



HAL
open science

Exploring a unique class of flavoenzymes: Identification and biochemical characterization of ribosomal RNA dihydrouridine synthase

Sabrina Toubdji, Quentin Thullier, Lea-Marie Kilz, V. Marchand, Yifeng Yuan, Claudia Sudol, Catherine Goyenvalle, Olivier Jean-Jean, Simon Rose, Stephen Douthwaite, et al.

► To cite this version:

Sabrina Toubdji, Quentin Thullier, Lea-Marie Kilz, V. Marchand, Yifeng Yuan, et al.. Exploring a unique class of flavoenzymes: Identification and biochemical characterization of ribosomal RNA dihydrouridine synthase. *Proceedings of the National Academy of Sciences of the United States of America*, 2024, 121 (32), pp.e2401981121. 10.1073/pnas.2401981121 . hal-04734945

HAL Id: hal-04734945

<https://hal.science/hal-04734945v1>

Submitted on 15 Oct 2024

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.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

1 **Exploring a Novel Class of Flavoenzymes:**
2 **Identification and Biochemical Characterization of Ribosomal RNA**
3 **Dihydrouridine Synthase**

4
5 Sabrine Toubdji^{1,2}, Quentin Thullier^{3,4}, Lea-Marie Kilz⁵, Virginie Marchand^{3,4}, Yifeng Yuan⁶,
6 Claudia Sudol^{1,2}, Catherine Goyenvalle¹, Olivier Jean-Jean¹, Simon Rose⁷, Stephen
7 Douthwaite⁷, Léo Hardy⁸, Zeynep Baharoglu⁸, Valérie de Crécy-Lagard^{6,9}, Mark Helm⁵, Yuri
8 Motorin^{3,4}, Djemel Hamdane*² and Damien Brégeon*¹
9

10 ¹ Sorbonne Université, CNRS, Institut de Biologie Paris Seine, Biology of Aging and
11 Adaptation, IBPS-B2A, F-75252 Paris cedex 05, France

12 ² Collège De France, Sorbonne Université, CNRS, Laboratoire de Chimie des Processus
13 Biologiques, LCPB, F-75231 Paris Cedex 05, France.

14 ³ Université de Lorraine, CNRS, INSERM, IBSLor, EpiRNA-Seq Core Facility, F-54000
15 Nancy, France.

16 ⁴ Université de Lorraine, CNRS, IMoPA, F-54000 Nancy, France.

17 ⁵ Institut für Pharmazeutische und Biomedizinische Wissenschaften (IPBW), Johannes
18 Gutenberg-Universität, D-55128 Mainz, Germany

19 ⁶ University of Florida, Department of Microbiology and Cell Science, FL-32611Gainesville,
20 Florida, USA.

21 ⁷ University of Southern Denmark, Department of Biochemistry and Molecular Biology, DK-
22 5230 Odense M, Denmark.

23 ⁸ Institut Pasteur, Université Paris Cité, CNRS UMR3525, Unité Plasticité du Génome
24 Bactérien, Paris, France

25 ⁹ University of Florida, Genetics Institute, Gainesville, FL-32610Gainesville, Florida, USA.

26
27 *To whom correspondence should be addressed:

28 Damien Brégeon, Sorbonne Université, CNRS UMR-8256, Institut de Biologie Paris Seine,
29 Biology of Aging and Adaptation, F-75252 Paris cedex 05, France.

30 Email : damien.bregeon@sorbonne-universite.fr
31

32 Djemel Hamdane, Laboratoire de Chimie des Processus Biologiques, CNRS-UMR 8229,
33 Collège de France, 11 place Marcelin Berthelot, F-75231 Paris Cedex 05, France.

34 Email : djemel.hamdane@college-de-france.fr
35
36
37
38
39

1 **Abstract (226 words)**

2 Dihydrouridine (D), a prevalent and evolutionarily conserved base in the transcriptome,
3 primarily resides in tRNAs and, to a lesser extent, in mRNAs. Notably, this modification is
4 found at position 2449 in the *E. coli* 23S rRNA, strategically positioned near the ribosome's
5 peptidyl transferase site. Despite the prior identification, in *E.coli* genome, of three
6 dihydrouridine synthases (DUS), a set of NADPH and FMN-dependent enzymes known for
7 introducing D in tRNAs and mRNAs, characterization of the enzyme responsible for D2449
8 deposition has remained elusive. This study introduces a rapid method for detecting D in
9 rRNA, involving reverse transcriptase-blockage at the rhodamine-labeled D2449 site,
10 followed by PCR amplification (RhoRT-PCR). Through analysis of rRNA from diverse *E. coli*
11 strains, harboring chromosomal or single-gene deletions, we pinpoint the *yhiN* gene as the
12 ribosomal dihydrouridine synthase, now designated as RdsA. Biochemical characterizations
13 uncovered RdsA as a novel class of flavoenzymes, dependent on FAD and NADH, with a
14 complex structural topology. *In vitro* assays demonstrated that RdsA dihydrouridylates a
15 short rRNA transcript mimicking the local structure of the peptidyl transferase site. This
16 suggests an early introduction of this modification before ribosome assembly. Phylogenetic
17 studies unveiled the widespread distribution of the *yhiN* gene in the bacterial kingdom,
18 emphasizing the conservation of rRNA dihydrouridylation. In a broader context, these
19 findings underscore nature's preference for utilizing reduced flavin in the reduction of uridines
20 and their derivatives.

21 **Significance (120 words)**

22 Post-transcriptional modifications of rRNA play a crucial role in ribosome biogenesis and
23 fine-tuning of protein synthesis. These are predominantly concentrated at decoding and
24 peptidyl transferase center (PTC) sites. To date, 35 of the 36 identified ribosomal
25 modifications have been characterized, although the biosynthetic pathway for U2449
26 dihydrouridylation within the PTC has remained elusive. To fill this gap in knowledge, we
27 developed a technique for monitoring dihydrourylation reaction of 23S rRNA. This method
28 involves chemical labeling of dihydrouridine and RT-PCR blockage. By integrating this
29 approach with epitranscriptomic, genetic, and biochemical methodologies, we successfully
30 identified a novel class of flavin-dependent enzymes involved in rRNA dihydrouridylation.
31 This discovery opens up new avenues for exploring the biological significance of PTC
32 dihydrouridylation in ribosomal function.

1 Introduction (632 words)

2 Post-transcriptional maturation processes shape RNA molecules into functional
3 biomolecules. Among these processes, the incorporation of chemical groups at the
4 nucleoside level, primarily within the bases, is catalyzed by modification enzymes. Over 170
5 modifications have been identified to date, with tRNAs and rRNAs being the most extensively
6 modified RNA species (1). The chemical diversity of these modifications can significantly
7 alter the physicochemical properties of nucleosides, exerting a profound impact on local and
8 even global RNA structures. For the majority of modifications, the effect is stabilizing in
9 nature (2, 3). For instance, the introduction of methyl groups tends to strengthen base
10 stacking. One of the most abundant modifications across the transcriptome is dihydrouridine
11 (D) (4, 5) (**Figure 1A**). This modified base is evolutionarily conserved and is predominantly
12 found in bacterial and eukaryotic tRNAs. Recently, D has also been detected in certain
13 eukaryotic mRNAs (6–8) and long non-coding RNAs (6, 7, 9). In the case of rRNAs, sporadic
14 observations of D have been reported, notably at position 2449 (D2449) in domain V of the
15 23S ribosomal RNA of *Escherichia coli*, in the peptidyl transferase center (10, 11) (**Figure 1B**
16 **& 1C**). D2449 has also been identified at positions 2449 and 2500 in *Clostridium sporogenes*
17 (with D2449 being partially methylated to m⁵D) (12) and at position 1211 or 1212 in the 16S
18 rRNA of *Clostridium acetobutylicum* (13). Comprehensive studies to investigate the presence
19 of dihydrouridine in the rRNAs of bacteria and eukaryotes remain to be made.

20 The study of D synthases involved in tRNA dihydrouridylation has revealed that the
21 conversion of U to D is a redox-dependent reaction relying on NADPH as a reducing agent
22 (4, 14–21) (**Figure 1A**). Dus enzymes are flavoenzymes, utilizing FMN as a coenzyme for
23 redox reactions. In the initial step, the flavin is reduced to an anionic hydroquinone (FMNH⁻),
24 which then transfers a hydride to the C5-C6 double bond of uridine, forming an enolate
25 intermediate (4, 22). This intermediate accepts a proton from the strictly conserved catalytic
26 cysteine, leading to the breakdown of D's aromaticity, resulting in a non-planar base and a
27 more flexible ribose conformation. The significance of D's base flexibility is highlighted by the
28 fact that psychrophilic organisms generally have more D in their tRNAs than thermophiles
29 (23). Structurally, D destabilizes small stem-loops, promoting the formation of larger loops
30 (24). The presence of D2449 in the central large loop of domain V in *E. coli*'s 23S rRNA
31 could potentially serve this purpose, although addressing this issue is challenging due to the
32 unidentified gene encoding ribosomal Dus. Beyond its presumed structural role, the
33 biological significance of D remains obscure. Recent discoveries of D residues in yeast and
34 human mRNAs, also introduced by tRNA dihydrouridine synthases, suggest potential roles in
35 meiosis and translational efficiency through yet-to-be-determined molecular mechanisms (6,
36 8). These studies indicate that Dus enzymes are promiscuous although, at least in the case
37 of *E. coli*, they do not act on rRNA.

38 The *E. coli* rRNA encompasses 36 modifications, and to date, all modification enzymes for *E.*
39 *coli* rRNA have been identified, except for the rRNA D synthase (1) (**Table S1**). In this study,
40 we identified this elusive enzyme, hereafter referred to as RdsA (Ribosomal Dihydrouridine
41 Synthase A), responsible for the biosynthesis of D2449 in 23S rRNA. Biochemical
42 characterizations of RdsA unveiled its novel identity as a flavin-dependent D synthase with
43 redox activity reliant on FAD and NADH. The proposed chemical mechanism of
44 dihydrouridylation, derived from our findings, outlines a likely hydride transfer mechanism,
45 underscoring nature's preference for utilizing flavin to reduce uridine or its derivatives.
46 Moreover, our phylogenetic data demonstrate that the RdsA gene is widely distributed
47 among bacteria, including many pathogens, while absent in eukaryotes. This observation
48 suggests that ribosomal D may be pervasive in microbial rRNAs, rendering it a potential
49 target for selective antibiotic intervention.

1 Results (2053 words)

2 Development of a Rapid D-ribosomal Detection Test

3 To facilitate the swift detection of D2449, we have developed a methodology termed RhoRT-
4 PCR. This innovative approach involves the blocking of reverse transcriptase (RT) through
5 dihydrouridine labelling with rhodamine 110, following a well-established mechanism (**Figure**
6 **S1**). Subsequently, this step was followed by PCR amplification of RT-products employing a
7 primer pair to target a specific region (**Figure 2A**). The primer pair 2F/3R was meticulously
8 designed to amplify the region downstream of position 2449 in the 23S rRNA, yielding a 248
9 bp fragment. It functions as a positive control for RT activity. Conversely, the region
10 encompassing position 2449 was amplified using primers 1F/3R, generating a 429 bp
11 product. Therefore, the absence of rhodamine labeling on D2449 or the lack of
12 dihydrouridylation at position 2449 (indicative of U2449) should result in the presence of both
13 248 and 429 bp fragments, discernible on an agarose gel. In contrast, the presence of D2449
14 would result solely in the 248 bp band. The validity of this approach was tested using purified
15 rRNA from wild-type strains of *E. coli* (known to contain D2449) and *Bacillus subtilis* (known
16 to harbor a uridine at the corresponding rRNA position). As depicted in **Figure 2B**, in the
17 case of *E. coli*, the absence of Rho labeling showed the anticipated two fragments. However,
18 pre-labeling D2449 with Rho resulted in the disappearance of the 429 bp band, strongly
19 indicating the presence of D2449. Conversely, in the case of *B. subtilis*, whether labeled or
20 not, both bands persisted, suggesting the absence of D in this region of rRNA. Recently, it
21 has been demonstrated that the three Dus enzymes in *E. coli* (DusA, B, and C) are not
22 responsible for the synthesis of D2449 (7). To confirm this result, we applied our RhoRT-
23 PCR method to purified rRNA from the *E. coli* triple mutant strain $\Delta dusABC$. As shown in
24 **Figure S2A**, the result revealed a single band at 248 bp, clearly demonstrating the presence
25 of D2449 in this mutant. This indicates that the classical Dus enzymes are not involved in the
26 dihydrouridylation of *E. coli* 23S rRNA. Thus, the RhoRT-PCR technique emerges as an
27 effective method for the precise detection of D2449.

28 Identification of the Gene Encoding rRNA Dihydrouridine Synthase

29 In the quest to pinpoint the gene responsible for dihydrouridine synthesis in rRNA, we
30 applied the RhoRT-PCR technique to rRNA extracted from the *E. coli* strain ME5125 (25,
31 26). This strain is distinguished by 15 major deletions in its chromosomal DNA, making it an
32 ideal candidate due to its ability to maintain the smallest genome among *E. coli* strains while
33 retaining sufficient viability for rRNA isolation. As illustrated in **Figure 3A**, the deletions
34 exhibited both fragments, whether treated with rhodamine or not, strongly suggesting the
35 probable absence of D. To substantiate these findings, we subjected BW25113 and ME5125
36 rRNA to MALDI mass spectrometry after RNase T1 treatment. The analysis revealed the
37 $^{2448}\text{ADAACAGp}^{2454}$ fragment with an m/z of 2293.3 in the case of BW25113, while in
38 ME5125, the corresponding fragment displayed an m/z of 2291.3, indicative of the non-
39 dihydrouridylated fragment (**Figure 3B**). Moreover, these results were validated by
40 AlkAniline-Seq, underscoring the presence of D2449 in WT rRNA and its conspicuous
41 absence in the ME5125 strain (**Figure 3C**). Individual testing of the 15 deletion strains using
42 RhoRT-PCR and AlkAniline-Seq showed that only the ME5079 strain, harboring a genomic
43 deletion encompassing 77 genes (**Figure S3**), lacked D2449 (**Figure 3A and 3C**). A similar
44 approach was used for the analysis of 10 strains bearing single-gene knockouts that were
45 annotated with unknown functions but potentially linked to redox processes or ribosome
46 biogenesis (**Table S2 and Figure S2**). Comprehensive testing across all these strains
47 unequivocally demonstrated that only the rRNA from the $\Delta yhiN$ strain lacked D2449 (**Figure**
48 **3C and S2A**). Notably, *yhiN* is annotated as a putative oxidoreductase with an FAD/NAD(P)H
49 domain and operates outside an operon (**Table S2**). However, it is situated in proximity to

1 two rRNA methyltransferase genes, namely *rsmJ* and *rlmJ*, and a putative ribosome-
2 associated ATPase gene (*rbbA*) (Table S2 & Figure S3).

3 The Gene *yhiN* Encodes for an FAD Binding Protein

4 The *E. coli yhiN* gene encodes a 400-amino acid protein with a theoretical mass of 43.7 kDa
5 (Figure S4). To elucidate the structural organization of YhiN, an AlphaFold model was
6 generated (Figure 4A). Remarkably, YhiN showcases a distinct fold compared to classical
7 Dus proteins (4, 15–17, 19, 20). A PDB search revealed homology between YhiN and
8 various flavoproteins, with crystallographic structures available for holoprotein forms (2GQF,
9 3V76, 4CNJ, 2I0Z), including *Streptococcus oligofermentis* L-aminoacetone oxidase.
10 Structural alignments of the YhiN model with homologous structures revealed RMSD values
11 below 2.3 Å (Figure S5), signifying significant structural conservation within this flavoprotein
12 class. YhiN comprises three distinct domains with an intricately complex topology (Figure
13 4A). The Rossmann fold domain (FBD), housing the flavin and putatively the catalytic site,
14 accommodates the insertion of the other two domains between G190 and A337, resulting in
15 the Nt and Ct ends being borne by this domain. The second domain (ins1D), homologous to
16 the Ct domain of *E. coli* EF-Tu, is interrupted at residue L257 by the insertion of the third
17 domain (ins2D), exhibiting structural homology with the Helix-2-turn-helix domain of the
18 topoisomerase VI-B subunit of *Sulfolobus shibatae*. Ins2D spans L258 to L322, succeeded
19 by a segment of ins1D and FBD. Structural alignment of YhiN with flavoprotein homologs
20 facilitated the identification of the conserved FAD binding site (Figure 4B & S5). No clashes
21 were observed between residues forming the FAD pocket and the coenzyme. The FAD
22 assumes an elongated conformation upheld by a series of interactions, with the
23 isoalloxazine, the redox center, situated within a solvent-accessible crevice. To assess the
24 flavoprotein nature of *E. coli* YhiN (RdsA), recombinant protein production was undertaken.
25 As depicted in Figure 4C and D, the UV/visible spectrum of freshly purified YhiN indicated
26 the presence of oxidized flavin in 61% of holoprotein. The addition of SDS induced
27 coenzyme released, yielding a UV/visible spectrum akin to free FAD. Additionally, mutating
28 key residues in this identified FAD binding pocket led to a lower ratio of holoproteins ranging
29 from 37%, when affecting residues interacting with the adenine moiety, to almost 0% when
30 affecting residues in the vicinity of the ribityl (Figure 4D). These findings conclusively
31 establish YhiN as a non-covalently bound FAD flavoprotein characterized by a sophisticated
32 structural topology.

33 The flavoprotein YhiN: Unraveling its Role as the rRNA Dihydrouridine Synthase

34 The conversion of uridine to dihydrouridine is a redox reaction requiring 2 electrons + 2H⁺ (or
35 hydride + H⁺)(Figure 1A) (4). In the case of classical Dus proteins, their RNA substrates are
36 reduced by FMNH⁻, produced through the prior reduction of FMN by NADPH (4, 22). Initially,
37 we investigated whether NADPH could serve as a reducer of flavin by monitoring the
38 spectrophotometric NADPH oxidase activity at 340 nm of recombinant YhiN. However, no
39 NADPH consumption was observed. Conversely, we observed efficient NADH oxidation with
40 a k_{cat} of 0.027 +/- 0.003 s⁻¹ and a K_M of 6.5 +/- 1.5 μM (Figure 4E). With a conclusive
41 reducing source identified, our focus shifted to detecting potential dihydrouridylation activity
42 of rRNA by YhiN. In our initial endeavors, we aimed to reconstitute *in vitro* activity using
43 rRNA isolated from an *E. coli* strain with the *yhiN* deletion. This involved co-incubating the
44 rRNA with recombinant YhiN and an excess of NADH. Subsequently, the generated rRNA
45 product was enzymatically digested to its nucleoside level, and the presence of D was
46 analyzed using LC/MS. Despite our concerted efforts, definitive detection of activity proved
47 challenging, primarily due to notable contamination of D most probably originating from tRNA
48 (Figure S6). To overcome this challenge, we designed and produced an *in vitro* transcribed
49 40-mer RNA (Mini-PTC-U), encompassing nucleotides 2438-2459 and mimicking stems and

1 loops of *E. coli*'s 23S rRNA (**Figure 4F**), serving as a refined substrate in subsequent activity
2 assays. Our results unequivocally established that YhiN dihydrouridylates this mini-RNA,
3 strategically mimicking a segment of the ribosomal peptidyl-transferase center (**Figure 4G**).
4 The D level increased with prolonged incubation time (**Figure 4G**) and, after 24 h of
5 incubation, 0.04% of D per quantified adenosine was formed (%mod/A). As the synthetic
6 RNA contains 10 adenosines, the number of D per molecule is then 0.4%. However, this
7 artificial substrate might adopt alternative structures prone to dihydrouridylation of other
8 uridines found in this Mini-PTC as a much lower, but detectable, activity was measured when
9 changing the target uridine by a cytidine (Mini-PTC-C in **Figure 4F** and **4G**). Nonetheless,
10 these findings conclusively position YhiN as the dihydrouridine synthase for *E. coli*'s
11 ribosome, catalyzing the reduction of U2449 to D2449, leveraging FAD as a redox coenzyme
12 and NADH as a source of reducing equivalents.

13 **Predicted RdsA are widespread in Bacteria and members of a non-isofunctional** 14 **superfamily**

15 The *E. coli* YhiN protein is a member of the COG2081 and PF03486 (HI0933-like protein)
16 families. An initial analysis of the family in the EggNog database
17 (<http://eggnog6.embl.de/search/ogs/COG2081/>) suggested it was not isofunctional. Indeed,
18 the 9,168 proteins listed in this family are from 6,723 species, hence 2,245 are paralogs
19 (**Table S3**). The only two experimentally characterized protein in this family are 3-dehydro-
20 bile acid delta(4,6)-reductase (BaiN) from *Clostridium scindens* involved in bile degradation
21 (27) and aminoacetone oxidase (AAO) from *Streptococcus cristatus*, involved in resistance to
22 endogenous metabolites-generated ROS and mutagens (28, 29). Hence, the RdsA activity of
23 YhiN must be shared by only a subset of the COG2081/PF03486 members. To identify the
24 RdsA subgroup more precisely, we constructed a Sequence Similarity Networks (SSN) of
25 PF03486 with different Alignment Score Thresholds (AST, see Supplementary methods). An
26 AST of 80 allowed the separation of the cluster containing YhiN_{Ec} from the cluster containing
27 bona-fide BaiN (indicated by a yellow line in **Figure 5A**) and the cluster containing AAO_{So}
28 (indicated by a green line in **Figure 5A**). Here, we analyzed the rRNA of *Vibrio cholerae*
29 species by AlkalineSeq and showed that its 23S rRNA indeed harbors D2432 while
30 carrying a RdsA (**Figure S7**). These bacteria do encode proteins highly similar to YhiNEc
31 (cyan nodes in **Figure 5A**), supporting their function as RdsA enzymes. In contrast, *Bacillus*
32 *subtilis* 168 rRNA does not show any D residues (1) but its genome does encode a PF03486
33 family protein (YtfP or BSU30060). This protein is not located in the same SSN cluster as
34 YhiNEc (**Figure 5A**) and the gene is in an operon with the *opuD* gene encoding the glycine
35 betaine/arsenobetaine transporter (**Figure 5A and S8A**). Most *Acinetobacter* species encode
36 two members of the PF03486 family (**Table S4**). These separate nicely into two groups in the
37 SSN (**Figure 5A**). Group 1 proteins (YhiN1) are part of the YhiNEc cluster and hence would
38 be predicted to catalyze the formation of D in rRNA. Group 2 proteins (YhiN2) are part of
39 separate cluster, but the corresponding genes are located downstream *trmB* genes encoding
40 tRNA (guanine(46)-N(7))-methyltransferase (**Figure 5B**) still indicating a possible link to RNA
41 modifications for this YhiN2 group. Other physical association with RNA modifications were
42 observed: *yhiN* homolog are located next to *dusA* genes in organisms such as
43 *Synechococcus elongatus* (yellow nodes in **Figure 5** and **S9**), or to genes encoding
44 ribosomal pseudouridine synthase RsuA-like proteins (orange nodes in **Figure 5** and **Figure**
45 **S8B**). In summary, the COG2081/PF03486 family is clearly not isofunctional and, because
46 gene neighborhood information does not give definitive functional clues, further experimental
47 characterization of the proteins from the different subgroups is required to establish whether
48 they harbor RdsA activity.

1 To survey the taxonomic distribution of the PF03486 proteins conservatively predicted to be
2 isofunctional with YhiNEc, we further analyzed the proteins within the same cluster (boxed in
3 **Figure 5A** and **Table S5**) in an SSN with a more stringent alignment score threshold (**Figure**
4 **S9**) and constructed a phylogenetic tree of the same protein set (**Figure S10**). These
5 analyses show that members of the YhiNEc subgroup are mostly found in bacteria with only
6 two eukaryotic homologs from dinoflagellates (Uniprot IDs : A0A812V8P8 and A0A9P1GL58)
7 and a small group found in Archaea (mainly Methanobacteriales and Methanomicrobiales,
8 **Table S5**). The protein tree diverges from the species tree in several places (for example,
9 the inclusion of Thermodesulfobacteriota in the middle of the gammaproteobacteria proteins)
10 suggesting horizontal gene transfer events. The gene neighborhood information does not link
11 YhiN to translation or RNA modifications as genes encoding universal stress protein,
12 transporters and RNA helicase-like proteins are mostly found surrounding these sets of YhiN
13 encoding genes (**Figure S10**).

14 **Discussion (555 words)**

15 In the course of this investigation, we uncovered a novel dihydrouridine synthase, referred to
16 as RdsA, responsible for catalyzing the NADH and FAD-dependent reduction of U2449 in the
17 23S rRNA of *E. coli*. Traditional Dus enzymes, encompassing bacterial DusA, B, C, and
18 eukaryotic Dus 1 to 4, also operate as flavoenzymes but rely on FMN and NADPH (4). This
19 parallels their counterparts, dihydropyrimidine deshydrogenase (DHPD) (30), and
20 dihydroorotate deshydrogenases (DODH) (31), which facilitate the reduction of uracil to
21 dihydrouracil and orotate to dihydroorotate (a uracil derivative), respectively. This implies that
22 nature utilizes reduced flavin for uridine reduction. However, in the tertiary structure
23 generated by AlphaFold, RdsA distinguishes itself from these FMN-dependent flavoenzyme
24 systems due to its predicted distinct structure, displaying a unique topology. Unlike Dus,
25 DHPD, and DODH, where the catalytic site binding FMN is carried by a TIM BARREL,
26 RdsA's flavin is held by a Rossmann fold domain (FBD) housing two additional inserted
27 domains (ins1D and ins2D). Surface charge representation on the AlphaFold model unveils
28 several patches of positive charges concentrated on the three domains, arranging around the
29 active site pocket and likely defining the RNA binding site (**Figure S11**). This model, if proved
30 to be correct, would suggest that FBD, ins1D, and ins2D probably play a role in substrate
31 recognition. RNA modifications follow a non-random, stepwise progression at distinct stages
32 during ribosome biogenesis. Despite the complexity of this process, our observation of
33 dihydrouridylation activity on a mini-ARN containing a small PTC-like region by RdsA should
34 imply a potential role for this enzyme in acting on a ribosome precursor.

35 Phylogenetic analysis reveals the widespread distribution of YhiN among bacteria, implying
36 the presence of D in the rRNA of numerous bacterial species. Indeed, our analyses of *V.*
37 *cholerae* rRNA demonstrate, for the first time, the presence of D2432 in this organism
38 (corresponding to the D2449 in *E. coli*), and furthermore this species harbors a gene
39 encoding RdsA. Nevertheless, our study brings attention to a perplexing observation: the
40 absence of YhiN in the genome of *Clostridium sporogenes*, despite the detection of
41 m5D2449 and D2500 in its 23S rRNA (12). This suggests the potential existence of another
42 unidentified type of Dus. In contrast, *B. subtilis* harbors a *yhiN* gene, yet its rRNA
43 modification profile does not show any D residues (1), hinting at potential alternative
44 biological functions for this flavoprotein. This proposition gains support from our findings,
45 which reveal that RdsA belongs to a non-isofunctional superfamily. This is further
46 underscored by the identification of a homologous gene to *E. coli yhiN*, operating as an
47 aminoacetone oxidase in *Streptococcus oligofermentas* (28), albeit with measured activities
48 that are notably weak.

1 Our genetic studies show that a functional RdsA enzyme is not essential for *E. coli* survival,
2 and this is in keeping with the majority of rRNA modification enzymes. However, the
3 biological role of ribosomal D remains elusive. Multiple studies have demonstrated that the
4 lack of rRNA modifications can result in changes of ribosomal active sites (32), leading to a
5 reduction in translation speed, an increase in mRNA reading errors by the ribosome (33),
6 reduced virulence in pathogenic bacterial species and a disrupted response to metabolites
7 and antibiotics (34, 35). Limited to microbes, including certain pathogens, our research lays
8 the groundwork for exploring the impact of dihydrouridylation on bacterial ribosome
9 maturation and the potential targeting of this process by inhibitors.

10 **Materials and methods (478 words)**

11 **Strains media and oligonucleotides used in this study**

12 All strains are listed in Table S6. BW25113 were obtained from the Keio collection and
13 genomic deletion strains were obtained from NBRP (Japan). Cells were usually grown in LB
14 medium at 37°C. Oligonucleotides used for RhoRT-PCR and MALDI-TOF analysis are listed
15 in Table S7.

16 **Rhodamine labelling, reverse transcription and PCR**

17 Method for rhodamine-110 labelling was derived from the one described in (7). From the
18 RNA stock, 50 µg was extracted, to which 16 µl of 1 M Tris HCl pH 7.5 and 40 µl of 100
19 mg/ml NaBH₄ diluted in 10 mM KOH were added, reaching a final volume of 400 µl. After
20 one hour of incubation at room temperature (open tubes in the dark), the reaction was
21 neutralized by adding 80 µl of 6 M acetic acid. The treated RNAs were then precipitated with
22 0.3 M NaCl and two volumes of 100% ethanol. The rRNAs were subsequently resuspended
23 in 550 µl of 0.1 M formate buffer pH 3.15, containing 0.5 mM Rhodamine 110, and incubated
24 at 37°C for 1.5 hours for fluorochrome conjugation. The reaction was stopped by adding 0.2
25 M Tris-HCl, pH 8.5, and the excess fluorochrome was removed by acidic phenol extraction
26 (pH 4.3). Finally, the RNA was precipitated and resuspended in RNase free H₂O. Reverse-
27 transcription and PCR conditions are detailed in supplementary information.

28 **Analysis of ribosomal dihydrouridylation at position 2449**

29 Dihydrouridine quantification was done by tandem LC-MS as detailed in supplementary
30 information. Ribosomal modification profiles were analyzed by the AlkAniline-Seq method
31 (36). This method exploits the instability of the D-ring under alkaline conditions leading to its
32 cleavage and formation of β-ureidopropionic acid (21). This instability leads to aniline-driven
33 RNA cleavage generating a 5'-phosphate group (5'-P) on the neighboring N+1 residue, which
34 serves as an input for highly selective ligation of sequencing adapters.

35 **Biochemical characterization of RdsA**

36 *E. coli yhiN* gene was cloned in pET15b for production and purification in BL21-DE3 *E. coli*
37 cells. The NADH oxidase activity was determined by following decrease of A₃₄₀ under aerobic
38 conditions and dihydrouridylation assays were performed by incubating the recombinant
39 RdsA with the target RNA in a buffer containing 50 mM HEPES pH 7.5, 150 mM NaCl, 5 mM
40 DTT, 10 mM MgCl₂, 100 µM FAD and 15% glycerol. Detailed procedure for LC-MS analysis
41 of the reaction is detailed in supplementary information.

42 **General Bioinformatic resources**

43 For literature and sequence retrievals, the resources at NCBI (<https://www.ncbi.nlm.nih.gov/>)
44 (37), UniProt (<https://www.uniprot.org>) (38), EggNog (<http://eggnog6.embl.de/>) (39) and BV-

1 BRC (<https://www.bv-brc.org>) (40) were routinely used. PaperBlast was used to find
2 published papers on members of the COG2018 family (papers.genomics.lbl.gov/) (38). The
3 SubtiWiki/CoreWiki (<http://corewiki.uni-goettingen.de/>) (41), the String database
4 (<https://string-db.org/>) (42) and the EFI Gene Neighborhood webtool (43) were used to
5 explore physical clustering. Protein sequences were aligned using MUSCLE v5.1 (44) and
6 visualized using Weblogo3 (45). Sequence similarity networks and phylogenetic analysis are
7 detailed in supplementary information.

8

9 **Data, Materials, and Software Availability**

10 All study data are included in the article and/or SI Appendix.

11 **Acknowledgments**

12 This research is funded by ANR/DFG grant DERASE (#20-CE92-0030), by grant GM132254
13 to V.d.C.L. M.H. was funded by the Deutsche Forschungsgemeinschaft (DFG, German
14 Research Foundation) –Project-ID 439669440 – TRR-319–TP C01 and, HE 3397/21-1. S.D.
15 was funded by a grant from the Danish Research Agency (FNU-rammebevilling 10-084554).

16 **Author contributions**

17 D.B. and D.H. designed the research project; D.B., D.H., S.D., V.D.C, M.H., Z.B. and Y.M
18 designed the experiments, S.T., Q.T., L.M.K., V.M., Y.Y., C.S., C.G, O.J.J. and S.R.
19 performed the research; all authors analyzed data; all authors contribute to the writing of the
20 paper.

21 **Competing interests**

22 The authors declare no competing interest.

1 References

- 2 1. A. Cappannini, *et al.*, MODOMICS: a database of RNA modifications and related
3 information. 2023 update. *Nucleic Acids Res.* **52**, D239–D244 (2024).
- 4 2. C. Lorenz, C. E. Lünse, M. Mörl, tRNA Modifications: Impact on Structure and Thermal
5 Adaptation. *Biomolecules.* **7**, 35 (2017).
- 6 3. Y. Motorin, M. Helm, tRNA stabilization by modified nucleotides. *Biochemistry.* **49**,
7 4934–4944 (2010).
- 8 4. D. Brégeon, *et al.*, Dihydrouridine in the Transcriptome: New Life for This Ancient RNA
9 Chemical Modification. *ACS Chem. Biol.* **17**, 1638–1657 (2022).
- 10 5. O. Finet, C. Yague-Sanz, F. Marchand, D. Hermand, The Dihydrouridine landscape
11 from tRNA to mRNA: a perspective on synthesis, structural impact and function. *RNA*
12 *Biol.* **19**, 735–750 (2022).
- 13 6. W. Dai, *et al.*, Activity-based RNA-modifying enzyme probing reveals DUS3L-mediated
14 dihydrouridylation. *Nat. Chem. Biol.* **17**, 1178–1187 (2021).
- 15 7. O. Finet, *et al.*, Transcription-wide mapping of dihydrouridine reveals that mRNA
16 dihydrouridylation is required for meiotic chromosome segregation. *Mol. Cell.* **82**, 404-
17 419.e9 (2022).
- 18 8. A. S. Draycott, *et al.*, Transcriptome-wide mapping reveals a diverse dihydrouridine
19 landscape including mRNA. *PLoS Biol.* **20**, e3001622 (2022).
- 20 9. X. Chen, S. Sim, E. J. Wurtmann, A. Feke, S. L. Wolin, Bacterial noncoding Y RNAs are
21 widespread and mimic tRNAs. *RNA.* **20**, 1715–1724 (2014).
- 22 10. J. A. Kowalak, E. Bruenger, J. A. McCloskey, Posttranscriptional modification of the
23 central loop of domain V in *Escherichia coli* 23 S ribosomal RNA. *J. Biol. Chem.* **270**,
24 17758–17764 (1995).
- 25 11. S. A. Fromm, *et al.*, The translating bacterial ribosome at 1.55 Å resolution generated by
26 cryo-EM imaging services. *Nat. Commun.* **14**, 1095 (2023).
- 27 12. F. Kirpekar, *et al.*, Mapping of ribosomal 23S ribosomal RNA modifications in
28 *Clostridium sporogenes*. *RNA Biol.* **15**, 1060–1070 (2018).
- 29 13. G. Emmerechts, S. Barbé, P. Herdewijn, J. Anné, J. Rozenski, Post-transcriptional
30 modification mapping in the *Clostridium acetobutylicum* 16S rRNA by mass
31 spectrometry and reverse transcriptase assays. *Nucleic Acids Res.* **35**, 3494–3503
32 (2007).
- 33 14. F. Yu, *et al.*, Molecular basis of dihydrouridine formation on tRNA. *Proc. Natl. Acad. Sci.*
34 *U.S.A.* **108**, 19593–19598 (2011).
- 35 15. H. Kusuba, *et al.*, In vitro dihydrouridine formation by tRNA dihydrouridine synthase
36 from *Thermus thermophilus*, an extreme-thermophilic eubacterium. *J. Biochem.* **158**,
37 513–521 (2015).
- 38 16. C. Bou-Nader, *et al.*, An extended dsRBD is required for post-transcriptional
39 modification in human tRNAs. *Nucleic Acids Res.* **43**, 9446–9456 (2015).

- 1 17. R. T. Byrne, *et al.*, Major reorientation of tRNA substrates defines specificity of
2 dihydrouridine synthases. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 6033–6037 (2015).
- 3 18. A. C. Bishop, J. Xu, R. C. Johnson, P. Schimmel, V. de Crécy-Lagard, Identification of
4 the tRNA-dihydrouridine synthase family. *J. Biol. Chem.* **277**, 25090–25095 (2002).
- 5 19. C. Bou-Nader, D. Brégeon, L. Pecqueur, M. Fontecave, D. Hamdane, Electrostatic
6 Potential in the tRNA Binding Evolution of Dihydrouridine Synthases. *Biochemistry.* **57**,
7 5407–5414 (2018).
- 8 20. C. Bou-Nader, *et al.*, Unveiling structural and functional divergences of bacterial tRNA
9 dihydrouridine synthases: perspectives on the evolution scenario. *Nucleic Acids Res.*
10 **46**, 1386–1394 (2018).
- 11 21. F. Xing, S. L. Hiley, T. R. Hughes, E. M. Phizicky, The specificities of four yeast
12 dihydrouridine synthases for cytoplasmic tRNAs. *J. Biol. Chem.* **279**, 17850–17860
13 (2004).
- 14 22. L. W. Rider, M. B. Ottosen, S. G. Gattis, B. A. Palfey, Mechanism of dihydrouridine
15 synthase 2 from yeast and the importance of modifications for efficient tRNA reduction.
16 *J. Biol. Chem.* **284**, 10324–10333 (2009).
- 17 23. J. J. Dalluge, T. Hashizume, A. E. Sopchik, J. A. McCloskey, D. R. Davis,
18 Conformational flexibility in RNA: the role of dihydrouridine. *Nucleic Acids Res.* **24**,
19 1073–1079 (1996).
- 20 24. N. Dyubankova, *et al.*, Contribution of dihydrouridine in folding of the D-arm in tRNA.
21 *Org. Biomol. Chem.* **13**, 4960–4966 (2015).
- 22 25. M. Hashimoto, *et al.*, Cell size and nucleoid organization of engineered *Escherichia coli*
23 cells with a reduced genome. *Mol. Microbiol.* **55**, 137–149 (2005).
- 24 26. J. Kato, M. Hashimoto, Construction of consecutive deletions of the *Escherichia coli*
25 chromosome. *Mol. Syst. Biol.* **3**, 132 (2007).
- 26 27. S. C. Harris, *et al.*, Identification of a gene encoding a flavoprotein involved in bile acid
27 metabolism by the human gut bacterium *Clostridium scindens* ATCC 35704. *Biochim*
28 *Biophys. Acta. Mol. Cell. Biol. Lipids.* **1863**, 276–283 (2018).
- 29 28. G. Molla, *et al.*, Aminoacetone oxidase from *Streptococcus oligofermentans* belongs to
30 a new three-domain family of bacterial flavoproteins. *Biochem. J.* **464**, 387–399 (2014).
- 31 29. P. Zhou, L. Liu, H. Tong, X. Dong, Role of operon *aaoSo-mutT* in antioxidant defense in
32 *Streptococcus oligofermentans*. *PLoS One.* **7**, e38133 (2012).
- 33 30. D. Dobritzsch, G. Schneider, K. D. Schnackerz, Y. Lindqvist, Crystal structure of
34 dihydropyrimidine dehydrogenase, a major determinant of the pharmacokinetics of the
35 anti-cancer drug 5-fluorouracil. *EMBO J.* **20**, 650–660 (2001).
- 36 31. R. A. G. Reis, F. A. Calil, P. R. Feliciano, M. P. Pinheiro, M. C. Nonato, The
37 dihydroorotate dehydrogenases: Past and present. *Arch. Biochem. Biophys.* **632**, 175–
38 191 (2017).
- 39 32. H. Demirci, *et al.*, Modification of 16S ribosomal RNA by the KsgA methyltransferase
40 restructures the 30S subunit to optimize ribosome function. *RNA.* **16**, 2319–2324
41 (2010).

- 1 33. A. Baudin-Baillieu, *et al.*, Nucleotide modifications in three functionally important regions
2 of the *Saccharomyces cerevisiae* ribosome affect translation accuracy. *Nucleic Acids*
3 *Res.* **37**, 7665–7677 (2009).
- 4 34. C. P. van Buul, W. Visser, P. H. van Knippenberg, Increased translational fidelity
5 caused by the antibiotic kasugamycin and ribosomal ambiguity in mutants harbouring
6 the *ksgA* gene. *FEBS Lett.* **177**, 119–124 (1984).
- 7 35. S. Okamoto, *et al.*, Loss of a conserved 7-methylguanosine modification in 16S rRNA
8 confers low-level streptomycin resistance in bacteria. *Mol. Microbiol.* **63**, 1096–1106
9 (2007).
- 10 36. V. Marchand, *et al.*, AlkAniline-Seq: Profiling of m7 G and m3 C RNA Modifications at
11 Single Nucleotide Resolution. *Angew. Chem. Int. Ed. Engl.* **57**, 16785–16790 (2018).
- 12 37. E. W. Sayers, *et al.*, Database resources of the National Center for Biotechnology
13 Information. *Nucleic Acids Res.* **52**, D33–D43 (2024).
- 14 38. UniProt Consortium, UniProt: the Universal Protein Knowledgebase in 2023. *Nucleic*
15 *Acids Res.* **51**, D523–D531 (2023).
- 16 39. A. Hernández-Plaza, *et al.*, eggNOG 6.0: enabling comparative genomics across
17 12 535 organisms. *Nucleic Acids Res.* **51**, D389–D394 (2023).
- 18 40. R. D. Olson, *et al.*, Introducing the Bacterial and Viral Bioinformatics Resource Center
19 (BV-BRC): a resource combining PATRIC, IRD and ViPR. *Nucleic Acids Res.* **51**,
20 D678–D689 (2023).
- 21 41. L. A. Flórez, S. F. Roppel, A. G. Schmeisky, C. R. Lammers, J. Stülke, A community-
22 curated consensual annotation that is continuously updated: the *Bacillus subtilis* centred
23 wiki SubtiWiki. *Database (Oxford)*. **2009**, bap012 (2009).
- 24 42. D. Szklarczyk, *et al.*, The STRING database in 2023: protein-protein association
25 networks and functional enrichment analyses for any sequenced genome of interest.
26 *Nucleic Acids Res.* **51**, D638–D646 (2023).
- 27 43. R. Zallot, N. Oberg, J. A. Gerlt, The EFI Web Resource for Genomic Enzymology Tools:
28 Leveraging Protein, Genome, and Metagenome Databases to Discover Novel Enzymes
29 and Metabolic Pathways. *Biochemistry*. **58**, 4169–4182 (2019).
- 30 44. R. C. Edgar, MUSCLE: multiple sequence alignment with high accuracy and high
31 throughput. *Nucleic Acids Res.* **32**, 1792–1797 (2004).
- 32 45. G. E. Crooks, G. Hon, J.-M. Chandonia, S. E. Brenner, WebLogo: a sequence logo
33 generator. *Genome Res.* **14**, 1188–1190 (2004).

34

1 **Figure legends**

2 **Figure 1: Dihydrouridylation of the peptidyl transferase center of the *E. coli* ribosome:**
3 (A) Biosynthetic reaction of dihydrouridine involves the reduction of the C5=C6 double bond
4 of uridine. This redox reaction requires 2 electrons + 2 protons. (B) Secondary structure of
5 the V domain of *E. coli* 23S rRNA. D2449 is highlighted in red, and the question mark
6 indicates that the associated modification enzyme was unknown before our current study. (C)
7 High-resolution structure of the *E. coli* ribosome in complex with tRNA (magenta). The PDB
8 associated with this structure is 8B0X (11). The inset on the right provides a close-up of the
9 peptidyl transferase center (PTC), showing the positioning of D2449 (pink) in proximity to
10 A2451 and m²G2445.

11

12 **Figure 2: Principle of RhoRT-PCR**

13 (A) Diagram showing the main steps of RhoRT-PCR and the expected results depending on
14 the presence of D in the analyzed region. The red scrambled oval depicts specific D-labelling
15 with rhodamine-110. (B) Representative results obtained with rRNA containing D at position
16 2449 (*E. coli* BW25113, 14 repeats) or U at the same position (*B. subtilis* W168, 2 repeats).
17 RT was performed with species specific 4R oligonucleotides. PCR amplification was
18 performed with species specific 1F/3R (1F) and 2F/3R (2F) oligonucleotides.

19

20 **Figure 3: Identification of D2449 deficient *E. coli* strains**

21 (A) Representative RhoRT-PCR results with the indicated strains (BW25113 and ME5125,
22 14 repeats; ME5079 and BW25113 $\Delta yhiN::kan$, 3 repeats). Primers used are the same as in
23 figure 2A. Absence of 1F/3R amplification with BW25113 rRNA as a substrate indicates the
24 presence of D2449, which was confirmed by MALDI-TOF analysis of RNase digestion
25 fragments (2 repeats) (B) and AlkAniline-Seq (C) (for the latest, BW25113, ME5125 and
26 ME5079, 3 repeats; BW25113 $\Delta yhiN::kan$, 2 repeats).

27

28 **Figure 4: Structural and Biochemical Characterization of *E. coli* rRNA Dihydrouridine**
29 **Synthase.** (A) Structural organization of *E. coli* RdsA. At the top, a schematic diagram
30 depicting the modularity of RdsA and delineating various domains – Rosman fold domain in
31 pink, Ins1 domain in blue, and Ins2 domain in green. At the bottom, an AlphaFold model of *E.*
32 *coli* RdsA in a cartoon representation with the same color code as above. The Ins1 domain is
33 homologous to the C-terminal domain of *E. coli* EF-Tu (PDB: 1EFC, colored in blue), while
34 the Ins2 domain is homologous to the helix-2-turn-helix domain of *S. shibatae* DNA
35 topoisomerase VI-B subunit (PDB: 1MU5, in green). (B) Holoprotein model of RdsA showing
36 the binding of the FAD coenzyme (ball sticks, yellow) at the Rossmann fold domain and key
37 residues potentially interacting with FAD. (C) UV-Visible spectrum of native (red) and 0.1%
38 SDS-denatured (blue) recombinant RdsA protein from *E. coli*. (D) FAD/protein molar ratio of
39 recombinant RdsA harboring no change (WT) or mutations changing key residues interacting
40 with FAD shown in B. This ratio \pm SEM of 3 repeats was determined by spectrophotometric
41 quantification of released FAD from a given amount of SDS-treated protein. For each mutant
42 the targeted interaction with FAD is indicated. (E) Michaelis-Menten curve for the NADH
43 oxidation by RdsA. (F) Sequence and potential structure of the *in vitro* transcript Mini-PTC-U
44 (top) and Mini-PTC-C (bottom) established with MXFold2 server. Nucleotide at position 21 is
45 the anticipated target of dihydrouridylation. (G) Dihydrouridylation activity of RdsA on a Mini-
46 PTC-U (blue) and Mini-PTC-C (green). D was quantified by LC-MS after incubation in

1 reaction buffer of the designated transcript without or with 25 μ M of RdsA for the indicated
2 time. Results are expressed as the percentage of quantified D per quantified A \pm SEM of 3 to
3 4 repeats (%mod/A).

4 **Figure 5. Protein sequence similarity network (SSN) of PF03486 family proteins**

5 (A) SSN of 7,930 PF03486 family proteins with the cutoff of alignment score of 80. Each
6 node in the network represents one or multiple PF03486 family proteins that share 90% or
7 more identity. An edge (represented as a line) is drawn between two nodes with a BLAST *E*-
8 value cutoff of better than 10^{-80} . The nodes are colored based on taxonomy or the gene
9 neighborhood context of the corresponding genomes. The cluster containing YihN_{Ec} is boxed.
10 (B) Examples of gene neighborhoods organization. Uniprot IDs of PF03486 proteins:
11 *Clostridium scindens* BaiN, B0NAQ4; *Streptococcus cristatus* AAO, E8JVH0; *Bacillus subtilis*
12 YhiN, Q795R8; *Acinetobacter calcoaceticus* YhiN1, A0A0A8XKZ3; *Acinetobacter*
13 *calcoaceticus* YhiN2, A0A446ZNN4; *Peptococcaceae bacterium* YhiN, A0A942G053;
14 *Synechococcales bacterium* YhiN, A0A354AWX7.