Primitive Genetic Polymers

Aaron E. Engelhart and Nicholas V. Hud

School of Chemistry and Biochemistry, Georgia Institute of Technology, Atlanta, Georgia 30332

Correspondence: hud@chemistry.gatech.edu

Since the structure of DNA was elucidated more than 50 years ago, Watson-Crick base pairing has been widely speculated to be the likely mode of both information storage and transfer in the earliest genetic polymers. The discovery of catalytic RNA molecules subsequently provided support for the hypothesis that RNA was perhaps even the first polymer of life. However, the de novo synthesis of RNA using only plausible prebiotic chemistry has proven difficult, to say the least. Experimental investigations, made possible by the application of synthetic and physical organic chemistry, have now provided evidence that the nucleobases (A, G, C, and T/U), the trifunctional moiety ([deoxy]ribose), and the linkage chemistry (phosphate esters) of contemporary nucleic acids may be optimally suited for their present roles—a situation that suggests refinement by evolution. Here, we consider studies of variations in these three distinct components of nucleic acids with regard to the question: Is RNA, as is generally acknowledged of DNA, the product of evolution? If so, what chemical and structural features might have been more likely and advantageous for a proto-RNA?

In contemporary life, nucleic acids provide the amino acid sequence information required for protein synthesis, while protein enzymes carry out the catalysis required for nucleic acid synthesis. This mutual dependence has been described as a "chicken-or-the-egg" dilemma concerning which came first. However, requiring that these biopolymers appeared strictly sequentially may be an overly restrictive preconception—nucleic acids and *noncoded* peptides may have arisen independently and only later become dependent on each other. Nevertheless, the requirements for the chemical emergence of life would appear simplified if one polymer was initially able to store and transfer information as well as perform selective chemical catalysis—two essential features of life.

The discovery of catalytic RNA molecules in the early 1980s (Kruger et al. 1982; Guerrier-Takada et al. 1983) created widespread interest in an earlier proposal (Woese 1967; Crick 1968; Orgel 1968) that nucleic acids were the first biopolymers of life, as nucleic acids transmit genetic information and could have once been responsible for catalyzing a wide range of reactions. The ever-increasing list of processes that involve RNA in contemporary life continues to strengthen this view (Mandal and Breaker 2004; Gesteland and Atkins 2006). Furthermore, the rule-based one-to-one pairing of

Advanced Online Article. Cite this article as *Cold Spring Harb Perspect Biol* doi: 10.1101/cshperspect.a002196

A.E. Engelhart and N.V. Hud

complementary bases in a Watson-Crick duplex (Fig. 1) provides a robust mechanism for information transfer during replication that could have been operative from the advent of oligonucleotides. In contrast, there is no obvious and general mechanism by which the amino acid sequence of a polypeptide can be transferred to a new polypeptide as part of a replication process.

If we accept that nucleic acids must have appeared without the aid of coded proteins, we are still faced with the question of how the first nucleic acid molecules came to be. Broadly defined, there are two schools of thought regarding the origin of the earliest nucleic acids. In one school, it is proposed that abiotic chemical processes initially gave rise to nucleotides (i.e., phosphorylated nucleosides), which were then coupled together to yield polymers identical in chemical structure to contemporary RNA. In support of this model, Sutherland presents in his article current progress toward discovering possible chemical pathways for the prebiotic synthesis of RNA mononucleotides, as well as methods for their protein-free polymerization (Sutherland 2010).

A second school of thought, discussed in this article, considers RNA to be a product of evolution, and that a different RNA-like polymer (or proto-RNA) was used by the earliest forms of life. Just as the deoxyribose sugar of DNA was likely the product of Darwinian evolution (selected for the hydrolytic stability it provides this long-lived biopolymer), so, too, might the sugar, phosphate, and bases of RNA have been refined by evolution. In this scenario, a proto-RNA is more likely to have spontaneously formed than RNA, because a proto-RNA could have had more favorable *chemical* characteristics (e.g., greater availability of precursors and ease of assembly), but such a polymer was eventually replaced, through evolution, by RNA (potentially after several incremental changes), based on *functional* characteristics (e.g. nucleoside stability, versatility in forming catalytic structures). Thus, contemporary RNA may possess chemical traits that, although optimally suited for contemporary life, may have been ill-suited for the earliest biopolymers, with the converse being true for proto-RNA.

BACKGROUND

Challenges to "Reinventing" Proto-RNA

If proto-RNA came before RNA, possibly comprised of different bases and a different backbone, then the chemical space of proto-RNA candidates seems almost limitless. Leslie Orgel once called this scenario a "gloomy prospect" with regard to solving the origin of life (Orgel
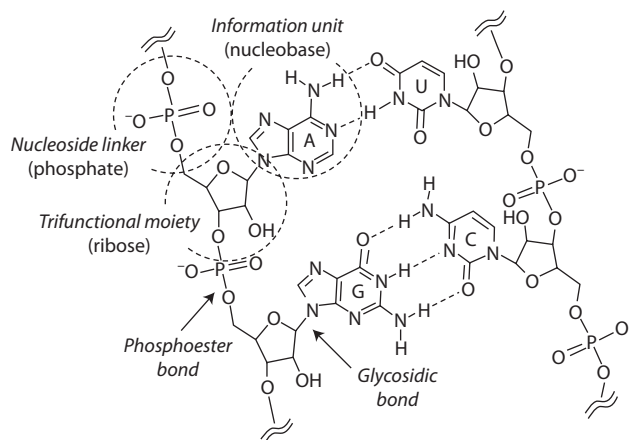

Figure 1. Two base-paired RNA dinucleotide steps with functional units discussed in the text annotated. In contemporary life, the nucleoside linker is phosphate, and the information unit is one of the canonical nucleobases (A, G, C, and U). The contemporary trifunctional moiety, ribose, is coupled via N,O-acetals to the informational unit and via phosphoesters to the nucleoside linker.

Advanced Online Article. Cite this article as *Cold Spring Harb Perspect Biol* doi: 10.1101/cshperspect.a002196

1998). However, we do not see the prospect of proto-RNAs as reason for pessimism. The possibility that RNA had one or more predecessors also implies the possibility that there was a polymer with more facile chemistry of assembly, which researchers could show to spontaneously assemble from simple precursors. Additionally, the chemical space within which proto-RNA was formed is constrained by the molecules that could have been present on the prebiotic Earth. Although we will never know the precise contents of the *prebiotic chemical inventory*, model prebiotic reactions continue to provide valuable information regarding plausible prebiotic molecules (Miyakawa et al. 2002; Plankensteiner et al. 2004; Saladino et al. 2004; Cleaves et al. 2006). Additionally, studies of our neighboring planets and their moons, meteorites, comets and even interstellar space (Schwartz and Chang 2002; Hollis et al. 2004; Pizzarello 2006; Martins et al. 2008) also provide clues to the types of molecules from which a proto-RNA might have emerged.

Sugars and nucleobases, two essential building blocks of nucleic acids, have been identified as products in a rather wide range of model prebiotic reactions. However, feasible prebiotic reactions have not been shown for the efficient coupling of ribose to all four nucleobases to create the canonical nucleosides, and only limited evidence has been presented that preformed nucleotides might be able to polymerize without resorting to nonprebiotic chemical activation. Thus, a conceptual gap still exists between our ideas for how the building blocks of RNA might have formed from the simple molecules that are ubiquitous in the universe versus the chemical processes that gave rise to the first RNA-like polymers.

All the present challenges to bridging this conceptual gap between the "small molecule world" and the "polymer world" can be classified as belonging to one of two major problems: molecular selection and polymerization reactions. There was likely a wide variety of molecules in the prebiotic chemical inventory, so the first formidable challenge is to find an abiotic mechanism by which "useful" building blocks were selected from a complex mixture.

The second formidable challenge is to find reactions by which these building blocks could have been "correctly" joined to create "useful" (i.e., self-replicating, catalytic or both) proto-biopolymers. In the following sections, we will elaborate on the challenge of RNA building block selection and assembly, as well as possible solutions to "reinventing," to use Albert Eschenmoser's term (Eschenmoser 2007), a process for the de novo synthesis of an RNA-like polymer from plausible prebiotic molecules and reactions.

Some Possible Constraints to Guide the Search for Proto-RNA

Synthetic chemists have prepared numerous nucleic acid analogs, some as part of origin of life studies (Schneider and Benner 1990; Pitsch et al. 1993; Herdewijn 2001b; Benner 2004; Mittapalli et al. 2007a; Mittapalli et al. 2007b) and many as part of a widespread search for therapeutic agents (Kurreck 2003). These analogs have included changes to both the nucleobases and to the polymer backbone. With regard to origin of life investigations, some proposals for the difference between RNA and an earlier proto-RNA have been subtle, such as the proposal that hypoxanthine once functioned in the place of guanine (these nucleobases differ by only a single exocyclic amino group) (Crick 1968). Other proposals are much more radical, such as the suggestion that the nucleobases were once connected by a completely different and uncharged backbone, as in the synthetic polymer known as peptide nucleic acid (Nelson et al. 2000), and intermediate proposals, in which the backbone was a peptide, but still negatively charged (Mittapalli et al. 2007a; Mittapalli et al. 2007b) (Fig. 2). Physical and chemical studies of these and other nucleic acid analogs have provided invaluable information regarding which alternative polymers might have predated RNA. An exhaustive review of all proposed predecessors of RNA is beyond the scope of this article, and several excellent reviews have been published considering possible variations in sugars and bases, to which the reader is referred (Benner 2004; Benner et al. 2004; Eschenmoser 2007).
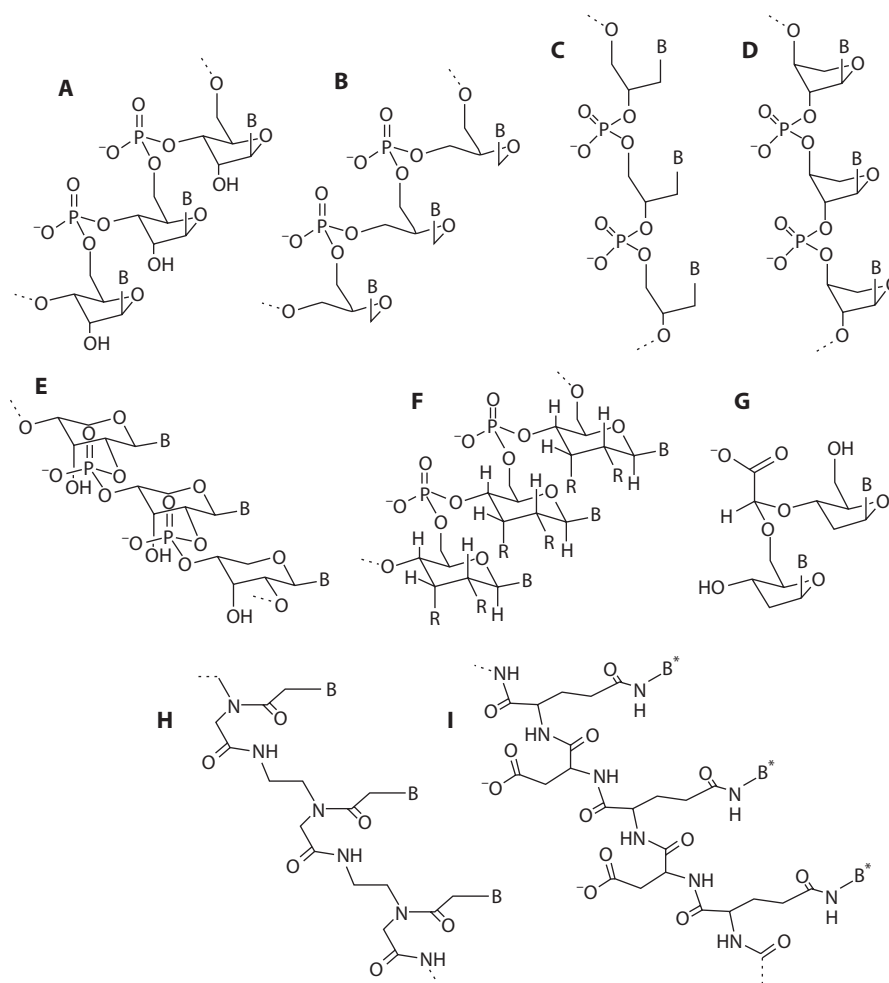
A.E. Engelhart and N.V. Hud

Figure 2. Structures of RNA and selected analogs discussed in the text. (*A*) RNA; (*B*) FNA or "flexible nucleic acid," originally named glycerol nucleic acid; (*C*) GNA, glycerol nucleic acid; (*D*) TNA, α-threofuranosyl nucleic acid; (*E*) pRNA, β-pyranosyl-RNA; (*F*) homo-DNA (R = H) and β-allopyranosyl (R = OH); (*G*) gaNA, glyoxylate-linked dinucleotide; (*H*) PNA, peptide nucleic acid; (*I*) poly-(L-Asp-L-Glu) with the γ-carboxyl function of Glu conjugated to a nucleobase by an isopeptide linkage. For all structures shown, B represents one of the canonical nucleobases, and B*, in structure I, represents a noncanonical base.

In this article, we focus on a subset of potential candidates for proto-RNA polymers that, in our opinion, would be closest to RNA in structure and physical properties but are composed of alternative building blocks that would have made prebiotic assembly more feasible. One could consider these structures as candidates for the later predecessors of RNA that resulted from proto-RNA evolution or, optimistically, as potential candidates for the first proto-RNA, if only a few changes were required to reach

RNA. Our selection of proposed proto-RNA candidates on which to focus is guided by three hypotheses:

(1) *The original molecular components of proto-RNA may have been different from, but still similar to, those found in RNA.* Three distinct molecular components comprise RNA: the nucleobases, ribose and phosphate (Fig. 1). Each component has specific roles for which it appears to be very well (perhaps

optimally) suited. However, because of the limited availability of some components, or the difficulty in forming covalent bonds between others, there is good reason to entertain the possibility that alternative components were initially used by proto-RNA. Each of these three components could have been replaced on different timescales over the course of evolution. Although this criterion may seem obvious regarding what would be recognized as RNA-like polymers, it does exclude from our discussion the more radical models of "genetic takeover" (Cairns-Smith 1982), in which an inorganic or decidedly non-RNA-like system gave rise to RNA, models which come with the added burden of deciphering how a physicochemical transition could have transpired between completely different molecular systems.

(2) *The covalent bonds that connected the three molecular components of proto-RNA were (periodically) reversible.* In the early stages of life, it would have been highly advantageous for chemical building blocks to be available for reuse. Mechanisms for recycling are easy to imagine, if the covalent bonds that joined the molecular components of proto-biopolymers could have been formed, broken and re-formed repeatedly. Such a process would have allowed proto-biopolymers to be created that were thermodynamically favored structures (Li et al. 2002; Hud et al. 2007; Ura et al. 2009) and for "errors" in synthesis to be corrected (e.g., replacement of non- or mispaired nucleobases with pairing ones). Proto-RNA monomers might have even been repeatedly recycled into polymers with different nucleotide sequences (and corresponding functions) as survival pressures changed (Ura et al. 2009). Furthermore, without some regular cycle by which RNA or a proto-RNA could have been depolymerized, any monomer incorporated into a polymer that was not functionally useful would be wasted—a condition under which there simply may

have not been enough material to get life started.[1]

(3) *Condensation-dehydration reactions were, from the beginning, integral to the formation of proto-RNA.* A wide range of biological bond-forming reactions involve the loss of water, including amide linkages in peptides, acetal linkages in polysaccharides and ester linkages in phosphoglycerols. Pertinent to the present discussion are the reactions required for nucleic acid formation—glycosylation of a nucleobase to form a nucleoside, phosphorylation of a nucleoside to form a nucleotide and condensation of nucleotides to form oligonucleotides—all of which are dehydration reactions. Such reactions are very appealing from a prebiotic standpoint, as their equilibria can be modulated via water activity (i.e., through drying-wetting cycles), suggesting that many polymerization reactions observed in contemporary life may have ancient roots in drying reactions driven by periodically fluctuating water activity—perhaps even dating back to the earliest stages of life.

We now will discuss the three molecular components of nucleic acid polymers and consider the implications of the three hypotheses presented earlier in considering alternative candidates for proto-RNA.

RECENT RESULTS

Selection of the Nucleobases: Was it C, U, A, and G from the Start?

Although nucleic acid chemists have provided some insights regarding what might have

---

[1]If the earliest effective ribozyme was 50nt in length (similar to the hammerhead ribozyme), and the earliest genetic code contained C, U, A, and G, this ribozyme would be one of $4^{50} \approx 10^{30}$ possible sequences. A collection of *one* molecule of every possible 50mer would require a carbon mass of around 10^7 kg. Assuming each polymer was produced in equal yield, and by an irreversible process, production of 1 picomole of the ribozyme would require 10^{19} kg of carbon (roughly the total weight of carbon in the Earth's crust).

predated RNA, as we hope we have conveyed herein, we concede that it is still not obvious by what mechanism the nucleobases were originally selected. The ability to form Watson-Crick base pairs should not be assumed to have been the sole criterion. One observation that runs counter to this assumption is that the free nucleobases and their nucleosides *do not* form Watson-Crick base pairs in aqueous solution (Ts'o et al. 1963), but rather, they form co-planar stacks, with their Watson-Crick edges interacting with the solvent. The hydrogen bonds formed with water are of comparable enthalpy to those that would be formed between paired bases. The local preorganization of nucleobases in oligonucleotides, on the other hand, promotes intra- and interstrand stacking of base pairs, which contributes substantial free energy to duplex stability, whereas Watson-Crick hydrogen bonds have been reported to contribute virtually no net free energy (De Voe and Tinoco 1962; Yakovchuk et al. 2006).

The observation that free nucleobases do not form Watson-Crick base pairs in aqueous solution inspired us to articulate a "paradox of base pairing" with regard to the emergence of RNA (Hud and Anet 2000; Hud et al. 2007). *Why would the nucleobases have been selected for inclusion in RNA or proto-RNA polymers for the purpose of Watson-Crick base pairing if they did not form base pairs before being linked by a common backbone?* To argue that the nucleobases formed nucleotides with ribose and phosphate *before* being involved in base pairing seems to imply that nature either predicted that the nucleobases would be useful for pairing polymers in the future (which would violate the nonpredictive property of evolution); that nature was simply lucky, and a series of chemical reactions coincidentally created rather complex molecules (i.e., nucleotides) that happened to form base pairs after polymerization; or that such molecules existed in a metabolism-first origin of life scenario, and these molecules were later incorporated into a genetic polymer. Serious conceptual challenges have been raised against metabolism-first scenarios (Anet 2004; Orgel 2008; Vasa et al. 2009). In any case, arguments for the prepolymer production and use

of nucleotides by an abiotic metabolic cycle remain largely philosophical, as we simply do not know at what point in evolution (of both metabolism and genetic polymers) that the specific components of contemporary nucleotides were selected. It is feasible, for example, that phosphate was used first as a metabolic cofactor before being incorporated into genetic polymers.

As one possible solution to the paradox of base pairing, we have proposed that there was a template molecule, or a class of molecules, present in the prebiotic chemical inventory, that formed stacks with the nucleobases that were energetically favored over stacks with themselves. Each template, or "molecular midwife," would have been large enough to accommodate two or more nucleobases, if the bases were arranged in a particular hydrogen-bonded structure (e.g., a Watson-Crick base pair) (Hud and Anet 2000; Hud et al. 2007). The stacking of multiple midwife molecules, interleaved with nucleobases, would have locally concentrated and organized the bases, thereby aiding in the polymerization of the earliest proto-RNA. These midwife molecules would have no longer been necessary for polymer synthesis (and replication) once evolution had produced a superior means for carrying out these reactions (e.g., catalysis via protein or RNA enzymes). We envision that the midwife molecules would have been similar to the small planar molecules known to intercalate the bases of contemporary RNA and DNA (Ihmels and Otto 2005). Beyond a purely speculative model, dye molecules that intercalate DNA and RNA base pairs have been shown to promote the template-directed synthesis of nucleic acids. For example, intercalators can provide a >1000-fold enhancement to the ligation of short oligonucleotides (i.e., tri- and tetranucleotides) (Jain et al. 2004), and promote a specific base pairing (Hud, Jain et al. 2007; Horowitz et al. 2010).

Should we assume that Watson-Crick base pairs (i.e., adenine paired with uracil, guanine paired with cytosine) were used by the earliest proto-RNA? The successful incorporation of nonnatural base pairs into nucleic acid duplexes has given ample reason to seriously consider

this question (Switzer et al. 1989; Piccirilli et al. 1990; Geyer et al. 2003). Most of the nonnatural base pairs studied thus far have maintained the form of a base comprised of a six-membered ring (like a pyrimidine) paired with a base comprised of fused five- and six-membered rings (like a purine), which can be accommodated in a helix also containing canonical Watson-Crick base pairs. However, duplexes that contain all noncanonical base pairs have also been shown, including duplexes with only purine–purine base pairs (Groebke et al. 1998; Battersby et al. 2007; Heuberger and Switzer 2008b; Engelhart et al. 2009), pyrimidine–pyrimidine base pairs (Mittapalli et al. 2007a; Mittapalli et al. 2007b), and even duplexes with tricyclic bases (Krueger et al. 2007).

If we restrict our consideration of potential proto-RNA base pairs to those with a Watson-Crick pairing geometry, there are still twelve possible types of nucleobases that can form six types of base pairs with distinct hydrogen bond donor–acceptor arrangements (Fig. 3) (Benner et al. 2004), of which only two types

are commonly found in contemporary life. Some of these base pairs, at first glance, seem like reasonable candidates as proto-RNA alternatives to the canonical base pairs, such as isoC paired with isoG (Fig. 3, bottom left). However, even this close analog of the guanine-cytosine base pair comes with significant chemical challenges that could have prevented it from being a component of proto-RNA (Jaworski et al. 1985). In particular, isoguanine exists to about 10% as its enol tautomer in water (Sepiol et al. 1976; Voegel et al. 1993; Seela et al. 1995), which has a hydrogen-bonding pattern that is complementary to U/T. Moreover, some of the four noncanonical Watson-Crick-like base pairs could only be realized experimentally by using a nonpyrimidine or a nonpurine base, or by the connection of the base to the sugar through a carbon-carbon bond (C-nucleosides) (Fig. 3) (Benner 2004). The use of non- purine or pyrimidine heterocycles may not be so problematic from a prebiotic standpoint. Although purines are abundant in model reactions of the prebiotic chemical inventory using, for instance,
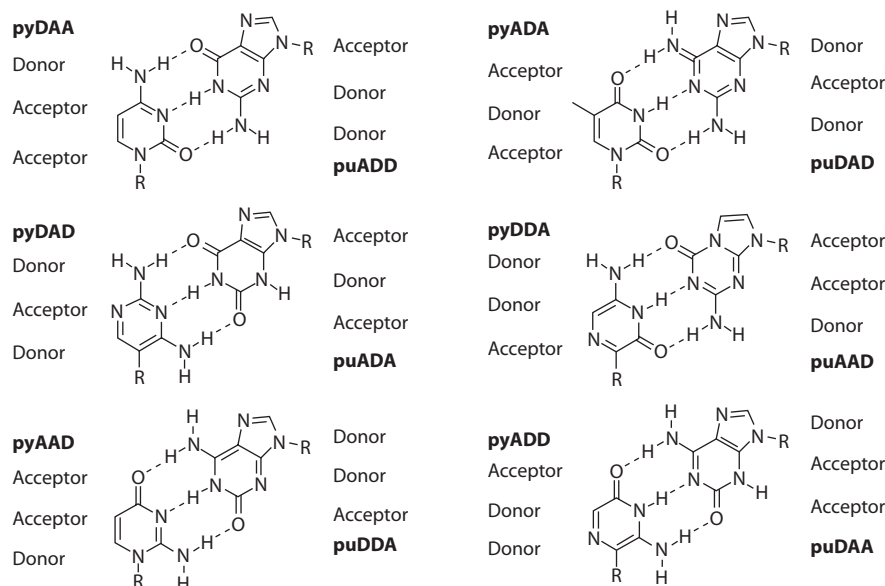

Figure 3. The six possible hydrogen-bonding patterns that can join 12 different nucleobases with Watson-Crick geometry. Nomenclature: pu, a [6,5] fused ring system; py, a six-membered ring. The hydrogen-bonding pattern of acceptor (A) and donor (D) groups from the major to the minor groove is indicated. For example, the standard nucleobase cytosine is pyDAA. R indicates the point of attachment of the backbone. Reprinted with permission from Benner et al., 2004. Copyright Elsevier Science Ltd.

A.E. Engelhart and N.V. Hud

formamide and HCN, other, helix-compatible parent heterocycles were surely present in the prebiotic chemical inventory (Mittapalli et al. 2007a; Mittapalli et al. 2007b), and these should be considered. With respect to C-nucleosides, although they are well-known, they have been less well-examined from a prebiotic standpoint than the canonical N,O-acetal linked nucleosides. Regardless of any potential difficulties in prebiotic synthesis, the demonstration of all six possible base pairs with a Watson-Crick paring geometry, made possible through the application of synthetic and physical organic chemistry, has provided invaluable constraints regarding what should be considered more or less likely as possible ancestral proto-RNA building blocks.

Among the proposals for proto-RNA noncanonical base pairs, that of purine-purine base pairs is one of the oldest (Crick 1968). The abiotic formation of proto-RNA containing only purine bases is appealing for several reasons. First, an all purine system requires only one type of heterocycle, and the variety of conditions under which purine production has been shown suggests that purines may be formed more easily in model prebiotic reactions than pyrimidines (Oro 1961; Sanchez et al. 1966; Saladino et al. 2004). Second, nucleoside formation, as discussed later, might have occured more efficiently with purine bases. Third, the purine bases have much more favorable stacking interactions in aqueous solution than the pyrimidine bases (Ts'o, Melvin et al. 1963; Inoue and Orgel 1983). The potential prebiotic obstacles presented by poor pyrimidine stacking in aqueous solution is illustrated by the inhibition of template-directed polymerization reactions that require the coupling of successive pyrimidine nucleotides (Joyce 1987). With respect to the molecular midwife hypothesis discussed earlier, purine−purine base pairs would also be expected to form assemblies with midwife molecules at lower concentrations and over a wider temperature range than Watson-Crick base pairs.

Crick was among the first to propose that life originally used purine-purine base pairs. In particular, he proposed that A·I base pairs

might have come first (Fig. 4) (Crick 1968). However, duplexes formed by homo-A and homo-I polymers proved to be relatively unstable, and these polymers were prone to forming triplexes (Howard and Miles 1977). Eschenmoser and coworkers have subsequently shown that homo-purine oligonucleotides (with G·isoG base pairs) can form very stable duplexes when the standard ribofuranose nucleosides are replaced by either a ribopyranose or a dideoxyhexopyranose sugar (Krishnamurthy et al. 1996; Groebke et al. 1998) (other sugar modifications are discussed later). More recently, evidence has been presented that standard oligodeoxynucleotides can form duplexes with mixed G·isoG and diaminopurine·xanthine base pairs (Fig. 4) that are of comparable stability to Watson-Crick duplexes (Heuberger and Switzer 2008a). In light of these recent advances, the decades-old idea of purine-purine
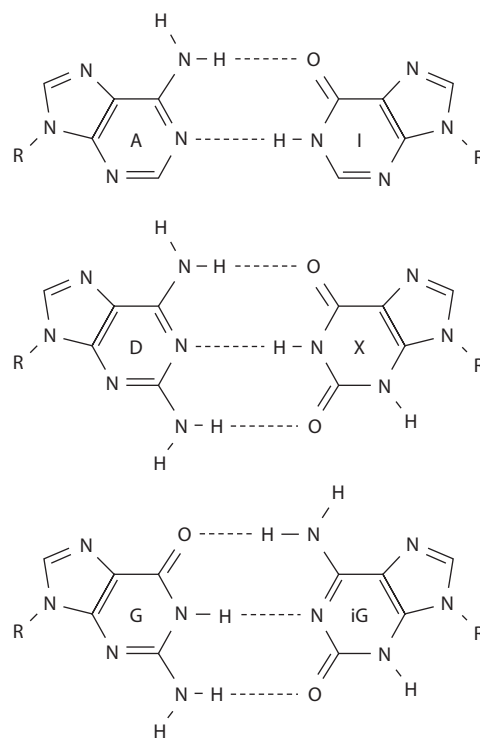

Figure 4. Chemical structures of purine−purine base pairs. A · I, adenine with hypoxanthine (the base of inosine); D · X, diaminopurine with xanthine; G · iG, guanine with isoguanine.

Advanced Online Article. Cite this article as *Cold Spring Harb Perspect Biol* doi: 10.1101/cshperspect.a002196

base pairs in proto-RNA now looks even more attractive.

The Trifunctional Moiety: Was It Ribose from the Start?

In RNA, ribose serves as the branch point between the polymer backbone and the side chains (i.e., the nucleobases). This role requires a chemical moiety that is, at least, trifunctional. Sugars with three or more carbons meet this minimum criterion and have been studied; among those investigated thus far, ribose appears optimal (Eschenmoser 1999). The conformational landscape of the β-furanose anomer of ribose is well accommodated in a helical arrangement when its nucleosides are linked by phosphate. Although some nucleic acids synthesized with alternative sugars form more stable Watson-Crick duplexes than RNA (Eschenmoser 1999), ribose appears to provide an optimal balance between duplex stability and the ability to adopt the more globular structures required for creating a catalytically active ribozyme, as exemplified by the variety of RNA structures observed in the ribosome (Ban et al. 2000; Wimberly et al. 2000).

The optimal functionality of ribose suggests that its inclusion in RNA is more the result of evolutionary refinement than prebiotic availability. The yield of ribose in the formose reaction, which is often cited as a prebiotically plausible synthetic route to sugars, is low with respect to the numerous other sugars produced in this reaction (Decker et al. 1982). Possible solutions to these challenges have included proposals that prebiotic ribose was preferentially stabilized by adduct formation with cyanamide (Springsteen and Joyce 2004), borate (Ricardo et al. 2004) or phosphate (Müller et al. 1990), each of these adducts being more stable than the unmodified sugar.

If ribose does not stand out for its ease of abiotic synthesis or chemical stability, then what sugars would have been more likely to have come before ribose in proto-RNA? Twenty years ago, Joyce et al. thoughtfully discussed this question (Joyce et al. 1987). With the goal of identifying a predecessor to ribose with enhanced prebiotic availability and tolerance for incorporation of mixed stereoisomers (i.e., D- and L-sugars), these authors proposed a proto-RNA with acyclic analogs of ribose, including an acyclic flexible nucleic acid (FNA) backbone that is tantamount to an RNA backbone with "deletion" of the 2′ carbon (Fig. 2). Subsequently, the Benner laboratory synthesized oligonucleotides containing these linkages and found that each substitution within a DNA oligonucleotide resulted in significantly depressed duplex stability, and complementary thirteen-base oligonucleotides containing eleven FNA thymine residues failed to hybridize to a DNA template (Schneider and Benner 1990). Later work by Merle and colleagues showed weak interaction of FNA oligomers with DNA—the strongest pairing being observed for a homo-FNA-adenine oligomer with the homo-dT complement, although this pairing was still depressed relative to the analogous DNA-DNA duplex (Merle et al. 1995). Thus, in addition to being well-accommodated in a Watson-Crick duplex, the conformational preorganization of ribose nucleosides is clearly an important component of duplex stability.

The Eschenmoser laboratory has performed an extensive and systematic investigation of how changes to the natural β-furanosyl ribonucleosides of RNA, to nucleosides with a sugar of different stereochemical configuration, phosphate connectivity, carbon chain length or ring form (i.e., pyranose or furanose), affect the ability of nucleic acids to base pair (Eschenmoser 2007). A few examples are shown as Figure 1 D−F. Among their findings was the observation that oligomers incorporating the pyranosyl form of nucleosides with a pentose sugar (including ribose) can form more stable duplexes than their furanosyl forms. This observation is intriguing, particularly taken in light of the observation that pyranosyl nucleosides are apparently formed along with furanosyl nucleosides in heating-drying reactions (Bean et al. 2007), suggesting that both forms could have been available for proto-RNA. Eschenmoser has interpreted the intermediate helical stability of the contemporary furanose nucleosides as an indication that the contemporary

form of RNA prevailed due, in part, to an optimal, *intermediate* helical stability, not *maximal* helical stability. He suggests that the strong hybridization of pentopyranosyl backbones could have come at a price in a proto-RNA world—in the form of product inhibition and depressed sequence fidelity in replication.

The four carbon sugar threose is of particular interest as a potential ancestor of ribose (Schöning et al. 2000; Herdewijn 2001a; Wilds et al. 2002). The nucleosides of this threose can be connected by phosphate esters to create a backbone with a five-atom repeat (Fig. 2D), one atom shorter than the RNA backbone. Experiments with 3′,2′-linked threofuranose nucleic acid, or TNA, have generated significant excitement in the origin of life community. In addition to the potential relative ease of formation of threose in a prebiotic environment, duplexes of TNA oligonucleotides are of comparable stability to RNA and DNA duplexes, and TNA forms hybrid duplexes with RNA and DNA (Schöning et al. 2000; Wilds et al. 2002). It has often been speculated that cross-hybridization with RNA is an essential feature for an alternative nucleic acid to be considered a possible ancestor of RNA (Orgel 2000), as this ability would have been necessary for transfer of sequence information between the two polymers over the course of evolution.

A backbone containing even fewer atoms, but of the same polymer repeat length as TNA, has been termed GNA (Fig. 2C). The structure and stability of duplexes formed by oligonucleotides with this backbone have been examined in detail by Meggers and coworkers (Zhang and Meggers 2005; Schlegel et al. 2009). These investigators have shown that GNA oligonucleotides with Watson-Crick complementary sequences form antiparallel duplexes that are even more stable than DNA or RNA duplexes of the same nucleobase sequence (Zhang and Meggers 2005). In contrast, little or no hybridization is observed between complementary GNA and DNA oligonucleotides (relative to the DNA homoduplex) and base pairing with RNA is of roughly equal stability relative to homoduplex RNA. Meggers and coworkers have also provided solution-state evidence (by CD and

NMR) for the preorganization of GNA (on the single-strand and single-residue levels) into a helix-compatible conformer. These data illustrate that even an acyclic backbone can be conformationally preorganized. A subsequent X-ray crystal structure revealed that the GNA duplex makes extensive interstrand stacking contacts and minimal intrastrand contacts (in contrast to RNA and DNA, for which the converse is true).

The Szostak laboratory recently showed that a GNA region of a DNA-GNA chimeric strand could act as a template for *Bst* polymerase (Tsai et al. 2007). Despite little or no hybridization between GNA and DNA, full-length polymerization of the dodecamer GNA template was achieved. Although it is unlikely that polymerase-catalyzed replication was contemporaneous with the first self-replicating polymers (which these authors do not suggest), their results show that, even if cross-hybridization is disfavored, the presence of a duplex-stabilizing molecule (a protein or RNA polymerase, or a small molecule) can relax this requirement for information transfer. A parallel line of evidence for this principle comes from the Switzer laboratory, who have shown that 2′,5′-linked DNA and FNA can also act as substrates for natural DNA polymerases, even though 2′,5′-linked DNA and FNA hybridize weakly or not at all to natural DNA (Sinha et al. 2004; Heuberger and Switzer 2008b).

Overall, synthetic organic chemists have shown that numerous alternative sugars from among the tetroses, pentoses, and hexoses, and alternative conformations of ribose (e.g., β-pyranosyl) can produce backbones that form double-stranded structures with Watson-Crick base pairs, with some of these being even more stable than duplex RNA (Bolli et al. 1997; Eschenmoser 1999; Schöning, Scholz et al. 2000; Eschenmoser 2004; Egli et al. 2006). These results support the proposal that proto-RNA backbones may have been constructed from a sugar (or a mixture of sugars) other than ribose. The search through sugar space is by no means complete. As one example, the ketose sugars have not been explored as possible ancestors to ribose, but they should also be

considered as potentially simpler ancestors of ribose given the remarkably high yield with which the ketohexoses, such as fructose, are formed from DL-glyceraldehyde in a model prebiotic reaction (Weber 1992).

Finally, from a reactivity perspective, amino sugars are also an attractive possibility. As discussed later, the formation of phosphodiester bonds in aqueous solution is problematic. Recent work by Szostak and coworkers (Mansy et al. 2008; Chen et al. 2009; Schrum et al. 2009), as well as earlier work by Orgel and coworkers (Zielinski and Orgel 1985, 1987; Sievers and von Kiedrowski 1994), have shown the use of amino sugars in promoting polymerization because of the greater nucleophicity of amines compared with alcohols. The amino-substituted phosphoimidazole-activated nucleosides used in these studies were, obviously, prepared using classical synthetic organic techniques. Nevertheless, given the promising polymerization results that have been achieved with such nucleosides, as well as the presence of amino sugars in contemporary life (e.g., glucosamine), potential prebiotic reactions for the production of amino-substituted sugars are an intriguing area for future investigation.

Nucleoside Formation: How Did the Nucleobase and Trifunctional Moiety First Connect?

The prebiotic origin of the glycosidic bond (an N,O-acetal) that connects the nucleobase to ribose in a nucleoside has turned out to be one of the most vexing problems facing the origin of RNA. Given that the synthesis of adenine by HCN polymerization was one of the earliest successes of prebiotic chemistry (Oró 1960), and sugar formation from formaldehyde was known for a century prior (Butlerow 1861), many have speculated that prebiotic nucleoside formation resulted from the heating and drying of preexisting nucleobases and ribose. Orgel and coworkers first reported the formation of adenosine, inosine and guanosine when samples of free adenine, hypoxanthine and guanine, respectively, were dried and heated with ribose in the presence of magnesium salts (Fuller

et al. 1972a, b). However, yields were low, and the synthesis of guanosine was especially problematic, because of the low solubility of guanine. More troublesome is the observation that the canonical pyrimidine bases of RNA, uracil and cytosine, do not form nucleosides when dried and heated with ribose (Orgel 2004). It is worth noting that even the previously reported formation of adenosine in heating-drying reactions has been described as difficult to reproduce (Zubay and Mui 2001).

Several alternative hypotheses for prebiotic nucleoside formation have emerged as a result of the difficulty of glycosidic bond formation. These hypotheses include the possibility that the nucleobases were formed on a preexisting sugar, or the converse, that the sugar was built off of the nucleobase. The earliest exploration of these alternative routes was again reported by the Orgel laboratory (Sanchez and Orgel 1970). Starting with 5′-phosphorylated ribose, followed by multiple steps, including chemical reagent addition, UV anomerization, and hydrolysis, Sanchez and Orgel showed the synthesis of cytidine monophosphate. The Sutherland laboratory has recently revisited this hypothesis with great vigor and has made impressive strides toward the presentation of a comprehensive pathway for the stepwise abiotic formation of β-DL-cytidine-2′,3′-cyclic phosphate, using cyanamide, cyanoacetylene, glycolaldehyde, glyceraldehyde and inorganic phosphate as starting molecules (Ingar et al. 2003; Anastasi et al. 2006; Crowe and Sutherland 2006; Powner et al. 2009; Sutherland 2010).

The possibility that the nucleoside sugar was originally synthesized on the nucleobases, the converse of the previously described nucleosidation hypothesis, has also received some experimental support. Simply heating formamide in the presence of various mineral catalysts, as performed with a great variety of minerals by the Saladino and DiMauro laboratories (Saladino et al. 2007), have produced all four canonical RNA bases, as well as purine acyclonucleosides (albeit with rudimentary "sugar" moieties) (Saladino et al. 2003). Given the plausibly prebiotic nature of formamide as a starting material and solvent (Benner et al.

2004), the one-pot synthesis of a proto-nucleoside is a remarkable observation that certainly merits serious consideration as an alternative route to nucleosides.

The concept of nucleic acid evolution also opens the door to that possibility that proto-RNA contained nucleobases and/or a trifunctional moiety that more easily formed nucleosides. The merit of this hypothesis was first demonstrated by Miller and co-workers, who showed that urazole, a five-membered heterocycle with the same Watson-Crick hydrogen bonding pattern as uracil, reacts spontaneously with ribose to give α- and β- pyrano- and furanosides (Kolb et al. 1994; Dworkin and Miller 2000). More recently, Bean et al. investigated nucleoside formation with ribose and 2-pyrimidinone (Bean et al. 2007), a pyrimidine with the same 2-oxo functional group as uracil and cytosine, but with a proton instead of oxo- or amino-substitution at the 4 position. This base was shown to form nucleosides in heating-drying reactions, in which up to 12% of the remaining base was converted to the β-furanoside, possibly the highest conversion of a base to a nucleoside observed to date in a simple heating-drying reaction. Computational studies of this reaction have confirmed the importance of the in-plane lone pair of 2-pyrimidinone in nucleoside formation, which is also present in the imidazole ring of purines, and the importance of divalent cations in lowering the activation barrier to glycosidation (Bean, Sheng et al. 2007; Sheng et al. 2009). Together, the results obtained for urazole and 2-pyrimidinone illustrate that nucleoside formation from a preexisting base and sugar may not have been such a difficult step in proto-RNA formation, if the proto-nucleobases differed somewhat from those of contemporary RNA.

Selection of the Nucleoside Linker: When Did Nature Choose Phosphate?

The phosphate backbone is a chemical hallmark of contemporary nucleic acids. As a triacid, phosphate provides nucleic acids with a means to connect two alcohols, in the form of sugars, via ester linkages, while maintaining a negative charge (a key contributor to nucleic acid solubility in water). As Westheimer pointed out in his insightful paper entitled "Why Nature Chose Phosphates" (Westheimer 1987), it is difficult to conceive of a chemical arrangement that would provide better functional characteristics. The ester linkages provide connectivity that is sufficiently labile so that enzymes can hydrolyze the polymer (for proofreading or monomer salvage), whereas the negative charge provides sufficient screening of the phosphorus center from nucleophilic attack to stabilize the polymer against hydrolysis. Additionally, the intrastrand electrostatic repulsion afforded by the regularly-spaced negative charges may be crucial to an extended helical conformation and for information transfer that is rigorously dependent on Watson-Crick base pairing (contrast proteins, in which self-assembly via molecular recognition occurs, but not by the rule-based Watson-Crick system observed in nucleic acids). Westheimer noted that phosphate is unique here, as well: as a polyvalent acid with all acidic protons in close proximity, its negative charge will afford greater protection from nucleophilic attack and intraresidue repulsion than would a polyvalent acid with its ionizable moieties more distant from one another (e.g., citrate).

Benner and coworkers have experimentally explored the importance of backbone charge on the ability for RNA-like molecules to form duplexes with Watson-Crick base pairs. As a neutral structural analog of the phosphodiester-linkage, they studied dimethylenesulfone-linked nucleosides (in which the O3′ and O5′ sugar atoms are replaced by methyelene groups and phosphorus is replaced by sulfur) (Huang et al. 1991; Richert et al. 1996). The results from these studies were intriguing, but less than straightforward to interpret. Although short sulfone-linked DNA analogs (sNAs) were shown to support Watson-Crick base pairing (Roughton et al. 1995), longer oligosulfones appeared to have somewhat compromised pairing abilities (Huang et al. 1991; Richert et al. 1996) and small changes in nucleobase oligosulfone sequence resulted in appreciable changes in oligomer solubility, folding and

aggregation (Eschgfaller et al. 2003; Schmidt et al. 2003). These authors interpreted the results of these studies as evidence that (1) charged linkages are important for molecular recognition by providing a repulsive energetic term between the two backbones of a duplex, ensuring molecular recognition is largely a function of the nucleobases, (2) that the regular repeating charge limits intramolecular folding of oligomers, thereby allowing these polymers to function well as linear duplexes and as templates for the same during replication, and (3) that, given the dominance of charge in governing the physical properties of these nucleic acids in aqueous solution, changes in DNA/RNA nucleobase sequence have only a second-order effect on polymer properties (i.e., sequence exerts more effect on the molecular recognition of a complementary strand and less effect on helical parameters), thereby allowing the use of virtually any possible nucleotide sequence (Benner et al. 2004).

Studies from the fields of biological and medicinal chemistry have also provided some insight regarding the importance of backbone charge on nucleic acid base pairing. For example, methylphosphonates, in which one phosphate oxygen is replaced by a methyl group, provide a neutral linkage that is among the closest possible structural analogs of the natural backbone (Miller et al. 1981). Oligonucleotides with methylphosphonate substitutions support Watson-Crick base pairing in water, as homoduplexes and as hybrid duplexes with natural nucleic acids (Miller et al. 1981; Kiblerherzog et al. 1991; Schweitzer and Engels 1999). Surprisingly, removal of negative charges does not result in a substantial increase in duplex stability, as might be expected because of the reduced Coulombic repulsion between backbones compared with DNA and RNA. In contrast, two neutral polymers with more radical structural changes, "morpholinos" and PNAs, have been shown to form hybrid duplexes with RNA and DNA that are of greater stability than homo-DNA or homo-RNA duplexes (Summerton and Weller 1989; Egholm et al. 1993; Brown et al. 1994; Summerton and Weller 1997), and PNA will form homoduplexes with Watson-Crick

molecular recognition (Rasmussen et al. 1997; He et al. 2008). The hybridization of nucleic acids is the result of a complex interplay of numerous factors, including hydrogen bonds, nucleobase stacking interactions, backbone conformational landscapes, electrostatic interactions and hydration effects. Thus, the sensitivity of duplex stability to backbone charge is consistent with the predicted importance of the phosphate backbone in facilitating base pairing, but the observation of several successful neutral polymers suggest we presently lack sufficient information to understand the precise origin of the observed charge-dependent differential stabilities.

Although phosphate esters are *functionally* superior linkages, they show several *chemical* characteristics that may have made it difficult to incorporate them into the earliest biopolymers. The first is solubility. In the presence of divalent cations, free phosphate tends to form insoluble minerals; outside of living organisms, the phosphate resources on Earth even today are found as insoluble mineral deposits (Keefe and Miller 1995). Indeed, phosphate is a limiting reagent in much of contemporary life. Although recent hypotheses have been proposed for the availability of reduced forms of phosphorus on the earth which are more soluble (De Graaf and Schwartz 2000; Bryant and Kee 2006; Schwartz 2006; Pasek 2008), the low aqueous solubility of the contemporary (V) oxidation state of phosphorus, in the presence of divalent cations, and the positive ΔH associated with phosphoester formation present problems for prebiotic phosphate chemistry.

Another chemical characteristic of phosphate that would have made polymer formation challenging is the kinetic barrier to spontaneous phosphate ester formation. The negative charge of phosphate above pH ≈ 1 presents a barrier to ester formation, which, like hydrolysis, must proceed via nucleophilic attack on the phosphorus atom. It is possible to activate phosphate for ester formation (e.g., using methylimidazolides). However, such activated phosphates are necessarily high-energy compounds and, as would be expected, hydrolysis competes with polymerization.

A.E. Engelhart and N.V. Hud

Phosphate activation poses problems for both de novo production of oligomers from monomers and the incorporation of oligomers into higher polymers. In the case of monomer polymerization, Ferris and coworkers have shown that activated imidazolide derivatives of adenosine nucleotides will polymerize on mineral surfaces, but a large proportion of the monomer is incorporated into cyclic dimers (Miyakawa et al. 2006). In attempting to polymerize tiled half-complementary hexanucleotides into higher oligo- and polymers by carbodiimide activation, Kawamura and Okamoto also observed substantial amounts of cyclized starting material (Kawamura and Okamoto 2001). In both of these examples, a significant amount of starting material is irreversibly incorporated into an undesired side product. In addition to being a nuisance in the lab, these side products would have amounted to a fatal and committed step in the synthesis of a nascent proto-RNA. This problem illustrates a difficulty in nonenzymatic polymerization that must be taken into account when considering how the nature of the synthetic routes to and structural identities of early genetic polymers: irreversible linkages are adaptive for an informational polymer *only when mechanisms exist to make them conditionally reversible* (so as to allow proofreading).

The difficulties described earlier confound the efficient prebiotic chemical synthesis of phosphate esters—ligation of activated substrates yields mixed regioisomers and poor replication fidelity because of strand cyclization, base misincorporation, and premature product chain termination (Joyce and Orgel 1999; Kawamura and Okamoto 2001; Miyakawa al. 2006). Work by Usher and colleagues has provided encouraging evidence that the $2',5'$ linkage is more labile than the contemporary $3',5'$ regioisomer, suggesting that thermodynamic selection may be helpful in addressing a subset of the aforementioned problems with phosphodiester linkages (Usher and McHale 1976). However, all these difficulties can be circumvented by the use of a linker chemistry that initially forms a low-energy, reversible bond, allowing for selection of the thermodynamic product (i.e., in the case of RNA, the Watson-Crick base-paired product).

One remarkable example of the power of reversible linkages was provided by Lynn and coworkers, who showed that $5'$-deoxy,$5'$-amino, $3'$-deoxy, $3'$-formylmethyl-dT could polymerize via reductive amination in aqueous solution (Li et al. 2002). Polymerization occurred only in the presence of a $d(A_8)$ template; without a template, no oligomer product was detected. Interestingly, only linear polymers were observed. Here, polymerization decreases the entropic cost of hybridization, driving base pairing (which is disfavored at the mononucleotide level in water) (Yakovchuk et al. 2006). In turn, hybridization increases the local concentration of amines and aldehydes, driving the formation of imine linkages (which are otherwise disfavored in 55 M water). Since the irreversible reduction step requires the imine linkage to be present, and imine formation was template dependent, only linear polymers were formed. In this system, the subtle interplay between base pairing and linkage formation drives the simultaneous formation of reversible covalent and noncovalent interactions which are otherwise disfavored. Several other intriguing examples of self-assembling reversible polymers have been shown, particularly by the Lehn laboratory (Sreenivasachary and Lehn 2005; Sreenivasachary et al. 2006; Sreenivasachary and Lehn 2008).

Acetal linkages have the potential to provide reversible bonds, thereby providing the benefits showed by the reversible imine chemistry of Lynn. For example, the $3'$, $5'$-formacetal linkage (i.e., phosphate replaced by $-O-CH_2-O-$), has been examined by a number of laboratories (albeit not in fully modified oligonucleotides). Phosphorus (V) esters and acetals are both tetrahedral and approximately isosteric, and it is known that aNAs can base pair with RNAs (Matteucci and Bischofberger 1991). Specifically, point substitutions of this linkage impart a modest destabilization to DNA-DNA and DNA–RNA duplexes and show salt-dependent effects on RNA-RNA duplexes (stabilization at low salt, destabilization at high salt) (Jones et al. 1993; Rice and Gao 1997; Rozners et al. 2007; Kolarovic et al. 2009).

Advanced Online Article. Cite this article as *Cold Spring Harb Perspect Biol* doi: 10.1101/cshperspect.a002196

Investigators have invoked compatibility with the B-versus A-form helices, electrostatic, and hydration effects for these observed differential effects. Further, some evidence has been presented that the destabilization associated with this substitution in DNA–DNA duplexes is sequence-dependent (Pitulescu et al. 2008). Finally, the 3′-thioformacetal linkage has been shown to slightly destabilize DNA-DNA duplexes and slightly stabilize DNA-RNA duplexes (Jones et al. 1993). The latter stabilization was explained by a combination of sugar pucker, bond lengths, and accessible torsion angles resulting from the 3′-thio substitution. This sugar pucker effect was later confirmed by NMR spectroscopy (Rice and Gao 1997).

With the proper choice of aldehyde or ketone, these linkages can provide a negative charge that has proven critical to the phosphodiester backbone. For these reasons, acetal-linked nucleic acids (aNAs) are an intriguing potential ancestral polymer (Hud and Anet 2000). Acetal bond formation is favored over hydrolysis on water removal (Wiberg et al. 1994). The concentration of substrates that occurs during dry-down, driving bond formation, would also favor the hybridization of short, complementary sequences, providing an accessible means by which to drive polymerization.

The functionalization of an acetal with a negatively charged group, such as carboxylate group in the case of the aldehyde glyoxylate, would provide a localized charge near the backbone acetal linkage, thereby providing many of the favorable electrostatic properties afforded by phosphate. Glyoxylate has been shown to form in a prebiotically plausible reaction from glycolaldehyde (Weber 2001). We have speculated that glyoxylate could have preceded phosphate in an early proto-RNA, in the form of glyoxylate-acetal nucleic acids (gaNAs, Fig. 2G) (Bean et al. 2006). Based on energy minimized molecular models, it appears that a gaNA duplex would show helical properties very similar to RNA, suggesting that information transfer could occur between the two polymers. The formation of gaNA dinucleotides has been shown on heating and drying of a neutral solution containing nucleosides, glyoxylate and mono- or divalent cations (Bean et al. 2006). These acetal linkages formed by glyoxylate and nucleosides are surprisingly stable to hydrolysis—no detectable decomposition occurs after two weeks at room temperature and neutral pH. Importantly, gaNAs can be hydrolyzed at moderately elevated temperature in the presence of salt (i.e., 60 $^\circ$C, 1 M $MgCl_2$). An accessible means to modulate polymer stability (in this case, temperature and salt concentration), which provides an attractive means by which to "turn on and off" thermodynamic control of polymerization—a crucial feature in the chemical synthesis of an early informational polymer.

It is possible that other, similar, molecules could form suitable linkages. For instance, pyruvate, like glyoxylate, provides a negative charge and tetrahedral geometry, and ketal bond formation is still expected to be more facile than phosphodiester bond formation. Further, pyruvate is ubiquitous in life and formed in excellent yield in model prebiotic reactions (Weber 2001).

It is important to note that additional promising models have presented for the prebiotic formation of oligonucleotides that begin with either a reduced form of phosphorus or mononucleotides activated as cyclic monophosphates. Reduced phosphorus species, such as phosphite, in addition to showing the aforementioned enhanced solubility relative to phosphates, enjoy enhanced reactivity. Schwartz and coworkers reported high yields of uridine 5′-H-phosphonate (as high as 44% in the presence of urea) in 60 $^\circ$C drying reactions of aqueous ammonium phosphite, uridine and, optionally, urea. Additionally, when dried at 110 $^\circ$C, these authors reported production of small amounts of putative higher-order products, the most abundant of which was produced in 4% yield, which they tentatively assigned as the 5′,5′ H-phosphonate-linked dinucleotide. This species could be oxidized with iodine to a second species, assigned as the phosphate-linked dinucleotide (De Graaf and Schwartz 2005). Like acetals, H-phosphonate diesters are a conditionally reversible, low-energy linkage.

Although they are uncharged and hydrolytically labile, upon oxidation, they become negatively charged, comparatively hydrolytically inert phosphate diesters, making phosphonates another attractive, conditionally reversible backbone linkage. Finally, DiMauro and coworkers have recently reported that 3′, 5′-cyclic AMP and GMP will spontaneously form homo-A and homo-G oligonucleotides, respectively, in aqueous solution (Costanzo et al. 2009). These authors propose that nucleotides activated as cyclic monophosphates are potentially prebiotic, based on their earlier demonstration of 3′, 5′-cyclic AMP formation in formamide solutions containing phosphate minerals (Costanzo et al. 2007).

As Westheimer, Benner, and others have elegantly discussed, phosphate affords RNA numerous positive characteristics (Westheimer 1987; Benner et al. 2004). When considering some of the adaptive traits of RNA as a polymer—hydrolytic resistance, water solubility, extended conformation, and heritability of an arbitrary sequence because of rule-based molecular recognition—it is tempting to speculate that *phosphate might be the optimal linker for all possible RNA-like polymers*. The roles of phosphate in contemporary life, however, are possible only because of a highly optimized suite of catalysts, which enable the transfer of energy between disparate substrates and the "phosphate economy" (i.e., via the agency of ATP and similar compounds). Absent such catalysis, the chemical synthesis (and subsequent proofreading) of phosphate esters and anhydrides is problematic. *Many of the same attributes of phosphate that are adaptive in contemporary life could have been maladaptive in the proto-biopolymer world.* We therefore suggest that other chemistries, including those discussed here, be considered as possible linkers for the nucleosides of proto-RNA. These linkers could have incorporated some of the favorable traits of phosphate (e.g., electrostatic benefits, significant hydrolytic stability) while addressing some of the problematic ones (e.g., disfavored bond formation, lack of a proofreading mechanism in chemical synthesis) (Westheimer 1987; Bean, Anet et al. 2006).

CHALLENGES AND FUTURE DIRECTIONS

Our primary objective in this article has been to present some of the known challenges for the prebiotic synthesis of RNA, as well as potential solutions to these challenges that become available if we accept the possibility that one or more proto-RNAs predated RNA. More than a mere academic exercise, the discovery (or "reinvention") of a self-assembling RNA-like polymer would have tremendous potential for applications in biotechnology, medicine, and materials science.

Although our current examples of self-assembling polymers with reversible linkages are limited, the properties of these polymers are impressive and indicative of the promise held by the development of this area of polymer science. For example, dynamic covalent assemblies have been examined by the Lehn laboratory in the form of guanosine hydrogels, and they have shown that these self-assembling materials can act as controlled release media for bioactive small molecules (Sreenivasachary and Lehn 2005, 2008). Similarly, these investigators, as well as Leibler and coworkers, have also shown that polymers with the capacity for constitutional reorganization show a number of intriguing material properties, including self-healing and chemoresponsive characteristics (Sreenivasachary et al. 2006; Cordier et al. 2008).

Most recently, Ghadiri and coworkers have showed the ability for peptide nucleic acid polymer with reversible linkages between a polycysteine backbone and thioester acyclonucleosides to undergo a form of evolution in response to the presence of different DNA templates (Ura et al. 2009). Almost simultaneously, the Liu group reported a similar system, involving amidation and reductive amination of a PNA backbone with acyclonucleosides containing carboxylate and aldehyde functional groups (Heemstra and Liu 2009). Acyclonucleosides formed in formamide mixtures recently reported by Di Mauro and coworkers share at least some structural similarity with those used by the Ghadiri and Liu laboratories, suggesting that prebiotic routes may exist to nucleosides that multiple investigators have

now shown could couple to varied preformed backbones. One can now imagine the development of RNA-like polymers following a similar theme, creating functional polymers that are environmentally responsive, perhaps even evolving new functions in response to unanticipated environmental signals, which is completely feasible for polymers with reversible linkages that can assume the thermodynamically most favored state, as defined by their local environment.

Finally, the prospect of proto-RNA is reason for optimism that we will yet discover RNA-like polymers with self-assembling properties that are superior to any currently known. The chemical space that still must be searched is vast, but it is not infinite, and the studies discussed earlier have certainly provided clues to where we will most likely find such polymers, and perhaps even the origin of RNA.

ACKNOWLEDGMENTS

REFERENCES

Anastasi C, Crowe MA, Powner MW, Sutherland JD. 2006. Direct assembly of nucleoside precursors from two- and three-carbon units. *Angew Chem Int Ed Engl* **45:** 6176–6179.

Anet FAL. 2004. The place of metabolism in the origin of life. *Curr Opin Chem Biol* **8:** 654–659.

Ban N, Nissen P, Hansen J, Moore PB, Steitz TA. 2000. The complete atomic structure of the large ribosomal subunit at 2.4 Ångstrom resolution. *Science* **289:** 905–920.

Battersby TR, Albalos M, Friesenhahn MJ. 2007. An unusual mode of DNA duplex association: Watson-Crick interaction of all-purine deoxyribonucleic acids. *Chem Biol* **14:** 525–531.

Bean HD, Anet FAL, Gould IR, Hud NV. 2006. Glyoxylate as a backbone linkage for a prebiotic ancestor of RNA. *Orig Life Evol B* **36:** 39–63.

Bean HD, Sheng YH, Collins JP, Anet FAL, Leszczynski J, Hud NV. 2007. Formation of a β-pyrimidine nucleoside by a free pyrimidine base and ribose in a plausible prebiotic reaction. *J Am Chem Soc* **129:** 9556–9557.

Benner SA. 2004. Understanding nucleic acids using synthetic chemistry. *Acc Chem Res* **37:** 784–797.

Benner SA, Ricardo A, Carrigan MA. 2004. Is there a common chemical model for life in the universe? *Curr Opin Chem Biol* **8:** 672–689.

Bolli M, Micura R, Eschenmoser A. 1997. Pyranosyl-RNA: chiroselective self-assembly of base sequences by ligative oligomerization of tetra nucleotide-2′,3′-cyclophosphates (with a commentary concerning the origin of biomolecular homochirality). *Chem Biol* **4:** 309–320.

Brown S, Thomson S, Veal J, Davis D. 1994. NMR solution structure of a peptide nucleic acid complexed with RNA *Science* **265:** 777–780.

Bryant DE, Kee TP. 2006. Direct evidence for the availability of reactive, water soluble phosphorus on the early Earth. H-Phosphinic acid from the Nantan meteorite. *Chem Commun*: 2344–2346.

Butlerow A. 1861. Bildung einer zuckerartigen substanz durch synthese. *Liebigs Ann Chem* **120:** 295–298.

Cairns-Smith AG. 1982. *Genetic takeover and the mineral origins of life*. Cambridge University Press, Cambridge.

Chen JJ, Cai X, Szostak JW. 2009. N2′→P3′ phosphoramidate glycerol nucleic acid as a potential alternative genetic system. *J Am Chem Soc* **131:** 2119–2121.

Cleaves HJ, Nelson KE, Miller SL. 2006. The prebiotic synthesis of pyrimidines in frozen solution. *Naturwissenschaften* **93:** 228–231.

Cordier P, Tournilhac F, Soulié-Ziakovic C, Leibler L. 2008. Self-healing and thermoreversible rubber from supramolecular assembly. *Nature* **451:** 977–980.

Costanzo G, Saladino R, Crestini C, Ciciriello F, Di Mauro E. 2007. Nucleoside phosphorylation by phosphate minerals. *J Biol Chem* **282:** 16729–16735.

Crick FHC. 1968. The origin of the genetic code. *J Mol Biol* **38:** 367–379.

Crowe MA, Sutherland JD. 2006. Reaction of cytidine nucleotides with cyanoacetylene: support for the intermediacy of nucleoside-2′, 3′-cyclic phosphates in the prebiotic synthesis of RNA. *ChemBioChem* **7:** 951–956.

De Graaf RM, Schwartz AW. 2000. Reduction and activation of phosphate on the primitive earth. *Orig Life Evol B* **30:** 405–410.

De Graaf RM, Schwartz AW. 2005. Thermal synthesis of nucleoside H-phosphonates under mild conditions. *Orig Life Evol B* **35:** 1–10.

De Voe H, Tinoco I. 1962. The stability of helical polynucleotides: base contributions. *J Mol Biol* **4:** 500–517.

Decker P, Schweer P, Pohlmann R. 1982. Identification of formose sugars, presumable prebiotic metabolites, using capillary gas chromatography/gas chromatography-mass spectroscopy of n-butoxime trifluoroacetates on OV-225. *J Chromatogr* **244:** 281–291.

Dworkin JP, Miller SL. 2000. A kinetic estimate of the free aldehyde content of aldoses. *Carbohydr Res* **39:** 359–365.

Egholm M, Buchardt O, Christensen L, Behrens C, Freier S, Driver D, Berg R, Kim S, Norden B, Nielsen P. 1993. PNA hybridizes to complementary oligonucleotides obeying the Watson–Crick hydrogen-bonding rules. *Nature* **365:** 566–568.

Egli M, Pallan PS, Pattanayek R, Wilds CJ, Lubini P, Minasov G, Dobler M, Leumann CJ, Eschenmoser A. 2006. Crystal structure of homo-DNA and Nature's choice of pentose

over hexose in the genetic system. *J Am Chem Soc* **128:** 10847–10856.

Engelhart AE, Morton TH, Hud NV. 2009. Evidence of strong hydrogen bonding by 8-amino-guanine. *Chem Commun*: 647–649.

Eschenmoser A. 1999. Chemical etiology of nucleic acid structure. *Science* **284:** 2118–2124.

Eschenmoser A. 2004. The TNA-family of nucleic acid systems: Properties and prospects. *Orig Life Evol B* **34:** 277–306.

Eschenmoser A. 2007. The search for the chemistry of life's origin. *Tetrahedron* **63:** 12821–12844.

Eschgfaller B, Schmidt JG, Konig M, Benner SA. 2003. Synthesis and properties of oligodeoxynucleotide analogs with bis(methylene) sulfone bridges. *Helv Chim Acta* **86:** 2959–2997.

Fuller WD, Sanchez RA, Orgel LE. 1972a. Studies in prebiotic synthesis. VI. Synthesis of purine nucleosides. *J Mol Biol* **67:** 25–33.

Fuller WD, Sanchez RA, Orgel LE. 1972b. Studies in prebiotic synthesis: VII. Solid-state synthesis of purine nucleosides. *J Mol Evol* **1:** 249–257.

Gesteland R, Atkins JF. 2006. *The RNA world: The nature of modern RNA suggests a prebiotic RNA world* (3rd ed.) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Geyer CR, Battersby TR, Benner SA. 2003. Nucleobase pairing in Watson-Crick-like genetic expanded information systems. *Structure* **11:** 1485–1498.

Groebke K, Hunziker J, Fraser W, Peng L, Diederichsen U, Zimmermann K, Holzner A, Leumann C, Eschenmoser A. 1998. Why pentose- and not hexose-nucleic acids? Purine-purine pairing in homo-DNA: guanine, isoguanine, 2,6-diaminopurine, and xanthine. *Helv Chim Acta* **81:** 375–474.

Guerrier-Takada C, Gardiner K, Marsh T, Pace N, Altman S. 1983. The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell* **35:** 849–857.

He W, Hatcher E, Balaeff A, Beratan D, Gil R, Madrid M, Achim C. 2008. Solution structure of a peptide nucleic acid duplex from NMR data: Features and limitations. *J Am Chem Soc* **130:** 13264–13273.

Heemstra JM, Liu DR. 2009. Templated synthesis of peptide nucleic acids via sequence-selective base-filling reactions. *J Am Chem Soc* **131:** 11347–11349.

Herdewijn P. 2001a. TNA as a potential alternative to natural nucleic acids. *Angew Chem Int Ed Engl* **40:** 2249–2251.

Herdewijn P. 2001b. TNA as a potential alternative to natural nucleic acids. *Angew Chem Int Ed Eng* **40:** 2249–2251.

Heuberger BD, Switzer C. 2008a. An alternative nucleobase code: Characterization of purine-purine DNA double helices bearing guanine-isoguanine and diaminopurine-7-deaza-xanthine base pairs. *ChemBioChem* **9:** 2779–2783.

Heuberger BD, Switzer C. 2008b. A nonRNA candidate revisited: Both enantiomers of flexible nucleoside triphosphates are DNA polymerase substrates. *J Am Chem Soc* **130:** 412–413.

Hollis JM, Jewell PR, Lovas FJ, Remijan A. 2004. Green bank telescope obserations of interstellar glycolaldehyde: Low-temperature sugar. *Astrophys J* **613:** L45–L48.

Horowitz ED, Engelhart AE, Chen MC, Quarles KA, Smith MW, Lynn DG, Hud NV. 2010. Intercalation as a means to suppress cyclization and promote polymerization of base-pairing oligonucleotides in a prebiotic world. *Proc Natl Acad Sci* **107:** 5288–5293.

Howard FB, Miles HT. 1977. Interaction of poly(A) and poly(I), a reinvestigation. *Biochemistry* **16:** 4647–4650.

Huang Z, Schneider KC, Benner SA. 1991. Building-blocks for oligonucleotide analogs with dimethylene sulfide, sulfoxide, and sulfone groups replacing phosphodiester linkages. *J Org Chem* **56:** 3869–3882.

Hud NV, Anet FAL. 2000. Intercalation-mediated synthesis and replication: A new approach to the origin of life. *J Theor Biol* **205:** 543–562.

Hud NV, Jain SS, Li X, Lynn DG. 2007. Addressing the problems of base pairing and strand cyclization in template-directed synthesis – A case for the utility and necessity of 'molecular midwives' and reversible backbone linkages for the origin of proto-RNA. *Chem Biodiver* **4:** 768–783.

Ihmels H, Otto D. 2005. Intercalation of organic dye molecules into double-stranded DNA: General principles and recent developments. *Top Curr Chem* **258:** 161–204.

Ingar A-A, Luke RWA, Hayter BR, Sutherland JD. 2003. Synthesis of cytidine ribonucleotides by stepwise assembly of the heterocycle on a sugar phosphate. *ChemBioChem* **4:** 504–507.

Inoue T, Orgel LE. 1983. A non-enzymatic RNA polymerase model. *Science* **219:** 859–862.

Jain SS, Anet FAL, Stahle CJ, Hud NV. 2004. Enzymatic behavior by intercalating molecules in a template-directed ligation reaction. *Angew Chem Int Ed Engl* **43:** 2004–2008.

Jaworski A, Kwiatkowski JS, Lesyng B. 1985. Why isoguanine and isocytosine are not the components of the genetic code. *Int J Quantum Chem* **28, S12:** 209–216.

Jones RJ, Lin K-Y, Milligan JF, Wadwani S, Matteucci MD. 1993. Synthesis and binding properties of pyrimidine oligonucleotide analogs containing neutral phosphodiester replacements: The formacetal and 3′-thioformacetal internucleoside linkages. *J Org Chem* **58:** 2983–2991.

Joyce GF. 1987. Nonenzymatic template-directed synthesis of informational macromolecules. *Cold Spring Harbor Symp Quant Biol* **52:** 41–51.

Joyce GF, Orgel LE. 1999. Prospects for understanding the origin of the RNA world. In The RNA World, Second Edition: The Nature of Modern RNA Suggests a Prebiotic RNA World (Eds. J.F. Atkins, R.F. Gesteland), 49–77. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Joyce GF, Schwartz AW, Miller SL, Orgel LE. 1987. The case for an ancestral genetic system involving simple analogs of the nucleotides. *Proc Natl Acad Sci* **84:** 4398–4402.

Kawamura K, Okamoto F. 2001. Cyclization and dimerization of hexanucleotides containing guanine and cytosine with water-soluble carbodiimide. *Viva Origino* **29:** 162–167.

Keefe AD, Miller SL. 1995. Are polyphosphates or phosphate esters prebiotic reagents? *J Mol Evol* **41:** 693–702.

Kiblerherzog L, Zon G, Uznanski B, Whittier G, Wilson WD. 1991. Duplex stabilities of phosphorothioate,

methylphosphonate, and RNA analogs of 2 DNA 14-mers. *Nucleic Acids Res* **19:** 2979–2986.

Kolarovic A, Schweizer E, Greene E, Gironda M, Pallan P, Egli M, Rozners E. 2009. Interplay of structure, hydration and thermal stability in formacetal modified oligonucleotides: RNA may tolerate nonionic modifications better than DNA. *J Am Chem Soc* **131:** 14932–14937.

Kolb VM, Dworkin JP, Miller SL. 1994. Alternative bases in the RNA world: the prebiotic synthesis of urazole and its ribosides. *J Mol Evol* **38:** 549–557.

Krishnamurthy R, Pitsch S, Minton M, Miculka C, Windhab N, Eschenmoser A. 1996. Pyranosyl-RNA: base pairing between homochiral oligonucleotide strands of opposite sense of chirality. *Angew Chem Int Ed Engl* **35:** 1537–1541.

Krueger AT, Lu HG, Lee AHF, Kool ET. 2007. Synthesis and properties of size-expanded DNAs: Toward designed, functional genetic systems. *Acc Chem Res* **40:** 141–150.

Kruger K, Grabowski PJ, Zaug AJ, Sands J, Gottschling DE, Cech TR. 1982. Self-splicing RNA: Autoexcision and autocyclization of the ribosomal RNA intervening sequence of *Tetrahymena*. *Cell* **31:** 147–157.

Kurreck J. 2003. Antisense technologies—Improvement through novel chemical modifications. *Eur J Biochem* **270:** 1628–1644.

Li X, Zhan Z-YJ, Knipe R, Lynn DG. 2002. DNA-catalyzed polymerization. *J Am Chem Soc* **124:** 746–747.

Mandal M, Breaker RR. 2004. Gene regulation by riboswitches. *Nat Rev Mol Cell Biol* **5:** 451–463.

Mansy SS, Schrum JP, Krishnamurthy M, Tobe S, Treco DA, Szostak JW. 2008. Template-directed synthesis of a genetic polymer in a model protocell. *Nature* **454:** 122–125.

Martins Z, Botta O, Fogel ML, Sephton MA, Glavin DP, Watson JS, Dworkin JP, Schwartz AW, Ehrenfreund P. 2008. Extraterrestrial nucleobases in the Murchison meteorite. *Earth Planet Sci Lett* **270:** 130–136.

Matteucci MD, Bischofberger N. 1991. Sequence-defined oligonucleotides as potential therapeutics. *Annu Rep Med Chem* **26:** 287–296.

Merle Y, Bonneil E, Merle L, Sági J, Szemzö A. 1995. Acyclic oligonucleotide analogues. *Int J Biol Macromol* **17:** 239–246.

Miller PS, McParland KB, Jayaraman K, Tso POP. 1981. Biochemical and biological effects of nonionic nucleic acid methylphosphonates. *Biochem Cell Biol* **20:** 1874–1880.

Mittapalli GK, Osornio YM, Guerrero MA, Reddy KR, Krishnamurthy R, Eschenmoser A. 2007a. Mapping the landscape of potentially primordial informational oligomers: oligodipeptides tagged with 2,4-disubstituted 5-aminopyrimidines as recognition elements. *Angew Chem Int Ed Engl* **46:** 2478–2484.

Mittapalli GK, Reddy KR, Xiong H, Munoz O, Han B, De Riccardis F, Krishnamurthy R, Eschenmoser A. 2007b. Mapping the landscape of potentially primordial informational oligomers: Oligodipeptides and oligodipeptoids tagged with triazines as recognition elements. *Angew Chem Int Ed Engl* **46:** 2470–2477.

Miyakawa S, Joshi PC, Gaffey MJ, Gonzalez-Toril E, Hyland C, Ross T, Rybij K, Ferris JP. 2006. Studies in the mineral and salt-catalyzed formation of RNA oligomers. *Origins Life Evol B* **36:** 343–361.

Miyakawa S, Yamanashi H, Kobayashi K, Cleaves HJ, Miller SL. 2002. Prebiotic synthesis from CO atmospheres: Implications for the origins of life. *Proc Natl Acad Sci* **99:** 14628–14631.

Müller D, Pitsch S, Kittaka A, Wagner E, Wintner CE, Eschenmoser A. 1990. Chemie von α-Aminonitrilen. Aldomerisierung von Glycoladehyd-phosphat zu racemischen Hexose-2,4,6-triphosphaten und (in Gegenwart von Formaldehyd) reacemischen Pentose-2,4-diphosphaten: *rac*-Allose-2,4,6-triphosphat und *rac*-Ribose-2,4-diphosphat sind die Reaktionshauptprodukte. *Helv Chim Acta* **73:** 1410–1468.

Nelson KE, Levy M, Miller SL. 2000. Peptide nucleic acids rather than RNA may have been the first genetic molecule. *Proc Natl Acad Sci* **97:** 3868–3871.

Orgel L. 2000. A simpler nucleic acid. *Science* **290:** 1306–1307.

Orgel LE. 1968. Evolution of the genetic apparatus. *J Mol Biol* **38:** 381–393.

Orgel LE. 1998. The origin of life – a review of facts and speculations. *TIBS* **23:** 491–495.

Orgel LE. 2004. Prebiotic chemistry and the origin of the RNA world. *Crit Rev Biochem Mol Biol* **39:** 99–123.

Orgel LE. 2008. The implausibility of metabolic cycles on the prebiotic Earth. *Plos Biology* **6:** 5–13.

Oro J. 1961. Mechanism of synthesis of adenine from hydrogen cyanide under possible primitive Earth conditions. *Nature* **191:** 1193–&.

Oró J. 1960. Synthesis of adenine from ammonium cyanide. *Biochem Biophys Res Comm* **2:** 407–412.

Pasek MA. 2008. Rethinking early Earth phosphorus geochemistry. *Proc Natl Acad Sci* **105:** 853–858.

Piccirilli JA, Krauch T, Moroney SE, Benner SA. 1990. Enzymatic incorporation of a new base pair into DNA and RNA extends the genetic alphabet. *Nature* **343:** 33–37.

Pitsch S, Wendeborn S, Jaun B, Eschenmoser A. 1993. Why pentose- and not hexose-nucleic acids? Part VII. Pyranosyl-RNA('p-RNA'). *Helv Chim Acta* **76:** 2161–2183.

Pitulescu M, Grapp M, Krätzner R, Knepel W, Diederichsen U. 2008. Synthesis of formacetal-linked dinucleotides to facilitate dsDNA bending and binding to the homeodomain of PAX6. *Eur J Org Chem*: 2100–2106.

Pizzarello S. 2006. The chemistry of life's origin: A carbonaceous meteorite perspective. *Acc Chem Res* **39:** 231–237.

Plankensteiner K, Reiner H, Schranz B, Rode BM. 2004. Prebiotic formation of amino acids in a neutral atmosphere by electric discharge. *Angew Chem Int Ed Engl* **43:** 1886–1888.

Powner MW, Gerland B, Sutherland JD. 2009. Synthesis of activated pyrimidine ribonucleotides in prebiotically plausible conditions. *Nature* **459:** 239–242.

Rajamani S, Vlassov A, Benner S, Coombs A, Olasagasti F, Deamer D. 2008. Lipid-assisted synthesis of RNA-like polymers from mononucleotides. *Orig Life Evol B* **38:** 57–74.

A.E. Engelhart and N.V. Hud

Rasmussen H, Kastrup J, Nielsen J, Nielsen J, Nielsen P. 1997. Crystal structure of a peptide nucleic acid (PNA) duplex at 1.7 Å resolution. *Nat Struct Biol* **4:** 98–101.

Ricardo A, Carrigan MA, Olcott AN, Benner SA. 2004. Borate minerals stabilize ribose. *Science* **303:** 196–196.

Rice JS, Gao X. 1997. Conformation of formacetal and 3′-thioformacetal nucleotide linkers and stability of their antisense RNA.DNA hybrid duplexes. *Biochemistry* **36:** 399–411.

Richert C, Roughton AL, Benner SA. 1996. Nonionic analogs of RNA with dimethylene sulfone bridges. *J Am Chem Soc* **118:** 4518–4531.

Roughton AL, Portmann S, Benner SA, Egli M. 1995. Crystal-structure of a dimethylene sulfone-linked ribodinucleotide analog. *J Am Chem Soc* **117:** 7249–7250.

Rozners E, Katkevica D, Strömberg R. 2007. Oligoribonucleotide analogues containing a mixed backbone of phosphodiester and formacetal internucleoside linkages, together with vicinal 2′-O-methyl groups. *Chembiochem* **8:** 537–545.

Saladino R, Ciambecchini U, Crestini C, Costanzo G, Negri R, Di Mauro E. 2003. One-pot TiO$_2$-catalyzed synthesis of nucleic bases and acyclonucleosides from formamide: Implications for the origin of life. *ChemBioChem* **4:** 514–521.

Saladino R, Crestini C, Cicirielloc F, Costanzo G, Di Mauro E. 2007. Formamide chemistry and the origin of informational polymers. *Chem Biodiv* **4:** 694–720.

Saladino R, Crestini C, Costanzo G, DiMauro E. 2004. Advances in the prebiotic synthesis of nucleic acids bases: implications for the origin of life. *Curr Org Chem* **8:** 1425–1443.

Sanchez R, Ferris J, Orgel LE. 1966. Conditions for purine synthesis—Did prebiotic synthesis occur at low temperatures? *Science* **153:** 72–73.

Sanchez RA, Orgel LE. 1970. Studies in prebiotic synthesis. V. Synthesis and photoanomerization of pyrimidine nucleosides. *J Mol Biol* **47:** 531–543.

Schlegel MK, Xie XL, Zhang LL, Meggers E. 2009. Insight into the high duplex stability of the simplified nucleic acid GNA. *Angew Chem Int Ed Engl* **48:** 960–963.

Schmidt J, Eschgfaller B, Benner SA. 2003. A direct synthesis of nucleoside analogs homologated at the 3′- and 5′-positions. *Helv Chim Acta* **86:** 2937–2958.

Schneider KC, Benner SA. 1990. Oligonucleotides containing flexible nucleoside analogs. *J Am Chem Soc* **112:** 453–455.

Schöning KU, Scholz P, Guntha S, Wu X, Krishnamurthy R, Eschenmoser A. 2000. Chemical etiology of nucleic acid structure: The α-threofuranosyl-(3′→2′) oligonucleotide system. *Science* **290:** 1347–1351.

Schrum JP, Ricardo A, Krishnamurthy M, Blain JC, Szostak JW. 2009. Efficient and rapid template-directed nucleic acid copying using 2′-amino-2′,3′-dideoxyribonucleoside–5′-phosphorimidazolide monomers. *J Am Chem Soc* **131:** 14560–14570.

Schwartz AW. 2006. Phosphorus in prebiotic chemistry. *Philos Trans R Soc London B* **361:** 1743–1749.

Schwartz AW, Chang S. 2002. From Big Bang to primordial planet: Setting the stage for the origin of life. In *Life's Origin: The Beginnings of Biological Evolution* (Eds.) W. Schopf, 46–77. University of California, Berkeley.

Schweitzer M, Engels JW. 1999. Sequence specific hybridization properties of methylphosphonate oligodeoxynucleotides. *J Biomol Struct Dyn* **16:** 1177–1188.

Seela F, Wei C, Kazmierczuk Z. 1995. Substituent reactivity and tautomerism of isoguanosine and related nucleosides. *Helv Chim Acta* **78:** 1843–1854.

Sepiol J, Kazmierczuk Z, Shugar D. 1976. Tautomerism of iso-guanosine and solvent-induced keto-enol equilibrium. Zeitschrift für Naturforschung C-A. *J Biosci* **31:** 361–370.

Sheng Y, Bean HD, Mamajanova I, Hud NV, Leszczynski J. 2009. A comprehensive investigation of the energetics of pyrimidine nucleoside formation in a model prebiotic reaction. *J Am Chem Soc*: in press.

Sievers D, von Kiedrowski G. 1994. Self replication of complementary nucleotide-based oligomers. *Nature* **369:** 221–224.

Sinha S, Kim PH, Switzer C. 2004. 2′,5′-linked DNA is a template for polymerase-directed DNA synthesis. *J Am Chem Soc* **126:** 40–41.

Springsteen G, Joyce GF. 2004. Selective derivatization and sequestration of ribose from a prebiotic mix. *J Am Chem Soc* **126:** 9578–9583.

Sreenivasachary N, Hickman DT, Sarazin D, Lehn JM. 2006. DyNAs: Constitutional dynamic nucleic acid analogues. *Chem-Eur J* **12:** 8581–8588.

Sreenivasachary N, Lehn JM. 2005. Gelation-driven component selection in the generation of constitutional dynamic hydrogels based on guanine-quartet formation. *Proc Natl Acad Sci* **102:** 5938–5943.

Sreenivasachary N, Lehn JM. 2008. Structural selection in G-quartet-based hydrogels and controlled release of bioactive molecules. *Chem-Asian J* **3:** 134–139.

Summerton J, Weller D. 1989. Uncharged morpholino-based polymers having achiral intersubunit linkages. *US Patent #5,034,506*

Summerton J, Weller D. 1997. Morpholino antisense oligomers: design, preparation, and properties. *Antisense Nucleic Acid Drug Dev* **7:** 187–195.

Sutherland JD. 2010. Ribonucleotides. *Cold Spring Harb Perspect Biol* **2:** a005439.

Switzer C, Moroney SE, Benner SA. 1989. Enzymatic incorporation of a new base pair into DNA and RNA. *J Am Chem Soc* **111:** 8322–8323.

Ts'o P, Melvin I, Olson A. 1963. Interaction and association of bases and nucleosides in aqueous solutions. *J Am Chem Soc* **85:** 1289–1296.

Tsai CH, Chen JY, Szostak JW. 2007. Enzymatic synthesis of DNA on glycerol nucleic acid templates without stable duplex formation between product and template. *Proc Natl Acad Sci* **104:** 14598–14603.

Ura Y, Beierle JM, Leman LJ, Orgel LE, Ghadiri MR. 2009. Self-assembling sequence-adaptive peptide nucleic acids. *Science* **325:** 73–77.

Usher DA, McHale AH. 1976. Hydrolytic stability of helical RNA—Selective advantage for natural 3′,5′-bond. *Proc Natl Acad Sci* **73:** 1149–1153.

Cold Spring Harbor Perspectives in Biology

www.cshperspectives.org

Vasas V, Szathmáry E, Santos M. 2010. Lack of evolvability in self-sustaining autocatalytic networks constraints metabolism-first scenarios for the origin of life. *Proc Natl Acad Sci* **107:** 1470–1475.

Voegel JJ, Altorfer MM, Benner SA. 1993. The donor-acceptor-acceptor purine analog-transformation of 5-aza-7-deaza-1H-isoguanine (= 4-aminoimidazo-[1,2-α]-1,3,5-triazin-2(1H)-one) to 2′-deoxy-5-aza-7-deaza-isoguanosine using purine nucleoside phosphorylase. *Helv Chim Acta* **76:** 2061–2069.

Weber AL. 1992. Prebiotic sugar synthesis: hexose and hydroxy acid synthesis from glyceraldehyde catalyzed by iron(III) hydroxide oxide. *J Mol Evol* **35:** 1–6.

Weber AL. 2001. The sugar model: Catalysis by amines and amino acid products. *Orig Life Evol B* **31:** 71–86.

Westheimer FH. 1987. Why Nature chose phosphates. *Science* **235:** 1173–1178.

Wiberg KB, Morgan KM, Maltz H. 1994. Thermochemistry of carbonyl reactions. 6. A study of hydration equilibria. *J Am Chem Soc* **116:** 11067–11077.

Wilds CJ, Wawrzak Z, Krishnamurthy R, Eschenmoser A, Egli M. 2002. Crystal structure of a B-form DNA duplex containing (L)-α- threofuranosyl (3′ → 2′) nucleosides: A four-carbon sugar is easily accommodated into the backbone of DNA. *J Am Chem Soc* **124:** 13716–13721.

Wimberly BT, Brodersen DE, Clemons WM, Morgan-Warren RJ, Carter AP, Vonrhein C, Hartsch T, Ramakrishnan V. 2000. Structure of the 30S ribosomal subunit. *Nature* **407:** 327–339.

Woese C. 1967. The evolution of the genetic code. In *The genetic code* (Eds.), 179–195. Harper & Row, New York.

Yakovchuk P, Protozanova E, Frank-Kamenetskii MD. 2006. Base-stacking and base-pairing contributions into thermal stability of the DNA double helix. *Nucleic Acids Res* **34:** 564–574.

Zhang LL, Meggers E. 2005. An extremely stable and orthogonal DNA base pair with a simplified three-carbon backbone. *J Am Chem Soc* **127:** 74–75.

Zielinski WS, Orgel LE. 1985. Oligomerization of activated derivatives of 3′-amino-3′-deoxyguanosine on poly(C) and poly(dG) templates. *Nucleic Acids Res* **13:** 2469–2484.

Zielinski WS, Orgel LE. 1987. Oligoaminonucleoside phosphoramidates. Oligomerization of dimers of 3′-amino-3′-deoxy-nucleotides (GC and CG) in aqueous solution. *Nucleic Acids Res* **15:** 1699–1715.

Zubay G, Mui T. 2001. Prebiotic synthesis of nucleosides. *Origins Life Evol B* **31:** 87–102.

Cold Spring Harbor Perspectives in Biology

www.cshperspectives.org

Primitive Genetic Polymers

Aaron E. Engelhart and Nicholas V. Hud

Subject Collection   The Origins of Life

For additional articles in this collection, see http://cshperspectives.cshlp.org/cgi/collection/