Polyphosphate-Peptide Synergy and the Organic Takeover at the Emergence of Life

E. James Milner-White, Ph.D. 1, and Michael J. Russell, Ph.D. 2,

1 Institute of Biomedical and Life Sciences, University of Glasgow, G12 8QQ, UK
2 Jet Propulsion Laboratory, Planetary Science, Section 3220, MS: 183-601, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA, and Department of Physics, California State University, Fullerton, CA, USA.

Abstract

A number of species are known in which pyrophosphate is used in place of ATP for some of the housekeeping enzymes of metabolism. Pyrophosphates derived from volcanic gasses and dissolved in the earliest ocean as orthophosphate could—when recharged by protons to pyro- and tri-phosphate in a hydrothermal mound—have constituted the main energy storage molecules at an early period in evolution of life, before the advent of nucleic acids and their encoding of amino acids. Triphosphates have been shown to phosphorylate glycines and other amino acids in the laboratory to form cyclic acylphosphoramidates that react with further amino acids to produce dipeptides. The dipeptides then react with such cyclic acylphosphoramidates to generate tripeptides and so on, forming higher oligopeptides. Sequence-independent glycine-rich oligopeptides of eight or more residues are surmised to adopt a conformation in the presence of pyro- and tri-phosphate that both binds them and acts enzymically to catalyze their interconversions with phosphates and pyrophosphates. Oligopeptides and pyro/triphosphates are potentially synergistic as they encourage each other’s synthesis. This synergy would have been a key feature of this early stage in evolution before the advent of the ribosome. A hydrothermal mound forming over a hottish (≤100ºC) alkaline (pH 10-11) spring at the bottom of the carbonic and thereby acidulous (PH 5.5-6) Hadean Ocean affords an appropriate setting for this synergy. The peptide-bound phosphate molecules became arranged in inorganic membranes comprising the outer margins of the hydrothermal mound in such a way that the protons needed for this reaction come from outside of the existing large pH gradient across an inorganic divide, driving pyro/triphosphate, and hence also oligopeptide, synthesis.
1. Introduction

"One basic evolutionary principle is that every living organism or structure or function had ancestors very similar to itself, but simpler". Eck and Dayhoff 1966

In several bacteria, archaea and eukaryotes pyrophosphate (PPi) can replace ATP as an energy source, such that certain enzymes are specific for PPi rather than ATP (Muller et al., 2001; Huang et al., 2004). Many of the organisms utilising PPi this way are anaerobic. Moreover, certain steps in the metabolism of several plants change from ATP as the high-energy cofactor in anaerobic environments to PPi instead. Knowing that the world was essentially anaerobic up to the generation of oxygen in large amounts by oxygenic photosynthesis leads to the idea that enzymes using PPi instead of ATP might be of ancient origin (Wood, 1985; Farquhar & Wing, 2003). Transmembrane pyrophosphatases (PPases) that give rise to, or create, proton gradients directly are known in bacteria, archaea and also in plant vacuoles (Baltscheffsky and Baltscheffsky, 1992; Baltscheffsky et al., 1999; McIntosh and Vaidhya, 2002; Hirono et al., 2007). Many of these PPases have been sequenced and sixteen transmembrane domains are clearly distinguishable, although no structures are yet known. Other enzymes employing PPi as energy source are phosphoenol pyruvate carboxytransphosphorylase, pyrophosphate-phosphofructokinase and pyruvate-phosphate dikinase (Wood et al., 1977; Wood, 1985).

Long unbranched polyphosphates (polyP) are known to occur in eukaryotes, archaea and bacteria and to be essential molecules in many, and are likely to have played a key role in early evolution (Kulaev and Kulakovskaya, 2000; Rao et al., 2009). H⁺-pyrophosphatases are commonly associated with polyP-rich acidic organelles called acidocalcisomes in several bacteria and protozoa (Luo et al., 2001; Seufferheld et al., 2004; Docampo et al., 2005). Hydrolysis of the terminal phosphates of such molecules have energies comparable to that of PPi, which in turn are comparable to that of ATP, so it is not so surprising that some enzymic reactions are known which prefer, or can utilize, polyphosphate, rather than PPi or ATP. Apart from polyP kinases and phosphatases, these include polyP:AMP phosphotransferases, and certain NAD kinases and glucokinases. Triphosphate (PPPi) is a short polyphosphate that can act as an effective substrate for some of the enzymes that act on polyP in general. However, we are not aware of any that act exclusively on PPPi. Here we draw attention to the interesting finding (Rabinowitz, et al., 1969; Yamanaka et al., 1988; Imai et al., 1999; Ni et al., 2009). that PPi and PPPi react with amino acids to form cyclic acylphosphoramidates (CAPAs) which react further with amino acids or di- or tri-peptides to generate higher oligopeptides. These sequence-independent glycine-rich oligopeptides are presumed here to adopt a conformation in the presence of PPi and PPPi that binds them and acts enzymically to catalyze their interconversions with Pi and PPi. Hence short polyphosphates and peptides promote each other’s synthesis.

2. Initial Conditions

The emergence of life and early evolution is thought to have occurred at the outlets of hottish (40-100C) submarine hydrothermal vents, where reduced and alkaline water enriched in H₂, HS⁻, CH₄ and accompanied by minor concentrations of alkyl sulphides, and traces of tungsten and molybdenum, issued into an acidulous carbonic ocean—a reservoir of Fe²⁺ and Ni²⁺ ions, phosphates and free energy in the form of protons (Yamagata et al., 1991; Russell et al., 1994; Hagan et al., 2006; Russell and Hall, 2006; Martin et al., 2008). Phosphate from volcanic activity would also have dissolved in the carbonic ocean as FeH₂PO₄⁻ at a concentration of between 6 and 9 mmol/kg (Yamagata et al., 1991; Macleod et al., 1994; Russell and Hall, 1997; Hagan et al., 2007). Phosphates and iron (nickel) sulphides are dispersed within large compartmentalized mounds composed of silicates, carbonates and hydroxides precipitated at the mouths of the alkaline vents (Russell and Hall, 2006). If it could be harnessed, the consistent pH gradient across the vents would induce a continuous and reliable supply of free energy to do metabolic work. We suppose that simple metabolic pathways were beaten from carbon dioxide to organic molecules such as acetate, partly based on this permanent local energy supply, additional to the thermodynamic energy released during the geochemical reactions between CO₂ and H₂ (Russell et al., 1994; Nitschke and Russell, 2009, 2010; Lane et al., 2010). This use of an ambient proton-motive force on the inner margins of the submarine hydrothermal mound is what distinguishes early biochemistry from the slow geochemical hydrogenation of carbon dioxide that takes place during hydrothermal convection in the oceanic crust (Russell and Martin, 2004; Martin and Russell, 2007).
3. Peptide-Phosphate Interactions

At neutral or alkaline pH and room temperature the deprotonated amino groups (the pKa of an \(-\text{amino}\) group lies between 9 and 9.5) of L-amino acids nucleophilically attack polyphosphates giving rise to cyclic acylphosphoramidates (CAPAs), which possess a reactive carboxyl-phosphoryl mixed anhydride group. CAPA reacts further via nucleophilic attack of a second glycine at the carboxyl group of CAPA to produce N-phosphorylated dipeptides. Dephosphorylation via hydrolysis at higher pH generates the dipeptide as illustrated in Figure 1. If the same reaction occurs, but a dipeptide takes the place of the second glycine, tripeptides are produced. Repetition of the process gives rise to higher polypeptides. The most reactive polyphosphate for this purpose is the cyclic trimetaphosphate, but linear pyro-, tri- and poly-phosphates are all effective (Rabinowitz, et al., 1969; Rabinowitz and Hampai, 1970; Chung et al., 1971; Yamanaka et al., 1988; Yamagata and Inomata, 1997; Imai et al., 1999; Gao et al., 2008; Ni et al., 2009).

Figure 1. How triphosphates are thought to assist polypeptide formation in the alkaline to neutral conditions within the hydrothermal mound. The \(\alpha\)-amino group of an amino acid (pKa = 9-9.5) attacks triphosphate to produce a CAPA which is in turn attacked by the \(\alpha\)-amino group of a second amino acid at the carboxyl group of the CAPA, resulting in an N-phosphorylated dipeptide. Hydrolysis removes the phosphate to leave the dipeptide. The key intermediates, CAPAs, are cyclic acylphosphoramidates.

Figure 2. Peptide Nests. Pymol images of a) a nest bound to an oxygen atom; b) three overlapping nests belonging to a P-loop in the G-protein P21\(^{ras}\) bound to a phosphate; c) four overlapping nests bound to an \(\text{Fe}_3\text{S}_4\) center in archael ferredoxin. Oxygen atoms are red, nitrogen blue, carbon green, phosphorous orange, iron brown and sulfur yellow. The peptide side chains are omitted.

Nests are simple features of proteins and peptides in which the main chain NH groups of alternating peptide bonds are bridged by forming hydrogen bonds with an anionic or \(\delta\)-charged atom. The concave shape of the main chain part of the three amino acid
residues of the peptide is termed a nest (Watson and Milner-White, 2002; Milner-White et al., 2004; Milner-White and Russell, 2005, 2008). One is illustrated in Figure 2a. Seven percent of all amino acid residues in proteins are involved in nests. If more than three consecutive CONH groups adopt the same main chain conformation then overlapping nests form, since the NH groups all point roughly towards the center of a single, wider, concavity. If so, there is the possibility of their binding a group of anionic atoms, as in Figures 2b and 2c. In this conformation the nest, in combination with the ionic "egg", acts as an entropy trap and sometimes also as a proximity catalyst (cf., Hsiao and Williams, 2009). Another relevant aspect is that in such peptides the main chain conformations of adjacent residues are approximately enantiomeric, so that glycines are fairly common at alternate positions.

The commonest ATP or GTP-binding feature in proteins is the P-loop (Saraste et al., 1990; Vetter and Wittinghofer, 1999; Leipe et al., 2002), illustrated in Figure 2b. It is often known by its consensus sequence motif GxxxxGKx. A key feature is created by three overlapping nests at residues xxGKx. The mainchain NH groups of these residues form several hydrogen bonds to the oxygen atoms of the β-phosphate of ATP or GTP. In most P-loop proteins the enzyme reaction catalyzed is removal, or addition, of the γ-phosphate group, either by means of an ATPase, ATP synthase or GTPase, or via phosphoryl transfer from ATP to another substrate in a kinase reaction. However, it is of particular interest for the present work that the H⁺-PPases have a sequence signature characteristic of a P-loop so, even though its 3D structure has not been determined, phosphate is expected to bind to nests in much the same way as in the other P-loop proteins (Drozdowicz et al., 2003; Hedlund et al, 2006). This is illustrated in the left half of Figures 3a and 3b.
Apart from the P-loop, several other peptides consisting of overlapping nests are well-known that may have been common in early evolution. About a half of Fe\(_2\)S\(_2\) and Fe\(_4\)S\(_4\) centers are bound by nests, as in Figure 2c; they are mentioned later. Most is known about phosphate binding and one aspect that stands out is that, while P-loops are thought to be particularly ancient features, the protein binding sites for more recent phosphorylated metabolites like IP\(_3\) and phosphorylated proteins usually consist of positively charged lysine, arginine or histidine side chains, sometimes as clusters (Yaffe and Smerdon, 2004). This leads to the idea that the P-loop and other overlapping nest motifs evolved during the early (pre-genetic code) period before amino acid side chains were reliable whereas, once positively charged side chains were genetically encoded, they were employed instead (Milner-White and Russell, 2005; Caetano-Anollés et al., 2007).

### 4. How P-Loops May Act As A Pyrophosphatase

Given the likely ancient nature of P-loop proteins and that most of them catalyze transfer of a phosphate either to a substrate or to a water molecule, a version of a P-loop peptide is presumed to have been functionally important. We suppose that the H\(^+\)-PPase is a model for an early protein-like peptide. However at this pre-genetic code period it is likely that the participating amino acids would have been uncertain. Glycines were probably common at an early stage in evolution and the other amino acids would have consisted of mixtures of L and D forms (Hennet et al., 1992; Holm et al., 1995; Huber and Wachterhauser, 2003). However, from peptides in the Cambridge Data Base, it is evident that alternating heterochiral peptides have a particular propensity to form nests (Milner-White et al., 2004). So do peptides with a high glycine content. Both amino acid compositions are compatible with enantiomeric peptides like nests. Thus a short peptide of this type is likely to form readily into a P-loop and operate as an H\(^+\)-PPase whatever the membrane composition and thickness.

Just as the atomic groups of the amino acid side chains are not integral to the function of nests, the same argument should apply to the groups on the enzyme employed catalytically. We expect that a positively charged group might be needed, as happens in most P-loop proteins, to improve binding to the phosphate, so, when reliable side chains were not available, the N-terminal amino group was used. On the other hand a nucleophilic group is needed to act as a base catalyst to remove a proton from a water molecule so the C-terminal carboxylate was used as that group. The resulting hydroxide ion attacks the other phosphate, causing hydrolysis of pyrophosphate. It is likely that a magnesium ion chelates the pyrophosphate to offset the repulsion between the anions, since a magnesium-pyrophosphate complex is the substrate for H\(^+\)-pyrophosphatases (McIntosh and Vaishya, 2002; cf. Hsiao and Williams, 2009). Magnesium is likely to have been present in Hadean Ocean water, contributed by low temperature.

Figure 3. **Pyrophosphatase Mechanism.** How the main chain atoms of a short peptide may catalyze a. the pyrophosphatase reaction PPi -> Pi + Pi and b. the reverse reaction, pyrophosphate synthase.
hydrothermal solutions (Barnes and O’Neil, 1969; Palandri and Reed 2004; Hopkinson et al., 2004). The mechanism is shown in Figure 3a.

The pyrophosphatase peptide could act to synthesize as well as hydrolyze pyrophosphate by catalyzing the nucleophilic attack of one phosphate molecule upon another. This pyrophosphate synthase reaction is the reverse of the pyrophosphatase in the sense that the C-terminal carboxylate in the PPase acts as an acidic rather than a basic group; protonation of one phosphate renders it more susceptible to nucleophilic attack by the other phosphate. The reaction is inherently pH-dependent, with low pH tending to favour the synthase (Fig 3b) and high pH the pyrophosphatase (Fig 3a). A magnesium ion plays a similar role in the forward and reverse reactions by compensating for the negative charges on the two phosphate anions. However, the synthase reaction is decidedly endergonic at neutral pH and it seems likely that some means arose whereby energy is utilized for its synthesis (and see de Zwart et al., 2002). For example the peptide could be situated in a membrane such that the inherent energy of the proton gradient across it is harnessed to favour pyrophosphate synthesis. Precisely how this occurs is difficult to tell, except that it happens in present-day H⁺-pyrophosphatases (Baltscheffsky et al., 1999). The diagonal arrows at the left hand side of Figure 4 indicate the requirement for this coupling.

By analogy with the majority of P-loop proteins which act on triphosphates, the peptidic enzyme in Figure 3 would be expected to catalyze PPPi, as well as PPI, hydrolysis and also the reverse synthase reaction, PPI + Pi → PPPi. For such a reaction the left-hand phosphate in Figures 3a and 3b is replaced by pyrophosphate. The peptide would also be liable to catalyze the corresponding reactions for the longer polyphosphates to some degree, as mentioned later.

![Figure 4. Synergy between short peptides and polyphosphates. The decapeptide aa-aa-aa-aa-aa-aa-aa-aa-aa-aa-aa acts as an enzyme that, driven by a proton gradient, phosphorylates phosphate to form pyrophosphate. The same enzyme phosphorylates pyrophosphate to form triphosphate. Triphosphate reacts with amino acids and phosphorylates them, giving rise to polypeptides after their dephosphorylation. In doing so triphosphate becomes hydrolyzed to phosphate and pyrophosphate. The effect is synergy between peptide and polyphosphates. Polyphosphates are required for making polypeptides which in turn catalyze polyphosphate formation. For the sake of simplicity, longer polyphosphates such as PPPPi and PPPPPP are not illustrated, but they may have played an active part in the process, just as longer peptides do.](http://journalofcosmology.com/Abiogenesis100.html)

5. Organic Takeover: Synergy Between Peptides and Polyphosphates

The diagram in Figure 4 shows how simple peptide elongation based on sequence-independent peptides and polyphosphates may have functioned at an early stage in evolution at alkaline hydrothermal vents under the ocean, before the advent of the ribosomal peptidyl transferase center. Without self-replicating molecules there has to be some means for molecules to enhance their
synthesis and mutual synergy may be the precursor mechanism. Polyphosphates react with amino acids chemically causing the formation of peptides while the peptides catalyze the formation of polyphosphates. The polyphosphates in Figure 4 are limited to P Pi and P Pi, but longer chains may well act both as energy sources and to enhance peptide synthesis. Because the peptides employ main chain atoms for their functions they are to some degree independent of their amino acid side chains and their chirality. The reader will appreciate that the process in Figure 4 has a certain dynamic. Once it gets going the peptides will tend to become longer and begin to engage in more complex activities. For example some of the peptides will, assisted by alkyl thiolates from the hydrothermal fluid, bind the Fe$_2$S$_2$ and Fe$_4$S$_4$ clusters that otherwise would go to comprise the crystal lattices of mackinawite [FeS]$_n$ and greigite [as {[SFeS][Fe$_4$S$_4$][SNiS]}$_n$] (Eck and Dayhoff, 1966; Bonomi et al., 1985; Stevens and Kurtz; 1985; Milner-White and Russell, 2005, 2008). And shorter peptides have been shown to bind single ions of nickel or cobalt directly (Bossu et al., 1978; Hawkins and Kelso, 1982). These various complexes will develop the catalytic power to engage in more ambitious forms of metabolism. Moreover, the longer peptides would act as binding agents in the membrane, affording it more flexibility and cell-wall-like strength while retaining catalytic sites and permitting channel formation. We imagine peptides, and beyond that proteins, to have expedited the eventual organic takeover from mineral compartments and catalysts, while the active inorganic clusters were tuned to become ever more efficient in their energy-providing and catalytic roles (Zhang et al., 1993; Baymann et al, 2002; Milner-White and Russell, 2008; McGlynn et al., 2009). Later still in evolutionary time came RNA molecules with all the possibilities they have brought in terms of catalysis, self-replication and encoding of amino acids via the ribosome (Martin and Russell, 2007; Hsiao et al., 2009; Yarus, 2010).

**Acknowledgments:** We thank Loren Williams, Shawn McGlynn, Isik Kanik and Richard Kidd for discussions and support. MJR's research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration: Exobiology and Evolutionary Biology and supported by NASA’s Astrobiology Institute (Icy Worlds). Copyright 2010 California Institute of Technology, Government sponsorship acknowledged.

**References**


Nitschke, W., Russell, M.J. (2009). Hydrothermal focusing of chemical and chemiosmotic energy, supported by delivery of catalytic Fe, Ni, Mo/W, Co, S and Se, forced life to emerge. Journal Molecular Evolution, 69, 481-496.


Watson, J.D., Milner-White, E.J. (2002). A novel main-chain anion-binding site in proteins: The Nest. A particular combination of phi, psi values in successive residues gives rise to anion-binding sites that occur commonly and are found often at functionally important regions. Journal of Molecular Biology, 315, 171-182.


http://journalofcosmology.com/Abiogenesis100.html