Subsystem: Archaeosine and queuosine biosynthesis. (discovering missing genes and pathways).

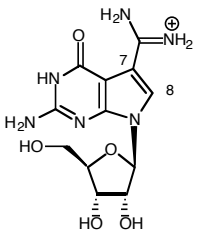
Valérie de Crécy-Lagard,[1] and Dirk Iwata-Reuyl[2]
[1]Department of Microbiology and Department of Microbiology and Cell Science, University of Florida, P.O. Box 110700, Gainesville, FL 32611-0700. [2]Department of Chemistry, Portland State University, PO Box 751, Portland, OR 97207

I. Introduction

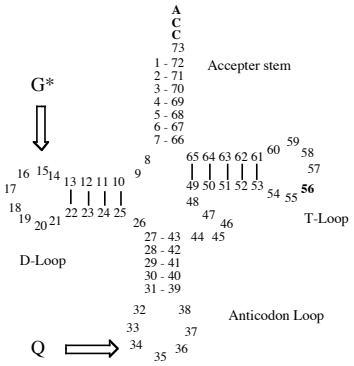
Comparative genomics can be used not only to find missing enzymes of known pathways but also to discover novel pathways. One such example described below is the discovery of the pathways leading to the synthesis and incorporation of the modified bases of tRNA Queuosine and Archaeosine (G*).

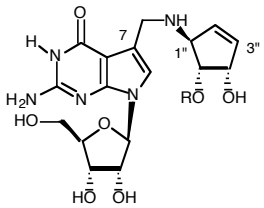
Queuosine (Q) and its derivatives occur exclusively in Bacteria and Eukaryotes at position 34 (the wobble position) in the anticodons of tRNAs coding for the amino acids asparagine, aspartic acid, histidine, and tyrosine1 . Archaeosine (G*) is present only in Archaea, where it is found in the majority of tRNA species, specifically at position 15 in the dihydrouridine loop (D-loop) 2, a site not modified in any tRNA outside of the archaeal domain.

Subsystem diagram including the list and abbreviations of functional roles and pathway intermediates is provided in Figure 1. A representative section of subsystem spreadsheet is shown in Figure 2 (modified from the full display available in SEED). Brief notes and comments on some of the revealed problems and conjectures are provided in Section II “Subsystem Notes”. Section III contains a summary of pathway discovery illustrating the use of comparative genomics



Archaeosine (G*)





Queuosine (Q) R = H
Mann Q R = β-D-mannose
Gal Q R = β-D-galactose

II. Subsystem notes

Subsystem variants:

The discovery of the missing Q/G* genes allowed us to project the encoded subsystem over a variety of genomes and to analyze the different biologically relevant variants.

- The signature enzyme of the pathway is TGT. Several organisms, such as *S. cerevisiae* and Mycoplasma, lack TGT (variant -1) in agreement with the well-known absence of queuosine 22 in their tRNA.
- Most Bacteria such as *E. coli* contain the Q-de novo pathway (Variant 211: 1 or 2,3,4,5,6, 7, 9)
- Some bacteria have only the preQ1 salvage pathway (Variant 011)
- Most Archaea have the G* de novo pathway (Variant 120), but some have just the preQ0 salvage pathway (Variant 020)
- Most eukaryotes have the q (queuine) salvage pathway (variant 010) This variant is also found in some bacteria suggesting that in these organisms the TGT enzymes exchange the q-base (like eukaryotes) and not the preQ1-base (like most bacteria).

Variant codes: "XXX"

First number: {0}, no preQ0/preQ1 biosynthesis; {1} preQ0 biosynthesis; {2} preQ1 biosynthesis.

Second number: : {0}, no tgt, {1}, bacterial/eukaryotic tgt; {2}, archaeal tgt

Third number: {0}, no queA; {1} queA present.

Variant "-1" no pathway

Variant "0" unresolved

Open questions, missing genes and gene candidates.

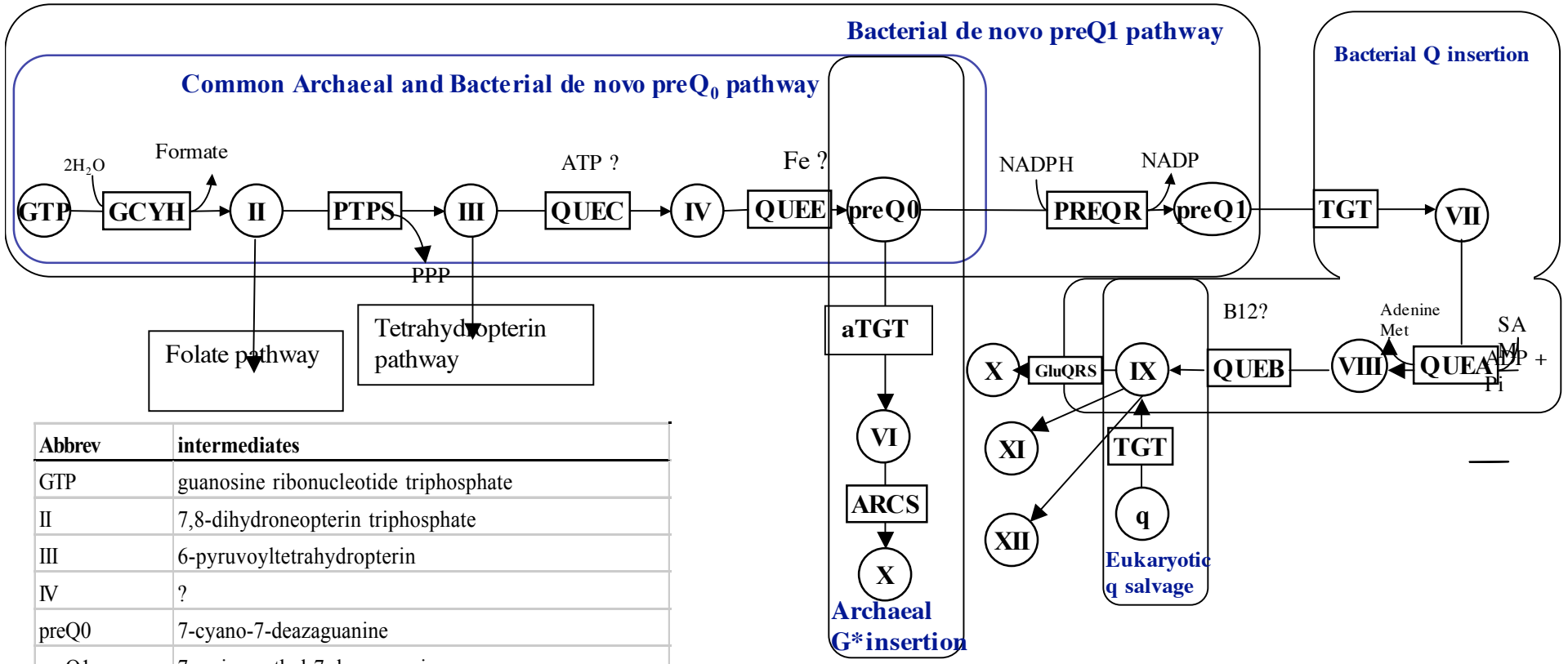
Two genes are still missing for the respective last steps of Q and G* biosynthesis.

Nothing is know about transporters of the pathway but transporters for the q-base must be present in eukaryots and some bacteria, as well as transporters for preQ1 or preQ0 in organisms that have only the bacterial salvage pathway.

Figure 1. Subsystem diagram

Subsystem: Archaeosine and queuosine biosynthesis

Queuosine and Archaeosine Biosynthesis



Abbrev	intermediates
GTP	guanosine ribonucleotide triphosphate
II	7,8-dihydroneopterin triphosphate
III	6-pyruvoyltetrahydropterin
IV	?
preQ0	7-cyano-7-deazaguanine
preQ1	7-aminomethyl-7-deazaguanine
VI	preQ0-tRNA15
VII	preQ1-tRNA34
VIII	epoxyqueuosine-tRNA
IX	queuosine-tRNA
X	archaeosine-tRNA
q	queuine
XI	glutamyl-Q-tRNA
XII	mannosyl-Q-tRNA

Abbrev	Functional Role
GCYHI1	GTP cyclohydrolase I (EC 3.5.4.16) type 1
GCYHI2	GTP cyclohydrolase I (EC 3.5.4.16) type 2
PTPS	6-pyruvoyl tetrahydrobiopterin synthase (EC 4.2.3.12)
QUEC	QueC ATPase
QUEE	QueE Radical SAM
PREQR	NADPH dependent preQ0 reductase
TGT	tRNA-guanine transglycosylase (EC 2.4.2.29)
ATGT	archaeosine tRNA-ribosyltransferase (EC 2.4.2.-)
QUEA	S-adenosylmethionine:tRNA ribosyltransferase-isomerase (EC 5.-.-.-)
QUEB	B12 dependent reductase
ARCS	formamidine synthase

Figure 2. Subsystem sprteadsheet (fragment)

Subsystem: Archaeosine and queuosine biosynthesis

biosynthesis of		$preQ_0$					preQ1	Q		G*	Glu-Q
Organism	Variant Code	GCYHI1	GCYHI2	PTPS	queC	queE	PREQR	qTGT	QUEA	aTGT	GluQRS
Saccharomyces cerevisiae [E]	-1	2304									
Corynebacterium diphtheriae NCTC 13129 [B]	010	1923						232			233
Homo sapiens [E]	010	398		549				13168			
Lactobacillus plantarum WCFS1 [B]	011	2687						1902	1903		
Rhodobacter capsulatus SB1003 [B]	011		4355					3598	2487		
Ferroplasma acidarmanus [A]	120	1041		1042	1680	1040				1306, 1817	
Halobacterium sp. NRC-1 [A]	120		1638	2489, 974	2487	2488				1682, 1683, 505	
Bacillus anthracis str. Ames [B]	211	1411		1246	5	7	1248	2	4293		
Escherichia coli K12 [B]	211	2128		2721	441	3	2750	403	402		144
Staphylococcus aureus NCTC 8325 [B]	211		2486	408	409	407	2279	1070	1071		

III. Summary and a current status of the pathway discovery project

The biosynthesis of Q was only partially understood when we began this analysis. Whole organism incorporation experiments established that GTP is the probable primary precursor in the biosynthesis of queuosine [3]. The common intermediate in the queuosine and archaeosine pathway is 7-cyano-7-deazaguanine ($preQ_0$) [4].

In bacteria $preQ_0$ undergoes reduction to 7-aminomethyl-7-deazaguanine ($preQ_1$) which is subsequently inserted into the tRNA by the enzyme tRNA-guanine transglycosylase (TGT), a reaction in which the genetically encoded base (guanine) is eliminated [5, 6]. The remainder of queuosine biosynthesis occurs at the level of the tRNA, and involves the construction of an epoxycyclopentandiol ring [7-9] by the S-adenosylmethionine:tRNA ribosyltransferase-isomerase (EC 5.-.-.-) (QueA) to give epoxyqueuosine (oQ), followed by an apparent B12-dependent step in which the epoxide in oQ is reduced to give queuosine [10].

In higher eukaryotes, a mannosyl-group or galactosyl-group is further added on the cyclopentene diol of Q-$tRNA^{Asp}$ and Q-$tRNA^{Tyr}$, respectively, by as yet uncharacterized specific glycosyl-Q transferase(s). Recently, it was shown that a family of enzymes similar to glutamyl-tRNA synthetases glutamylates Q of $tRNA^{Asp}$.(see [11] for review)

Only Bacteria are capable of *de novo* queuosine biosynthesis. Eukaryotes acquire queuosine as a nutrient factor and/or from the intestinal flora1, and insert queuine, the free base of queuosine, directly into the appropriate tRNAs [12] by a eukaryotic TGT.

In Archaea, $preQ_0$ is the substrate for an archaeosine tRNA-ribosyltransferase (aTGT, EC 2.4.2.-) [13, 14]. The formation of archaeosine can then in principle occur through the formal addition of ammonia to the nitrile of $preQ_0$ after incorporation into the polynucleotide.

Only three genes of the pathway have been previously identified. The *tgt* gene and *queA* genes of *E. coli* [15, 16] and the archaeal *tgt* family [13, 14]. We have classified archaeal TGT homologs in three subfamilies, one not containing a PUA domain (type 1), another, containing a PUA domain (type 2), and the third one, one containing just the PUA domain (type 3). Additional analysis is required to decipher functional roles of these subfamilies.

Predicting the preQ1 pathway by comparative genomics.

A combination of phylogenetic occurrence, clustering on the chromosome and biochemical knowledge led to the hypothesis that the *ykvJKLM* genes of *B. subtilis* are involved in Q biosynthesis. These candidate genes were experimentally tested using an Acinetobacter ADP1 model [17]. tRNA from all four *Acinetobacter ykvJ,K,L,M* mutants lacked Q 18. Homologs of YkvJKL are found in most Archaea, and we propose that these genes are involved in the synthesis of $preQ_0$. YkvM is specific to bacteria, and while sequence homology suggested that this enzyme catalyzed GTP cyclohydrolase-like chemistry, our biochemical and genetic data clearly established that YkvM is not a GTP cyclohydrolase, but instead catalyzes the reduction of $preQ_0$ to $preQ_1$, and thus represents a new class of oxido-reductase that carries out the unprecedented reduction of a nitrile group to a primary amine [19].

All the experimental evidence generated on the biosynthesis of queuosine and other 7-deazapurine natural products point to a GTP cyclohydrolase(GCYHI) or cyclohydrolase-like reaction as the first step in the biosynthesis. While we demonstrated that YkvM was not the expected cyclohydrolase enzyme, functional coupling analysis performed on the *folE* gene encoding GTP cyclohydrolase I showed that it clustered with the *ykvJKLM* genes. The analysis of co-distribution of the *ykvJKL* and *folE* genes shows that many organisms containing both, *ykvJKL* genes and folate biosynthesis genes (*folBKCA*), lack a *folE* homolog. This observation suggests that another protein family is catalyzing the same reaction in these organisms. By combining phylogenetic occurrence profiles and chromosomal clustering analysis, a candidate for the missing gene family (COG1469) was identified. We are currently testing the hypothesis that *folE* is involved in Q synthesis, and that COG1469 represents an alternative GCYHI.

The *ykvK* family (COG0720) has been annotated as 6-pyruvoyl-tetrahydropterin synthase (PTPS) involved in tetrahydropterine (BH4) biosynthesis in higher animals [20]. BH4 is not found in most bacteria, and the physiological role of members of this family in *E. coli* or *B. subtilis* is unknown. Recently, the *E. coli ygcM* homolog was shown to encode an enzyme having PTPS activity (8.7% of the mammal counterpart). [21]. Our finding that a *ΔykvK* mutant is deficient in queuosine biosynthesis, suggests that YkvK is the first dedicated step of $preQ_0$ biosynthesis. Our current working hypothesis for the biosynthesis of $preQ_0$ requires the 4 enzymes, FolE, YkvK (PTPS), YkvJ, and YkvL. We propose that, following the conversion of GTP to 6-pyruvoyltetrahydropterin by FolE and YkvK, YkvJL catalyze the conversion of 6-pyruvoyltetrahydropterin to $preQ_0$ via a still unknown intermediate.

References.

1. Kersten, H.; Kersten, W., Biosynthesis and Function of Queuine and Queuosine tRNAs. In Chromatography and Modification of Nucleosides Part B, ed.; Gehrke, C. W.; Kuo, K. C. T., 'Ed.'^'Eds.' Elsevier: Amsterdam, 1990; 'Vol.' p^pp B69-B108.
2. Sprinzl, M.; Dank, N.; Nock, S.; Schon, A., Compilation of tRNA Sequences and Sequences of tRNA Genes. Nuc. Acids Res. 1991, 19, (Suppl.), 2127-2171.
3. Kuchino, Y.; Kasai, H.; Nihei, K.; Nishimura, S., Biosynthesis of the Modified Nucleoside Q in Transfer RNA. Nucleic Acids Research 1976, 3, 393-398.
4. Okada, N.; Noguchi, S.; Nishimura, S.; Ohgi, T.; Goto, T.; Crain, P. F.; McCloskey, J. A., Structure Determination of a Nucleoside Q Precursor Isolated from E. coli tRNA: 7-(aminomethyl)-7-deazaguanosine. Nucleic Acids Research 1978, 5, 2289-2296.
5. Okada, N.; Noguchi, S.; Kasai, H.; Shindo-Okada, N.; Ohgi, T.; Goto, T.; Nishimura, S., Novel Mechanism of Post-transcriptional Modification of tRNA. J Biol Chem 1979, 254, (8), 3067-3073.
6. Okada, N.; Noguchi, S.; Nishimura, S.; Ohgi, T.; Goto, T.; Crain, P. F.; McCloskey, J. A., Structure Determination of a Nucleoside Q Precursor Isolated from E. coli tRNA: 7-(aminomethyl)-7-deazaguanosine. Nucleic Acids Res 1978, 5, 2289-2296.

7. Kinzie, S. D.; Thern, B.; Iwata-Reuyl, D., Mechanistic studies of the tRNA-modifying enzyme QueA: a chemical imperative for the use of AdoMet as a "ribosyl" donor. Organic Letters 2000, 2, (9), 1307-1310.
8. Slany, R. K.; Bosl, M.; Crain, P. F.; Kersten, H., A New Function of S-Adenosylmethionine: The Ribosyl Moiety of AdoMet Is the Precursor of the Cyclopentenediol Moiety of the tRNA Wobble Base Queuine. Biochemistry 1993, 32, 7811-7817.
9. Slany, R. K.; Bosl, M.; Kersten, H., Transfer and isomerization of the ribose moiety of AdoMet during the biosynthesis of queuosine tRNAs, a new unique reaction catalyzed by the QueA protein from Escherichia coli. Biochimie 1994, 76, (5), 389-93.
10. Frey, B.; McCloskey, J. A.; Kersten, W.; Kersten, H., New Function of Vitamin B12: Cobamide-Dependent Reduction of Epoxyqueuosine to Queuosine in tRNAs of Escherichia coli and Salmonella typhimurium. J Bacteriol 1988, 170, (5), 2078-2082.
11. Grosjean, H.; de Crecy-Lagard, V.; Bjork, G. R., Aminoacylation of the anticodon stem by a tRNA-synthetase paralog: relic of an ancient code? Trends Biochem Sci 2004, 29, (10), 519-22.
12. Shindo-Okada, N.; Okada, N.; Ohgi, T.; Goto, T.; Nishimura, S., Transfer Ribonucleic Acid Guanine Transglycosylase Isolated from Rat Liver. Biochemistry 1980, 19, 395-400.
13. Bai, Y.; Fox, D. T.; Lacy, J. A.; Van Lanen, S. G.; Iwata-Reuyl, D., Hypermodification of tRNA in Thermophilic archaea. Cloning, overexpression, and characterization of tRNA-guanine transglycosylase from Methanococcus jannaschii. Journal of Biological Chemistry 2000, 275, (37), 28731-8.
14. Watanabe, M., et al., Biosynthesis of Archaeosine, a Novel Derivative of 7-Deazaguanosine Specific to Archaeal tRNA, Proceeds via a Pathway Involving Base Replacement of the tRNA Polynucleotide Chain. J. Biol. Chem. 1997, 272, (32), 20146-20151.
15. Noguchi, S.; Nishimura, Y.; Hirota, Y.; Nishimura, S., Isolation and Characterization of an Escherichia coli Mutant Lacking tRNA-Guanine-Transglycosylase. Journal of Biological Chemistry 1982, 257, (11), 6544-6550.
16. Reuter, K.; Slany, R.; Ullrich, F.; Kersten, H., Structure and Organization of E. coli Genes Involved in Biosynthesis of the Deazaguanine Derivative Queuine, a Nutrient Factor for Eukaryotes. J Bacteriol 1991, 173, (7), 2256-2264.
17. Metzgar, D.; Bacher, J. M.; Pezo, V.; Reader, J.; Doring, V.; Schimmel, P.; Marliere, P.; Crecy-Lagard, V. d., Acinetobacter sp. ADP1: an ideal model organism

for genetic analysis and genome engineering. Nucleic Acids Research 2004, in press.

18. Reader, J. S.; Metzgar, D.; Schimmel, P.; de Crecy-Lagard, V., Identification of four genes necessary for biosynthesis of the modified nucleoside queuosine. J Biol Chem 2004, 279, (8), 6280-5.
19. Van Lanen, S. G.; Reader, J. S.; Swairjo, M. A.; de Crecy-Lagard, V.; Lee, B.; Iwata-Reuyl, D., From cyclohydrolase to oxidoreductase: Discovery of nitrile reductase activity in a common fold. Proc Natl Acad Sci U S A 2005, 102, (12), 4264-9.
20. Thony, B.; Auerbach, G.; Blau, N., Tetrahydrobiopterin biosynthesis, regeneration and functions. Biochem J 2000, 347 Pt 1, 1-16.
21. Woo, H. J.; Hwang, Y. K.; Kim, Y. J.; Kang, J. Y.; Choi, Y. K.; Kim, C. G.; Park, Y. S., Escherichia coli 6-pyruvoyltetrahydropterin synthase ortholog encoded by ygcM has a new catalytic activity for conversion of sepiapterin to 7,8-dihydropterin. FEBS Lett 2002, 523, (1-3), 234-8.
22. Katze, J. R.; Basile, B.; McCloskey, J. A., Queuine, a modified base incorporated posttranscriptionally into eukaryotic