



# Pyrophosphate and Irreversibility in Evolution, or why $PP_i$ Is Not an Energy Currency and why Nature Chose Triphosphates

Jessica L. E. Wimmer<sup>1\*</sup>, Karl Kleinermanns<sup>2</sup> and William F. Martin<sup>1\*</sup>

<sup>1</sup>Institute for Molecular Evolution, Department of Biology, Heinrich Heine University Duesseldorf, Duesseldorf, Germany,

<sup>2</sup>Institute for Physical Chemistry, Department of Chemistry, Heinrich Heine University Duesseldorf, Duesseldorf, Germany

## OPEN ACCESS

### Edited by:

Christiane Dahl,  
University of Bonn, Germany

### Reviewed by:

Jose Roman Perez-Castineira,  
Sevilla University, Spain  
Adrian Goldman,  
University of Helsinki, Finland

### \*Correspondence:

Jessica L. E. Wimmer  
jessica.wimmer@hhu.de  
William F. Martin  
bill@hhu.de

### Specialty section:

This article was submitted to  
Microbial Physiology and Metabolism,  
a section of the journal  
Frontiers in Microbiology

**Received:** 16 August 2021

**Accepted:** 15 September 2021

**Published:** 05 October 2021

### Citation:

Wimmer JLE, Kleinermanns K and  
Martin WF (2021) Pyrophosphate and  
Irreversibility in Evolution, or why  $PP_i$   
Is Not an Energy Currency and why  
Nature Chose Triphosphates.  
*Front. Microbiol.* 12:759359.  
doi: 10.3389/fmicb.2021.759359

The possible evolutionary significance of pyrophosphate ( $PP_i$ ) has been discussed since the early 1960s. Lipmann suggested that  $PP_i$  could have been an ancient currency or a possible environmental source of metabolic energy at origins, while Kornberg proposed that  $PP_i$  vectorializes metabolism because ubiquitous pyrophosphatases render  $PP_i$  forming reactions kinetically irreversible. To test those ideas, we investigated the reactions that consume phosphoanhydride bonds among the 402 reactions of the universal biosynthetic core that generates amino acids, nucleotides, and cofactors from  $H_2$ ,  $CO_2$ , and  $NH_3$ . We find that 36% of the core's phosphoanhydride hydrolyzing reactions generate  $PP_i$ , while no reactions use  $PP_i$  as an energy currency. The polymerization reactions that generate ~80% of cell mass – protein, RNA, and DNA synthesis – all generate  $PP_i$ , while none use  $PP_i$  as an energy source. In typical prokaryotic cells, aminoacyl tRNA synthetases (AARS) underlie ~80% of  $PP_i$  production. We show that the irreversibility of the AARS reaction is a kinetic, not a thermodynamic effect. The data indicate that  $PP_i$  is not an ancient energy currency and probably never was. Instead,  $PP_i$  hydrolysis is an ancient mechanism that imparts irreversibility, as Kornberg suggested, functioning like a ratchet's pawl to vectorialize the life process toward growth. The two anhydride bonds in nucleoside triphosphates offer ATP-cleaving enzymes an option to impart either thermodynamic control ( $P_i$  formation) or kinetic control ( $PP_i$  formation) upon reactions. This dual capacity explains why nature chose the triphosphate moiety of ATP as biochemistry's universal energy currency.

**Keywords:** energetics, bioenergetics, chemical evolution, origin of life, early evolution, metabolism, kinetics, thermodynamics

## INTRODUCTION

Starting in the 1960s, thoughts on the possible evolutionary significance of inorganic pyrophosphate ( $PP_i$ ) have centered around two main concepts: irreversibility and energy. Kornberg, who worked on nucleic acid polymerization, recognized that  $PP_i$  producing biochemical steps confer the property of irreversibility upon reactions under physiological conditions because ubiquitous pyrophosphatases constantly degrade  $PP_i$  in the cytosol of cells (Kornberg, 1962). His reasoning

was straightforward: By degrading  $PP_i$ , a substrate required for the enzymatic back reaction of the  $PP_i$  producing step, the rate of the back reaction effectively approaches zero. In this way, pyrophosphatases would render  $PP_i$  producing reactions irreversible by means of kinetics, rather than thermodynamics. Though Kornberg's mechanism of irreversibility was later called into question because  $PP_i$  concentrations in exponentially growing cells were reported to be too high for this principle to work (Kukko and Heinonen, 1982), as soon as cells leave the exponential growth phase, Kornberg's principle immediately applies, as we will see during the course of this paper, because  $PP_i$  production is strictly linked to growth, while  $PP_i$  hydrolysis is not. Kornberg's list of such irreversible  $PP_i$  producing reactions included nucleic acid polymerization, translation, and cofactor biosynthetic routes (Kornberg, 1962) and this function, irreversibility, was seen as harboring the significance of  $PP_i$ .

Lipmann, who worked on high energy bonds, suggested that  $PP_i$  could have served as a possible energy currency in primordial metabolism, and that modern  $PP_i$ -dependent enzymes represent fossils from a time in which prebiotic metabolism extracted energy from environmentally available phosphate minerals (Lipmann, 1965). In that view, the evolutionary significance of  $PP_i$  is sought in its possible role as a source of biochemical energy in prebiotic chemical reactions resembling those of physiology. Aspects of both Kornberg's and Lipmann's views are germane to Schramm's proposal that environmental polyphosphates could have powered early nucleic acid synthesis (Schramm et al., 1962).

In 1966, Baltscheffsky reported a membrane-associated pyrophosphatase (mPPase) that reversibly couples proton translocation to  $PP_i$  hydrolysis (Baltscheffsky et al., 1966), thereby linking  $PP_i$  to Mitchell's then new chemiosmotic theory of ATP synthesis involving ion gradients and electron transfer chains (Mitchell, 1961). That finding, together with Reeves' report of a  $PP_i$ -dependent glycolytic enzyme (Reeves, 1968), now called pyruvate orthophosphate dikinase, seemed to support an ancient bioenergetic role behind the possible evolutionary significance of  $PP_i$ . Based on such findings, the view that  $PP_i$ 's evolutionary significance resides in primordial energetics established a long tradition that is still widely embraced (de Duve, 1991; Russell and Hall, 1997; Russell et al., 2013, 2014; Wang et al., 2019; Piast et al., 2020) though seldom critically inspected (Martin, 2020).

Comparatively few enzymatic reactions involve  $PP_i$ . Kornberg (1962) listed 35 enzymatic reactions that release  $PP_i$  in the physiological reaction. Heinonen (2001) listed 173  $PP_i$  producing reactions. By contrast, Kyoto Encyclopedia of Genes and Genomes (KEGG) list 194 reactions among prokaryotes that involve ATP. Since the book of Heinonen (2001), some new  $PP_i$  producing reactions have been reported (Nagata et al., 2018), yet the precise roles of  $PP_i$  in physiology and evolution are still discussed (Heinonen, 2001; Pérez-Castiñeira et al., 2021). Soluble pyrophosphatases (sPPases) are ubiquitous in distribution (Lahti, 1983). Ion-pumping mPPases are found in various microbes and plants (Serrano et al., 2007), and  $PP_i$ -dependent glycolytic enzymes occur as alternatives of ATP-dependent forms, at the phosphofructokinase (PFK) and

pyruvate kinase (PYK) steps (Heinonen, 2001; Siebers and Schönheit, 2005; Bräsen et al., 2014; Holwerda et al., 2020). Though  $PP_i$ -dependent glycolysis is often interpreted as an adaptation that reduces ATP expense (Heinonen, 2001) or that salvages energy from  $PP_i$  produced from translation (Reeves, 1968),  $PP_i$ -utilizing glycolytic enzymes have a conspicuous tendency to occur among microbes that have specialized to sugar-rich environments. Such specialists include human parasites, such as *Entamoeba* (Reeves, 1984), *Giardia* (Müller et al., 2012), and trypanosomes (Michels et al., 2006), as well as non-parasitic cellulose-, saccharose-, and sugar-degrading bacteria (Bielen et al., 2010; Holwerda et al., 2020) and archaea (Bräsen et al., 2014). In addition,  $PP_i$ -dependent enzymes are particularly common in the strictly sugar-based carbon metabolism of plants (Serrano et al., 2007). This pattern of occurrence might be suggestive of an ecological rather than energetic basis behind the distribution of  $PP_i$ -dependent glycolytic pathways.

In line with that view, the use of  $PP_i$ -dependent glycolytic enzymes generally coincides with loss of allosteric regulation through the pathway (Siebers and Schönheit, 2005; Bräsen et al., 2014). In the well-studied example of trypanosomes, loss of regulation allows flux through the pathway to be governed by sugar concentrations in the medium (blood sugar), an ecological adaptation of growth rates to substrate availability, not energetic efficiency, especially as trypanosomes excrete the energy rich compound pyruvate as a metabolic end product (Michels et al., 2006). Even in the well-studied glucose fermenting bacterium *Clostridium thermocellum*, which also excretes pyruvate, a clear energetic advantage of its  $PP_i$ -dependent glycolysis is not evident (Holwerda et al., 2020). Moreover, deletion of *C. thermocellum*'s mPPase has no impact on growth (Holwerda et al., 2020), a finding that is hard to reconcile with a central role for energy conservation *via*  $PP_i$  in energy metabolism of the bacterium, although sPPase activity was not reported in the mPPase mutant. By contrast, deletion mutants of sPPases are lethal in *Escherichia coli* (Chen et al., 1990) and in yeast (Pérez-Castiñeira et al., 2002). This finding is very notable because from an energetic standpoint, because sPPases effectively "waste" phosphoanhydride bonds *via* rapid  $PP_i$  hydrolysis, raising the question: why should elimination of the "energy wasting" reaction catalyzed by sPPase be lethal? The growth inhibiting phenotype of sPPase deletion mutants is, however, readily reconciled with Kornberg's kinetic view of  $PP_i$  function, because sPPase knockouts in *E. coli* and yeast yield cells that cannot grow mainly because protein synthesis comes to a halt through product inhibition *via*  $PP_i$  accumulation at the amino acyl tRNA synthesis step.

Our present interest in  $PP_i$  stems from comparative physiological investigations into the energetics of primordial metabolism (Martin and Russell, 2007; Sousa et al., 2013; Sousa and Martin, 2014; Preiner et al., 2020). We reasoned that if  $PP_i$  had played any role in primordial energetics, as is widely assumed (de Duve, 1991; Russell et al., 2013, 2014; Wang et al., 2019; Piast et al., 2020), evidence for that role should be preserved in the conserved core of metabolism within modern cells. This is the same conventional logic that is used

to interpret other aspects of physiology as relicts of ancient metabolism: metal sulfide clusters in proteins (Eck and Dayhoff, 1966; Wächtershäuser, 1992; Heinen and Lauwers, 1997), the use of organic cofactors as catalysts (White, 1976), carbon metal bonds in enzyme active sites (Martin, 2019), thioesters as energy currencies (Semenov et al., 2016; Kitadai et al., 2021), or anaerobic chemolithoautotrophy (Mereschkowsky, 1910; Decker et al., 1970). Though this line of reasoning (comparative physiology) can be questioned, it is the same reasoning that underlies the view that  $PP_i$  is an ancient energy currency. The conserved core of metabolism is a set of roughly 400 reactions that generates the 20 canonical amino acids, the four bases of RNA and DNA, and the cofactors required for their synthesis from  $H_2$ ,  $CO_2$ ,  $NH_3$ ,  $H_2S$ ,  $P_i$ , and inorganic salts (Wimmer et al., 2021). Because of its universally conserved nature, this biosynthetic core of chemical reactions (though not necessarily all of its enzymes) was present in the last universal common ancestor, LUCA, and has persisted in all lineages throughout evolution over the last 4 billion years since their divergence from LUCA (Weiss et al., 2016). The universal core thus harbors insights not only into LUCA's physiology, but also into the primordial set of reactions that gave rise to the building blocks from which LUCA was assembled. Because our present investigation probes metabolism itself, our insights into the role of  $PP_i$  in evolution differ from those based in the study of phosphorous minerals (Pasek et al., 2017). And because our investigation is based in the comparative physiology of living cells, our insights into the role of thermodynamics and kinetics in evolution differ from those based in studies of chemical nucleic acid synthesis (Pascal et al., 2013). A fresh look at the role of  $PP_i$  in ancient metabolism suggests that Lipmann was probably wrong, that Kornberg was probably right, and furthermore reveals why nature chose triphosphates as the universal energy currency.

## MATERIALS AND METHODS

### Reactions Involving Inorganic Pyrophosphate

The 36 metabolic reactions involving inorganic pyrophosphate ( $PP_i$ ; **Supplementary Table S1**) in the biosynthetic core were taken from supplemental data of Wimmer et al. (2021). Reaction R00720 involving IMP synthesis was removed from the core because it is not essential. The reactions were initially collected from the KEGG (Kanehisa and Goto, 2000), version December 2020 and polarized in the direction of cell synthesis.

### ATP Hydrolysis Among Prokaryotes

Reactions involving ATP hydrolysis of the reaction scheme  $X + ATP \leftrightarrow Y + ADP + P_i$  were obtained from KEGG (Kanehisa and Goto, 2000).  $X$  and  $Y$  are placeholders for variable compounds. Additional compounds on both sides can be present. A total of 15,339 KEGG reactions were searched for the reaction scheme in the forward and back direction since KEGG reactions are not polarized in general. A total of 131 reactions involving

ATP hydrolysis were obtained in the data, whereas 61 reactions are specific to prokaryotes (**Supplementary Table S2**). The domain check was performed by parsing the Enzyme Commission (EC) numbers of each reaction, gathering a list of genes and their respective organisms leading to the domain.

### Collection of Michaelis–Menten Constants for Pyrophosphatases

Michaelis–Menten constants ( $K_m$ ) for inorganic pyrophosphatase activity in *E. coli* wildtypes were obtained from BRENDA (Schomburg et al., 2002) via EC number 3.6.1.1. *Escherichia coli* mutants were removed from the data (**Supplementary Table S3**).

### Kinetic Effect of $PP_i$ in Translation

To investigate the effect of pyrophosphate hydrolysis on the product yield (adenylated amino acid) in aminoacyl-tRNA synthetase (AARS) reactions more quantitatively, kinetic simulations of substrate binding and activation of isoleucine by adenylation in isoleucyl-tRNA synthetase were performed using Mathcad 2001 (Mathsoft Engineering & Education, Inc.). The underlying kinetic scheme is taken from Pope et al. (1998) with hydrolysis of  $PP_i$  added (see **Supplementary Figure S1A**). The rate equations were used to obtain the concentration vs. time profiles by numerical integration (see **Supplementary Figure S1B**). Experimental rate constants were obtained from Pope et al. (1998) and Stockbridge and Wolfenden (2011). Initial concentrations of 1 mM amino acid, enzyme, and ATP were used, and integration was carried out up to 20 s. These calculations provide an empirical basis for the intuitive effect of product removal during the  $PP_i$  forming step of translation (see **Supplementary Figure S2**).

## RESULTS

### Pyrophosphate Polarized LUCA's Core Biosynthetic Metabolism

To see whether  $PP_i$  might have had a role in primordial energetics, the reactions of the core (Wimmer et al., 2021) that involve  $PP_i$  or ATP were identified and polarized in the biosynthetic direction, that is, from  $H_2$  and  $CO_2$  toward cell mass synthesis. In Lipman's view,  $PP_i$  was an environmental energy source, a substrate that assumes a thermodynamic role as an educt residing on the left side of an enzymatic reaction, while in Kornberg's view  $PP_i$  is synthesized in metabolism via ATP hydrolysis and assumes a kinetic role as a product that is removed from the right side of the reaction. Writing the reactions from left to right in the direction of  $CO_2$  to products as the pathways are mapped in KEGG (Kanehisa and Goto, 2000) brings the role of  $PP_i$  in the core into focus. Among the 36 reactions of the core in which  $PP_i$  occurs, it is always a reaction product occurring on the right side of the reaction, serving as an energy source in zero reactions (**Supplementary Table S1**). The reactions of the biosynthetic core of metabolism thus speak 36:0 in favor of  $PP_i$  conferring irreversibility, as Kornberg (1962)

suggested, and harbor no traces of Lipmann's proposal for an ancient energetic or thermodynamic role for  $PP_i$ .

In the metabolism of modern cells,  $PP_i$  is always produced from ATP by reaction sequences that sum to  $ATP + H_2O \rightarrow AMP + PP_i$  ( $\Delta G_o' = -46 \text{ kJ}\cdot\text{mol}^{-1}$ ), slightly more exergonic than  $ATP + H_2O \rightarrow ADP + P_i$  ( $\Delta G_o' = -32 \text{ kJ}\cdot\text{mol}^{-1}$ ). This opens the possibility that  $PP_i$  formation might have played an energetic role in the core, but not as a source of high energy bonds. Were the role of  $PP_i$  in the core thermodynamic, it could have readily been replaced in evolution by compounds with a similar or higher free energy of hydrolysis, such as acetyl phosphate ( $\Delta G_o' = -43 \text{ kJ}\cdot\text{mol}^{-1}$ ), 1,3-bisphosphoglycerate ( $\Delta G_o' = -52 \text{ kJ}\cdot\text{mol}^{-1}$ ), or phosphoenolpyruvate ( $\Delta G_o' = -62 \text{ kJ}\cdot\text{mol}^{-1}$ ), the high energy bonds in all three of which are synthesized in metabolism using one ATP each (the same cost as  $PP_i$  hydrolysis). Because ATP hydrolysis to adenosine monophosphate (AMP) and  $PP_i$  is not replaced by alternative energy currencies with a higher free energy of hydrolysis, and because  $PP_i$  is always a product in the core, not an educt, the function of  $PP_i$  in the core can hardly be thermodynamic.

Keeping in mind that Kornberg's suggestion for the role of  $PP_i$  was based on nucleic acid polymerization and translation, the occurrence of  $PP_i$  in the core solely as a product suggests that its role is kinetic, lowering the rate of back reactions, rather than thermodynamic. Is this true more generally in metabolism, that is, outside the core? We consulted KEGG. If  $PP_i$  had any role during early evolution as an energy currency, then some reactions should persist in which  $PP_i$  hydrolysis is coupled to an otherwise endergonic reaction. Though a handful of  $PP_i$  consuming reactions phosphorylate substrates in the physiological reaction (Nagata et al., 2018), among 15,339 reactions in KEGG, we found no  $PP_i$  hydrolyzing, non-phosphorylating reactions at all that provide energetic coupling to an otherwise thermodynamically unfavorable reaction. That is, there were no reactions of the type  $X + PP_i \leftrightarrow Y + 2 P_i$ , whereby we note that the pyrophosphatase reaction, KEGG reaction number R00004, employs  $H_2O$  as X but has no Y component. This is a noteworthy observation. It indicates that  $PP_i$  serves at best as a phosphorylating agent in metabolism, but never as a source of pure thermodynamic impetus to help push unfavorable reactions forward *via* coupling to  $PP_i$  hydrolysis. By contrast, a number of metabolic reactions (61 prokaryote specific reactions among 15,339 total reactions in KEGG; **Supplementary Table S2**) go forward because they are coupled to non-phosphorylating ATP hydrolysis in reactions of the type  $X + ATP \leftrightarrow Y + ADP + P_i$ . The lack of such reactions for  $PP_i$  in KEGG clearly indicates that  $PP_i$  is not a dedicated energy currency in biosynthesis, notwithstanding the existence of  $PP_i$ -dependent glycolytic pathways, as outlined in the introduction. Note that KEGG does not include the myriad reactions in which ATP (or GTP) phosphorylates proteins, and we know of no examples in which  $PP_i$  is used to phosphorylate proteins as true energy currencies do. These findings indicate that  $PP_i$  is not a dedicated energy currency and that by inference, in the simplest interpretation, it never has been.

$PP_i$  producing reactions are generally seen as being irreversible under physiological conditions because of the

ubiquitous presence of high activities of sPPases in cells (Lahti, 1983; Danchin et al., 1984; Heinonen, 2001), which catalyze the reaction  $PP_i + H_2O \rightarrow 2P_i$  ( $\Delta G_o' = -21 \text{ kJ}\cdot\text{mol}^{-1}$ ), thereby continuously removing a substrate for  $PP_i$  producing reactions in the reverse direction. Notably, aqueous  $Mg^{2+}$  ions alone accelerate the rate of spontaneous  $PP_i$  hydrolysis by three orders of magnitude in water and  $PP_i$  hydrolysis in dimethyl sulfoxide/water by six orders of magnitude (Stockbridge and Wolfenden, 2011), such that inorganically catalyzed  $PP_i$  hydrolysis might have been a mechanism of irreversibility even before the advent of enzymes. Irreversibility at 36  $PP_i$ -dependent enzymatic reactions in the core – nine in amino acid pathways, three in nucleotide synthesis, and 19 in cofactor synthesis (**Supplementary Table S1**) – functions in modern metabolism as a system of check valves (valves that close to prevent backward flow) that, individually and in concert, act as a ratchet's pawl, inching the reactions of the core unidirectionally forward toward product synthesis. We suggest that this has been the case since the availability of ATP as the universal energy currency.

## Pyrophosphate Polarized Metabolism *in toto* Throughout All of Evolution

In the metabolism of LUCA,  $PP_i$  forced the reactions of the core forward in the direction of monomer synthesis for cell mass synthesis. The effect of  $PP_i$ , however, extended well beyond LUCA's core biosynthesis because  $PP_i$  renders nucleic acid and protein synthesis irreversible (Kornberg, 1962), and because LUCA possessed the genetic code and was able to synthesize RNA, DNA, and proteins (Weiss et al., 2016). To get a better picture of the polarizing role of  $PP_i$  in the central dogma of molecular biology, we generated estimates for its quantitative contribution to the overall ATP budget based on the classical estimates of Stouthamer (1978), which are still in wide use today. Protein synthesis requires activated amino acids, rRNA, tRNA, and mRNA. In a modern cell growing from  $H_2$ ,  $CO_2$ , and  $NH_3$ , the synthesis of protein comprises the combined energetic cost of making RNA and protein, consuming roughly 76% of the biosynthetic ATP budget (**Table 1**). The quantitative contribution of  $PP_i$  forming reactions to the cellular energy budget is surprisingly large. In amino acid biosynthesis, 47% of the ATP consuming reactions (8/17) generate  $PP_i$ , whereas in nucleotide synthesis 13.6% (3/22) generate  $PP_i$ . In polymerization reactions, the contributions are greater.

Guanosine triphosphate (GTP) hydrolyzing reactions are not uncommon in LUCA's biosynthetic core (**Figure 1**), in line with its ancient role in metabolism (Martin and Russell, 2007) and the observation that in some organisms where it has been investigated, GTP is readily used as a substrate in reactions that are typically regarded as ATP dependent (Holwerda et al., 2020). About 26% of a cell's energy budget is consumed in the GTP-dependent steps of translation. The main biosynthetic ATP expense in protein synthesis is translation, which consumes four ATP per peptide bond (Stouthamer, 1978). Peptide chain elongation at the ribosome has two  $P_i$  forming GTP hydrolysis steps catalyzed by EF-Tu and EF-G (Satpati et al., 2014), while

**TABLE 1** | ATP expense per gram of cells during biosynthetic processes<sup>a</sup> [mol·10<sup>4</sup>].

| Component                   | Monomer synthesis <sup>b</sup> |                           | Polymerization <sup>c</sup> |                           |
|-----------------------------|--------------------------------|---------------------------|-----------------------------|---------------------------|
|                             | Total                          | PP <sub>i</sub> - forming | Total                       | PP <sub>i</sub> - forming |
| Protein                     | 14                             | 6.6                       | 191                         | 91.3                      |
| RNA                         | 34                             | 4.6                       | 23                          | 20.3                      |
| DNA                         | 9                              | 1.2                       | 2                           | 1.2                       |
| Lipid <sup>d</sup>          | 1                              | –                         | –                           | –                         |
| Polysaccharide <sup>d</sup> | 21                             | –                         | –                           | –                         |
| Import <sup>e</sup>         | 52                             | –                         | –                           | –                         |
| Sum                         | 131                            | 12.4                      | 216                         | 113                       |

Energetic cost of protein synthesis incl. Ribosome biosynthesis: 262/347 = 76%

Energetic contribution of PP<sub>i</sub> forming steps in cell biosynthesis: 125/347 = 36%

Energetic contribution of GTP-dependent biosynthetic steps: 98.5/347 = 28%

<sup>a</sup>Values are for *Escherichia coli* from Stouthamer (1978) as tabulated by Harold (1986). Lever et al. (2015) calculate  $\Delta G'_0$  for the synthesis of monomers from H<sub>2</sub>, CO<sub>2</sub>, and NH<sub>3</sub> based on the values of Neidhardt et al. (1990) but not the ATP expense per monomer. Neidhardt et al. (1990) estimate ATP expense for monomer synthesis as 42 ATP per 20 amino acids and 40 ATP per four nucleotides in *E. coli*. A dash (–) indicates that the value is zero or negligible. Note that these calculations entail only biosynthetic costs and do not consider energy spilling (Tempest and Neijssel, 1984) or maintenance energy, which can exceed biosynthesis by a factor of 3 in exponential growth (Russell and Cook, 1995; Russell, 2007), and by more under energy limitation (Lever et al., 2015).

<sup>b</sup>The proportion of PP<sub>i</sub> forming steps in monomer biosynthesis was calculated as the total cost for the monomer multiplied by the fraction of PP<sub>i</sub> forming steps among ATP hydrolyzing steps en route to nucleoside monophosphates (Wimmer et al., 2021).

<sup>c</sup>The proportion of PP<sub>i</sub> forming steps in polymerization takes the costs of proofreading, assembly and modification from Lever et al. (2015) into account. These are not PP<sub>i</sub>-forming reactions.

<sup>d</sup>LPS, lipopolysaccharide. These values are for *E. coli*. There is a PP<sub>i</sub> forming component of lipid synthesis in archaea that is neglected here. Archaea also lack LPS and possess no murein, though sometimes pseudomurein (Albers and Meyer, 2011), and have a larger protein component in the cell wall (S-Layer).

<sup>e</sup>Stouthamer (1978) calculates the cost of import for precursors, mainly ammonium, Lever et al. (2015) neglect import. If one considers a functional core before the origin of free-living cells, no costs for import are incurred.

<sup>f</sup>In the core, 13% of the triphosphate expense for amino acid and NMP/dNMP monomer synthesis (55.10<sup>-4</sup> mol ATP per gram of cells) is GTP-dependent (7.2.10<sup>-4</sup> mol ATP per gram of cells), plus two GTP-dependent steps in translation (91.3.10<sup>-4</sup> mol ATP per gram of cells) yield ca. 98.5.10<sup>-4</sup> mol GTP per gram of cells.

amino acyl tRNA synthesis requires the expense of two ATP through amino acid activation by amino acyl tRNA synthetase (AARS) enzymes. Activation generates aminoacyl adenylate and PP<sub>i</sub>, followed by aminoacylation of tRNA and AMP release (Gomez and Ibbá, 2020). Because half of the energy cost for translation resides in the PP<sub>i</sub> producing nature of the AARS reactions, roughly 26% of the cell's total biosynthetic ATP expense (91/347, **Table 1**) is incurred to pay the price of irreversibility at the formation of aminoacyl tRNAs for translation.

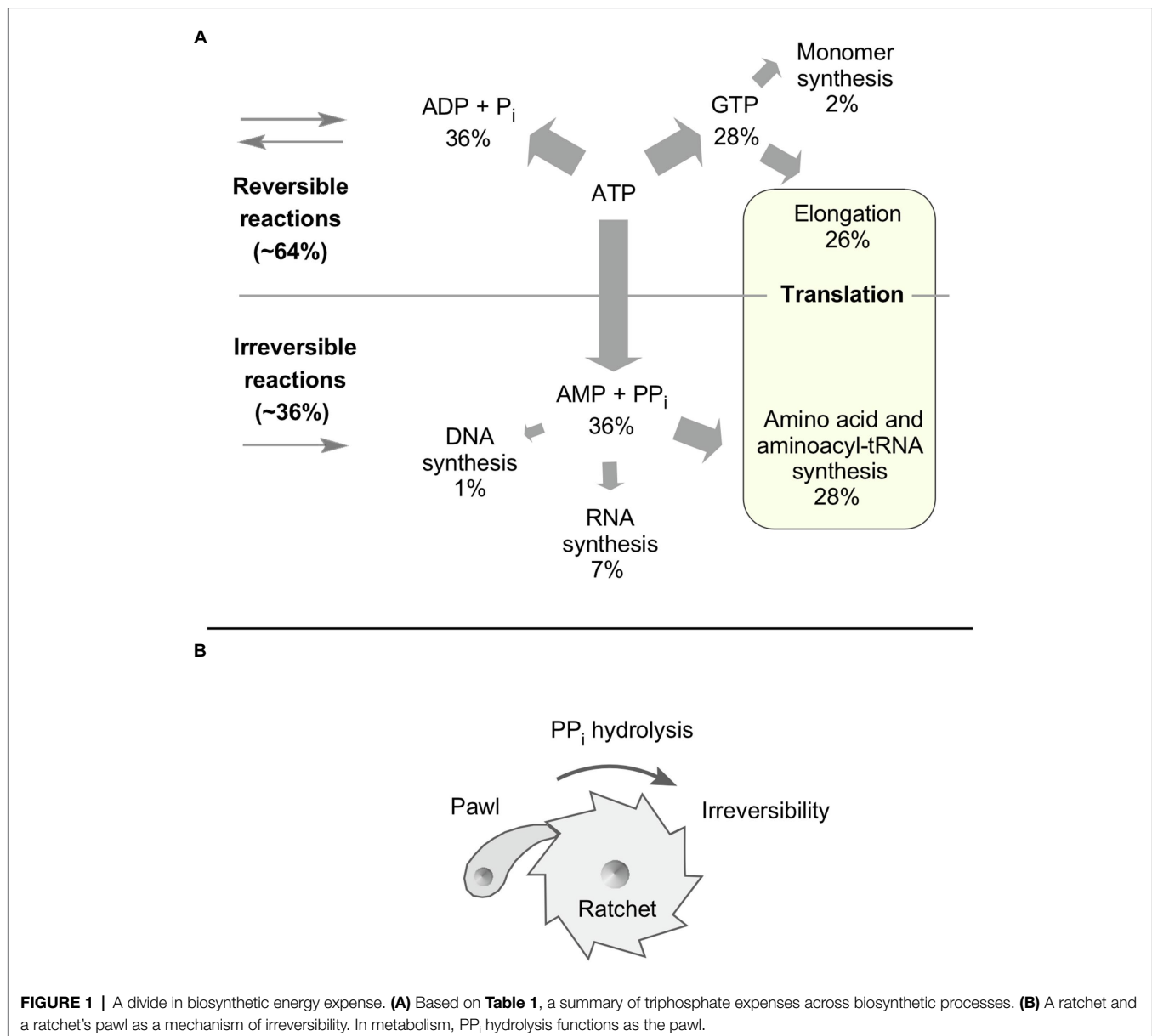
PP<sub>i</sub> has an ancient and conserved function in metabolism as a mediator of irreversibility (Kornberg, 1962) that clearly traces to LUCA (**Figure 1**). By contrast, not a single reaction in the core uncovers a role for PP<sub>i</sub> as an energy currency in primordial metabolism. Based upon the conserved nature of the core across all modern lineages, we can infer that PP<sub>i</sub> generating reactions have vectorialized monomer synthesis of ABC compounds throughout evolution, acting as a ratchet's pawl (**Figure 1B**), rendering monomer synthesis unidirectional, even in low energy environments.

PP<sub>i</sub> producing steps in RNA monomer and polymer synthesis account for about 7% of the overall biosynthetic energy budget (**Figure 1**). As with translation, thermodynamics do not strictly demand PP<sub>i</sub> generation and hydrolysis, as transcription can operate with NDPs *in vitro* (Gottesman and Mustaev, 2019). PP<sub>i</sub> production during DNA synthesis only accounts for 1% of the cell's energy budget (**Figure 1**), whereby some DNA polymerases can also operate with NDP substrates (Burke and Lupták, 2018). In addition, some polymerases use the irreversible effect of PP<sub>i</sub> hydrolysis by possession of a pyrophosphatase domain that cleaves PP<sub>i</sub> as the enzyme moves forward (Kottur and Nair, 2018).

PP<sub>i</sub> generation and hydrolysis render both the reactions of the core and polymerization reactions during replication, transcription and translation irreversible under the physiological conditions of the cell. In the core, and in the cytosol, this ratchet has locked biochemistry in the direction of cell synthesis during the 4 billion years since metabolic origin. The insight of Kornberg (1962, p. 261) “Hydrolysis of the latter [pyrophosphate] by inorganic pyrophosphatases promotes the irreversibility of the synthetic route to coenzymes, nucleic acids, proteins, and structural carbohydrates and lipids” still stands.

## The Effect of PP<sub>i</sub> in Translation Is Demonstrably Kinetic

The vectorializing effect of PP<sub>i</sub> in translation is not thermodynamic, it is the same as in the biosynthetic core: a kinetic ratchet afforded by ubiquitous pyrophosphatases that render protein synthesis unidirectional toward growth. We were able to demonstrate this effect by calculating the kinetics of aminoacyl tRNA synthesis using published rate constants obtained from Pope et al. (1998) and from Stockbridge and Wolfenden (2011). The equations are given in **Supplementary Figure S1**, into which we introduced values from the literature to obtain the kinetics (the time dependence of reactant and product concentrations) for the isoleucyl-tRNA synthetase reaction in *E. coli* in the presence of inorganic pyrophosphatase. The result is shown in **Supplementary Figure S2**. The *E. coli* enzyme belongs to the family I or type I PPase, which typically have rate constants on the order of 200–400 s<sup>-1</sup> (Kajander et al., 2013).



Using the *E. coli* sPPase I rate constant of  $570\text{ s}^{-1}$  provided by Stockbridge and Wolfenden (2011) in our calculations,  $PP_i$  hydrolysis is essentially complete after 0.4 s (**Supplementary Figure S2**). This is a clear result: Pyrophosphatase activity drives the overall reaction of aminoacyl tRNA synthesis forward by removing  $PP_i$  at a high rate relative to other steps of the reaction such that the adenylated amino acid is formed irreversibly. The reaction kinetics provides a clear empirical basis for the intuitive effect of product removal during the  $PP_i$  forming step of translation.

Different PPases have, however, different rate constants for  $PP_i$  hydrolysis. In particular, the membrane bound mPPases are extremely slow; hence, it was of interest to see if they could still provide a similar kinetic effect in the AARS reaction.

The mPPases occur in roughly 25% of prokaryotes; the enzyme is also common among protists and is ubiquitous among land plants, where it couples  $PP_i$  hydrolysis to the pumping of ions ( $\text{Na}^+$  or  $\text{H}^+$ ) out of the cell or, in the case of vacuolar PPases (vPPases), from the cytosol into vacuoles or acidocalcisomes, organelles rich in calcium and polyphosphate (Kajander et al., 2013). The mPPases and vPPases are one to two orders of magnitude slower than type I sPPases, with rates on the order of  $3.5\text{--}20\text{ s}^{-1}$  (Kajander et al., 2013). Thus, we lowered the rate constant of  $PP_i$  hydrolysis by a factor of 100 in our calculations ( $k_7 = 5.7\text{ s}^{-1}$ ; **Supplementary Figure S2**). The kinetic effect was basically the same: Nearly complete conversion (96%) of  $PP_i$  to  $P_i$  is reached at 20 s ( $\tau_{1/2} = 900\text{ ms}$ ). Hence, even very slow pyrophosphatases such as mPPases or vPPases can still drive amino acid activation by AARS enzymes

to near completion. The reaction takes slightly longer than in the case of the type I sPPase, but is still complete in well under a minute. The reasons why the rate constants of the membrane bound PPases are so low is not known (Kajander et al., 2013), and it might be because their ion pumps work reversibly, synthesizing  $PP_i$  from  $2P_i$  when the cations flow back through the membrane.

There are also type II sPPase (type II sPPase) that are much faster, and they hydrolyze  $PP_i$  with rate constants on the order of  $1,700\text{--}3,000\text{ s}^{-1}$  (Kajander et al., 2013). While all type I sPPases use  $Mg^{2+}$  ions to bind  $PP_i$  and water to negatively charged amino acids like aspartic acid and polarize water for  $PP_i$  hydrolysis by nucleophilic  $OH^-$  attack, type II sPPases additionally use  $Mn^{2+}$  (or  $Co^{2+}$ ) for binding and polarization, which probably relates to their higher rate constants. Type II sPPases have only been found in prokaryotes so far; they are common among clostridia and bacilli. Using the rate of  $3,000\text{ s}^{-1}$  for a type II sPPase (Kajander et al., 2013), we find that  $PP_i$  hydrolysis is 96% complete at 0.35 s and 98% complete at 0.5 s ( $\tau_{1/2} = 36\text{ ms}$ ). The reaction rate of type II sPPase drives  $PP_i$  hydrolysis to completion in less than second, removing all  $PP_i$  substrate for the AARS back reaction. PPases I and II release the  $PP_i$  hydrolysis energy of  $-20$  to  $-25\text{ kJ}\cdot\text{mol}^{-1}$  as heat in the cytosol. Hence, it is not pyrophosphatase thermodynamics, but its kinetics which drive the amino acid adenylation to high yield.

We point out that that in *E. coli* with,  $PP_i$  can accumulate to transient concentrations on the order of 1 mM (Kukko and Heinonen, 1982) in exponentially growing cells, which would seem to create a conflict with the idea of a kinetic effect. Yet in natural environments, exponential growth is rarely if ever attained, as discussed in more detail in the next section. In the context of weighing kinetic vs. thermodynamic effects of  $PP_i$  production, we also recall the “uncomfortable” observation that deletion of *C. thermocellum* mPPase has no impact upon exponential growth (Holwerda et al., 2020), whereby chemostat cultures of *C. thermocellum*, which have high  $PP_i$  concentrations in the cytosol and a  $PP_i$ -dependent glycolytic pathway, show clear signs of increased reversibility at the AARS reactions in the form of high concentrations of excreted amino acids (Holwerda et al., 2020). Whether such altruistic amino acid excretion into the environment *via* the AARS reaction would be manifested or sustainable in natural cellulose degrading environments over evolutionary timescales as opposed to chemostat growth conditions designed for biofuel yield is currently not known. It is also noteworthy that cells expend four ATP to generate a peptide bond even though one ATP would suffice, as peptide synthesis from aminoacyl phosphates (Kachalsky and Paecht, 1954) or non-ribosomal peptide synthesis (Martínez-Núñez and López y López, 2016) shows. The energetic difference between the one ATP required to form a peptide bond in solution vs. the four ATP that cells expend to make peptide bonds during translation can be seen as the energetic cost of structural information that is specified within a protein sequence (Haber and Anfinsen, 1962) plus the cost of its irreversible synthesis (Kornberg, 1962).

## DISCUSSION

### Irreversibility in the Long-Term Evolution of Cells in Nature

What are the consequences of a  $PP_i$  irreversibility ratchet over geological timescales? From a physiological and energetic standpoint, the function of  $PP_i$  is always subordinate to ATP, because the source of the anhydride bond in  $PP_i$  in modern metabolism is always ATP, generated either *via* ion gradients or *via* substrate level phosphorylation. A critic might interject that thylakoid pyrophosphatases might be able to conserve energy as  $PP_i$  (Jiang et al., 1997), but if they do, it would be at the expense of one ATP per  $PP_i$  formed.

A critic might interject that Heinonen (2001) has summarized evidence to suggest that the measured cellular  $PP_i$  levels on the order of 1 mM in logarithmically growing *E. coli* cells are too high to exert a kinetic effect of the kind that Kornberg had in mind. This issue can be illustrated with a passage from Kukko and Heinonen (1982), who measured the intracellular  $PP_i$  concentration of *E. coli* grown in batch culture with a doubling time of roughly 1 h. They found that the  $PP_i$  concentration was constant at about 0.5 mM during exponential growth. From this, they concluded that “[...] the metabolic role of  $PP_i$  has been clouded by the widespread belief that  $PP_i$  formed in the metabolism is rapidly hydrolyzed in cells to inorganic phosphate and the concentration of  $PP_i$  thus approaches zero in the cytoplasm. This view must be in error.” We do not doubt their observations; their interpretation is the issue. Kukko and Heinonen (1982) cite several other papers where  $PP_i$  concentrations on the order of 0.1–2 mM are reported, always from exponentially growing cells. Why are  $PP_i$  concentrations in exponentially growing cells misleading in the context of irreversibility?

When growth stops, so does  $PP_i$  production in the cytosol. But even after growth-dependent production of  $PP_i$  has ceased,  $PP_i$  continues to be hydrolyzed by pyrophosphatase activity. In a rare report, Danchin et al. (1984) measured  $PP_i$  after blocking ATP (hence  $PP_i$ ) synthesis in *E. coli*, and they found that  $PP_i$  levels dropped exponentially to  $100\text{ }\mu\text{M}$  within a minute and to  $10\text{ }\mu\text{M}$  within 10 min, in line with our kinetic calculations (Supplementary Figure S2). Bielen et al. (2010) also noted that  $PP_i$  levels dropped when cells ceased exponential growth. Why do  $PP_i$  levels drop when growth is arrested? It is because  $PP_i$  is produced by growth processes (Klemme, 1976), but is hydrolyzed to phosphate by pyrophosphatases continuously, also in resting cells, independent of growth. This leads to a rapid drop in  $PP_i$  concentrations, which do in fact approach zero in the cytoplasm, once  $PP_i$  production is halted. This is why Danchin et al. (1984) observed a precipitous drop in  $PP_i$  concentrations once  $PP_i$  production was arrested.

The sources of  $PP_i$  in metabolism in typical cells (here, typical means cells that lack  $PP_i$ -dependent glycolysis) have been known for decades. Klemme (1976) summarized the main sources of  $PP_i$  production in growing *E. coli* and the rates are which it is produced. In the units of  $\mu\text{mol}$  per 100 mg biomass, the contributions to  $PP_i$  production in exponentially growing *E. coli* were: synthesis of protein (545), nucleic acids (67),

polysaccharides (60), and lipids (60) for a total of 740. Using the fixed relationship between  $PP_i$  synthesis and growth in either rich or minimal medium, he was able to obtain good estimates for the rate of  $PP_i$  synthesis for several bacteria, which he set in relationship to the measured PPase activity for the same bacteria, allowing him to calculate the ratio of rates ( $\mu\text{moles}\cdot\text{h}^{-1}\cdot\text{mg protein}^{-1}$ ) for  $PP_i$  production and  $PP_i$  hydrolysis. For eight exponentially growing bacterial species (six Gram negative, two Gram positive), he found that the ratio of  $PP_i$  hydrolysis to  $PP_i$  synthesis was 79, 73, 57, 33, 14, 10, 8, and 1 (Table II of Klemme, 1976). The average value was 35; the value for *E. coli* was 14. That is, the rate of *E. coli*  $PP_i$  hydrolysis is 14 times higher than the rate of  $PP_i$  production from growth processes. Even if the ratio is only 1, when exponential growth is arrested, pyrophosphatase activity remains, which will relentlessly gnaw away at  $PP_i$  concentrations until they essentially reach zero or until rapid growth is resumed, such that rates of cytosolic  $PP_i$  production can exceed  $PP_i$  hydrolysis. A number of microbes possess  $Mn^{2+}$ - or  $Co^{2+}$ -dependent sPPases, called family II sPPases, that have a catalytic rate higher than that of typical sPPases (Kajander et al., 2013).

The foregoing raises two important points. The first is that  $PP_i$  levels in exponentially growing cells are not a good proxy for the function of  $PP_i$  over evolutionary time. This is because sustained exponential growth is never attained for microbes in the environment or during evolution, as they are mainly starved for nutrients in the wild. In the largest microbial community known, marine sediment, cells do not actually grow, they just slowly die as organic nutrients become limiting (Orsi et al., 2020). In such environments, the standard concept of doubling times does not apply to growth or survival, as cell mass never doubles, it just turns over from one living cell to another, with turnover times on the order of tens to thousands of years (Hoehler and Jørgensen, 2013). In starved cells as they exist in sediment, ATP synthesis is orders of magnitude slower than in exponentially growing cells and  $PP_i$  production is governed by the rate of protein synthesis, meaning that on time scales of days, months, and years, trace pyrophosphatase activity will hold the cytosolic  $PP_i$  concentration close to zero, even if the enzyme's affinity for  $PP_i$  is comparatively low. However, measured values of  $K_m$  (the substrate concentration at half maximal enzymatic reaction rate) for  $PP_i$  for sPPases are not high, and they tend to be in the range of  $1\mu\text{M}$  to  $1\text{mM}$  in *E. coli* (Supplementary Table S3) and values of catalytic rate tend to be on the order of  $200\text{--}400\text{s}^{-1}$  (Avaeva, 2000; Kajander et al., 2013), meaning that over long time scales, PPases keep  $PP_i$  levels in cells too low to permit the enzymatic reactions of translation or nucleic acid polymerization from running backwards, especially at the extremely slow  $PP_i$  production rates of starved microbial communities.

We say "extremely slow  $PP_i$  production rates." How slow is slow? We can provide an estimate. Starved cells are small. An exponentially growing *E. coli* cell has about 2 million proteins with on average 300 amino acids each (Milo et al., 2010). If a starved cell is half that size, roughly 300 million peptide bonds are required for its formation. There are 32 million seconds in a year, such that if the turnover time of starved

cell is on the order of 10 years, one peptide bond per second is formed, on average, during the formation of the cell. That might not sound too slow, because a ribosome can form about 10 peptide bonds per second. But a small *E. coli* cell has on the order of 10,000 ribosomes, such that in a starved cell with a 10-year turnover time, an individual ribosome might perform an elongation step on the order of roughly once every 3h or  $10^{-4}$  peptide bonds per second. In terms of molecular processes that are extremely slow, the rate of  $PP_i$  production is 100,000 times slower than from translation during exponential growth. This example underscores the value to the cell of translation (aminoacyl tRNA synthesis) being an irreversible process over evolutionary timescales, and the essential function that irreversibility plays by acting as a ratchet's pawl to prohibit the back reaction of aminoacyl tRNA formation, quantitatively the most important energetic expense a cell encounters (Figure 1).

A second important point concerns (rare) examples in which sPPases are lacking in the genome such that  $PP_i$  reaches high cytosolic concentrations rendering translation theoretically reversible. This can occur in cells that have  $PP_i$ -based glycolysis, such as *C. thermocellum*, where ion-pumping membrane-bound PPases (mPPases) are present (Holwerda et al., 2020). In such cells, which are sugar specialists,  $PP_i$  levels can exceed 20mM during chemostat growth, whereby the metabolic source of such high  $PP_i$  levels is still unclear and deletion of the *C. thermocellum* mPPase has no impact upon growth (Holwerda et al., 2020). The existence of cells with  $PP_i$ -dependent carbon metabolism lacking high activities of sPPases in chemostat growth on very rich medium do not invalidate Kornberg's principle of irreversibility over evolutionary timescales. Rather they constitute an evolutionarily derived special case of adaptation to growth on sugar. As outlined in the introduction,  $PP_i$  utilizing glycolytic enzymes tend to occur among microbes that have specialized to sugar-rich environments, including human parasites such as *Entamoeba*, *Giardia*, and trypanosomes, or cellulose- and saccharose-degrading bacteria and archaea, but also in plants with their specialized sugar synthesizing compartment, the plastid. Proton-pumping mPPases are often associated with acidocalcisomes, membrane-bounded compartments that occur in some prokaryotes and in some eukaryotes including parasites and plants (Docampo et al., 2005). The functions discussed for acidocalcisomes include among other things storage of cations, phosphate and polyphosphate, calcium signaling, and osmoregulation but no evidence for an involvement in energy metabolism of acidocalcisomes or their mPPase has emerged so far (Docampo and Huang, 2016).

In the bigger picture of microbial evolution, sugar-dependent lifestyles cannot be ancestral, and they have to be derived. The early earth was barren and offered  $\text{CO}_2$ , not glucose, as the main environmental carbon source (Schönheit et al., 2016). In the modern crust (Smith et al., 2019) and marine sediments (Orsi et al., 2020), where most cells on Earth have always resided (McMahon and Parnell, 2018), net growth is almost non-existent due to nutrient limitations (Orsi et al., 2020). Particularly in low-energy environments,  $PP_i$  irreversibility at the quantitatively dominant (in terms of



PP<sub>i</sub> synthesis) AARS reaction acts like a ratchet's pawl that keeps aminoacyl tRNAs moving in solely in the direction of translation, even if translation is slow for reasons of substrate limitation or prolonged starvation.

## CONCLUSION

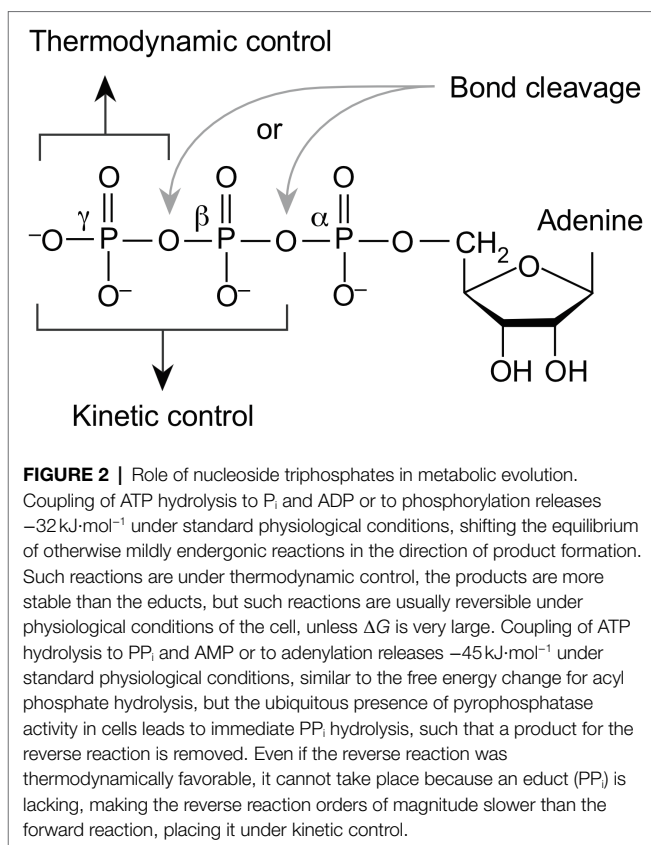
### Why Nature Chose Triphosphates

The role of PP<sub>i</sub> in evolution raises a question similar to Westheimer's "why phosphate," namely "why triphosphates?" Westheimer (1987) proposed that phosphates became energy carriers because of the metastability of the various bonds that phosphate forms with organic compounds under physiological conditions. By examining the role of PP<sub>i</sub> in the core and in the central dogma, we found that the central function of PP<sub>i</sub> producing reactions is not that of an energy currency in any case. In metabolism, PP<sub>i</sub> is always generated from nucleoside triphosphates. This is also true for mPPases, where the ion gradients that are required for PP<sub>i</sub> synthesis are generated at ATP expense. Formulated directly, we find no evidence that PP<sub>i</sub> served as a primordial energy currency or that it serves as an energy currency today. Rather, it appears that the role of PP<sub>i</sub> is to impart direction upon the most essential operations of life: biosynthesis of cofactors, the biosynthesis of the monomeric building blocks of proteins and nucleic acids, and the polymerization of those building blocks into the catalysts and information carriers of cells (Figure 1).

Why did nature specifically choose nucleoside triphosphates as the universal energy currency? That is a fundamentally different question from why nature chose phosphate (Westheimer, 1987; Liu et al., 2019) because many biological compounds harbor phosphate bonds with a large free energy of hydrolysis (Decker et al., 1970), but only triphosphates are the universal energy currency in all lineages today. Irreversibility provides the answer. Triphosphates can generate either P<sub>i</sub> or PP<sub>i</sub>. This subtle property reveals why triphosphates became the universal energy currency in cells and why they have not been replaced in 4 billion years of evolution (Figure 2). How so?

Nucleoside triphosphates such as ATP have two phosphoanhydride bonds. The β-phosphate in ATP can be cleaved on either side (Figure 2). ATP-dependent enzymatic reactions that release P<sub>i</sub> utilize ATP as a currency of thermodynamic control, making ΔG of the reaction sufficiently negative (or the net activation energy sufficiently low) to allow the reaction to go forward. ATP-dependent reactions that release PP<sub>i</sub> also have a thermodynamic component, but the irreversibility of the reaction conferred by PP<sub>i</sub> hydrolysis under physiological conditions places the reaction under kinetic rather than thermodynamic control.

No biochemical energy currency other than (nucleoside) triphosphates offers, within the same compound, the alternative of exerting either thermodynamic or kinetic control over a reaction. This property is specific to triphosphates. It allowed primordial enzymes to exert either kinetic control or thermodynamic control over catalyzed reactions, depending upon which anhydride bond of the β-phosphate was cleaved.



This in turn imparted the option of evolutionary refinement of an initial catalytic activity among ATP utilizing enzymes according to the prevailing selective forces in a given cellular environment. An early onset of PP<sub>i</sub>-dependent irreversibility in metabolism would not require the presence of a pre-existing inorganic pyrophosphatase enzyme activity at the site of origins, because the reaction can be catalyzed by inorganic ions alone, such as Mg<sup>2+</sup> (Stockbridge and Wolfenden, 2011), which catalyzes hydrolysis in the active site of many modern pyrophosphatase enzymes (Varfolomeev and Gurevich, 2001).

This, in turn, is the reason why nucleoside triphosphates became fixed in both monomer and polymer biosynthesis in the metabolism of LUCA and have not been displaced since. From an ancestral state in which acyl phosphates provided thermodynamic impetus and a means of energetic coupling in enzymatic reactions, the advent of nucleoside triphosphates changed the nature of early biochemical evolution by introducing the option of kinetic irreversibility. Triphosphates offered primordial enzymes a means to exert either kinetic control or thermodynamic control over catalyzed reactions with one and the same energy currency. The only evident alternative solutions would have been (i) to maintain two distinct energetic currencies in the cell, one for energetic and one for kinetic purposes (an event for which there is no evidence) or (ii) to abandon one of the functions (which is not a viable option over evolutionary time). The ability of triphosphates to function in roughly 2/3 of phosphoanhydride bond expenditure as a currency of energy (thermodynamic drive) and in roughly 1/3 of phosphoanhydride

bond expenditure as a currency of irreversibility (kinetic drive; **Figure 1**) is the reason they became – and remained – life’s universal energy currency.

## DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**; further inquiries can be directed to the corresponding author/s.

## AUTHOR CONTRIBUTIONS

JW collected and analyzed data, participated in project design, and revised the manuscript. KK performed the kinetic calculations and contributed in data interpretation. WM wrote the first manuscript draft, performed literature research, visualization, and conceived and supervised the study. All authors contributed to the article and approved the submitted version.

## REFERENCES

- Albers, S. V., and Meyer, B. H. (2011). The archaea cell envelope. *Nat. Rev. Microbiol.* 9, 414–426. doi: 10.1038/nrmicro2576
- Avaeva, S. M. (2000). Active site interactions in oligomeric structures of inorganic pyrophosphatases. *Biochemistry* 65, 361–372.
- Baltscheffsky, H., von Stedingk, L. V., Heldt, H. W., and Klingenberg, M. (1966). Inorganic pyrophosphate: formation in bacterial photophosphorylation. *Science* 153, 1120–1122. doi: 10.1126/science.153.3740.1120
- Bielen, A. A. M., Willquist, K., Engman, J., van der Oost, J., van Niel, E. W. J., and Kengen, S. W. M. (2010). Pyrophosphate as a central energy carrier in the hydrogenproducing extremely thermophilic Caldicellulosiruptor saccharolyticus. *FEMS Microbiol. Lett.* 307, 48–54. doi: 10.1111/j.1574-6968.2010.01957.x
- Bräsen, C., Esser, D., Rauch, B., and Siebers, B. (2014). Carbohydrate metabolism in archaea: current insights into unusual enzymes and pathways and their regulation. *Microbiol. Mol. Biol. Rev.* 78, 89–175. doi: 10.1128/MMBR.00041-13
- Burke, C. R., and Lupták, A. (2018). DNA synthesis from diphosphate substrates by DNA polymerases. *Proc. Natl. Acad. Sci. U. S. A.* 115, 980–985. doi: 10.1073/pnas.1712193115
- Chen, J., Brevet, A., Fromant, M., Lévêque, F., Schmitter, J.-M., Blanquet, S., et al. (1990). Pyrophosphatase is essential for growth of *Escherichia coli*. *J. Bacteriol.* 172, 5686–5689. doi: 10.1128/jb.172.10.5686-5689.1990
- Danchin, A., Dondon, L., and Daniel, J. (1984). Metabolic alterations mediated by 2-ketobutyrate in *Escherichia coli* K12. *Mol. Gen. Genet.* 193, 473–478. doi: 10.1007/BF00382086
- de Duve, C. (1991). *Blueprint for a Cell: The Nature and Origin of Life*. Burlington, USA: Neil Patterson Publishers
- Decker, K., Jungerman, K., and Thauer, R. K. (1970). Energy production in anaerobic organisms. *Angew. Chem. Int. Ed.* 9, 138–158. doi: 10.1002/anie.197001381
- Docampo, R., de Souza, W., Miranda, K., Rohloff, P., and Moreno, S. N. (2005). Acidocalcisomes—conserved from bacteria to man. *Nat. Rev. Microbiol.* 3, 251–261. doi: 10.1038/nrmicro1097
- Docampo, R., and Huang, H. (2016). Acidocalcisomes of eukaryotes. *Curr. Opin. Cell Biol.* 41, 66–72. doi: 10.1016/j.cob.2016.04.007
- Eck, R. V., and Dayhoff, M. O. (1966). Evolution of the structure of ferredoxin based on living relics of primitive amino acid sequences. *Science* 152, 363–366. doi: 10.1126/science.152.3720.363
- Gomez, M. A. R., and Ibba, M. (2020). Aminoacyl-tRNA synthetases. *RNA* 26, 910–936. doi: 10.1261/rna.071720.119

## FUNDING

This work was supported by the European Research Council (Advanced Grants eMicrobevol and EcolMetabOrigin to WM), the Deutsche Forschungsgemeinschaft (Ma 1426/21-1 to WM), and the Volkswagen Foundation (VW 96742 to WM).

## ACKNOWLEDGMENTS

We thank Harun Tüysüz, William Orsi, Martina Preiner, Joana Xavier, Delfina Pereira, and Andrey do Nascimento Vieira for helpful discussions and the team of the central computing facility at the University of Düsseldorf (ZIM) for their support.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2021.759359/full#supplementary-material>

- Gottesman, M. E., and Mustaev, A. (2019). Ribonucleoside-5'-diphosphates (NDPs) support RNA polymerase transcription, suggesting NDPs may have been substrates for primordial nucleic acid biosynthesis. *J. Biol. Chem.* 294, 11785–11792. doi: 10.1074/jbc.RA119.009074
- Haber, B., and Anfinsen, C. B. (1962). Side-chain interactions governing the pairing of half-cystine residues in ribonuclease. *J. Biol. Chem.* 237, 1839–1844. doi: 10.1016/S0021-9258(19)73945-3
- Harold, F. M. (1986). *The Vital Force: A Study of Bioenergetics*. New York, USA: WH Freeman
- Heinen, W., and Lauwers, A. M. (1997). The iron-sulfur world and the origins of life: abiotic synthesis from metallic iron, H<sub>2</sub>S and CO<sub>2</sub>: A comparison of the thiol generating FeS/HCl(H<sub>2</sub>S)/CO<sub>2</sub>-system and its Fe<sup>0</sup>/H<sub>2</sub>S/CO<sub>2</sub>-counterpart. *Proc. K. Ned. Akad. van Wet.* 100, 11–25.
- Heinonen, J. K. (2001). *Biological Role of Inorganic Pyrophosphate*. New York, USA: Springer US
- Hoehler, T. M., and Jørgensen, B. B. (2013). Microbial life under extreme energy limitation. *Nat. Rev. Microbiol.* 11, 83–94. doi: 10.1038/nrmicro2939
- Holwerda, E. K., Zhou, J., Hon, S., Stevenson, D. M., Amador-Noguez, D., Lynd, L. R., et al. (2020). Metabolic fluxes of nitrogen and pyrophosphate in chemostat cultures of *Clostridium thermocellum* and *Thermoanaerobacterium saccharolyticum*. *Appl. Environ. Microbiol.* 86, e01795–e01720. doi: 10.1128/AEM.01795-20
- Jiang, S. S., Fan, L. L., JingYang, S., Kuo, S. Y., and Pan, R. L. (1997). Purification and characterization of thylakoid membrane-bound inorganic pyrophosphatase from *Spinacia oleracea* L. *Arch. Biochem. Biophys.* 346, 105–112. doi: 10.1006/abbi.1997.0279
- Kachalsky, A., and Paecht, M. (1954). Phosphate anhydrides of amino acids. *J. Biol. Chem. Soc.* 76, 6042–6044. doi: 10.1021/ja01652a053
- Kajander, T., Kellosalo, J., and Goldman, A. (2013). Inorganic pyrophosphatases: one substrate, three mechanisms. *FEBS Lett.* 587, 1863–1869. doi: 10.1016/j.febslet.2013.05.003
- Kanehisa, M., and Goto, S. (2000). KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 28, 27–30. doi: 10.1093/nar/28.1.27
- Kitadai, N., Nakamura, R., Yamamoto, M., Okada, S., Takahagi, W., Nakano, Y., et al. (2021). Thioester synthesis through geoelectrochemical CO<sub>2</sub> fixation on Ni sulfides. *Commun. Chem.* 4:37. doi: 10.1038/s42004-021-00475-5
- Klemme, J.-H. (1976). Regulation of intracellular pyrophosphatase-activity and conservation of the phosphoanhydride-energy of inorganic pyrophosphate in microbial metabolism. *Z. Naturforsch.* 31, 544–550. doi: 10.1515/znc-1976-9-1011
- Kornberg, A. (1962). “On the metabolic significance of phosphorylytic and pyrophosphorolytic reactions,” in *Horizons in Biochemistry*. eds. H. Kasha and B. Pullman (New York, USA: Academic Press), 251–264.

- Kottur, J., and Nair, D. T. (2018). Pyrophosphate hydrolysis is an intrinsic and critical step of the DNA synthesis reaction. *Nucleic Acids Res.* 46, 5875–5885. doi: 10.1093/nar/gky402
- Kukko, E., and Heinonen, J. (1982). The intracellular concentration of pyrophosphate in the batch culture of *Escherichia coli*. *Eur. J. Biochem.* 127, 347–349. doi: 10.1111/j.1432-1033.1982.tb06878.x
- Lahti, R. (1983). Microbial inorganic pyrophosphatases. *Microbiol. Rev.* 47, 169–178. doi: 10.1128/mr.47.2.169-178.1983
- Lever, M. A., Rogers, K. L., Lloyd, K. G., Overmann, J., Schink, B., Thauer, R. K., et al. (2015). Life under extreme energy limitation: a synthesis of laboratory- and field-based investigations. *FEMS Microbiol. Rev.* 39, 688–728. doi: 10.1093/femsre/fuv020
- Lipmann, F. (1965). “Projecting backward from the present stage of evolution of biosynthesis,” in *The Origin of Prebiological Systems and of their Molecular Matrices*. ed. S. W. Fox (New York, USA: Academic Press), 259–280.
- Liu, Z., Rossi, J. C., and Pascal, R. (2019). How prebiotic chemistry and early life chose phosphate. *Lifestyles* 9:26. doi: 10.3390/life9010026
- Martin, W. F. (2019). Carbon-metal bonds: rare and primordial in metabolism. *Trends Biochem. Sci.* 44, 807–818. doi: 10.1016/j.tibs.2019.04.010
- Martin, W. F. (2020). Older than genes: The acetyl CoA pathway and origins. *Front. Microbiol.* 11:817. doi: 10.3389/fmicb.2020.00817
- Martin, W., and Russell, M. J. (2007). On the origin of biochemistry at an alkaline hydrothermal vent. *Philos. Trans. R. Soc. B* 362, 1887–1925. doi: 10.1098/rstb.2006.1881
- Martínez-Núñez, M. A., and López y López, V. E. (2016). Nonribosomal peptide synthetases and their applications in industry. *Sustain. Chem. Process.* 4:13. doi: 10.1186/s40508-016-0057-6
- McMahon, S., and Parnell, J. (2018). The deep history of earth's biomass. *J. Geol. Soc.* 175, 716–720. doi: 10.1144/jgs2018-061
- Mereschkowsky, C. (1910). Theorie der zwei Plasmaarten als Grundlage der Symbiogenese, einer neuen Lehre von der Entstehung der Organismen. *Biol. Centralbl.* 30, 353–367.
- Michels, P. A. M., Bringaud, F., Herman, M., and Hannaert, V. (2006). Metabolic functions of glycosomes in trypanosomatids. *Biochim. Biophys. Acta* 1763, 1463–1477. doi: 10.1016/j.bbamcr.2006.08.019
- Milo, R., Jorgensen, P., Moran, U., Weber, G., and Springer, M. (2010). BioNumbers—the database of key numbers in molecular and cell biology. *Nucleic Acids Res.* 38, D750–D753. doi: 10.1093/nar/gkp889
- Mitchell, P. (1961). Coupling of phosphorylation to electron and hydrogen transfer by a chemi-osmotic type of mechanism. *Nature* 191, 144–148. doi: 10.1038/191144a0
- Müller, M., Mentel, M., van Hellemond, J. J., Henze, K., Woehle, C., Gould, S. B., et al. (2012). Biochemistry and evolution of anaerobic energy metabolism in eukaryotes. *Microbiol. Mol. Biol. Rev.* 76, 444–495. doi: 10.1128/MMBR.05024-11
- Nagata, R., Fujihashi, M., Sato, T., Atomi, H., and Miki, K. (2018). Identification of a pyrophosphate-dependent kinase and its donor selectivity determinants. *Nat. Commun.* 9:1765. doi: 10.1038/s41467-018-04201-z
- Neidhardt, F. C., Ingraham, J. L., and Schaechter, M. (1990). *Physiology of the Bacterial Cell*. Sunderland, MA, USA: Sinauer Associates
- Orsi, W. D., Schink, B., Buckel, W., and Martin, W. F. (2020). Physiological limits to cell mass conversion in anoxic seafloor sediment. *FEMS Microbiol. Rev.* 44, 219–231. doi: 10.1093/femsre/fuaa004
- Pascal, R., Pross, A., and Sutherland, J. D. (2013). Towards an evolutionary theory of the origin of life based on kinetics and thermodynamics. *Open Biol.* 3:130156. doi: 10.1098/rsob.130156
- Pasek, M. A., Gull, M., and Herschy, B. (2017). Phosphorylation on the early earth. *Chem. Geol.* 475, 149–170. doi: 10.1016/j.chemgeo.2017.11.008
- Pérez-Castiñeira, J. R., Docampo, R., Ezawa, T., and Serrano, A. (2021). Pyrophosphates and polyphosphates in plants and microorganisms. *Front. Plant Sci.* 26:653416. doi: 10.3389/fpls.2021.653416
- Pérez-Castiñeira, J. R., Lopez-Marques, R. L., Villalba, J. M., Losada, M., and Serrano, A. (2002). Functional complementation of yeast cytosolic pyrophosphatase by bacterial and plant H<sup>+</sup>-translocating pyrophosphatases. *Proc. Natl. Acad. Sci. U. S. A.* 99, 15914–15919. doi: 10.1073/pnas.242625399
- Piast, R. W., Garstka, M., Misicka, A., and Wiczorek, R. M. (2020). Small cyclic peptide for pyrophosphate dependent ligation in prebiotic environments. *Lifestyles* 10:103. doi: 10.3390/life10070103
- Pope, A. J., Moore, K. J., McVey, M., Mensah, L., Benson, N., Osbourne, N., et al. (1998). Characterization of isoleucyl-tRNA synthetase from *Staphylococcus aureus* II. Mechanism of inhibition by reaction intermediate and pseudomonic acid analogues studied using transient and steady-state kinetics. *J. Biol. Chem.* 273, 31680–31690. doi: 10.1074/jbc.273.48.31691
- Preiner, M., Igarashi, K., Muchowska, K. B., Yu, M., Varma, S. J., Kleinermanns, K., et al. (2020). A hydrogen dependent geochemical analogue of primordial carbon and energy metabolism. *Nat. Ecol. Evol.* 4, 534–542. doi: 10.1038/s41559-020-1125-6
- Reeves, R. E. (1968). A new enzyme with the glycolytic function of pyruvate kinase. *J. Biol. Chem.* 243, 3202–3204. doi: 10.1016/S0021-9258(18)93395-8
- Reeves, R. E. (1984). Metabolism of *Entamoeba histolytica* Schaudinn, 1903. *Adv. Parasitol.* 23, 105–142. doi: 10.1016/s0065-308x(08)60286-9
- Russell, J. B. (2007). The energy spilling reactions of bacteria and other organisms. *J. Mol. Microbiol. Biotechnol.* 13, 1–11. doi: 10.1159/000103591
- Russell, M. J., Barge, L. M., Bhartia, R., Bocanegra, D., Bracher, P. J., Branscomb, E., et al. (2014). The drive to life on wet and icy worlds. *Astrobiology* 14, 308–343. doi: 10.1089/ast.2013.1110
- Russell, J. B., and Cook, G. M. (1995). Energetics of bacterial growth: balance of anabolic and catabolic reactions. *Microbiol. Rev.* 59, 48–62. doi: 10.1128/mr.59.1.48-62.1995
- Russell, M. J., and Hall, A. J. (1997). The emergence of life from iron monosulphide bubbles at a submarine hydrothermal redox and pH front. *J. Geol. Soc. Lond.* 154, 377–402. doi: 10.1144/gsjgs.154.3.0377
- Russell, M. J., Nitschke, W., and Branscomb, E. (2013). The inevitable journey to being. *Philos. Trans. R. Soc. B Biol. Sci.* 368:20120254. doi: 10.1098/rstb.2012.0254
- Satpati, P., Sund, J., and Åqvist, J. (2014). Structure-based energetics of mRNA decoding on the ribosome. *Biochemistry* 53, 1714–1722. doi: 10.1021/bi5000355
- Schomburg, I., Chang, A., and Schomburg, D. (2002). BRENDA, enzyme data and metabolic information. *Nucleic Acids Res.* 30, 47–49. doi: 10.1093/nar/30.1.47
- Schönheit, P., Buckel, W., and Martin, W. (2016). On the origin of heterotrophy. *Trends Microbiol.* 24, 12–25. doi: 10.1016/j.tim.2015.10.003
- Schramm, G., Grottsch, H., and Pollmann, W. (1962). Nicht-enzymatische Synthese von Polysacchariden, Nucleosiden und Nucleinsäuren und die Entstehung selbst-vermehrungsfähiger Systeme. *Angew. Chem.* 74, 53–92. doi: 10.1002/ange.19620740202
- Semenov, S. N., Kraft, L. J., Ainla, A., Zhao, M., Baghbanzadeh, M., Campbell, V. E., et al. (2016). Autocatalytic, bistable, oscillatory networks of biologically relevant organic reactions. *Nature* 537, 656–660. doi: 10.1038/nature19776
- Serrano, A., Pérez-Castiñeira, J. R., Baltscheffsky, M., and Baltscheffsky, H. (2007). H<sup>+</sup>-PPases: yesterday, today and tomorrow. *IUBMB Life* 59, 76–83. doi: 10.1080/15216540701258132
- Siebers, B., and Schönheit, P. (2005). Unusual pathways and enzymes of central carbohydrate metabolism in archaea. *Curr. Opin. Microbiol.* 8, 695–705. doi: 10.1016/j.mib.2005.10.014
- Smith, A. R., Kieft, B., Mueller, R., Fisk, M. R., Mason, O. U., Popa, R., et al. (2019). Carbon fixation and energy metabolisms of a seafloor olivine biofilm. *ISME J.* 13, 1737–1749. doi: 10.1038/s41396-019-0385-0
- Sousa, F., and Martin, W. F. (2014). Biochemical fossils of the ancient transition from geoenergetics to bioenergetics in prokaryotic one carbon compound metabolism. *Biochim. Biophys. Acta* 1837, 964–981. doi: 10.1016/j.bbabi.2014.02.001
- Sousa, F. L., Thiery, T., Landan, G., Nelson-Sathi, S., Pereira, I. A. C., Allen, J. F., et al. (2013). Early bioenergetic evolution. *Philos. Trans. R. Soc. B* 368:20130088. doi: 10.1098/rstb.2013.0088
- Stockbridge, R. B., and Wolfenden, R. (2011). Enhancement in the rate of pyrophosphate hydrolysis by nonenzymatic catalysts and by inorganic pyrophosphatase. *J. Biol. Chem.* 286, 18538–18546. doi: 10.1074/jbc.M110.214510
- Stouthamer, A. H. (1978). “Energy-yielding pathways,” in *The Bacteria*. ed. I. C. Gunsalus (New York, USA: Academic Press), 389–462.
- Tempest, D. W., and Neijssel, O. M. (1984). The status of Y<sub>ATP</sub> and maintenance energy as biologically interpretable phenomena. *Annu. Rev. Microbiol.* 38, 459–486. doi: 10.1146/annurev.mi.38.100184.002331
- Varfolomeev, S. D., and Gurevich, K. G. (2001). Enzyme active sites: bioinformatics, architecture, and mechanisms of action. *Russ. Chem. Bull.* 50, 1709–1717. doi: 10.1023/A:1014353724442

- Wächtershäuser, G. (1992). Groundworks for an evolutionary biochemistry—the iron sulfur world. *Prog. Biophys. Mol. Biol.* 58, 85–201. doi: 10.1016/0079-6107(92)90022-X
- Wang, Q., Barge, L. M., and Steinbock, O. (2019). Microfluidic production of pyrophosphate catalyzed by mineral membranes with steep pH gradients. *Chem. Eur. J.* 25, 4732–4739. doi: 10.1002/chem.201805950
- Weiss, M. C., Sousa, F. L., Mrnjavac, N., Neukirchen, S., Roettger, M., Nelson-Sathi, S., et al. (2016). The physiology and habitat of the last universal common ancestor. *Nat. Microbiol.* 1:16116. doi: 10.1038/nmicrobiol.2016.116
- Westheimer, F. H. (1987). Why nature chose phosphates. *Science* 235, 1173–1178. doi: 10.1126/science.2434996
- White, H. B. 3rd (1976). Coenzymes as fossils of an earlier metabolic state. *J. Mol. Evol.* 7, 101–104. doi: 10.1007/BF01732468
- Wimmer, J. L. E., Vieira, A. D. N., Xavier, J. C., Kleinermanns, K., Martin, W. F., and Preiner, M. (2021). The autotrophic core: An ancient network of 404 reactions converts H<sub>2</sub>, CO<sub>2</sub>, and NH<sub>3</sub> into amino acids, bases, and cofactors. *Microorganisms* 9:458. doi: 10.3390/microorganisms9020458

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Wimmer, Kleinermanns and Martin. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.