FOCUS ARTICLE

ChemComm
www.rsc.org/chemcomm

The quest for the chemical roots of life

Why did Nature choose nucleic acids derived from a pentose rather than a hexose or tetrose? Nina Hall discusses **Albert Eschenmoser's** work investigating the chemistry of alternative potential genetic oligomers that may throw light on the problem of how life as we know it now got started.

One of the greatest and most inspirational challenges in science today is to discover how simple molecular building blocks, created in planetary and possibly astrophysical processes, might have self-organised into the complex, replicating structures of life. Yet, despite the fact that during the past half-century, the study of the form and function of life's information processors—DNA and RNA (Fig. 1)—has rapidly evolved into a major research discipline, virtually nothing is known of how they originated. Indeed, although there is much 'coffee-break' discussion about the origin of life, there are only a handful of experimental programmes around the world that meticulously examine the chemical imperatives underpinning the emergence of a genetically viable molecular system.

Albert Eschenmoser leads just such a programme at ETH in Zurich, Switzerland and, together with Ram Krishnamurthy, at the Scripps Research Institute in California. For the past one and a half decades, he has adopted an experimental strategy that explores in a systematic and stepwise manner the structures of certain analogues of DNA and RNA, with reference to their capacity to undergo Watson–Crick base pairing and their potential for self-replication. The chemical structures of these alternative systems were all taken from the close structural neighbourhood of RNA; they are oligomeric systems that, according to chemical reasoning, could have formed by the same type of chemical processes that had led to RNA and, therefore, could have been evolutionary competitors in Nature's choice of the molecular basis of genetic function.[1] Eschenmoser's comparative approach has been to establish whether the genetic properties of the natural oligomers are truly unique—and if not, whether there is a degree of contingency, perhaps driven by environmental factors, in the type of life that developed on our planet. The strategy also plays a crucial role in pinning down speculation about the possible chemical pathways leading to RNA and/or

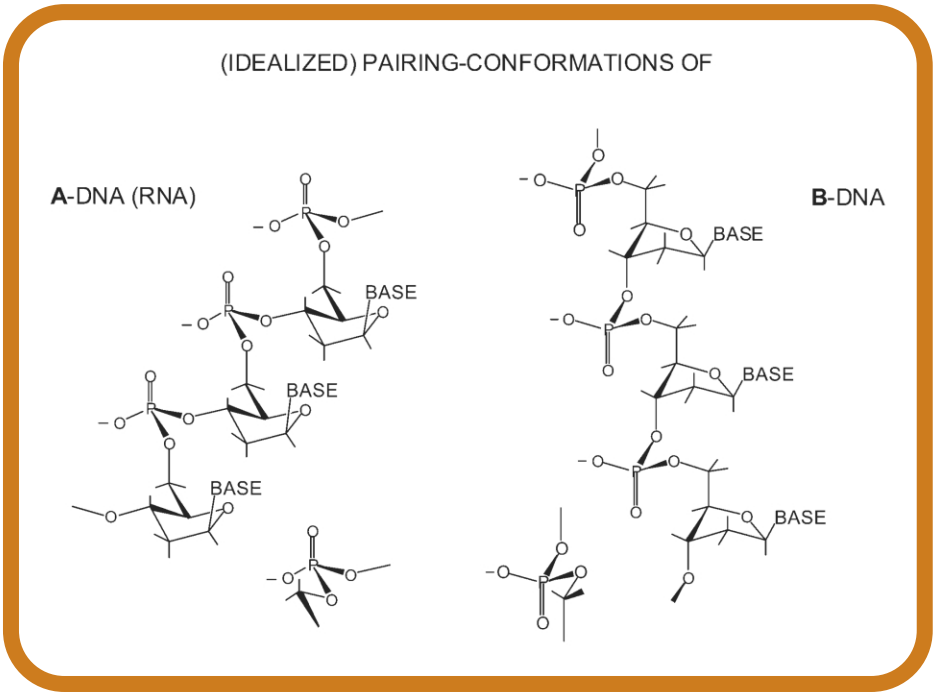


Fig. 1 Idealised pairing conformations of A-DNA (RNA) and B-DNA.

Albert Eschenmoser has been Emeritus Professor at ETH Zurich since 1992. He currently directs a research group working on the etiology of nucleic acid structure at the Skaggs Institute for Chemical Biology, part of the Scripps Research Institute in California. Over the past 50 years or so, he has made contributions to organic and bioorganic chemistry through studies of natural products, mechanisms and stereochemistry of organo-chemical and biochemical reactions, problems related to chemical bonding and structure, new methods for chemical synthesis, and total synthesis of complex biomolecules.

DOI: 10.1039/b401124b

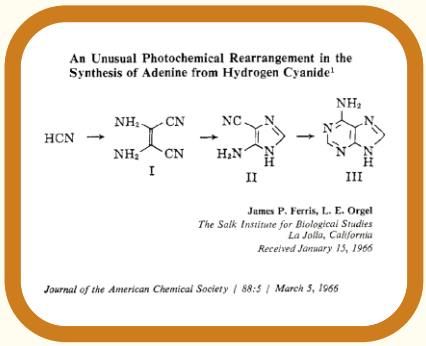
its precursors, be it in either a prebiotic or a biotic environment. These issues lie at the heart of understanding the origin of the type of life that now exists on Earth.

Eschenmoser's interest in these issues grew out of his work on the synthesis of vitamin B_{12} and, more specifically, of his subsequent synthetic studies on the chemical underpinnings of the biosynthesis of this structurally most complex of all vitamins.[2] The results of these latter studies led him to conclude that the structural complexity of the B_{12} molecule is, from a generational point of view, only an apparent complexity. He found that generating the corrinoid type of structure (corrin is the name of the central ligand chromophore system of vitamin B_{12}) can be surprisingly straightforward, provided the target is approached from the 'right' direction. The 'right' direction is that which corresponds to the chemical pathway deemed responsible for the very existence of the target structure in Nature. Such a path most often is, but does not necessarily have to be, closely related to the path of a biomolecule's contemporary biosynthesis. Seeing an 'intrinsic simplicity' in the molecular architecture of vitamin B_{12} became just a special (admittedly extreme) example of a more general postulate: that all truly fundamental biomolecules—proteins, nucleic acids, sugars, co-factors and so on—are intrinsically simple from a generational point of view.

Origin of RNA

It is now agreed that DNA could not have arisen as a self-replicator abiotically. The contemporary assembly of DNA (the blueprint for biological organisation) and proteins (the catalysts for the processes involved) is an interdependent synthetic cycle that ensures mutually faithful reproduction of their own precise oligomeric sequences, as well as the cell's other molecular components; it is difficult to see how such complex entities could have arisen both simultaneously and separately. RNA, on the other hand, has long been conjectured (by Leslie Orgel and Francis Crick) to represent an older structure than DNA, one that might have once existed in a form that could catalyse its own replication. This conjecture became more credible when, in the 1970s, Thomas Cech at the University of Colorado at Boulder and Sydney Altman at Yale University discovered RNA variants, ribozymes, that could catalyse their own cleavage. The finding has led researchers to hypothesise that now extinct sequences of catalytic RNA could have been the genetic key that opened the door to the so-called 'RNA world', a precursor life-form of the one we know today. This raises the question of how RNA arose. The simplest (and oldest) idea is that the nucleic acid bases, known to derive from potentially primordial constituents such as hydrogen cyanide, cyanamide, urea and cyanoacetylene, combined with ribose and phosphate to form nucleotides which, after activation, polymerised. However, there always were and still are a plethora of chemical and conceptual difficulties with such a simplistic scenario.[3]

These difficulties do not necessarily refer to the generation of nucleic acid bases. Juan Oró at the University of Houston had already discovered in 1961 that adenine can assemble from aqueous ammonium cyanide[4] and, subsequently, Jim Ferris and Leslie Orgel at the Salk Institute in La Jolla demonstrated a remarkable photochemical pathway from the HCN tetramer to adenine (Fig. 2).[5]


An Unusual Photochemical Rearrangement in the Synthesis of Adenine from Hydrogen Cyanide[1]

HCN → I → II → III

James P. Ferris, L. E. Orgel
The Salk Institute for Biological Studies
La Jolla, California
Received January 15, 1966

Journal of the American Chemical Society / 88:5 / March 5, 1966

Fig. 2 Adenine from HCN (facsimile from Ferris and Orgel's paper). Reprinted with permission from *J. Am. Chem. Soc.*, 1966, **5**, 1074, Copyright 2004 American Chemical Society.[5]

Later, it was shown that RNA's four canonical nucleobases can all be generated under a variety of potentially prebiotic conditions containing primordial constituents such as those mentioned above. What for some time appeared to be more critical was the ribose problem. The long-known formose reaction—the autocatalytic oligomerisation of formaldehyde initiated by glycolaldehyde—produces a messy mixture of aldo- and ketosugars in which (racemic!) ribose is only a very minor product.

Whenever Nature handles sugars biochemically, she does so with sugars in their phosphorylated forms. This induced Eschenmoser to consider a modification of the formose reaction that would lead directly to phosphorylated sugars. It involved the base-catalysed aldol reaction of glycolaldehyde phosphate in both the presence and absence of formaldehyde. Eschenmoser's group found that in the absence of formaldehyde, the eight possible diastereomeric *hexo*-aldose-2,4,6-triphosphates were formed (with the allose derivative—the 'ribose of the hexose series'—as the major product), whereas in the presence of formaldehyde under favourable conditions, the product mixture predominantly consisted of the four diastereomeric *pento*-aldose-2,4-diphosphates, with the ribose derivative as the major product (which is not the most thermodynamically stable isomer).[6] Eschenmoser thinks that this kind of intrinsic diastereoselection under conditions that allow reactions to proceed under kinetic control could be significant in understanding how Nature chose specific stereoisomers. Later, it was also shown that glycolaldehyde can be very efficiently phosphorylated in aqueous solution and at extremely low concentrations by a novel phosphorylation process using amidotriphosphate (the ammonolysis product of metatriphosphate) as a sugar-selective phosphorylation agent (Fig. 3).[7]

Why a pentose and not a hexose?

Playing experimentally with the formation of phosphorylated sugars from glycolaldehyde phosphate led Eschenmoser to ask why Nature has

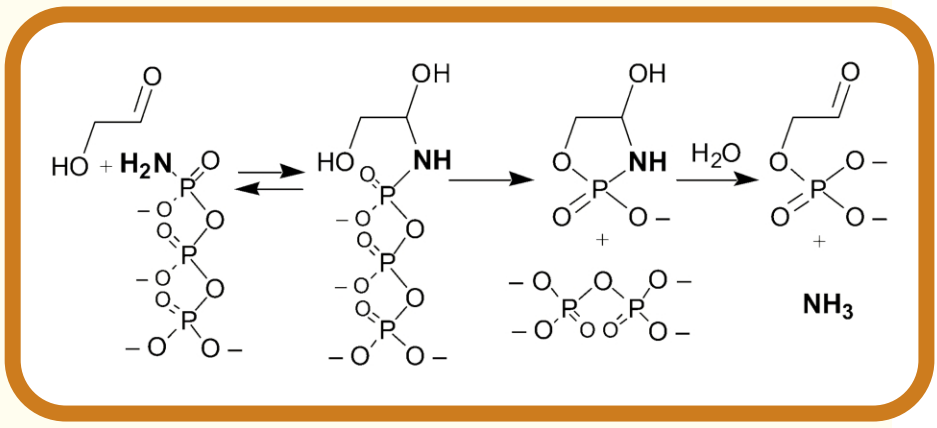


Fig. 3 Phosphorylation of glycolaldehyde aqueous solution in the presence of magnesium ions.[7]

chosen to build her genetic molecules containing a pentose rather than a hexose unit, which would seem, from a chemical point of view, an equally likely candidate. His reaction to this question was not the proposal of some quick hypothesis, but rather the initiation of an experimental research programme (starting in 1986) to screen, through synthesis, the structural neighbourhood of RNA in terms of modifying the sugar phosphate backbone, with the aim of seeing whether there were

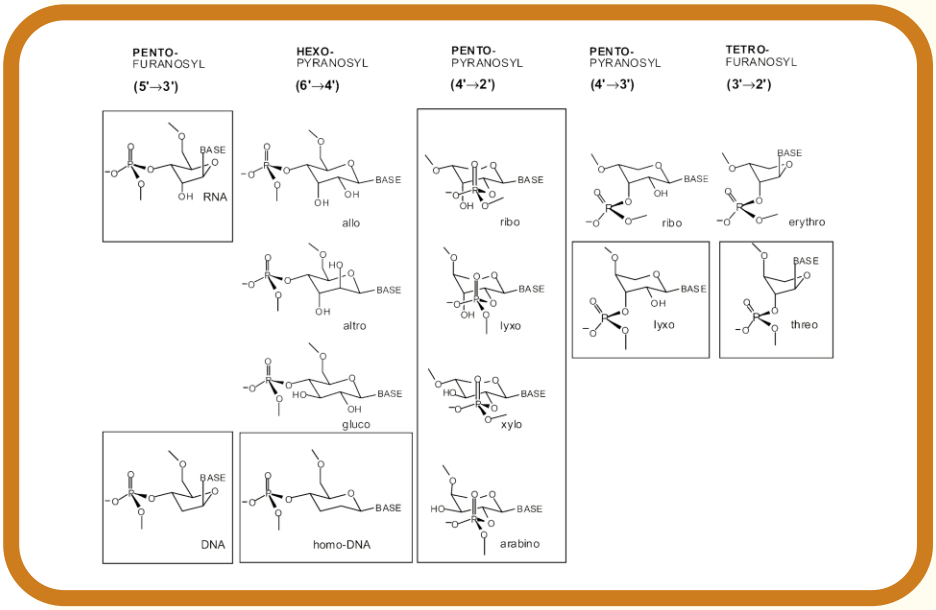


Fig. 4 Survey of monomer units of nucleic acid alternatives chosen from RNA's structural neighbourhood.[1] Figure previously published in *Orig. Life Evol. Biosphere*, 2004, **34**, 277. Reproduced with kind permission of Kluwer Academic Publishers.

any RNA analogues that could have been competitors along the evolutionary route to RNA (Fig. 4). From its very beginning, this project focused on function rather than on the formation of molecules and, therefore, was thought to be immune to the pitfalls into which projects in prebiotic chemistry notoriously fall when trying to reconstruct what might have happened billions of years ago.

The project uncovered a number of insights and some surprises! The first experiments were based on the strategy of substituting a hexose (pyranose) for the pentose (furanose) in DNA. In other words, it involved simply inserting an additional CH_2 group into the sugar ring (with everything else unchanged) in order to see the effect of ring size on the ability of an oligonucleotide system to undergo Watson–Crick base pairing. This first system, a 'homo-DNA', was just a model system, since it lacked hydroxyl groups at positions 2′ and 3′ in the pyranose ring and, therefore, could not be considered a potentially natural nucleic acid alternative. Interestingly, a qualitative conformational analysis indicated that homo-DNA might very well be a base-pairing system and, above all, that its duplexes would have a quasi-linear rather than helical structure, implying that it could be the five-membered nature of the sugar ring that is responsible for the shape of the famous double-helical structure of DNA.[8] Homo-DNA was indeed found to be a pairing system, one whose base pairing was, surprisingly, even stronger than that of DNA,[9] and one whose duplexes have, according to an NMR structure analysis,[10] a quasi-linear ladder structure. Most importantly, this model system showed that Watson–Crick base pairing can occur in systems whose duplex conformations are quite different from the canonical double helix. Furthermore, there are Watson–Crick base-pairing systems that can 'speak languages' which are different from that of the natural nucleic acids: homo-DNA pairs with itself, but does not cross-pair with DNA.

When the researchers switched to the synthesis of corresponding fully hydroxylated hexopyranosyl oligonucleotides—molecules deemed to be potentially natural nucleic acid alternatives, since they relate to natural hexoses in the way RNA relates to ribose—they found that their capacity for Watson–Crick base pairing is negligible. Additional model studies indicated that the pairing conformation of such fully hydroxylated hexopyranosyl oligonucleotide strands are destabilised by steric clashes between hydroxyl groups and neighbouring nucleobases. Fully hydroxylated hexopyranoses are apparently too bulky to allow for Watson–Crick pairing in corresponding RNA analogues. This led to the conclusion that hexoses could not have competed in Nature with pentoses as key components in a burgeoning genetic system.

Why ribofuranose and not ribopyranose?

If it is the smaller bulk of ribose relative to a corresponding hexose that is important for RNA's capability of Watson–Crick base pairing, why then ribofuranose and not ribopyranose? After all, the latter is the more stable form of ribose. The question led to the construction of the pyranosyl isomer of RNA ('p-RNA'), which is built from the very same building blocks as RNA, but contains the ribose unit in its pyranose form, with phosphodiester linkages spanning the ribose units between the 4′→2′ instead of the 5′→3′ positions (Fig. 5).

Pyranosyl RNA turned out to be not only a much stronger base-pairing system than either RNA (or DNA), but also one that is more selective with respect to base-pairing mode; it pairs exclusively in the Watson–Crick mode (no isomeric Hoogsteen pairing).[11] In fact, the whole family of diastereomeric pentopyranosyl 4′ → 2′ systems, containing β-ribo-, β-xylo, α-lyxo- and α-arabinopyranosyl rings, all showed stronger Watson–Crick pairing than RNA, with α-arabinopyranosyl exhibiting the strongest interaction.[12] They all cross-pair amongst themselves, though again not with natural RNA (or DNA). NMR studies in the p-RNA series revealed an overall duplex structure very different from that of RNA and DNA: a ladder that is slightly twisted to the left. "This family of nucleic acids alternatives taught us that Watson–Crick pairing is not unique to the natural type of system, but rather widespread among potentially natural alternatives from RNA's structural neighbourhood and, furthermore, that Nature did not choose her genetic system by the criterion of maximal base-pairing strength," says Eschenmoser (Fig. 6).

Extensive studies on the chemical properties of p-RNA delineated the scope of this RNA isomer to perform as a replicating system. p-RNA sequences can be efficiently copied by template-controlled ligation of short p-RNA sequences containing phosphate groups that are activated in a very simple way, namely, as 2′,3′-cyclophosphates.[15] Importantly, such copying processes, while proceeding with high regio- and chiroselectivity, could not be induced to proceed autocatalytically, due to product inhibition. Duplex stabilities at room temperature and template concentrations

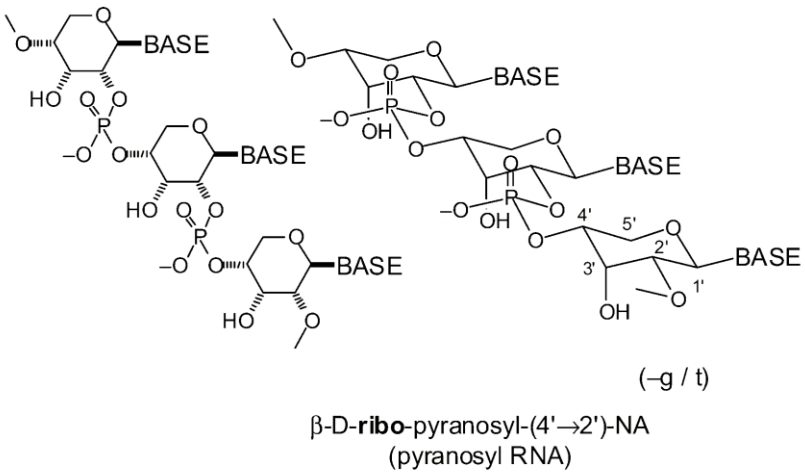


Fig. 5 Constitutional and conformational formulae of the pyranosyl isomer of RNA. Figure previously published in *Helv. Chim. Acta*, 2003, **86**, 4270. Reproduced by permission of the *Verlag Helvetica Chimica Acta*.

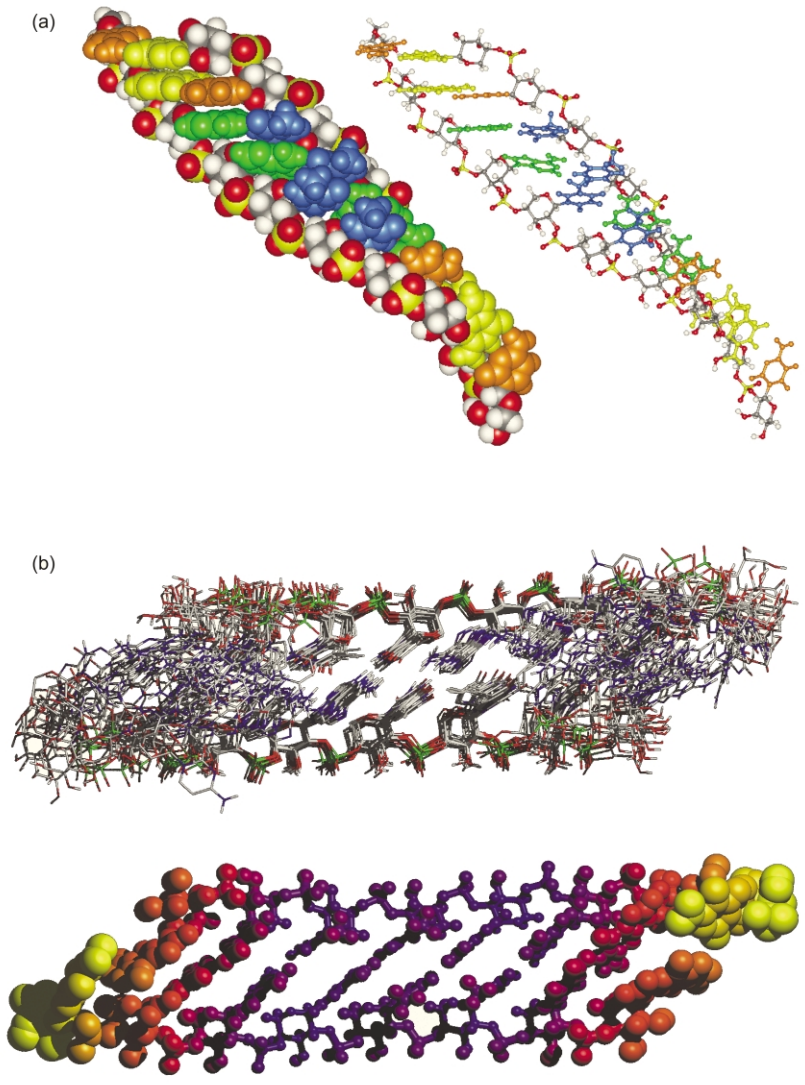


Fig. 6 NMR structures of duplexes derived from the self-complementary base sequence (4′)CGAATTCG(2′) in the D-β-ribopyranosyl- (a)[13] and L-α-arabinopyranosyl- (b)[14] RNA series. Part (a) is part of a figure reprinted with permission from *Science*, 1999, **284**, 2118 (www.sciencemag.org). Copyright 1999 AAAS. Part (b) was previously published in *Helv. Chim. Acta*, 2002, **85**, 4055. Reproduced by permission of the *Verlag Helvetica Chimica Acta*.

high enough to allow ligation to be monitored analytically are too high. On the other hand, p-RNA sequences were shown to be capable of autonomous and continuous growth within duplexes by way of self-templating ligation of activated smaller pieces, such as semi-complementary tetramer-2′,3′-cyclophosphates.[16]

So why did Nature choose the furanosyl structure of RNA and not its pyranosyl isomer? After all, both systems are composed of the very same building blocks and pyranosyl RNA can be presumed to be thermodynamically more stable. While Eschenmoser cautions against making facile conclusions based on limited knowledge, his work clearly points to backbone flexibility and pairing strength as being major influences. High base-pairing strength can facilitate template-controlled growth of oligomer strands, but is detrimental to turnover in replication and, therefore, will hamper a system's potential to evolve.

Even four carbons will do

The Swiss team was soon to discover an extraordinary RNA analogue. In pinning down the regioselectivity of template-controlled ligation in the p-RNA series and in exploring the structure–function relationships within the family of pentopyranosyl (4′→2′)-oligonucleotides, the Eschenmoser group looked at pentose-based oligonucleotides in which the phosphodiester junction sits between the 4′ and 3′ positions in the ring, rather than between 4′ and 2′.[17] In such systems, the repeating backbone unit contains only five covalent bonds, instead of the usual six present in all previously investigated systems, including RNA and DNA. As expected, due to the shortening of the backbone, the (4′→3′) isomer of p-RNA (containing axial–equatorial phosphodiester bridges) no longer shows base pairing. However, it was found that L-α-lyxopyranosyl (4′→3′)-oligonucleotides (containing axial–axial phosphodiester bridges) not only do pair, though more weakly than the (4′→2′) isomers, but must be also helically oriented in such a way that they (weakly) cross-pair with DNA and RNA, thus illustrating how significant functional changes can be induced through subtle modifications of the shape of the sugar–phosphate backbone (Fig. 7).[17]

This surprising observation led Eschenmoser to focus attention on a family of sugars that originally was not thought to deserve consideration in the systematic search for potential evolutionary competitors to RNA. These were oligonucleotides derived from sugars with only four carbon atoms, tetroses. Preparing the conformational tetrose

analogues of L-α-lyxopyranosyl (4′→3′)-oligonucleotides, namely, α-threofuranosyl (3′→2′)-oligonucleotides, or TNA, produced stunning results. Not only did they show efficient and specific base pairing, but could also efficiently cross-pair with DNA and RNA (Fig. 8 and 9).[18]

"TNA is a thought-provoking nucleic acid alternative not only because of its RNA-like properties," says Eschenmoser, "but also because of its structural simplicity. With its four-carbon sugar building block, it is, in principle, a generationally simpler type of molecular system than RNA. The simplest pathway for the generation of the four-carbon sugar threose requires one single starting material (glycolaldehyde), whereas a five-carbon sugar requires at least two. What is more, TNA talks the same base-pairing language as RNA and DNA."

Eschenmoser warns, however, that much more work needs to be done to recognise the full range of chemically possible candidate systems and the processes that could assemble them. He recently extended the work to look at oligonucleotides with nitrogen-linked C_4 backbone units (*i.e.* with phosphoramidate instead of phosphodiester links) which derive from nitrogen-containing starting materials.[20] These systems were found to behave quite similarly to TNA. Another variation was to replace adenine with 2,6-diaminopurine, which in TNA dramatically enhances pairing and the efficiency of template-controlled ligations as a consequence of its extra hydrogen bond to uracil or thymine.[21]

Could any of these systems have been a primitive replicator? It is too early to tell, according to Eschenmoser: "Much too early, because what we need to know are the boundaries of the structure space within which potentially natural informational oligomers can be conceived and demonstrated to exist. As many as possible of them should be made and studied." He is expanding his own group's research to so-far unknown types of oligomer systems within the constraints of the criterion that a system needs to have had a chance to assemble itself geochemically in a world without life.[22]

Here, though, lies yet another set of unknowns. No one is sure of the geochemical conditions on the Earth some three to four billion years ago. It now seems likely that the atmosphere was not as reducing as originally proposed by Oparin and Haldane in the 1930s (and later used in the famous Miller–Urey experiment of 1953 to show that life's organic precursors could be created from primordial atmospheric components of hydrogen, methane and ammonia).[23] Other processes may prove to be significant in driving the chemistry, including comet and meteorite bombardment and volcanic activity. We may find out more soon when, in early 2005, the NASA/ESA Cassini-Huygens space mission reaches the Saturnian moon Titan, which is thought to resemble prebiotic Earth.

In the meantime, Eschenmoser exhorts chemists to pursue the various chemical problems associated with the quest for the chemical roots of life—problems referring by no means just to the constitutional

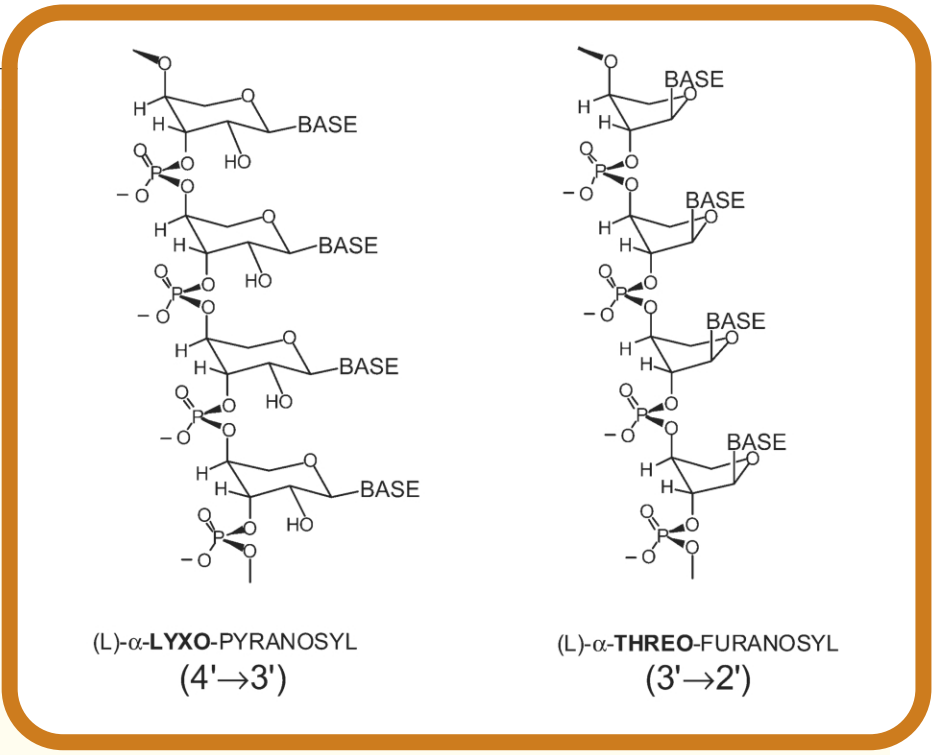


Fig. 7 Conformational formulae (idealised) of L-α-lyxopyranosyl (4′ → 3′)- and L-α-threofuranosyl (3′ → 2′)-oligonucleotides.

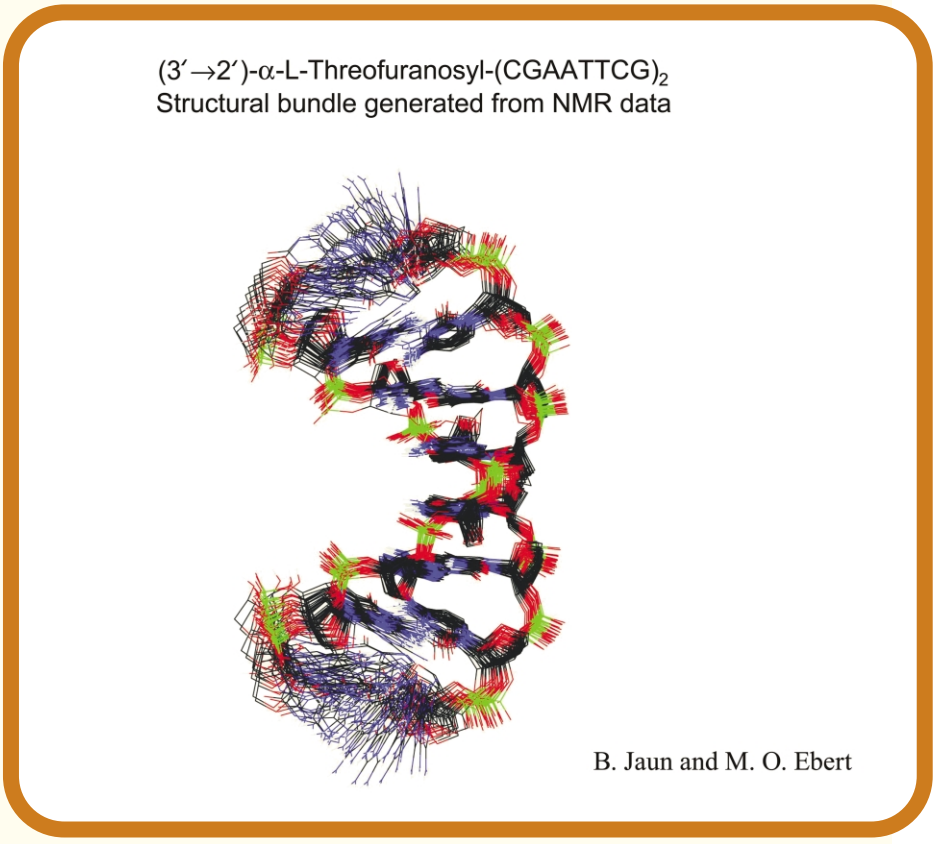


Fig. 8 NMR structure of the TNA duplex derived from the self-complementary base sequence (3′)CGAATTCG(2′).[19] Thanks are due to Professor Bernhard Jaun (ETH Zurich) for permission to reproduce the structure here prior to its publication.

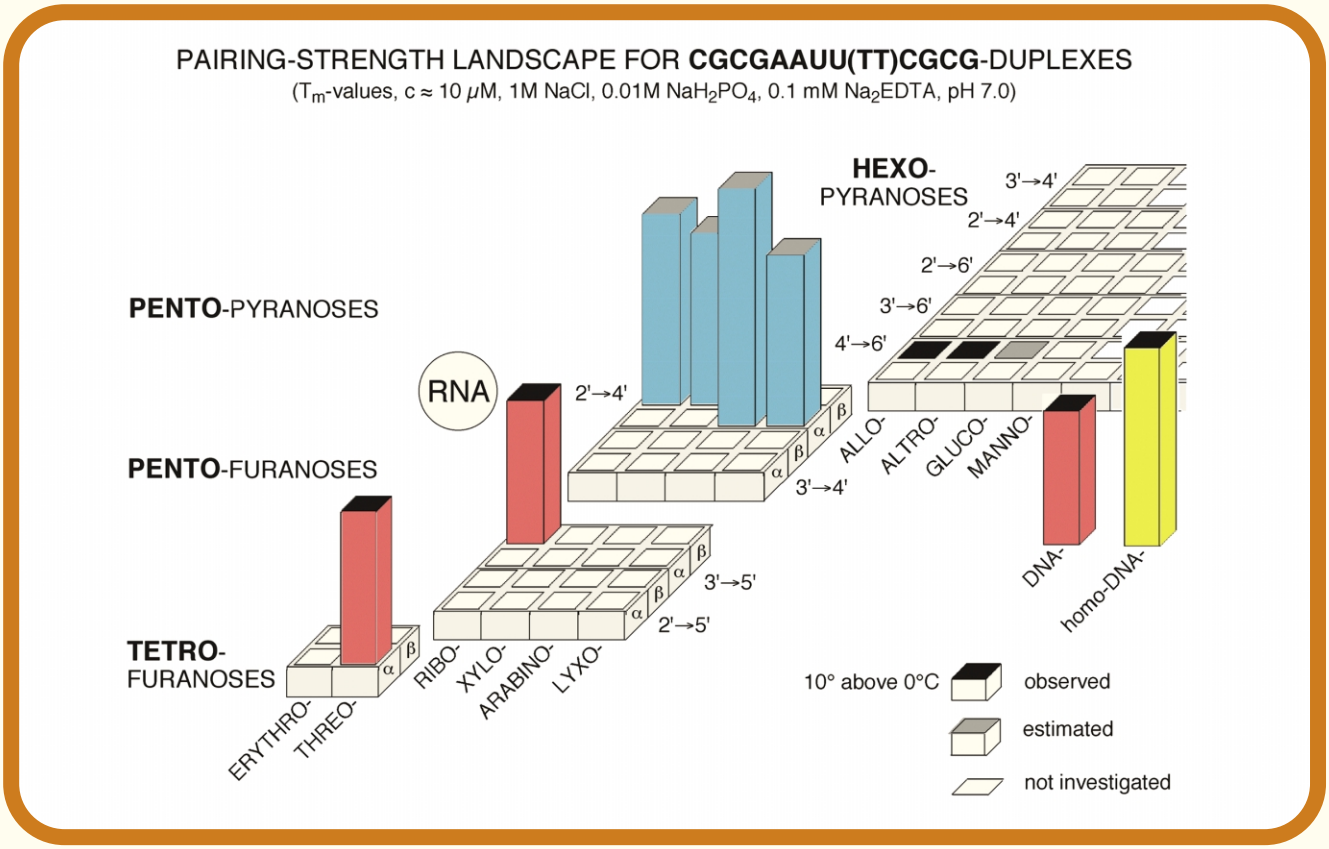


Fig. 9 Pairing strength landscape of nucleic acid alternatives investigated so far.[22] Figure previously published in *Orig. Life Evol. Biosphere*, 2004, **34**, 277. Reproduced with kind permission of Kluwer Academic Publishers.

diversity of replicating systems and their informational chemistry, but covering aspects such as the phenomenon of autocatalysis in general, the search for its experimental substantiation in not only 'genetic', but also potentially 'metabolic' autocatalytic cycles, and the operation of such cycles in cellular compartments. Rewarding new chemistry (and perhaps new applications) will be uncovered. "It is chemistry as a whole that is challenged to arrive at an understanding of what is perhaps the most extraordinary property of chemical matter—its potential to undergo a transition from non-living to living", concludes Eschenmoser.

Nina Hall

Acknowledgement

The author thanks Professor Eschenmoser for his help with this article.

Notes and references

1 A. Eschenmoser, *Science*, 1999, **284**, 2118.
2 A. Eschenmoser, *Angew. Chem., Int. Ed. Engl.*, 1988, **27**, 5.
3 G. Joyce and L. E. Orgel, in *The RNA World*, ed. R. Gesteland, T. Cech and J. Atkins, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 2nd edn., 1999, pp. 49–77.
4 J. Oró, *Biochem. Biophys. Res. Commun.*, 1960, **2**, 407.
5 J. P. Ferris and L. E. Orgel, *J. Am. Chem. Soc.*, 1966, **88**, 1074.
6 D. Müller, S. Pitsch, A. Kittaka, E. Wagner, C. Wintner and A. Eschenmoser, *Helv. Chim. Acta*, 1990, **73**, 1410.
7 R. Krishnamurthy, G. Arrhenius and A. Eschenmoser, *Origins Life Evol. Biosphere*, 1999, **29**, 333.
8 A. Eschenmoser and M. Dobler, *Helv. Chim. Acta*, 1992, **75**, 218.
9 J. Hunziker, H.-J. Roth, M. Böhringer, A. Giger, U. Diederichsen, M. Göbel, R. Krishnan, B. Jaun, C. Leumann and A. Eschenmoser, *Helv. Chim. Acta*, 1993, **76**, 259.
10 G. Otting, M. Billeter, K. Wüthrich, H.-J. Roth, C. Leumann and A. Eschenmoser, *Helv. Chim. Acta*, 1993, **76**, 2701.
11 S. Pitsch, S. Wendeborn, B. Jaun and A. Eschenmoser, *Helv. Chim. Acta*, 1993, **76**, 2161; S. Pitsch, S. Wendeborn, R. Krishnamurthy, A. Holzner, M. Minton, M. Bolli, C. Miculka, N. Windhab, R. Micura, M. Stanek, B. Jaun and A. Eschenmoser, *Helv. Chim. Acta*, 2003, **86**, 4270.
12 M. Beier, F. Reck, T. Wagner, R. Krishnamurthy and A. Eschenmoser, *Science*, 1999, **283**, 699; O. Jungmann, H. Wippo, M. Stanek, H. K. Huynh, R. Krishnamurthy and A. Eschenmoser, *Org. Lett.*, 1999, **1**, 1527.
13 I. Schlönvogt, S. Pitsch, C. Lesueur, A. Eschenmoser, B. Jaun and R. M. Wolf, *Helv. Chim. Acta*, 1996, **79**, 2316.
14 M.-O. Ebert, A. Luther, H. K. Huynh, R. Krishnamurthy, A. Eschenmoser and B. Jaun, *Helv. Chim. Acta*, 2002, **85**, 4055.
15 M. Bolli, R. Micura, S. Pitsch and A. Eschenmoser, *Helv. Chim. Acta*, 1997, **80**, 1901.
16 M. Bolli, R. Micura and A. Eschenmoser, *Chem. Biol.*, 1997, **4**, 309.
17 H. Wippo, F. Reck, R. Kudick, M. Ramaseshan, G. Ceulemans, M. Bolli, R. Krishnamurthy and A. Eschenmoser, *Bioorg. Med. Chem.*, 2001, **9**, 2411.
18 K.-U. Schöning, P. Scholz, S. Guntha, X. Wu, R. Krishnamurthy and A. Eschenmoser, *Science*, 2000, **290**, 1347; K.-U. Schöning, P. Scholz, X. Wu, S. Guntha, G. Delgado, R. Krishnamurthy and A. Eschenmoser, *Helv. Chim. Acta*, 2002, **85**, 4111.
19 M.-O. Ebert and B. Jaun, unpublished results.
20 X. Wu, S. Guntha, M. Ferencic, R. Krishnamurthy and A. Eschenmoser, *Org. Lett.*, 2002, **4**, 1279.
21 X. Wu, G. Delgado, R. Krishnamurthy and A. Eschenmoser, *Org. Lett.*, 2002, **4**, 1283.
22 A. Eschenmoser, *Origins Life Evol. Biosphere*, 2004, **34**, 277.
23 S. L. Miller, *Science*, 1953, **117**, 528.