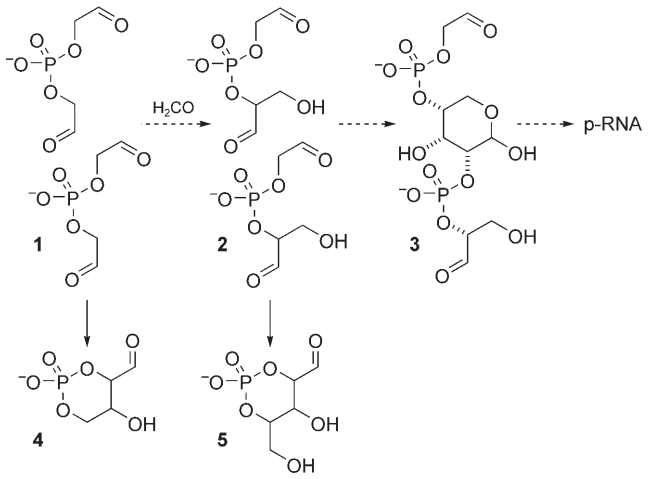
DOI: 10.1002/cbic.200500267

Aldolisation of Bis(glycolaldehyde) Phosphate and Formaldehyde

James M. Smith and John D. Sutherland*[a]

In the seminal work of Eschenmoser and co-workers on pyranosyl-RNA (pRNA),[1] it was pointed out that the backbone of this potentially prebiotic nucleic acid can be formally derived from bis(glycolaldehyde) phosphate **1** and formaldehyde by aldolisation (Scheme 1).[1,2] This derivation, involving hypothetical and formaldehyde to the *p*-RNA backbone **3**.

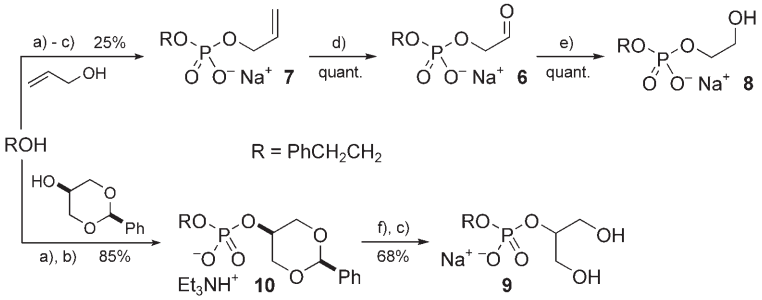


Scheme 1. Constitutional relationship of bis(glycolaldehyde) phosphate (**1**)

[a] *J. M. Smith, Prof. Dr. J. D. Sutherland*
School of Chemistry, The University of Manchester
Oxford Road, Manchester, M13 9PL (UK)
Fax: (+44) 161-275-4939
E-mail: john.sutherland@manchester.ac.uk

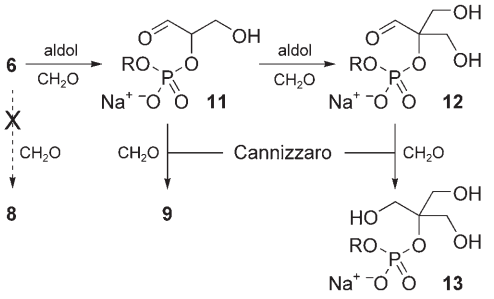
Supporting information for this article is available on the WWW under http://www.chembiochem.org or from the author.

hydroxymethylation to give **2** followed by oligomerisation to the ribose-2,4-phosphate backbone **3**, closely paralleled the documented conversion of glycolaldehyde phosphate and formaldehyde to ribose-2,4-diphosphate.[3] In considering an experimental investigation of this aldol oligomerisation, however, it was evident from the outset that intramolecular aldolisation would be a major competing reaction—indeed the cyclisation of **1** to a mixture of the tetrose-2,4-cyclic phosphates **4** had been shown to occur very readily.[2,4] We recently discovered that **2**, as prepared by conventional synthesis, does not undergo intermolecular aldolisation to **3** to a detectable extent, but undergoes easy cyclisation, and, at pH 9.5, ribose- and xylose-2,4-cyclic phosphates **5** were both obtained, each in 30% yield.[5] The weakly alkaline conditions required to effect this cyclisation contrast with the harsh conditions required to induce aldolisation in the phosphate monoester series.[3] Although the oligomerisation of **2** could not be demonstrated, the mild and stereoselective formation of the phosphorylated ribose and xylose derivatives **5** by intramolecular aldolisation of **2** appeared to us to be etiologically relevant. Since the diester **1** has been shown to be prebiotically plausible, we then wondered if it would be possible to demonstrate the formation of ribose- and xylose-2,4-cyclic phosphates **5** from **1** in one reaction sequence.[2] However, the great ease with which the cyclisations of **1** and **2** proceed suggested that the intermolecular reaction between **1** and formaldehyde might only be possible at high formaldehyde:**1** ratios. We were concerned that the excess of formaldehyde thought necessary to drive the formation of **2** kinetically might also result in dihydroxymethylation of **1** and competing Cannizzaro chemistry, so we first focused on establishing conditions for intermolecular aldolisation in a model system. The mixed phosphate diester **6** was selected as a model substrate for these studies (Scheme 2); the phenylethyl group providing a hydrophobic



Scheme 2. Synthesis of model substrate **6** and standards **8** and **9**. a) $POCl_3$, Et_3N, THF; b) H_2O, THF, Et_3N; c) Na^+-*Dowex*-50®, H_2O; d) O_3/O_2, then Me_2S, MeOH, −78°; e) $NaBH_4$, H_2O; f) H^+-*Dowex*-50®, H_2O.

and chromophoric handle for the synthesis of **6** and certain product standards.[6] The allyl phosphate diester **7** was prepared by standard phosphorus(v) coupling, and then subjected to ozonolysis to liberate the sensitive aldehyde group of **6**.[7] Standard **8** was prepared by reduction of **6**, whilst **9** was produced by hydrolysis of the protected diester **10**. With these standards in hand, we proceeded to investigate the behaviour of **6** (Scheme 3). At pH 9.5 and at the concentrations used, we observed no self-aldolisation of **6** in the absence of formaldehyde, but were gratified to see smooth hydroxymethylation in



Scheme 3. Model study of the reaction of a glycolaldehyde phosphate diester **6** with formaldehyde.

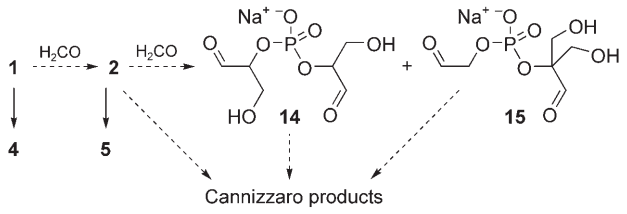
its presence (Table 1). The yield of the monohydroxymethylated product **11** was highest in the experiment with 1 equivalent of formaldehyde. Using 2 or 3 equivalents of formaldehyde resulted in the formation of the dihydroxymethylated product **12**, and decreased the yield of **11**. When the ratio of formaldehyde:**6** was increased to 5, Cannizzaro chemistry was also observed, but interestingly only the formaldehyde adducts **11** and **12** were seen to be reduced—to **9** and **13**, respectively. No ^1H NMR signals for **8**, the reduction product of **6**, were apparent in samples in which subsequent spiking with **8** revealed that its production from **6** would have been readily detectable. The Cannizzaro chemistry was more extensive at higher pH values (data not shown), and it is possible that it might have been caused in the experiments described here by locally elevated hydroxide concentrations during the pH adjustment process. This caveat notwithstanding, the lack of reduction of **6**, given the reduction of **11** and **12**, is somewhat surprising. It is possible that **11** and **12** are reduced intramolecularly, via mixed hemiacetals with formaldehyde, in a process akin to the aldol-Tishchenko reaction (see Supporting Information).[8] The formate esters that would result from such a process would be rapidly hydrolysed under the conditions of the reaction.[4]

Table 1. Aldolisation of **6** in the presence of varying amounts of formaldehyde.

Equiv CH_2O	**6** [%]	**8** [%]	**11** [%]	**9** [%]	**12** [%]	**13** [%]
1	15[a]	0	80	0	5	0
2	15	0	65	0	20	0
3	0	0	50	0	50	0
5	5	0	10	20	50	15

[a] Yields are based on integration of ^1H NMR spectra from individual experiments, and are given to the nearest 5%.

The behaviour of **6** did not bode well for the attempted conversion of **1** to the ribose and xylose phosphates **5** (Scheme 4). For successful production of **5**, the hydroxymethylation of **1** would have, at least, to compete with its cyclisation



Scheme 4. By-products and competing reactions anticipated for the aldolisation of **1** in the presence of formaldehyde.

to **4**, and yet the cyclisation of any **2** produced would be expected to be slower than the cyclisation of **1** due to steric hindrance. Further hydroxymethylation of **2** to give **14** would presumably be similar to the hydroxymethylation of **1** on steric grounds, and the production of **12** in the model series suggested that the alternate dihydroxymethylation of **2** to give **15** would also be possible (Scheme 4). The probability of Cannizzaro reactions further increased the number of likely by-products. This analysis suggested that our only hope of demonstrating the production of **5** in significant amounts was by using 1 equivalent of formaldehyde, and only then if the intrinsic rate constant for the hydroxymethylation of **1** turned out to be sufficiently high to allow this reaction mode to compete with cyclisation of **1** at low formaldehyde concentrations. As a control for the intermolecular reaction of **1** with formaldehyde, we carried out a reaction at pH 9.5 in the absence of formaldehyde, and observed efficient production of the tetrose-2,4-cyclic phosphates **4**, as previously described.[2,4] Both products exist as hydrates (**h**) in solution in D_2O, and were easily identified by ^{1}H NMR. Lack of by-products and signal overlap enabled the relative amounts of the two diastereoisomers to be determined, and the *threo*-isomer was found to predominate (*threo*-**4** 63%; *erythro*-**4** 37%). We then carried out a reaction in the presence of 1 equivalent of formaldehyde. The ^{1}H NMR spectrum was more complicated, and again we observed the production of **4**, but we could also detect signals that we thought were due to the hydrated aldehyde forms of *ribo*-**5** and *xylo*-**5** (Figure 1). Sample spiking with authentic standards of both pentose derivatives[5] confirmed these assignments. In D_2O, *xylo*-**5** exists exclusively in the hydrated aldehyde form, but *ribo*-**5** also exists in pyranose forms (*ribo*-**5 h**, 78–80%; *ribo*-**5 β-p** 18–20%; *ribo*-**5 α-p** 0–4%).[5] This structural complexity, and the presence of unidentified materials amongst the products of the reaction of **1** and formaldehyde, makes accurate quantitation difficult, but the relative abundance of identifiable tetrose and pentose derivatives is *threo*-**4** $\gg$ *erythro*-**4** $\sim$ *ribo*-**5** $>$ *xylo*-**5**. The absolute combined yield of these four products is of the order of 70–80%, and only one unknown (signal at $\delta =$ 5.17) contributes $>$10%. Given the results of the model study and the number of by-products possible, the diastereoselective production of *ribo*-**5** at the same (high) level as *erythro*-**4** is remarkable. Given the chemistry of threose nucleic acid (TNA) recently described by Eschenmoser and co-workers however, the predominance of *threo*-**4** is possibly of greater etiological relevance.[9]

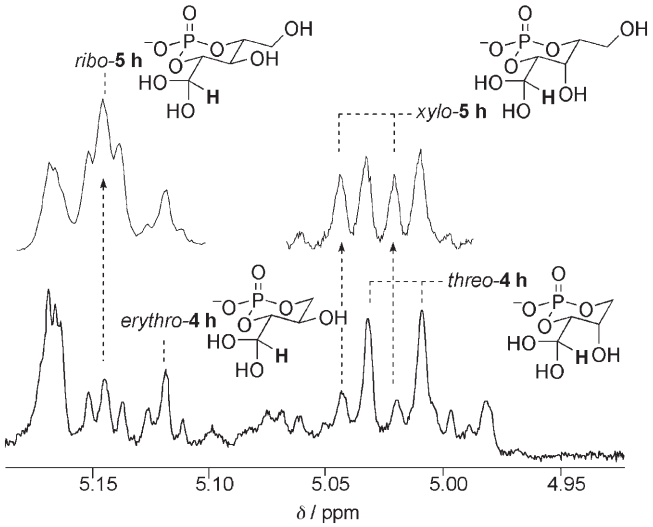


Figure 1. ^{1}H NMR analysis of the aldolisation of **1** and 1 equiv of formaldehyde. Expansion of the region of the spectrum showing signals due to aldehyde hydrates with spectra from spiked samples (inset).

Acknowledgements

This work was funded by the Engineering and Physical Sciences Research Council. We thank Prof. A. Eschenmoser (ETH-Zürich and Scripps) and Prof. S. Pitsch (EPF-Lausanne) for helpful discussions.

Keywords: aldol reaction · Cannizzaro reaction · cyclization · nucleic acids · phosphates

[1] S. Pitsch, S. Wendeborn, R. Krishnamurthy, A. Holzner, M. Minton, M. Bolli, C. Miculka, N. Windhab, R. Micura, M. Stanek, B. Jaun, A. Eschenmoser, *Helv. Chim. Acta* **2003**, *86*, 4270.
[2] S. Pitsch, E. Pombo-Villar, A. Eschenmoser, *Helv. Chim. Acta* **1994**, *77*, 2251.
[3] D. Müller, S. Pitsch, A. Kittaka, E. Wagner, C. E. Wintner, A. Eschenmoser, *Helv. Chim. Acta* **1990**, *73*, 1410.
[4] S. D. Cook, J. D. Sutherland, *Tetrahedron Lett.* **1996**, *37*, 5779.
[5] J. M. Smith, V. Borsenberger, J. Raftery, J. D. Sutherland, *Chem. Biodiversity* **2004**, *1*, 1418.
[6] Experimental procedures and characterisation data for the standards and the aldolisation experiments are given in the Supporting Information.
[7] The powerful electron-withdrawing effect of the phosphate diester group causes the aldehyde group of **6** to be extensively (according to ^{1}H NMR analysis $>$95%) hydrated in aqueous solution. Similar behaviour is shown by other glycolaldehyde phosphate derivatives, such as **1**.
[8] F. J. Villiani, F. F. Nord, *J. Am. Chem. Soc.* **1947**, *69*, 2605.
[9] K.-U. Schöning, P. Scholz, S. Guntha, X. Wu, R. Krishnamurthy, A. Eschenmoser, *Science* **2000**, *290*, 1347.

Received: June 27, 2005
Published online on September 19, 2005

