Origins of Life and Evolution of Biospheres (2006) 36: 39–63

GLYOXYLATE AS A BACKBONE LINKAGE FOR A PREBIOTIC ANCESTOR OF RNA

HEATHER D. BEAN[1], FRANK A. L. ANET[2], IAN R. GOULD[3] and NICHOLAS V. HUD[1,*]

[1]*School of Chemistry and Biochemistry, Parker H. Petit Institute for Bioengineering and Bioscience, Georgia Institute of Technology, Atlanta, GA 30332, U.S.A.;* [2]*Department of Chemistry and Biochemistry, University of California, Los Angeles, CA 90095, U.S.A.;* [3]*Department of Chemistry, Imperial College London, London SW7 2AZ, U.K.*

*(*author for correspondence, e-mail: hud@chemistry.gatech.edu)*

(Received 26 May 2005; accepted in revised form 15 July 2005)

Abstract. The origin of the first RNA polymers is central to most current theories for the origin of life. Difficulties associated with the prebiotic formation of RNA have lead to the general consensus that a simpler polymer preceded RNA. However, polymers proposed as possible ancestors to RNA are not much easier to synthesize than RNA itself. One particular problem with the prebiotic synthesis of RNA is the formation of phosphoester bonds in the absence of chemical activation. Here we demonstrate that glyoxylate (the ionized form of glyoxylic acid), a plausible prebiotic molecule, represents a possible ancestor of the phosphate group in modern RNA. Although in low yields (∼1%), acetals are formed from glyoxylate and nucleosides under neutral conditions, provided that metal ions are present (e.g., Mg^{2+}), and provided that water is removed by evaporation at moderate temperatures (e.g., 65 °C), i.e. under "drying conditions". Such acetals are termed *ga*-dinucleotides and possess a linkage that is analogous to the backbone in RNA in both structure and electrostatic charge. Additionally, an energy-minimized model of a *ga*RNA duplex predicts a helical structure similar to that of A-form RNA. We propose that glyoxylate-acetal linkages would have had certain advantages over phosphate linkages for early self-replicating polymers, but that the distinct functional properties of phosphoester and phosphodiester bonds would have eventually lead to the replacement of glyoxylate by phosphate.

Keywords: acetal, backbone analog, glyoxylic acid, nucleotide analog, prebiotic synthesis, RNA ancestor, proto-RNA

1. Introduction

The persistent lack of a feasible solution to the prebiotic synthesis of RNA has lead most researchers to conclude that RNA was preceded by one or more ancestral RNA-like polymers (Orgel, 2004). Recently, both peptide nucleic acid (PNA) and threose nucleic acid (TNA) have garnered much attention as possible ancestors to RNA (Böhler *et al.*, 1995; Chaput *et al.*, 2003; Herdewijn, 2001; Joyce, 2002; Miller, 1997; Piccirilli, 1995; Schwartz, 1997). However, the PNA backbone is so chemically distinct from RNA, including a lack of negative charges, that it is difficult to imagine how PNA could have ever functioned in place of RNA (Benner,

2004; Westheimer, 1987). TNA, which has a simpler sugar than ribose, still includes phosphodiester bonds, and thus suffers from some of the same difficulties associated with the prebiotic synthesis of RNA. In particular, phosphodiester bond formation requires chemical activation, and suitable phosphate precursors were likely scarce on the early Earth (Keefe and Miller, 1995).

Glyoxylic acid exists in water at neutral pH in equilibria with its ionized form and also with its hydrate form, but almost entirely as the ionized aldehyde hydrate, a form that for simplicity we will refer to as *glyoxylate*. The structural similarity of glyoxylate and dihydrogen phosphate is clear (Figure 1). Close structural analogs of modern nucleotides would be created by the formation of hemiacetal bonds between glyoxylate and nucleoside sugars, which we term *ga*-nucleotides (Figure 1). Furthermore, nucleosides connected through glyoxylate by acetal bonds

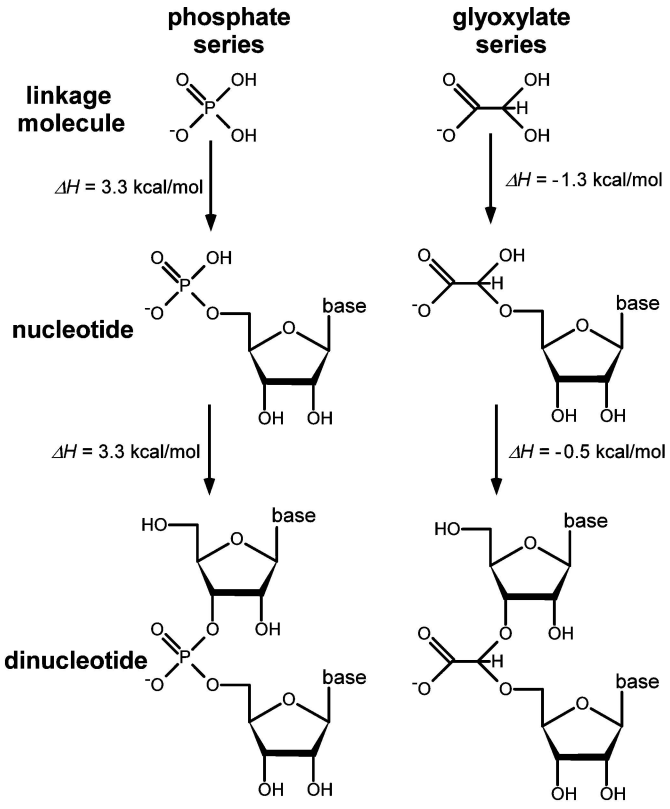


Figure 1. Comparison of phosphate and glyoxylate chemical structures, and their associated nucleotides and dinucleotides. ΔH values represent enthalpies of phosphoester bond formation (3.3 kcal/mol) (Ould-Moulaye *et al.*, 2002), glyoxylic acid-methyl hemiacetal bond formation (-1.3 kcal/mol) and glyoxylic acid-dimethyl acetal bond formation (-0.5 kcal/mol). The experiments used to determine glyoxylic acid hemiacetal and acetal bond formation enthalpies are described in Section 2.1 of Methods.

would likewise represent close structural analogs of RNA and DNA, termed collectively as *ga*NA (Figure 1). Nucleotide analogs with a formaldehyde moiety (i.e. $–OCH_2O–$) replacing the natural phosphodiester group have previously been shown to adopt a helical structure similar to that of natural nucleic acids (Rice and Gao, 1997). Thus, it is expected that the glyoxylate-acetal backbone of *ga*NA would have structural and electrostatic properties similar to that of modern RNA and DNA.

Although there are structural similarities between natural nucleotides and *ga*-nucleotides, there are several important differences. First, we note that glyoxylate is more likely to have been available on the prebiotic Earth than free phosphate. Amines and amino acid products catalyze the formation of glyoxylate from glycolaldehyde, a widely accepted prebiotic molecule (Weber, 2001). Second, the chemistry of *ga*NA versus modern nucleic acids is functionally different. Hemiacetal bond formation between an aldehyde hydrate, such as glyoxylate, and an alcohol (e.g. an OH of a nucleoside) is thermodynamically favored, as is acetal bond formation between a hemiacetal, such as a *ga*-nucleotide, and an alcohol (Wiberg *et al.*, 1994). Since the formation of a phosphodiester bond is thermodynamically "uphill", previous studies of protein-free template-directed syntheses of phosphodiester-linked oligonucleotides have required the use of chemically activated nucleotides, such as 5′-phospho-2-methylimidazolide nucleotides (2-MeImpN) (Joyce *et al.*, 1984). Although activated mononucleotides will spontaneously form oligonucleotides along existing nucleic acid templates or when dried onto certain mineral surfaces (Ferris *et al.*, 1996), chemically activated nucleotides are of questionable relevance to the origin of life as their synthesis by a plausible prebiotic chemistry has not been demonstrated (Orgel, 2004). Although yields are low (∼1%), the method described here for creating *ga*-nucleotide linkages represents, to the best of our knowledge, the first demonstration that an RNA-like linkage can be created with plausible prebiotic chemistry.

2. Methods

2.1. Determination of enthalpies of hemiacetal and acetal bond formation

The enthalpy of hemiacetal formation by glyoxylate (Figure 1) was determined by measuring the equilibrium constant (based upon peak integration in NMR spectra) of glyoxylic acid hydrate and the glyoxylic acid hemiacetal (*R* and *S* forms) formed with methanol in a 100 mM glyoxylic acid solution, 20% w/v methyl-d3 alcohol-d/80% w/v D_2O (pD 1) at 65 °C ($K = 9.4$; $\Delta G_{65\,^\circ C} = -1.5$ kcal/mol) and 85 °C ($K = 8.4$; $\Delta G_{85\,^\circ C} = -1.5$ kcal/mol). The free energy of acetal formation by glyoxylic acid (from its hemiacetal form) was determined by measuring the equilibrium constant of the glyoxylic acid-methyl hemiacetal (*R* and *S* forms) and the glyoxylic acid-dimethyl acetal in a 100 mM glyoxylic acid solution, 20% w/v methyl-d3

alcohol-d/80% w/v D_2O (pD 1) at 65 °C (K = 1.4; $\Delta G_{65\,^\circ C} = -0.7$ kcal/mol). The enthalpy of glyoxylic acid acetal formation was derived from the free energy measurement using the entropy measured for the hydrate/hemiacetal reaction in the same sample, which also coincides with the addition of one methanol and the liberation of one water molecule. For both free energy derivations reported, peak integration in NMR spectra was taken over the glyoxylic acid and the methyl-ester forms of glyoxylic acid that exist in equilibrium exchange at pD 1.

2.2. Synthesis of *ga*-dinucleotides with metal ion hydrates

ga-Dinucleotide yields presented in the text were measured from reaction samples that contained 20 μmol dT, 10 μmol glyoxylate (pH 6.4), and 100 μmol of the chloride salt for each metal ion. In the case of $MgCl_2$, *ga*-dinucleotide yield was determined to be highest at a relative concentration of salt to glyoxylate of approximately 10:1 (data not shown). This same ratio was used for all metal ion species tested. The samples were dried at 22 °C under vacuum or at 65 or 85 °C at atmospheric pressure, and then resuspended in a 0.25 N NH_4OH solution.

2.3. Mass spectrometry

Mass spectra of *ga*-nucleotides and *ga*-dinucleotides were collected in negative ion mode on a Micromass Quattro LC mass spectrometer and ionized under electrospray conditions (capillary voltage, 2.5 kV; cone voltage, 40 V). Accurate mass analyses, used to confirm *ga*-nucleotide and *ga*-dinucleotide formation, were accomplished using the negative ion mode on an Applied Biosystems QSTAR XL hybrid Quadrupole-TOF mass spectrometer under electrospray conditions (capillary voltage, −4.2 kV; cone voltage, −70 V).

2.4. HPLC and NMR spectroscopy

HPLC separation of dT/glyoxylate reaction products was achieved on a Phenomenex ODS analytical column (4.6 × 250 mm, 5 μ) with mobile phase conditions of 10–30% B (0–5 min), 30–50% B (5–30 min), flow rate = 1 ml/min. Mobile phase A = 0.1 M triethyl ammonium acetate (TEAA), B = 25% (v/v) acetonitrile in 0.1 M TEAA. dT-*ga*-dT *ga*-dinucleotides for the NMR study shown in Figure 4B were prepared in bulk by drying under vacuum at room temperature ten solutions of 20 μmol dT nucleoside, 10 μmol glyoxylic acid, and 15 μmol HCl in 200 μl dH_2O. Reaction products were resuspended in 0.25 N NH_4OH, pooled and purified by the same HPLC gradient with a Phenomenex ODS semi-preparative column (10 × 250 mm, 5 μ). The process of *ga*-dinucleotide isolation and purification resulted in less than 15% degradation into monomer dT, as determined by NMR. Products of bulk dT-*ga*-dT *ga*-dinucleotide preparations were correlated with products of $MgCl_2$-catalyzed reactions by simultaneous injection

onto an analytical HPLC column. NMR spectra of *ga*-dinucleotides were recorded on a Bruker DRX500 spectrometer in D_2O at 25 °C.

2.5. Molecular modeling

Partial electrostatic charges for the glyoxylate atoms and ribose atoms of *ga*-oligonucleotides were determined by fitting the electrostatic surface potential obtained from a high-level quantum mechanical calculation of ribose-3′-glyoxylate-5′-ribose as described in the Cornell force field paper (Cornell *et al.*, 1995). The Amber suite of programs (Case *et al.*, 2004) was used for generation, energy minimization, and molecular dynamics simulations of a *ga*RNA duplex. The model *ga*RNA duplex with the self-complementary sequence A_6U_6 was constructed by replacing the phosphate groups of an RNA duplex $[A_6U_6]_2$ in a canonical A-form helix with glyoxylate. *R* and *S* isomers of the glyoxylate-acetal linkages between nucleotides were varied along the two strands as: 5′-A-*Rga*A-*Sga*A-*Sga*A-*Rga*A-*Rga*A-*Sga*U-*Rga*U-*Sga*U-*Rga*U-*Rga*U-*Sga*U-3′ and 5′-A-*Sga*A-*Rga*A-*Sga*A-*Rga*A-*Sga*A-*Rga*U-*Rga*U-*Rga*U-*Sga*U-*Rga*U-*Sga*U-3′, respectively. The energy minimization and molecular dynamics simulation protocol used was adapted from that described previously for nucleic acids simulations (Cheatham and Kollman, 1997). Briefly, prior to energy minimization, 22 Na^+ counterions and 4142 solvent water molecules were added around the *ga*RNA duplex in a truncated octahedral solvent box that extended at least 8 Å from all solute atoms. Solvent water molecules and cations were energy minimized by 500 steps of steepest descent and 500 steps of conjugate gradient, while holding the *ga*RNA duplex atoms fixed. A second minimization was then carried out with no constraints on the duplex or solvent. The model duplex and solvent molecules were heated to 300 K over the course of 20 ps, and then subjected to 1 ns of molecular dynamics simulation.

3. Results

3.1. Enthalpies of glyoxylate hemiacetals and acetals

Based upon NMR analysis of glyoxylic acid dissolved in a water-methanol mixture, we have determined that a glyoxylic acid-methyl hemiacetal is enthalpically favored over its hydrate species by 1.3 kcal/mol, and a glyoxylic acid-dimethyl acetal is enthalpically favored over its hemiacetal by an additional 0.5 kcal/mol. These enthalpies of hemiacetal and acetal bond formation between glyoxylic acid and methanol are used as estimates for the bond formation enthalpies of glyoxylate-nucleoside hemiacetal and acetal species (Figure 1). It is expected that these enthalpy values are quite reliable, as the enthalpies of hemiacetal and acetal formation have previously been shown to only change by a few percent with the identity of the

alcohol (Wiberg *et al.*, 1987). These enthalpies of bond formation are in contrast to the 3.3 kcal/mol of enthalpy that must be provided to phosphorylate a nucleoside, and the 3.3 kcal/mol required for the subsequent formation of a phosphodiester linkage (Ould-Moulaye *et al.*, 2002). The negative enthalpies of formation of hemiacetals and acetals suggest that *ga*-nucleotides, *ga*-dinucleotides, and *ga*NA polymers will form spontaneously under appropriate conditions.

3.2. Formation of glyoxylate hemiacetals with nucleosides

The formation of a hemiacetal by an alcohol (e.g. the OH of ribose) and an aldehyde hydrate (e.g. glyoxylate) is a dehydration reaction that releases one molecule of water and takes place in acidic, basic, or neutral pH conditions. The existence of *ga*-nucleotides in neutral aqueous solutions containing a mixture of glyoxylate and each of the RNA and DNA nucleosides is supported by mass spectrometry. Samples containing 100 mM glyoxylate and 100 mM nucleoside (pH 6.9) were prepared at room temperature and injected directly into the mass spectrometer. In each case a prominent peak in the mass spectrum confirmed formation of the *ga*-nucleotide (Table I). *ga*-rU and *ga*-dT nucleotides were further characterized by high-resolution mass spectrometry, which confirmed that the observed masses for these nucleotides match their theoretical masses within 3 ppm error (Table III).

Although all eight *ga*-nucleotides were observed by mass spectrometry, the electrospray ion source of the mass spectrometer desolvates samples prior to ion separation and detection. This desolvation process within the mass spectrometer creates dehydrating conditions that could promote hemiacetal (e.g. *ga*-nucleotide) formation. To confirm that *ga*-nucleotides are also formed in the solution state, a sample containing 1 M glyoxylate (pH 6) and 1 M 1-O-Methyl-2-deoxy-D-ribose (dOM), a model of a nucleoside sugar, was analyzed by solution state NMR spectroscopy (Figure 2). New chemical species were observed to form upon mixing glyoxylate

TABLE I
Mass spectrometry of hemiacetal peaks

ga-Nucleotide	Observed m/z[a]
ga-rU	317
ga-rC	316
ga-rA	340
ga-rG	356
ga-dT	315
ga-dC	300
ga-dA	324
ga-dG	340

[a]Data collected in negative ion mode as described in Methods

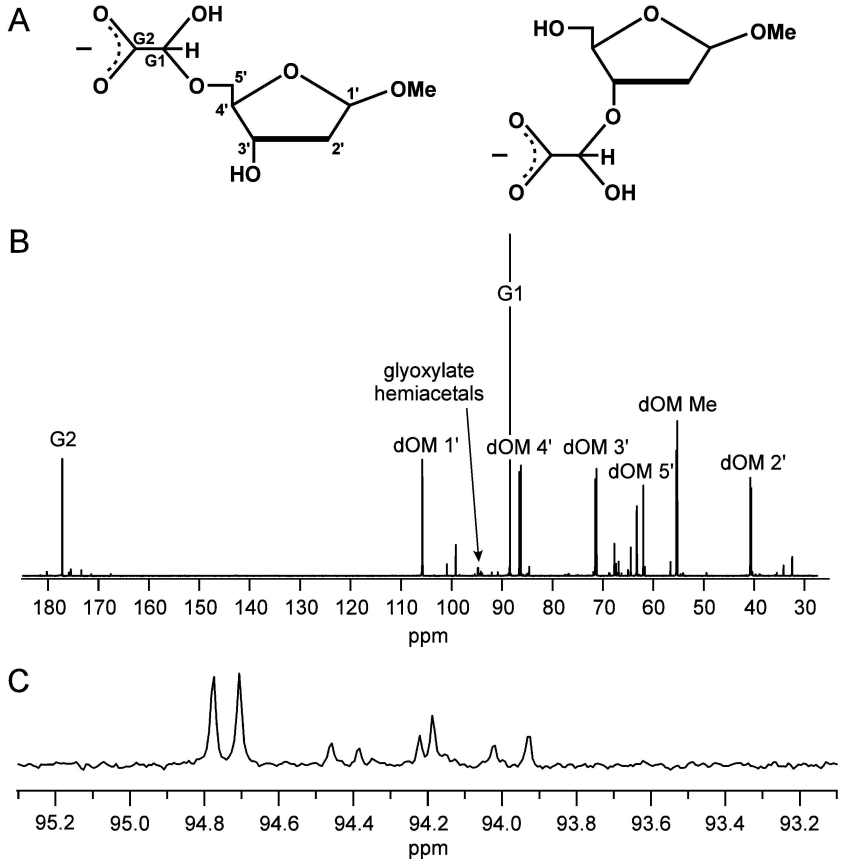


Figure 2. Verification of *ga*-nucleotide formation in solution by glyoxylate and 1–O-Methyl-2-deoxy-D-ribose (dOM). **(A)** Structures of the 3′ and 5′ hemiacetals. The dOM exists in both α and β forms, and the glyoxylate moiety is prochiral, giving rise to eight possible hemiacetal structures. **(B)** Natural abundance ^{13}C NMR spectrum of 1 M glyoxylate and 1 M dOM in D_2O (pH 6). At this approximate ratio of 55:1 D_2O:dOM, the dominant signals in the NMR spectrum correspond to free dOM and free glyoxylate. The smaller peaks indicated by arrow correspond to the glyoxylate CH (G1) resonances of 3′ and 5′ *ga*-dOM hemiacetals. Resonance assignments are based upon correlations observed in 2D HMQC and HMBC spectra (not shown). **(C)** Expanded region of the NMR spectrum (B) that contains the G1 resonances of the *ga*-dOM hemiacetals.

and dOM, which are consistent with the formation of *ga*-dOM nucleotide analogs in solution. The use of dOM instead of any single nucleoside was done for solubility reasons. The equilibrium constant of the glyoxylate hydrate-hemiacetal system in H_2O-ROH is approximately 10 (Methods). Thus, in a solution with D_2O present at a nominal concentration of 55 M, confirmation of *ga*-nucleotide formation by NMR spectroscopy in the presence of free dOM and the glyoxylate hydrate requires a dOM concentration that is beyond the solubility limit of the natural nucleosides (i.e. approximately 1 M) and is therefore difficult to study in the solution state.

3.3. FORMATION OF GLYOXYLATE-LINKED DINUCLEOTIDES

The formation of an acetal bond in the reaction of a hemiacetal and an alcohol, such as that required for the formation of a *ga*-dinucleotide, also represents a dehydration reaction (loss of a water molecule), but requires acid catalysis. It does not take place under neutral or basic conditions, in contrast to the reaction for hemiacetal formation. Acetal bond formation would typically be carried out in a synthetic laboratory using a strong acid catalyst and an anhydrous medium. However, life likely emerged from an aqueous environment within the pH range of 6–9 (Bengston, 1994). Given these restraints, we sought to catalyze acetal bond formation using cationic metal hydrates as general acids and evaporation of water by drying as a means to reduce the water activity to a very low level. Similar conditions have previously proven fruitful in the prebiotic synthesis of nucleosides from the free bases and ribose (Fuller *et al.*, 1972).

Samples containing 50 mM glyoxylate, 100 mM nucleoside, and 500 mM $MgCl_2$ (pH 6.4) were dried at 85 °C in an open-air test tube and then resuspended in dH_2O at room temperature. Using mass spectrometry we have verified that nucleosides dT, dC, rU, rC, and rA will couple under these conditions to form homo *ga*-dinucleotides (Table II). The *ga*-dinucleotides dT-*ga*-dT and rU-*ga*-rU were further characterized by high-resolution mass spectrometry (Table III). The remaining nucleosides, dA, dG, and rG, did not produce *ga*-dinucleotides in detectable yields, and no dinucleotide formation was observed without sample drying. A correlation exists between nucleoside solubility and *ga*-dinucleotide production, strongly suggesting that nucleosides dA, dG, and rG crystallize from solution before acetal bond formation can occur.

We have also investigated the effects of temperature and cationic metal species on *ga*-dinucleotide formation during sample drying. For this study, samples of dT and glyoxylate were prepared with eight different metal chlorides and dried at

TABLE II

Mass spectrometry of acetal peaks

ga-Dinucleotide	Observed m/z[a]
rU-*ga*-rU	543
rC-*ga*-rC	541
rA-*ga*-rA	589
rG-*ga*-rG	nd[b]
dT-*ga*-dT	539
dC-*ga*-dC	509
dA-*ga*-dA	nd[b]
dG-*ga*-dG	nd[b]

[a]Data collected in negative ion mode as described in Methods
[b]Not detected

TABLE III
High resolution mass spectrometry data for selected hemiacetal and acetal species

Species	Molec. formula[a]	Observed m/z[b]	Calculated mass	ppm error
ga-rU	$C_{11}H_{13}N_2O_9$	317.0616	317.0621	1.6
ga-dT	$C_{12}H_{15}N_2O_8$	315.0837	315.0828	2.7
rU-*ga*-rU	$C_{20}H_{23}N_4O_{14}$	543.1177	543.1211	6.2
dT-*ga*-dT	$C_{22}H_{27}N_4O_{12}$	539.1593	539.1626	6.0

[a]Formula of the negatively ionized species
[b]Data collected in negative ion mode as described in Methods

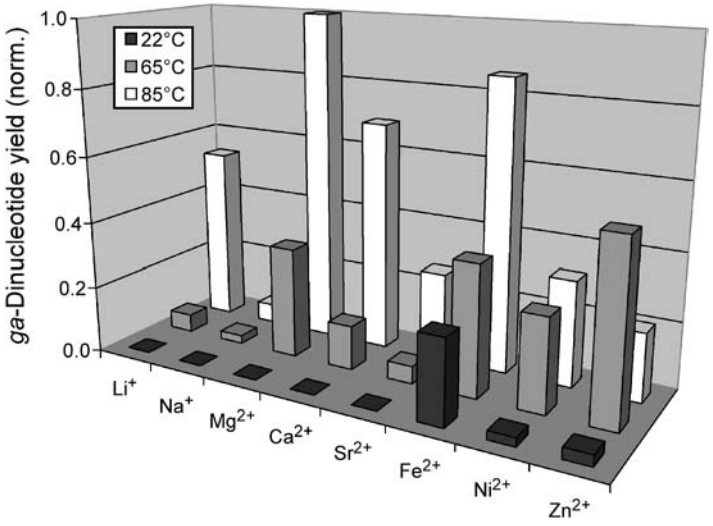


Figure 3. Effects of metal ion species and temperature on the formation of dT-*ga*-dT *ga*-dinucleotides. Each sample contained 20 μmol dT, 10 μmol glyoxylate (pH 6.4), and 100 μmol of the chloride salt for each metal ion. The 65 and 85 °C reaction mixtures were dried at atmospheric pressure; the room temperature samples were dried under vacuum. Five identical samples (initially 200 μL) were dried and analyzed by HPLC for each salt/temperature combination. Yields are based upon the average integrated HPLC peak intensities for the five samples of each reaction condition. All yields are normalized with respect to the yield of the $MgCl_2$ samples dried at 85 °C.

22, 65, or 85 °C (Figure 3). At 22 °C only the divalent metal ions Fe^{2+}, Ni^{2+}, and Zn^{2+} produced detectable yields of dT-*ga*-dT *ga*-dinucleotides. At 65 and 85°C all metal ions tested produced appreciable amounts of *ga*-dinucleotides. The highest yield was observed for the solution containing Mg^{2+} that was dried at 85 °C, and corresponded to the conversion of approximately 1% of starting dT nucleoside into *ga*-dinucleotides. All other values shown in Figure 3 are normalized to this reaction, and lie roughly within an order of magnitude for all metal ions tested at 65 and 85 °C. Thus, acetal bond formation by glyoxylate and nucleosides requires a metal ion catalyst, but does not strongly depend on the elemental species of the metal

ion. From the perspective of prebiotic chemistry, Mg^{2+} was likely to have been widely distributed in liquid water on the prebiotic Earth, as it is today. Furthermore, temperatures around 65 °C would have likely been a common occurrence on the early Earth (Bengston, 1994). Hence, nucleosides on the prebiotic Earth could have been coupled through glyoxylate-acetal linkages without the need for any additional catalysts, energy sources, or chemical activation, provided that nucleoside organization occurred before reactant solubility became limiting (see Section 4.1).

3.4. Characterization of dT-*ga*-dT dinucleotide linkages

Deoxyribose nucleosides and glyoxylate have the potential to form acetal-linked dinucleotides involving all combinations of the 3′-OH and 5′-OH groups of deoxyribose (i.e. 3′-*ga*-5′, 5′-*ga*-5′ and 3′-*ga*-3′ linkages), and ribonucleosides can also form linkages involving the 2′-OH group of ribose. Additionally, the CH carbon of glyoxylate is prochiral and can give rise to *R*and *S* isomers when the two acetal groups are different from one another. The possible structural implications of coexisting *R* and *S* isomers are addressed below (Section 3.5). In the case of *ga*-dinucleotide formation, acetal bonds involving the 5′-OH, a primary alcohol, would be expected to be more favored than bonds involving one or more of the secondary alcohols (i.e. 2′-OH and 3′-OH) (De Proft *et al.*, 1995; Wiberg *et al.*, 1987). To discern which of the possible *ga*-dinucleotides are actually synthesized, we conducted a more in-depth analysis of the products formed when dT is dried from a neutral solution containing glyoxylate and $MgCl_2$. HPLC analysis of reaction products exhibited several peaks with relative mobilities close to that of the natural dinucleotide dT-p-dT (Figure 4A). Mass spectrometry confirmed that several of these elution peaks have a mass corresponding to a dT-*ga*-dT dinucleotide. HPLC peaks labeled A, B, and C were also present in a bulk acid-catalyzed preparation, allowing the compounds in these peaks to be obtained in sufficient quantity for characterization by NMR.

In Figure 4B proton NMR spectra for the *ga*-dinucleotide HPLC peaks A, B, and C are shown along with an NMR spectrum of the natural dT-p-dT dinucleotide. The proton resonance in the 4.8–5.0 ppm region of the *ga*-dinucleotide spectra belongs to the single CH proton of the glyoxylate moiety. 2D correlation ^{1}H-^{13}C spectroscopy (HMBC) revealed peak C to be largely the 5′-*ga*-5′ dinucleotide, and peaks A and B to be largely 3′-*ga*-5′ linked dinucleotides (Tables IV–IX). Assignment of the 3′-*ga*-5′ dinucleotide HPLC peaks A and B to specific *R* or *S* isomers at the glyoxylate moiety is currently underway. Analysis of additional HPLC elution peaks gave no indication of dinucleotides with 3′-*ga*-3′ linkages, which is consistent with the reaction of two secondary alcohols with glyoxylate being the least favored product of the four possible linkages. We note that the observation of more than three peaks in the dinucleotide region of the HPLC trace implies the creation of more products than the 5′-*ga*-5′ and 3′-*ga*-5′ dinucleotides.

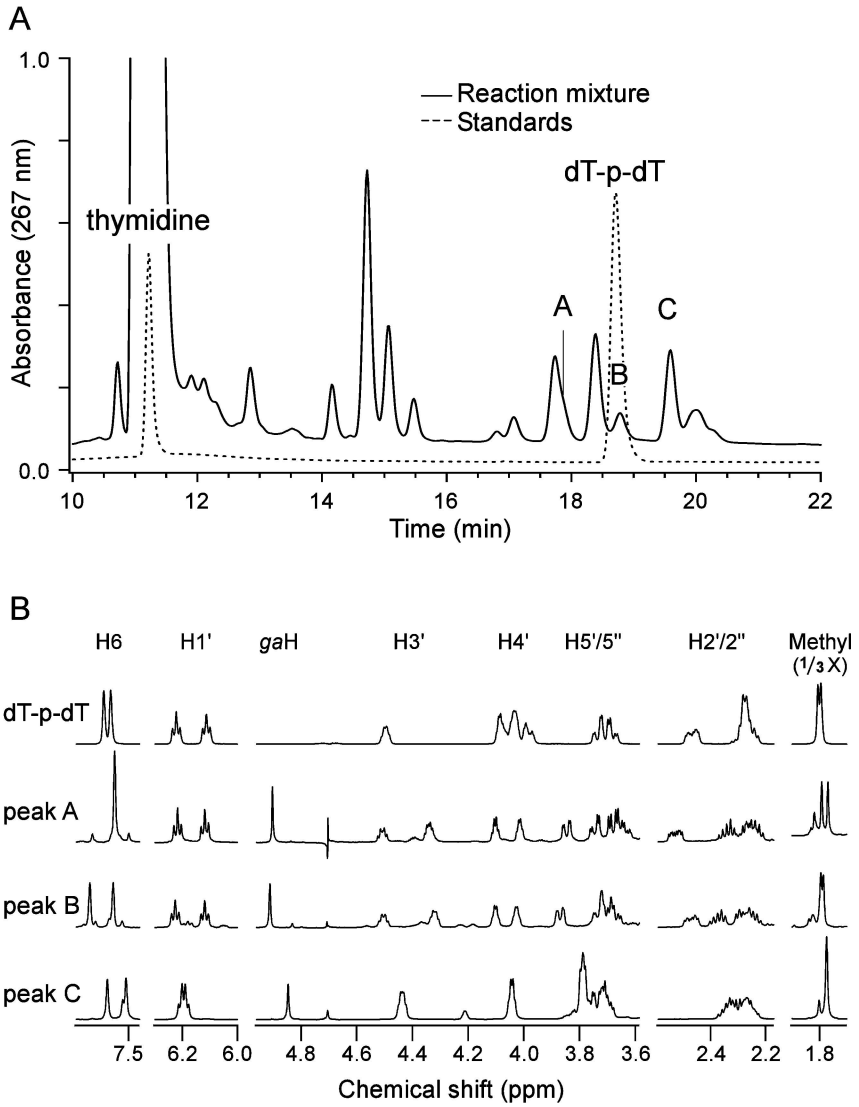


Figure 4. Chromatographic and spectroscopic analyses of *ga*-dinucleotide products of deoxythymidine reacted with glyoxylate under drying conditions. **(A)** Overlay of an HPLC chromatograph of a glyoxylate/thymidine reaction mixture (solid line) with a chromatograph of thymidine and dT-p-dT standards (dashed line). Seven peaks with relative mobilities similar to dT-p-dT, including those marked A, B, and C, were confirmed to be *ga*-dinucleotides by mass spectrometry. **(B)** Comparison of the ^{1}H spectrum of the dT-p-dT dinucleotide with spectra of the *ga*-dinucleotide products designated as HPLC peaks A, B, and C. The H3′ resonance of the 5′ nucleoside of dT-p-dT lies beneath the solvent peak (4.7 ppm), which was suppressed by a presaturation pulse.

TABLE IV
NMR data for peak A in D_2O

Group[e]	Position	δ_C, ppm	δ_H, ppm[a]
dT1	2	151.8	
	6	137.7	7.53 (s)
	1′	85.9	6.10(t)
	2′[c]	37.2	2.23 (m)
			2.51 (m)
	3′	77.4	4.32 (m)
	4′	85.1	4.08 (m)
	5′[d]	61.4	3.65 (d,d)
			3.73 (d,d)
Glyoxylate	G1 (CH)	100.4	4.89 (s)
	G2 (COO^-)	173.9	
dT2	2	152.0	
	6	137.9	7.53 (s)
	1′	84.8	6.20 (t)
	2′[c]	38.7	2.23 (m)
			2.31 (m)
	3′	70.5	4.49 (m)
	4′	85.0	4.00 (m)
	5′[d]	64.7	3.64 (d,d)
			3.83 (d,d)
Unassigned[b]	4	166.4	
	4	166.6	
	5	111.6	
	5	111.8	
	M	11.9	1.77 (s)
	M	11.9	1.74 (s)

[a] Multiplicities are indicated in parentheses
[b] Unable to assign these peaks to a particular nucleoside
[c] H2′ is highfield, H2″ is lowfield (Wood, *et al.*, 1974)
[d] H5′ is lowfield, H5″ is highfield (Wood, *et al.*, 1974)
[e] Group and position numbers refer to structure in Table V

It is possible that some of these products represent *ga*-dinucleotides linked through one or both bases by iminal-type bonds.

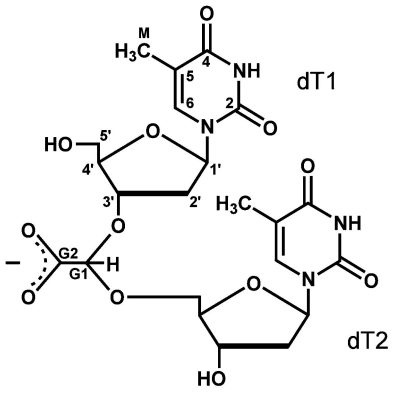
3.5. Stability of dT-*ga*-dT dinucleotide linkages

Acetal bonds are known to be stable in neutral and basic pH conditions, but readily labile in an aqueous solution of acidic pH. Nevertheless, we have found that the acetal bonds formed by glyoxylate and nucleosides are relatively stable

TABLE V
HMBC correlations for peak A

Group	Proton	Carbon
dT1	6	2, 4, M
	1′	2, 6
	2′	1′, 3′
	2″	3′
	4′	3′
	5′	3′
	5″	3′
Glyoxylate	G1	G2, dT2 5′, dT1 3′
dT2	6	2, 4, M
	1′	2, 6
	2′	1′, 3′
	2″	3′
	4′	3′
Unassigned[a]	M	4, 5, 6
	M	4, 5, 6

[a] Unable to assign these peaks to a particular nucleoside



in aqueous solution, even at pH 1. When dissolved in water acidified by HCl to pH 1, no appreciable degradation of dT-*ga*-dT dinucleotides was observed after a two-week incubation at 22 °C (Figure 5). Under the same acidic conditions at 60 °C, considerable degradation of dT-*ga*-dT into its components was observed. The addition of 1 M $MgCl_2$ to the acidic *ga*-dinucleotide solution was found to accelerate degradation at 60 °C, but had no visible effect at 22 °C. Under neutral conditions (pH 7) neither heating to 60°C nor incubation with 1 M $MgCl_2$ for two-weeks resulted in appreciable *ga*-dinucleotide degradation (Figure 5).

TABLE VI
NMR data for peak B in D_2O

Group[e]	Position	δ_C, ppm	δ_H, ppm[a]
dT1	2	152.2	
	4	167.4	
	5	112.4	
	6	138.4	7.53 (s)
	1′	86.6	6.09 (t)
	2′[c]	38.3	2.20 (m)
			2.44 (m)
	3′	76.4	4.29 (m)
	4′	85.9	4.08 (m)
	5′[d]	62.1	3.65 (m)
			3.72 (m)
Glyoxylate	G1 (CH)	100.4	4.89 (s)
	G2 (COO^-)	174.3	
dT2	2	152.4	
	4	167.0	
	5	112.4	
	6	138.8	7.62 (s)
	1′	85.9	6.20 (t)
	2′[c]	39.1	2.25 (m)
			2.34 (m)
	3′	71.4	4.48 (m)
	4′	85.6	4.00 (m)
	5′[d]	66.7	3.69 (m)
			3.84 (d,d)
Unassigned[b]	M	11.9	1.77 (s)
	M	11.9	1.78 (s)

[a] Multiplicities are indicated in parentheses
[b] Unable to assign these peaks to a particular nucleoside
[c] H2′ is highfield, H2″ is lowfield (Wood, *et al.*, 1974)
[d] H5′ is lowfield, H5″ is highfield (Wood, *et al.*, 1974)
[e] Group and position numbers refer to structure in Table VII

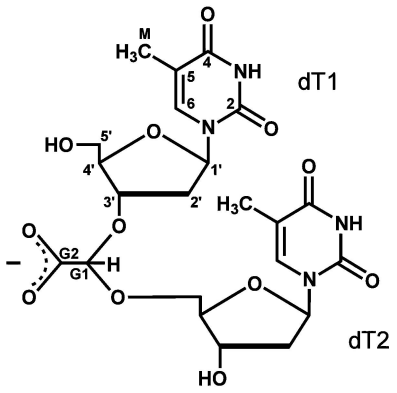
3.6. Modeling studies of *ga*RNA electrostatics and duplex structure

As mentioned above, the CH carbon of glyoxylate is prochiral, and therefore 3′-*ga*-5′ linked *ga*-dinucleotides can form as either the *R* or *S* isomers (Figure 6). In the absence of a mechanism that specifically enhances the formation of one isomer over the other, one would expect little selectivity in a prebiotic world. Formation of *ga*-dinucleotides with non-identical components raises the possibility that mixing glyoxylate-acetal *R* and *S* isomers *external* to a *ga*-oligonucleotide

TABLE VII
HMBC correlations for peak B

Group	Proton	Carbon
dT1	6	2, 4, M, 1′
	1′	2, 6
	2′	1′, 3′
	2″	3′
	3′	4′, G1
	4′	1′, 3′
	5′	3′, 4′
	5″	3′, 4′
Glyoxylate	G1	G2, dT2 5′, dT1 3′
dT2	6	2, 4, M, 1′
	1′	2, 6
	2′	3′
	2″	1′, 3′
	3′	4′, 5′
	4′	3′
	5′	3′
	5″	3′
Unassigned[a]	M	4, 5, 6
	M	4, 5, 6

[a] Unable to assign these peaks to a particular nucleoside



backbone could destabilize duplex formation, but this is very different from mixing ribose stereoisomers *internal* to the backbone, where dramatic effects are to be expected. For example, Watson-Crick base pair formation is not possible for a single L-ribose nucleotide incorporated within an otherwise D-ribose nucleotide duplex, as the Watson-Crick face of the base attached to the L-ribose is rotated approximately 100° away from the duplex helical axis (Avetisov and Goldanskii, 1996). In contrast,

TABLE VIII
NMR data for peak C in D_2O

Group[e]	Position	δ_C, ppm	δ_H, ppm[a]
Glyoxylate	G1 (CH)	100.1	4.83 (s)
	G2 (COO^-)	173.6	
Unassigned[b]	2	151.7	
	2	151.7	
	4	166.7	
	4	166.7	
	5	111.7	
	5	111.7	
	6	137.8	7.49 (s)
	6	138.0	7.56 (s)
	M	12.1	1.76 (s)
	M	12.1	1.76 (s)
	1′	85.5	6.17 (t)
	1′	85.5	6.18 (t)
	2′[c]	38.8	2.23 (m)
			2.32 (m)
	2′[c]	38.8	2.25 (m)
			2.30 (m)
	3′	71.5	4.42 (m)
	3′	71.5	4.42 (m)
	4′	85.3	4.03 (m)
	4′	85.3	4.03 (m)
	5′[d]	66.1	3.71 (m)
			3.77 (m)
	5′[d]	66.4	3.72 (m)
			3.74 (m)

[a] Multiplicities are indicated in parentheses
[b] Unable to assign these peaks to a particular nucleoside
[c] H2′ is highfield, H2″ is lowfield (Wood, *et al.*, 1974)
[d] H5′ is lowfield, H5″ is highfield (Wood, *et al.*, 1974)
[e] Group and position numbers refer to structure in Table IX

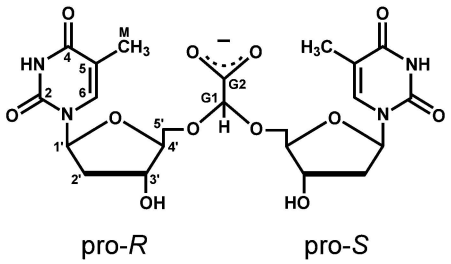
switching between the *R* and *S* isomers of glyoxylate is not expected to directly affect bond angles between the backbone and the nucleoside bases.

Molecular modeling was used to investigate the potential for mixed *R* and *S* isomers of glyoxylate-acetal linkages to interfere with *ga*NA duplex formation. Although determining the specific structural details of a *ga*-oligonucleotide duplex is beyond the present capabilities of computer-based molecular modeling, molecular dynamics simulations have successfully reproduced experimentally observed changes in the helical structures of nucleic acids, such as the DNA B-form to A-form transition (Cieplak *et al.*, 1997). Thus, we have performed modeling studies with

TABLE IX
HMBC correlations for peak C

Group	Proton	Carbon
Glyoxylate	G1	G2, pro-*R* 5′, pro-*S* 5′
Unassigned[a]	6	2, 4, 5, M, 1′
	M	4, 5, 6
	1′	2, 6, 3′
	2′	1′, 3′
	2″	1′, 3′
	3′	1′, 5′
	4′	1′, 3′
	5′	G1, 3′, 4′
	5″	G1, 3′, 4′

[a] Unable to assign these peaks to a particular nucleoside



the Amber force field (Cornell *et al.*, 1995) as an initial means to test if a *ga*RNA duplex containing mixed *R* and *S* isomers of the glyoxylate-acetal linkage would be expected to maintain a double helical structure resembling that of A-form RNA.

Partial charges for the atoms of the glyoxylate-acetal linkage were required to perform energy minimization and molecular dynamics simulations with the Amber force field (Cornell *et al.*, 1995). These partial atom charges were determined by an *ab initio* quantum mechanical calculation (Methods). The partial charges of the four oxygen atoms of a glyoxylate-acetal linkage are predicted to be very close to those of the corresponding atoms of a phosphodiester linkage (Figure 6A). Additionally, the two carbon atoms of glyoxylate and the CH hydrogen in this linkage are predicted to have partial positive charges with a sum total that is approximately equal to that of the phosphorus atom of a phosphodiester linkage (Figure 6A).

An initial model of a *ga*RNA duplex was constructed by changing the phosphate groups to glyoxylate in a model RNA double helix in the canonical A-form. The duplex was twelve base pairs in length, with the self-complementary sequence A_6U_6. *R* and *S* isomers of the glyoxylate-acetal linkages between nucleotides were varied along the two strands (Methods). In a periodic solvent box with explicit water molecules and Na^+ counterions, the *ga*RNA duplex model was energy minimized

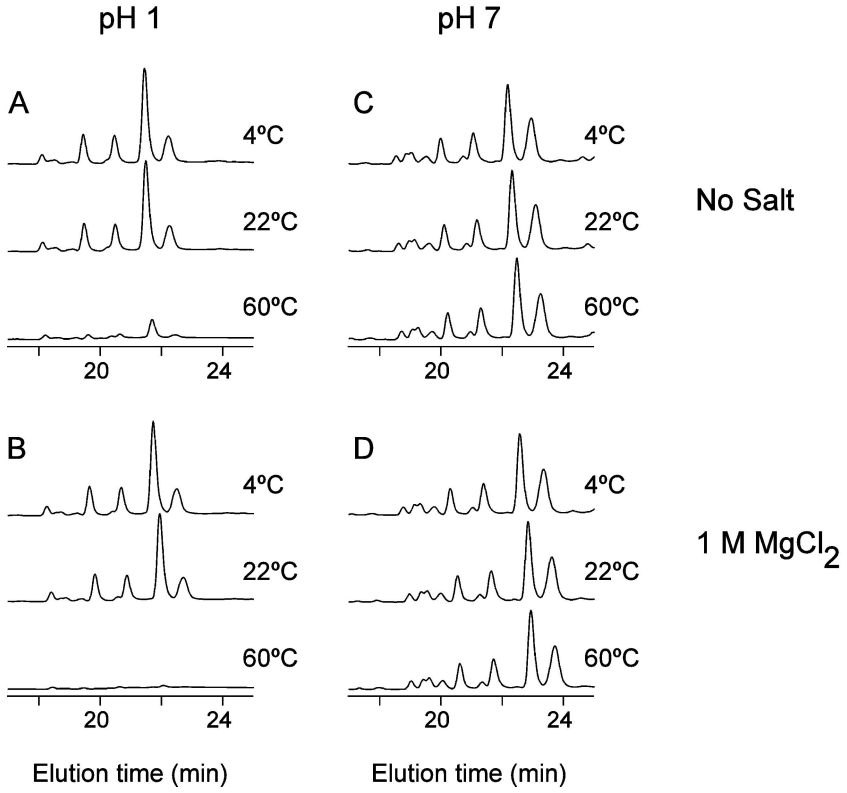


Figure 5. HPLC chromatographs illustrating the relative stability of dT-*ga*-dT *ga*-dinucleotides under various solution conditions. Dinucleotides were prepared by drying under vacuum at room temperature solutions of 20 μmol dT, 10 μmol glyoxylic acid, and 15 μmol HCl in 200 μL dH_2O. After drying, dT-*ga*-dT reaction solutions were resuspended at pH 1 or pH 7, in the presence or absence of 1 M $MgCl_2$, and then incubated for two weeks at 4, 22, or 60 °C. At day 14, samples were diluted five-fold by a 0.25 N NH_4OH solution prior to HPLC analysis. **(A)** Samples resuspended at pH 1 without $MgCl_2$. **(B)** Samples resuspended at pH 1 with 1 M $MgCl_2$. **(C)** Samples resuspended at pH 7 without $MgCl_2$. **(D)** Samples resuspended at pH 7 with 1 M $MgCl_2$.

and subjected to 1 ns of molecular dynamics simulation (Methods). The model retained a double helix structure similar to that of RNA, except for a partial unwinding of the helix (Figures 6B and C). This modeling study demonstrated that helix structure and base pair geometry within a *ga*-oligonucleotide duplex does not depend significantly on the particular isomer of a glyoxylate-acetal linkage, and is not affected by the coexistence of *R* and *S* linkages within the same backbone (Figure 6E). The similarities between a *ga*RNA duplex and a natural RNA duplex also suggests that it should be possible for *ga*RNA and RNA strands to form hybrid duplexes, which would be essential for the continuity of genetic transfer from one nucleic acid system to another during the evolutionary process.

4. Discussion

The RNA world hypothesis presents an attractive scenario for a very early stage of life (Gesteland and Atkins, 1999). However, wide acceptance of the RNA world hypothesis will ultimately require the discovery of a process by which RNA polymers, or RNA-like polymers, form spontaneously from plausible prebiotic molecules. The formation of the nucleotide bases and ribose from simple precursors (Joyce and Orgel, 1999), and the coupling of at least the purine bases and ribose upon heating in the presence of divalent cations (Fuller *et al.*, 1972), give reason to believe that the nucleosides are plausible prebiotic compounds. In contrast, the abiotic coupling of nucleotides into RNA polymers has only been achieved with chemically activated mononucleotides, and the activation chemistries employed are hard to imagine in the prebiotic world (Orgel, 2004). If RNA was preceded by another informational polymer, it stands to reason that this polymer must have been easier to form under prebiotic conditions. Furthermore, the structure of this polymer would have been sufficiently close to RNA that molecular evolution would have allowed it to eventually and continuously evolve into RNA. By replacing phosphate with glyoxylate we have introduced a thermodynamically favorable and plausibly prebiotic ancestor of the phosphodiester linkage of RNA. The glyoxylate group is also remarkably close in structure and electrostatic charge distribution to that of a phosphate group. Here we discuss the additional reasons why *ga*NA is attractive as a prebiotic ancestor of RNA.

4.1. Prospects for *ga*NA polymer formation

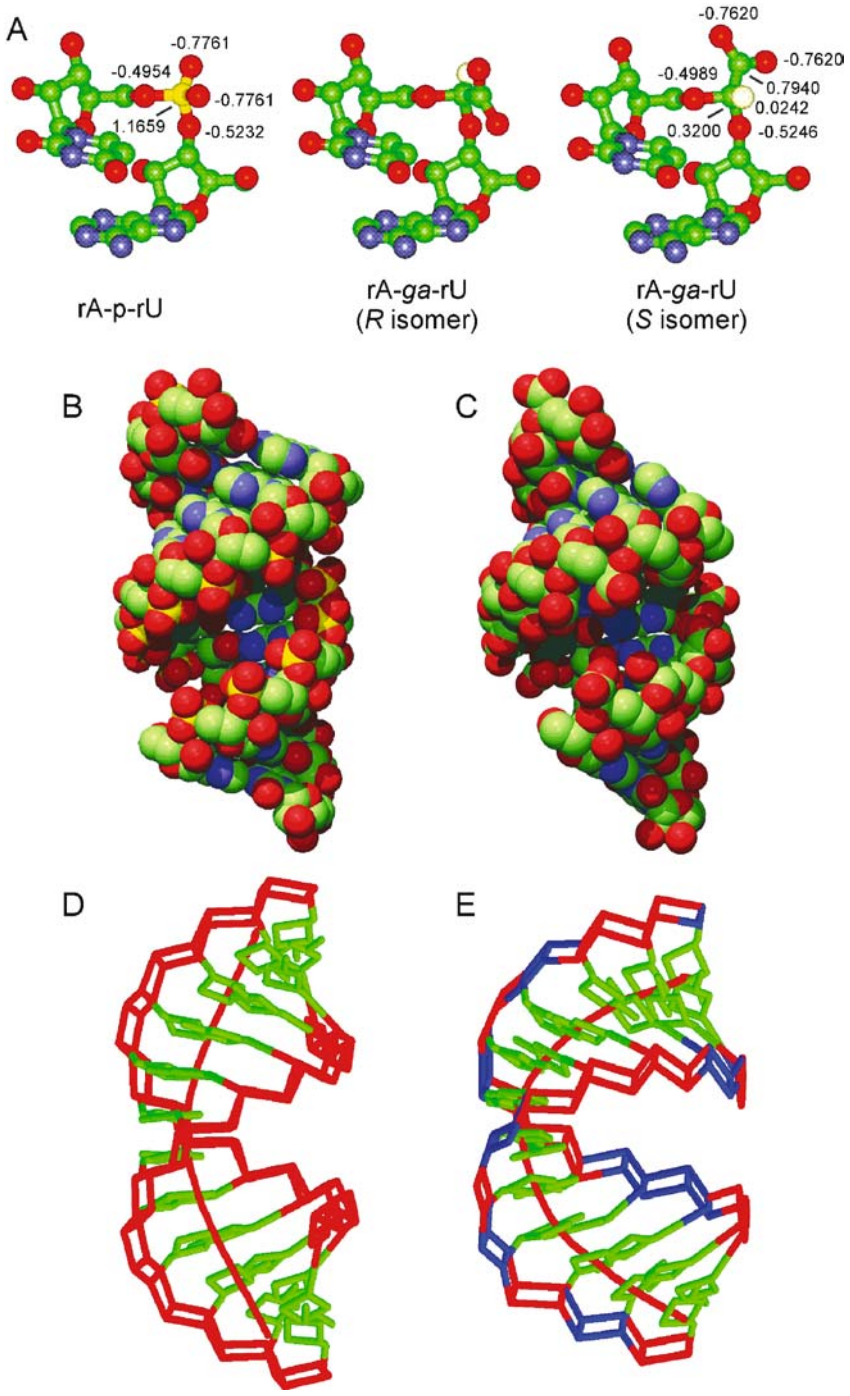
We have shown that drying nucleosides and glyoxylate from an aqueous solution containing divalent metal cations yields approximately 1% *ga*-dinucleotides. While these experiments demonstrate a plausible prebiotic method for the formation of an RNA-like linkage, we have not yet realized the spontaneous formation of *ga*-oligonucleotides. Experiments in which samples containing nucleosides and glyoxylate were subjected to repeated drying-rehydration cycles did not result in increased *ga*-dinucleotide yield or the growth of *ga*-oligonucleotides (data not shown). It appears that, under the conditions of the experiments presented, the observed 1% yield represents the equilibrium amount of *ga*-dinucleotide formed shortly before the point at which one component of the reaction mixture precipitates or crystallizes from solution.

The thermodynamic parameters we have determined for glyoxylate-acetal formation indicate that *ga*-oligonucleotide formation should proceed spontaneously if *ga*-nucleotides are preorganized prior to the point of sample drying when acetal bond formation is catalyzed by cationic metal hydrates. Solution state experiments with glyoxylic acid in H_2O/methanol mixtures revealed the equilibrium constant between a hemiacetal and an acetal to be 1.5, corresponding to a free energy of -0.5 kcal/mol at 65 °C (Methods). Thus, for free *ga*-nucleotides in solution,

formation of acetals between two nucleotides only becomes favored over the *ga*-nucleotide hemiacetal state when nucleotide concentration becomes equal to or greater than that of water. However, if a surface or a macromolecular assembly ordered *ga*-nucleotides during the course of sample drying, then the effective local concentration of ribose OH groups to form acetal linkages would be increased dramatically. Preorganization of *ga*-nucleotides, for example by base stacking on an intercalating molecule (Hud and Anet, 2000), would eliminate the entropic penalty of bringing together two *ga*-nucleotides in solution prior to acetal formation while the entropic gain of releasing a water molecule during acetal formation at each *ga*-dinucleotide step is retained. Additionally, under drying conditions, any water molecules formed upon acetal bond formation could be lost to evaporation, essentially blocking the reverse reaction of acetal hydrolysis and driving the equilibrium towards acetal formation. The sudden addition of water after the acetal has been formed, particularly if the addition takes place at a lower temperature, will quench the reaction. The acetal will not hydrolyze because the hydrolysis rate is then too low, and equilibrium conditions are no longer operative, unless inordinately long times are considered. The possibility of quenching acetal bond hydrolysis by rehydration at a lower temperature than that at which bond formation was catalyzed is illustrated by the fact that *ga*-dinucleotide stability is significantly greater at 22 °C versus 60 °C under otherwise identical solution conditions (Figure 5).

Ferris and co-workers have demonstrated that drying the chemically activated nucleoside 5′-phosphorimidazolide-adenosine (ImpA) onto the surface of a montmorillonite clay leads to the elongation of a pre-absorbed decanucleotide by up to 30 additional nucleotides, whereas the same decanucleotide is not elongated by ImpA in solution or when dried in the absence of the clay surface (Ferris *et al.*, 1996). A similar catalytic inorganic surface may exist that facilitates the polymerization of

Figure 6. Comparison of computer generated models of RNA and *ga*RNA. **(A)** Ball and stick models of dinucleotides connected by a phosphate (RNA), an *R*-glyoxylate-acetal, and an *S*-glyoxylate-acetal linkage, respectively. For clarity, no hydrogen atoms are shown, except the CH hydrogen of glyoxylate. Partial electrostatic atomic charges, based upon quantum mechanical calculation (Materials), are shown for the atoms of the phosphate linkage and the *S*-glyoxylate-acetal linkage. Partial electrostatic atomic charges assigned to the *R*-glyoxylate-acetal linkage were the same as those shown for the *S*-glyoxylate-acetal linkage. **(B)** Space-filling representation of the RNA duplex $[A_6U_6]_2$ in a canonical A-form helix. **(C)** Space-filling representation of an energy minimized model (Materials) of the *ga*RNA duplex with the self-complementary sequence A_6U_6, with a mixture of the *R*- and *S*-glyoxylate-acetal linkages along each strand (Methods). **(D and E)** Vector model of backbone and base pair geometries generated by the Curves program (Lavery and Sklenar, 1989) for the models shown in B and C. For the *ga*RNA model (E), backbone linkages of the *R*-glyoxylate-acetal nucleotides are shown in red, backbone linkages of the *S*-glyoxylate-acetal nucleotides are shown in blue. Note that the helix of *ga*RNA (models C and E) is slightly underwound from the canonical A-form helix (models B and D). Also note that the backbone and base pair vectors of the *ga*RNA duplex indicate that the helical form of *ga*RNA does not vary appreciably between *R*- and *S*-glyoxylate-acetal linkages. Models in D and E are rotated 90° from those shown in B and C.


A
-0.7761
-0.4954
-0.7761
1.1659
-0.5232
rA-p-rU
rA-ga-rU
(R isomer)
-0.7620
-0.7620
-0.4989
0.7940
0.0242
0.3200
-0.5246
rA-ga-rU
(S isomer)
B
C
D
E

ga-nucleotides. We have previously proposed that during prebiotic times planar, organic molecules, similar to those known to intercalate the bases of DNA and RNA, may have organized the nucleoside bases into assemblies that facilitated the formation of RNA-like polymers (Hud and Anet, 2000; Jain *et al.*, 2004). The chemistry we have described here for *ga*NA synthesis would also be compatible with this suggested mode of polymer assembly and replication by "molecular midwives" (Hud and Anet, 2000). Additionally, preorganization of the bases or nucleosides into columnar stacks of base pairs or base tetrads prior to backbone synthesis would protect the bases from iminal bond formation with glyoxylic acid.

4.2. The advantage of reversible backbone linkages in early evolution

Template-directed synthesis using activated nucleoside monophosphates (e.g. 5′-phosphorimidazolides) is highly prone to nucleotide misincorporation and the production of truncation products (Joyce and Orgel, 1999). Such processes with high rates of mutation would have been catastrophic to early life (Joyce and Orgel, 1999). In contemporary life, the proofreading mechanism of polymerase enzymes excises misincorporated nucleotides (i.e. errors) from a product strand that do not form Watson-Crick base pairs with the template strand, providing the polymerase with another opportunity to attach the correct nucleotide. Since proofreading protein polymerases likely appeared after the era of the RNA world, it would have been advantageous for early self-replicating systems to have had some other mechanism by which replication errors were corrected. Lynn and co-workers have shown that backbone linkages with reversible covalent bonds can provide a form of thermodynamically-controlled proofreading during template-directed synthesis (Li *et al.*, 2002). Correct base pairing between a template strand and substrates (e.g. mononucleotides, dinucleotides) is ultimately selected from incorrect pairings based upon the lower energy of a template paired with a product strand containing all Watson-Crick base pairs. With this approach, Lynn and co-workers have achieved both sequence-specific and chain-length-specific reading of DNA templates into product strands using DNA-like mononucleosides with backbone coupling chemistries that form reversible imine bonds, prior to conversion to stable amine linkages upon reduction (Li *et al.*, 2002).

The reversible formation of acetal bonds in an aqueous environment could have likewise provided a means for proofreading during the template-directed synthesis of *ga*-oligonucleotides (Hud and Anet, 2000). Based upon the *ga*-dinucleotide formation and degradation data presented in the current study, glyoxylate-acetal bond formation and cleavage at neutral or basic pH in a drying sample is apparently only catalyzed by divalent cations close to the point at which liquid water is completely removed. Therefore, an existing *ga*NA strand could be long-lived in an aqueous solution containing metal hydrates. As a solution of *ga*NA strands, nucleosides, and glyoxylate progressed towards dryness, inorganic surfaces or intercalating

molecular midwives could have promoted base pairing between free nucleotides and the parent *ga*NA polymer before acetal bond cleavage/formation commenced. If this multimolecular assembly remained intact as liquid water was completely removed, then acetal bond formation could have been catalyzed along the backbone of the daughter strand nucleotides without loss of the parent *ga*NA polymer.

4.3. *ga*NA AND RNA EVOLUTION

A molecule as functionally versatile as RNA has likely undergone more than one stage of molecular evolution. The proposed role of glyoxylate as a precursor to phosphate in modern nucleic acids would, as illustrated above, alleviate some serious difficulties associated with the prebiotic synthesis of RNA. Nevertheless, the origin of nucleosides, including problems associated with ribose synthesis and glycosidic bond formation, remain separate and formidable questions regarding the origin of RNA. Eschenmoser and co-workers have suggested that alternative sugars may have preceded ribose in an ancestor of RNA (Eschenmoser, 1999; Schöning *et al.*, 2000). Orgel and co-workers have even proposed that RNA may have evolved from a polymer containing glycerol-derived acyclonucleosides, a suggestion aimed at alleviating perceived problems associated with RNA synthesis from a racemic mixture of ribonucleotides (Joyce *et al.*, 1987). Our proposal of glyoxylate as an ancestor of phosphate is compatible with these and other proposals for the molecular evolution of RNA, as long as the proposed nucleosides contain at least two alcohol groups for the formation of acetal linkages.

Even if glyoxylate-acetal linkages are ancestors to the phosphodiester linkage, there may have been even earlier linkage molecules than glyoxylate. Glycolaldehyde, a widely accepted prebiotic molecule, would be a possible ancestor to glyoxylate. Just as we have shown that glyoxylate can link nucleosides in a plausible prebiotic reaction, glycolaldehyde and other aldehydes would be expected to form similar linkages under similar reaction conditions. It is even possible that the earliest RNA-like polymers had backbones that contained a mixture of aldehyde linkage groups and a variety of nucleoside sugars (Hud and Anet, 2000) before Darwinian evolution gradually resulted in the appearance of polymers with backbones of uniform structure and chemical composition.

It is important to note the difference between the kinetic stabilities of *ga*-nucleotides and nucleotide monophosphates, and the possible role that this could have played in evolution. The phosphoester bond of a monophosphate nucleotide is kinetically stable to hydrolysis, with a half-life in the absence of catalysis on the order of a thousand years at room temperature (Kirby and Jencks, 1965; O'Brien and Herschlag, 2001). In contrast, the hemiacetal bond of a *ga*-nucleotide is freely labile in the presence of water (at neutral, acidic, or alkaline pH), and a *ga*-nucleotide in aqueous solution will exist in equilibrium with its nucleoside and glyoxylate. Thus, a *ga*-nucleotide could not function in the many roles that are played by nucleotide monophosphates (or polyphosphates) in modern life. For example, the kinetic

stabilities of phosphoester and polyphosphate linkages are of paramount importance for the transfer and utilization of energy within all living cells. Thus, the positive attributes of nucleotide kinetic stability would have been a strong selective pressure that favored replacement of glyoxylate by phosphate.

Acknowledgements

This work was supported by the Exobiology program of the National Aeronautics and Space Administration (NNG04GJ32G). We thank Dr. Leslie Gelbaum, Prof. Loren Williams, and Mr. Derrick Watkins for helpful discussions.

References

Avetisov, V. and Goldanskii, V.: 1996, Mirror Symmetry Breaking at the Molecular Level, *Proc. Natl. Acad. Sci. USA* **93**, 11435–11442.

Bengston, S.: 1994, *Early Life on Earth*, Columbia University, New York.

Benner, S.A.: 2004, Understanding Nucleic Acids Using Synthetic Chemistry, *Acc. Chem. Res.* **37**, 784–797.

Böhler, C., Nielsen, P. E. and Orgel, L. E.: 1995, Template Switching Between PNA and RNA Oligonucleotides, *Nature* **376**, 578-581.

Case, D. A., Darden, T. A., Cheatham, T. E., III, Simmerling, C. L., Wang, J., Duke, R. E., Luo, R., Merz, K. M., Wang, B., Pearlman, D. A. *et al.*: 2004, University of California, San Francisco.

Chaput, J. C., Ichida, J. K. and Szostak, J. W.: 2003, DNA Polymerase-Mediated DNA Synthesis on a TNA Template, *J. Am. Chem. Soc.* **125**, 856–857.

Cheatham, T. E., III and Kollman, P. A.: 1997, Molecular Dynamics Simulations Highlight the Structural Differences among DNA:DNA, RNA:RNA, and DNA:RNA Hybrid Duplexes, *J. Am. Chem. Soc.* **119**, 4805–4825.

Cieplak, P., Cheatham, T. E., III and Kollman, P. A.: 1997, Molecular Dynamics Simulations Find That 3′ Phosphoramidate Modified DNA Duplexes Undergo a B to A Transition and Normal DNA Duplexes an A to B Transition, *J. Am. Chem. Soc.* **119**, 6722–6730.

Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W. and Kollman, P. A.: 1995, A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules, *J. Am. Chem. Soc.* **117**, 5179–5197.

De Proft, F., Langenaeker, W. and Geerlings, P.: 1995, Acidity of Alkyl Substituted Alcohols: Are Alkyl-Groups Electron-Donating or Electron-Withdrawing? *Tetrahedron* **51**, 4021–4032.

Eschenmoser, A.: 1999, Chemical Etiology of Nucleic Acid Structure, *Science* **284**, 2118–2124.

Ferris, J. P., Aubrey R. Hill, J., Liu, R. and Orgel, L. E.: 1996, Synthesis of Long Prebiotic Oligomers on Mineral Surfaces, *Nature* **381**, 59–61.

Fuller, W. D., Sanchez, R. A. and Orgel, L. E.: 1972, Studies in Prebiotic Synthesis. VII Solid-State Synthesis of Purine Nucleosides, *J. Mol. Evol.* **1**, 249–257.

Gesteland, R. and Atkins, J. F. (eds.): 1999, *The RNA World, Second Edition: The Nature of Modern RNA Suggests a Prebiotic RNA World*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor.

Herdewijn, P.: 2001, TNA as a Potential Alternative to Natural Nucleic Acids, *Angew. Chem., Int. Ed. Engl.* **40**, 2249–2251.

Hud, N. V. and Anet, F. A. L.: 2000, Intercalation-Mediated Synthesis and Replication: A New Approach to the Origin of Life, *J. theor. Biol.* **205**, 543–562.

Jain, S. S., Anet, F. A. L., Stahle, C. J. and Hud, N. V.: 2004, Enzymatic Behavior by Intercalating Molecules in a Template-Directed Ligation Reaction, *Angew. Chem. Int. Ed. Engl.* **43**, 2004–2008.

Joyce, G. F.: 2002, The Antiquity of RNA-Based Evolution, *Nature* **418,** 214–221.

Joyce, G. F., Inoue, T. and Orgel, L. E.: 1984, Non-Enzymatic Template-Directed Synthesis on RNA Random Copolymers. Poly(C, U) Templates, *J. Mol. Biol.* **176**, 279–306.

Joyce, G. F. and Orgel, L. E.: 1999, Prospects for Understanding the Origin of the RNA World in *The RNA World*, Gesteland, R. F., Cech, T. R. and Atkins, J. F. (eds), 2nd ed. *Cold Spring Harbor Laboratory Press, Cold Spring Harbor*, pp. 49–77.

Joyce, G. F., Schwartz, A. W., Miller, S. L. and Orgel, L. E.: 1987, The Case for an Ancestral Genetic System Involving Simple Analogues of the Nucleotides, *Proc. Natl. Acad. Sci. USA* **84**, 4398–4402.

Keefe, A. D. and Miller, S. L.: 1995, Are Polyphosphates or Phosphate Esters Prebiotic Reagents? *J. Mol. Evol.* **41**, 693–702.

Kirby, A. J. and Jencks, W. P.: 1965, The Reactivity of Nucleophilic Reagents Toward the p-Nitrophenyl Phosphate Dianion, *J. Am. Chem. Soc.* **87**, 3209–3216.

Lavery, R. and Sklenar, H.: 1989, Defining the Structure of Irregular Nucleic Acids – Conventions and Principles, *J. Biomol. Struct. Dyn.* **6**, 655-667.

Li, X., Zhan, Z.-Y. J., Knipe, R. and Lynn, D. G.: 2002, DNA-Catalyzed Polymerization, *J. Am. Chem. Soc.* **124**, 746–747.

Miller, S. L.: 1997, Peptide Nucleic Acids and Prebiotic Chemistry, *Nat. Struct. Biol.* **4**, 167–169.

O'Brien, P. J. and Herschlag, D.: 2001, Functional Interrelationships in the Alkaline Phosphatase Superfamily: Phosphodiesterase Activity of *Escherichia Coli* Alkaline Phosphatase, *Biochemistry* **40**, 5691–5699.

Orgel, L. E.: 2004, Prebiotic Chemistry and the Origin of the RNA World, *Crit. Rev. Biochem. Mol. Biol.* **39,** 99–123.

Ould-Moulaye, C. B., Dussap, C. G. and Gros, J. B.: 2002, A Consistent Set of Formation Properties of Nucleic Acid Compounds. Nucleosides, Nucleotides and Nucleotide-Phosphates in Aqueous Solution, *Thermochimica Acta* **387**, 1–15.

Piccirilli, J. A.: 1995, Origin of Life. RNA Seeks Its Maker, *Nature* **376**, 548–549.

Rice, J. F. and Gao, X.: 1997, Conformation of Formacetal and 3′-Thioformacetal Nucleotide Linkers and Stability of Their Antisense RNA· DNA Hybrid Duplexes, *Biochemistry* **36**, 399–411.

Schöning, K. U., Scholz, P., Guntha, S., Wu, X., Krishnamurthy, R. and Eschenmoser, A.: 2000, Chemical Etiology of Nucleic Acid Structure: The Alpha-Threofuranosyl-(3′→2′) Oligonucleotide System, *Science* **290**, 1347–1351.

Schwartz, A. W.: 1997, Speculation on the RNA Precursor Problem, *J. theor. Biol.* **187**, 523–527.

Weber, A. L.: 2001, The Sugar Model: Catalysis by Amines and Amino Acid Products, *Origins Life Evol. B.* **31**, 71–86.

Westheimer, F. H.: 1987, Why Nature Chose Phosphates, *Science* **235**, 1173–1178.

Wiberg, K. B., Bader, R. F. W. and Lau, C. D. H.: 1987, Theoretical Analysis of Hydrocarbon Properties. 2. Additivity of Group Properties and the Origin of Strain-Energy, *J. Am. Chem. Soc.* **109**, 1001–1012.

Wiberg, K. B., Morgan, K. M. and Maltz, H.: 1994, Thermochemistry of Carbonyl Reactions. 6. A Study of Hydration Equilibria, *J. Am. Chem. Soc.* **116**, 11067–11077.

Wood, D. J., Hruska, F. E. and Ogilvie, K. K.: 1974, Proton Magnetic Resonance Studies of 2′-Deoxythmidine, Its 3′- and 5′-Monophosphates and 2′-Deoxythymidylyl-(3′,5′)-2′-Deoxythymidine in Aqueous Solution, *Can. J. Chem.* **52**, 3353–3366.