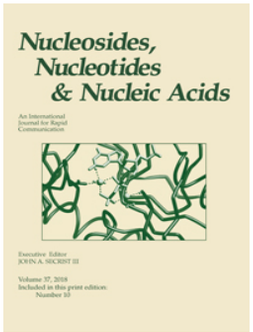


Nucleosides, Nucleotides & Nucleic Acids

ISSN: 1525-7770 (Print) 1532-2335 (Online) Journal homepage: https://www.tandfonline.com/loi/lncn20

Synthesis and base-pairing properties of pyrazine nucleic acids

Maura Gately, Sarah Wong, Jennifer Peoples, Davin Galamay, Guillermo Delgado, Arthur L. Weber & Tammy Campbell

To cite this article: Maura Gately, Sarah Wong, Jennifer Peoples, Davin Galamay, Guillermo Delgado, Arthur L. Weber & Tammy Campbell (2020): Synthesis and base-pairing properties of pyrazine nucleic acids, Nucleosides, Nucleotides & Nucleic Acids, DOI: 10.1080/15257770.2020.1711525

To link to this article: https://doi.org/10.1080/15257770.2020.1711525

Published online: 04 Mar 2020.

NUCLEOSIDES, NUCLEOTIDES AND NUCLEIC ACIDS
https://doi.org/10.1080/15257770.2020.1711525

Synthesis and base-pairing properties of pyrazine nucleic acids

Maura Gately[a]*, Sarah Wong[a]*, Jennifer Peoples[a], Davin Galamay[a], Guillermo Delgado[c], Arthur L. Weber[b], and Tammy Campbell[a]

[a]Department of Chemistry and Biochemistry, California Polytechnic State University, San Luis Obispo, California, USA; [b]SETI Institute, NASA Ames Research Center, Moffett Field, California, USA; [c]Instituto de Química de la Universidad Nacional Autónoma de México, Coyoacán, México

ABSTRACT

The diversity of backbone modifications in the study of primitive informational polymers is partly limited by the plausible formation of their prebiotic starting components. In this paper, we synthesize pyrazine nucleic acid, an acyclic polymer, with the nucleoside derivable from a prebiotic one-pot synthesis containing alanine amide and D-ribose. Pyrazine nucleic acid (PzNA) which has a backbone structurally similar to glycerol nucleic acid (GNA), contain pyrazine derived nucleosides as informational elements that may exhibit base pairing properties similar to the pyrimidines present in RNA.[1] We found that insertion of pyrazinone nucleotides into DNA oligonucleotide sequences is not well-tolerated, and that homogenous sequences of PzNA are unable to form duplexes with RNA or DNA. Reasons for our results may be attributed to the pyrazine-2-one moiety, which is purposed to be a thymine analog, but has a lower pK_a (p$K_a \sim 8.5$) than thymine and uracil. Additionally, we discovered an "apparent" regioselective protection of pyrazine-2-one derivatives in the presence of a secondary hydroxyl group that proved crucial in the preparation of the pyrazine-2-one phosphoramidite. The regioselectivity observed is proposed to be of general interest in the context of heterocyclic chemistry. In the larger context of origins of life studies, it points to the importance of keto-enol preferences of the canonical nucleobases versus pyrazine heterocycles in functioning as recognition elements.

ARTICLE HISTORY
Received 29 August 2019
Accepted 31 December 2019

KEYWORDS
nucleosides; nucleic acids; pyrazinone; origins of life; prebiotic

Introduction

The nucleic acid structures of modern organisms host purines and pyrimidines as coding elements for the continuum of life. However, there are many complexities that afflict the plausible prebiotic formation of both

CONTACT Tammy Campbell tcampb06@calpoly.eduarthur; l.weber@nasa.gov Department of Chemistry and Biochemistry, California Polytechnic State University, 1 Grand Ave, San Luis Obispo, CA 93407, USA.
*These authors contributed equally to this work.
Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/lncn.
Supplemental data for this article is available online at https://doi.org/10.1080/15257770.2020.1711525.

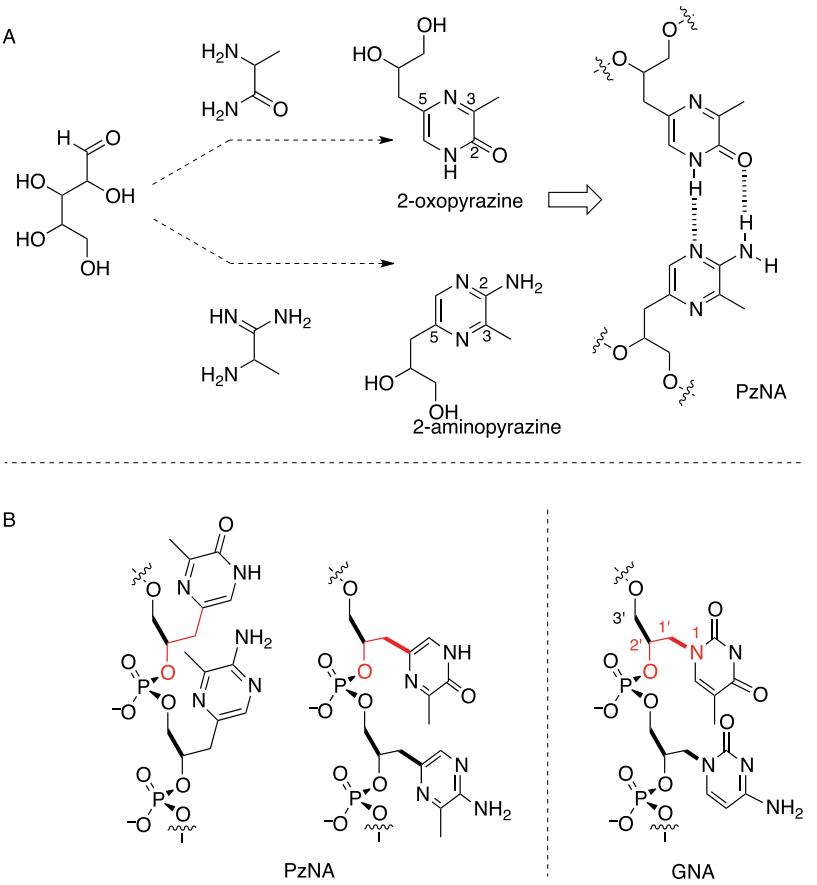


Figure 1. (a) Proposed generation of pyrazine tagged sugar backbones derived from pentose and alanine amide and alanine amidine. (b) Structural comparisons of pyrazine nucleic acid (PzNA) to glycerol nucleic acid (GNA). The red highlights indicate the interaction between the C–O and the C–N (C–C) bonds in GNA (and PzNA). The gauche effect in PzNA restricts the conformation around the C_1'–N_1 bond, while such an effect is absent in PzNA allowing access to both conformations.

RNA and DNA, one of them being the condensation of nucleobases and sugars in an aqueous environment. Broadening the scope in which nucleosides may be formed prebiotically, pyrazine nucleic acid (PzNA) has been proposed by Weber et al. to be a more chemically feasible alternative informational system that could have preceded RNA (Figure 1a).[1] This was based on observations that reaction of glyceraldehyde with alanine amide (or ammonia) gave rise to pyrazinone derivatives, under potentially prebiotic conditions; these pyrazinones could also be, in principle, derived starting from the pentose sugars (Figure 1a).[1]

The structure of PzNA is similar to another oligomer derived from a glycerol backbone known as GNA,[2] however with a major difference (Figure 1b): PzNA contains a *C*-nucleoside (as opposed to a *N*-nucleoside in GNA). The *N*-nucleoside connection in GNA results in a limited

interaction of the (*S*)-GNA with RNA while preventing interactions of (*S*)-GNA with DNA,[3] a strong backbone–base-pair inclination in GNA (due to a conformational restriction caused by the gauche effect between *C*1′-*N*1 and *C*2′–*O*2′) has been suggested to be a reason for this limited cross-pairing behavior.[6b] However, in (*S*)-PzNA, the *C*-nucleoside arrangement is expected to remove this restriction and allow for other orientations and, therefore, for the possibility that (*S*)-PzNA may pair with both RNA and DNA (Figure 1b).

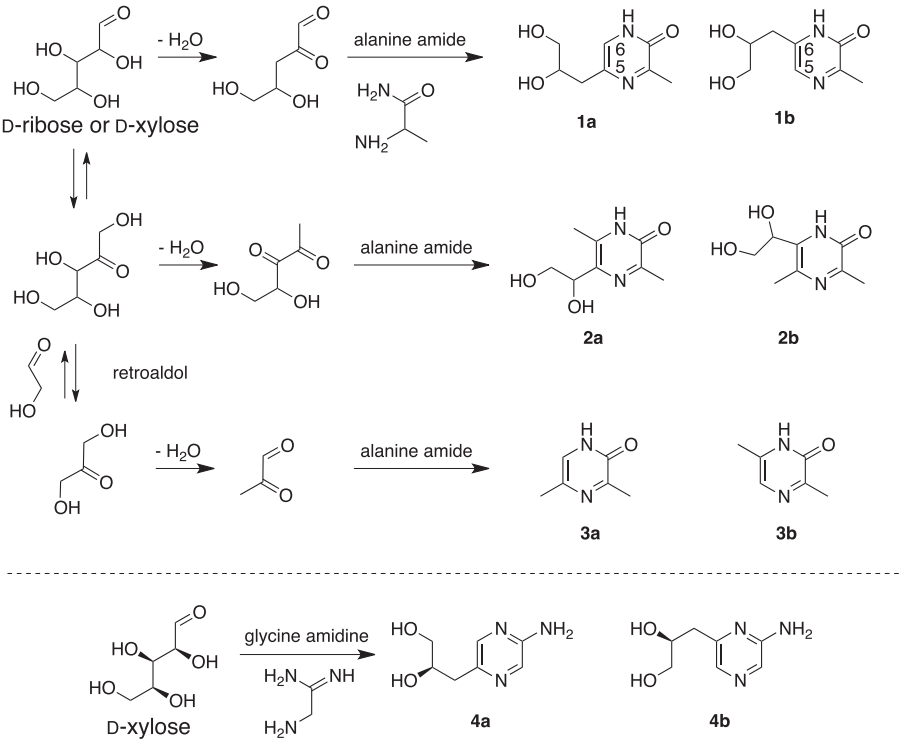
With the aim to synthesize PzNA and investigate its pairing properties, we employed synthetic methods to access the monomeric building blocks required for automated PzNA oligomer synthesis. While the synthesis and efficiencies of synthetic routes toward cyclic pyrazine nucleosides are known,[4] there is still a need to establish an approach toward the formation of acyclic pyrazine nucleosides and their conversion into phosphoramidite monomers that are compatible with automated oligonucleotide synthesis.

In the course of our work, we anticipated that the scale-up of a prebiotically formed nucleoside would prove challenging according to literature, but were also intrigued to discover that the amide portion of the heterocyclic ring (in its iminol form) was found to act as a nucleophilic site in competition with the 2′OH of the propanediol chain. We explored whether this competition could be influenced and established a regioselective protection protocol of a pyrazine amide in the presence of a secondary hydroxyl group (and vice versa); and exploited this selective functionalization of the pyrazine-2-one moiety for the efficient preparation of the required phosphoramidites in the pyrazine-2-one series. We report here on our progress and the selective functionalization that may be of general interest in pyrazine chemistry. We also disclose the synthesis of 2-(1*H*)-pyrazinone phosphoramidites for homogenous, single-insertion and alternating sequence incorporations into oligonucleotides.

Results and discussion

Prebiotic synthesis

Our initial plan for the synthesis of the propanediol-pyrazines (Scheme 1) was based on the potential simplicity provided by the condensation of alanine amide with the keto-aldehyde intermediated (Scheme 1), which in turn was inspired by the potential prebiotic condensation of glyceraldehyde with alanine amide.[1] A retrosynthetic analysis of the desired keto-aldehyde intermediate indicated that we could, in principle, start with any of the D-pentoses, since there is no diastereomeric information necessary (or retained) in the propanediol-pyrazine skeleton (e.g., **3** or **4**).



Scheme 1. Generation of pyrazin-2-ones (**1**-**3**) and pyrazin-2-ones (**4**) starting from aldopentoses (ribose and xylose), alanine amide and glycine amidine.

Initially pyrazinone synthesis was examined by reacting 0.1-0.2 M concentrations of ribose and alanine amide at 65-85 °C in 0.25 M buffers (pH 4.5-8.0) [Buffers: sodium acetate (pH 4.5-5.5), succinate (pH 6.0) and phosphate (pH 6.4-8.0)]. Reverse phase HPLC of the reaction mixtures showed a single strong UV absorbing peak when monitored at 330 nm, the maximum absorbance region for pyrazin-2-ones.[1] The peak area of this UV absorbing product increased as the reaction pH was raised from 4.5 to 7.4, with the largest increase achieved in phosphate buffers above pH 7.4 reacted at 85 °C for 2-4 days. Reverse phase HPTLC analysis of the reaction product resolved the single HPLC peak into two major fluorescent products **1** and **2**. Two minor fluorescent products were present. One of them was identified as 3,5(6)-dimethyl pyrazin-2-one **3a**,**b** by coincident 2 D-TLC with standards synthesized previously;[1] the other minor product was not identified. Product **2** was crystallized after purification by ethanol extraction and column chromatography. The preparation was purified using a gravity flow column packed with BioRad AG-1 (X8) (acetate form) anion exchange resin. LC/MS and proton NMR structural analysis of crystalline material is consistent with 3,6-dimethylpyrazin-2-one 5-ethanediol **2a** or its isomer 3,5-dimethylpyrazin-2-one 6-ethanediol **2b**. As expected, other aldopentoses could be used (in place of ribose) to generate the pyrazinones

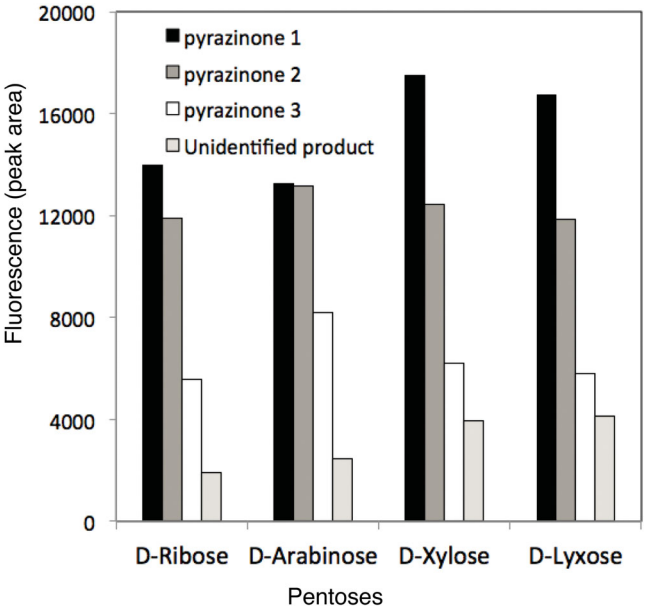


Figure 2. Synthesis of pyrazin-2-ones from each of the four D-aldopentoses from 0.20 m pentose, 0.40 m alanine amide, 0.25 m phosphate buffer, pH 7.5) at 85 °C for 2 days. **1a**/**1b** = pyrazinone 1; **2a**/**2b** = pyrazinone 2; **3a**/**3b** = pyrazinone 3.

(Figure 2); the stereocenters at the C2 and C3 positions of the pentoses were not of consequence since they are lost upon conversion to the 3-deoxyribo-2-ulose.

Further studies aimed at optimizing the yield of product **1** revealed that replacing ribose with xylose, increased the yield of **1** from 38% to 47% of the total for the four TLC-separated fluorescent products.[1] As shown in Figure 3 increasing the reactant concentrations to 2 molal and using a 4 molal phosphate buffer dramatically increased the yield of **1** to 81% of the total for the four fluorescent products. Under these conditions glycerol-3-phosphate as buffer gave higher yields when compared to inorganic phosphate buffer. In addition, as shown in Figure 4, at high reactant concentration the yield and ratio of **1** increased as the pH decreased from 8.5 to 6.4 while those of products **2**-**3** decreased to near zero. Using the high concentration and pH 6.4 conditions, **1b** was obtained and purified by ethyl acetate/isopropyl alcohol extraction. LC/MS and proton NMR structural analysis of this purified **1b** showed it was 3-methylpyrazin-2-one 6-propanediol. The crude mixture contained 6.3% yield of **1** based on alanine amide.[2]

With a view to access the complementary partner of pyrazin-2-one **1**, the 2-amino-pyrazine **4**, the reaction of xylose with glycine amidine was explored under the same conditions used for pyrazin-2-one synthesis (high reactant concentration at pH 6.4). This reaction yielded a single major fluorescent product that was purified using a C-18 solid phase extraction.

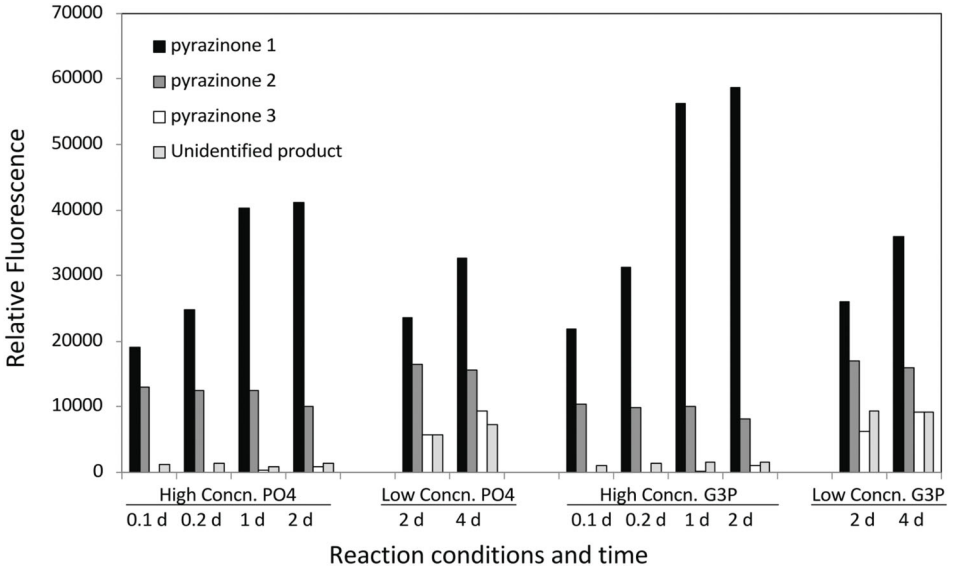


Figure 3. Formation of pyrazin-2-ones at high concentration (2.2 m xylose, 2.0 m alanine amide, 4.0 m phosphate buffer, pH 7.5) and low concentration (0.20 m xylose, 0.20 m alanine amide, 0.40 m phosphate buffer, pH 7.5) at 85 °C. G3P = glycerol-3-phosphate buffer. PO4 = sodium phosphate buffer. **1a**/**1b** = pyrazinone 1; **2a**/**2b** = pyrazinone 2; **3a**/**3b** = pyrazinone 3.

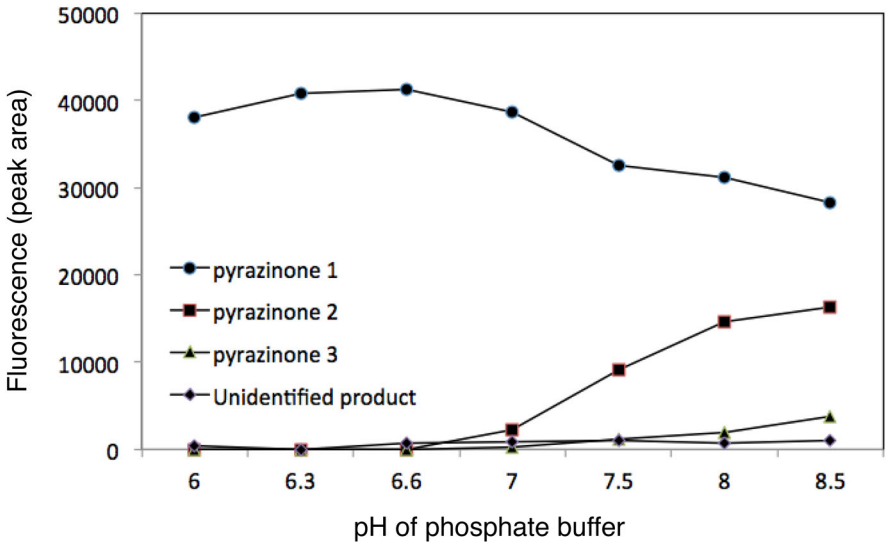
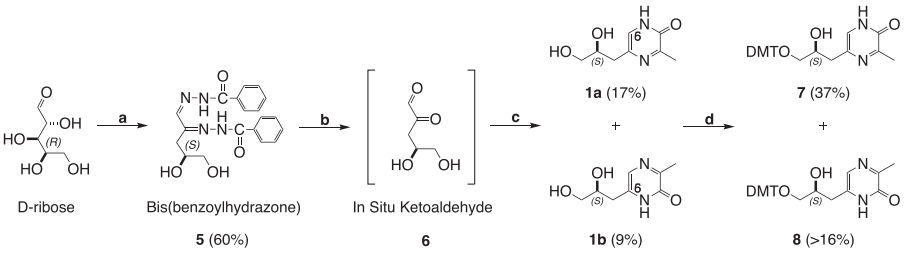


Figure 4. Formation of pyrazin-2-ones as a function of pH at high concentration (2.0 m xylose, 2.0 m alanine amide, 4.0 m phosphate buffers) at 85 °C for 2 days. **1a**/**1b** = pyrazinone 1; **2a**/**2b** = pyrazinone 2; **3a**/**3b** = pyrazinone 3.

LC/MS and proton NMR structural analysis of SPE purified material showed it was 2-aminopyrazine 5(6)-propanediol **4a**,**b**.

As shown in reaction scheme 1, the formation of pyrazin-2-ones **1**-**3**, is thought to proceed via dicarbonyl intermediates that are known to react



Scheme 2. Synthesis of pyrazine-2-one **7** and **8** starting from D-Ribose. *Reagents and conditions:* (a) Benzoylhydrazine, *p*-toluidine, acetic acid; (b) Benzaldehyde, acetic acid, 65 °C, 48 h, EtOH/H_2O (5:9); (c) Alanine amide, NaOH (1 M), H_2O, 0 °C, 3 h; (d) DMTCl, DIPEA, Pyridine.

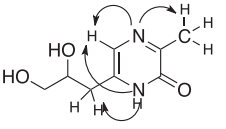
with amino acid amides to give pyrazin-2-ones and 2-aminopyrazines.[3] The formation of **1a**,**b** is thought to occur by direct β-dehydration of the aldopentose to give a 1,2-dicarbonyl intermediate that reacts further with alanine amide. The synthesis of 2-aminopyrazine **4a**,**b** follows a similar course, except that the final condensation with glycine amidine yields 2-aminopyrazine. The formation of **2a**,**b** proceeds by isomerization of the aldopentose to a ketopentose, which then undergoes a loss of water to give a 2,3-dicarbonyl intermediate, which subsequently reacts with alanine amide. **3a**,**b** is rationalized to form by reaction of alanine amide with methylgloxyal, which in turn is formed from the starting pentose by a retroaldol reaction yielding a triose that dehydrates (to the methylglyoxal).

Traditional synthesis

Demonstrating that pyrazine derivatives can be formed by the simple condensation reaction of D-pentoses with alanine amide derivatives, we turned our attention to pyrazine nucleic acid (PzNA). Based on preliminary investigations we came to the conclusion that the simple condensation methods shown above, though convenient, would not provide the large amounts of easily isolatable pyrazine-2-one compounds needed for the preparation of the phosphoramidites for automated synthesis of PzNA. Therefore, we considered a traditional synthetic approach.

Commercially available D-ribose was converted to (bis)benzoylhydrazone **5,** followed by an in situ oxidation to ketoaldeyde **6** (Scheme 2). Cyclocondensation with alanine amide led to the formation of two 2-(1 *H*)-pyrazinone nucleoside regioisomers **7** and **8**.[4–6] The pyrazine nucleoside was then tritylated following a standard protocol to afford regioisomers **7a** and **7b**.

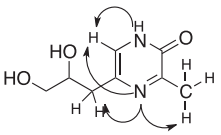
Obtaining a pure compound at this stage gave us the opportunity to know which regioisomer we had in hand using 2 D NMR, since attempts to recrystallize compound **1a** for X-ray analysis were unsuccessful. We removed the trityl group (of the major isomer) since its aromatic signals overlap with the C(6)-H proton of the pyrazinone, and could interfere with



3,6-regioisomer

Two sets of correlation

one set - **TWO** ^{15}N - CH correlation
another set - **TWO** ^{15}N - CH correlation



3,5-regioisomer

Two sets of correlation

one set - **THREE** ^{15}N - CH correlation
another set - **ONE** ^{15}N - CH correlation

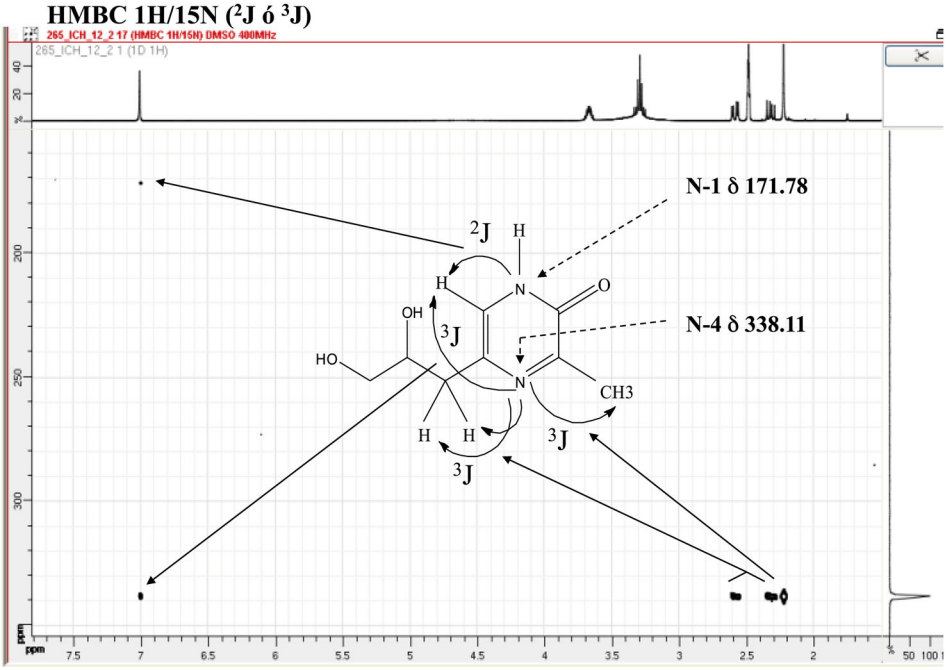
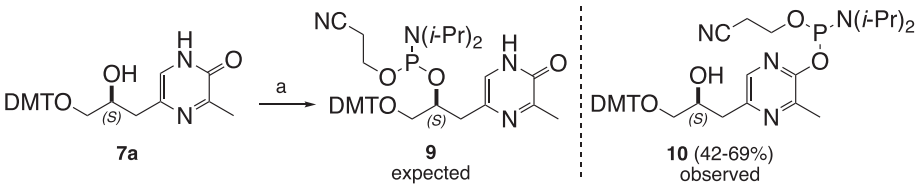


Figure 5. ^{15}N-^{1}H HMBC NMR spectroscopy confirming the 3,5-regioisomer **1a**. The top figures show what is to be expected with the bottom NMR image showing the actual measurement.

any correlation spectroscopy experiment. Accordingly, compound **7** (the major isomer) was detritylated in 3% trifluoroacetic acid in 1,2-dichloroethane, and the regiochemistry of the resulting compound investigated by ^{15}N-^{1}H HMBC NMR spectroscopy. For the 3,5-regioisomer, a correlation between N-4 and H-(C1′), H-(Ar), and H-(CH_3) was expected in conjunction with a correlation between N-1 and H-(Ar). On the contrary, the 3,6-regioisomer was expected to have correlations between N-4 and H-(CH_3) and H-(Ar), and N-1 with H-(Ar) and H-(C1′). The ^{15}N-^{1}H HMBC shown in Figure 5 confirmed the assignment for **1a**, which in turn confirmed the constitution in the tritylated parent compound **7** as the 3,5-regioisomer (Figure 5).

With the tritylated intermediate **7** in hand we proceeded to the phosphitylation step, which turned out to be not trivial. In nucleoside chemistry, unprotected uracil, thymine and guanine (at the C(6)-O position) nucleosides have been phosphitylated successfully (on the sugar hydroxyl)



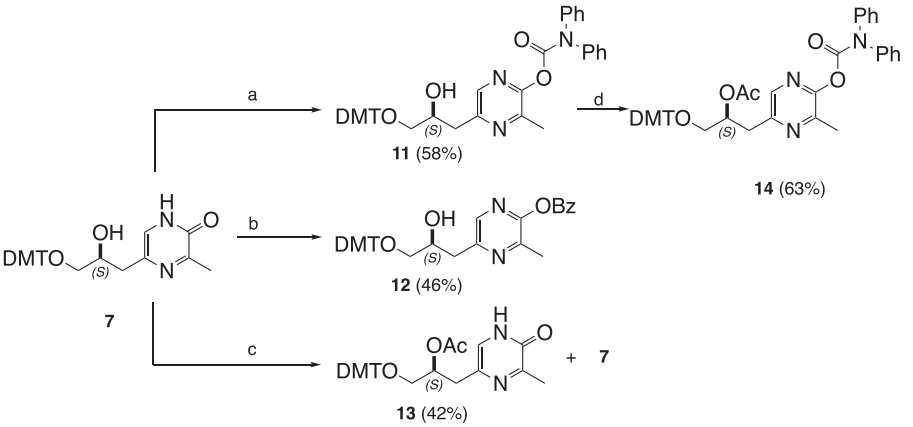
Scheme 3. Phosphitylation of tritylated pyrazinone **7**. *Reagents and conditions* (a) 2-cyanoethyl *N,N,N,N*-tetraisopropyl phosphoroamidite, DCI, DCM, −10 °C to 25 °C.

with minimal interference from the oxo-groups on the nucleobases. Therefore, we followed the same protocols. However, initial efforts toward the preparation of 2′-*O*-phosphoramidite using the 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphine reagent with Hünig's base in methylene chloride were unsuccessful as we observed by ^{31}P NMR (600 MHz) that there were 6 peaks in the crude reaction mixture (two) at 146 and (four) at 150 ppm, indicating the formation of more than one phosphitylated product; this was suggestive of an over-reaction at the pyrazin-2-one moiety apart from the desired reaction at 2′OH position. Use of the less reactive reagent, 2-cyanoethyl-*N,N,N*′,*N*′-tetraisopropyl-phosphorodiamidite under different activation conditions, led to isolation of a product (which we assigned as **9**) that had only two diastereomeric peaks at 146 ppm (^{31}P NMR) with yields ranging from 42–69% (Scheme 3).

This was encouraging, and we proceeded further with this 'material' to synthesize the PzNA oligomers. However we ran into serious difficulties in terms of efficiency on the automated synthesizer (as the compound proved unstable under conditions for automated DNA synthesis) as well as in the step of HPLC purification of the oligomer material produced with this "phosphoramidite". This led us to question whether phosphitylation had occurred at the desired 2′OH, and whether we had misinterpreted the ^{31}P NMR characterization of this compound. The fact that we had peaks at 146 ppm (and not at the usual 150 ppm) came back to trouble us[10]. We conjectured that perhaps phosphitylation of **7** may be occurring preferentially at the 2-oxo- position of the pyrazinone structure (compound **10**) and not at the 2′OH of the propanediol backbone as we had proposed.

If this indeed was the case, we needed to selectively protect the 2-oxo group on the pyrazinone while leaving the 2′OH free for the phosphitylation reaction. To test this possibility, we subjected compound **7** to three different reactions: (a) with *N*,*N*-diphenylcarbamoyl chloride (DPCC), (b) with benzoyl chloride, and (c) with acetic anhydride to see if we could differentiate between the nucleophilicity of the secondary hydroxyl group and the oxo-group of the pyrazinone (Scheme 4).

^{1}H NMR was useful in interpreting the results from these three reactions: a downfield shift of the C(2′)-H from 4.1 ppm to around 5 ppm was indicative of reaction at the 2′OH, while a downfield shift of the C(6)-H



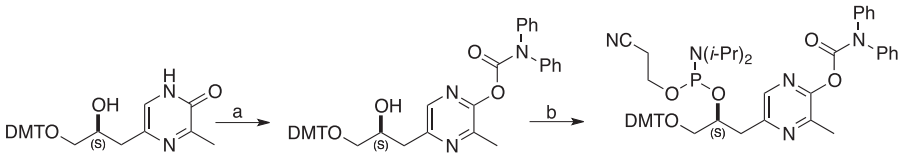
Scheme 4. Exploration of the regioselective protection of **7**. *Reagents and conditions* (a) 1.5 eq. DPCC, DIPEA, Pyridine; (b) 1.5 eq. BzCl, 0.2 eq. DMAP, Pyridine; (c) 1.5 eq. Ac_2O, 0.2 eq. DMAP, Pyridine.

Table 1. Selected ^{1}H NMR shifts ($CDCl_3$, 600 MHz) documenting the chemical shifts that reflect the position of acylation of 2-oxo compared to acylation of the 2′OH of compound **9a**.

Compound	CH_2 (H_1′) ppm	CH (H_2′) ppm	CH_2 (H_3′) ppm	H (C_6) Ar ppm
7	2.7, 2.6	4.1	3.7, 3.1	No peak at 8
11	2.89, 2.92	4.15	3.09, 3.17	*8.01*
12	3.00, 3.05	4.19	3.64, 3.77	*8.18*
13	2.84, 2.89	*5.31*	3.10, 3.27	No peak at 8
14	3.10, 3.11	*5.40*	3.11,3.28	*8.01*

of the pyrazinone from 7.2 to around 8.0 ppm reflected substitution at the 2-oxo- group of the pyrazinone ring (Table 1). Examination of the crude reaction mixture, by ^{1}H NMR, revealed that benzoylation and carbamoylation had occurred selectively at the 2-oxo site of the pyrazinone ring. On the contrary, acetylation was found to occur only at the 2′OH; ^{1}H NMR of the crude showed that only about 50% of 2′-*O*-acetylated product had formed along with unreacted starting material. We confirmed that carbamoylation was occurring at the pyrazin-2-one in the case of **11** by further acetylating **11** and showing that the C(2′)-H in the acetylated derivative **14** had shifted correspondingly downfield to 5.4 ppm.

Based on such control over the regioselectivity of the 2-oxo- of pyrazinone versus the 2′OH group, we were able to re-route our synthesis (Scheme 5). The selective protection at the 2-oxo-position on the pyrazinone ring of **7** by reaction with 3.5 equivalents of diphenylcarbamoyl chloride (DPCC) gave us higher yields (than in Scheme 5) of **11**. Now we could selectively phosphitylate the 2′OH group, under standard conditions, giving good yields of **15**, confirmed by the ^{31}P NMR peaks at 150 ppm (and not at 146 ppm). With the phosphoramidite **15** in hand we are, now, in a position to move forward toward the automated synthesis of PzNA oligomers and investigate their base-pairing properties.



7 **11** (58%) **15** (69%)

Scheme 5. Synthesis of phosphoramidite 15. *Reagents and conditions:* (a) 3.5 eq. DPCC, DIPEA, pyridine; (b) 2-cyanoethyl *N*, *N*, *N*′, *N*′-tetraisopropylphosphorodiamidite, DCI, DCM, −10 °C.

Table 2. Single- and multi-insertions of pyrazine nucleotides into synthesized DNA oligonucleotides.

	Oligonucleotide sequences
Oligonucleotide-**1**	$^{5'}$-(dA)$_5$-*g*PzO-(dA)$_6$-$^{3'}$
Oligonucleotide-**2**	$^{3'}$-(*g*PzO-dA)$_6$-$^{3'}$
Oligonucleotide-**3**	$^{5'}$-(*g*PzO)$_{11}$-dT-$^{3'}$
Oligonucleotide-**4**	$^{5'}$-(dAdT)$_2$-dA*g*PzO-(dAdT)$_3$-$^{3'}$
Oligonucleotide-**5****	$^{5'}$-dT-(*g*PzO)$_{16}$-dT-$^{3'}$

**Mixture of homogenous sequences of lengths 10 to 16 pyrazine units were isolated by HPLC.

Oligonucleotide synthesis

Five oligonucleotide sequences were synthesized. The oligonucleotides **1-5** were synthesized using the conventional phosphoramidite reagents, dA-CE Phosphoramidite, and dT-CE Phosphoramidite (Table 2). Phosphoramidite **15** required 1,2-dichloroethane as the dissolving solvent. The protecting groups were removed from the oligonucleotides during the deprotection procedures by treatment with 28–30% NH_4OH in EtOH (3:1, v/v) at 40 °C for 20 h. PzNA insertion sequences **1-4** were successfully isolated, desalted and characterized by mass spectrometry.

Since the absorption maximum of pyrazinone occurs at 230 and 325 nm (see Supporting Information, Figure S24), we investigated the duplex formation of the oligonucleotides with temperature-dependent UV spectroscopy at 230, 260, and 325 nm. The UV-T_m data of the modified DNA sequences and their potentially complementary strands of DNA, RNA, and self-pairing systems are listed in Table 3 and shown in Figure 6. An equal molar concentration of a homoadenine sequence containing a single insertion of a pyrazine nucleotide and the unmodified complementary strand showed sigmoidal melting curves with melting temperature (T_m) of 33.6 °C (entry 7) and 60.5 °C (entry 10), indicating duplex formation despite the single mismatch in base-pairing and backbone constitution (Figure 6a). However, in comparison to canonical strands, destabilization of the duplex occurs by 8 °C for the dodecamer sequences and 10 °C for the poly-(dT) sequences (entries 8 and 9).

Hyperchromicity was observed, but not in Watson-Crick base-pairing fashion for an alternating self-pairing purine-pyrimidine sequence with

Table 3. UV-T_m data of intra- and inter-system pairing of pyrazine nucleic acid (PzNA), with DNA and RNA.

Entry	Oligonucleotide pairing system	T_m [°C]	Conditions (c μM), Buffer
1	$^{3'}$-(*g*PzO-dA)$_6$-$^{3'}$	36.1 (320 nm)**	10 mM, phosphate[b]
2	$^{3'}$-(*g*PzO-dA)$_6$-$^{3'}$ + $^{5'}$-(dT)$_{12}$-$^{3'}$	37.8 (262 nm)**	5 + 5 mM, phosphate[b]
3	$^{3'}$-(*g*PzO-dA)$_6$-$^{3'}$ + poly-(rU)	——	5 + 5 mM, phosphate[b]
4	$^{3'}$-(*g*PzO-dA)$_6$-$^{3'}$ + poly-(dT)	——	5 + 5 mM, phosphate[b]
5	$^{5'}$-(dAdT)$_6$-$^{3'}$	**29.1**	10 mM, phosphate[b]
6	$^{5'}$-(dA)$_5$-*g*PzO-(dA)$_6$-$^{3'}$ + poly-(rU)	22.0 (260 nm)	5 + 5 mM, phosphate[b]
7	$^{5'}$-(dA)$_5$-*g*PzO-(dA)$_6$- $^{3'}$ + $^{5'}$-(dT)$_{12}$-$^{3'}$	33.6 (260 nm)	5 + 5 mM, phosphate[b]
8	$^{5'}$-d(A)$_{12}$-$^{3'}$ + $^{5'}$-(dT)$_{12}$-$^{3'}$	**41.6 (260 nm)**	5 + 5 mM, phosphate[b]
9	$^{5'}$-(dA)$_{12}$-$^{3'}$ + poly-(dT)	**70.4 (260 nm)**	5 + 5 mM, phosphate[b]
10	$^{5'}$-(dA)$_5$-*g*PzO-(dA)$_6$- $^{3'}$ + poly-(dT)	60.5 (260 nm)	5 + 5 mM, phosphate[b]
11	$^{5'}$-(dAdT)$_2$-dA*g*PzO-(dAdT)$_3$ $^{-3'}$	——	10 mM, phosphate[b]
12	$^{5'}$-(dAdT)$_2$-dAgPzO-(dAdT)$_3$-$^{3'}$	——	10 mM, phosphate[d]
13	$^{5'}$-dT-(*g*PzO)$_{16}$-dT$^{-3'}$ + poly-(rA)	——	2 + 2 mM, phosphate[b]
14	$^{5'}$-dT-(*g*PzO)$_{12}$-dT$^{-3'}$ + poly-(rA)	——	2 + 2 mM, phosphate[f]
15	$^{5'}$-(*g*PzO)$_{11}$-dT$^{-3'}$ + poly-(rA)	——	2 + 2 mM, phosphate[g]

Results for oligonucleotide standards highlighted in bold. [a]gPzO = Pyrazinone nucleic acid with glycol backbone; d = DNA; r = RNA. T_m measurements were made in [b]aqueous phosphate buffer: 10 mM NaH_2PO_4, pH. 7.0 in the presence of 1 M NaCl and 0.1 mM Na_2EDTA; [c]100 mM NaOAc, pH 5.0 in the presence of 1 M NaCl and 0.1 mM Na_2EDTA. [d]aqueous phosphate buffer in the presence of 150 mM NaCl and 0.1 mM Na_2EDTA; [d]aqueous phosphate buffer: 10 mM NaH_2PO_4, pH. 7.0 in the presence of 150 mM $MgCl_2$; [e]aqueous Tris buffer: 10 mM Tris, pH. 8.5 in the presence of 150 mM $MgCl_2$; [f]aqueous phosphate buffer: 10 mM NaH_2PO_4, pH. 6.0 in the presence of 1 M NaCl and 0.1 mM Na_2EDTA; [g]aqueous phosphate buffer: 10 mM NaH_2PO_4, pH. 8.8 in the presence of 1 M NaCl and 0.1 mM Na_2EDTA. Entry 11, previously established result. Estimated error of T_m determination was ± 1.0 °C. —no sigmoidal behavior and UV- T_m detected. **Non-reproducible result.

a single pyrazine insertion in the center (entry 11; Figure 6b). A self-pairing strand containing alternating DNA and PzNA backbones also showed no duplex formation (entry 1; Figure 6a). The homogenous PzNA sequence, oligonucleotide **3**, was sensitive to hydrolysis of the phosphodiester bond perhaps by the intramolecular cyclization starting from the 5' end of the nucleoside-sequence. We modified the future sequence by capping the end with a dT-protecting group (oligonucleotide **5**). Homogenous sequences (entries 13-15) were mixed with complementary DNA and RNA strands under different pH conditions and showed no duplex formation (entries 13-15).

Spectral properties of 2-(1*H*)-pyrazinone nucleoside indicate a tentative pK_a of approximately 8.3 (λ_{max} (nm): λ_{225} to λ_{230}; $\varepsilon_{225} = 8000\,M^{-1}cm^{-1}$). Literature values show 2-hydroxypyrazine to have a pKa of 8.23 (λ_{max} = 316, ε_{316} = 5640; λ_{max} = 227, ε_{227} = 11, 240).[8] The literature values indicate that 2-(1*H*)-pyrazinone might behave similarly to adenine rather than thymine, by tautomerizing to the iminol structure. Our initial supposition was reinforced by a regioselective nucleophilic preferences at the amide position of the heterocycle in the presence of a secondary hydroxyl group. Melting curve studies under slightly acidic conditions of pH 6 in phosphate buffer was expected to show base-pairing for oligonucleotide **4** and oligonucleotide **5** if tautomerization of the amide to an iminol form occurred under neutral to basic pH. However, we did not observe a sigmoidal curve to indicate that tautomerization was the primary reason for why our system did not form duplexes (See SI).

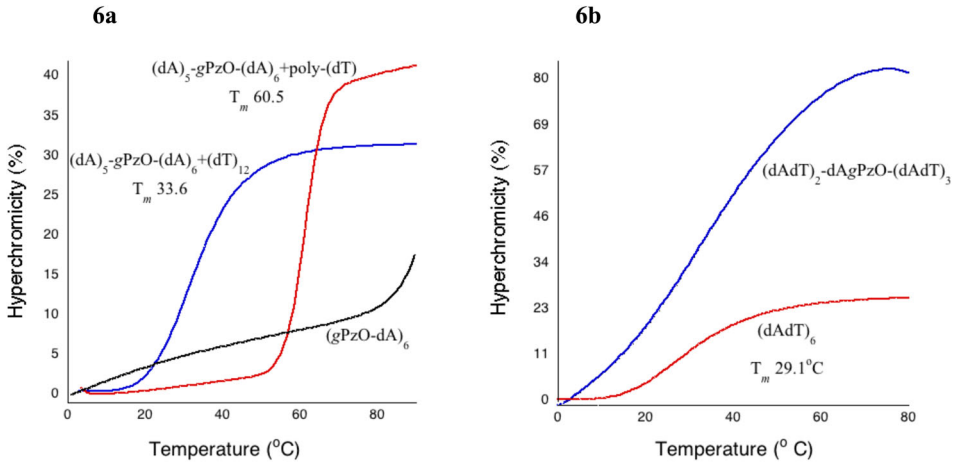


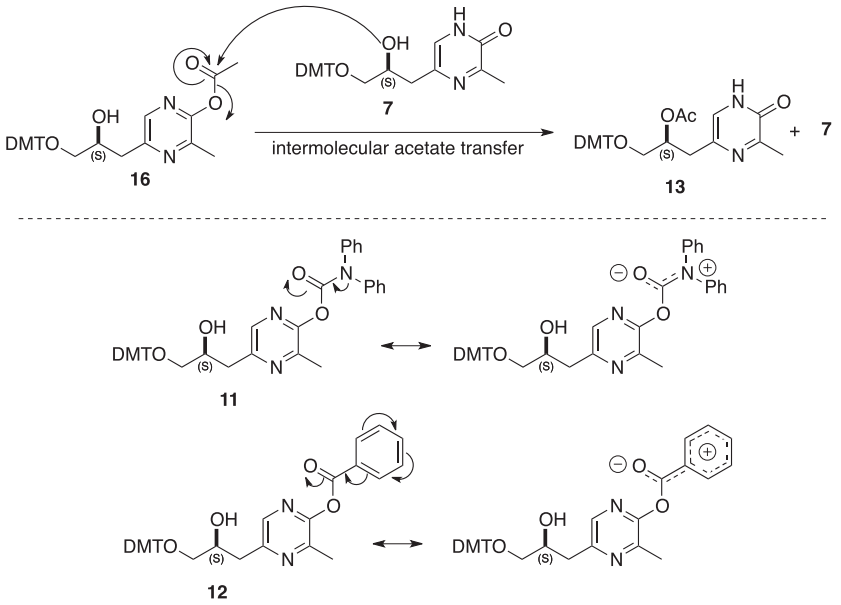
Figure 6. UV melting curves monitored at 260 nm. (a) Pairing was observed for homogenous purine sequence with single **PzO** insertion (($^{5'}$-(dA)$_5$-*g*PzO-(dA)$_6$-$^{3'}$) with complementary DNA sequence (left graph; blue curve = dT; red curve = poly-(dT); table entries 7 and 10). Pairing was not observed for alternating backbone, **6 PzO insertions**, with self or with complementary RNA or DNA sequences (left graph, black curve; table entry 1). (b) Base-pairing was observed between self-pairing alternating purine-pyrimidine sequence $^{5'}$-(dAdT)$_6$-$^{3'}$ (right graph; red curve; entry 5). Watson-Crick base-pairing was not observed for self-pairing alternating purine-pyrimidine sequence with single **PzO** insertion (right graph; blue curve; entry 11).

Discussion

Regioselectivity

The selectivity of the reactions observed in Scheme 4 at first glance is not straightforward. While the reaction of electrophiles at the 2-oxo-moiety of pyrazinones seems to have precedence in the literature[4], all those pyrazinones do not have competing nucleophilic (e.g., hydroxyl) centers, which would interfere (like the 2′OH group in this case). From the point of view of the reactivity of the reagents, it was not straightforward to predict the outcomes: acetic anhydride is supposedly the least reactive of the three acylating reagents used and it still seems to prefer the 2′OH position over the 2-oxo-position of the pyrazinone. For the more reactive acyl chlorides it is vice versa. Thus, the results shown in Scheme 4 pointed to some other mechanism that must be responsible for this dichotomous selectivity.

We hypothesize that acetic anhydride also reacts first at the 2-oxo-position of the pyrazinone to give the acylated-pyrazinone intermediate **16** (similar to benzoyl chloride and diphenylcarbamoyl chloride) (Scheme 6). But this acylated-pyrazinone intermediate **16** (an iminol-acetate) is itself reactive and has the potential to act as an acetylating agent. Thus, **16** could transfer the acetyl group (inter molecularly) to the 2′OH position of another molecule **7** to form the stable acetate (an ester) derivative **13** and regenerate (Scheme 6).[5] Therefore, the apparent "selective" reaction

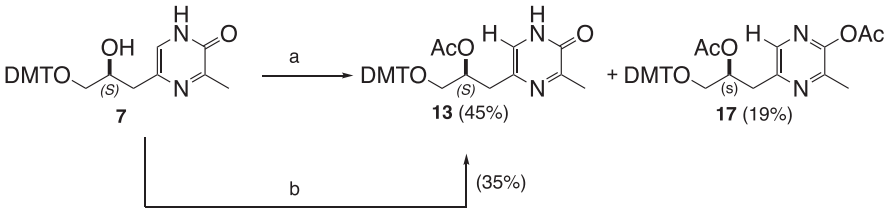


Scheme 6. A possible explanation for the "selective" acetylation of **7** at the 2′OH position.

observed at the 2′OH position of **7** seems to be, in actuality, the consequence of a "relay transfer of acetyl group". This is possible because (a) of formation of a reactive iminol-acetate intermediate **16** and (b) having an internal nucleophile, 2′OH, that under the basic conditions can get acetylated by this iminol-acetate **16**.

This immediately suggested that the benzoyl and diphenylcarbamoyl derivatives at the 2-oxo- position of pyrazinone (**11** and **12**) are, perhaps, less prone to this intermolecular transfer reaction under the reaction conditions, which may be the case: the additional resonance stabilization from the aromatic group (in case of **12**) and the amide resonance (in the case of the carbamoyl group in **11**) render these intermediates more stable and less susceptible to nucleophilic attack (Scheme 6). Such stabilization is absent for the enol-acetate in **16**, which makes it more prone to nucleophilic attack, in this case by the 2′OH group of **7**.[8]

In order to test this, we subjected **7** to 3.5 equivalents of Ac_2O and checked the ^1H NMR of the crude reaction mixture (without work-up), which showed complete conversion to the monoacetylated **13** (at the 2′OH position) and the bisacetylated product **17** (Scheme 7), with none of the starting material **7** remaining (Figure S21). The reaction mixture was subject to aqueous work-up wherein ^1H NMR showed **13** with **17** still present. After column purification, compound **13** was isolated in 45% yield along with **17** in 19% yield (Figure S21 and S22). Such a result has to be contrasted with the reaction of **7** with 3.5 equivalents of diphenylcarbamoyl chloride (Scheme 5) where no 2′OH-acylated product was isolated; rather



Scheme 7. Acetylation of **7**. (a) 3.5 eq. Ac_2O, 0.4 eq. DMAP, Pyridine. (b) 1.5 eq. AcCl, 0.2 eq. DMAP, Pyridine.

58% of acylated-pyrazinone product **11** was isolated (in one case as high as 80%). We also checked whether the nature of the acetylating agent influenced the outcome: reacting **7** with 1.5 equivalent of acetyl chloride (instead of acetic anhydride) afforded compound **13** with none of **17** discernible in the crude NMR (Figure S22). Thus, these series of experiments seemed to reaffirm the hypothesis that reaction is taking place at the 2-oxo-group of the pyrazinone moiety and not at the 2′OH group.

Keto-enol preferences

Preferential reactions at the 2-oxo-group over the hydroxyl group could be directly correlated with the keto-enol equilibrium of the pyrazin-2-one moiety. The UV-spectrum of the parent pyrazin-2-one reveals that the keto form dominates in aqueous solution with a λ_{max} around 221 nm and 317 nm.[6] Our own UV measurements showed an ultraviolet absorption maxima at 230 nm and 325 nm for the 3-methylpyrazinone 5-propanediol (**1a**) in H_2O at a solution pH 7 (Figure S24). Thus our measurements are indicative of the keto form of the pyrazin-2-one predominating in aqueous solution. However, gas phase calculations comparing both theoretical and experimental free energy differences between the enol and amide tautomers of the parent 2-pyrazinone agrees well with an experimental ratio of 14:1 enol to keto form.[7] This is far greater when compared to that for the canonical nucleobases; for e.g., guanine has an estimated 4×10^{-7} population of the enol form.[8] Additionally, the tautomerization of the pyrazine-2-one to the enol form can be correlated with the pK_a of 8.5 for pyrazine-2-ones in aqueous solution,[1] which is relatively lower in comparison to the corresponding pK_a values of canonical nucleobases (pK_a: Uracil = 9.2, Guanine = 9.2)[9]; this may translate to an increase in the amount of the enol form for the pyrazine-2-one (when compared to the canonical nucleobases) and thus may be correlated with the preferential reaction of electrophiles with the 2-oxo-position of pyrazinones in less polar solvents.

The enhanced reactivity observed for the 2-oxo-moiety on the pyrazine ring, when compared with the oxo-positions on the canonical pyrimidine (uracil or thymine) or purine (guanine) is intriguing and can be put into a

larger framework of 2-oxo- and 2-amino-pyrazines functioning as recognition elements in the context of an information system: the keto-enol tautomerization is believed to play a role in the origins of genetic selection and mutation. For example, the overwhelming preference of keto-forms of the guanine, uracil, thymine and cytosine canonical nucleobases is proposed to be at the heart of the strict Watson-Crick pairing and fidelity of replication.[10] The pyrazinone system in the keto form mimics the acceptor-donor hydrogen bonding behavior of uracil and is proposed to pair with adenine.[1] However, if the enol form interferes, then the pyrazinone could also act as a hydrogen bond donor-acceptor and may prefer other interactions (e.g., with thymine, uracil or guanine). Although the proton exchange between N and O is fast, the equilibrium between tautomers is dependent on the polarity of the medium and the ability of the medium to hydrogen bond with each form,[11] which in turn affects the participation of the most prevalent form. In addition, the difference between the pH of a neutral medium (≈7) and the pKa of the parent pyrazin-2-one (≈8.5) is 1.5 units, which falls below the pH-pKa criteria of ≈2 units that is thought to be optimal pairing between complementary bases.[12] While such assessments do cast a doubt as to whether the pyrazinone moiety can function effectively (as an thymine/uracil mimic) in base pairing with the corresponding complementary partner under contemporary conditions, it may not be the case for a primordial informational system if the pH of early prebiotic environments were acidic (pH ≈ 6) as proposed.[13] On the other hand, by the same arguments, the 2-amino-pyrazine ($pK_a \approx 3$)[9a] moiety is expected to function as an adenine analog, and to form base-pairs with complementary thymine and uracil. Future studies are expected to reveal how these factors may affect the base-pairing behavior of PzNA.

Tautomerization of bases in nucleic acids is of great interest owing to the fact that conversion of a normal hydrogen bond donor:acceptor (amide) to an acceptor:donor (iminol) alters the base coding properties of nucleic acids.[11] Thus, tautomerization events in proto-nucleic acids, if left unrepaired, could result in genetic alteration depending upon the pH of the environment and on the identity of the nucleoside inserted opposite the tautomerized base by various primitive polymerization mechanisms.

The gauche effect in backbone conformation

While destabilization of the amide into an iminol functional group may be a reason for why our system does not base-pair, another contributor to duplex destabilization may be the absence of gauche interactions between the nucleobases and backbone. While it may have impaired the ability of an acyclic system such as GNA to base-pair with DNA, the gauche

conformation may be an essential component to proper communication between homogenous (S-GNA:S-GNA) and heterogeneous systems (S-GNA:RNA), as it restricts bond rotation and backbone flexibility.

Acyclic systems with strong backbone inclinations assume inter-strand base stacking as a driving force toward duplex formation.[3b] Additional restrictions placed by gauche interactions may enable methylene groups to be organized in set repetitive conformations adding to the hydrophobic effect.[3a] The lack of a gauche restriction in the PzNA system may eliminate the additional stabilization derived from the methylene groups due to their ability to rotate into other conformations, rendering the backbone "disorganized", while also changing the position of the nucleobase with respect to the backbone.

An unrestricted rotation of the nucleobases with the respect to the backbone may enable either a syn- or anti- positions as is the case with isoGNA.[13] Furthermore, the isoGNA system has nucleobases that are positioned one bond away from the backbone. While not having the additional stabilization of a methylene group, the orientation of the bases (syn- or anti-) and gauche restrictions with respect to the backbone axis enables an organized system capable of base pairing between complementary strands.

The observation that the alternating sequence **2** does not self-pair, suggests that the alternating backbone may exist in a pattern of inclination and non-inclination due to the *C*-nucleoside nature of the PzO insertions. Such a pattern would render the bases unable to pair, since canonical units dA would be planar and alternating with non-canonical units of PzO, which may be inclined along the axis. Even when shifted by one base pair in order to align the planar-planar adenines and inclined-inclined pyrazinones between strands, the result would be non-base pairing between adenines, and very weak, if any, hydrogen-bonding amongst the amide and iminol forms of the pyrazinone.

Conclusions

Our work highlights how a problem inspired by prebiotic chemistry observations[1] was translated to synthetic organic chemistry. In our pursuit we aimed to align the complexities associated with nucleic acid synthesis directly to the emergence of its monomeric units from the prebiotic environment of sugars and amino acids. Our results suggest that PzNA does not have a probable role as a sustainable primitive genetic system due to the potential tautomerization of the amide functional group responsible for stable hydrogen bonding interactions, and the potential disorganization of the backbone. Further work on the spectral and electronic properties of the pyrazine heterocycle will help to assess whether pyrazines in general can be

useful information elements when incorporated into alternative backbones capable of duplex formation with canonical nucleobases.

Experimental section

Abbreviations used

DMTCl: 4,4′-dimethoxytriphenylmethyl chloride; DIPEA: diisopropylethylamine; CEPCl: 2-cyanoethoxydiisopropyl-amino chlorophosphine; DCM: dichloromethane; DMAP: 4-dimethylaminopyridine; Ac_2O: acetic anhydride; BzCl: benzoyl chloride; DPCC: diphenylcarbamoyl chloride.

Reagents and materials

All reagents and solvents were obtained from commercial sources and used without purification. Anhydrous solvents were purchased from *EMD chemicals* or *Alpha Aesar*. D(+)-Ribose was purchased from *AMRESCO*. Benzoylhydrazine was purchased from *Alpha Aesar*. Benzoyl chloride, acetic anhydride and *p*-toluidine were obtained from *Acros*. Hünig's base, diphenylcarbamoyl chloride, benzaldehyde, D(+)-xylose, D(+)-arabinose, and D(+)-lyxose were purchased from *Sigma Aldrich*. Alanine amide was obtained from *AK Scientific Inc* and alanine amide hydrochloride from *Bachem*. Dimethoxytrityl chloride was purchased from *Combi-blocks Inc*. 2-cyanoethyl-*N,N,N′,N′*-tetraisopropylphosphorodiamidite was obtained from *D-L Chiral Chemicals*. Pre-coated flexible silica gel TLC F_{254} plates were obtained from *Whatman Ltd* and Merck HPTLC RP-18W plates were purchased from *VWR*. Flash chromatography was performed using silica gel 60 (40–60 μm) from *Fisher Scientific Co*. ^{1}H and ^{13}C NMR spectra were recorded on a Bruker DRX-600 equipped at 600 MHz for proton and 150 MHz for carbon, respectively. ^{1}H and ^{13}C NMR chemical shifts in $CDCl_3$ were referenced to chloroform at 7.26 ppm and 77.23 ppm, respectively. ^{1}H NMR in D_2O was referenced to HOD at 4.80 ppm. ^{31}P NMR chemical shifts were recorded in ppm relative to an external probe (85% H_3PO_4) referenced at 0.0 ppm. Mass analysis was performed on Agilent ESI-TOF mass spectrometer at an ESI voltage of 4000 V and a flow rate of 200 μL/minute. UV measurements were taken on a Thermo Scientific Nanodrop 2000 C Spectrophotometer. Sample volume was 2 μL. Sample concentrations of compound were at 1 mM and 0.1 mM of compound **8**. Path length was 0.1 cm and wavelengths were at 230 and 325 nm.

Synthesis of pyrazin-2-one phosphoramidite building blocks

3-deoxy-D-glycero-pentos-2-ulose bis(benzoylhydrazone)(5)[7]

D-ribose (10 g, 66.6 mmol) and p-toluidine (5.5 g, 51.3 mmol) in a mixture of 95% ethanol (225 mL) and acetic acid (11 mL) was boiled under reflux for 30 min. To this solution was added benzoylhydrazine (16.5 g, 121 mmol) and the reflux continued for 7 h. The solution was allowed to cool to 25 °C with stirring. The product that separated after 12 h was collected by filtration, washed with methanol (3 × 50 mL), ether (3 × 50 mL), and methanol (3 × 50 mL) and air dried to give a light brown solid. Recrystallization from boiling in ethanol and then cooling to room temperature overnight afforded (**5**) (14.6 g, 60%) as a spongy white solid. ^{1}H-NMR of pure **5** (600 MHz, DMSO-d_6): δ 8.18 (s, 1 H,), 7.9 (m, 4 H, Ar*H*), 7.62 (m, 2 H, Ar*H*), 7.51-7.55 (m, 4 H, CH_2OH), 6.25 (b, 1 H, CHO*H*), 4.97 (b, 1 H, CH$_2$O*H*, 3.9 (b, 1 H, C*H*OH), 3.39 (m, 1 H, CH_2OH), 3.02 (dd, 1 H, CH_2CH), 2.90 (dd, 1 H, CH_2CH). ^{13}C-NMR (150 MHz, DMSO-d_6): 163.4 (C=O), 163.0 (C=O), 154.0 (C=N), 148.2 (C=N), 133.1, 132.0, 128.7, 127.7 (aromatics), 70.70 (CH-OH), 65.00 (CH_2-OH), 29.42 (CH_2), 18.54 157.14 (C(3)-CH_3), 136.73 (C(5)-CH2), 123.67 (C(6)-H), 36.58 (CH_2-Ar), 19.46 (CH_3).

3-methylpyrazin-2-one 5- (or 6-) propanediol (1a and 1 b)

To a 6.23 g (16.9 mmol) solution of bis(benzoylhydrazone) **5**, ethanol (99 mL) and water (162 mL) was added acetic acid (1.02 mL) and benzaldehyde (4.98 mL). The solution was refluxed at 95 °C for 5 h (or at 65 °C for 2 days), cooled to room temperature and filtered. The filtrate was extracted with ethyl acetate (3 × 150 mL) and the aqueous phase collected and placed under reduced pressure to 150 mL (to remove ethanol). The solution was then brought back up to 261 mL with addition of deionized water and the flask cooled between −5 to −10 °C in an ice-salt bath. Alanine amide (1 g, 8.1 mmol) was added and the solution stirred for 15 min, followed by the addition of sodium hydroxide (10.4 g, 261 mmol). The solution was stirred for one hour before being brought back to room temperature for two hours. The solution temperature was brought down to −5 to −10 °C in an ice-salt bath and hydrochloric acid (conc.) (20.9 mL) was slowly added until the solution reached pH ∼6. The solvent was evaporated under vacuum and the residue desalted by dissolving in ethanol and filtering the salt. The mother liquid was evaporated again under vacuum and the desalting procedure repeated two more times. The final residue was purified by short column chromatography (silica gel eluting with 10% methanol in dichloromethane) quickly to avoid decomposition of product to afford 0.50 g of pure **1a** as light yellow oil for a yield of 24% (71a and **1 b** is 17% yield).

TLC of mixture of **1a** + **1b** (SiO_2, 20% methanol in ethyl acetate): R_f 0.26 ^{1}H-NMR of **1a** (600 MHz, D_2O): δ 7.24 (s, 1 H, Ar*H*), 3.98-4.02 (m, 1 H, C*H*OH), 3.69 (dd, $J = 7.86$, 6 Hz, 1 H, CH_2OH), 3.56 (dd, $J = 6.59$, 5.22 Hz, 1 H, CH_2OH), 2.76 (dd, $J = 4.44$, 10 Hz, 1 H, CH_2), 2.63 (dd, $J = 8.94$, 5.55 Hz, 1 H, CH_2, 2.40 (s, 3 H, Ar-CH_3); ^{13}C-NMR (150 MHz, D_2O): 157.58 (C(2)=O), 157.14 (C(3)-CH_3), 136.73 (C(5)-CH2), 123.67 (C(6)-H), 71.14 (CH-OH), 65.22 (CH_2-OH), 36.58 (CH_2-Ar), 19.46 (CH_3). ESI (pos., MeOH); calcd for $C_8H_{12}N_2O_3$ $(M + Na)^+$, *m/z* 207.10; found, *m/z* 207.10 ^{1}H-NMR signals for **7** assigned from the mixture (600 MHz, D_2O): δ 7.32 (s, 1 H, Ar*H*), 3.90-3.96 (m, 1 H, C*H*OH), 3.69 (1 H, CH_2OH), 3.54 (1 H, CH_2OH), 2.79 (dd, $J = 4.44$, 10 Hz, 1 H, CH_2), 2.66 (dd, $J = 8.94$, 5.55 Hz, 1 H, CH_2), 2.38 (s, 3 H, Ar-CH_3).

5-(3′-O-dimethoxytriphenylmethyl)-2-hydroxypropyl-3-methylpyrazin-2(1H)-one (7 and 8)

The crude sample of pyrazinone isomers were azeotroped three times with pyridine and placed under vacuum over night. The isomers were dissolved in anhydrous pyridine (0.44 M) under an argon atmosphere and stirred for 15 minutes at 0 °C. *N*, *N*-diisopropylethylamine (0.7 mL, 3.99 mmol) was added to the solution, followed by 4,4′-Dimethoxytritylchloride. The reaction mixture was warmed to room temperature after 20 min and stirred over night. The reaction mixture was poured over saturated aqeous $NaHCO_3$ (15 mL) and extracted with CH_2Cl_2 (3 × 10 mL). The organic layer was dried over anhydrous Na_2SO_4, filtered, and the filtrate evaporated. The resulting residue was purified over silica gel eluting with 0.05–0.1% methanol in dichloromethane and 1% triethylamine affording isomer **8** (0.16 g, 12%, contaminated with some trityl derivative), mixtures of isomers **7** and **8** (0.05 g, 4%), and pure **7** (0.25 g, 19%). ^{1}H-NMR of **7** (600 MHz, $CDCl_3$): δ 7.35 (m, 2 H, Ar*H*), 7.24 (m, 4 H, Ar*H*), 7.20 (m, 4 H, Ar*H*), 7.12 (t, 7.2 Hz, 1 H, Ar*H*), 6.89 (s, 1 H, Ar*H*), 4.07-4.04 (m, 1 H, C*H*OH), 3.70 (s, 6 H, OCH_3), 3.11 (dd, $J = 5.92$, 3.41 Hz, 1 H, CH_2OH), 3.06 dd (dd, $J = 5.14$, 3.41 Hz, 1 H, CH_2OH), 2.66 (dd, $J = 10.9$, 3.69 Hz, 1 H, CH_2), 2.58 (dd, $J = 8$, 6.46 Hz, 1 H, CH_2), 2.33 (s, 3 H, CH_3); ^{13}C-NMR (150 MHz, $CDCl_3$): 158.6, 157.3 (C(2)=O), 156.9 (C(3)-CH_3), 144.9, 136.1, 133.4,113.24 130.1 (C(5)-CH_2), 128.2, 127.9, 126.9, 122.3 (C(6)-H), 113.24 (aromatic) 86.20 (C), 70.49 (CH-OH), 66.7 (CH_2-OH), 55.32 (OCH_3) 37.04 (CH_2), 20.52 (CH_3). ESI (pos., MeOH); calcd. for $C_{29}H_{30}N_2O_5$ $(M + H)^+$, *m/z* 487.22; found, *m/z* 487.22, (31%). In an alternative (improved) procedure starting from the mixture of (0.487 g, 2.65 mmol) of purified pyrazinone isomers **1a** + **1b** using the same tritylation procedure described above we obtained (0.47 g, 36.5%) of pure compound **7** and (0.443 g, 16.7%) of a mixture of compounds **7** and **8**.

5-(3′-O-dimethoxytriphenylmethyl)-2-hydroxypropyl-3-methylpyrazin-2(1H)-phosphoramidite (10)[7]

To a solution of compound **7** (0.105 g, 0.216 mmol) dissolved in DCM (3 mL) was added 1 mL of a solution containing 2-cyanoethyl *N*, *N*, *N*′,*N*′-tetraisopropylphosphorodiamidite (0.2 mL, 0.618 mmol), dicyanoimidazole (0.036, 0.618 mmol) in CH_2Cl_2 (3 mL) at 0 °C. After 10 min, the solution was brought to room temperature, 3 Å molecular sieves was added, and then shaken. After 1 hour, the reaction mixture was diluted with CH_2Cl_2 (4 mL), poured into saturated aqueous $NaHCO_3$ (6 mL) and followed by extraction with CH_2Cl_2 (3 × 6 mL). The organic layer was dried over anhydrous Na_2SO_4 and evaporated to dryness with the resulting residue purified by chromatography over silica gel eluting with hexanes:ethyl acetate:triethylamine 4:1:0.01 affording a clear oil of compound **10** (0.134 g, 69%). ^{1}H-NMR (600 MHz, $CDCl_3$): δ 7.78 (s, 1 H, Ar*H*), 7.42-7.44 (m, 2 H Ar*H*), 7.31-7.33 (m, 4 H, Ar*H*), 7.27-7.29 (m, 3 H, Ar*H*), 7.19-7.21 (m, 1 H, Ar*H*), 6.81-6.83 (m, 4 H, Ar*H*), 4.16-4.20 (m, 1 H, C*H*OH), 4.07-4.12 (m, 2 H, C*H*$(CH_3)_2$), 3.86 (dd, $J = 4.15$, 6.81 Hz, 1 H, OCH_2), 3.79 (1 H, OCH_2), 3.79 (s, 6 H, OCH_3), 3.19-3.2 (m, 1 H, CH_2CH), 3.10-3.13 (m, 1 H, CH_2CH), 2.91 (dd, $J = 12.37$, 2.83 Hz, 1 H, CH_2CH), 2.84 (dd, $J = 5.86$, 8.45 Hz, CH_2CH), 2.64 (q, $J = 6.60$, 6.26 2 H, CH_2CN), 2.43 (s, 3 H, CH_3), 1.25-1.28 (m, 12 H, CH_3). ^{31}P-NMR (600 MHz, $CDCl_3$): 145.8 (s), 145.7 (s).

5-(3′-O-dimethoxytriphenylmethyl)-2-hydroxypropyl-3-methylpyrazin-2(1H)-diphenylcarbamate (11)

To compound **7** (0.454 g, 0.934 mmol) in dry pyridine (0.1 M, 9 mL) and *N*, *N*-diisopropylethylamine (1.1 mL, 6.54 mmol) at 0 °C was added diphenylcarbamoyl chloride (0.757 g, 3.27 mmol). After 10 min, solution was brought to room temperature and stirred over night. The reaction mixture was poured into saturated aqueous $NaHCO_3$ (15 mL) and extracted with CH_2Cl_2 (3 × 15 mL). The organic layer was dried over anhydrous Na_2SO_4, evaporated in vacuo and the resulting residue purified by silica gel chromatography eluting with 10-24% ethylacetate in hexanes and 1% triethylamine to afford compound **11** (0.62 g, 58%) as a light orange oil. ^{1}H-NMR (600 MHz, $CDCl_3$): δ 7.96 (s, 1 H, Ar*H*), 7.34-6.72 (m, 24 H, Ar*H*), 4.13-4.08 (m, 1 H, CHO*H*), 3.68 (s, 6 H, OCH_3), 3.13 (dd, $J = 5.82$, 3.54 Hz, 1 H, CH_2OH), 3.05 (dd, $J = 5.27$, 4.21, 1 H, CH_2OC), 2.89 (dd, $J = 10.9$, 4.05 Hz, 1 H, CH_2), 2.85 (dd, $J = 8.10$, 6.99 Hz, 1 H, CH_2), 2.37 (s, 3 H); ^{13}C-NMR (150 MHz, $CDCl_3$): 158.6, 156.1 (C(5)), 152.3 (C(2)=O), 151.7 (C=O), 146.4, 145.0, 140.1(C(6)-H), 136.1, 130.2, 129.3, 128.2, 127.9, 126.9, 113.3 (aromatic) 86.26 (C) , 70.43 (CH-OC), 66.76 (CH_2) 55.32 (OCH_3), 37.78 (CH_2), 19.39 (CH_3). ESI (pos., MeOH): calcd for $C_{42}H_{39}N_3O_6$ $(M + Na)^+$, *m/z* 704.27; found, *m/z* 704.27 (88%).

5-(3′-O-dimethoxytriphenylmethyl)-2-hydroxypropyl-3-methylpyrazin-2(1H)-benzoate (12)

To compound **7** (0.195 g, 0.40 mmol) in dry pyridine (11 mL) and *N*, *N*-dimethylaminopyridine (0.015 g, 0.12 mmol) at 0 °C was added benzoyl chloride (0.07 mL, 0.60 mmol). After 10 min, solution was brought to room temperature and stirred over night. The reaction mixture was poured into saturated aqueous $NaHCO_3$ (15 mL) and extracted with CH_2Cl_2 (3 × 15 mL). The organic layer was dried over anhydrous Na_2SO_4, evaporated in vacuo and the resulting residue purified by silica gel chromatography eluting with 10–24% ethylacetate in hexanes and 1% triethylamine to afford compound **12** (0.072 g, 31%) as a light orange oil. ^{1}H-NMR (600 MHz, $CDCl_3$): δ 8.13 (s, 1 H, Ar*H*), 7.66-7.69 (m, 2 H, Ar*H*), 7.52-7.55 (m, 2 H, Ar*H*), 7.45-7.47 (m, 2 H, Ar*H*), 7.34-7.35 (m, 4 H, Ar*H*), 7.28-7.31 (m, 2 H, Ar*H*), 7.20-7.23 (m, 1 H, Ar*H*), 6.83-6.85 (m, 4 H, Ar*H*), *4*.25-4.29 (m, 1 H, CHO*H*), 3.79 (s, 6 H, OCH_3), 3.26 (dd, $J = 5.62$, 3.74 Hz, 1 H, CH_2OC), 3.20 (dd, $J = 4.94$, 4.49, 1 H, CH_2OC), 3.05 (dd, $J = 11.2$, 3.74 Hz, 1 H, CH_2), 2.99 (dd, $J = 8.09$, 6.88 Hz, 1 H, CH_2), 2.48 (s, 3 H); ^{13}C-NMR (150 MHz, $CDCl_3$): 158.6, 152.6 (C(2)=O), 152.2 (C(3)), 149.9 (C = O), 144.9, 145.0, 140.1(C(6)-H), 136.1 (C(5)), 134.3, 130.2, 130.5, 130.2, 128.8. 128.6, 128.2, 127.9, 126.9, 123.3, 113.2 (aromatic) 86.26 (C) , 69.56 (CH-OH), 65.81 (CH_2) 54.67 (OCH_3), 37.19 (CH_2), 18.29 (CH_3). ESI (pos., MeOH): calcd for $C_{36}H_{34}N_2O_6$ $(M + Na)^+$, *m/z* 613.23; found, *m/z* 613.23 (46%).

5-(3′-O-dimethoxytriphenylmethyl)-2-hydroxypropyl-3-methyl-3-methylpyrazin-2(1H)-acetate (13)

To compound **7** (0.121 g, 0.249 mmol) in dry pyridine (7 mL) and *N*,*N*-dimethylaminopyridine (0.006 g, 0.05 mmol) at 0 °C was added acetic anhydride (0.04 mL, 0.37 mmol). After 10 min, solution was brought to room temperature and stirred over night. The reaction mixture was poured into saturated aqueous $NaHCO_3$ (15 mL) and extracted with CH_2Cl_2 (3 × 15 mL). The organic layer was dried over anhydrous Na_2SO_4, evaporated in vacuo and the resulting residue purified by silica gel chromatography eluting with 10–24% ethylacetate in hexanes and 1% triethylamine to afford compound **13** (0.06 g, 42%) as a light orange oil. ^{1}H-NMR (600 MHz, $CDCl_3$): δ 7.35-7.36 (m, 2 H, Ar*H*), 7.19-7.25 (m, 6 H, Ar*H*), 7.09-7.14 (m, 2 H, Ar*H*), 6.83 (s, 1 H, Ar*H*), 6.74-6.75 (m, 4 H, Ar*H*), 5.23–5.26 (m, 1 H, CHO*H*), 3.70 (s, 6 H, OCH_3), 3.20 (dd, $J = 7.1$, 2.90 Hz, 1 H, CH_2OC), 3.03 (dd, $J = 5.25$, 4.79, 1 H, CH_2OC), 2.82 (dd, $J = 8.49$, 5.41 Hz, 1 H, CH_2), 2.77 (dd, $J = 7.72$, 7.1 Hz, 1 H, CH_2), 2.32 (s, 3 H); ^{13}C-NMR (150 MHz, $CDCl_3$): 170.3, 158.1 (C(2)=O), 152.2 (C(3)), 146.9 (C = O), 144.9, 145.0, 140.1(C(6)-H), 139.0 (C(5)), 129.5, 128.7, 127.4,

127.3, 126.6, 122.1, 112.7 (aromatic) 80.97 (C) , 73.45 (CH-OH), 62.60 (CH_2) 54.77 (OCH_3), 33.84 (CH_2), 19.90 (CH_3).

Reaction with 3.5 equivalents acetic anhydride. The above reaction was repeated but with 3.5 equivalents of acetic anhydride in place of 1.5 equivalents. NMRs (a) of crude reaction without work-up, (b) of crude after work up and (c) after column purification were recorded. Work up and isolation was identical as above to yield 45% of **13** and 19% of the bis-acetylated product **17** (See supporting information Figure S22).

Reaction with 1.5 equivalent acetyl chloride. The above reaction was repeated but with 1.5 equivalents of acetyl chloride in place of acetic anhydride. NMRs (a) of crude reaction without work-up and (b) after column purification were recorded. Work up and isolation was identical as above to afford 35% of **14** (with none of bis-acetylated product isolated).

(S)-1-(bis(4-methoxyphenyl)(phenyl)methoxy)-3-(5-((diphenylcarbamoyl)oxy)-6-methylpyrazin-2-yl)propan-2-yl acetate (14)

To compound **11** (0.018 g, 0.026 mmol) in dry pyridine (0.9 mL) and *N*, *N*-dimethylaminopyridine (0.013 g, 0.01 mmol) at 0 °C was added acetic anhydride (0.09 mL, 0.092 mmol). After 10 min, solution was brought to room temperature and stirred over night. The reaction mixture was poured into saturated aqueous $NaHCO_3$ (1.5 mL) and extracted with CH_2Cl_2 (3 × 1.5 mL). The organic layer was dried over anhydrous Na_2SO_4, evaporated in vacuo and the resulting residue purified by silica gel chromatography eluting with 5% ethylacetate in hexanes and 1% triethylamine to afford compound **17** (0.012 g, 63%) as a light orange oil. ^{1}H-NMR (600 MHz, $CDCl_3$): δ 8.01 (s, 1 H, Ar*H*), 7.39-6.79 (m, 26 H, Ar*H*), 5.38-5.42 (m, 1 H, CHO*H*), 3.76 (s, 6 H, OCH_3), 3.27 (dd, $J = 7.17$, 3.07 Hz, 1 H, CH_2OC), 3.12 (m, 2 H, CH_2OC), 3.10 (dd, $J = 4.44$, 4.44 Hz, 1 H, CH_2), 2.41 (s, 3 H).

5-((2S)-3-(bis(4-methoxyphenyl)(phenyl)methoxy)-2-(((2-cyanoethoxy)(diisopropylamino)phosphanyl)oxy)propyl)-3-methylpyrazin-2-yl diphenylcarbamate (15)

To a solution of compound **12** (0.15 g, 0.22 mmol) and dicyanoimidazole (0.039, 0.33 mmol) in CH_2Cl_2 (0.08 M, 2.7 mL) was added 2-cyanoethyl *N*, *N*, *N*′,*N*′-tetraisopropylphosphorodiamidite (0.21 mL, 0.66 mmol) at 0 °C. After 10 min, the solution was brought to room temperature, 3 Å molecular sieves was added, and then shaken. After 1 hour, the reaction mixture was diluted with CH_2Cl_2 (4 mL), poured into saturated aqueous $NaHCO_3$ (6 mL) and followed by extraction with CH_2Cl_2 (3 × 6 mL). The organic

layer was dried over anhydrous Na_2SO_4 and evaporated to dryness with the resulting residue purified by chromatography over silica gel eluting with hexanes:ethyl acetate:triethylamine 4:1:0.01 affording a clear oil of compound **15** (0.134 g, 69%). ^{1}H-NMR of a diastereomeric mixture (600 MHz, $CDCl_3$): δ 8.07 (s, 1 H, Ar*H*), 8.02 (s, 1 H Ar*H*), 7.18-7.46 (m, >24 H impure Ar*H*) 6.80-6.82 (m, 9 H, Ar*H*), 4.39-4.48 (m, 2 H, CHO*H*), 3.76 (m, 12 H, OCH_3) 3.60-3.67 (m, 2 H, OCH_2CH_2CN), 3.51-3.54 (m, 2 H, $CH(CH_3)_2$), 3.43-3.47 (m, 2 H, $CH(CH_3)_2$), 3.35-3.39 (m, 2 H, OCH_2CH_2CN), 3.29-3.32 (dd, $J = 4.48$, 8.96 Hz, 1 H, CH_2), 3.25-3.28 (m, 2 H, OCH_2CH), 3.17-3.20 (m, 1 H, OCH_2CH), 3.14-3.15 (m, 1 H, CH_2), 3.06-3.10 (m, 1 H, CH_2), 2.99-3.03 (dd, $J = 4.83$, 8.09 Hz, 1 H, CH_2) 2.48 (s, 3 H, CH_3), 2.43 (s, 3 H, CH_3), 2.41 (t, $J = 5.94$, 6.92 Hz, 2 H, CH_2CN), 2.38 (t, $J = 5.94$, 7.25 Hz, 2 H, CH_2CN), 1.13 (d, $J = 6.68$ Hz, 3 H, $CHCH_3$), 1.10 (d, $J = 6.48$ Hz, 3 H, $CHCH_3$), 1.02 (d, $J = 6.74$ Hz, 3 H, $CHCH_3$), 1.00 (d, $J = 6.67$ Hz, 3 H, $CHCH_3$); ^{13}C-NMR of a diastereomeric mixture (600 MHz, $CDCl_3$): 158.5 (C(2)=O), 151.9 (2 × C(5)), 145.0 and 146.5 (C = O), 140.3 and 140.6 (C(6)), 136.2 (2 × C(3)), 129.2, 128.2, 127.9, 127.8, 127.7, 126.8, 113.1, (aromatic) 86.10 (C), 73.40 (CH-OH, $J = 15.37$ Hz) and 72.90 (CH-OH, $J = 18.62$ Hz), 65.89 and 66.33 (OCH_2CH), 58.39 (CH_2, $J = 17.8$, 22.7 Hz) 55.28 (OCH_3), 43.23 (CH_2, $J = 9.28$ Hz and 43.14 (CH_2, $J = 9.28$ Hz), 24.5 and 24.7 ($CH(CH_3)_2$), 20.2 (2 × Ar-CH_3), 19.4 (2 × CH_2CN). ^{31}P-NMR (600 MHz, $CDCl_3$): 150.7 (s), 149.9 (s); HRMS (ESI-TOF high-acc.) calcd for $C_{51}H_{56}N_5O_7P$ $(M + H)^+$: 882.3917, found: 882.3997.

Notes

1. Yields determined by measuring and plotting product fluorescence using intensity using NIH's ImageJ image analysis program (see TLC Methods in supporting information).
2. The crude reaction mixture was shown to contain a 6.3 mole percent yield of **1b** (based on alanine amide) by relating TLC fluorescence measurements of the crude product to that of purified **1b** of known concentration (established by NMR).
3. Barlin, G. B. *The Chemistry of Heterocyclic Compounds, The pyrazines*; John Wiley & Sons, Inc: New York, 1982; Vol. 41, pp 28–35, 158–161, 172–175, 213–214.
4. (a) Trehan, S. B.; MacDiarmid, J. E.; Bardos, T. J.; Cheng, Y. Synthesis of 5-ethyl-2-pyrimidinone-2′-deoxyriboside as a new potential antiviral agent. *Med. Chem. Res.* **1991**, *1*, 30-36. (b) Schaaf, K. H.; Spoerri, P. E. Syntheses in the Pyrazine Series: The Proof of the Structure and the Reactions of 2,6-Dibromopyrazine. *J. Am. Chem. Soc.* **1949**, *71*, 2043–2047. (c) Jin, S.; Wessig, P.; Liebscher, J. Intermolecular and Intramolecular Diels–Alder Cycloadditions of 3-Ylidenepiperazine-2,5-diones and 5-Acyloxy-2(1H)-pyrazinones. *J. Org. Chem.* **2001**, *66*, 3984–3997.
5. Such an intermolecular transfer of acetate (and even benzoate) in pyrazinones has been documented in a *metal induced* acylation to prepare ketones: Abe, K.; Nakamura, N.; Sakan, T. *Chemistry Lett.* **1977**, 645–646.
6. (a) Cheeseman, G.W.H. *J. Chem. Soc.* **1960**, 242-247. (b) Dutcher, J.D. *J. Biol. Chem.* **1958**, *232*, 785–796. (c) Mason, S.F. *J. Chem. Soc.* **1957**, 1253–1262. (d) Sato, N. *Science of Synthesis*, **2003**, *16*, 751–759.

7. Gerega, A.; Lapinski, L.; Nowak, M.; Furmanchuk, M; Leszczynski, J. Systematic effect of benzo-annelation on oxo–hydroxy tautomerism of heterocyclic compounds. experimental matrix-isolation and theoretical study. *J. Phys. Chem A* **2007**, *111*, 4934–4943.
8. Jang, Y. H.; Goddard, W. A.; Noyes, K. T.; Sowers, L. C.; Hwang, S.; Chung, D. S. pK_a Values of guanine in water: density functional theory calculations combined with poisson–boltzmann continuum–solvation model. *J. Phys. Chem. B.* **2003**, *107*, 344–357.
9. (a) Shapiro, R. Reactions with purines and pyrimidines. *Ann. N. Y. Acad. Sci.* **1969**, *163*, 624-630. (b) Wilcox, J.L.; Ahlulwalia, A.K.; Bevilacqua, P.C. Charged nucleobases and their potential for RNA catalysis. *Acc. Chem. Res.* **2011**, *44*, 1270–1279.
10. See the discussion (pages 438-445) in Groebke, K.; Hünziker, J.; Fraser, W.; Peng, L.; Diederichsen, U.; Zimmermand, K.; Holzner, A.; Leumann, C.; Eschenmoser, A. *Helv. Chim. Acta* **1998**, *81*, 375–474.
11. Katritzky, A.R.; Hall, C.D.; El-Gendy, B.E; Draghici, B. *J. Comput. Aided Mol. Des.* **2010**, *24*, 475–484.
12. Krishnamurthy, R. *Acc. Chem. Res.* **2012**, *45*, 2035–2044.
13. (a) Walker, J. C. G. *Nature*, **1983**, *302*, 518-520. (b) Russel, M. J.; Hall, A. J. *J. Geol. Soc. Lond.* **1997**, *154*, 377-402. (c) Kua, J.; Bada, J. L. *Orig. Life Evol. Biosph.* **2011**, *41*, 553–558.

Funding

The work was supported jointly by NSF and the NASA Astrobiology Program under the NSF Center for Chemical Evolution, Grant CHE-1504217. Dr. Tammy Campbell thanks Professor Ramanarayanan Krishnamurthy for providing support in whose lab at The Scripps Research Institute (TSRI) the project was initiated and conducted and the NASA Astrobiology Institute for a NASA Postdoctoral Fellowship (under Professor Ramanarayanan Krishnamurthy at TSRI).

ORCID

Tammy Campbell http://orcid.org/0000-0001-8264-0631

References

[1] Weber, A. L. Sugar-Driven Prebiotic Synthesis of 3,5(6)-Dimethylpyrazin-2-One: A Possible Nucleobase of a Primitive Replication Process. *Orig. Life Evol. Biosph.* **2008**, *38*, 279–292. DOI: 10.1007/s11084-008-9141-6.
[2] Zhang, L.; Peritz, A.; Meggers, E. A Simple Glycol Nucleic Acid. *J. Am. Chem. Soc.* **2005**, *127*, 4174–4175. DOI: 10.1021/ja042564z.
[3] (a) Meggers, E.; Zhang, L. Synthesis and Properties of the Simplified Nucleic Acid Glycol Nucleic Acid. *Acc. Chem. Res.* **2010**, *43*, 1092–1102. (b) Hernández-Rodríguez, M.; Xie, J.; Osornio, Y.M.; Krishnamurthy, R. Mapping the Landscape of Potentially Primoridial Informational Oligomers. *Chem. Asian J.* **2011**, *6*, 1252–1262. DOI: 10.1002/asia.201000828.
[4] (a) Mehta, V.P.; Sharma, A.; Van Hecke, K.; Van Meervelt, L; Van der Eycken, E. A Novel and Versatile Entry to Asymmetrically Substituted Pyrazines. *J. Org. Chem.*

2008, *73*, 2382–2388. (b) Jeong, J.U.; Sutton, S.C.; Kim, S.; Fuchs, P.L. Biomimetic total syntheses of (+)-cephalostatin 7, (+)-Cephalostatin 12, and (+)-Ritterazine K. *J. Am. Chem. Soc.* **1995**, 117, 10157–10161. (c) Helliwell, M.; Yun, Y.; Joule, J.A. Surprising orientation in ring synthesis of 3,5-dimethylpyrazin-2(1H)-one. Acta Cryst. **2006**, *E62*, 0955–o966. DOI: 10.1021/jo702656v.

[5] (a) El-khadem, H.; Horton, D.; Meshreki, M.; Nashed, M.A. New Route for the Synthesis of 3-Deoxyaldos-2-uloses. *Carbohydr. Res.* **1971**, *17*, 183–192. (b) Madson, M.A.; Feather, M.S. An Improved Preparation of 3-Deoxy-d-erythro-hexos-2-ulose via the Bis(benzoylhydrazone) and Some Related Constitutional Studies. *Carbohydr. Res.* **1981**, *94*, 183–191.

[6] (a) Bayne S. Aldosuloses (Osones). *Carbohydr. Chem.* **1963**, *2*, 421–423. (b) Karmas, G.; Spoerri, P. E. The Preparation of Hydroxypyrazines and Derived Chloropyrazines. *J. Am. Chem. Soc.* **1952**, *74*, 1580–1584. DOI: 10.1021/ja01126a070.

[7] The location of the phosphoramidite on either the oxygen or the nitrogen atom was unresolved by ^{1}H and ^{31}P NMR. We could only conclude that phosphitylation did not occur at the secondary hydroxyl group of the propane-diol chain though data indicated in Table 1 and the occurrence of the phosphorus peak at 146 ppm instead of the expected value at 150 ppm.

[8] (a) Cheeseman, G.W.H. The Synthesis and Tautomerism of Some 2-Substituted Pyrazines. *J. Chem. Soc.* **1960**, 242–247. (b) Dutcher, J.D. Aspergillic acid; an antibiotic substance produced by Aspergillus flavus. *J. Biol. Chem.* **1958**, *232*, 785–796. (c) Mason, S.F. The Electronic Spectra of N-Heteroaromatic System. Part II. Substituted Monocyclic Azones. *J. Chem. Soc.* **1957**, 1253–1262. (d) Sato, N. *Science of Synthesis*, **2003**, *16*, 751–759.

[9] See the discussion (pages 438–445) in Groebke, K.; Hünziker, J.; Fraser, W.; Peng, L.; Diederichsen, U.; Zimmermann, K.; Holzner, A.; Leumann, C.; Eschenmoser, A. Warum Pentose-und nicht Hexose-Nucleinsäuren??. Teil V. (Purin-Purin)-Basenpaarung in der homo-DNS-Reihe: Guanin, Isoguanin, 2, 6-Diaminopurin und Xanthin. *Helv. Chim. Acta* **1998**, *81*, 375–474. DOI: 10.1002/hlca.19980810302.

[10] Katritzky, A. R.; Hall, C. D.; El-Gendy, B. E.; Draghici, B. Tautomerism in Drug Discovery. *J. Comput. Aided Mol. Des.* **2010**, *24*, 475–484. DOI: 10.1007/s10822-010-9359-z.

[11] Krishnamurthy, R. Role of pK(a) of Nucleobases in the Origins of Chemical Evolution. *Acc. Chem. Res.* **2012**, *45*, 2035–2044. DOI: 10.1021/ar200262x.

[12] (a) Walker, J. C. G. Possible Limits on the Composition of the Archaean Ocean. *Nature*, **1983**, *302*, 518–520. (b) Russel, M. J.; Hall, A. J. The emergence of life from iron monosulphide bubbles at a submarine hydrothermal redox and pH front *J. Geol. Soc. Lond.* **1997**, *154*, 377–402. (c) Kua, J.; Bada, J. L. *Orig. Life Evol. Biosph.* **2011**, *41*, 553–558.

[13] Karri, P.; Punna, V.; Kim, K.; Krishnamurthy, R. Base-Pairing Properties of a Structural Isomer of Glycerol Nucleic Acid. Primordial ocean chemistry and its compatibility with the RNA world. *Angew. Chem. Int. Ed.* **2013**, *52*, 5840–5844. DOI: 10.1002/anie.201300795.