Origin of Life

DOI: 10.1002/anie.200503387

A DNA-Templated Aldol Reaction as a Model for the Formation of Pentose Sugars in the RNA World**

*Michael Oberhuber and Gerald F. Joyce**

An early stage in the history of life on Earth is thought to have involved RNA genomes and RNA catalysts, an era commonly referred to as the "RNA world".[1,2] RNA-based life would have required substantial amounts of the building blocks of RNA, such as the sugar ribose. The formation of ribose and other pentoses from simple precursors involves a substrate-specific cross-aldol reaction of glyceraldehyde and glycolaldehyde.[3] Without enzymatic control, however, a network of competing pathways, including self-aldol, isomerization, and retro-aldol reactions, gives rise to a complex mixture of compounds that includes only a small amount of pentoses and challenges even the most advanced analytical techniques.[4] While several laboratories have reported conditions that can lead to an increased yield of ribose and other pentoses from simple aldehydes,[5–9] a chemical system that allows direct and systematic investigation of the reaction of glyceraldehyde and glycolaldehyde has remained elusive. Such a system would enable examination of the kinetic and mechanistic features of this reaction, setting the stage for the development of RNA catalysts that can be used to synthesize ribose, analogous to those that may have existed in an RNA world.

The present study demonstrates a DNA-templated model system that allows the selective formation of pentoses by a cross-aldol reaction of glyceraldehyde and glycolaldehyde. DNA-templated chemistry parallels nature's approach to selectivity by exploiting the molecular recognition properties of biological macromolecules to direct the reactivity of small molecules.[10] The reactivity of the aldehydes was controlled by linking them to oligonucleotide "handles", which were recognized by Watson–Crick pairing to a complementary

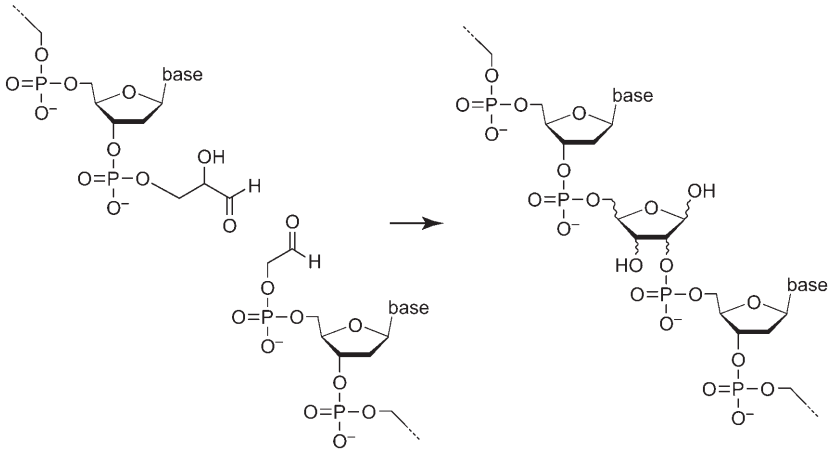
[*] Dr. M. Oberhuber,[+] Prof. Dr. G. F. Joyce
Departments of Chemistry and Molecular Biology and
The Skaggs Institute for Chemical Biology
The Scripps Research Institute
La Jolla, CA 92037 (USA)
Fax: (+1) 858-784-2943
E-mail: gjoyce@scripps.edu

[+] Current address:
Institute of Organic Chemistry and
Center for Molecular Biosciences
Leopold-Franzens University of Innsbruck
6020 Innsbruck (Austria)

[**] This work was supported by NASA and The Skaggs Institute for Chemical Biology at The Scripps Research Institute. M.O. was supported by postdoctoral fellowship J2296 from the Austrian National Science Fund (FWF) and the Max Kade Foundation.

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

DNA template. These handles also made it possible to use sensitive detection methods available for nucleic acids to overcome the analytical challenges. Each aldehyde was linked to a different synthetic oligonucleotide (Scheme 1), allowing



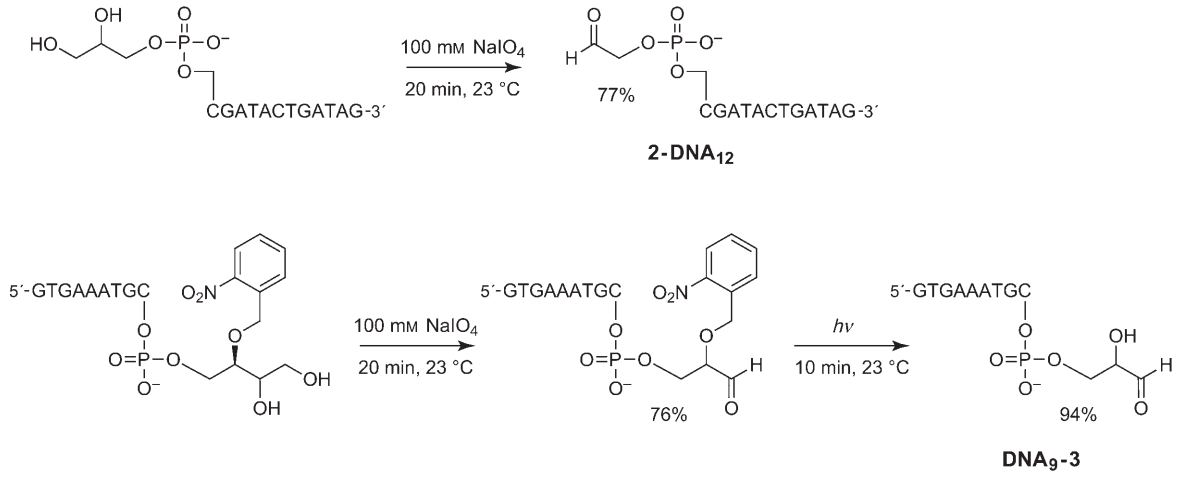
Scheme 1.

the two to be co-aligned in the desired orientation through binding to a common template. The template was designed to promote the cross-aldol reaction by increasing the effective concentration of the two juxtaposed aldehydes, at the same time suppressing undesired reactions. Over the past few years DNA templates have been shown to promote a remarkably diverse set of organic reactions.[10,11] A DNA-directed aldol reaction, however, has not yet been described.

A new class of oligonucleotide–aldehyde conjugates was synthesized by using a phosphoramidite-based approach. The synthetic strategy involved vicinal diols as aldehyde synthons, which were incorporated during solid-phase DNA synthesis and converted to the corresponding aldehyde through periodate oxidation.[12] The 5′-glyceryl-DNA conjugate was synthesized in the 5′-to-3′ direction, purified by denaturing polyacrylamide gel electrophoresis (PAGE), and oxidized to the glycolaldehyde-bearing DNA (**2-DNA$_{12}$**; Scheme 2, top).

The design of the diol precursor for the glyceraldehyde-DNA conjugate (**DNA$_9$-3**) involved a photolabile protection scheme to prevent formation of glycolaldehyde. The hydroxy group vicinal to the phosphate was protected with an *o*-NO_2-benzyl moiety, which is stable under the conditions of DNA synthesis and deprotection, and can be cleaved by irradiation with UV light following periodate oxidation.[13] The diol precursor for **DNA$_9$-3** was prepared by standard DNA synthesis using a glyceraldehyde synthon as the 3′-terminal residue. The precursor was purified by PAGE and converted to the corresponding aldehyde by oxidation and photolysis (Scheme 2, bottom). The identities of **DNA$_9$-3** and **2-DNA$_{12}$** were verified by matrix-assisted laser desorption time-of-flight (MALDI-TOF) mass spectrometry (see Supporting Information) and by DNA-templated reductive amination[11] (data not shown).

By employing the same approach, longer versions of the two DNA–aldehyde conjugates were synthesized and used to explore different formats for the template-directed reaction. The compound 5′-T_{15}GTGAAATGC-3′-glyceraldehyde (**DNA$_{24}$-3**) was efficiently [5′-^{32}P]-labeled by using [γ-^{32}P]-ATP and T4 polynucleotide kinase. It was mixed with an excess of glycolaldehyde-5′-CGATACTGATAGGACGAAGAGAT-GGCACC-3′ (**2-DNA$_{29}$**) and a complementary template that contained an unpaired thymidylate opposite the site of the juxtaposed aldehydes. Following incubation at pH 8.5 and 23 °C, the reaction mixture was analyzed by PAGE, which demonstrated formation of products with the electrophoretic mobility expected for the pentose-linked DNA (**DNA$_{24}$-5-DNA$_{29}$**) (Figure 1; lanes 2, 3, and M). There was no detectable reaction in the absence of the template or in the presence of a mismatched template (Figure 1; lanes 6 and 7, respectively). If an analogue of **DNA$_{24}$-3** was used that contained glycerol instead of glyceraldehyde, no product formation was observed (Figure 1; lane 8). Similarly, replacing **2-DNA$_{29}$** by



Scheme 2.

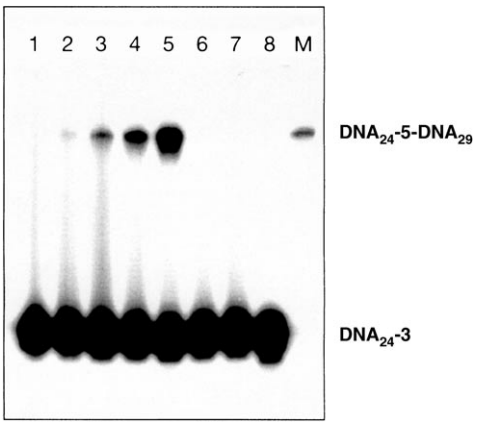


Figure 1. DNA-templated cross-aldol reaction. Reaction of 5 nM labeled **DNA_{24}-3** and 2 μM unlabeled **2-DNA_{29}** in the presence of 1 μM complementary template (lanes 2–5), no template (lane 6), or 1 μM mismatched template (lane 7). Lanes 2 and 3 correspond to reaction in the absence of L-Lys for 6 and 48 h, respectively; lanes 4 and 5 correspond to reaction in the presence of 50 mM L-Lys for 6 and 48 h, respectively; lanes 6 and 7 correspond to reaction in the absence of L-Lys for 48 h. Lane 1 contains unreacted labeled **DNA_{24}-3**; lane 8 corresponds to a control reaction (48 h) employing a labeled DNA-glycerol conjugate instead of **DNA_{24}-3**, and lane M contains a marker of the same composition as **DNA_{24}-5-DNA_{29}** but with thymidine in place of the newly-formed pentose. All reaction mixtures contained 10 mM $MgCl_2$, 100 mM NaCl, and 50 mM TAPS (pH 8.5). Reaction products were separated by denaturing 20% PAGE, an autoradiogram of which is shown.

the corresponding glycerol-containing DNA eliminated the observed reaction (data not shown).

In order to carry out mass-spectrometric analysis of the products, unlabeled **DNA_9-3** was incubated with a slight excess of **2-DNA_{12}** in the presence of a complementary template. MALDI-TOF mass spectrometry of the PAGE-purified products gave a mass consistent with the formation of a pentose linking the two substrate oligonucleotides (see Supporting Information). The **DNA_{24}-5-DNA_{29}** products were prepared similarly and analyzed by a primer extension reaction, which employed a 17 mer oligonucleotide that hybridized to the 3′ end of the product molecules and was extended by reverse transcriptase. A minor pause with substantial read-through was seen at the aldol junction (see Supporting Information), similar to the behavior of an abasic site,[14] and excluding the possibility of an unusual connectivity such as would result from the reaction of an aldehyde with a nucleobase. It was not possible to obtain sufficient material to determine the relative amounts of the four pentose sugars and their respective enantiomers. The 2′,5′-phosphodiester linkage that is formed at the aldol junction (Scheme 1) allows read-through by reverse transcriptase,[15] but the minor pause that occurs at the junction cannot be taken to indicate the presence of any particular sugar.

The rate of the DNA-templated aldol reaction was measured as a function of pH (Figure 2). Consistent with a base-catalyzed reaction, the observed pseudo-first-order rate constant increased at alkaline pH, but was nearly unchanged at neutral to acidic pH. At elevated pH and temperatures a degradation product of **DNA_{24}-3** also was detected, presumably arising from elimination of glyceraldehyde-2,3-cyclic-phosphate (data not shown). Mildly alkaline pH (8.5) and ambient temperature (23 °C) were chosen as the standard reaction conditions, for which the observed rate of reaction was about 4×10^{-6} min^{-1}.

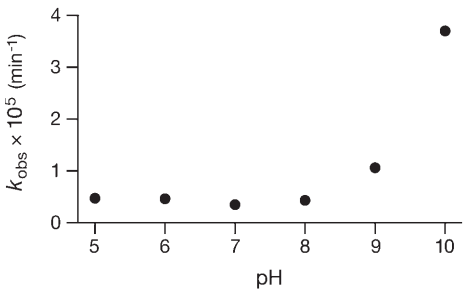


Figure 2. pH dependence of the DNA-templated aldol reaction. Reaction conditions were as described in the legend to Figure 1, except for differences in pH.

In view of the well-established role of amines as organocatalysts,[16] and with regard to their potential role as cofactors in RNA catalysis, the DNA-templated aldol reaction was studied in the presence of various amines and amino acids. A family of biological aldolases, as well as catalytic antibody aldolases that have been reported, utilize a lysine residue to catalyze aldol reactions via a Schiff-base intermediate.[17,18] This provided the impetus for experiments identifying L-Lys as a potential catalyst for the DNA-templated aldol reaction (Figure 1; lanes 4 and 5). Elevated concentrations of L-Lys accelerated the reaction up to 20-fold, but concentrations below 10 mM had no significant effect (Figure 3), suggesting a slow second-order rate of formation of the Schiff base.

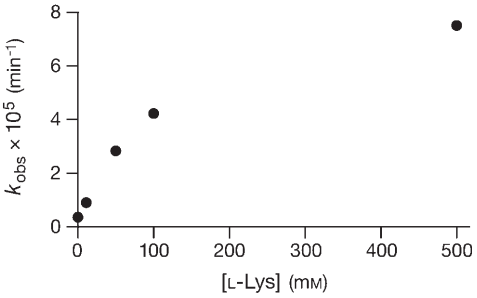


Figure 3. DNA-templated aldol reaction in the presence of varying concentrations of L-Lys. Reaction conditions were as described in the legend to Figure 1, except for the addition of 0–500 mM L-Lys.

The dipeptide Lys–Lys was similarly effective as L-Lys, but L-Pro, a well-known catalyst of aldol reactions in organic solvents,[16] was ineffective in the DNA-templated format, even at high concentrations (Figure 4). The aldol reaction also was accelerated by small aliphatic diamines, especially putrescine (Figure 4), suggesting the need for an extended aliphatic primary amine. As a negative control, the reaction was analyzed in the presence of 1M NaCl, excluding the possibility that the enhanced rate with certain amines was due to the high ionic strength of these solutions.

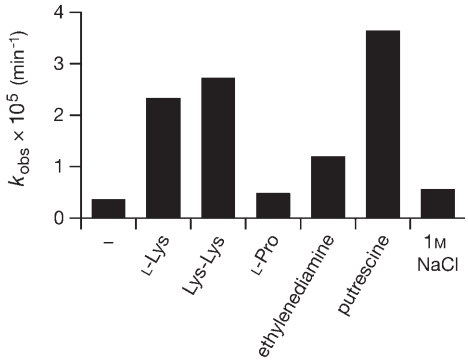


Figure 4. DNA-templated aldol reaction in the presence of various amine-containing cofactors. Reaction conditions were as described in the legend to Figure 1, except for the addition of no cofactor (lane –), 50 mM cofactor (indicated), or 1 M NaCl.

The influence of the template architecture on the aldol reaction was examined by replacing the unpaired thymidylate opposite the aldol junction with a hexaethylene glycol linker, or by placing the two aldehydes at the same end of hybridized DNA strands. Consistent with the reduced frequency of substrate encounter in these formats, the reaction rate decreased by about threefold compared to the standard format, but was similarly accelerated by the addition of 50 mM L-Lys. A variety of other template architectures were tested, but none resulted in an increased rate of reaction compared to the standard format.

In summary, this first reported example of a DNA-templated aldol reaction demonstrates how nucleic acids can selectively promote the cross-aldol reaction of glyceraldehyde and glycolaldehyde. This model system allows kinetic and mechanistic investigation of the template-directed formation of pentoses, relevant to the corresponding biosynthetic reaction in an RNA world. Aliphatic amines, such as L-Lys, were found to accelerate the reaction, emphasizing their role as potential cofactors in the RNA-catalyzed synthesis of ribose. The observed requirement for high amine concentrations may reflect the slow formation of a Schiff-base intermediate, indicating that precise positioning of the cofactor in the active site will be a crucial feature of enzymatic catalysis. A recent report by Famulok and colleagues[19] described the first example of an aldolase ribozyme, obtained by in vitro evolution, which catalyzes the reaction of a levulinic acid modified RNA and benzaldehyde-4-carboxamide. The ribozyme employs a Zn^{2+} cofactor, analogous to biological class II aldolases.[20] That work, together with the present study, suggest the plausibility of evolving aldolase ribozymes that catalyze the biosynthesis of ribose.

Experimental Section

Oligonucleotides were prepared by using a PerSeptive Biosystems Expedite 8909 Nucleic Acid Synthesis System, employing standard methods. Diol-containing oligonucleotides were synthesized by using either 3′-glyceryl CPG (Glen Research) or a solid-supported glyceraldehyde synthon (see Supporting Information). Oligonucleotides were cleaved from the solid support and deprotected by using either concentrated aqueous NH_3 at 55 °C for 2–20 h for standard DNA or 10 M $MeNH_2$ in H_2O:EtOH (1:1 v/v) at 23 °C for 14–20 h for diol-bearing DNA. All oligonucleotides were purified by PAGE. Diol precursor oligonucleotides were adsorbed on a Sep-Pak C18 resin (Waters) and oxidized with 100 mM $NaIO_4$ (ca. 3 mL), which was allowed to drip slowly through the cartridge at 23 °C over 15–20 min. After washing with H_2O (5 mL), then 10 % (v/v) MeOH (1.5 mL), the aldehyde-bearing DNA was eluted with 50 % (v/v) acetonitrile (5 mL). Photolysis was carried out at 23 °C for 10–15 min using a 100-W long-wavelength UV lamp.

DNA-templated reactions were carried out in the presence of 10 mM $MgCl_2$, 100 mM NaCl, and 50 mM 3-[(2-hydroxy-1,1-bis(hydroxymethyl)ethyl)amino]propanesulfonic acid (TAPS, pH 8.5) at 23 °C, unless otherwise stated. Aliquots were taken at 4–10 different time points and stored at −80 °C in the presence of 90 mM Tris-borate (pH 8.0), 20 mM Na_2EDTA, 10 % (w/v) sucrose, and 8 M urea. The samples were analyzed by denaturing PAGE, using a Molecular Dynamics PhosphorImager to quantify the radiolabeled material.

Supporting Information describes additional experimental procedures, including synthesis of the solid-supported glyceraldehyde synthon (Scheme S1), radiolabeling of oligonucleotides, mass-spectrometric analysis of the substrates and products of the DNA-templated aldol reaction (Figure S1), and primer extension analysis of the pentose-linked DNA (Figure S2).

Received: September 23, 2005
Published online: October 20, 2005

Keywords: enzyme models · nucleic acids · ribose · ribozymes · template synthesis

[1] W. Gilbert, *Nature* **1986**, *319*, 618.
[2] G. F. Joyce, *Nature* **2002**, *418*, 214.
[3] R. Breslow, *Tetrahedron Lett.* **1959**, 22.
[4] P. Decker, H. Schweer, R. Pohlman, *J. Chromatogr.* **1982**, *244*, 281.
[5] D. Müller, S. Pitsch, A. Kittaka, E. Wagner, C. E. Wintner, A. Eschenmoser, *Helv. Chim. Acta* **1990**, *73*, 1410.
[6] S. Pitsch, A. Eschenmoser, B. Gedulin, S. Hui, G. Arrhenius, *Origins Life Evol. Biosphere* **1995**, *24*, 297.
[7] A. Ricardo, M. A. Carrigan, A. N. Olcott, S. A. Benner, *Science* **2004**, *303*, 196.
[8] G. Springsteen, G. F. Joyce, *J. Am. Chem. Soc.* **2004**, *126*, 9578.
[9] J. Kofoed, J.-L. Reymond, T. Darbre, *Org. Biomol. Chem.* **2005**, *3*, 1850.
[10] X. Li, D. R. Liu, *Angew. Chem.* **2004**, *116*, 4956; *Angew. Chem. Int. Ed.* **2004**, *43*, 4848.
[11] Z. J. Gartner, M. W. Kanan, D. R. Liu, *Angew. Chem.* **2002**, *114*, 1874; *Angew. Chem. Int. Ed.* **2002**, *41*, 1796.
[12] H. Urata, M. Akagi, *Tetrahedron Lett.* **1993**, *34*, 4015.
[13] T. Tanaka, M. Orita, S. Uesugi, M. Ikehara, *Tetrahedron* **1988**, *44*, 4331.
[14] M. Takeshita, C.-N. Chang, F. Johnson, S. Will, A. P. Grollman, *J. Biol. Chem.* **1987**, *262*, 10171.
[15] J. R. Lorsch, D. P. Bartel, J. W. Szostak, *Nucleic Acids Res.* **1995**, *23*, 2811.
[16] J. Seayad, B. List, *Org. Biomol. Chem.* **2005**, *3*, 719.
[17] B. L. Horecker, O. Tsolas, C. Y. Lai, *Enzymes 3rd ed.* **1972**, *7*, 213.
[18] J. Wagner, R. A. Lerner, C. F. Barbas, *Science* **1995**, *270*, 1797.
[19] S. Fusz, A. Eisenführ, S. G. Srivatsan, A. Heckel, M. Famulok, *Chem. Biol.* **2005**, *12*, 941.
[20] W.-D. Fessner, A. Schneider, H. Held, G. Sinerius, C. Walter, M. Hixon, J. V. Schloss, *Angew. Chem.* **1996**, *108*, 2366; *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 2219.

