/U	7/9

47	acp ³ U/D/U	3/3/24
48	D/U	1/59

D is not abundant in rRNA, however, unlike tRNAs, the number of sequenced rRNAs remains low. D has so far only been observed at a single location, 2449, in the central loop of domain V in *E. coli* 23 S rRNA (77), two positions, 2449 and 2500, in 23S rRNA of *Clostridium sporogenes* (78)(**Figure 4B**), and one position, 1211 or 1212, in 16S rRNA of *Clostridium acetobutylicum*, but the exact location remains unclear (79). D has not been detected in other bacterial rRNAs, nor in any of the eukaryotic rRNAs sequenced to date. Interestingly, in 23S rRNA of *C. sporogenes*, D2449 was found to be methylated at the C5 atom to give m⁵D2449(78). Again, it is not known if there is a particular order in the m⁵D biosynthetic pathway. Is m⁵U formed first and then reduced to m⁵D or is dihydrouridine formed first and then its C5 is methylated? This last scenario is unprecedented so far but deserves further investigation. The enzyme(s) responsible for m⁵D biosynthesis remain presently unknown.

Presence of dihydrouridine in eukaryotic mRNAs

The first reports on chemical modifications in mRNAs date back to the 1970s, when development of poly(A) tail purification techniques made possible preparation of sufficiently enriched mRNAs, reducing eventual contamination by other abundant RNA species such as tRNA or rRNA. These studies revealed the presence of three *in situ* modifications, namely m⁶A, m⁵C, and G_m (**Figure 1**) (16). In recent years, many other modifications have been added to the mRNA repertoire such as pseudouridine (80, 81), m¹A (82) and ac⁴C (83, 84) (**Figure 1**). Very recently, large transcriptome analysis discovered D in *S. pombe* and human mRNAs, however no trace of this modification was detected in bacterial mRNAs (23, 24). D is distributed predominantly in the coding regions of conserved genes, supporting a biologically relevant function of dihydrouridylation, which may as well be a general feature of eukaryotic mRNAs. The methodologies and strategies that led to

mapping the D sites in the transcriptomes have been discussed in the preceding sections. Of the total D sites detected (372 sites) in the *S. pombe* transcriptome, 38% and 61% are found in mRNA and tRNA, respectively (**Figure 4C**). Of the 125 D-containing mRNAs in *S. pombe*, 87% have a single putative D, whereas only two mRNAs (encoding non-classical export protein and alanine-tRNA ligase) carry at least three distinct D sites (**Figure 4D**). In colon human cells (HCT116), 112 D-sites within mRNAs were also identified. The total D content in mRNAs, however, appears to be relatively low compared with that observed for other modifications. The exact numbers of modified residues vary between studies. The numbers of reproducible peaks seem to converge on approximately 13,000 sites in 5000–7000 mRNAs for m⁶A(85) (86, 87). Regarding pseudouridine, approximately 250 to 300 sites have been mapped in yeast mRNA at single base resolution (80, 81), however overlap between these sites is only modest. In human, the reported number of mRNA pseudouridine sites varies widely from 96 in one study, 23 to 353 in a second, 22 and up to 2084 in a third study (80, 88). For m¹A, the first study identified 7,154 peaks in 4,151 coding genes, while the other identified 887 peaks in 600 genes (89-92), while another study reported only a handful of sites (93). It is therefore important to keep in mind that the number of identified Ds in mRNAs can vary depending on the methods used for mapping. Dihydrouridylated mRNAs derived from genes with conserved functions as it was estimated that 73% of these are conserved in vertebrates (24). Notably, several mRNAs encoding cytoskeleton-related proteins have been identified as D-containing transcripts. For instance, in *S. pombe* *nda2* and *nda3* encoding the α - and β -tubulin are found to be a D-containing mRNAs presenting a single D at position 1133 in GDU valine codon and 586 in DCU serine codon, respectively (**Figure 4D**).

Dihydrouridine synthase enzymes: structure-function relationship

A large family of enzymes called dihydrouridine synthases (Dus) that uses flavin mononucleotide (FMN) as coenzyme catalyzes the synthesis of dihydrouridine in tRNAs (94), mRNAs and bacterial YrIA lncRNA. Nicotinamide adenine dinucleotide phosphate (NADPH) serves as a reducing agent for

flavins. This family of flavoenzymes is classified into three major groups and eight subfamilies, all of which evolved by independent duplications of an ancestral gene (95). The first group found in prokaryotes includes three Dus (Dus A, B and C) while the second is eukaryotic and contains four different enzymes (Dus 1 to 4). Archaea has a single Dus and it characterizes the last group. The bacterial enzyme, DusB, is considered as the oldest enzyme while Dus3 could be the ancestral eukaryotic enzyme, from which the other Dus were derived by gene duplication starting with Dus2, then Dus1 and finally Dus4.

Dihydrouridine site specificities

One of the experimental evidences for the enzymatic introduction of D into RNAs comes from the work of Grosjean's lab which systematically tested all enzymatic activities in cell-free yeast extract and showed formation of 11 of the 14 naturally occurring modifications present in mature yeast tRNA^{Phe}_{GAA}, including D17, on a radioactive tRNA transcript (96). It was shown that D17 biosynthesis is insensitive to the presence of the intron since the modification reaction proceeds with the same efficiency in both the intron-free and intron-containing tRNA^{Phe} precursor. Five years later, the first genes encoding yeast dihydrouridine synthases were discovered by screening the genomic library of *S. cerevisiae* GST-ORF proteins with pre-tRNA^{Phe} substrate whose dihydrouridylated form carries both D16 and D17, and showed both *in vivo* and *in vitro* that Dus1 catalyzes D17 synthesis in several tRNAs (97). The complete characterization of the site specificity of the 4 existing Dus in *S. cerevisiae* could eventually be achieved by the same group using 3 complementary techniques, namely (i) determination of the molar ratio of dihydrouridine in purified tRNAs from different dus mutants; (ii) microarray analysis of a large number of tRNAs based on differential hybridization of uridine- and dihydrouridine-containing tRNAs to complementary oligonucleotides; and (iii) the primer extension analysis (56). These *in vivo* studies established that Dus1, Dus2, Dus3 and Dus4 are responsible for the synthesis of D16-D17, D20, D47, and D20a-D20b, respectively (**Figure 4A & Table 2**). Furthermore, the Dus proteins faithfully retain their specificity in the absence of the other Dus,

indicating that they have non-redundant activities. In parallel, Bishop *et al.* used comparative genomics and computational methods to identify members of the orthologous gene cluster, COG0042, annotated in protein sequence databases as 'predicted TIM-barrel enzymes, possibly dehydrogenases, nifR3 family' as genes encoding dihydrouridine synthase and identified three members of the COG0042 family in *E. coli*, named DusA, DusB, and DusC, which are responsible for introducing all D content into tRNAs and that DusA is likely the D20-catalyzing enzyme in tRNA^{fMet}₂ (98). Recently, we and others have contributed to unravel the specificity of Dus enzymes from several model organisms, including those in humans. In *E. coli*, DusA, DusB and DusC catalyze the formation of D20-D20a, D17 and D16, respectively (52, 98, 99). *Thermus thermophilus* harbors a single Dus, which is a homolog of DusA and has been shown to catalyze the formation of D20-D20a (100, 101). DusA is also responsible for the biosynthesis of D76 in the lncRNA YrIA of *Salmonella typhimurium* (102). Indeed, D is observed at position 76 in YrIA from *S. typhimurium*, which is a long non-coding RNA that attaches the Ro60 protein to the polynucleotide phosphorylase, thus targeting this exoribonuclease for degradation of structured RNAs. In *M. capricolum*, the single DusB introduces all D content present in tRNAs, namely D17, D20, D20a (53). In *S. pombe* and *H. sapiens*, the Dus enzymes share the same site specificities for tRNAs as those of *S. cerevisiae* (23, 24, 103). However, it was shown that all four Dus enzymes are involved in modifying both tRNA and mRNA and in the case of Dus3, mRNAs seem to be its predominant target (23, 24). It is plausible that Dus-catalyzed mRNA dihydrouridylation is a conserved functional feature in eukaryotes, however no D residues were reported in *S. cerevisiae* mRNA (84). The involvement of Dus in the synthesis of D in both tRNA and mRNA is not unique as it has now been reported that most enzyme systems catalyzing mRNA marks are those involved in the modification of other RNAs, mainly tRNAs. The only known exception to date regards the biosynthesis of m⁶A in mRNA that is catalyzed by the dedicated METTL3/METTL4 complex (9). For example, NSUN2, the human SAM-dependent tRNA methyltransferase involved in the formation of m⁵C at positions 34, 48, 49, and 50 of tRNAs (104-106), also mediates m⁵C synthesis in nearly 300 mRNAs (107). For comparison, the orthologous yeast

enzyme Trm4 catalyzes m⁵C formation at positions 34, 40, 48 or 49 depending on the tRNA (108). The *S. cerevisiae* several pseudouridine synthases demonstrate multisite substrate specificity, Pus1 modifies tRNAs at multiple locations (109), as well as U2 snRNA (110), Pus4 forms a universally conserved ψ55 in tRNAs (111), but was also reported to modify mRNAs ((112, 113), while Pus7 modifies U13 in several tRNAs, U35 in pre-tRNA^{Tyr}, U35 in the small nuclear RNA U2, U50 in 5S rRNA (114, 115), and several U residues in mRNA (116). In human cells, the majority of cytoplasmic tRNAs carries the m¹A58 modification catalyzed by the heterotetrameric TRMT61A/TRMT6 enzyme complex, which is also responsible for the m¹A modification in mRNAs (93).

Today, the enzymes responsible of D incorporation in rRNAs remain to be identified since in the case of D2449 present at the peptidyl transferase site in *E. coli* 23S, the deletion of the 3 bacterial *Dus* genes does not abolish its formation (24), suggesting that another class of dihydrouridine synthase specific to rRNAs must exist in this bacteria. Finally, the enzymes that introduce D into mitochondrial tRNAs also remain to be identified. However, it might be possible that the cytosolic *Dus* are also responsible for the biosynthesis of the corresponding D in mitochondria. This dual specificity is common for other tRNA modification enzymes such as pseudouridines synthases (111, 117) or m⁵U54 tRNA methyltransferases (118). The mammalian *Dus2* was indeed detected in mitochondria in a study aiming to create a mitochondrial compendium of 1098 genes and their protein expression across 14 mouse tissues (119).

Table 2. Experimentally established specificities of dihydrouridine synthases

Enzymes	Organisms	Substrate(s)	Products in tRNA	Refs
DusA	<i>E. coli</i>	tRNA	D20, D20a	(52, 98)
	<i>T. thermophilus</i>	tRNA	D20, D20a	(101)
DusB	<i>E. coli</i>	tRNA	D17	(52)
	<i>M. capricolum</i>	tRNA	D17, D20, D20a	(53)

DusC	<i>E. coli</i>	tRNA	D16	(52, 99)
Dus1	<i>S. cerevisiae</i>	tRNA	D16, D17	(56, 97)
	<i>S. pombe</i>	tRNA, mRNA	D16, D17	(24)
	<i>H. sapiens</i>	tRNA, mRNA	D16	(24)
Dus2	<i>S. cerevisiae</i>	tRNA	D20	(56, 97)
	<i>S. pombe</i>	tRNA, mRNA	D20	(24)
	<i>H. sapiens</i>	tRNA	D20	(24, 103)
Dus3	<i>S. cerevisiae</i>	tRNA	D47	(56)
	<i>S. pombe</i>	tRNA, mRNA	D47	(24)
	<i>H. sapiens</i>	tRNA, mRNA	D47	(23, 24)
Dus4	<i>S. cerevisiae</i>	tRNA	D20a, D20b	(56)
	<i>S. pombe</i>	tRNA, mRNA	D20a, D20b	(24)
	<i>H. sapiens</i>	tRNA	D20a	(24)

Structural analysis of dihydrouridine synthases

The X-ray crystallographic structures of some Dus have defined a canonical fold for this family of enzymes (**Figure 5**). To date, the structures of the three bacterial Dus homologues, namely DusA from *T. thermophilus*, DusB and DusC from *E. coli* were solved (52, 99, 100, 120). However, the data are scarce for eukaryotic enzymes since only the structures of the isolated domains of human Dus2 (hDus2) recently solved by our group and that of Antson are available (**Figure 5**), seriously impeding our understanding on this subfamily of enzymes (103, 121). All Dus enzymes share a common catalytic N-terminal domain, which adopts a TIM-Barrel type fold. The latter carries an active site circumscribed within a solvent-accessible crevice providing a binding site for the FMN prosthetic group and uridine substrate. The catalytic domain is followed by a domain formed by four parallel helices organized in a bundle and which constitutes the helical domain (HD). Beyond these

domains, the canonical fold can accommodate additional domains that are added either on the N-terminal end or on the C-terminal end or even both, but this concerns essentially eukaryotic Dus (95, 98, 103). Obviously, this complex modularity could serve new biological purposes that do not exist in prokaryotic organisms, such as RNA substrate transport and localization, or specialization in the recognition of various RNA substrates, as is the case in yeast and human Dus involved in tRNA and mRNA modification.

The structures of *T. thermophilus* DusA and *E. coli* DusC in complex with tRNAs have elucidated the molecular basis of tRNA substrate recognition by bacterial enzymes (**Figure 6A & 6B**) (99, 100). Here, the two canonical domains provide the platform for RNA recognition, and involve numerous ionic interactions, including positively charged residues (i.e. Lys, Arg) that interact specifically with bases, ribose and phosphate constituting the substrate backbone. The recognition mode actually differs depending on whether the enzyme is specific for U16 or U20. While the structure of both proteins is globally conserved, these enzymes access the target uridine by positioning their substrate tRNAs differently on their respective surfaces involving a 160° rotation from tRNA on DusC to that of DusA (**Figure 6A & 6B**). DusA recognizes a more important portion of the tRNA than DusC does, in particular the anticodon stem that DusC does not bind. Both enzymes recognize the elbow region, the D- stem-loop and the T-loop. In the case of DusA, the D and T-loops are recognized only by the TIM-Barrel domain while the D and anticodon stems are recognized by the HD. In contrast, in the case of DusC, the D-loop is recognized by both canonical domains while the D-stem is recognized only by the TIM-Barrel domain and the HD recognizes the T-loop. In these complexes, the Dus enzymes appear to bind to tRNAs without disrupting the crucial interactions that maintain their tertiary structure. Thus, the tRNA elbow must be a quality control checkpoint that Dus scrutinizes before dihydrouridylation. Finally, the two enzymes flip their uridine substrate and stack it on the isoalloxazine to proceed with its reduction (see the mechanism of Dus in the next section).

We have recently shown that the recognition mechanism of the tRNA substrate by *Homo sapiens* Dus2 (hDus2 or Dus2L) is much more complex than that observed in bacterial enzymes (**Figure 6**) (122). Indeed, hDus2 has a structural insertion within the TIM-Barrel and an additional double stranded RNA binding domain (dsRBD) that is appended to the polypeptide just after the HD, both playing a role in tRNA recognition (**Figure 5 & 6**) (103, 122, 123). The dsRBD is a double-stranded RNA (dsRNA) recognition module and is mainly found in proteins involved in mRNA transport, processing or editing (124-127). Our structures of the hDus2 dsRBD in complex with a dsRNA (**Figure 6C**), as well as in-depth investigations by site-directed mutagenesis, nuclear magnetic resonance (NMR) and small angle X-ray scattering (SAXS) of the interaction of the dsRBD with human tRNA^{Lys}₃ revealed how this domain binds to tRNA. Indeed, this domain has a particular mode of tRNA recognition involving, in addition to the well-known canonical interactions between dsRBDs and dsRNAs (ribose and phosphate recognition), specific interactions with RNA, i.e. hydrogen bonds between some residues of dsRBD and RNA bases. The model of the hDus2/tRNA complex showed that in addition to engaging the TIM-Barrel and to a lesser extent the HD, the dsRBD provides an important substrate recognition platform by binding to the long double-stranded region formed by the junction between the acceptor and T stems of tRNA (**Figure 6D**). Interestingly, we can infer that mammalian Dus2 encompasses almost the entire tRNA, where only the anticodon loop seems to be left out. Although we have no experimental data at this stage, we believe that the dsRBD may also be important for mRNA with dsRNAs structures recognition but this remains to be tested. For instance, the dsRBDs of adenosine deaminase ADAR2, which acts on mRNA to recode genomic information by the site-selective deamination of adenosine, binds to a stem-loop pre-mRNA encoding the R/G editing site of GluR-2 by recognizing the shape and sequence of the dsRNA (128).

3D-models of human dihydrouridine synthases

Recently a revolution in the world of structural biology has occurred with the arrival of artificial intelligence in the high precision prediction of 3D protein models generated by AlphaFold

(<https://alphafold.ebi.ac.uk/>) (129), opening a new area in protein modeling (129-131). This accuracy applies not only to the prediction of folding but also to the positioning of residues within active sites for example. By applying AlphaFold2, Tunyasuvunakool *et al.* markedly expand the structural coverage of the proteome at a scale that covers almost the entire human proteome (98.5% of human proteins) (132). They also provide some case studies to illustrate how high-quality predictions could be used to generate biological hypotheses. We have seized this unique opportunity to obtain accurate 3D models of the four human Dus as shown in **Figure 7**. The resulting models predict, as expected, the conservation of the canonical domains in all four human Dus with, however, some peculiarities that can be noted. For example, the beta-sheet inserted into the TIM-Barrel of hDus2 is absent in the other three enzymes, making it a unique feature of this enzyme subfamily. While the TIM-Barrel structure is largely maintained in all four Dus, Dus3 has an HD that is different from the others (see below for details). With the exception of Dus4, which has no additional domain, all the others have additional structural elements added to their respective canonical structures. Dus1 has an extension of 154 residues on the C-ter side, part of which forms three independent helices (368-380, 412-423, and 432-448), while the remained sequence are organized into loops and unstructured regions. Regarding Dus2, the model is in perfect agreement with our structures wherein as indicated above, the HD is followed by the dsRBD. The protein ends with a predicted unfolded 50-amino acid extension, probably an intrinsically disordered region as we recently showed whose truncation does not abrogate tRNA dihydrouridylation (103). It should be noted that a structural genomics group has published in the PDB the solution structure of the isolated dsRBD domain of mouse Dus2 under the code 1WHN and annotated "Solution structure of the dsRBD from hypothetical protein BAB26260". In this NMR models, the long C-terminal extension is disordered likely due to the lack of restraints, a consequence of the intrinsic flexibility of this region (**Figure 5**). Dus3 is the enzyme that is clearly distinguished from the three others by the presence of several features that we list as follow: (i) it is the Dus orthologue with the largest size, with a polypeptide of 650 amino acids, i.e. twice as large as Dus4 for example; (ii) the enzyme has a much more complex modularity with the addition of several

structural elements that are appended to the N-terminal side of the sequence thereby extending several regions of the canonical domains in the 3D space. Specifically, residues 36-60 form a helix that flanks the HD, extending that domain to the right, while residues 211-236 adopt two helices that cover the TIM Barrel. An additional zinc finger domain (ZnD) is inserted between these structural elements and is positioned above the HD. A zinc finger is a small protein structural motif that is characterized by the coordination of one or more zinc ions (Zn^{2+}) to stabilize the fold. However, despite the wide variety of these motifs, the vast majority typically functions as interaction modules that bind DNA or RNA and structural variations serve primarily to alter the binding specificity of a particular protein. In the case of hDus3, the ZnD carries a C161-X8-C171-X5-C178-X5-H183 motif, with perfect pre-organization of the Zn^{2+} binding site; as if the cysteines and histidine ligands were pre-oriented to readily coordinate the metal. (iii) Finally, the four helices of the HD no longer form the bundle that characterizes this domain, while it is interspersed with an insertion of about 20 amino acids that form a large loop.

All these structural features of eukaryotic Dus raise an important question: What exactly is the utility of having these extensions knowing that, for instance, Dus4 can obviously function only with the canonical architecture and counts tRNAs and mRNAs as well among its known substrates. These structural elements may play a functional role, such as cooperative participation with the canonical domains in substrate recognition, as we recently demonstrated for the dsRBD of hDus2 (103, 122). In the case of the zinc finger of Dus3, the region of this domain that faces the HD bears a positive electrostatic surface suggesting its likely involvement in RNA binding. Thus, all these models offer interesting perspectives to evaluate in the future the role of these extensions and domains in the dihydrouridylation reactions of RNA substrates.

Molecular mechanism of dihydrouridine biosynthesis and inhibition

Chemical mechanism of the Dus-catalyzed dihydrouridylation

The chemical mechanism of Dus has been primarily elucidated through the study of yeast Dus2 (133), subsequently corroborated by crystallographic structures of several Dus active sites, including those obtained in complex with tRNA. Overall, the catalytic cycle of these enzymes is composed, like for most flavoenzymes, of a first reductive step that involves the reduction of FMN by the natural reductant followed by a second step that involves the oxidation of the reduced flavin by the second substrate (**Figure 8A**). The fast kinetic stopped-flow approach by Rider *et al.* showed that NADPH rapidly reduces FMN (2.5 s^{-1} at 4°C) to give rise to a flavin hydroquinone (133). We also showed, by monitoring the oxidase activity of hDus2, DusB from *E. coli* and *M. capricolum*, that these enzymes prefer NADPH to NADH. Although the latter may be a potential reductant, kinetic evidences argue against its utilization while strengthening general use of NADPH as the physiological substrate by all Dus enzymes (52, 53, 103). Only a structure in complex with NADPH will eventually identify potential residues involved in the physiological discrimination of the flavin reducing agent. The reduction of FMN to hydroquinone by NADPH occurs via a hydride transfer reaction and it was shown that Dus2 is specific for the proR hydrogen of NADPH (133) (**Figure 8B**). We propose that the form of hydroquinone produced in this reaction is FMNH^- and not FMNH_2 as previously proposed (100, 133). This is supported by the fact that all the Dus structures and models strictly conserves a lysine residue (K147, K155, K435, K158, K132, K139 in hDus1, hDus2, hDus3, hDus4, *T. thermophilus* DusA, *E. coli* DusC, respectively) in their active sites that faces the N1-isoalloxazine and that is perfectly poised to stabilize the eventual negative charge built on this flavin nitrogen via hydrogen bonding (**Figure 8C & 9A**). Confirmation of its importance in the biosynthesis of D was evidenced by its replacement by an alanine in *E. coli* DusA causing a collapse of D-level in a triple-Dus *E. coli* strain complemented with *dusA-K153A* mutant compared to the same strain complemented with a wild-type *dusA* (134). Another strictly conserved active site residue, namely a glutamine (Q79, Q87, Q365, Q87, Q63, Q68 in hDus1, hDus2, hDus3, hDus4, *T. thermophilus* DusA, *E. coli* DusC, respectively), engages two H-bond with C2=O and N3H of FMN, to assist in the stabilization of FMNH^- (**Figure 8C & 9A**). Thus, this structural information implies that FMNH^- is likely to be the RNA reducing entity. The structure of

bacterial Dus in complex with tRNA and that of hDus2 also reveal that there is not enough space to accommodate both NADPH and the target uridine at the same time. Hence, after the reduction of FMN by NADPH, the first product of the reaction, namely NADP^+ , will have to leave the active site to allow the accommodation of the second substrate, which is consistent with a ping-pong type enzymatic mechanism. In a second step, FMNH^- transfers its hydride to the pyrimidine substrate uridine at C6 followed by a subsequent protonation step at C5 breaking eventually the unsaturation character of the base (**Figure 9B**) (133). A conserved cysteine in Dus acts as a general acid in the protonation step (C108, C116, C396, C116, C93, and C98 in hDus1, hDus2, hDus3, hDus4, *T. thermophilus* DusA, *E. coli* DusC, respectively) (**Figure 8C & 9A & 9B**). This redox mechanism is validated by the structures of *T. thermophilus* DusA:tRNA^{phe} and *E. coli* DusC:tRNA^{phe} complexes (99, 100) where we clearly distinguish the substrate uridines stacked against the *si*-face of the isoalloxazine with their C5 pointing within 3.5 angstroms of the hydride donor, i.e. the N5-FMN (**Figure 9A**). This productive orientation of uracil is firmly maintained by 2 hydrogen bonds between, on the one hand the C2=O of the pyrimidine with the side chain of a polar amino acid (R178 and Y176 for DusA and DusC, respectively) and, on the other hand between the C4=O of uracil and the side chain of an asparagine strictly conserved in Dus (**Figure 9A**). It is worth mentioning that similar mechanisms are also observed in dihydropyrimidine and dihydroorotate dehydrogenases, both of which share a TIM-Barrel catalytic domain homologous to that of Dus enzymes (98, 135, 136).

Dihydrouridylation and modifications interdependence

Remarkably, the study by Rider *et al.* showed that the rate of U20 reduction by Dus2 is dependent on the presence of other modifications (133). Indeed, U20 of yeast tRNA^{Leu} is reduced 600 times faster when the substrate is matured, i.e. it bears all its modifications. Similarly, we have shown that the dihydrouridylation activity of hDus2 is more efficient on modified bulk tRNAs than on naked transcripts (103). This functional behavior does not seem to be a peculiarity of the fungal and mammalian enzymes because *T. thermophilus* DusA also behaves in the same way. These studies

raise the possibility of modification interdependencies that may translate into a precise order in the biosynthesis of certain modifications and D20 could be among those that appear late in the maturation pathway. In this context, Barraud *et al.* have ingeniously developed a recent methodology allowing them to follow the maturation of labeled tRNA^{Phe} in yeast cell extracts using time-resolved NMR (137). Remarkably, they observed that modifications are introduced in a predefined sequential order and that this timing seems to be controlled by interdependencies between modification elements. Among these events, the introduction of D by Dus1 occurs after the biosynthesis of Ψ55 and m⁵U54 in the T-arm. While D has a beneficial effect on the biosynthesis of m¹A58 in the T-arm, it is shown to have a negative effect on the introduction of m²G10 in the D-arm, revealing a complex connected circuit that could be specific to each tRNA or even each isoacceptors. It is important to mention that such hierarchical circuits not only concern tRNA core modifications but are quite frequent for modifications affecting the anticodon loop region. In these cases, the modification circuits can enhance the specificity of the modification enzymes by using the first modification in the circuit as an additional recognition identity factor for the following modification (75). Nevertheless, beyond this concept, the structures of Dus/tRNA complexes may indeed provide convincing rationale for these phenomena at least for D biosynthesis. As previously noted, dihydrouridylation appears to require the architectural integrity of the tRNA elbow, so it is quite logical to expect that any modification that stabilizes tRNA tertiary interactions would have a direct positive impact on the efficiency of D biosynthesis. Interestingly, Cavaille *et al.* observed that D20 biosynthesis in tRNA incubated in yeast cell-free extracts is completely abolished in tRNA mono-mutants affecting substrate architecture (G18C, G19C, C56G) (138). Consistent with this analysis, D is introduced after isomerization of U55 and methylation of U54, both of which are known to stabilize the elbow structure. It will therefore be interesting to validate this hypothesis by determining the dihydrouridylation activity on tRNAs lacking these two modifications.

Activity-based inhibition of dihydrouridine synthases

Dai et al. recently developed a chemo-proteomic strategy based on an RNA-mediated protein profiling approach to map the *in vivo* interactions existing in human cells between C5-pyrimidine-RNA methyltransferases and mRNAs by directly feeding the living cells with 5-fluorocytidine (5FC) or 5-fluorouridine (5FU), both analogues of C or U, respectively (23). The principle is to use these fluorinated nucleotides as baits, thanks to their potential incorporation into RNA especially at the sites targeted by these enzymes, and to generate dead-end enzyme/substrate covalent complexes formed during catalysis. This strategy has been widely used *in vitro* to trap covalent complexes between C5-pyrimidine-RNA methyltransferases and miniRNAs containing fluorinated substrate analogues with a fluorine atom at the C5 of the target pyrimidines during reaction with the natural carbon donor (*S*-adenosylmethionine for SAM-dependent m^5C or m^5U methyltransferases, and 5-methylenetetrahydrofolate for folate-dependent m^5U methyltransferases) (139-142). The mechanistic principle was established a long time ago and consists of an activity-based inhibition. Like all C5-pyrimidine methyltransferases, the catalytic mechanism requires, as a first step, pre-activation of the base by addition of a cysteine present in the active site to the C6 carbon via a nucleophilic attack (Mickael's addition type of reaction). This leads to a transient enolate with an activated C5. The latter then attacks the methyl donor allowing the transfer of the methyl group onto the RNA. Finally, labile H5 proton abstraction by a base residue triggers beta-elimination and dissociation of the RNA from the protein. However, when F, Br or I replaces the H5, as the general base cannot abstract the halogen atom the covalent RNA-Enzyme species is the final stable product. This method has unexpectedly led to the capture of an mRNA-hDus3 covalent complex (23). This cross-linking was not observed when the potential general acid in the dihydrouridylation mechanism, namely C396, was replaced by an alanine suggesting that in the presence of a 5-halodihydrouridine, this residue likely becomes a reactive nucleophile. An activity based mechanism for this Dus-RNA trapping can be formulated as proposed in **Figure 9C**, which is in theory a strategy applicable to all Dus. (i) Reduction of FMN to FMNH⁻ by NADPH, (ii) hydride transfer from FMNH⁻ to C6 of 5FU, (iii) protonation of C5 by the conserved cysteine and formation of 5-fluorodihydrouridine (5FD), (iv)

formation of the covalent bond between cysteine and dihydrouridine by nucleophilic substitution of fluorine leading to an RNA-Dus adduct.

Dihydrouridine and biological implications

Biological D-relevance in RNA folding and architectural stability

The relevance of D in the conformational dynamics of RNAs was established early on by structural approaches (see previous sections). On the other hand, this function took on its full biological meaning when MacCloskey analyzed the quantitative composition of post-transcriptional modifications in tRNAs isolated from psychrophilic organisms having the capacity to grow under extremely low temperatures ranging from -5°C to 12°C (26, 27). These organisms have implemented evolutionary strategies to counteract the restriction of molecular mobility and to maintain a form of resilience in the face of low temperature by incorporating into their biomolecule, biochemical components that have the capacity to maintain molecular flexibility. Among these compounds, the observed high incorporation of unsaturated fatty acids in the lipid bilayers tends to fluidify the cell membranes while the limitation of hydrophobic clusters or salt bridges, known for their stabilizing effect on proteins. Similarly, these organisms show much less post-transcriptional modifications in their tRNAs, although they retain some of them such as pseudouridine, m⁵U, and m⁷G at normal levels. In dramatic contrast, D levels are found to be between 40 and 70% higher than those found in mesophilic organisms, such as *E. coli*. Thus, these biological data together with structural information corroborate the role of D in promoting the local fluctuation and mobility of nucleic acids.

Depletion of D by deletion of *dus* genes does not cause significant defects in growth phenotype of model organisms such as eubacteria *T. thermophilus* (101), *E. coli* (98) or *S. cerevisiae* (56). Similarly, in the absence of identified enzyme catalyzing rRNA dihydrouridine synthesis, O'Connor *et al.* removed D2449 and evaluated its effect on *E. coli* physiology and ribosome function by direct mutation of U2449 to C2449. The results obtained in this study pointed out that ribosomal D2449 is dispensable to the cell (143). These results are not surprising by themselves because many

so-called "non-essential" modifications, which are located mainly in the body of the RNA molecule, produce only minor phenotypic impact following their removal. On the contrary, this phenomenon makes sense given the fact that these modifications are part of an interconnected network where compensation phenomena, or functional redundancy, may occur. The biological relay of these modifications becomes relevant when this network is disturbed under particular stressing events or beyond the simple loss of a single modification. It is in this context that Phizicky uncovers the importance of D in combination with m⁷G46 in yeast tRNA^{Val}_{AAC} (144). Indeed, the double mutants *dus/trm8*, and particularly *dus3/trm8* produce severe growth defects. Molecular analysis has revealed that this growth defect coincides with a rapid decrease at the steady-state level of the pool of this tRNA via its rapid intracellular degradation, which approaches the degradation rates of mRNAs. Hence, by maintaining the functional folding of RNAs in cooperation with its relatives, D acts as a kind of quality control mark for RNAs.

Dihydrouridine and protein translation

Finet *et al.* recently showed that D has a critical role in the control of cell cytoskeleton dynamics in *S. pombe* via its presence in a codon of the *nda2* and *nda3* mRNAs encoding alpha and beta-tubulin, respectively (24). The absence of D in these mRNAs, obtained either by deletion of *Dus3* or by replacement of D by C using mutagenesis, leads to the same result, namely a cell growth defect observed in the presence of the depolymerized tubulin drug 2-(4-thiazolyl)benzimidazole. This growth defect appears to be attributed to a meiosis problem caused by an excess of tubulin. Here, the function of D is to slow down the translation rate of the alpha and beta tubulin genes to allow a controlled accumulation of the alpha/beta-tubulin pool compatible with functional concentrations preventing any imbalance of these components, which is known to be detrimental to cell's life. Conversely, the absence of D provokes overexpression of these cytoskeleton proteins, which ultimately leads to altered chromosome segregation and reduced gamete viability. D in the human transcriptome seems to have an opposite behavior on translation since it has been shown that its

absence obtained via hDus3 deletion impairs translational efficiency leading to a strong decrease in cell viability (23). We speculate that D could therefore be considered as a key mark in the control of the translational homeostasis of certain genes.

Implication of dihydrouridine and Dus enzymes in cancers

Because of their importance in translation, defects in post-transcriptional RNA modifications and in enzymes that catalyze them are often associated with severe human diseases (145-151). Situations where an over-representation of certain modifications can also occur in some cancers, however, the molecular mechanisms that link aberrant RNA modifications to human diseases are largely unknown. The case of overexpression of modifying enzymes in some cancers can also be observed (145). Both types of situation are observed in some cancers for the case of dihydrouridine, i.e. (i) increased D content, (ii) overexpression of Dus (see below) and (iii) Dus activity (specific or not). However, the link between (i), (ii), and (iii) has never been established and therefore it cannot be concluded whether there is a correlation between overexpression of the enzyme and increased tRNA dihydrouridylation activity.

(i) Increase of D levels

Kuchino & Borek reported in the late 1970s excessively abnormal levels of D, together with m⁵C, in tumor-specific phenylalanine tRNA isolated from Novikoff hepatoma and Ehrlich ascites cancerous tissues (152). These pathogenic tissues showed an increase in D content of up to 50% compared to healthy tissues. Most of the subsequent work focused only on the observed increased methylation yet the increase in D levels raises interesting questions that are yet to be addressed. Of note, human tRNA^{Phe}_{GmAA} has three D residues at positions 16, 17 and 47 on its sequence (1). Given this information, one can ask whether U16, U17 and U47 are fully modified in tRNAs from healthy cells, and if not, this could possibly suggest that the increase in D content observed in tumorigenic tissue is the result of complete dihydrouridylation of vacant D-sites as it was known that D modification is rarely stoichiometric (153).

(i) Dus overexpression

By using screening for up-regulated genes in cancer cells based on genetic information obtained on cDNA microarrays, combined with high-throughput screening of their effects on cell growth, Kato *et al.* found that hDus2 is frequently overexpressed in clinical lung cancer samples and non-small lung cancer cell lines, and that the overexpressed hDus2 is necessary for survival/growth of lung cancer cells (154). The contribution of hDus2 to lung carcinogenesis was revealed using siRNA to suppress hDus2 expression, showing a reduced dihydrouridylation of total RNA and a growth suppression of these pathogenic cells. However, it is important to note that this observation does not mean that Dus2 activity is per se responsible for tumor activity since it is expected that the decrease in hDus2 biosynthesis by siRNA would decrease its physiological activity, i.e. D20 biosynthesis (see below). Specific interaction of hDus2 with cellular partner has also been documented. Using immunoprecipitations assays, a multifunctional glutamyl-prolyl-tRNA synthetase, which catalyzes the aminoacylation of glutamic acid and proline tRNA species, has been found to interact with hDus2 (154), but the exact function of such an interaction is still obscure. Mittelstadt *et al.* also reported the capacity of hDus2 to interact with other cellular partner, namely protein kinase R (PKR) and Interferon-inducible double-stranded RNA-dependent protein kinase activator A (PACT), using yeast two-hybrid screen and immunoprecipitation assays (155). The interferon-induced, dsRNA-activated protein kinase PKR, a serine/threonine kinase, is a major mediator of the antiproliferative and antiviral actions of interferon (IFN) (156-158). Although induced at the transcriptional level by IFNs, PKR is present at a low, basal level in most cell types until it binds to its activators, including the protein activator PACT. Once activated, PKR phosphorylates eukaryotic translation initiation factor 2 (eIF2 α) leading to an inhibition of protein synthesis. Binding of hDus2 to PKR resulted in an inhibition of PKR activity both *in vitro* as well as in mammalian cells. Moreover, overexpression of hDus2 seems to inhibit stress-induced apoptosis indicating that it acts as an important negative regulator of PKR activity in cells (155). How hDus2 enhance the rate of translation is not clear but inhibition of eIF2 α phosphorylation brought about by activated PKR could be a possibility for efficient translation.

Recent advances in transcriptome sequencing have facilitated identification of novel fusion transcripts in human gastric cancer, including a Dus4-Bcap29 fusion transcript present in most of the analyzed gastric cancer tissues (159), Bcap29 being a B-cell receptor associated protein (160). Dus4-Bcap29 fusion transcript exists also in a variety of normal tissues notably in noncancerous epithelial and fibroblast cell lines (161, 162). Suppression of the highly expressed Dus4-Bcap29 transcript without affecting expression of Dus4 stopped cell proliferation while a siRNA specifically targeting Dus4 did not. Induced overexpression of this transcript in SNU-638 cell lines promotes cell proliferation in a time-dependent manner suggesting that the Dus4-Bcap29 is tumorigenic in gastric cancer.

(iii) Dihydrouridylation

As mentioned above, there has never been a clear link between RNA dihydrouridylation activity and the possibility of leading to tumor processes. However, we would like to discuss this aspect, even if it is more speculation than established evidence. Human Dus2 protein appears to impact translation efficiency since in vitro rabbit-reticulocyte lysate system increased protein production in the presence of hDus2, reminiscent of the recent results by Dai et al. showing that hDus3 is important for translational efficiency (22). This impact on translation has not been rationalized. In that respect, an action on mRNA is not excluded given the presence of D in mRNAs and their effect on the regulation of certain genes, including those involved in the cytoskeleton have been confirmed. Hence, the increase of dihydrouridylation activity in the transcriptome could also target mRNAs providing additional tricks used by cancer cells to stimulate protein translation and thereby their cellular metabolism. Another reasonable explanation that we propose here is that additional D sites could appear at non-canonical positions due to, for example, non-specific Dus activity. This nonspecific activity could occur during protein overexpression events, as is often the case in cancer cells (see below). On closer inspection, the sequence of the Dus4 transcript includes residues 1 to 237 of the protein, which corresponds to the majority of the TIM-barrel deleted from its last beta-strand based

on our hDus4 3D-model (**Figure 7**). In our opinion, this form of Dus4 is certainly non-functional especially in the absence of its HD domain. We can therefore legitimately think that Dus4-Bcpa29 is devoid of RNA dihydrouridylation function thus excluding in that specific case the involvement of Dus activity in the tumorigenic process. However, the dihydrouridylation activity of Dus in general and of Dus4 in particular could obviously be relevant in other types of cancer, in particular in lung cancer. Indeed, a recent study has proven the tumorigenic role of hDus4 in lung cancers (163). Overall, special attention should be given to a more detailed examination of the relevance of D in cancer biology in the future.

Conclusion

This is the first comprehensive review devoted to dihydrouridine in the transcriptome in which we have addressed all aspects related to its physicochemical and structural properties, its distribution in the transcriptome as well as its biosynthesis and functional and pathological implications. Although in our opinion we are just beginning to glimpse its potential biological role in protein translation, there are still many unknowns that will need to be addressed in the future. It is undeniable that this modified base carries a structural attribute by preventing and/or promoting the formation of certain RNA folds. The consequences are specific to the nature of the substrate since D seems to stabilize tRNAs whereas it destabilizes dsRNAs. This point needs to be further investigated, in particular by trying to apply it to specific mRNA sequences that have been identified as carrying this modification. Beyond its structuring role and its impact on translation, it is quite legitimate to wonder whether D could fulfill other potential functions as a local identity element for yet-to-be-discovered cellular partners as is already the case for other types of modifications. But to better appreciate the importance of this modified base, more precise mapping of D in the transcriptome wide is required, which will require breaking through common roadblocks imposed by large sequencing techniques. Although the recent Rho-seq technique is an obvious advance in the field since it does not rely on the generation of abasic sites but on a specific tagging of the D-base, there is still room to improve this

method. For example demethylation steps of the RNA sample by treatment with AlkB demethylases could be introduced to remove m⁶A, m¹A and m⁵C. Finally, we would like to emphasize that this review has highlighted the extent of what remains to be accomplished to fully understand the enzymology of D synthesis. Certainly, the enzymes of Dus have been identified and their site specificity determined, the full understanding of nucleotide level determinants beyond elbow integrity for tRNA substrates at least is yet to be determined, particularly for eukaryotic enzymes that target both tRNAs and mRNAs. Besides, the fact that mRNA-modifying enzymes usually also have tRNA as a substrate, as is the case for Dus, seriously complicates the interpretation of genetic targeting, encouraging us to elucidate the molecular basis behind the substrate discrimination of these enzymes. Likewise, the complex modularity of eukaryotic Dus enzymes needs to be better characterized and AlphaFold can help us considerably in that endeavor. Have these eukaryotic specific domains been acquired during evolution for functional purposes of recognition, substrate discrimination, transport, cellular localization, are questions begging for answers. These points will certainly be substantial to advance our understanding of the role of D and its enzymes in carcinogenic mechanisms. Finally, the abundance of D in tRNAs and more recently in mRNAs should not make us forget rRNAs, certainly the least understood dihydrouridylation substrates. The fact that *bona fide* Dus enzymes do not appear to be involved in D biosynthesis in rRNAs suggests that a fascinating new enzyme system for D biosynthesis exists in nature and is just waiting to be brought to light.

Author information

Corresponding Author

Email : djemel.hamdane@college-de-france.fr. Telephone : +33-(0)1-44271645.

Notes

The authors declare no competing financial interest.

Acknowledgments

The authors thank all past and present members of our laboratories who participated in the dihydrouridine project. This research is funded by ANR/DFG grant DERASE (#20-CE92-0030), LABEX DYNAMO”, ANR-11-LABX-0011, by grant GM132254 to V.d.C.L. M.H. was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) –TRR-319–TP A05, SPP1784, HE 3397/13-2, HE 3397/21-1.

References

1. Boccaletto, P., and Baginski, B. (2021) MODOMICS: An Operational Guide to the Use of the RNA Modification Pathways Database, *Methods Mol Biol* 2284, 481-505.
2. Spenkuch, F., Motorin, Y., and Helm, M. (2014) Pseudouridine: still mysterious, but never a fake (uridine)!, *RNA biology* 11, 1540-1554.
3. Frye, M., Jaffrey, S. R., Pan, T., Rechavi, G., and Suzuki, T. (2016) RNA modifications: what have we learned and where are we headed?, *Nature reviews. Genetics* 17, 365-372.
4. Jia, G., Fu, Y., and He, C. (2013) Reversible RNA adenosine methylation in biological regulation, *Trends in genetics : TIG* 29, 108-115.
5. Cozen, A. E., Quartley, E., Holmes, A. D., Hrabeta-Robinson, E., Phizicky, E. M., and Lowe, T. M. (2015) ARM-seq: AlkB-facilitated RNA methylation sequencing reveals a complex landscape of modified tRNA fragments, *Nature methods* 12, 879-884.
6. Ma, C. J., Ding, J. H., Ye, T. T., Yuan, B. F., and Feng, Y. Q. (2019) AlkB Homologue 1 Demethylates N(3)-Methylcytidine in mRNA of Mammals, *ACS chemical biology* 14, 1418-1425.
7. Hrabeta-Robinson, E., Marcus, E., Cozen, A. E., Phizicky, E. M., and Lowe, T. M. (2017) High-Throughput Small RNA Sequencing Enhanced by AlkB-Facilitated RNA de-Methylation (ARM-Seq), *Methods Mol Biol* 1562, 231-243.
8. Dai, Q., Zheng, G., Schwartz, M. H., Clark, W. C., and Pan, T. (2017) Selective Enzymatic Demethylation of N(2) ,N(2) -Dimethylguanosine in RNA and Its Application in High-Throughput tRNA Sequencing, *Angew Chem Int Ed Engl* 56, 5017-5020.
9. Roundtree, I. A., Evans, M. E., Pan, T., and He, C. (2017) Dynamic RNA Modifications in Gene Expression Regulation, *Cell* 169, 1187-1200.
10. Helm, M., and Motorin, Y. (2017) Detecting RNA modifications in the epitranscriptome: predict and validate, *Nature reviews. Genetics* 18, 275-291.
11. El Yacoubi, B., Bailly, M., and de Crecy-Lagard, V. (2012) Biosynthesis and function of posttranscriptional modifications of transfer RNAs, *Annual review of genetics* 46, 69-95.
12. Jackman, J. E., and Alfonzo, J. D. (2013) Transfer RNA modifications: nature's combinatorial chemistry playground, *Wiley interdisciplinary reviews. RNA* 4, 35-48.
13. Helm, M., and Alfonzo, J. D. (2014) Posttranscriptional RNA Modifications: playing metabolic games in a cell's chemical Legoland, *Chemistry & biology* 21, 174-185.
14. Lorenz, C., Lunse, C. E., and Morl, M. (2017) tRNA Modifications: Impact on Structure and Thermal Adaptation, *Biomolecules* 7.
15. Barraud, P., and Tisne, C. (2019) To be or not to be modified: Miscellaneous aspects influencing nucleotide modifications in tRNAs, *IUBMB life* 71, 1126-1140.

16. Motorin, Y., and Helm, M. (2022) RNA nucleotide methylation: 2021 update, *Wiley interdisciplinary reviews. RNA* 13, e1691.
17. Grosjean, H., de Crecy-Lagard, V., and Marck, C. (2010) Deciphering synonymous codons in the three domains of life: co-evolution with specific tRNA modification enzymes, *FEBS letters* 584, 252-264.
18. Pan, T. (2013) N6-methyl-adenosine modification in messenger and long non-coding RNA, *Trends in biochemical sciences* 38, 204-209.
19. Liu, J., and Jia, G. (2014) Methylation modifications in eukaryotic messenger RNA, *Journal of genetics and genomics = Yi chuan xue bao* 41, 21-33.
20. Zhao, B. S., Roundtree, I. A., and He, C. (2017) Post-transcriptional gene regulation by mRNA modifications, *Nature reviews. Molecular cell biology* 18, 31-42.
21. Franco, M. K., and Koutmou, K. S. (2022) Chemical modifications to mRNA nucleobases impact translation elongation and termination, *Biophysical chemistry* 285, 106780.
22. Huang, H., Weng, H., Sun, W., Qin, X., Shi, H., Wu, H., Zhao, B. S., Mesquita, A., Liu, C., Yuan, C. L., et al. (2018) Recognition of RNA N(6)-methyladenosine by IGF2BP proteins enhances mRNA stability and translation, *Nature cell biology* 20, 285-295.
23. Dai, W., Li, A., Yu, N. J., Nguyen, T., Leach, R. W., Wuhr, M., and Kleiner, R. E. (2021) Activity-based RNA-modifying enzyme probing reveals DUS3L-mediated dihydrouridylation, *Nature chemical biology* 17, 1178-1187.
24. Finet, O., Yague-Sanz, C., Kruger, L. K., Tran, P., Migeot, V., Louski, M., Nevers, A., Rougemaille, M., Sun, J., Ernst, F. G. M., et al. (2022) Transcription-wide mapping of dihydrouridine reveals that mRNA dihydrouridylation is required for meiotic chromosome segregation, *Molecular cell* 82, 404-419 e409.
25. House, C. H., and Miller, S. L. (1996) Hydrolysis of dihydrouridine and related compounds, *Biochemistry* 35, 315-320.
26. Dalluge, J. J., Hamamoto, T., Horikoshi, K., Morita, R. Y., Stetter, K. O., and McCloskey, J. A. (1997) Posttranscriptional modification of tRNA in psychrophilic bacteria, *Journal of bacteriology* 179, 1918-1923.
27. Noon, K. R., Guymon, R., Crain, P. F., McCloskey, J. A., Thomm, M., Lim, J., and Cavicchioli, R. (2003) Influence of temperature on tRNA modification in archaea: Methanococcus burtonii (optimum growth temperature [Topt], 23 degrees C) and Stetteria hydrogenophila (Topt, 95 degrees C), *Journal of bacteriology* 185, 5483-5490.
28. Best, A. N. (1978) Composition and Characterization of tRNA from Methanococcus vanniellii, *Journal of bacteriology* 133, 240-250.
29. Holley, R. W., Apgar, J., Everett, G. A., Madison, J. T., Marquisee, M., Merrill, S. H., Penswick, J. R., and Zamir, A. (1965) Structure of a Ribonucleic Acid, *Science* 147, 1462-1465.
30. Huang, R. C., and Bonner, J. (1965) Histone-bound RNA, a component of native nucleohistone, *Proceedings of the National Academy of Sciences of the United States of America* 54, 960-967.
31. Sundaralingam, M., Rao, S. T., and Abola, J. (1971) Molecular conformation of dihydrouridine: puckered base nucleoside of transfer RNA, *Science* 172, 725-727.
32. Kim, S. H., Suddath, F. L., Quigley, G. J., McPherson, A., Sussman, J. L., Wang, A. H., Seeman, N. C., and Rich, A. (1974) Three-dimensional tertiary structure of yeast phenylalanine transfer RNA, *Science* 185, 435-440.
33. Suddath, F. L., Quigley, G. J., McPherson, A., Sneden, D., Kim, J. J., Kim, S. H., and Rich, A. (1974) Three-dimensional structure of yeast phenylalanine transfer RNA at 3.0 angstroms resolution, *Nature* 248, 20-24.
34. Westhof, E., Dumas, P., and Moras, D. (1985) Crystallographic refinement of yeast aspartic acid transfer RNA, *Journal of molecular biology* 184, 119-145.
35. Basavappa, R., and Sigler, P. B. (1991) The 3 A crystal structure of yeast initiator tRNA: functional implications in initiator/elongator discrimination, *The EMBO journal* 10, 3105-3111.

36. Dalluge, J. J., Hashizume, T., Sopchik, A. E., McCloskey, J. A., and Davis, D. R. (1996) Conformational flexibility in RNA: the role of dihydrouridine, *Nucleic acids research* *24*, 1073-1079.
37. Dyubankova, N., Sochacka, E., Kraszewska, K., Nawrot, B., Herdewijn, P., and Lescrinier, E. (2015) Contribution of dihydrouridine in folding of the D-arm in tRNA, *Organic & biomolecular chemistry* *13*, 4960-4966.
38. Davis, D. R., Griffey, R. H., Yamaizumi, Z., Nishimura, S., and Poulter, C. D. (1986) 15N-labeled tRNA. Identification of dihydrouridine in Escherichia coli tRNA^{fMet}, tRNA^{Lys}, and tRNA^{Phe} by 1H-15N two-dimensional NMR, *The Journal of biological chemistry* *261*, 3584-3587.
39. Kawai, G., Yamamoto, Y., Kamimura, T., Masegi, T., Sekine, M., Hata, T., Imori, T., Watanabe, T., Miyazawa, T., and Yokoyama, S. (1992) Conformational rigidity of specific pyrimidine residues in tRNA arises from posttranscriptional modifications that enhance steric interaction between the base and the 2'-hydroxyl group, *Biochemistry* *31*, 1040-1046.
40. Shigi, N., Suzuki, T., Tamakoshi, M., Oshima, T., and Watanabe, K. (2002) Conserved bases in the TΨC loop of tRNA are determinants for thermophile-specific 2-thiouridylation at position 54, *The Journal of biological chemistry* *277*, 39128-39135.
41. Davis, D. R. (1995) Stabilization of RNA stacking by pseudouridine, *Nucleic acids research* *23*, 5020-5026.
42. Nomura, Y., Ohno, S., Nishikawa, K., and Yokogawa, T. (2016) Correlation between the stability of tRNA tertiary structure and the catalytic efficiency of a tRNA-modifying enzyme, archaeal tRNA-guanine transglycosylase, *Genes Cells* *21*, 41-52.
43. Sipa, K., Sochacka, E., Kazmierczak-Baranska, J., Maszewska, M., Janicka, M., Nowak, G., and Nawrot, B. (2007) Effect of base modifications on structure, thermodynamic stability, and gene silencing activity of short interfering RNA, *RNA* *13*, 1301-1316.
44. Grosjean, H., Keith, G., and Droogmans, L. (2004) Detection and quantification of modified nucleotides in RNA using thin-layer chromatography, *Methods Mol Biol* *265*, 357-391.
45. Hossain, M., and Limbach, P. A. (2007) Mass spectrometry-based detection of transfer RNAs by their signature endonuclease digestion products, *RNA* *13*, 295-303.
46. Su, D., Chan, C. T. Y., Gu, C., Lim, K. S., Chionh, Y. H., McBee, M. E., Russell, B. S., Babu, I. R., Begley, T. J., and Dedon, P. C. (2014) Quantitative analysis of ribonucleoside modifications in tRNA by HPLC-coupled mass spectrometry, *Nat Protoc* *9*, 828-841.
47. Solivio, B., Yu, N. X., Addepalli, B., and Limbach, P. A. (2018) Improving RNA modification mapping sequence coverage by LC-MS through a nonspecific RNase U2-E49A mutant, *Analytica chimica acta* *1036*, 73-79.
48. Douthwaite, S., and Kirpekar, F. (2007) Identifying modifications in RNA by MALDI mass spectrometry, *Methods in enzymology* *425*, 3-20.
49. Hamdane, D., Guerineau, V., Un, S., and Golinelli-Pimpaneau, B. (2011) A catalytic intermediate and several flavin redox states stabilized by folate-dependent tRNA methyltransferase from Bacillus subtilis, *Biochemistry* *50*, 5208-5219.
50. Sirand-Pugnet, P., Bregeon, D., Beven, L., Goyenvalle, C., Blanchard, A., Rose, S., Grosjean, H., Douthwaite, S., Hamdane, D., and Crecy-Lagard, V. (2020) Reductive Evolution and Diversification of C5-Uracil Methylation in the Nucleic Acids of Mollicutes, *Biomolecules* *10*.
51. Hamdane, D., Guelorget, A., Guerineau, V., and Golinelli-Pimpaneau, B. (2014) Dynamics of RNA modification by a multi-site-specific tRNA methyltransferase, *Nucleic acids research* *42*, 11697-11706.
52. Bou-Nader, C., Montemont, H., Guerineau, V., Jean-Jean, O., Bregeon, D., and Hamdane, D. (2018) Unveiling structural and functional divergences of bacterial tRNA dihydrouridine synthases: perspectives on the evolution scenario, *Nucleic acids research* *46*, 1386-1394.
53. Faivre, B., Lombard, M., Fakroun, S., Vo, C. D., Goyenvalle, C., Guerineau, V., Pecqueur, L., Fontecave, M., De Crecy-Lagard, V., Bregeon, D., et al. (2021) Dihydrouridine synthesis in tRNAs is under reductive evolution in Mollicutes, *RNA biology* *18*, 2278-2289.

54. Magrath, D. I., and Shaw, D. C. (1967) The occurrence and source of beta-alanine in alkaline hydrolysates of sRNA: a sensitive method for the detection and assay of 5,6-dihydrouracil residues in RNA, *Biochemical and biophysical research communications* 26, 32-37.
55. Molinaro, M., Sheiner, L. B., Neelon, F. A., and Cantoni, G. L. (1968) Effect of chemical modification of dihydrouridine in yeast transfer ribonucleic acid on amino acid acceptor activity and ribosomal binding, *The Journal of biological chemistry* 243, 1277-1282.
56. Xing, F., Hiley, S. L., Hughes, T. R., and Phizicky, E. M. (2004) The specificities of four yeast dihydrouridine synthases for cytoplasmic tRNAs, *The Journal of biological chemistry* 279, 17850-17860.
57. Cerutti, P., and Miller, N. (1967) Selective reduction of yeast transfer ribonucleic acid with sodium borohydride, *Journal of molecular biology* 26, 55-66.
58. Wintermeyer, W., and Zachau, H. G. (1974) Replacement of odd bases in tRNA by fluorescent dyes, *Methods in enzymology* 29, 667-673.
59. Wintermeyer, W., and Zachau, H. G. (1971) Replacement of Y base, dihydrouracil, and 7-methylguanine in tRNA by artificial odd bases, *FEBS letters* 18, 214-218.
60. Yang, C. H., and Soll, D. (1974) Studies of transfer RNA tertiary structure of singlet-singlet energy transfer, *Proceedings of the National Academy of Sciences of the United States of America* 71, 2838-2842.
61. Betteridge, T., Liu, H., Gamper, H., Kirillov, S., Cooperman, B. S., and Hou, Y. M. (2007) Fluorescent labeling of tRNAs for dynamics experiments, *RNA* 13, 1594-1601.
62. Marchand, V., Ayadi, L., Ernst, F. G. M., Hertler, J., Bourguignon-Igel, V., Galvanin, A., Kotter, A., Helm, M., Lafontaine, D. L. J., and Motorin, Y. (2018) AlkAniline-Seq: Profiling of m(7) G and m(3) C RNA Modifications at Single Nucleotide Resolution, *Angew Chem Int Ed Engl* 57, 16785-16790.
63. Kaur, J., Raj, M., and Cooperman, B. S. (2011) Fluorescent labeling of tRNA dihydrouridine residues: Mechanism and distribution, *RNA* 17, 1393-1400.
64. Suzuki, T. (2014) A complete landscape of post-transcriptional modifications in mammalian mitochondrial tRNAs, *Nucleic acids research* 42, 7346-7357.
65. Bohnsack, M. T., and Sloan, K. E. (2018) The mitochondrial epitranscriptome: the roles of RNA modifications in mitochondrial translation and human disease, *Cellular and molecular life sciences : CMLS* 75, 241-260.
66. Krog, J. S., Espanol, Y., Giessing, A. M., Dziergowska, A., Malkiewicz, A., Ribas de Pouplana, L., and Kirpekar, F. (2011) 3-(3-amino-3-carboxypropyl)-5,6-dihydrouridine is one of two novel post-transcriptional modifications in tRNA^{Lys}(UUU) from *Trypanosoma brucei*, *The FEBS journal* 278, 4782-4796.
67. Esberg, B., Leung, H. C., Tsui, H. C., Bjork, G. R., and Winkler, M. E. (1999) Identification of the miaB gene, involved in methylthiolation of isopentenylated A37 derivatives in the tRNA of *Salmonella typhimurium* and *Escherichia coli*, *Journal of bacteriology* 181, 7256-7265.
68. Pierrel, F., Douki, T., Fontecave, M., and Atta, M. (2004) MiaB protein is a bifunctional radical-S-adenosylmethionine enzyme involved in thiolation and methylation of tRNA, *Journal of Biological Chemistry* 279, 47555-47563.
69. Gerber, A., Grosjean, H., Melcher, T., and Keller, W. (1998) Tad1p, a yeast tRNA-specific adenosine deaminase, is related to the mammalian pre-mRNA editing enzymes ADAR1 and ADAR2, *Embo Journal* 17, 4780-4789.
70. Maas, S., Gerber, A. P., and Rich, A. (1999) Identification and characterization of a human tRNA-specific adenosine deaminase related to the ADAR family of pre-mRNA editing enzymes, *Proceedings of the National Academy of Sciences of the United States of America* 96, 8895-8900.
71. Brule, H., Elliott, M., Redlak, M., Zehner, Z. E., and Holmes, W. M. (2004) Isolation and characterization of the human tRNA-(N1G37) methyltransferase (TRM5) and comparison to the *Escherichia coli* TrmD protein, *Biochemistry* 43, 9243-9255.

72. Macbeth, M. R., Schubert, H. L., VanDemark, A. P., Lingam, A. T., Hill, C. P., and Bass, B. L. (2005) Inositol hexakisphosphate is bound in the ADAR2 core and required for RNA editing, *Science* 309, 1534-1539.
73. Grosjean, H., Constantinesco, F., Foiret, D., and Benachenhou, N. (1995) A novel enzymatic pathway leading to 1-methylinosine modification in *Haloferax volcanii* tRNA, *Nucleic acids research* 23, 4312-4319.
74. Rubio, M. A. T., Gaston, K. W., McKenney, K. M., Fleming, I. M. C., Paris, Z., Limbach, P. A., and Alfonzo, J. D. (2017) Editing and methylation at a single site by functionally interdependent activities, *Nature* 542, 494-+.
75. Han, L., and Phizicky, E. M. (2018) A rationale for tRNA modification circuits in the anticodon loop, *RNA* 24, 1277-1284.
76. Armengod, M. E., Moukadiri, I., Prado, S., Ruiz-Partida, R., Benitez-Paez, A., Villarroya, M., Lomas, R., Garzon, M. J., Martinez-Zamora, A., Meseguer, S., et al. (2012) Enzymology of tRNA modification in the bacterial MnmEG pathway, *Biochimie* 94, 1510-1520.
77. Kowalak, J. A., Bruenger, E., and McCloskey, J. A. (1995) Posttranscriptional modification of the central loop of domain V in *Escherichia coli* 23 S ribosomal RNA, *The Journal of biological chemistry* 270, 17758-17764.
78. Kirpekar, F., Hansen, L. H., Mundus, J., Tryggedsson, S., Teixeira Dos Santos, P., Ntokou, E., and Vester, B. (2018) Mapping of ribosomal 23S ribosomal RNA modifications in *Clostridium sporogenes*, *RNA biology* 15, 1060-1070.
79. Emmerechts, G., Barbe, S., Herdewijn, P., Anne, J., and Rozenski, J. (2007) Post-transcriptional modification mapping in the *Clostridium acetobutylicum* 16S rRNA by mass spectrometry and reverse transcriptase assays, *Nucleic acids research* 35, 3494-3503.
80. Carlile, T. M., Rojas-Duran, M. F., Zinshteyn, B., Shin, H., Bartoli, K. M., and Gilbert, W. V. (2014) Pseudouridine profiling reveals regulated mRNA pseudouridylation in yeast and human cells, *Nature* 515, 143-146.
81. Schwartz, S., Bernstein, D. A., Mumbach, M. R., Jovanovic, M., Herbst, R. H., Leon-Ricardo, B. X., Engreitz, J. M., Guttman, M., Satija, R., Lander, E. S., et al. (2014) Transcriptome-wide mapping reveals widespread dynamic-regulated pseudouridylation of ncRNA and mRNA, *Cell* 159, 148-162.
82. Dominissini, D., Nachtergaele, S., Moshitch-Moshkovitz, S., Peer, E., Kol, N., Ben-Haim, M. S., Dai, Q., Di Segni, A., Salmon-Divon, M., Clark, W. C., et al. (2016) The dynamic N(1)-methyladenosine methylome in eukaryotic messenger RNA, *Nature* 530, 441-446.
83. Arango, D., Sturgill, D., Alhusaini, N., Dillman, A. A., Sweet, T. J., Hanson, G., Hosogane, M., Sinclair, W. R., Nanan, K. K., Mandler, M. D., et al. (2018) Acetylation of Cytidine in mRNA Promotes Translation Efficiency, *Cell* 175, 1872-1886 e1824.
84. Tardu, M., Jones, J. D., Kennedy, R. T., Lin, Q., and Koutmou, K. S. (2019) Identification and Quantification of Modified Nucleosides in *Saccharomyces cerevisiae* mRNAs, *ACS chemical biology* 14, 1403-1409.
85. Jia, G., Fu, Y., Zhao, X., Dai, Q., Zheng, G., Yang, Y., Yi, C., Lindahl, T., Pan, T., Yang, Y. G., et al. (2011) N6-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO, *Nature chemical biology* 7, 885-887.
86. Dominissini, D., Moshitch-Moshkovitz, S., Schwartz, S., Salmon-Divon, M., Ungar, L., Osenberg, S., Cesarkas, K., Jacob-Hirsch, J., Amariglio, N., Kupiec, M., et al. (2012) Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq, *Nature* 485, 201-206.
87. Meyer, K. D., Saletore, Y., Zumbo, P., Elemento, O., Mason, C. E., and Jaffrey, S. R. (2012) Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons, *Cell* 149, 1635-1646.
88. Li, X. Y., Zhu, P., Ma, S. Q., Song, J. H., Bai, J. Y., Sun, F. F., and Yi, C. Q. (2015) Chemical pulldown reveals dynamic pseudouridylation of the mammalian transcriptome, *Nature chemical biology* 11, 592-U593.

89. Dominissini, D., Nachtergaele, S., Moshitch-Moshkovitz, S., Peer, E., Kol, N., Ben-Haim, M. S., Dai, Q., Di Segni, A., Salmon-Divon, M., Clark, W. C., et al. (2016) The dynamic N-1-methyladenosine methylome in eukaryotic messenger RNA, *Nature* 530, 441-+.
90. Li, X. Y., Xiong, X. S., Wang, K., Wang, L. X., Shu, X. T., Ma, S. Q., and Yi, C. Q. (2016) Transcriptome-wide mapping reveals reversible and dynamic N-1-methyladenosine methylome, *Nature chemical biology* 12, 311-+.
91. Li, X., Xiong, X., Zhang, M., Wang, K., Chen, Y., Zhou, J., Mao, Y., Lv, J., Yi, D., Chen, X. W., et al. (2017) Base-Resolution Mapping Reveals Distinct m(1)A Methylome in Nuclear- and Mitochondrial-Encoded Transcripts, *Molecular cell* 68, 993-1005 e1009.
92. Wang, X., Lu, Z., Gomez, A., Hon, G. C., Yue, Y., Han, D., Fu, Y., Parisien, M., Dai, Q., Jia, G., et al. (2014) N6-methyladenosine-dependent regulation of messenger RNA stability, *Nature* 505, 117-120.
93. Safra, M., Sas-Chen, A., Nir, R., Winkler, R., Nachshon, A., Bar-Yaacov, D., Erlacher, M., Rossmanith, W., Stern-Ginossar, N., and Schwartz, S. (2017) The m1A landscape on cytosolic and mitochondrial mRNA at single-base resolution, *Nature* 551, 251-255.
94. Lombard, M., and Hamdane, D. (2017) Flavin-dependent epitranscriptomic world, *Archives of biochemistry and biophysics* 632, 28-40.
95. Kasprzak, J. M., Czerwoniec, A., and Bujnicki, J. M. (2012) Molecular evolution of dihydrouridine synthases, *BMC bioinformatics* 13, 153.
96. Jiang, H. Q., Motorin, Y., Jin, Y. X., and Grosjean, H. (1997) Pleiotropic effects of intron removal on base modification pattern of yeast tRNA^{Phe}: an in vitro study, *Nucleic acids research* 25, 2694-2701.
97. Xing, F., Martzen, M. R., and Phizicky, E. M. (2002) A conserved family of *Saccharomyces cerevisiae* synthases effects dihydrouridine modification of tRNA, *RNA* 8, 370-381.
98. Bishop, A. C., Xu, J., Johnson, R. C., Schimmel, P., and de Crecy-Lagard, V. (2002) Identification of the tRNA-dihydrouridine synthase family, *The Journal of biological chemistry* 277, 25090-25095.
99. Byrne, R. T., Jenkins, H. T., Peters, D. T., Whelan, F., Stowell, J., Aziz, N., Kasatsky, P., Rodnina, M. V., Koonin, E. V., Konevega, A. L., et al. (2015) Major reorientation of tRNA substrates defines specificity of dihydrouridine synthases, *Proceedings of the National Academy of Sciences of the United States of America* 112, 6033-6037.
100. Yu, F., Tanaka, Y., Yamashita, K., Suzuki, T., Nakamura, A., Hirano, N., Yao, M., and Tanaka, I. (2011) Molecular basis of dihydrouridine formation on tRNA, *Proceedings of the National Academy of Sciences of the United States of America* 108, 19593-19598.
101. Kusuba, H., Yoshida, T., Iwasaki, E., Awai, T., Kazayama, A., Hirata, A., Tomikawa, C., Yamagami, R., and Hori, H. (2015) In vitro dihydrouridine formation by tRNA dihydrouridine synthase from *Thermus thermophilus*, an extreme-thermophilic eubacterium, *Journal of biochemistry* 158, 513-521.
102. Chen, X., Sim, S., Wurtmann, E. J., Feke, A., and Wolin, S. L. (2014) Bacterial noncoding Y RNAs are widespread and mimic tRNAs, *RNA* 20, 1715-1724.
103. Bou-Nader, C., Pecqueur, L., Bregeon, D., Kamah, A., Guerineau, V., Golinelli-Pimpaneau, B., Guimaraes, B. G., Fontecave, M., and Hamdane, D. (2015) An extended dsRBD is required for post-transcriptional modification in human tRNAs, *Nucleic acids research* 43, 9446-9456.
104. Brzezicha, B., Schmidt, M., Makalowska, I., Jarmolowski, A., Pienkowska, J., and Szweykowska-Kulinska, Z. (2006) Identification of human tRNA:m5C methyltransferase catalysing intron-dependent m5C formation in the first position of the anticodon of the pre-tRNA Leu (CAA), *Nucleic acids research* 34, 6034-6043.
105. Squires, J. E., Patel, H. R., Nousch, M., Sibbritt, T., Humphreys, D. T., Parker, B. J., Suter, C. M., and Preiss, T. (2012) Widespread occurrence of 5-methylcytosine in human coding and non-coding RNA, *Nucleic acids research* 40, 5023-5033.
106. Auxilien, S., Guerineau, V., Szweykowska-Kulinska, Z., and Golinelli-Pimpaneau, B. (2012) The Human tRNA m(5)C methyltransferase Misu is multisite-specific, *RNA biology* 9, 1331-1338.

107. Huang, T., Chen, W. Y., Liu, J. H., Gu, N. N., and Zhang, R. (2019) Genome-wide identification of mRNA 5-methylcytosine in mammals, *Nature structural & molecular biology* 26, 380-+.
108. Motorin, Y., and Grosjean, H. (1999) Multisite-specific tRNA:m⁵C-methyltransferase (Trm4) in yeast *Saccharomyces cerevisiae*: identification of the gene and substrate specificity of the enzyme, *RNA* 5, 1105-1118.
109. Motorin, Y., Keith, G., Simon, C., Foiret, D., Simos, G., Hurt, E., and Grosjean, H. (1998) The yeast tRNA:pseudouridine synthase Pus1p displays a multisite substrate specificity, *RNA* 4, 856-869.
110. Massenet, S., Motorin, Y., Lafontaine, D. L., Hurt, E. C., Grosjean, H., and Branlant, C. (1999) Pseudouridine mapping in the *Saccharomyces cerevisiae* spliceosomal U small nuclear RNAs (snRNAs) reveals that pseudouridine synthase pus1p exhibits a dual substrate specificity for U2 snRNA and tRNA, *Molecular and cellular biology* 19, 2142-2154.
111. Becker, H. F., Motorin, Y., Planta, R. J., and Grosjean, H. (1997) The yeast gene YNL292w encodes a pseudouridine synthase (Pus4) catalyzing the formation of psi55 in both mitochondrial and cytoplasmic tRNAs, *Nucleic acids research* 25, 4493-4499.
112. Lovejoy, A. F., Riordan, D. P., and Brown, P. O. (2014) Transcriptome-wide mapping of pseudouridines: pseudouridine synthases modify specific mRNAs in *S. cerevisiae*, *PLoS one* 9, e110799.
113. Begik, O., Lucas, M. C., Prysycz, L. P., Ramirez, J. M., Medina, R., Milenkovic, I., Cruciani, S., Liu, H., Vieira, H. G. S., Sas-Chen, A., et al. (2021) Quantitative profiling of pseudouridylation dynamics in native RNAs with nanopore sequencing, *Nature biotechnology* 39, 1278-1291.
114. Urban, A., Behm-Ansmant, I., Branlant, C., and Motorin, Y. (2009) RNA sequence and two-dimensional structure features required for efficient substrate modification by the *Saccharomyces cerevisiae* RNA:{Psi}-synthase Pus7p, *The Journal of biological chemistry* 284, 5845-5858.
115. Decatur, W. A., and Schnare, M. N. (2008) Different mechanisms for pseudouridine formation in yeast 5S and 5.8S rRNAs, *Molecular and cellular biology* 28, 3089-3100.
116. Safra, M., Nir, R., Farouq, D., Vainberg Slutskin, I., and Schwartz, S. (2017) TRUB1 is the predominant pseudouridine synthase acting on mammalian mRNA via a predictable and conserved code, *Genome research* 27, 393-406.
117. Lecointe, F., Simos, G., Sauer, A., Hurt, E. C., Motorin, Y., and Grosjean, H. (1998) Characterization of yeast protein Deg1 as pseudouridine synthase (Pus3) catalyzing the formation of psi 38 and psi 39 in tRNA anticodon loop, *The Journal of biological chemistry* 273, 1316-1323.
118. Hopper, A. K., Furukawa, A. H., Pham, H. D., and Martin, N. C. (1982) Defects in modification of cytoplasmic and mitochondrial transfer RNAs are caused by single nuclear mutations, *Cell* 28, 543-550.
119. Pagliarini, D. J., Calvo, S. E., Chang, B., Sheth, S. A., Vafai, S. B., Ong, S. E., Walford, G. A., Sugiana, C., Boneh, A., Chen, W. K., et al. (2008) A mitochondrial protein compendium elucidates complex I disease biology, *Cell* 134, 112-123.
120. Chen, M., Yu, J., Tanaka, Y., Tanaka, M., Tanaka, I., and Yao, M. (2013) Structure of dihydrouridine synthase C (DusC) from *Escherichia coli*, *Acta crystallographica. Section F, Structural biology and crystallization communications* 69, 834-838.
121. Whelan, F., Jenkins, H. T., Griffiths, S. C., Byrne, R. T., Dodson, E. J., and Antson, A. A. (2015) From bacterial to human dihydrouridine synthase: automated structure determination, *Acta crystallographica. Section D, Biological crystallography* 71, 1564-1571.
122. Bou-Nader, C., Barraud, P., Pecqueur, L., Perez, J., Velours, C., Shepard, W., Fontecave, M., Tisne, C., and Hamdane, D. (2019) Molecular basis for transfer RNA recognition by the double-stranded RNA-binding domain of human dihydrouridine synthase 2, *Nucleic acids research* 47, 3117-3126.

123. Bou-Nader, C., Pecqueur, L., Barraud, P., Fontecave, M., Tisne, C., Sacquin-Mora, S., and Hamdane, D. (2019) Conformational Stability Adaptation of a Double-Stranded RNA-Binding Domain to Transfer RNA Ligand, *Biochemistry* 58, 2463-2473.
124. Tian, B., Bevilacqua, P. C., Diegelman-Parente, A., and Mathews, M. B. (2004) The double-stranded-RNA-binding motif: interference and much more, *Nature reviews. Molecular cell biology* 5, 1013-1023.
125. Masliah, G., Barraud, P., and Allain, F. H. (2013) RNA recognition by double-stranded RNA binding domains: a matter of shape and sequence, *Cellular and molecular life sciences : CMLS* 70, 1875-1895.
126. Barraud, P., Emmerth, S., Shimada, Y., Hotz, H. R., Allain, F. H., and Buhler, M. (2011) An extended dsRBD with a novel zinc-binding motif mediates nuclear retention of fission yeast Dicer, *The EMBO journal* 30, 4223-4235.
127. Barraud, P., Banerjee, S., Mohamed, W. I., Jantsch, M. F., and Allain, F. H. (2014) A bimodular nuclear localization signal assembled via an extended double-stranded RNA-binding domain acts as an RNA-sensing signal for transportin 1, *Proceedings of the National Academy of Sciences of the United States of America* 111, E1852-1861.
128. Stefl, R., Oberstrass, F. C., Hood, J. L., Jourdan, M., Zimmermann, M., Skrisovska, L., Maris, C., Peng, L., Hofr, C., Emeson, R. B., et al. (2010) The solution structure of the ADAR2 dsRBM-RNA complex reveals a sequence-specific readout of the minor groove, *Cell* 143, 225-237.
129. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Zidek, A., Potapenko, A., et al. (2021) Highly accurate protein structure prediction with AlphaFold, *Nature* 596, 583-589.
130. Cramer, P. (2021) AlphaFold2 and the future of structural biology, *Nature structural & molecular biology* 28, 704-705.
131. Varadi, M., Anyango, S., Deshpande, M., Nair, S., Natassia, C., Yordanova, G., Yuan, D., Stroe, O., Wood, G., Laydon, A., et al. (2022) AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models, *Nucleic acids research* 50, D439-D444.
132. Tunyasuvunakool, K., Adler, J., Wu, Z., Green, T., Zielinski, M., Zidek, A., Bridgland, A., Cowie, A., Meyer, C., Laydon, A., et al. (2021) Highly accurate protein structure prediction for the human proteome, *Nature* 596, 590-596.
133. Rider, L. W., Ottosen, M. B., Gattis, S. G., and Palfey, B. A. (2009) Mechanism of dihydrouridine synthase 2 from yeast and the importance of modifications for efficient tRNA reduction, *The Journal of biological chemistry* 284, 10324-10333.
134. Savage, D. F., de Crecy-Lagard, V., and Bishop, A. C. (2006) Molecular determinants of dihydrouridine synthase activity, *FEBS letters* 580, 5198-5202.
135. Dobritzsch, D., Schneider, G., Schnackerz, K. D., and Lindqvist, Y. (2001) Crystal structure of dihydropyrimidine dehydrogenase, a major determinant of the pharmacokinetics of the anti-cancer drug 5-fluorouracil, *The EMBO journal* 20, 650-660.
136. Dobritzsch, D., Ricagno, S., Schneider, G., Schnackerz, K. D., and Lindqvist, Y. (2002) Crystal structure of the productive ternary complex of dihydropyrimidine dehydrogenase with NADPH and 5-iodouracil. Implications for mechanism of inhibition and electron transfer, *The Journal of biological chemistry* 277, 13155-13166.
137. Barraud, P., Gato, A., Heiss, M., Catala, M., Kellner, S., and Tisne, C. (2019) Time-resolved NMR monitoring of tRNA maturation, *Nature communications* 10, 3373.
138. Cavaille, J., Chetouani, F., and Bachellerie, J. P. (1999) The yeast *Saccharomyces cerevisiae* YDL112w ORF encodes the putative 2'-O-ribose methyltransferase catalyzing the formation of Gm18 in tRNAs, *RNA* 5, 66-81.
139. Kealey, J. T., and Santi, D. V. (1991) Identification of the catalytic nucleophile of tRNA (m5U54)methyltransferase, *Biochemistry* 30, 9724-9728.

140. Liu, Y., and Santi, D. V. (2000) m⁵C RNA and m⁵C DNA methyl transferases use different cysteine residues as catalysts, *Proceedings of the National Academy of Sciences of the United States of America* 97, 8263-8265.
141. Lee, T. T., Agarwalla, S., and Stroud, R. M. (2005) A unique RNA Fold in the RumA-RNA-cofactor ternary complex contributes to substrate selectivity and enzymatic function, *Cell* 120, 599-611.
142. Hamdane, D., Argentini, M., Cornu, D., Myllykallio, H., Skouloubris, S., Hui-Bon-Hoa, G., and Golinelli-Pimpaneau, B. (2011) Insights into folate/FAD-dependent tRNA methyltransferase mechanism: role of two highly conserved cysteines in catalysis, *The Journal of biological chemistry* 286, 36268-36280.
143. O'Connor, M., Lee, W. M., Mankad, A., Squires, C. L., and Dahlberg, A. E. (2001) Mutagenesis of the peptidyltransferase center of 23S rRNA: the invariant U2449 is dispensable, *Nucleic acids research* 29, 710-715.
144. Alexandrov, A., Chernyakov, I., Gu, W., Hiley, S. L., Hughes, T. R., Grayhack, E. J., and Phizicky, E. M. (2006) Rapid tRNA decay can result from lack of nonessential modifications, *Molecular cell* 21, 87-96.
145. Boriack-Sjodin, P. A., Ribich, S., and Copeland, R. A. (2018) RNA-modifying proteins as anticancer drug targets, *Nature reviews. Drug discovery* 17, 435-453.
146. Delaunay, S., and Frye, M. (2019) RNA modifications regulating cell fate in cancer, *Nature cell biology* 21, 552-559.
147. Barbieri, I., and Kouzarides, T. (2020) Role of RNA modifications in cancer, *Nature reviews. Cancer* 20, 303-322.
148. Haruehanroengra, P., Zheng, Y. Y., Zhou, Y., Huang, Y., and Sheng, J. (2020) RNA modifications and cancer, *RNA biology* 17, 1560-1575.
149. Chujo, T., and Tomizawa, K. (2021) Human transfer RNA modopathies: diseases caused by aberrations in transfer RNA modifications, *The FEBS journal* 288, 7096-7122.
150. Zhou, J. B., Wang, E. D., and Zhou, X. L. (2021) Modifications of the human tRNA anticodon loop and their associations with genetic diseases, *Cellular and Molecular Life Sciences* 78, 7087-7105.
151. Suzuki, T. (2021) The expanding world of tRNA modifications and their disease relevance, *Nat Rev Mol Cell Bio* 22, 375-392.
152. Kuchino, Y., and Borek, E. (1978) Tumour-specific phenylalanine tRNA contains two supernumerary methylated bases, *Nature* 271, 126-129.
153. Dalluge, J. J., Hashizume, T., and McCloskey, J. A. (1996) Quantitative measurement of dihydrouridine in RNA using isotope dilution liquid chromatography mass spectrometry (LC/MS), *Nucleic acids research* 24, 3242-3245.
154. Kato, T., Daigo, Y., Hayama, S., Ishikawa, N., Yamabuki, T., Ito, T., Miyamoto, M., Kondo, S., and Nakamura, Y. (2005) A novel human tRNA-dihydrouridine synthase involved in pulmonary carcinogenesis, *Cancer research* 65, 5638-5646.
155. Mittelstadt, M., Frump, A., Khuu, T., Fowlkes, V., Handy, I., Patel, C. V., and Patel, R. C. (2008) Interaction of human tRNA-dihydrouridine synthase-2 with interferon-induced protein kinase PKR, *Nucleic acids research* 36, 998-1008.
156. Lee, Y. S., and Kunkeaw, N. (2020) Protein kinase R and its cellular regulators in cancer: An active player or a surveillant?, *Wiley interdisciplinary reviews. RNA* 11, e1558.
157. Bou-Nader, C., Gordon, J. M., Henderson, F. E., and Zhang, J. (2019) The search for a PKR code-differential regulation of protein kinase R activity by diverse RNA and protein regulators, *RNA* 25, 539-556.
158. Hull, C. M., and Bevilacqua, P. C. (2016) Discriminating Self and Non-Self by RNA: Roles for RNA Structure, Misfolding, and Modification in Regulating the Innate Immune Sensor PKR, *Accounts of chemical research* 49, 1242-1249.

159. Kim, H. P., Cho, G. A., Han, S. W., Shin, J. Y., Jeong, E. G., Song, S. H., Lee, W. C., Lee, K. H., Bang, D., Seo, J. S., et al. (2014) Novel fusion transcripts in human gastric cancer revealed by transcriptome analysis, *Oncogene* 33, 5434-5441.
160. Kim, K. M., Adachi, T., Nielsen, P. J., Terashima, M., Lamers, M. C., Kohler, G., and Reth, M. (1994) Two new proteins preferentially associated with membrane immunoglobulin D, *The EMBO journal* 13, 3793-3800.
161. Tang, Y., Qin, F., Liu, A., and Li, H. (2017) Recurrent fusion RNA DUS4L-BCAP29 in non-cancer human tissues and cells, *Oncotarget* 8, 31415-31423.
162. Tang, Y., Guan, F., and Li, H. (2020) Case Study: The Recurrent Fusion RNA DUS4L-BCAP29 in Noncancer Human Tissues and Cells, *Methods Mol Biol* 2079, 243-258.
163. Li, Z., Yin, C., Li, B., Yu, Q. Y., Mao, W. J., Li, J., Lin, J. P., Meng, Y. Q., Feng, H. M., and Jing, T. (2020) DUS4L Silencing Suppresses Cell Proliferation and Promotes Apoptosis in Human Lung Adenocarcinoma Cell Line A549, *Cancer management and research* 12, 9905-9913.

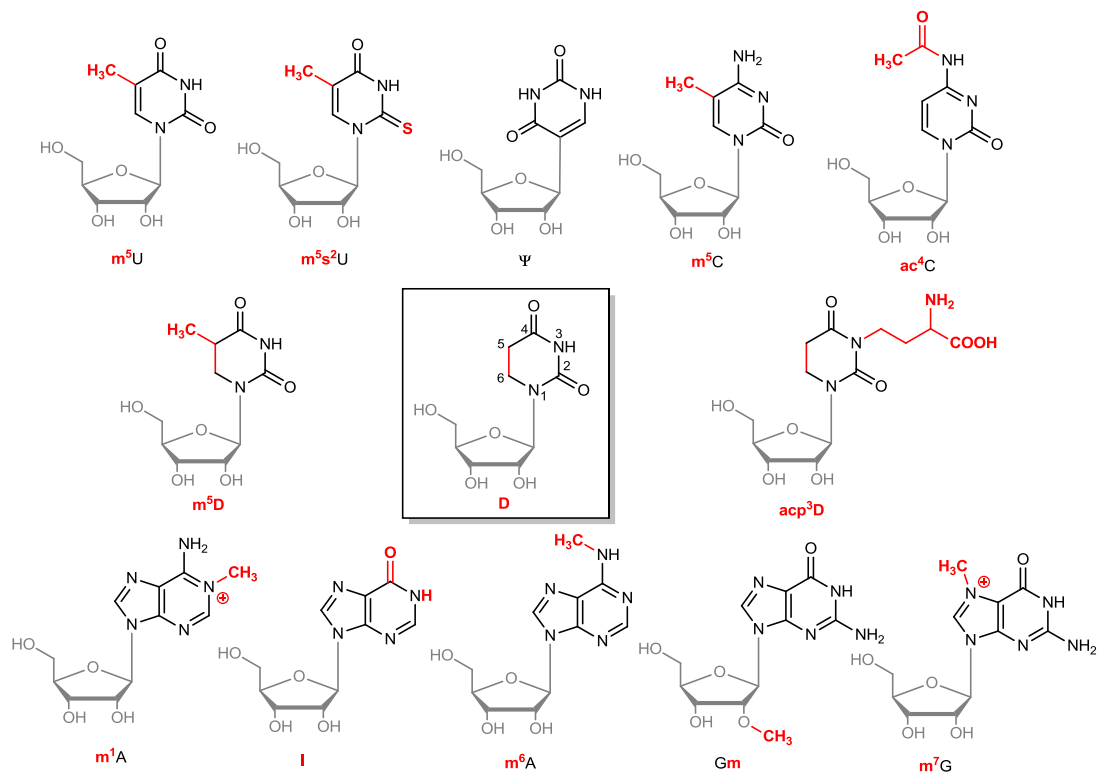


Figure 1: Chemical structure of some modified nucleosides found in different types of RNAs. The modification made to the canonical base or ribose is indicated in red. The chemical structure of dihydrouridine (D) is boxed and the atom numbering of the base is also indicated. The acronym of the modifications is indicated under the corresponding modified nucleoside.

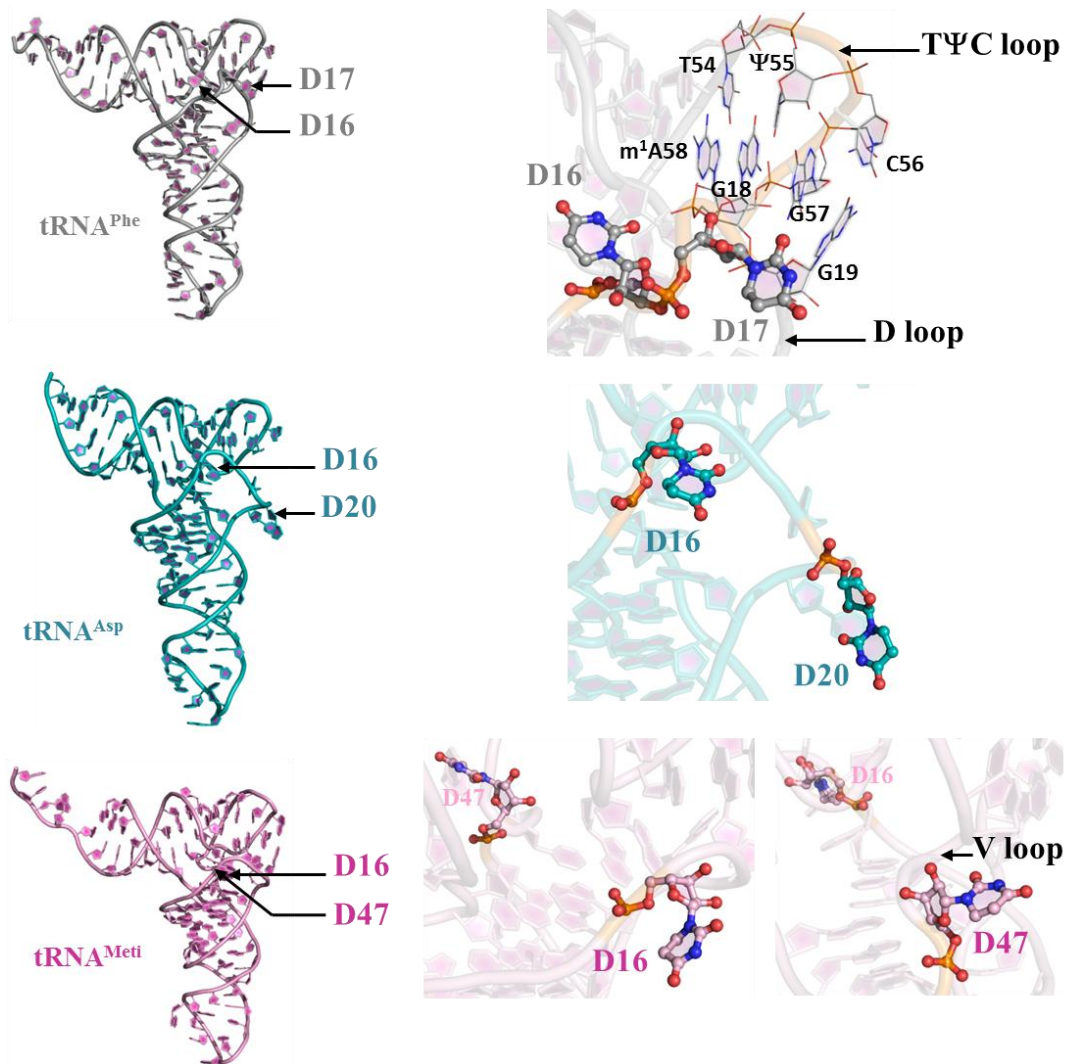


Figure 2: X-ray structures of three matured yeast tRNAs. The full-length tRNA structures Phe, Asp, and Met initiator are shown on the left side in gray, cyan, and pink colored cartoons, respectively. The pdb codes used are 1EVV, 3TRA and 1YFG for tRNA^{Phe}, tRNA^{Asp} and tRNA^{Met}, respectively. On the right side, the D-residues found in each of the tRNAs are represented in stick in its corresponding color. Nucleotide involved in tertiary interactions that are highly conserved in cytosolic tRNAs, including the Hoogsteen-reverse base pair T54-A58, the interloop base pairs G18-Ψ55, G19-C56, and the stack of four mutually interspersed purine bases, A58-G18-A/G57-G19 are shown in tRNA^{Phe}.

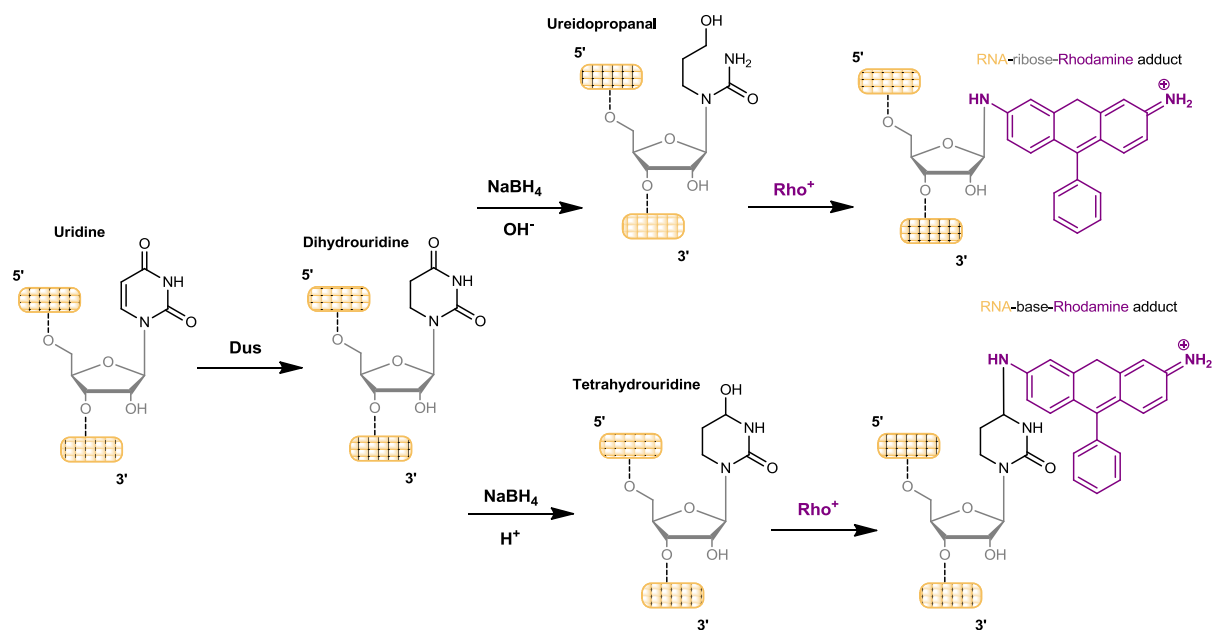


Figure 3: Strategy of chemical labeling of D sites with rhodamine. Reduction of D by NaBH_4 in basic condition leads to the opening of the pyrimidine ring and to a labeling of the ribose by elimination of the ureidopropanal moiety. On the other hand, reduction under acidic conditions produces the tetrahydrouridine which is directly labeled by replacement of its C4-hydroxyl group.

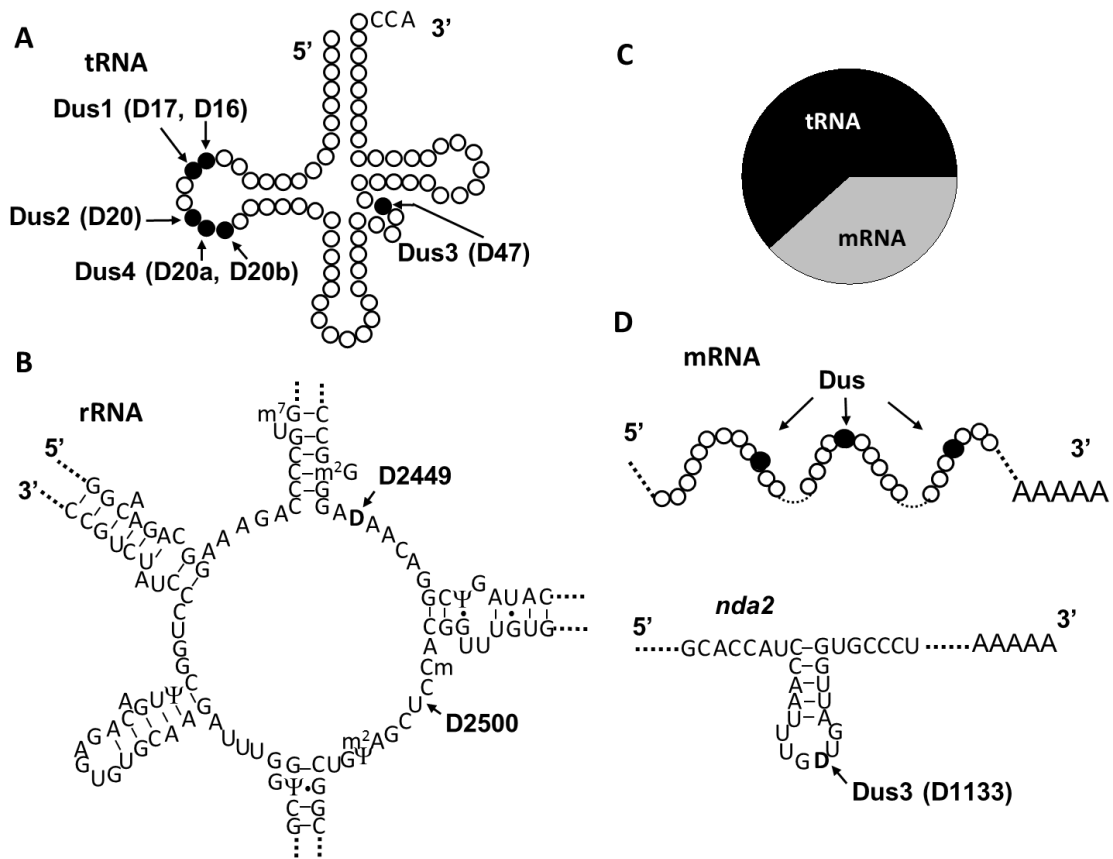


Figure 4: Distribution of D in the transcriptome. (A) Secondary tRNA structure showing the position of D residues observed in *S. cerevisiae* tRNAs. The Dus enzymes responsible for biosynthesis are also shown next to the D residues in parentheses. (B) *E. coli* 23S rRNA sequence showing the position of the unique dihydrouridine (D2449). The 23S subunit of bacterial rRNAs located at the peptidyltransferase center of the ribosome. Uridine 2500 is an unmodified uridine in *E. coli* whereas in *C. sporogenes* this uridine is modified to dihydrouridine, D2500. (C) Pie chart showing the distribution of D in the *S. pombe* transcriptome within tRNAs and mRNAs. (D) Putative Dus-catalyzed dihydrouridylated mRNAs carrying either 3 D-sites as *ala1* and *fhn1* or one as *nda2*. In the case of *nda2*, the sequence part of the mRNA shows the putative stem-loop structure containing the D1133 catalyzed by the enzyme Dus3.

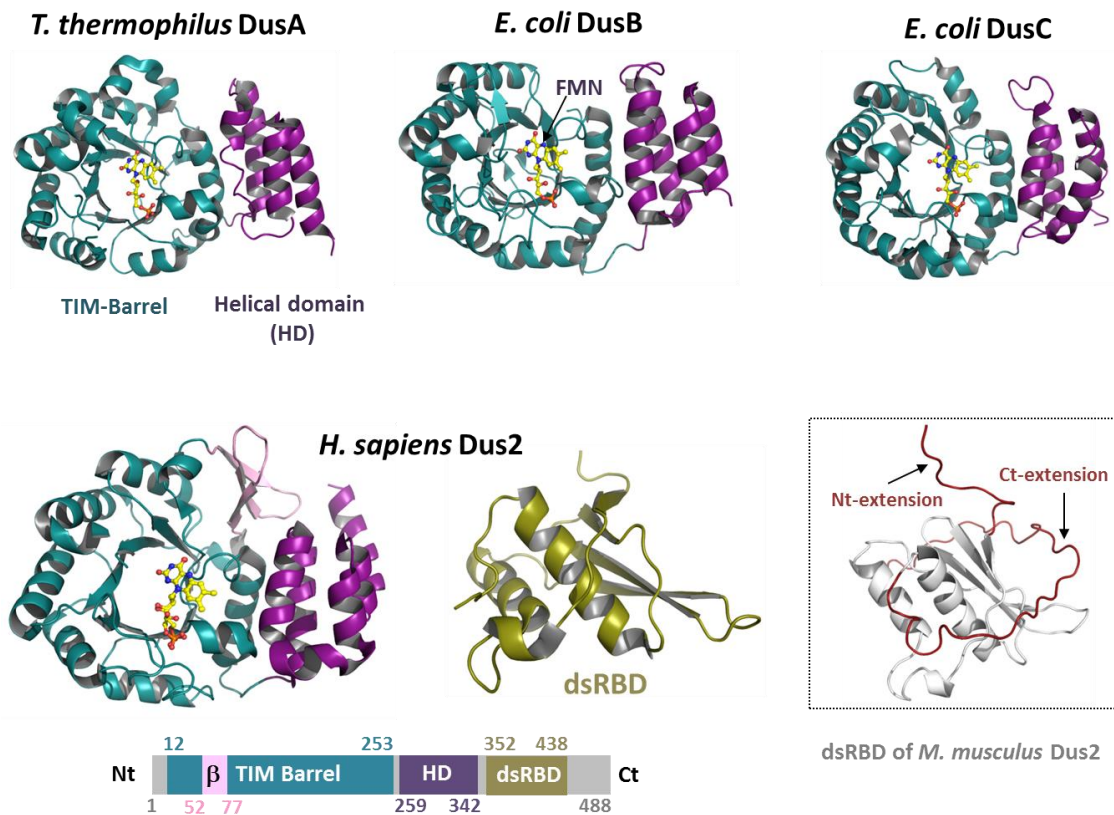


Figure 5: Crystallographic structures of bacterial and human Dus. The TIM-barrel, HD, and dsRBD domains are colored in deep-teal, purple, and olive, respectively. The prosthetic group, FMN, is in yellow stick. In pink is the inserted domain found in hDus2. Inset is the putative mouse Dus2 dsRBD whose structure was resolved by NMR and annotated in the PDB as dsRBD from hypothetical protein BAB26260. This solution structure shows extensions in N-terminal (Nt) and C-terminal (Ct) regions. The pdb codes used are 3B0P, 6EI9, 4BFA, 4WFS, 4WFT and 1WFN for *T. thermophilus* DusA, *E. coli* DusB, *E. coli* DusC, the canonical domains of hDus2, the dsRBD of hDus2, the dsRBD of BAB26260 protein, respectively. Below the crystallographic structure of hDus2 is shown the diagram of the modular organization of this enzyme as well as the delimitation of the respective domains.

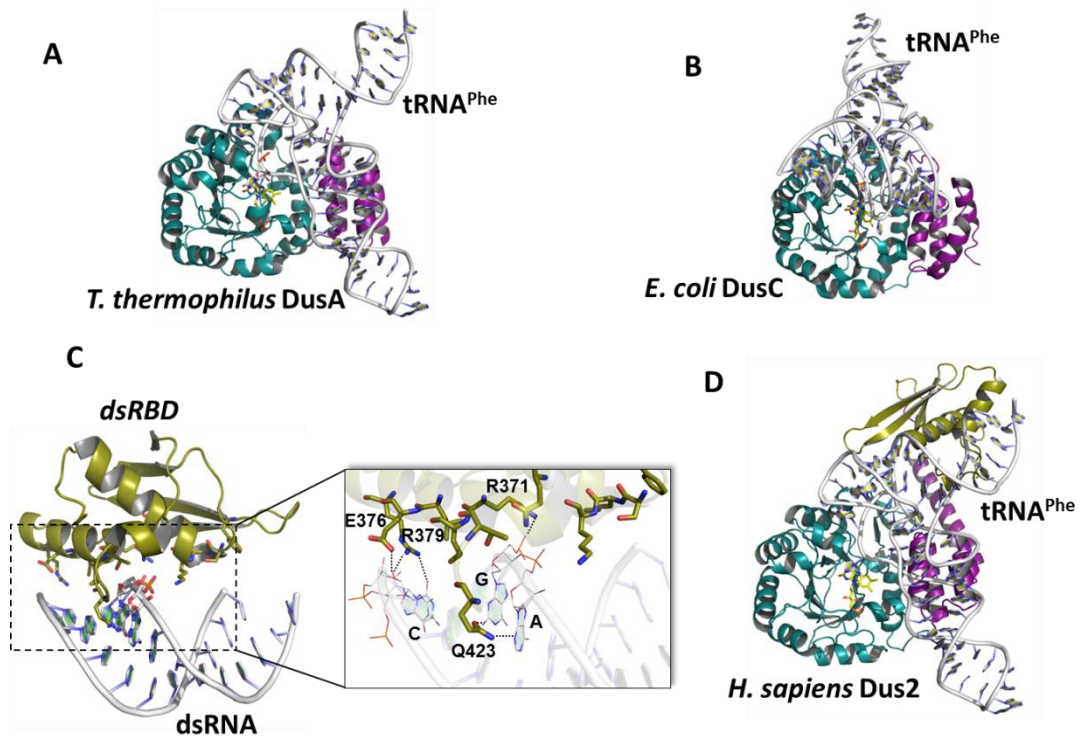


Figure 6: Molecular basis of tRNA substrate recognition by *T. thermophilus* DusA, *E. coli* DusC and human Dus2. The protein/RNA complexes, namely *T. thermophilus* DusA:tRNA^{Phe} (A), *E. coli* DusC:tRNA^{Phe} (B) and hDus2 dsRBD:dsRNA (C) are crystallographic structures whose PDB codes are 3B0V, 4YCO and 5OC6, respectively. (D) Molecular model of the hDus2:tRNA complex generated as described in Bou-nader *et al.* (122).

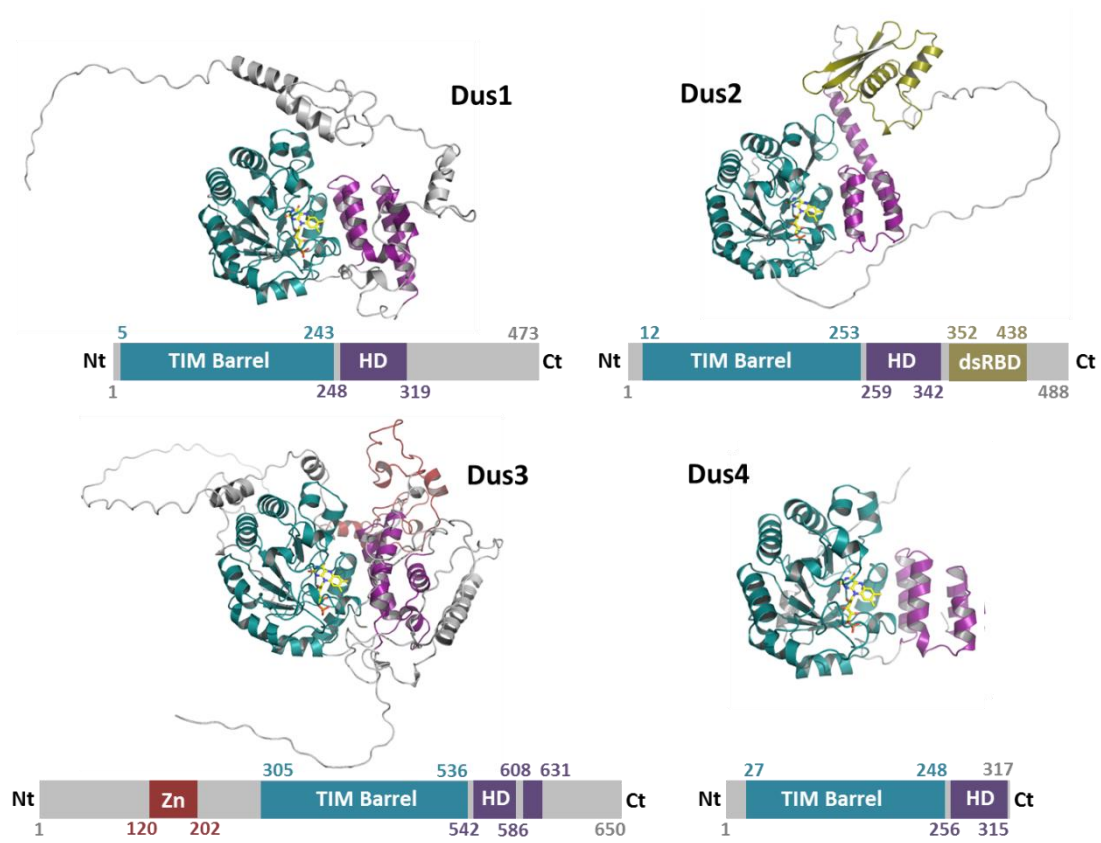


Figure 7: 3D-models of human Dus generated by Alpha-fold. The architectural organization and the delimitation of the domains are indicated under each corresponding Dus model.

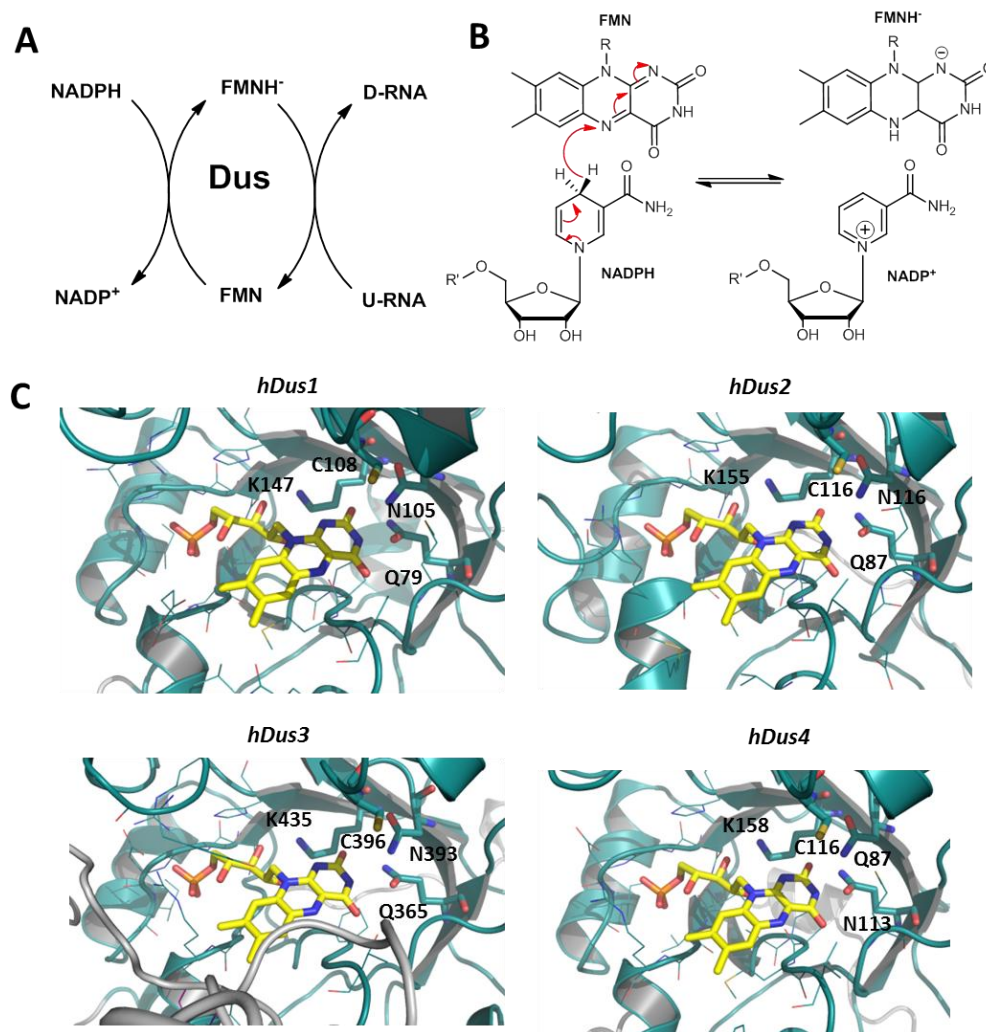


Figure 8: Chemical mechanism of D synthesis catalyzed by Dus. (A) General reaction scheme of the Dus catalytic cycle. (B) Proposed stereo-chemical mechanism of flavin reduction of Dus by NADPH. (C) Sections of the active sites of human Dus obtained from Alpha fold protein models. The FMN coenzyme is represented in yellow ball-sticks

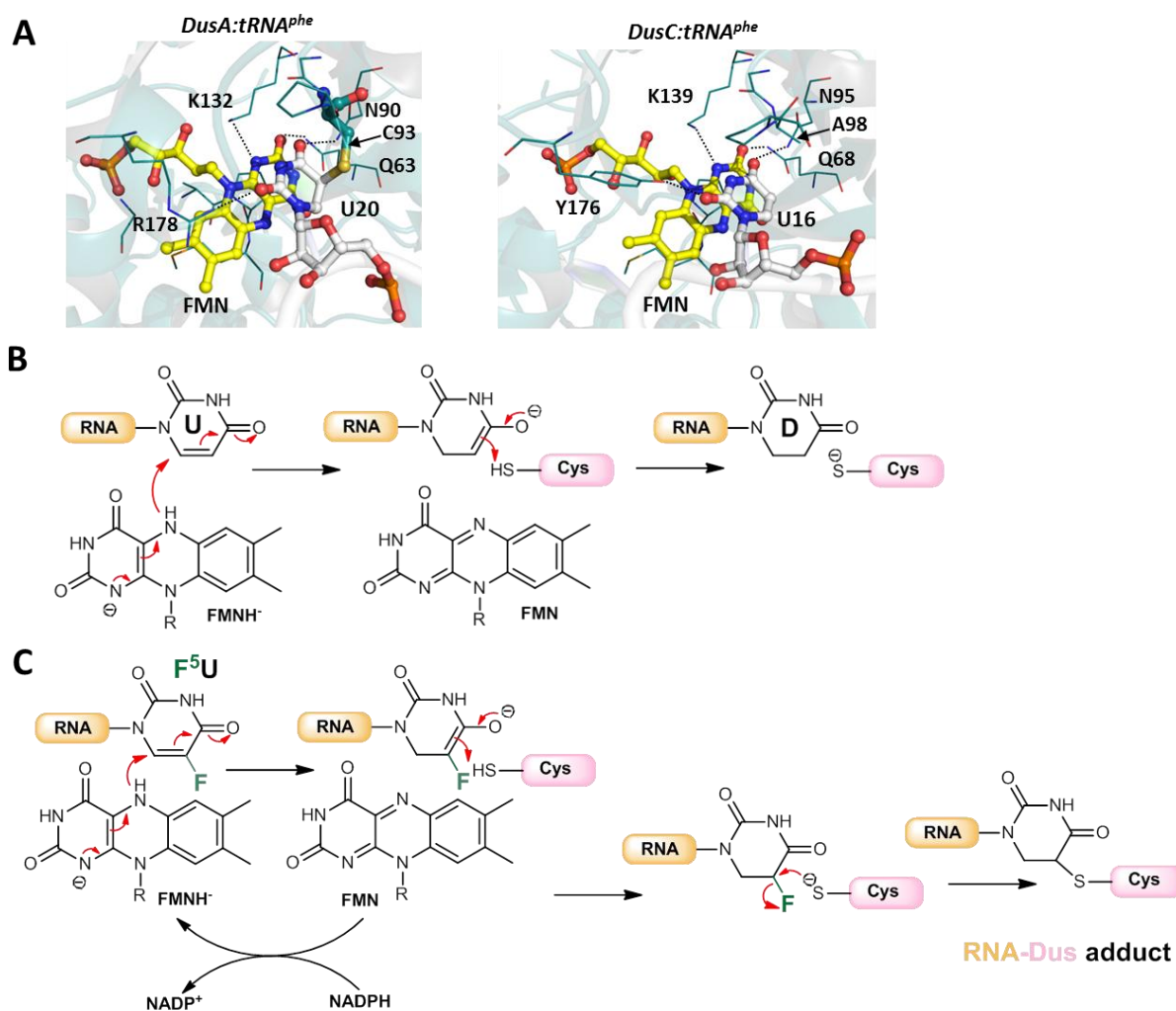


Figure 9: Proposed chemical mechanism of D biosynthesis and activity-based inhibition of Dus. (A) Focus on the active site of *T. thermophilus* DusA and *E. coli* DusC in complex with their respective tRNA substrates. The FMN is represented in yellow ball-sticks while the uridines are in white ball-sticks. (B) Postulated chemical mechanism of Dus enzymes. (C) Postulated chemical mechanism for the activity-based inhibition of Dus via formation of a covalent Dus/RNA complex in the presence of fluorinated uridine.

TOC

