Deazaguanine derivatives, examples of crosstalk between RNA and DNA modification pathways

Geoffrey Hutinet, Manal A. Swarjo, and Valérie de Crécy-Lagard

Introduction

DNA and RNA, the two cellular information polymers are made of very similar building blocks: nucleotides composed of a nitrogenous base, a five carbon sugar and a phosphate group. The seemingly small differences between the sugars (ribose in RNA and 2’-deoxyribose in DNA) and one of the four bases (thymine (T) in DNA and uracil (U) in RNA) do however dramatically change the properties of the two molecules. DNA is more stable as a double stranded helix, and has been recruited as the core matrix of genetic information. RNA is generally single stranded and more flexible, harboring different shapes and functions in the cell. Historically, the fields of DNA and RNA research had been kept quite separate with the two communities rarely interacting, but this has recently changed with 1) the popularity of the RNA world hypothesis, and the realization that DNA is a modified form of RNA; 2) the growing interest in RNA modifications that has revealed how similar many of the RNA and DNA modification machineries are (see reference [1] for a detailed discussion).

While the first nucleoside modifications were found in genomic DNA, with the discovery of 5-methyl-deoxycytosine (m5C) or 5mC in Mycobacteria in 1925,1 they are more diverse and chemically complex in RNA. Around twenty modifications have been found to date in DNA.2,3 Most of the known genomic DNA modifications are methylation, although more complex modifications do exist, particularly in phages.2,4 In contrast, more than one hundred modifications have been found in RNA,3 and some of these modifications are very complex, e.g., wybutosine which requires five enzymes for its synthesis.5 Modifications are mainly found in stable RNAs such as transfer-RNAs (tRNAs) and ribosomal-RNA (rRNAs),3 and a few are also found in messenger-RNA (mRNAs).6,7 The molecular and biochemical characterization of RNA modifying enzymes has revealed multiple cases of broad specificity for various RNA substrates. For example, some pseudouridine syntheses have been shown to modify both mRNA and tRNA substrates,6 whereas methylases of the RimD family methylate rRNAs in some organisms and tRNAs in others.9

More recently, examples of crosstalks have been observed between the RNA and DNA modification machineries. It is the case in the eukaryotic DNMT family of methyltransferases that is subdivided in three groups: DNMT1, 2 and 3. DNMT1 and DNMT3 methylate the carbon 5 of cytosines in DNA CpG sequences, playing key roles in epigenetic regulation.10 DNMT2, despite its high sequence similarity to DNMT1 and 3, modifies specific tRNAs, e.g., tRNA^{Asp}_{GUC}, at the position C38.11 Cross-talks also exist in the APOBEC family of proteins that deaminate cytosines into uridines. Although the first proteins discovered in the APOBEC family were mRNA modifying proteins, almost all members of the APOBEC family also modify single stranded DNA.12 Similarly, some members of the ADAT family, first discovered as a family of deaminases responsible for conversion of adenosine to inosine in tRNAs, deaminate cytosine into uridine in DNA.13 Finally, the AlkB proteins are a family of oxidative deamination enzymes that has been found to date in DNA and RNA, and recently expanded to include members of the FTO subgroup that are responsible for conversion of 2-hydroxylation of mcm5^U in tRNAs.15 In addition, members of the FTO subgroup were found to demethylate mRNAs, adding regulatory functions to the AlkB family.16

In retrospect, these examples of crosstalk are not surprising, as deamination and methylation reactions were already known for both RNA and DNA. The novelty mainly derives from the switch in specificity within an enzyme superfamily from a RNA
to a DNA substrate. However, the recent discovery that DNA contains 7-deazaguanine derivatives, modifications that were believed to exist exclusively in tRNA, is a novel and unexpected example of crosstalk between the RNA and DNA modification machineries, which is the focus of this review.

**Queuosine and archaeosine function in tRNA**

Queuosine (Q) and archaeosine (G\(^+\)) are long-known 7-deazaguanine derivatives that were thought to exist exclusively in tRNA molecules. Q is found in many eukaryotes and bacteria at the wobble position of tRNAs harboring G\(_{34}\)U\(_{35}\)N\(_{36}\) anticodon sequences. These tRNAs are responsible for the insertion of Asn, Asp, His and Tyr amino acids in proteins. A role for Q in decoding efficiency was demonstrated over 30 years ago and recent work by the Farabaugh laboratory has also shown that Q affects decoding accuracy in both directions depending on the codon. However, Q is not required for growth under most tested conditions and its physiological importance remains elusive, particularly as it was repeatedly lost in the course of evolution. In bacteria, the absence of Q does not seem to be critical in exponential growth but affects growth under stress conditions in a diverse set of organisms and recent work by the Farabaugh laboratory has also shown that Q affects decoding accuracy in both directions depending on the codon. However, Q is not required for growth under most tested conditions and its physiological importance seems to be critical in exponential growth but affects growth under stress conditions in a diverse set of organisms and recent work by the Farabaugh laboratory has also shown that Q affects decoding accuracy in both directions depending on the codon.

**PreQ\(_0\) synthesis in bacteria and archaea**

PreQ\(_0\) is synthesized de novo from GTP in many bacteria and archaea by a series of complex reactions catalyzed by four enzymes (Fig. 1, purple arrows). The first step of the preQ\(_0\) pathway, the formation of dihydroequoripent triphosphate (H\(_2\)NTP), is not a dedicated step but is shared with the tetrahydrofolate (THF) and biotin (BH\(_4\)) pathways. It is catalyzed by GTP cyclohydrolase I (EC 3.5.4.16). Most organisms use a zinc\(^2+\)-dependent GTP cyclohydrolase I (FolE), a member of the Tunnel-folate (THF) and biopterin (BH\(_4\)) pathways. It is catalyzed by GTP cyclohydrolase I (FolE2), another T-fold superfamily member that utilizes other enzymes (Fig. 1, red arrows). The preQ\(_0\) precursor is a DNA substrate. However, the recent discovery that DNA contains 7-deazaguanine derivatives, modifications that were believed to exist exclusively in tRNA, is a novel and unexpected example of crosstalk between the RNA and DNA modification machineries, which is the focus of this review.

**Q synthesis in bacteria**

Three more steps are required to synthesize Q in bacteria (Fig. 1, red arrows). The preQ\(_0\) precursor is first reduced to 7-aminomethyl-7-deazaguanine (preQ\(_1\)) by the NADPH-dependent 7-cyano-7-deazaguanine reductase (EC 1.7.1.13) enzyme QueF. QueF is closely related to FolE in primary structure, and the two families can be distinguished at the sequence level by a QueF-specific motif involved in NADPH binding and by the FoLE-specific zinc binding residues. QueF is also a T-fold enzyme and, like GTP cyclohydrolase I, is represented by the unimodular and bimodular T-fold subfamilies. In unimodular QueF, a homodecameric enzyme, catalysis occurs at the intersubunit interfaces; whereas in bimodular QueF, a homodimer, catalysis occurs at the intrasubunit interface between the two T-fold modules, not at the intersubunit interface as in the bimodular FolE.

The entire pathway to preQ\(_0\) is independent of tRNA. tRNA comes into play when the bacterial tRNA-guanine transglycosylase (EC 2.4.2.29, bTGT) enzyme removes the G base at position 34 from the target tRNAs and replaces the base with preQ\(_0\) (Fig. 1). Extensive biochemical and structural studies on bTGT have shown that the enzyme is a homodimer that binds one tRNA molecule with its U\(_{33}\)G\(_{34}\)U\(_{35}\) sequence as the recognition element. The monomer consists of an irregular (beta/alpha)\(_8\) TIM barrel domain that harbors specific loop insertions that guarantee proper tRNA recognition, and a tightly attached C-terminal zinc-binding domain involved in dimerization (Fig. 2). A reaction mechanism has been proposed in which G\(_{34}\) of tRNA first binds in the purine recognition pocket, constituted of the residues...
D102, S103, D156, Q203 and G230 (Zymomonas mobilis TGT protein numbering, Fig. 2). Nucleophilic attack by an active-site aspartate side chain (D280 in Z. mobilis TGT) on the ribosyl C1 of G34 detaches the guanine base from the RNA, and results in formation of a covalent intermediate between TGT and RNA. A subsequent conformational change in the active site facilitates the release of the detached guanine base from the pocket and the binding of preQ1 in the same pocket with L231 and A232 as specificity determinants. Deprotonation of N9 of the incoming preQ1 by another active-site aspartate (D102) allows nucleophilic attack by N9 on ribose C1' to form the preQ1-tRNA product.

Two subsequent reactions are required to synthesize the final Q product. First, S-adenosylmethionine tRNA ribosyltransferase-isomerase (EC 2.4.99.17, QueA) catalyzes the addition of an epoxycyclopentandiol ring to preQ1 to form the epoxyqueuosine (oQ). QueA uses the ribose moiety from S-adenosyl-methionine (SAM), and transfers it to the ammonium moiety of preQ1. Finally, the epoxyqueuosine reductase (EC

Figure 1. Deazaguanine derivative synthesis pathways. GTP is the preQ0 precursor in both bacteria and archaea (purple arrows). In most bacteria, four more enzymatic steps lead to the insertion of Q in tRNAs at position 34 (red arrows). In a few organisms, preQ0 can be transformed to secondary metabolites such as toyocamycin or sangivamycin antibiotics (red arrows, toy genes). In eukaryotes, queuine is salvaged (green circle) and directly transferred to tRNAs (green arrows). Bacteria salvage preQ0 (red circle), and both bacteria and archaea salvage preQ0 (purple circle). In archaea, preQ0 is transferred to position 15 of tRNA before being modified to G (blue arrows). PreQ0 and ADG have been found in bacterial DNA (dashed red arrows) and G in phage DNA (dashed yellow arrow). All dashed arrows represent uncharacterized reactions. All molecule abbreviations and protein names are described in the text.
1.17.99.6) reduces oQ into Q⁶⁰. In E. coli, this reaction is catalyzed by QueG,⁶¹ a B₁₂-dependent enzyme with iron-sulfur clusters. Because QueG is missing in many organisms that harbor Q in tRNA (Zallot and de Crécy-Lagard, unpublished), it is expected that non-orthologous enzymes catalyzing the same reaction are yet to be discovered in these organisms.

### G⁺ synthesis in archaea

ArcTGT is the archaeal bTGT homolog that inserts preQ⁰ at position 15 (or 13 in specific organisms) in the dihydrouridine arm (D-arm) of target tRNAs (Fig. 1, blue arrows). Like bTGT, arcTGT functions as a homodimer. The monomer is constituted of an N-terminal catalytic domain, and three arcTGT-specific C-terminal domains C₁, C₂ and C₃/PUA (PseudoUridine synthase and Archaeosine transglycosylase domain). The catalytic domain is built around an (α/β)₉ TIM barrel subdomain and a zinc-binding subdomain, making it very similar in structure to the bTGT monomer (Fig. 2). This catalytic domain also mediates the homodimerization of arcTGT through the zinc-binding sites, similar to bTGT homodimerization.⁶²,⁶³ The C₁ domain provides an additional homodimerization interface, while the C₂ and C₃/PUA domains provide a scaffold for tRNA binding. However, unlike the bTGT homodimer which binds one tRNA molecule, the arcTGT homodimer binds two tRNA molecules such that each tRNA interacts with both protein subunits. One subunit provides binding capacity through its C-terminal domains, while the other subunit performs the reaction.⁶³

ArcTGT distorts the tRNA structure from the canonical L-shaped form to the 1-form in which the D-arm is single stranded and protruded, allowing G₁₅ to be accessible for transglycosidation. The enzyme presumably positions the exposed G₁₅ in the active site by “counting” the nucleotides from G₁ to G₁₅ in the 1-form. Although the substrate binding pocket and catalytic residues are conserved between the bacterial and archaean enzymes, the replacement of L231/A232 in bTGTs by V197/V198 in arcTGTs (Fig. 2) allows for the change of substrate specificity from preQ₁ to preQ⁰.⁶⁴

Once preQ₁ has been inserted into the target archaean tRNA, one last amidotransferase step is required to produce the G⁺ modification. Surprisingly, a great diversity of amidotransferase enzymes have been recruited to catalyze this reaction (Fig. 1). The first to be discovered was archaeosine synthase (EC 2.6.1.97, ArcS).⁶⁴ ArcS catalyzes the conversion of preQ₀-tRNA to G⁺-tRNA in an ATP-independent manner and may use glutamine, asparagine as well as free ammonia as ammonium donors in vitro.⁶⁴ ArcS is evolutionarily and structurally related to arcTGT. It has retained the three tRNA binding domains found in arcTGT but has acquired an insertion of four α-helices and 3 β-strands in its C-terminal domain that confer a Rossmann fold architecture to the domain. This domain extension contains a conserved PCX,KPYX,SX,H motif proposed to be involved in catalytic activity. Like the TGT family, ArcS proteins form homodimers but are predicted to bind the L-form, not the 1-form tRNA.⁶⁴

Many archaea that synthesize G⁺ lack ArcS homologs and, in those organisms, two non-orthologous enzymes were found to be responsible for the formation of G⁺.⁶⁵ The first is QueF-Like (QueF-L) which converts preQ₀ base of preQ₀-modified tRNA to G⁺, unlike the bacterial preQ₁ synthase QueF which acts on the free preQ₀ base as substrate.⁶⁶ QueF-Like is also a homodimeric T-fold enzyme with a similar structure to that of QueF, except that its pentameric subunits are tightly wound, resulting in elimination of the QueF-specific NADPH binding site.⁶⁶ The second enzyme that replaces ArcS in some archaea is Gat-QueC, a homolog of QueC fused to a glutamine amidotransferase class-II domain (GATase). This enzyme is yet to be biochemically characterized, and it remains unclear whether it acts before or after integration of preQ₀ in tRNA and, consequently, whether the corresponding arcTGT integrates the G⁺ base or the preQ₀ base in the tRNAs of these organisms.

### Q salvage in eukaryotes

Most eukaryotes harbor Q in their tRNAs but do not produce the modification de novo. Instead, they salvage the queuine base (q) from their diet or microflora (see reference [67] for recent review) using transporters that are yet to be discovered. Members of the DUF219 family have recently been implicated in q salvage in eukaryotes.²² Structural modeling suggests a possible ribonucleotide hydrolase function for this family, but its exact role is yet to be determined.

The eukaryotic TGTs (eTGT) insert the q base directly in tRNAs (Fig. 1, green arrow). eTGTs are heterodimers composed of a catalytic subunit (queuine tRNA-ribosyltransferase catalytic subunit 1 or QTRT1) and an accessory subunit that lacks the active site Asp residue (queuine tRNA-ribosyltransferase accessory subunit 2 or QTRT2, previously called QRTD1).⁶⁸ Key differences between the eukaryotic and prokaryotic TGTs are in the substrate binding pocket where Val233 is replaced by a glycine and Cys158 is replaced by a valine. These changes enlarge the binding pocket and allow the binding of a larger substrate (q versus preQ₀ or preQ₁).⁶⁹
Seven-deazapurines salvage in bacteria

bTGT is the signature gene of the Q pathway. All organisms that harbor a bTGT homolog can harbor Q in tRNA and no organism lacking bTGT has ever been found to contain Q in tRNA (de Crécy-Lagard, unpublished). However, many organisms with bTGT homologs lack preQ0 synthesis enzymes and/or QueF.70 In these cases, the preQ0 or preQ1 precursor must be salvaged. preQ0 is certainly available in natural environments as an intermediate of the Q and G+ pathway as described above, but also as a secondary metabolite or secondary metabolite precursor.71-72 It also may be recycled from tRNA degradation products, because the free G+ base spontaneously deamidates to preQ0.73

Specific preQ0/preQ1 transporters of the ECF family (QueT and QrT) have been predicted to be involved in 7-deazapurine salvage but have never been experimentally validated as such, and their sparse distribution suggests that other transporter families exist. We recently reported that the COG1738/YhhQ family is a preQ0-specific transporter found in E. coli and is widespread in bacteria.75 Of note, some bacteria such as Chlamydia trachomatis not only lack the precursor pathway but also QueA homologs. Queine should therefore be salvaged in these bacteria. In summary, while the de novo Q pathway is well characterized, many open questions remain regarding the possibility of Q salvage in bacteria.

Discovery of dADG and dpreQ0 in bacterial DNA

Comparative genomic analyses led to the observation that two copies of the tgt gene are present in specific bacteria.17 One copy encodes the canonical bTGT, while the other copy encodes a larger, more divergent protein that exhibited common structural features with both arcTGT and bTGT enzymes. Careful sequence analysis showed that the catalytic residues and substrate binding pocket of arcTGT (e.g., D95, D249, V217, and V218 in P. horikoshii arcTGT numbers) are conserved in this protein (Fig. 2), suggesting that it may bind preQ0. In addition, the tRNA-binding C3/PUA domain characteristic of arcTGT is missing from this protein, and the zinc binding motif is present in a supplementary domain17 (Fig. 2).

Physical clustering analyses strongly linked this tgt paralog to DNA metabolism genes, leading to the prediction that this enzyme might be involved in inserting preQ0 or a derivative in DNA.17,76 This prediction was tested by analyzing genomic DNA, extracted from different bacteria that contained copies of the tgt paralog, by mass-spectrometry, specifically searching for the deoxyribose forms of the Q pathway intermediates. Deoxy-preQ0 (dpreQ0) was detected as predicted but seemed to be a minor product. The main deazapurine modification found in the DNA of these organisms was dADG, the deoxyribose form of the preQ0 precursor ADG. The discovery of these modifications led to renaming these tgt paralogs as dpdA, for DeoxyPurine in DNA.17

Discovery of dG+ in bacteriophage 9g DNA

Many bacteriophages (or phages) encode TGT-like proteins17 (Fig. 3). Early studies of the Streptococcus Dp-1 phage,77 and the Mycobacterium phage Rosebush78 had led to the proposal that these genes are involved in the synthesis of Q in tRNA, interfering with or enhancing the translation of the host or phage proteins. Another possible function was proposed by the Letarov group, who suggested that Q could be incorporated in the DNA of Escherichia coli phage 9g, causing its resistance to cleavage by several endonucleases.79 Consistent with this proposal is the observation that most phage TGT-like proteins are closer in sequence to DpdA than to the canonical TGT enzymes. Although the catalytic core and substrate recognition residues of arcTGT are conserved in this protein (Fig. 2), two out of the four residues constituting the zinc binding sites are not conserved (Fig. 2).

Physical clustering analysis of the 9g phage genome revealed that the dpdA gene clustered with a gat-queC homolog, suggesting that G+ might be inserted in this phage. This prediction was validated by mass-spectrometry and it was shown that 25% of the dG residues were modified to dG+ in the 9g DNA, a first occurrence of the G+ base outside the archaeal kingdom.17

The Dpd cluster and horizontal gene transfer

Similarly to tgt being the signature gene for Q in tRNA, dpdA is the signature gene for the presence of deazapurine derivatives in DNA. dpdA is part of a genomic island of about 20 kb in size (Fig. 3). Synteny analyses of closely related strains as well as co-distribution of cluster genes with dpdA helped define the island boundaries.17 Deletion of the whole dpd cluster in Salmonella enterica serovar Montevideo was shown to abolish the insertion of ADG and preQ0 in DNA.17 In the S. Montevideo genomes, the cluster is inserted at the leuX locus, a hypervariable region that is a known landing site for pathogenicity virulence and defense genes80 (Fig. 4). 70% of the available S. Montevideo sequences in the SEED database (SEED viewer version 2.081) contain the dpd cluster. This data, in combination with the phylogenetic analysis of the DpdA family that carries many deviations from the species tree17 and the wide and sporadic distribution of the dpd cluster around the bacterial tree,17 makes a strong case for the great mobility of this cluster and its propagation by horizontal gene transfer.

Three genes, dpdB, dpdC and dpdD, are nearly always associated with dpdA in the bacterial dpd clusters; while the dpdEFJK genes, all related to DNA metabolism, have a more variable distribution17 (Fig. 3). One or more preQ0 synthesis genes are also present in some clusters (Fig. 3).

Based on the sequence similarity with TGT, DpdA is the most promising candidate for the enzyme that inserts preQ0 and/or ADG in DNA. DpdB is found in 92% of the genomes that harbor dpdA. Sequence similarity analyses showed that DpdB is a member of the ParB nuclease superfamily,76 more specifically related to the DndB subgroup of proteins.17 DndB is part of the dnd cluster that introduces a phosphorothioate modification in the DNA backbone.82 DndB is not part of the core modification pathway but is a DNA binding protein with a regulatory role,83 suggesting that DpdB might bind DNA and could be involved in target recognition. dpdC was found in 88% and DpdD in 90% of the genome containing dpdA. Both encoded proteins were similar to proteins of unknown function, DUF328 for DpdC and DUF3225 for a small portion of dpdD,17 making any functional prediction difficult at this stage.
Function of deazapurines in DNA

Up to 25% of the G bases in 9g phage DNA are converted to G^+. The modification event in the bacterial genomes is more rare because only 0.01% of the G bases were observed to be modified to ADG/preQ_0 in the DNA analyzed. 9g DNA is also resistant to cleavage by a variety of endonucleases, hence dG might have an anti-restriction function allowing the phage to escape host restriction. S. Montevideo DNA on the other hand is not resistant to cleavage by restriction enzymes. However, at least one gene in the dpd cluster is involved in cutting DNA that lacks the deazapurine modification based on transformation efficiency assays, suggesting that the cluster carries a restriction/modification system that protects the bacteria from foreign DNA. The identity of the restriction enzyme remains an open question. Candidates includes genes like dpdD or dpdC that are of unknown function; or some of the other Dpd proteins, like DpdE that is part of the SNF-II superfamily that contains endonucleases from type I and III modification/restriction systems.

Predicted diversity of deazapurine modification of phages

G^+ is present in E. coli phage 9g DNA, and the presence of very similar gene clusters, including the gat-queC gene, suggests that G^+ incorporation should also be the case for E. coli phages JenK1, and JenP1/2. Analysis of the genomic context of DpdA-like encoding genes in all sequenced phages suggests that other deazapurine derivatives might be found in phages. Some phages, e.g., Mycobacteriophage Rosebush, encode all enzymes of the preQ_0 biosynthesis pathway (from FolE to QueC, Fig. 3), and are predicted to insert preQ_0 or ADG in their DNA.
phages like *Streptococcus* phage Dp-1 also encode QueF, and could therefore insert preQ1. Finally, some phages like *Mycobacteria* phage Orion only harbor a *dpdA*-like gene raising the possibility of salvage of Q precursors from the host.

**Conclusion**

It has become quite clear recently that 7-cyano-7-deazaguanine, or preQ0, is a key metabolite that is recruited in different pathways and is used in a variety of ways. PreQ0 can be the final product, as recently described in the *Streptomyces*, but its most frequent fate (in around 70% of sequenced bacteria and nearly all archaea) is to be inserted in tRNA where it is further modified to Q or G (Fig. 1). The preQ0 base is also used as precursor of secondary metabolites, such as toyocamycin and sangivamycin in several *Actinomycetes*, or can be inserted in DNA in organisms that harbor the *dnd* genomic island. It is still not clear yet if ADG is incorporated directly in DNA or if it is synthesized after preQ0 incorporation.

Some organisms will use the preQ0 precursor in multiple pathways like *S.* Montevideo that harbors both Q in tRNA and ADG in DNA, hence the preQ0 synthesis cluster is sometimes duplicated (Fig. 3). Finding the preQ0 transporter gene *yhhQ* imbedded in certain *dpd* clusters also suggests that salvaged preQ0 could be inserted in DNA.

The insertion of preQ0 in a nucleic acid polymer required the presence of a transglycosylase of the TGT family that must exchange the guanine base with the 7-deazaguanine derivative. The plasticity of the TGT family was well established as small variations in sequences allowed the different orthologs to favor preQ0, preQ1 or Q substrates. It had also been shown that the bTGT could modify mRNA in vitro, and artificially modify DNA when thymines were replaced by deoxyuracils. The discovery of the role of the TGT paralog DndA in modifying DNA has shown that this can happen in vivo and opens many questions on the identity of potential sequence specific endonuclease inhibited by the insertion of dADG, on the role of deaza-purine in protecting DNA from restriction, and on how the normal replication/transcription machineries can deal with these modifications.

**Disclosure of potential conflicts of interest**

No potential conflicts of interest were disclosed.

**Acknowledgments**

We thank Dirk Iwata-Reuyl for sharing unpublished results and Rémi Zalolet for editing the manuscript.

**Funding**

This work was supported by the National Institutes of Health (grant number R01 GM70641 to V.d.C.-L.)

**ORCID**

Manal A. Swarjo [http://orcid.org/0000-0003-1334-1763](http://orcid.org/0000-0003-1334-1763)  
Valérie de Crécy-Lagard [http://orcid.org/0000-0002-9955-3785](http://orcid.org/0000-0002-9955-3785)

**References**

1. Forterre P, Grosjean H. The interplay between RNA and DNA modifications. Back to the RNA world. In DNA and RNA Modification Enzymes: Structure, Mechanism, Function and Evolution, H Grosjean, Ed. Landes Bioscience, 2009; pp 259-274.
2. Weigele P, Raleigh EA. Biosynthesis and function of modified bases in Bacteria and their viruses. Chemical Rev 2016; 116(20):12655-87; PMID:27319741; [https://doi.org/10.1021/acs.chemrev.6b00114](https://doi.org/10.1021/acs.chemrev.6b00114)
3. Grosjean H. Nucleic Acids are not boring long polymers of only four types of nucleotides. In DNA and RNA Modification Enzymes: Structure, Mechanism, Function and Evolution, H Grosjean, Ed. Landes Bioscience, 2009; pp 1-18.
4. Warren RA. Modified bases in bacteriophage DNAs. Ann Rev Microbiol 1980; 34:137-58; PMID:7002022; [https://doi.org/10.1146/annurev.mi.34.100180.001033](https://doi.org/10.1146/annurev.mi.34.100180.001033)
5. Perche-Letuve P, Mollé T, Forouhar F, Mulliez E, Atta M. Wybutosine biosynthesis: structural and mechanistic overview. RNA Biol 2014; 11:1508-18; PMID:25629788; [https://doi.org/10.4161/15476286.2014.992271](https://doi.org/10.4161/15476286.2014.992271)
6. Gilbert WV, Bell TA, Schaening C. Messenger RNA modifications: Form, distribution, and function. Science 2016; 352:1408-1412; PMID:27320337; [https://doi.org/10.1126/science.aad8711](https://doi.org/10.1126/science.aad8711)
7. Dominissini D, Nachtergaele S, Moschitch-Moshkovitz S, Peer E, Kol N, Ben-Haim MS, Dai Q, Di Segni A, Salmon-Divon M, Clark WC, et al. The dynamic N1-methyladenosine methylome in eukaryotic messenger RNA. Nature 2016; 530:441-446; PMID:26863196; [https://doi.org/10.1038/nature16998](https://doi.org/10.1038/nature16998)
20. Manickam N, Nag N, Abbasi A, Patel K, Farabaugh PJ. Studies of
21. Manickam N, Joshi K, Bhatt MJ, Farabaugh PJ. Effects of tRNA modi-
22. Zallot R, Brochier-Armanet C, Gaston KC, Forouhar F, Limbach PA,
23. Thibessard A, Borges F, Fernandez A, Gintz B, Decarisis B, Leblond-
24. Thibessard A, Borges F, Fernandez A, Gintz B, Decarisis B, Leblond-
25. Durand JM, Dagberg B, Uhlén BE, Björk GR. Transfer RNA modifica-
26. Marchetti M, Capela D, Poincloux R, Benmeradji N, Aurica MC, Le Ru
27. Vinayak M, Pathak C. Queuosine modification of tRNA: its divergent
28. Marks T, Forkas WR. Effects of a diet deficient in tyrosine and
29. Rakovich T, Boland C, Bernstein I, Chikwana VM, Iwata-Reuyl D,
30. Zaborske JM, Bauer DuMont JL, Wallace EWJ, Pan T, Aquadro CF,
31. M
32. Gregson JM, Crain PF, Edmonds CG, Gupta R, Hashizume T, Phillip-
33. Kawamura T, Hirata A, Ohno S, Nomura Y, Nagano T, Nameki N,
34. Reader JS, Metzgar D, Schimmel P, de Cr
35. Blaby IK, Phillips G, Blaby-Haas CE, Gulig KS, El Yacoubi B, de
36. Reader JS, Metzgar D, Schimmel P, de Cr
37. Ohno S, Nomura Y, Nagano T, Nameki N, Yokogawa T, Hori H. Multisite-speci
38. Gregson JM, Crain PF, Edmonds CG, Gupta R, Hashizume T, Phillip-
39. Reader JS, Metzgar D, Schimmel P, de Cr
40. Kawamura T, Hirata A, Ohno S, Nomura Y, Nagano T, Nameki N,
41. Meier F, Suter B, Grosjean H, Keith G, Kubli E. Queuosine modi
42. Rubio MAT, Pastar I, Gaston KW, Ragone FL, Janzen CJ, Cross GAM,
43. van den Born E, Vagbo CB, Songe-Moller L, Leihne V, Lien GF, Leszc-
44. Fedeles BI, Singh V, Delaney JC, Li D, Essigmann JM. The AlkB family
45. Yang YG, et al. N6-methyladenosine in nuclear RNA is a major sub-
46. Barbier N. Identi
47. Zallot R, Brochier-Armanet C, Gaston KC, Forouhar F, Limbach PA,
48. Marchetti M, Capela D, Poincloux R, Benmeradji N, Aurica MC, Le Ru
49. Matsuo S. Novel mechanism of post-transcriptional modi
50. Marchetti M, Capela D, Poincloux R, Benmeradji N, Aurica MC, Le Ru
51. Van den Born E, Vagbo CB, Songe-Moller L, Leihne V, Lien GF, Leszc-
52. Barbier N. Identi
53. Zallot R, Brochier-Armanet C, Gaston KC, Forouhar F, Limbach PA,
54. Marchetti M, Capela D, Poincloux R, Benmeradji N, Aurica MC, Le Ru
55. Matsuo S. Novel mechanism of post-transcriptional modi
56. Barbier N. Identi
57. Zallot R, Brochier-Armanet C, Gaston KC, Forouhar F, Limbach PA,
58. Marchetti M, Capela D, Poincloux R, Benmeradji N, Aurica MC, Le Ru
59. Matsuo S. Novel mechanism of post-transcriptional modi
60. Barbier N. Identi
61. Zallot R, Brochier-Armanet C, Gaston KC, Forouhar F, Limbach PA,
62. Marchetti M, Capela D, Poincloux R, Benmeradji N, Aurica MC, Le Ru
63. Matsuo S. Novel mechanism of post-transcriptional modi
64. Barbier N. Identi
65. Zallot R, Brochier-Armanet C, Gaston KC, Forouhar F, Limbach PA,
66. Marchetti M, Capela D, Poincloux R, Benmeradji N, Aurica MC, Le Ru
67. Matsuo S. Novel mechanism of post-transcriptional modi
68. Barbier N. Identi
69. Zallot R, Brochier-Armanet C, Gaston KC, Forouhar F, Limbach PA,
70. Marchetti M, Capela D, Poincloux R, Benmeradji N, Aurica MC, Le Ru
71. Matsuo S. Novel mechanism of post-transcriptional modi
72. Barbier N. Identi
73. Zallot R, Brochier-Armanet C, Gaston KC, Forouhar F, Limbach PA,
74. Marchetti M, Capela D, Poincloux R, Benmeradji N, Aurica MC, Le Ru
75. Matsuo S. Novel mechanism of post-transcriptional modi
76. Barbier N. Identi
77. Zallot R, Brochier-Armanet C, Gaston KC, Forouhar F, Limbach PA,
78. Marchetti M, Capela D, Poincloux R, Benmeradji N, Aurica MC, Le Ru
79. Matsuo S. Novel mechanism of post-transcriptional modi
80. Barbier N. Identi
81. Zallot R, Brochier-Armanet C, Gaston KC, Forouhar F, Limbach PA,
82. Marchetti M, Capela D, Poincloux R, Benmeradji N, Aurica MC, Le Ru
83. Matsuo S. Novel mechanism of post-transcriptional modi
84. Barbier N. Identi
85. Zallot R, Brochier-Armanet C, Gaston KC, Forouhar F, Limbach PA,
86. Marchetti M, Capela D, Poincloux R, Benmeradji N, Aurica MC, Le Ru
87. Matsuo S. Novel mechanism of post-transcriptional modi
88. Barbier N. Identi
89. Zallot R, Brochier-Armanet C, Gaston KC, Forouhar F, Limbach PA,
90. Marchetti M, Capela D, Poincloux R, Benmeradji N, Aurica MC, Le Ru
91. Matsuo S. Novel mechanism of post-transcriptional modi
92. Barbier N. Identi
93. Zallot R, Brochier-Armanet C, Gaston KC, Forouhar F, Limbach PA,
94. Marchetti M, Capela D, Poincloux R, Benmeradji N, Aurica MC, Le Ru
95. Matsuo S. Novel mechanism of post-transcriptional modi
96. Barbier N. Identi
97. Zallot R, Brochier-Armanet C, Gaston KC, Forouhar F, Limbach PA,
98. Marchetti M, Capela D, Poincloux R, Benmeradji N, Aurica MC, Le Ru
99. Matsuo S. Novel mechanism of post-transcriptional modi
100. Barbier N. Identi
101. Zallot R, Brochier-Armanet C, Gaston KC, Forouhar F, Limbach PA,
102. Marchetti M, Capela D, Poincloux R, Benmeradji N, Aurica MC, Le Ru
103. Matsuo S. Novel mechanism of post-transcriptional modi
104. Barbier N. Identi
105. Zallot R, Brochier-Armanet C, Gaston KC, Forouhar F, Limbach PA,
106. Marchetti M, Capela D, Poincloux R, Benmeradji N, Aurica MC, Le Ru
107. Matsuo S. Novel mechanism of post-transcriptional modi
108. Barbier N. Identi
109. Zallot R, Brochier-Armanet C, Gaston KC, Forouhar F, Limbach PA,
110. Marchetti M, Capela D, Poincloux R, Benmeradji N, Aurica MC, Le Ru
111. Matsuo S. Novel mechanism of post-transcriptional modi
112. Barbier N. Identi
113. Zallot R, Brochier-Armanet C, Gaston KC, Forouhar F, Limbach PA,
114. Marchetti M, Capela D, Poincloux R, Benmeradji N, Aurica MC, Le Ru
115. Matsuo S. Novel mechanism of post-transcriptional modi
116. Barbier N. Identi
117. Zallot R, Brochier-Armanet C, Gaston KC, Forouhar F, Limbach PA,
118. Marchetti M, Capela D, Poincloux R, Benmeradji N, Aurica MC, Le Ru
119. Matsuo S. Novel mechanism of post-transcriptional modi
120. Barbier N. Identi
Zinc-independent folate biosynthesis: genetic, biochemical, and structural investigations reveal new metal dependence for GTP cyclohydrolase I. J Bacteriol 2009; 191:6936-49; PMID:19764255; https://doi.org/10.1128/JB.00287-09

41. McCarty RM, Somogyi Ard, Bandarian V. Escherichia coli QueD is a 6-Carboxy-5,6,7,8-tetrahydropterin synthase. Biochemistry 2009; 48:2301-2303; PMID:19231875; https://doi.org/10.1021/bi801013p

42. Phillips G, Grochowski LL, Benetto S, Xu H, Bailly M, Babyn-Haas C, El Yazoubi B, Iwata-Reuyl D, White RH, de Crey-Lagard V. Functional promiscuity of the COG0720 family. ACS Chem Biol 2012; 7:197-209; PMID:21999246; https://doi.org/10.1021/cb200329f

43. Miles ZD, Roberts SA, McCarty RM, Bandarian V. Biochemical and structural studies of 6-Carboxy-5,6,7,8-tetrahydropterin synthase reveal the molecular basis of catalytic promiscuity within the Tunnel-fold superfamily. J Biol Chem 2014; 289:23641-23652; PMID:24990950; https://doi.org/10.1074/jbc.M114.535560

44. Zallot R, Harrison K, Kolaczkowski B, de Crey-Lagard V. Functional annotations of paralogs: a blessing and a curse. Life 2016; 6:39; PMID:27618105; https://doi.org/10.3390/life6030039

45. Bandarian V, Drennan CL. Radical-mediated ring contraction in the mechanism. Chem Bio Chem 2005; 6:1926-1939; PMID:16206323; https://doi.org/10.1002/cbic.200500063

46. Biela I, Tidten-Luksh N, Immekus F, Glinca S, Nguyen TXP, Gerber H-D, Heine A, Klebe G, Reuter K. Investigation of specificity determinants in bacterial tRNA-guanine transglycosylase reveals queuine, the substrate of its eucaryotic counterpart, as inhibitor. PLoS ONE 2013; 8:e66420; PMID:23704982; https://doi.org/10.1371/journal.pone.0064240

47. Romier C, Reuter K, Suck D, Fincher R. Crystal structure of tRNA-guanine transglycosylase: RNA modification by base exchange. Embo J 1996; 15:2850-2857; PMID:8654383

48. Phillips G, Grochowski LL, Benetto S, Xu H, Bailly M, Babyn-Haas C, El Yazoubi B, Iwata-Reuyl D, White RH, de Crey-Lagard V. Functional promiscuity of the COG0720 family. ACS Chem Biol 2012; 7:197-209; PMID:21999246; https://doi.org/10.1021/cb200329f

49. Miles ZD, Roberts SA, McCarty RM, Bandarian V. Biochemical and structural studies of 6-Carboxy-5,6,7,8-tetrahydropterin synthase reveal the molecular basis of catalytic promiscuity within the Tunnel-fold superfamily. J Biol Chem 2014; 289:23641-23652; PMID:24990950; https://doi.org/10.1074/jbc.M114.535560

50. Phillips G, Chikwana VM, Maxwell A, El-Yacoubi B, Swairjo MA, Iwata-Reuyl D, Kondo H, Sekine M, Okada N, Nishimura S, et al. Crystal structure of archaeosome tRNA-guanine transglycosylase. J Mol Biol 2002; 318:665-77; PMID:12054814; https://doi.org/10.1002/smb.22836(02)0090-0

51. Miles ZD, McCarty RM, Molnar G, Bandarian V. Discovery of epoxy-queuino (oQ) reductase reveals parallels between halorespiration and tRNA modification. Proc Natl Acad Sci USA 2011; 108:7368-7372; PMID:21502530; https://doi.org/10.1073/pnas.101863108

52. Ishitani R, Nureki O, Fukui S, Kijimoto T, Nameki N, Watanabe M, Kondo H, Sekine M, Okada N, Nishimura S, et al. Crystal structure of archaeosome tRNA-guanine transglycosylase. J Mol Biol 2002; 318:665-77; PMID:12054814; https://doi.org/10.1002/smb.22836(02)0090-0

53. Auerbach G, Herrmann A, Bracher A, Bader G, Gutlich M, Fischer M, Neukamm M, Garrido-Franco M, Richardson J, Nr H, et al. Zinc plays a key role in human and bacterial GTP cyclohydrolase I. Proc Natl Acad Sci USA 2000; 97:13567-72; PMID:11087827; https://doi.org/10.1073/pnas.101863108

54. Phillips G, Swairjo MA, Gaston KW, Bailly M, Limbach PA, Iwata-Reuyl D, de Crey-Lagard V. Diversity of archaeosome synthesis in crenarchaeota. ACS Chem Biol 2012; 7:300-5; PMID:22032275; https://doi.org/10.1021/cb200361w

55. Miles ZD, McCarty RM, Molnar G, Bandarian V. Discovery of epoxy-queuino (oQ) reductase reveals parallels between halorespiration and tRNA modification. Proc Natl Acad Sci USA 2011; 108:7368-7372; PMID:21502530; https://doi.org/10.1073/pnas.101863108

56. Phillips G, Chikwana VM, Maxwell A, El-Yacoubi B, Swairjo MA, Iwata-Reuyl D, de Crey-Lagard V. Discovery and characterization of an amidinotransferase involved in the modification of archaeal tRNA. J Biol Chem 2010; 285:12706-12713; PMID:20129918; https://doi.org/10.1074/jbc.M110.102236

57. Phillips G, Swairjo MA, Gaston KW, Bailly M, Limbach PA, Iwata-Reuyl D, de Crey-Lagard V. Divergence of archaeosome synthesis in crenarchaeota. ACS Chem Biol 2012; 7:300-5; PMID:22032275; https://doi.org/10.1021/cb200361w

58. Mei X, Alvarez J, Ramos AB, Samanta U, Iwata-Reuyl D, Swairjo MA. Crystal structure of the archaeosome synthase QueF-Like - insights into amidino transfer and tRNA Recognition by the tunnel fold. Proteins 2016; 25884661; https://doi.org/10.3390/nut7042897

59. Fergus C, Barnes D, Aqlasem M, Kelly V. The Queoine micronutrient: charting a course from microbe to man. Nutrients 2015; 7:2897; PMID:25884661; https://doi.org/10.3390/nut7042897

60. Boland C, Hayes P, Santa-Maria I, Nishimura S, Kelly VP. Queoine formation in eukaryotic tRNA occurs via a mitochondria-localized heteromeric transglycosylase. J Biol Chem 2009; 284:18218-18227; PMID:19414587; https://doi.org/10.1074/jbc.M109.002477

61. Biela I, Tidten-Luksh N, Immekus F, Glinca S, Nguyen TX, Gerber HD, Heine A, Klebe G, Reuter K. Investigation of specificity determinants in bacterial tRNA-guanine transglycosylase reveals queuine, the substrate of its eucaryotic counterpart, as inhibitor. PLoS ONE 2013; 8:e66420; PMID:23704982; https://doi.org/10.1371/journal.pone.0064240

62. de Crey-Lagard V, Olson G. RNA modification subsystems in the SEED database In DNA and RNA editing enzymes: comparative structure, mechanism, functions, cellular interactions and evolution, H Grosjean, Ed. Landes Bioscience: Austin, 2009; pp 624-228.
archaeal tRNA, proceeds via a pathway involving base replacement on the tRNA polynucleotide chain. J Biol Chem 1997; 272:20146-51; PMID:9242689; https://doi.org/10.1074/jbc.272.32.20146

74. Rodionov DA, Hebbeln P, Eudes A, ter Beek J, Rodionova IA, Erkens GB, Slotboom DJ, Gelfand MS, Osterman AL, Hanson AD, et al. A novel class of modular transporters for vitamins in prokaryotes. J Bacteriol 2009; 191:42-51; PMID:18931129; https://doi.org/10.1128/JB.01208-08

75. Zallot R, De Crécy Lagard V. Comparative genomics of bacterial queuosine salvage identify COG1738 as a preQ0 transporter. Biomolecules 2017, submitted.

76. Iyer LM, Zhang D, Maxwell Burroughs A, Aravind L. Computational identification of novel biochemical systems involved in oxidation, glycosylation and other complex modifications of bases in DNA. Nucleic Acids Res 2013; 41:7635-7655; PMID:23814188; https://doi.org/10.1093/nar/gkt573

77. Sabri M, Häuser R, Ouellette M, Liu J, Dehbi M, Moeck G, García E, Titz B, Uetz P, Moineau S. Genome annotation and intraviral interactome for the Streptococcus pneumoniae virulent Phage Dp-1. J Bacteriol 2011; 193:551-562; PMID:21097633; https://doi.org/10.1128/JB.01117-10

78. Pedulla ML, Ford ME, Houtz JM, Karthikeyan T, Wadsworth C, Lewis JA, Jacobs-Sera D, Falbo J, Gross J, Pannunzio NR, et al. Origins of highly mosaic Mycobacteriophage genomes. Cell 2003; 113:171-182; PMID:12705866; https://doi.org/10.1016/S0092-8674(03)00233-2

79. Kulikov E, Golomidova A, Letarov M, Kostryukova E, Zelenin A, Prokhorov N, Letarov A. Genomic sequencing and biological characteristics of a novel Escherichia coli bacteriophage 9g, a putative representative of a new Siphoviridae genus. Viruses 2014; 6:5077; PMID:25533657; https://doi.org/10.3390/v6125077

80. Moreno Witt AI, den Bakker HC, Cummings CA, Rodriguez-Rivera LD, Govoni G, Raneiri ML, Degoricija L, Brown S, Hoelzer K, Peters JE, et al. Identification and characterization of novel Salmonella mobile elements involved in the dissemination of genes linked to virulence and transmission. PLoS One 2012; 7:e41247; PMID:22911766; https://doi.org/10.1371/journal.pone.0041247

81. Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, Edwards RA, Gerdes S, Parrello B, Shukla M, et al. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). Nucleic Acids Research 2014; 42:D206-14; PMID:24293654; https://doi.org/10.1093/nar/gkt1126

82. Wang L, Chen S, Xu T, Taghizadeh K, Wishnok JS, Zhou X, You D, Deng Z, Dedon PC. Phosphorothioation of DNA in bacteria by ddt genes. Nat Chem Biol 2007; 3:709-710; PMID:17934475; https://doi.org/10.1038/nchembio.2007.39

83. He W, Huang T, Tang Y, Liu Y, Wu X, Chen S, Chan W, Wang Y, Liu X, Chen S, et al. Regulation of DNA phosphorothioate modification in Salmonella enterica by DndB. Scientific Reports 2015; 5:12368; PMID:26190504; https://doi.org/10.1038/srep12368

84. Gorbalenya AE, Koonin EV. Endonuclease (R) subunits of type-I and type-III restriction-modification enzymes contain a helicase-like domain. FEBS Letters 1991; 291:277-81; PMID:1657645; https://doi.org/10.1016/0014-5793(91)81301-N

85. Carstens AB, Kot W, Hansen LH. Complete genome sequences of four novel Escherichia coli bacteriophages belonging to new phage groups. Genome Announc 2015; 3; PMID:26184932; https://doi.org/10.1128/genomeA.00741-15

86. McCarty RM, Bandarian V. Deciphering deazapurine biosynthesis: pathway for pyrrolopyrimidine nucleosides toyoacamycin and sangivamycin. Chem & Biol 2008; 15:790-8; https://doi.org/10.1016/j.chembiol.2008.07.012

87. Hurt JK, Olgen S, Garcia GA. Site-specific modification of Shigella flexneri virF mRNA by tRNA-guanine transglycosylase in vitro. Nucleic Acids Res 2007; 35:4905-13; PMID:17626052; https://doi.org/10.1093/nar/gkm473

88. Nonekowsk ST, Kung FL, Garcia GA. The Escherichia coli tRNA-guanine transglycosylase can recognize and modify DNA. J Biol Chem 2002; 277:7178-82; PMID:11751936; https://doi.org/10.1074/jbc.M111077200