



# Alarmones as Vestiges of a Bygone RNA World

Ricardo Hernández-Morales<sup>1</sup> · Arturo Becerra<sup>1</sup> · Antonio Lazcano<sup>1,2</sup>

Received: 8 November 2018 / Accepted: 15 December 2018 / Published online: 2 January 2019  
© Springer Science+Business Media, LLC, part of Springer Nature 2019

## Abstract

All known alarmones are ribonucleotides or ribonucleotide derivatives that are synthesized when cells are under stress conditions, triggering a stringent response that affects major processes such as replication, gene expression, and metabolism. The ample phylogenetic distribution of alarmones (e.g., cAMP, Ap(n)A, cGMP, AICAR, and ZTP) suggests that they are very ancient molecules that may have already been present in cellular systems prior to the evolutionary divergence of the Archaea, Bacteria, and Eukarya domains. Their chemical structure, wide biological distribution, and functional role in highly conserved cellular processes support the possibility that these modified nucleotides are molecular fossils of an epoch in the evolution of chemical signaling and metabolite sensing during which RNA molecules played a much more conspicuous role in biological catalysis and genetic information.

**Keywords** Alarmones · Ribonucleotide derivatives · Signaling molecules · RNA World

## Introduction

All organisms are endowed with sensory systems that allow them to respond to different environmental chemical signals. The quorum sensing signaling system is a cell-to-cell communication process mediated by molecules known as autoinducers (e.g., acylated homoserine lactones, small peptides, furanosyl borate diester). Modifications in the autoinducer concentration by variations in density population trigger intracellular signaling cascades (Bharati and Chatterji 2013). Some components involved in these processes are in fact intra- and extracellular second messengers, which have been defined as alarmones. Alarmones is a term first used by Stephens et al. (1975) to describe signaling molecules that are ribonucleotides or have a chemical structure similar to a ribonucleotide, such as ppGpp or cAMP, which activate or inhibit different cellular processes during stress

conditions. Variations in alarmones concentration levels alter gene expression or metabolism changes that allow organisms to adapt to environmental insults. Production of alarmones is increased when cells face a wide range of stress conditions, including amino acid, carbon, carbohydrate, and phosphate starvation, as well as limitation of fatty acids or metallic ions, changes in temperature, salinity, pH, oxidative stress, and the presence of antibiotics (Supplementary Table S1). Increase in alarmones concentrations triggers a response known as stringent response or control (Stent and Brenner 1961), that regulates different key cellular processes such as replication, transcription, translation, and metabolism, allowing organisms to switch from a growth mode to a survival one or, in some cases, promoting the appearance of flagella that allow migration to more favorable environments (Pesavento and Hengge 2009; Bharati and Chatterji 2013).

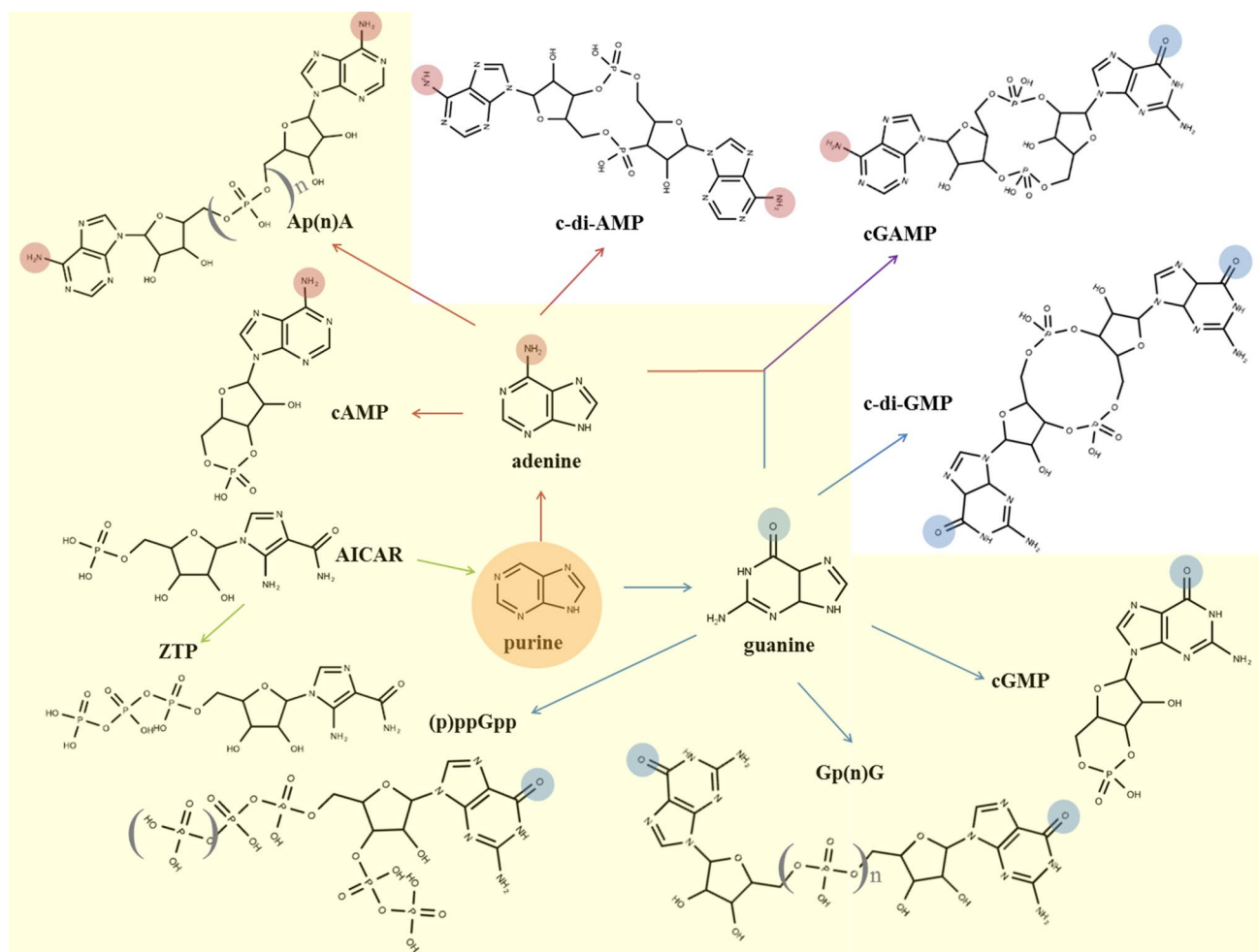
The chemical structures of all known alarmones include a purine moiety (Fig. 1). As summarized by Nelson and Breaker (Breaker 2010; Nelson and Breaker 2017), this bias can be explained by the fact that purines (a) are less soluble than pyrimidines, favoring their associations with RNA polymers; (b) contain a fused two-ring system and stack better than pyrimidines, allowing them a more favorable interaction with the biological receptors; and (c) are preferred by artificial ribozymes that recognize cyclic mononucleotides. However, other factors may be involved in this bias. For instance, it has been argued that delocalizable  $\pi$  electrons of

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s00239-018-9883-3>) contains supplementary material, which is available to authorized users.

✉ Antonio Lazcano  
alar@ciencias.unam.mx

<sup>1</sup> Facultad de Ciencias, Universidad Nacional Autónoma de México, Cd. Universitaria, 04510 Mexico City, Mexico

<sup>2</sup> Miembro de El Colegio Nacional, Donceles 104, Centro Histórico, 06000 Mexico City, Mexico



**Fig. 1** Alarmones divided in four main groups. The first group, adenine-based alarmones, is formed by cyclic AMP (cAMP), cyclic diadenylate (c-di-AMP), and diadenosine polyphosphates (Ap(n)A). All of them pointed with a red arrow. The second group is alarmones that have a guanine ring in their chemical structure, such as cyclic GMP (cGMP), cyclic diguanylate (c-di-GMP), guanosine (penta)-tetraphosphate ((p)ppGpp), diguanosine polyphosphates (Gp(n)G), which are pointed with a blue arrow. A third group, which is pointed with a green arrow, is formed by metabolic intermediates in the biosyn-

thesis of purines, such as the aminoimidazole carboxamide ribotide (AICAR), or derived from purines, such as aminoimidazole carboxamide ribotide triphosphate (ZTP). The last group is formed by joining guanine and adenine purines (pointed with a purple arrow), producing the cyclic guanosine monophosphate-adenosine monophosphate (cGAMP) alarmones. Alarmones with a universal cellular distribution are highlighted in yellow background. Red and blue circles represent the chemical differences between purines and pyrimidines, respectively. (Color figure online)

heterocyclic conjugated molecules could have played a key role in the selection of the chemical composition of living systems (Pullman 1972). As underlined by Pullman (1972), purines have a resonance energy higher than pyrimidines. This makes purines more stable, with adenine being the most stable of all the biochemically significant bases, and the one that exhibits the greatest value of resonance energy per  $\pi$  electron, which makes it more resistant to radiation damage. Therefore, it is reasonable to assume that this may have led to its selection as part of the basic skeleton of many compounds, including high energy molecules (e.g., ATP), coenzymes (e.g., NAD, FAD, CoA), alarmones (e.g., cAMP), and some antibiotics (e.g., lysylaminoadenosine).

Alarmones can be classified according to their nucleotide composition and the cyclic or lineal chemical linkage they form (Nelson and Breaker 2017) or, as shown in Fig. 1, by their chemical structure. In the latter case, four main groups can be recognized (Fig. 1): (a) adenine-ribonucleotide group; (b) guanine-ribonucleotide group; (c) alarmones derived from purine anabolism intermediates; and (d) those formed by guanine and adenine producing a cyclic dinucleotide. The adenine-alarmones group includes the well-known cyclic AMP (cAMP), together with cyclic diadenylate (c-di-AMP) and diadenosine polyphosphates (Ap(n)A). Alarmones endowed with a guanine moiety in their chemical structure include cyclic GMP (cGMP), cyclic diguanylate

(c-di-GMP), guanosine (penta-)tetraphosphate ((p)ppGpp), and diguanosine polyphosphates (Gp(n)G). A small third group is formed by intermediates of purine metabolism, which includes aminoimidazole carboxamide ribonucleotide (AICAR) and aminoimidazole carboxamide riboside triphosphate (ZTP). Finally, the last group is formed by the cyclic guanosine monophosphate-adenine monophosphate alarmone (cGAMP). Other modified purine [e.g., Xp(n)X] and pyrimidine ribonucleotides [e.g., cUMP, cCMP, Up(n)U] have been described, but their role in signaling cascades is poorly understood. The agonist effect against some uridine dinucleoside polyphosphates receptors has demonstrated that they affect proliferation, differentiation, phagocytosis, secretion, cell adhesion, and migration (Pendergast et al. 2001; Seifert et al. 2011). This suggests that these molecules could perhaps also function as alarmones or regulators (Nelson and Breaker 2017).

The key role of alarmones in the regulation of cellular processes, together with their biological distribution in the three major biological lineages, has been interpreted as evidence that signal mechanisms were already present in the last common ancestor (LCA) (i.e., the last universal common ancestor or LUCA) of all extant life forms (Lazcano et al. 2011). As reported here, analyses of the phylogenetic distribution of alarmone biosynthetic enzymes strongly support this possibility. Alarmone components (Fig. 1) are easily formed in prebiotic simulation experiments (Oró 1960; Ferris et al. 1978; Ritson and Sutherland 2012; Rios and Tor 2013), and are present in carbonaceous chondrites (Copper et al. 2001; Callahan et al. 2011), suggesting that they may have appeared during earlier stages of cellular evolution when RNA molecules and ribonucleotide derivatives played a much more conspicuous role in replication and metabolism (Breaker 2010; Lazcano et al. 2011; Lazcano 2014, 2018; Nelson and Breaker 2017). It has been argued that ribonucleotidyl coenzymes may also be vestiges of early stages of cellular evolution (Orgel 1968; Orgel and Sulston 1971) as discussed here: the same appears to be true for alarmones, and they appear to represent the oldest recognizable system of cell signaling involved in intra- and extracellular communication, environmental sensing, and interaction with organisms of the same and different species (Breaker 2010; Lazcano et al. 2011; Nelson and Breaker 2017).

## Results

### Phylogenetic Distribution of Biosynthetic Enzymes of Alarmones in Completely Sequenced Cellular Genomes

The chemical nature of alarmones cannot be used to reconstruct their evolutionary relationships, but the history of

these signal molecules can be understood in part by the sequence comparison and phylogenetic distribution of the enzymes involved in their biosynthesis and degradation. Therefore, search for homologs of biosynthetic and degradative enzymes of alarmones was carried out in completely sequenced cellular genomes (see “Materials and Methods”). Our results are summarized in Supplementary Table S2 and Fig. 2. Supplementary Table S2 displays the distribution of homologous enzymes (denoted as H) involved in the biosynthesis and in the degradation of alarmones. Letters in each column of the table represent the acronym for each organism’s genome from KEGG database. Their phylogenetic distribution, based on an rRNA evolutionary tree, is shown in Fig. 2.

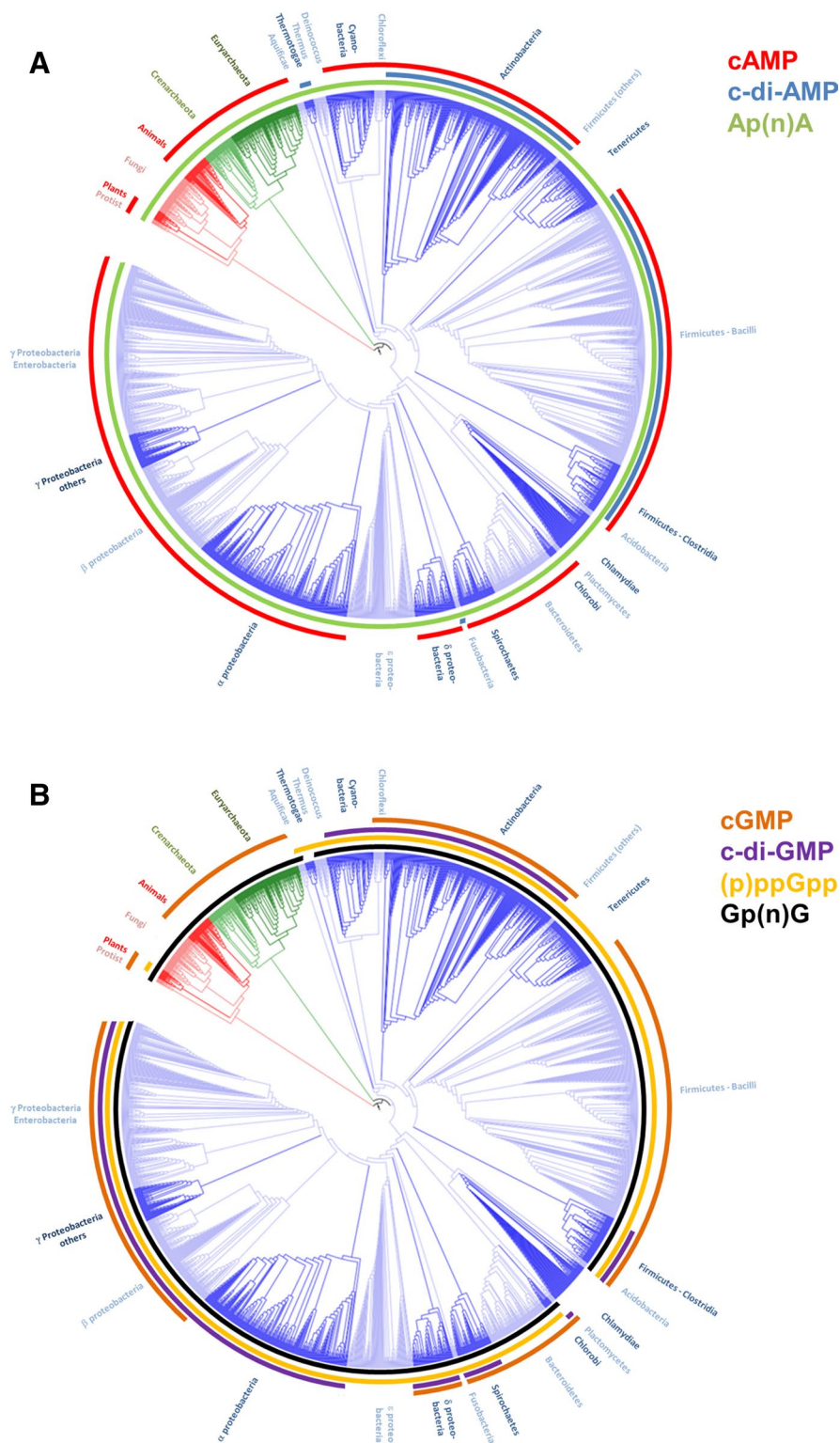
As shown in Fig. 2, the biosyntheses of five groups of alarmones were analyzed: (a) adenine-based alarmones (cAMP, c-di-AMP, and Ap(n)A); (b) alarmones that have a guanine ring (cGMP, c-di-GMP, (p)ppGpp, and Gp(n)G); (c) alarmones formed by intermediates in the purine metabolism (AICAR and ZTP); (d) alarmones formed by guanine and adenine purines in a cyclic structure (cGAMP); and (e) alarmone-like molecules (Xp(n)X, Up(n)U, cCMP, and cUMP). In Fig. 2, the color line drawn outside the tip of the branches in each clade of the phylogenetic trees shows the presence of alarmone biosynthetic enzymes. The enzymes were considered to be present in each clade if (1) more than fifty percent of the branches exhibited homologous hits, or if (2) such hits were distributed through the main basal branches of each clade.

As shown in the phylogenetic trees in Fig. 2, enzymes involved in biosynthesis of cAMP, Ap(n)A, cGMP, Gp(n)G, AICAR, and ZTP have a universal distribution, while enzymes involved in the synthesis of c-di-AMP, c-di-GMP, and (p)ppGpp are clearly restricted to the Bacteria domain. The absence of homologs in the Chlamydia and Tenericutes clades in most phylogenies can be understood as a secondary loss due to their intracellular parasitic lifestyle.

Since sequences of biosynthetic enzymes of Xp(n)X, Up(n)U, cCMP, and cUMP have not been reported, only their degradative enzymes were analyzed. All of them are present in almost all cellular lineages.

It has recently been proposed that other cyclic dinucleotides such as cyclic guanosine monophosphate-adenosine monophosphate (cGMP-AMP or cGAMP) (Fig. 2) can also function as an endogenous second messenger. In animals, it triggers the production of type I interferons in response to foreign DNA, e.g., DNA transfection or DVA virus infection (Wu et al. 2013; Sun et al. 2013; Hall et al. 2017), or an efficient intestinal colonization by *Vibrio cholera* (Davies et al. 2012). As shown in Fig. 2, the cGAMP biosynthetic enzyme is distributed in some  $\gamma$ -proteobacteria, essentially in the genus *Vibrio*, as well as in the animal group (Supplementary Table S2; Fig. 2). In our searches, the use of

**Fig. 2** Phylogenetic distribution of enzymes involved in the synthesis of alarmones. Five groups of alarmones were analyzed: **a** cAMP; c-di-AMP; Ap(n)A; **b** cGMP; c-di-GMP; (p)ppGpp; Gp(n)G; **c** AICAR; ZTP; **d** cGAMP; **e** Xp(n)X; Up(n)U; cCMP; cUMP. The outer circles of the phylogenies represent the presence of alarmones in each clade except for the tree E in which are represented the degradative enzymes. Red clade indicates the eukaryal branches, green clade the archaeal branches, and blue clade the bacterial branches. (Color figure online)



*V. cholerae* query sequences detected homologs only in  $\gamma$ -proteobacteria, while animal sequences used as queries found homologs exclusively within the animal group itself. However, tertiary structure comparisons of the two

groups of cGMP-AMP biosynthetic enzymes demonstrated their monophyletic origin (see below). No homologs were detected in Archaea domain using any of the above-mentioned sequences or three-dimensional structures.

Fig. 2 (continued)

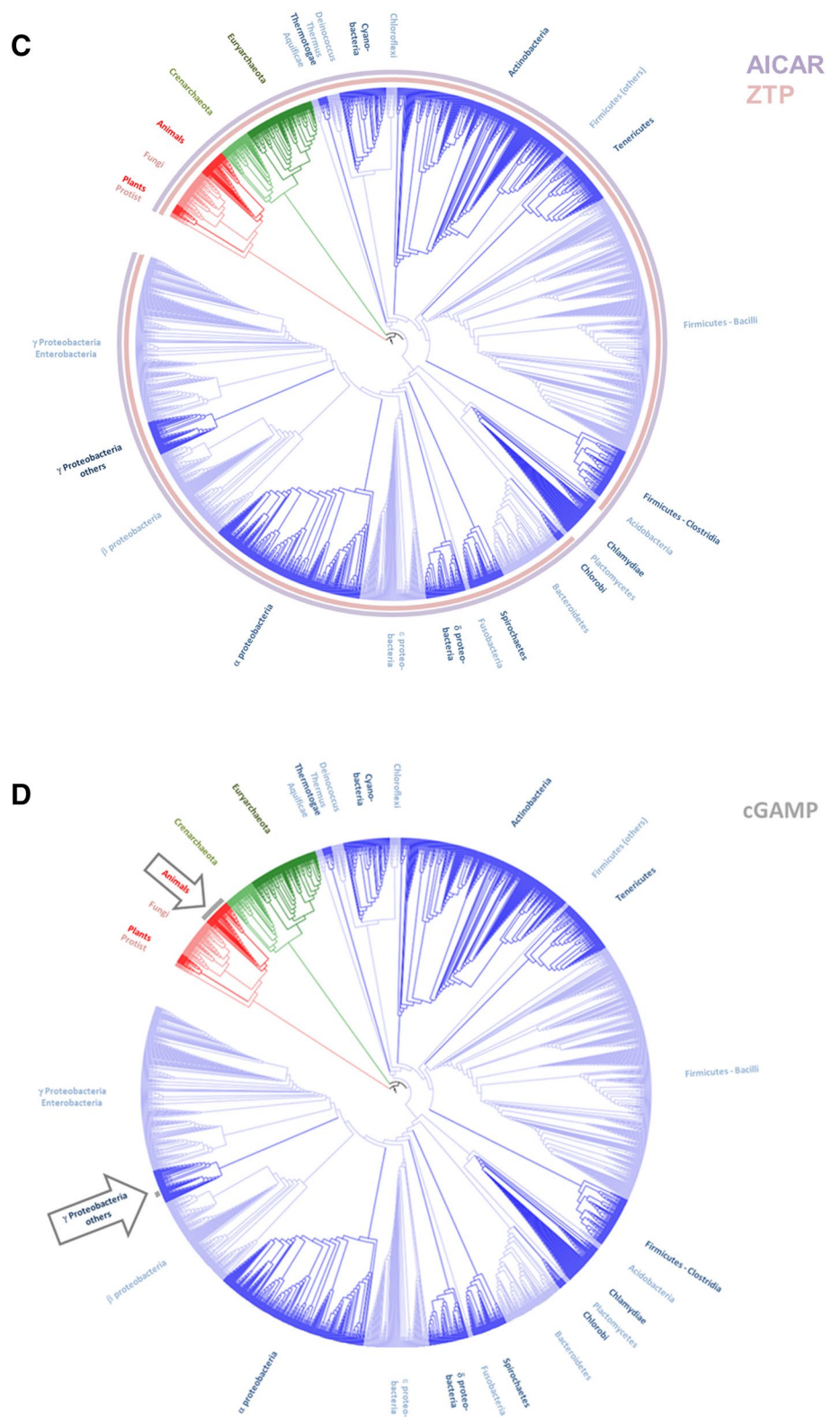
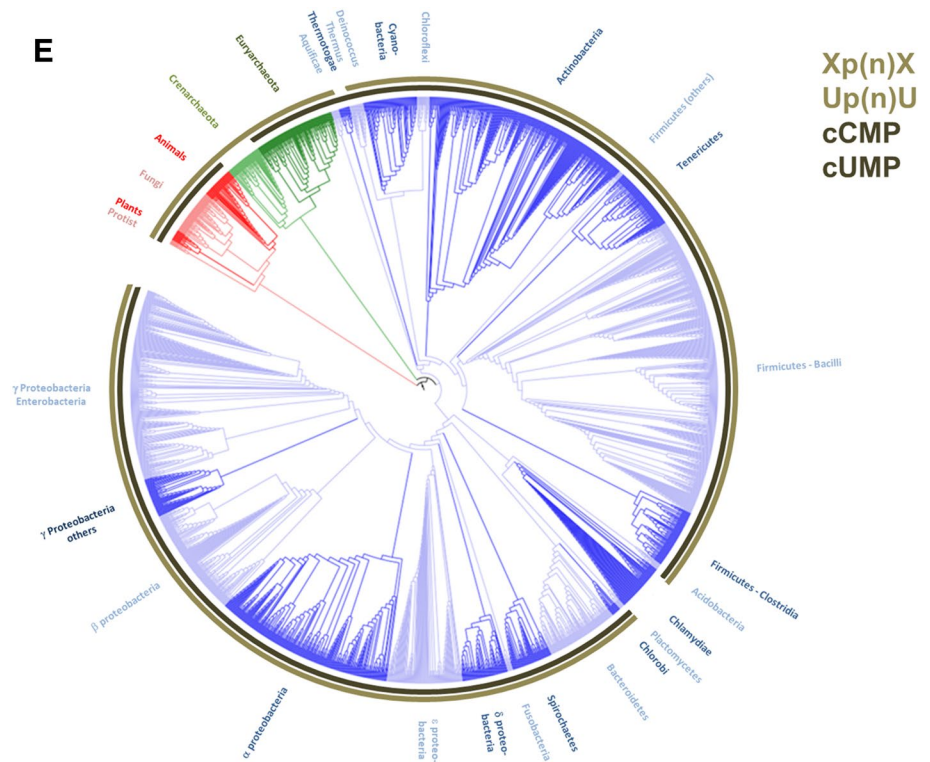


Fig. 2 (continued)

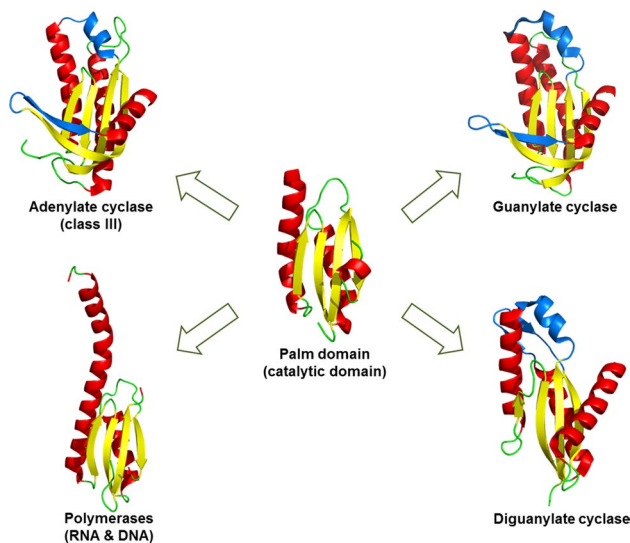


### Alarmones Biosynthetic and Degradative Enzymes and Their Homologous Sequences Present in DNA and RNA Viral Genomes

A search for homologs of alarmones biosynthetic and degradative enzymes was carried out in 5,691 DNA and RNA viral genomes. As shown in Supplementary Table S3, homologous sequences of five enzymes (EC: 3.6.1.29, 3.6.1.61, 3.6.1.17, 3.6.1.41, and 3.1.7.2) involved in the degradation of Ap(n)A, Gp(n)G, Xp(n)X, Up(n)U, and (p)ppGpp, as well as one biosynthetic enzyme (EC: 2.7.6.1) for ZTP, were found in dsDNA viruses that have no RNA stage in their biological cycle. Most of these DNA viruses belong to the Myoviridae and Siphoviridae families, whose hosts are  $\gamma$ -proteobacteria and Firmicutes and Actinobacteria, respectively. The presence of sequences encoding phosphoribosyl pyrophosphate (PRPP) synthetase (EC: 2.7.6.1) in viral genomes is probably due to the fact that it participates in the production of phosphoribosyl pyrophosphate (PRPP), but not of ZTP. No sequences encoding for alarmones biosynthetic and degradative enzymes were found in the RNA viral genomes included in our analysis.

### Structural Similarity Among Adenylate Cyclase (AC), Guanylate Cyclase (GC), Diguanylate Cyclase (DGC), and Polymerase Palm Domain

Cyclic nucleotides, such as cyclic AMP, cyclic GMP and cyclic di-GMP, are highly versatile second messengers that regulate different important cellular processes, including replication, transcription, motility, virulence, biofilm formation, cell cycle progression, and differentiation, among others (see Supplementary Table S1), and they are synthesized by adenylate cyclase (AC), guanylate cyclase (GC), and diguanylate cyclase (DGC), respectively. The structural homology of the catalytic core of ACs class III with the polymerase palm domain of DNA- and viral monomeric RNA polymerases is well established (Artymiuk et al. 1997; Bierger and Essen 2001; Bassler et al. 2018) (Fig. 3). This evolutionary relationship with the catalytic palm domain, which is the oldest and most conserved part of DNA- and monomeric viral RNA polymerases (Jácome et al. 2015), suggests that ACs may have emerged early in evolution of life, most probably during an RNA/protein World. There are other cyclases that synthesize alarmones (e.g., GC, DGC, DAC), whose origin is important to elucidate. As shown in Fig. 3, guanylate and diguanylate cyclases are also homologous with the



**Fig. 3** The palm domain structure found in different enzymes involved in alarmonone biosynthesis and present in nucleotide polymerases. The presence of the palm domain in different enzymes that synthesize alarmonones including adenylate cyclases (class III), guanylate cyclases, and diguanylate cyclases indicates that they are evolutionary related to an ancient nucleotide polymerase. In each structure,  $\alpha$ -helices are colored in red,  $\beta$ -strands in yellow, loops in green;  $\alpha$ -helices and  $\beta$ -strands colored in blue indicate structural segments not present in the palm domain. (Color figure online)

palm domain, and may be the outcome of ancient gene duplications and recruitment events (Supplementary Figure S1). The palm domain, which is constituted by four  $\beta$ -strands and three  $\alpha$ -helices, is similar both in element number and topology to the catalytic domain of different alarmonone biosynthetic enzymes, including adenylate cyclases (class III), guanylate cyclases, and diguanylate cyclases (Bierger and Essen 2001). In fact, nucleotide polymerization and the biosyntheses of cAMP, cGMP, and c-di-GMP are chemically equivalent processes and involve a nucleophilic attack from the 3'OH group of the ribose to the  $\alpha$ -phosphate of a nucleotide 5'-triphosphate, with the elimination of pyrophosphate (Artymiuk et al. 1997). In both cases, two conserved aspartate residues and the requirement of a metal cofactor, usually  $Mg^{2+}$  or  $Mn^{2+}$ , are necessary for catalysis.

The structural comparison between the palm domain and cGAMP synthase did not indicate high levels of primary structure similarity. However, tertiary structure comparisons (Supplementary Figure S2) demonstrate that the catalytic site of cGAMP synthase is homologous to the non-canonical palm domain present in DNA polymerase III or DNA polymerase  $\beta$  (Bailey et al. 2006; Lamers et al. 2006) (Supplementary Figure S2).

## Discussion

### Universal Distribution of Alarmonone Biosynthetic Routes

Comparisons of completely sequenced cellular genomes has shown that the biosynthetic and degradative enzymes of alarmonones, including cAMP, Ap(n)A, cGMP, AICAR, and ZTP, are widely distributed in all extant life forms (Fig. 2; Table S2). The universal phylogenetic distribution of biosynthetic and degradative enzymes of alarmonones, combined with their key role in the regulation of different conserved cellular processes such as replication, gene expression, and metabolism, suggests that they are indeed very ancient molecules. This supports the possibility that the LCA (LUCA) was already endowed with an elaborated sensory system comparable to that of extant cells that responded to these signal molecules (Lazcano et al. 2011; Becerra et al. 2007). This is consistent with the idea that biochemical pathways of earlier life forms were regulated by chemical signals derived from RNA molecules and ribonucleotides (Nelson and Breaker 2017).

Degradative enzymes of Gp(n)G, Xp(n)X, Up(n)U, cCMP, and cUMP molecules also have an universal distribution (Fig. 2), but the corresponding biosynthetic enzymes of these nucleotides are not detected in the available databases. There are several possible explanations for their absence: (a) metabolic accidents during stress conditions (e.g., cells under heat shock) could lead to accumulation of these unusual nucleotides. Enzymatic degradation of these products would explain the wide phylogenetic distribution of the corresponding enzymes; (b) Gp(n)G, Xp(n)X, Up(n)U, cCMP, and cUMP may be secondary products of enzymatic hydrolysis of RNA molecules; (c) they are synthesized by enzymes of broad substrate specificity; or (d) the degradative enzymes may be part of a salvage mechanism to recycle ribonucleotides.

As shown here, analysis of completely sequenced cellular genomes indicates that some alarmonones have an uneven phylogenetic distribution (e.g., c-di-AMP, c-di-GMP, (p)ppGpp, cGAMP). Comparisons of alarmonone biosynthetic enzymes have also shown that enzymes involved in the biosynthesis of c-di-AMP and c-di-GMP are not found in all completely sequenced cellular genomes. As shown in Fig. 2, sequences encoding for diadenylate cyclases (DACs) and diguanylate cyclases (DGCs) are present in most bacterial genomes. Indeed, most enzymes involved in synthesis and degradation of alarmonones are widely distributed in Bacteria, suggesting that this is the cellular domain that has exploited the widest range of these signal molecules. The same is true for sequences involved in (p)ppGpp biosynthesis. Nevertheless, other studies (Atkinson

et al. 2011) have reported the presence of (p)ppGpp synthase and hydrolase in some members of the archaeal (e.g., *Methanosarcina*, *Natronomonas*, and *Methanococoides*) and eukaryal (e.g., some algae, amoeba, other protists, and fungi) domains. The absence of homologous sequences reported here can be explained in part by the stringent search parameters we have employed, e.g., the percentage involved in the alignment in query and subject sequence. More flexible parameters allowed us to identify only (p) ppGpp hydrolase homologs. However, the robustness in our analysis resulted in a fewer false negatives, while less stringent parameters allowed us to obtain more hits, as well as protein domains or even sequence segments, increasing dramatically the number of false positives. The phylogenetic distribution of the cGAMP synthase in some  $\gamma$ -proteobacteria (e.g., *Vibrio cholera*) and in animals (Supplementary Table S2; Fig. 2) suggests that cyclic GMP-AMP alarmone is a very recent second messenger, and that the corresponding biosynthetic enzyme was horizontally transferred. Although *V. cholerae* and animal cGAMP synthase sequences have low similarity levels, their conserved tertiary structure suggest their common origin.

The ubiquitous distribution of AC class III and the apparent absence of lateral gene transport strongly support previous suggestions that cAMP was already present in the LCA (Lazcano et al. 2011). The wide variety of domains associated with the cyclase catalytic domain could reflect diverse strategies by which the additional domains could regulate different cellular processes (see ref. Bassler et al. 2018; Shenoy et al. 2004; Baker and Kelly 2004; Shenoy and Visweswariah 2004) (Supplementary Table S4). The evolutionary history of genes involved in cAMP synthesis appears to have been shaped by shuffling domain and gene duplication (Lazcano et al. 2011). As argued by Lazcano and coworkers (2011) and Nelson and Breaker (2017), cAMP and other alarmones can be considered “metabolic fossils” from early stages of biological evolution when ribonucleotides and RNA molecules played a more conspicuous role in chemical signaling, metabolite sensing, and storage of genetic information and catalysis.

The presence of the palm domain in different enzymes that synthesize alarmones, including adenylate cyclases (class III), guanylate cyclases, and diguanylate cyclases, indicates that they are all descended from a protein domain shared with an ancient nucleotide polymerase ancestor, which was duplicated and recruited many times during the evolutionary history of polymerases and cyclases (Fig. 3; Supplementary Figure S1). Although the canonical palm domain is not an homolog of cGAMP synthase, the structural comparison between the DNA polymerase X palm domain and the cGAMP synthase showed high levels of

similarity, suggesting the monophyletic origin of these two catalytic structures (Supplementary Figure S2).

A detailed bioinformatic search for alarmone biosynthetic and degradative sequences in DNA and RNA viral genomes also revealed that some double-stranded DNA bacteriophages encode sequences involved in alarmone synthesis or degradation (Supplementary Table S3). Homologous sequences of alarmone degradation were found in these same viral genomes. A possible explanation is that viruses control the intracellular alarmone concentration via degradation process as a strategy that allow them to remain hidden from other non-infected cells of the bacterial population.

## Alarmones and the RNA World

There are many different definitions of the RNA World, but all of them are based on the hypothesis that during early evolution of life replication and catalysis were mediated primarily by RNA molecules (Gilbert 1986). However, the RNA World should not be understood as a mere collection of catalytic and replicative polynucleotides (Lazcano 2014; Vázquez-Salazar and Lazcano 2018), but rather as a stage during which the interaction of RNA molecules, small molecules such as ribonucleotidyl cofactors (Handler 1961; Eakin 1963; Orgel 1968; White 1976), metallic ions, lipids, and small and simple peptides allowed the maintenance and evolution of a primitive replicative and metabolic apparatus (Lazcano 2014, 2018). Proposals of an RNA World are supported by the stunning expansion of the functional repertoire of synthetic ribozymes (Chen et al. 2007), which catalyze the same classes of chemical reactions as enzymes (Table 1), as well as by the direct participation of ribonucleotide cofactors and modified nucleotides in biological catalysis (White 1982; Lazcano 2014), the regulation of gene expression by riboswitches and other non-coding small RNAs (Breaker 2010; Lazcano 2014; Chen and Gottesman 2014; Nelson and Breaker 2017; Vázquez-Salazar and Lazcano 2018; González-Plaza 2018), and, as noted here, the activation or inhibition of cellular processes by alarmones.

Although in principle some cyclic deoxyribonucleotides (e.g., cyclic dAMP) could function as alarmones, this is not the case (Nelson and Breaker 2017). The fact that all known alarmones are derived from ribonucleotides but not from deoxyribonucleotides is certainly consistent with the possibility that at least some of them are remnants of an epoch prior to the evolutionary emergence of DNA cellular genomes (Lazcano et al. 2011; Nelson and Breaker 2017). How did they first evolve? Did alarmones first emerge in the RNA World? If so, what was their original role? Or did they first appear in the RNA/protein World? Did RNA/protein-based entities simply take advantage of what was already there? There are several alternative possibilities. Perhaps the simplest explanation is that they were by-products of the



**Table 1** Ribozyme-mediated catalysis

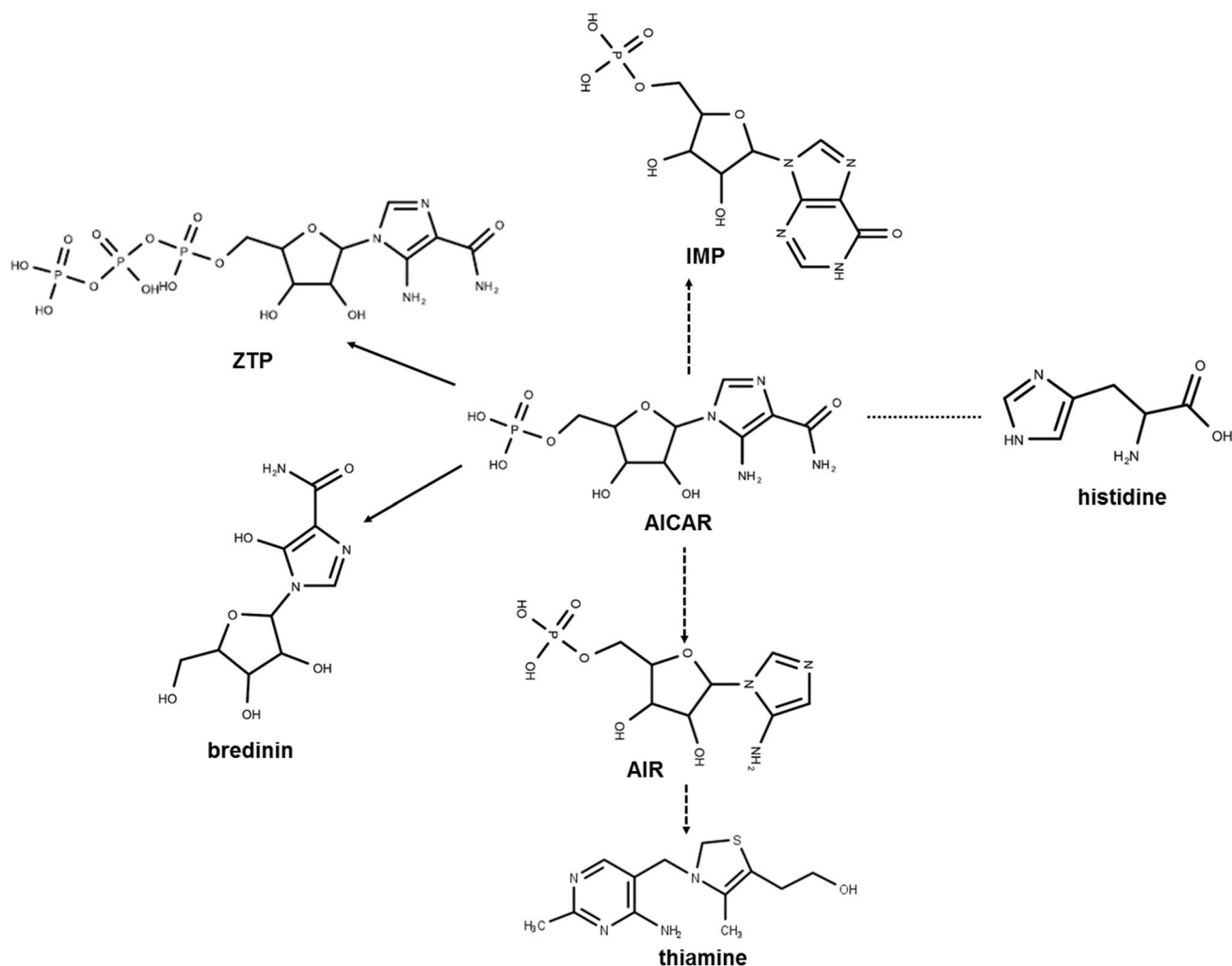
Classes	Enzymes	Ribozymes
EC1: oxidoreductases	Dehydrogenases Oxidases, peroxidases Reductases Monoxygenases Dioxygenases	Dehydrogenases Oxidases, peroxidases
EC2: transferases	C1-transferases Glycosyltransferases Aminotransferases Phosphotransferases	Methyltransferases Transketolase Aminoacyltransferases Pentosyltransferases Alkyltransferases Phosphotransferases Nucleotidyltransferases
EC3: hydrolases	Esterases Glycosidases Peptidases Amidases	Esterases Endodeoxyribonucleases Endoribonucleases Glycosylases Amidases Phosphatases
EC4: lyases (synthases)	C–C lyases C–O lyases C–N lyases C–S lyases	Carboxylases Aldehydelyases DNA photolyase Ferrolyase
EC5: isomerases	Epimerases <i>cis, trans</i> lyases Intramolecular transferases	Epimerases
EC6: ligases (synthetases)	C–C ligases C–O ligases C–N ligases C–S ligases	C–C ligases C–O ligases C–N ligases C–S ligases Phosphoric ester ligases

Ribozyme-mediated catalysis. The first column indicates the nomenclature and classification of enzymes by the reaction they catalyze following the International Union of Biochemistry and Molecular Biology (IUBMB) and Nomenclature Committee. Examples of enzymes are listed on the next column. The last column includes examples of ribozymes that can catalyze the same type of chemical reactions as protein enzymes

degradation and recycling of RNA molecules (Nelson and Breaker 2017), or resulted from side reactions of catalysts of a ribozyme-mediated underground metabolism. Underground metabolism has been defined as those reactions that occur when enzymes (or ribozymes, as discussed here) use chemically similar substrates that form part of endogenous metabolites. Although biochemical reactions are exceptionally precise, molecular errors may confer evolutionary advantages, such as an enhanced metabolic plasticity or the establishment of new pathways (D’Ari and Casadesús 1998). In both cases, the resulting alarmones precursors would be incorporated as signal mechanisms by an exaptation phenomenon prior to the evolutionary divergence of the Bacteria, Archaea, and Eukarya.

The results presented here cannot be extrapolated back in time beyond a period in which ribosome-mediated polypeptide synthesis had already evolved. However, *in vitro* selection experiments have shown that RNA molecules catalyze the formation of modified ribonucleotides such as CoA, NAD,

and FAD (Huang et al. 2000), supporting the possibility that cyclic or polyphosphate nucleotides such as alarmones could have also been synthesized by ribozymes. Since some of the reactions that form alarmones are equivalent to polynucleotide elongation reactions (Supplementary Table S5), the demonstration of ligase and polymerase activity of ribozymes, together with many other RNA-catalyzed chemical reactions, supports the assumption that these signal molecules could have been originally synthesized by catalytic RNA (Supplementary Table S6). Moreover, it has been reported that some (di)nucleotides, whose chemical structure resembles that of some alarmones (e.g., AppA or GppG), can be synthesized using a polyribonucleotide template (Puthenvedu et al. 2015; Majerfeld et al. 2016). It is therefore very likely that moieties of nucleotidyl-alarmones were available in an RNA World and in an RNA/protein World.



**Fig. 4** The AICAR metabolic intermediate and its role in other biological processes. This ribotide is a key intermediate in the biosynthesis of (1) IMP in nucleotide biosynthesis; (2) ATP to synthesize

histidine in amino acid biosynthesis; (3) the thiamine coenzyme; (4) signal molecules such as AICAR itself and AICAR triphosphate (ZTP); or (5) the bredinin antibiotic

### Alarmones as Examples of Exaptations in Early Evolution of Life

The evolutionary history of some alarmones such as 5-aminoimidazole-4-carboxamide ribonucleotide (AICAR) is probably best understood as an exaptation process, in which there was a shift during evolution of the original function of a molecule to another role (Fig. 4). It has been traditionally thought that intermediate forms or states do not have adaptive value by themselves and they can only be recognized when they are preserved in fossil record. However, at a sub-cellular level there are metabolic intermediates that have been recruited and perform several other functions due to their adaptive value. This appears to be the case of AICAR, which is a molecule used in different cellular processes (Fig. 4). AICAR is a metabolic intermediary in the biosynthesis of inosine monophosphate (IMP) in purine anabolism,

and is also formed in histidine anabolism (Vázquez-Salazar et al. 2017, 2018). Moreover, it is also a precursor of 5-aminoimidazole ribonucleotide (AIR) in the biosynthesis of thiamine cofactor (Bazurto et al. 2015), a pathway that is independent of the purine biosynthetic reactions. These three processes are ancient metabolic pathways. Quite surprisingly, AICAR can also function as an alarmon, which suggests that 5-aminoimidazole-4-carboxamide ribonucleotide was perhaps recruited from its original function as an intermediary in purine biosynthesis to a new role as a signal molecule. The triphosphate form of AICAR (ZTP) can also carry out completely different regulatory roles. Additionally, it is well known since the 1970s that the antibiotic bredinin (5'-hydroxy-1-β-ribofuranosyl-1*H*-imidazole-4-carboxamide) is an AICAR derivative (Suhadolnik 1979) that inhibits IMP dehydrogenase, blocking the conversion of IMP to GMP (Sakaguchi et al. 1976). It has been hypothesized that some

antibiotics could have had other more ancient roles such as signaling molecules. As noted by Yim et al. (2007), small concentrations of antibiotics can modulate gene transcription during the interactions of populations with the surrounding organisms. Like alarmones, antibiotics can affect the expression of genes related to virulence, colonization, motility, stress response, and biofilm formation (Romero et al. 2011), suggesting that they might act as signal molecules in natural environments, facilitating intra- or interspecies interactions within microbial communities.

## Alarmones and Riboswitches

As discussed in considerable detail by Nelson and Breaker (2017), alarmones can also regulate genes through riboswitches. A riboswitch is an RNA segment located in non-coding regions, which binds to a metabolite and controls the gene expression by changing the tridimensional structure of the mRNA, which results in the regulation of transcription or translation of the mRNA in which it is embedded. The discovery of riboswitches that regulate ribozyme activity (Winkler et al. 2004) led to the proposal this could be in fact an ancient control mechanism (Breaker 2010, 2012; Nelson and Breaker 2017). Since RNA can recognize by itself metabolites and control gene expression without the assistance of protein factors, this suggests that some riboswitches and their RNA-based regulatory elements are modern descendants of an ancient regulatory mechanism that evolved prior to the emergence of protein enzymes and receptors (Breaker 2010, 2012; Nelson and Breaker 2017). Both AICAR and ZTP metabolites bind to a conserved non-coding region of *pfl* mRNA and promote the transcription of genes involved in purine and folate biosynthesis (e.g., *purH*, *purJ*, and pyruvate formate lyase), i.e., in purine- and folate-deprived cells where they function as genetic “on” switches (Jones and Ferré-D’Amaré 2015). In addition, c-di-GMP and cAMP-GMP nucleotidyl metabolites can be sensed by GEMM, a conserved RNA motif that is characteristic of some riboswitches that control the expression of genes associated with cell differentiation, virulence, pilus formation, flagellum biosynthesis, and extracellular electron transfer in Bacteria (Sudarsan et al. 2008; Breaker 2012; Kellenberger et al. 2015). It has also been shown that c-di-AMP can be sensed by the riboswitch *ydaO*, which is implicated in cell wall metabolism, sporulation, and osmotic stress responses (Nelson et al. 2013). The recent discovery of a riboswitch that binds to the ppGpp alarmone has expanded the list of RNA-based signaling systems (Sherlock et al. 2018). The discovery of alarmones that function as metabolites that bind to riboswitches and control gene expression provides a good model of early stages where catalytic RNAs regulated their expression by sensing alarmone-like metabolites. A good example of a signal-receptor system composed solely

of RNA molecules is the complex c-di-GMP riboswitch and the self-splicing ribozyme from *Clostridium difficile* (Nelson and Breaker 2017). The presence of the alarmone c-di-GMP and its binding to the riboswitch facilitates the recognition of the 5′ splice site and the splicing reaction that produces a mature mRNA by the ribozyme. During early evolution of life these systems could regulate the ribozyme function, controlling both gene expression and metabolism (Nelson and Breaker 2017).

## Materials and Methods

### Sequences of Enzymes Involved in Synthesis and Degradation of Alarmones

Pathway maps from Kyoto Encyclopedia of Genes and Genomes (KEGG) (Kanehisa and Goto 2000; Kanehisa et al. 2016) were used to identify and collect the enzyme sequences involved in synthesis and degradation of alarmones. Four groups of alarmones were identified: (1) those derived from adenine, such as cAMP, c-di-AMP, and Ap(n)A; (2) those derived from guanine, such as cGMP, c-di-GMP, (p)ppGpp, and Gp(n)G; (3) those derived from intermediaries in purine metabolism, such as AICAR and ZTP; and (4) cyclic GMP-AMP or cGAMP. Additionally, two groups of possible alarmones were also considered: (1) those derived from other purines, such as Xp(n)X; and (2) those derived from a pyrimidine base, such as cCMP, cUMP, and Up(n)U.

### Searching for Biosynthetic and Degradative Enzymes of Alarmones in Completely Sequenced Cellular Genomes

A search for homologs for each enzyme involved in biosynthesis and degradation of alarmones was performed in archaeal ( $n=256$ ), bacterial ( $n=4392$ ), and eukaryal ( $n=401$ ) genomes. All completely sequenced cellular genomes analyzed were downloaded from the KEGG database, including the following archaeal genomes: Euryarchaeota (170), Crenarchaeota (60), Thaumarchaeota (15), Nanoarchaeota (2), Nanohaloarchaeota (1), Micrarchaeota (1), Korarchaeota (1), Bathyarchaeota (2), Lokiarchaeota (1), and other Archaea (3); bacterial genomes: Gammaproteobacteria (1043), Betaproteobacteria (316), Epsilonproteobacteria (157), Deltaproteobacteria (84), Alphaproteobacteria (489), Firmicutes (882), Tenericutes (126), Actinobacteria (547), Cyanobacteria (99), Melainabacteria (1), Chloroflexi (27), Deinococcus (27), Armatimonadetes (2), Terrabacteria (1), Chlamydiae (118), Verrucomicrobia (9), Planctomycetes (16), Kiritimatiellaeota (1), Spirochaetes (77), Acidobacteria (9), Elusimicrobia (3), Fusobacteria (14), Synergistetes (6), Fibrobacteres (2), Gemmatimonadetes (3),

Bacteroidetes (222), Chlorobi (17), Cloacimonetes (1), Aquificae (14), Thermotogae (30), Caldisei (1), Chrysiogenetes (1), Deferribacteres (4), Calditrichaeota (1), Dictyoglomi (2), Nitrospirae (9), Thermodesulfobacteria (4), Saccharibacteria (4), Peregrinibacteria (1), and unclassified bacteria (12); and eukaryal genomes: Animals (160), Plants (73), Fungi (119), and Protists (49). In order to determine the phylogenetic distribution of enzymes involved in the synthesis and degradation of alarmones, the query sequences of alarmones were compared against subject sequences from each genome using DELTA-BLAST program (Boratyn et al. 2012), version 2.2.29+. The parameters used in this work were a BLOSUM45 scoring matrix, an  $E$ -value  $\leq 0.001$ , and an identity percentage  $\geq 20\%$ . In order to obtain homologous sequences and no protein domains or protein segments, we considered the query coverage and subject coverage metric. These metrics are defined as the percent of the query and subject sequence that overlaps with the alignment. A query coverage  $\geq 70\%$  and a subject coverage  $\geq 70\%$  were used.

### Phylogenetic Analysis

Four phylogenetic trees were used to depict the distribution of homologs of alarmone biosynthetic enzymes. Phylogenetic trees based on small subunit of ribosomal RNA in Newick format were visualized with the Interactive Tree of Life (iTOL) program (Letunic and Bork 2016). Phylogenetic trees were grouped according to the chemical structure of alarmones: (A) cAMP; c-di-AMP; Ap(n)A; (B) cGMP; c-di-GMP; (p)ppGpp; Gp(n)G; (C) AICAR; ZTP; (D) cGAMP; and (E) Xp(n)X; Up(n)U; cCMP; cUMP.

The presence of biosynthetic enzymes was considered in each clade if (1) a few homologous hits were scattered in some branches of the clade, or (2) more than fifty percent of branches had homologous hits.

### Structural Comparisons of Polymerase Palm Domain and Different Kinds of Cyclase Enzymes

The tertiary structures of adenylate (1WC0), guanylate (2W01), diadenylate (4YVZ), and diguanylate cyclases (4H54) and palm domain of DNA polymerase (1D8Y) were obtained from Protein Data Bank (PDB) (Berman et al. 2000). The palm domain of the polymerases is homolog to the catalytic domain of adenylate cyclase (Artymiuk et al. 1997). In order to know if other cyclases (GC, DAC, and DGC) also share structural homology with the palm domain, a structural alignment was made overlapping the structures using PyMol CE package.<sup>1</sup> Moreover, their evolutionary

relationships were reconstructed by aligning all structures using PDBFold pairwise alignment (Krissinel and Henrick 2004) and calculating a geometric distance measure for each comparison using Structural Alignment Score (SAS) (Subbiah et al. 1993). A phylogenetic tree was made using Fitch, which is included in the PHYLIP package (Felsenstein 2005).

### Homologs of Biosynthetic and Degradative Enzymes of Alarmones in Viral Genomes

A search for homologs of alarmone biosynthetic and degradative enzymes was carried out in 5,691 viral genomes of 102 DNA and RNA viral families using DELTA-BLAST. The viral families analyzed were Ackermannviridae, Adenoviridae, Alloherpesviridae, Alphaflexiviridae, Alphasatellitidae, Alphetetraviridae, Alvernaviridae, Amalgaviridae, Ampullaviridae, Anelloviridae, Arenaviridae, Arteriviridae, Ascoviridae, Asfarviridae, Aspiviridae, Astroviridae, Bacilladnaviridae, Baculoviridae, Barnaviridae, Benyviridae, Betaflexiviridae, Bicaudaviridae, Bidnaviridae, Birnaviridae, Bornaviridae, Bromoviridae, Caliciviridae, Carmotetraviridae, Caulimoviridae, Chrysovriidae, Circoviridae, Closteroviridae, Coronaviridae, Corticoviridae, Cystoviridae, Dicistroviridae, Endornaviridae, Filoviridae, Fimoviridae, Flaviviridae, Flexiviridae, Fusariviridae, Fuselloviridae, Gammaflexiviridae, Geminiviridae, Genomoviridae, Globuloviridae, Hantaviridae, Hepadnaviridae, Hepeviridae, Herpesviridae, Hypoviridae, Hytrosaviridae, Iflaviridae, Inoviridae, Iridoviridae, Lavidaviridae, Leviviridae, Lipothrixviridae, Luteoviridae, Malacoherpesviridae, Marnaviridae, Marseilleviridae, Megabirnaviridae, Mesoniviridae, Microviridae, Mimiviridae, Mymonaviridae, Myoviridae, Nairoviridae, Nanoviridae, Narnaviridae, Nidovirales, Nodaviridae, Nudiviridae, Nyamiviridae, Orthomyxoviridae, Papillomaviridae, Paramyxoviridae, Partitiviridae, Parvoviridae, Peribunyaviridae, Permutotetraviridae, Phenuiviridae, Phycodnaviridae, Picobirnaviridae, Picornaviridae, Pithoviridae, Plasmaviridae, Pleolipoviridae, Pneumoviridae, Podoviridae, Polycipiviridae, Polydnaviridae, Polyomaviridae, Potyviridae, Poxviridae, Quadriaviridae, Reoviridae, Retroviridae, Rhabdoviridae, Roniviridae, Rudiviridae, Secoviridae, Siphoviridae, Smacoviridae, Solemoviridae, Solinviridae, Sphaerolipoviridae, Sunviridae, Tectiviridae, Togaviridae, Tolecusatellitidae, Tombusviridae, Tospoviridae, Totiviridae, Turriviridae, Tymoviridae, and Virgaviridae. Some relevant parameters included a BLOSUM45 scoring matrix, an  $E$ -value  $\leq 0.001$ , an identity percentage  $\geq 20\%$ , and query and subject coverage  $\geq 70\%$ .

**Acknowledgements** Ricardo Hernández Morales is a doctoral student from the Programa de Doctorado en Ciencias Biológicas, Universidad Nacional Autónoma de México (UNAM). Financial support

<sup>1</sup> The PyMol Molecular Graphics System, Version 2.0 Schrödinger, LLC.

from PAPIIT-UNAM (IN223916) is gratefully acknowledged. We are indebted to Samuel Ponce de León for many insightful discussions, Alberto Vázquez-Salazar for providing useful references for this work, and José Alberto Campillo Balderas, Rodrigo Jácome, and Adriana Benítez for technical support.

## References

- Abranches J, Martinez AR, Kajfasz JK, Chávez V, Garsin DA, Lemos JA (2009) The molecular alarmone (p)ppGpp mediates stress responses, vancomycin tolerance, and virulence in *Enterococcus faecalis*. *J Bacteriol* 191:2248–2256
- Artymiuk PJ, Poirrette AR, Rice DW, Willett P (1997) A polymerase I palm in adenyl cyclase? *Nature* 388:33–34
- Atkinson GC, Tenson T, Haurlyuk V (2011) The RelA/SpoT homolog (RSH) superfamily: distribution and functional evolution of ppGpp synthetases and hydrolases across the tree of life. *PLoS ONE* 6:e23479
- Bailey S, Wing RA, Steitz TA (2006) The structure of *T. aquaticus* DNA polymerase III is distinct from eukaryotic replicative DNA polymerase. *Cell* 126:893–904
- Baker DA, Kelly JM (2004) Structure, function and evolution of microbial adenyl and guanyl cyclases. *Mol Microbiol* 52:1229–1242
- Balodimos IA, Kashket ER, Rapaport E (1988) Metabolism of adenylated nucleotides in *Clostridium acetobutylicum*. *J Bacteriol* 170:2301–2305
- Bassler J, Schultz JE, Lupas AN (2018) Adenylate cyclases: receivers, transducers, and generators of signals. *Cell Signal* 46:135–144
- Battesti A, Bouveret E (2006) Acyl carrier protein/SpoT interaction, the switch linking SpoT-dependent stress response to fatty acid metabolism. *Mol Microbiol* 62:1048–1063
- Bazurto JV, Heitman NJ, Downs DM (2015) Aminoimidazole carboxamide ribotide exerts opposing effects on thiamine synthesis in *Salmonella enterica*. *J Bacteriol* 197:2821–2830
- Becerra A, Delaye L, Islas S, Lazcano A (2007) The very early stages of biological evolution and the nature of the last common ancestor of the three major cell domains. *Annu Rev Ecol Evol Syst* 38:361–379
- Berleman JE, Hasselbring BM, Bauer CE (2004) Hypercyst mutants in *Rhodospirillum centenum* identity regulatory loci involved in cyst cell differentiation. *J Bacteriol* 186:5834–5841
- Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shinyalov IN, Bourne PE (2000) The protein data bank. *Nucleic Acids Res* 28:235–242
- Bharati BK, Chatterji D (2013) Quorum sensing and pathogenesis: role of small signaling molecules in bacterial persistence. *Curr Sci* 105:643–656
- Bierger B, Essen LO (2001) Structural analysis of adenylate cyclases from *Trypanosoma brucei* in their monomeric state. *EMBO J* 20:433–445
- Bochner BR, Ames BN (1982) ZTP (5-amino 4-imidazole carboxamide riboside 5'-triphosphate): a proposed alarmone for 10-formyl-tetrahydrofolate deficiency. *Cell* 29:929–937
- Bochner BR, Lee PC, Wilson SW, Cutler CW, Ames BN (1984) AppppA and related adenylated nucleotides are synthesized as a consequence of oxidation stress. *Cell* 37:225–232
- Bonaventura C, Cashon R, Colacino JM, Hilderman RH (1992) Alteration of hemoglobin function by diadenosine 5',5"-PP1,P4-tetraphosphate and other alarmones. *J Biol Chem* 267:4652–4657
- Boratyn GM, Schäffer AA, Agarwala R, Altschul SF, Lipman DJ, Madden TL (2012) Domain enhanced lookup time accelerated BLAST. *Biol Direct* 7:1–14
- Breaker RR (2010) RNA second messengers and riboswitches: relics from the RNA World? *Microbe* 5:13–20
- Breaker RR (2012) Riboswitches and the RNA world. *Cold Spring Harb Perspect Biol* 4:a003566
- Callahan MP, Smith KE, Cleaves IHJ, Ruzicka J, Stern JC, Glavin DP, House CH, Dworkin JP (2011) Carbonaceous meteorites contain a wide range of extraterrestrial nucleobases. *Proc Natl Acad Sci USA* 108:13995–13998
- Cashel M, Gallant J (1974) Cellular regulation of guanosine tetraphosphate and guanosine pentaphosphate. In: Nomura M, Tissières A, Lengyel P (eds) *Ribosomes*. Cold Spring Harbor, New York, pp 733–745
- Chatterji D, Ojha AK (2001) Revisiting the stringent response, ppGpp and starvation signaling. *Curr Opin Microbiol* 4:160–165
- Chen J, Gottesman S (2014) Riboswitch regulates RNA. *Science* 345:876–877
- Chen Z-H, Schaap P (2012) The prokaryote messenger c-di-GMP triggers stalk cell differentiation in *Dictiostelium*. *Nature* 488:680–683
- Chen X, Li N, Ellington AD (2007) Ribozyme catalysis of metabolism in the RNA world. *Chem Biodivers* 4:633–655
- Copper G, Kimmich N, Belisle W, Sarinana J, Brabham K, Garrel L (2001) Carbonaceous meteorites as a source of sugar-related organic compounds for the early Earth. *Nature* 414:879–883
- Corrigan RM, Gründling A (2013) Cyclic di-AMP: another second messenger enters the fray. *Nat Rev Microbiol* 11:513–524
- Corrigan RM, Campeotto I, Jeganathan T, Roelofs KG, Lee VT, Gründling A (2013) Systematic identification of conserved bacterial c-di-AMP receptor proteins. *Proc Natl Acad Sci USA* 110:9084–9089
- D'Ari R, Casadesús J (1998) Underground metabolism. *Bioessays* 20:181–186
- Daignan-Fornier B, Pinson B (2012) 5-aminoimidazole-4-carboxamide-1-beta-D-ribofuranosyl 5'-monophosphate (AICAR), a highly conserved purine intermediate with multiple effects. *Metabolites* 2:292–302
- Dalebroux ZD, Svensson SL, Gaynor EC, Swanson MS (2010) ppGpp conjures bacterial virulence. *Microbiol Mol Biol Rev* 74:171–199
- Davies BW, Bogard RW, Young TS, Mekalanos JJ (2012) Coordinated regulation of accessory genetic elements produces cyclic dinucleotides for *V. cholerae* virulence. *Cell* 149:358–370
- De la Fuente-Núñez C, Refeuveille F, Haney EF, Straus SK, Hancock RE (2014) Broad-spectrum anti-biofilm peptide that targets a cellular stress response. *PLoS Pathog* 10:e1004152
- Eakin RE (1963) An approach to the evolution of metabolism. *Proc Natl Acad Sci USA* 49:360–366
- Felsenstein J (2005) PHYLIP (phylogeny inference package) version 3.6. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle
- Ferris JP, Joshi PC, Edelson EH, Lawless JG (1978) HCN: a plausible source of purines, pyrimidines and amino acids on the primitive earth. *J Mol Evol* 11:293–311
- Flårdh K, Axberg T, Albertson NH, Kjelleberg S (1994) Stringent control during carbon starvation of marine *Vibrio* sp. strain S14: molecular cloning, nucleotide sequence, and deletion of the relA gene. *J Bacteriol* 176:5949–5957
- Flores NA, Stavrou BM, Sheridan DJ (1999) The effects of diadenosine polyphosphate on the cardiovascular system. *Cardiovasc Res* 42:15–26
- Gallant J, Palmer L, Pao CC (1977) Anomalous synthesis of ppGpp in growing cells. *Cell* 11:181–185
- Garrison PN, Mathis SA, Barnes LD (1986) In vivo levels of diadenosine tetraphosphate and adenosine tetraphospho-guanosine in *Physarum polycephalum* during the cell cycle and oxidative stress. *Mol Cell Biol* 6:1179–1186

- Geiger T, Wolz C (2014) Intersection of the stringent response and the CodY regulon in low GC gram-positive bacteria. *Int J Med Microbiol* 304:150–155
- Geiger T, Francois P, Liebeke M, Fraunholz M, Goerke C, Krismer B, Schrenzel J, Lalk M, Wolz C (2012) The stringent response of *Staphylococcus aureus* and its impact on survival after phagocytosis through the induction of intracellular PSMs expression. *PLoS Pathog* 8:e1003016
- Gerdes K, Maisonneuve E (2015) Remarkable functional convergence: alarmone ppGpp mediates persistence by activating type I and II toxin–antitoxins. *Mol Cell* 59:1–3
- Gilbert W (1986) Origin of life: the RNA world. *Nature* 319:618
- Gomelsky M (2011) cAMP, c-di-GMP, c-di-AMP and now cGMP: bacteria use them all! *Mol Microbiol* 79:562–565
- González-Plaza JJ (2018) Small RNAs in cell-to-cell communications during bacterial infection. *FEMS Microbiol Lett* 365:1–9
- Hall J, Ralph EC, Shanker S, Wang H, Byrnes LJ, Horst R, Wong J, Brault A, Dumlao D, Smith JF, Dakin LA, Schmitt DC, Trujillo J, Vincent F, Griffor M, Aulabaugh AE (2017) The catalytic mechanism of cyclic GMP-AMP synthase (cGAS) and implications for innate immunity and inhibition. *Protein Sci* 26:2367–2380
- Handler P (1961) Evolution of the coenzymes. In: Oparin AI (ed) *Proceedings of the 5th International Congress of Biochemistry*. Macmillan, New York, pp 149–157
- Haseltine WA, Block R (1973) Synthesis of guanosine tetra- and pentaphosphate requires the presence of a codon-specific, uncharged transfer ribonucleic acid in the acceptor site of ribosome. *Proc Natl Acad Sci USA* 70:1564–1568
- Hauryliuk V, Atkinson GC, Murakami KS, Tenson T, Gerdes K (2015) Recent functional insights into the role of (p)ppGpp in bacterial physiology. *Nat Rev Microbiol* 13:298–309
- Hood EE, Armour S, Ownby JD, Handa AK, Bressan RA (1979) Effect of nitrogen starvation on the level of adenosine 3',5'-monophosphate in *Anabaena variabilis*. *Biochim Biophys Acta* 588:193–200
- Huang F, Bugg CW, Yarus M (2000) RNA-catalyzed CoA, NAD, and FAD synthesis from phosphopantetheine, NMN, and FMN. *Biochemistry* 39:15548–15555
- Jácome R, Becerra A, de León SP, Lazcano A (2015) Structural analysis of monomeric RNA-dependent polymerases: evolutionary and therapeutic implications. *PLoS ONE* 10:e0139001
- Jenal U, Reinders A, Lori C (2017) Cyclic di-GMP: second messenger extraordinaire. *Nat Rev Microbiol* 15:271–284
- Jones CP, Ferré-D'Amaré AR (2015) Recognition of the bacterial alarmone ZMP through long-distance association of two RNA subdomains. *Nat Struct Mol Biol* 22:679–685
- Kanehisa M, Goto S (2000) KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* 28:27–30
- Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M (2016) KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res* 44:457–462
- Kasahara M, Ohmori M (1999) Activation of a cyanobacterial adenylate cyclase, CyaC, by autophosphorylation and subsequent phosphotransfer reaction. *J Biol Chem* 274:15167–15172
- Kellenberger CA, Wilson SC, Hickey SF, Gonzalez TL, Su Y, Hallberg ZF, Brewer TF, Lavarone AT, Carlson HK, Hsieh YF, Hammond MC (2015) GEMM-I riboswitches from geobacter sense the bacterial second messenger cyclic AMP-GMP. *Proc Natl Acad Sci USA* 112:5383–5388
- Krissinel E, Henrick K (2004) Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr D Biol Crystallogr* 60:2256–2268
- Krol E, Becker A (2011) ppGpp in *Sinorhizobium meliloti*: biosynthesis in response to sudden nutritional downshifts and modulation of transcriptome. *Mol Microbiol* 81:1233–1254
- Lamers MH, Georgescu RE, Lee SG, O'Donnell M, Kuriyan J (2006) Crystal structure of the catalytic  $\alpha$  subunit of *E. coli* replicative DNA polymerase III. *Cell* 126:881–892
- Lazcano A (2014) The RNA World: stepping out of the shadows. In: Gabriel T (ed) *Why does evolution matter? The importance of understanding evolution*. Cambridge Scholars Publishing, Newcastle upon Tyne, pp 101–119
- Lazcano A (2018) Prebiotic evolution and self assembly of nucleic acids. *ACS Nano*. <https://doi.org/10.1021/acs.nano.8b07605>
- Lazcano A, Becerra A, Delaye L (2011) Alarmones. In: Margulis L, Asikainen CA, Krumbie WE (ed) *Chimeras and consciousness: evolution of the sensory self*. The MIT Press, Massachusetts Institute of Technology, Cambridge, pp 35–43
- Lee PC, Bochner BR, Ames BN (1983a) Diadenosine 5',5'''-P1,P4-tetraphosphate and related adenylylated nucleotides in *Salmonella typhimurium*. *J Biol Chem* 258:6827–6834
- Lee PC, Bochner BR, Ames BN (1983b) AppppA, heat-shock stress, and cell oxidation. *Proc Natl Acad Sci USA* 80:7496–7500
- Letunic I, Bork P (2016) Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res* 44:242–245
- Lori C, Ozaki S, Steiner S, Böhm R, Abel S, Dubey BN, Schirmer T, Hiller S, Jenal U (2015) Cyclic di-GMP acts as a cell cycle oscillator to drive chromosome replication. *Nature* 523:236–239
- Magnusson LU, Farewell A, Nyström T (2005) ppGpp: a global regulator in *Escherichia coli*. *Trends Microbiol* 13:236–242
- Maisonneuve E, Gerdes K (2014) Molecular mechanisms underlying bacterial persisters. *Cell* 157:539–548
- Majerfeld I, Puthenvedu D, Yarus M (2016) Cross-backbone templating: ribonucleotides made on poly(C). *RNA* 22:397–407
- McDonough KA, Rodriguez A (2011) The myriad roles of cyclic AMP in microbial pathogens: from signal to sword. *Nat Rev Microbiol* 10:27–38
- Mehne FM, Gunka K, Eilers H, Herzberg C, Kaefer V, Stülke J (2013) Cyclic di-AMP homeostasis in *Bacillus subtilis*: both lack in high level accumulation of the nucleotide are detrimental for cell growth. *J Biol Chem* 288:2004–2017
- Münzel T, Feil R, Mülsch A, Lohmann SM, Hofmann F, Walter U (2003) Physiology and pathophysiology of vascular signaling controlled by cyclic guanosine 3',5'-cyclic monophosphate-dependent protein kinase. *Circulation* 108:2172–2183
- Nelson JW, Breaker RR (2017) The lost language of the RNA World. *Sci Signal* 10:1–10
- Nelson JW, Sudarsan N, Furukawa K, Weinberg Z, Wang JX, Breaker RR (2013) Riboswitches in eubacteria sense the second messenger c-di-AMP. *Nat Chem Biol* 9:834–839
- Orgel LE (1968) Evolution of the genetic apparatus. *J Mol Biol* 38:381–393
- Orgel LE, Sulston JE (1971) Polynucleotide replication and the origin of life. In: Kimball AP, Oro J (ed) *Prebiotic and biochemical evolution*. North-Holland, Amsterdam, pp 89–94
- Oró J (1960) Synthesis of adenine from ammonium cyanide. *Biochem Biophys Res Commun* 2:407–412
- Pendergast W, Yerxa BR, Douglass JG III, Shaver SR, Dougherty RW, Redick CC, Sims IF, Rideout JL (2001) Synthesis and P2Y receptor activity of a series of uridine dinucleoside 5'-polyphosphates. *Bioorg Med Chem Lett* 11:157–160
- Pesavento C, Hengge R (2009) Bacterial nucleotide-based second messengers. *Curr Opin Microbiol* 12:170–176
- Pilz RB, Casteel DE (2003) Regulation of gene expression by cyclic GMP. *Circ Res* 93:1034–1046
- Poole K (2012) Bacterial stress responses as determinants of antimicrobial resistance. *J Antimicrob Chemother* 67:2069–2089
- Potrykus K, Cashel M (2008) (p)ppGpp: still magical? *Annu Rev Microbiol* 62:35–51

- Pullman B (1972) Electronic factors in biochemical evolution. In: Ponnamperuna C (ed) Exobiology. North Holland Publishing Company, Amsterdam, pp 136–169
- Puthenvedu D, Janas T, Majerfeld I, Illangasekare M, Yarus M (2015) Poly(u) RNA-templated synthesis of AppA. *RNA* 21:1818–1825
- Rapaport E, Zamecnik PC (1976) Presence of diadenosine 5',5'-P<sub>1</sub>P<sub>4</sub>-tetrphosphate (Ap<sub>4</sub>A) in mammals cells in levels varying widely with proliferative activity of the tissue: a possible positive "pleiotypic activator". *Proc Natl Acad Sci USA* 73:3984–3988
- Rios AC, Tor Y (2013) On the origin of the canonical nucleobases: an assessment of selection pressures across chemical and early biological evolution. *Isr J Chem* 53:469–483
- Ritson D, Sutherland JD (2012) Prebiotic synthesis of simple sugars by photoredox systems chemistry. *Nat Chem* 4:895–899
- Romero D, Traxler MF, López D, Kolter R (2011) Antibiotics as signal molecules. *Chem Rev* 111:5492–5505
- Ross P, Weinhouse H, Aloni Y, Michaeli D, Weinberger-Ohana P, Mayer R, Braun S, de Vroom E, van der Marel GA, van Boom JH, Benziman M (1987) Regulation of cellulose synthesis in *Ace-tobacter xylinum* by cyclic diguanylic acid. *Nature* 325:279–281
- Ryu MH, Moskvina OV, Siltberg-Liberles J, Gomelsky M (2010) Natural and engineered photoactivated nucleotidyl cyclases for optogenetic applications. *J Biol Chem* 285:41501–41508
- Sabina RL, Holmes EW, Becker MA (1984) The enzymatic synthesis of 5-amino-4-imidazolecarboxamide riboside triphosphate (ZTP). *Science* 16:1193–1195
- Saha S, Jia Z, Liu D, Misra HP (2011) The roles of cAMP and G protein signaling in oxidative stress-induced cardiovascular dysfunction. In: Basu S, Wiklund L (ed) Studies on experimental models. Humana Press, Totowa, pp 621–635
- Sakaguchi K, Tsujino M, Hayashi M, Kawai K, Mizuno K, Hayano K (1976) Mode of action of bredinin with guanylic acid on L5178Y mouse leukemia cells. *J Antibiot* 12:1320–1327
- Sanders KM, Ward SM (1992) Nitric oxide as a mediator of non-adrenergic, noncholinergic neurotransmission. *Am J Physiol* 262:379–392
- Seifert R, Beste K, Burhenne H, Voigt U, Wolter S, Hammerschmidt A, Reinecke D, Sandner P, Pich A, Schwede F, Genieser HG, Kaever V (2011) Cyclic CMP and cyclic UMP: new (old) second messengers. *BMC Pharmacol* 11:O34
- Shenoy AR, Visweswariah SS (2004) Class III nucleotide cyclases in bacteria and archaeobacteria: lineage-specific expansion of adenylyl cyclases and a dearth of guanylyl cyclases. *FEBS Lett* 561:11–21
- Shenoy AR, Sivakumar K, Krupa A, Srinivasan N, Visweswariah SS (2004) A survey of nucleotide cyclases in actinobacteria: unique domain organization and expansion of the class III cyclase family in *Mycobacterium tuberculosis*. *Comp Funct Genomics* 5:17–38
- Sherlock ME, Sudarsan N, Breaker RR (2018) Riboswitches for the alarmone ppGpp expands the collection of RNA-based signaling systems. *Proc Natl Acad Sci USA* 21:1–6
- Sidi Y, Mitchel BS (1985) Z-nucleotide accumulation in erythrocytes from Lesch–Nyhan patients. *J Clin Invest* 76:2416–2419
- Spira B, Silberstein N, Yagil E (1995) Guanosine 3',5'-bispyrophosphate (ppGpp) synthesis in cells of *Escherichia coli* starved for Pi. *J Bacteriol* 177:4053–4058
- Stent GS, Brenner S (1961) A genetic locus for the regulation of ribonucleic acid synthesis. *Proc Nat Acad Sci USA* 47:2005–2014
- Stephens JC, Artz SW, Ames BN (1975) Guanosine 5'-diphosphate 3'-diphosphate (ppGpp): positive effector for histidine operon transcription and general signal for amino-acid deficiency. *Proc Nat Acad Sci USA* 72:4389–4393
- Subbiah S, Laurents DV, Levitt M (1993) Structural similarity of DNA-binding domains of bacteriophage repressors and the globin score. *Curr Biol* 3:141–148
- Sudarsan N, Lee ER, Weinberg Z, Moy RH, Kim JN, Link KH, Breaker RR (2008) Riboswitches in eubacteria sense the second messenger cyclic di-GMP. *Science* 18:411–413
- Suhadolnik RJ (1979) Naturally occurring nucleoside and nucleotide antibiotics. *Prog Nucleic Acid Res Mol Biol* 22:193–291
- Sun L, Wu J, Du F, Chen F, Chen ZJ (2013) Cyclic GMP-AMP synthase is a cytosolic DNA sensor that activates the type I interferon pathway. *Science* 339:786–791
- Sunahara RK, Beuve A, Tesmer JJ, Sprang SR, Garbers DL, Gilman AG (1998) Exchange of substrate and inhibitor specificities between adenylyl and guanylyl cyclases. *J Biol Chem* 273:16332–16338
- Sureka K, Choi PH, Precit M, Delince M, Pensinger DA, Huynh TN, Jurado AR, Goo YA, Sadilek M, Lavarone AT, Sauer JD, Tong L, Woodward JJ (2014) The cyclic dinucleotide c-di-AMP is an allosteric regulator of metabolic enzyme function. *Cell* 158:1389–1401
- Tamayo R, Pratt JT, Camilli A (2007) Role of cyclic diguanylate in the regulation of bacterial pathogenesis. *Annu Rev Microbiol* 61:131–148
- Thomas CB, Meade JC, Holmes EW (1981) Aminoimidazole carboxamide ribonucleoside toxicity: a model for study of pyrimidine starvation. *J Cell Physiol* 107:335–344
- Vázquez-Salazar A, Lazcano A (2018) Early life: embracing the RNA World. *Curr Biol* 5:220–222
- Vázquez-Salazar A, Tan G, Stockton A, Fani R, Becerra A, Lazcano A (2017) Can an imidazole be formed from an alanyl-seryl-glycine tripeptide under possible prebiotic conditions? *Orig Life Evol Biosph* 47:345–354
- Vázquez-Salazar A, Becerra A, Lazcano A (2018) Evolutionary convergence in the biosyntheses of the imidazole moieties of histidine and purines. *PLoS ONE* 13:e0196349
- Vinella D, Albrecht C, Cashel M, D'Ari R (2005) Iron limitation induces SpoT-dependent accumulation of ppGpp in *Escherichia coli*. *Mol Microbiol* 56:958–970
- Warner TD, Mitchell JA, Sheng H, Murad F (1994) Effects of cyclic GMP on smooth muscle relaxation. *Adv Pharmacol* 26:171–194
- White HB (1976) Coenzymes as fossils of an earlier metabolic state. *J Mol Evol* 7:101–104
- White HB (1982) Evolution of coenzymes and the origin of pyridine nucleotides. In: Everse J, Anderson B, You KS (ed) The pyridine nucleotide coenzymes. Academic Press, Cambridge, pp. 1–17
- Winkler WC, Nahvi A, Roth A, Collins JA, Breaker RR (2004) Control of gene expression by a natural metabolite-responsive ribozyme. *Nature* 428:281–286
- Witte G, Hartung S, Büttner K, Hopfner KP (2008) Structural biochemistry of a bacterial checkpoint protein reveals diadenylate cyclase activity regulated by DNA recombination intermediates. *Mol Cell* 30:167–178
- Wu J, Sun L, Chen X, Du F, Shi H, Chen C, Chen ZJ (2013) Cyclic GMP-AMP is an endogenous second messenger in innate immune signaling by cytosolic DNA. *Science* 339:826–830
- Yim G, Wang HH, Davies J (2007) Antibiotics as signaling molecules. *Philos Trans R Soc Lond B Biol Sci* 362:1195–1200
- Zhang L, Li W, He ZG (2013) DarR, a TetR-like transcriptional factor, is a cyclic di-AMP-responsive repressor in *Mycobacterium smegmatis*. *J Biol Chem* 288:3085–3096