Chemistry of Potentially Prebiological Natural Products*

Albert Eschenmoser
Laboratory of Organic Chemistry, ETH-Zentrum, CH-8092 Zürich, Switzerland
Eli Loewenthal
Department of Chemistry, Technion, Haifa 32000, Israel

The title of this lecture is meant to convey that it will describe work inspired by the question of the origin of life, a topic that certainly has much to do with 'Pattern Formation and Pattern Recognition', the main theme of this symposium. It should also make clear that the lecture will deal with sober experimental facts and not with conjectures on how life could have begun. The very nature and immensity of that problem, as against how little we do know and can know about it, behoves chemists interested in this field to proceed with the utmost caution.

The chemical study of potentially prebiological natural products intends to meet a challenge directed specifically at synthetic bioorganic chemistry. It is this branch of chemistry that is expected to contribute a central part of the factual knowledge which natural science will require for a comprehensive and experimentally based conception of the origin of life on this planet. Kinetic and thermodynamic implications of a self-organization of organic matter have been reasoned out in profound theoretical analyses;[1] and on the experimental side prebiotic chemistry[2] has established a solid basis on which to build. What is to be done is to elicit by experimental means chemical structures, reactions, and pathways that were possible vehicles of such a process;[3] eventually there will be the quest for nothing less than the design and experimental verification of structural models of self-organization of organic matter. Studies in this direction will to a large extent be an exercise in research on the specificity of molecular recognition, a topic which as such will in any case remain in the foreground of interest in organic chemistry as well as biology, for a long time to come.

On the subject of 'Pattern Formation and Pattern Recognition', an event unique in the history of science should be recalled, one which turned out to be nothing short of revolutionary for biologists and chemists alike and one where the recognition of a *pattern* was of basic and exceptional importance: the discovery of the constitutional principle underlying the structure of DNA by Watson and Crick in 1953.[4] Natural product chemists tend to look back on that event with somewhat mixed feelings, and for good reason. Here we had two non-chemists, or at any rate two who by way of their expertise or lack of it were decidedly non-organic chemists, but what they succeeded in solving was quite probably the most significant problem that Nature had held in store for natural product chemists in the field of determining the constitution of the products of Nature, namely the problem whose solution became the constitutional concept of purine–pyrimidine base pairing. The decisive step in that discovery was one of *pattern recognition* in its most elementary sense: not a carefully detailed chemical study of the structure and reactivity of the four nucleic acid bases, but a single-, open-, and (ingeniously) simple-minded search for a *pattern*, using rough cardboard cut-outs of sketched constitutional formulae of the four bases (Figures 1),[5,6] and that after having just been given a basic lesson by Jerry Donahue about using the right tautomers, and, somewhat earlier, one about Chargaff's rules.[6,7] The outcome was the recognition of those two base-pairs with their superimposable geometry. It was left to Pauling, the chemist, three years later, to point out that the guanine–cytosine pair is held together by a three- and not two-pronged hydrogen-bridge system (Figure 2).[8]

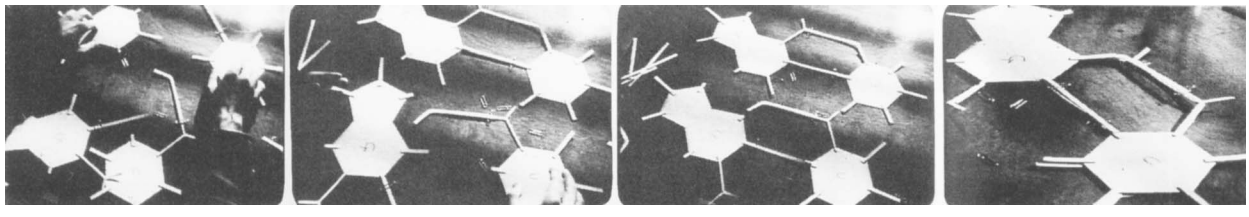


Figure 1 Stills from the BBC film 'Life Story'.
(Reproduced by permission of the BBC.)

'Prebiotic chemistry'[2] began with Stanley Miller's classic experiment[9] in 1953 and rests primarily on three important chemical discoveries that constitute the pillars on which current views on the prebiological origin of the building blocks of nucleic acids and proteins are based. They are (Figure 3): the formation of proteinogenic amino acids by the action of an electric discharge through anaerobic mixtures containing methane, ammonia, hydrogen, and water vapour (the Miller experi-

Albert Eschenmoser was born in Erstfeld, Uri (Switzerland) in 1925. He obtained a Dipl.sc.nat., Abteilung für Naturwissenschaften, from the Swiss Federal Institute of Technology (ETH), Zürich, in 1949 and gained a Dr.sc.nat. in Organic Chemistry (Prof. L. Ruzicka and Dr. H. Schinz, ETH) in 1951. He was appointed Privatdozent in Organic Chemistry at ETH in 1956, a.o. Professor in Organic Chemistry in 1960, and o. Professor for General Organic Chemistry in 1965.

Eli Loewenthal was born in Hamburg, Germany, in 1925. From 1938 to 1948 he lived in the U.K., and received his B.Sc. (Hons., External, University of London) in 1947. In 1948 he moved to Israel, and obtained his Ph.D. at the Weizmann Institute of Science under Prof. Ernst D. Bergmann in 1952. In 1952/3 he was I.C.I. Fellow at the University of Glasgow, and then for 2 years a postdoctoral Fellow with Prof. G. Stork at Columbia University, New York. Since 1956 he has been at the Israel Institute of Technology, Haifa, where he has been Professor since 1965.

* This is a slightly extended English version of a lecture presented by A.E. to a largely non-chemical audience as part of the Symposium 'Musterbildung und Mustererkennung' at the Jahresversammlung der Deutschen Akademie der Naturforscher Leopoldina on April 6—9 (1991) in Halle (Saale); the lecture has been translated from German by E.L. The German version is published in *Nova Acta Leopoldina*, 1991, Vol. **67**, No. 281. Permission for the publication of an English version was provided by the Academy and is kindly acknowledged.

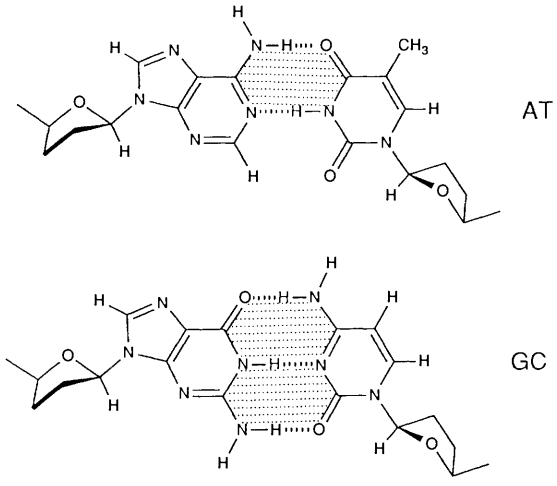


Figure 2

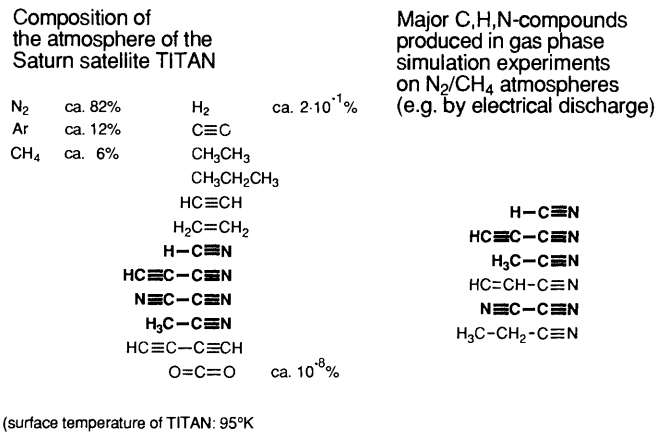


Figure 4
(F. Raulin, D. Gautier, and W. H. Ip, 'Exobiology and the Solar System: The Cassini Mission to Titan'. Reproduced by permission from *Origins of Life*, 1984, **14**, 817.)

ment[9]); the formation of adenine on heating an aqueous solution of ammonium cyanide (the Oró experiment[10]) and its photochemical variant (the Ferris-Orgel reaction[11]); and finally – in retrospect – the 'formose reaction', the formation of sugars from formaldehyde, discovered by Butlerow[12] and later elucidated by Emil Fischer.[13]

Miller's amino acids, Oró's adenine as well as the hexoses and pentoses contained in 'formose' (ribose included[14,15]) have an *elementary* molecular structure. This statement is based on the criterion of whether a molecular structure can be shown to have the capacity for constitutional self-assembly under (potentially) natural conditions. The starting materials for natural self-assembly are considered to be but a handful of the simplest organic compounds whose existence under natural and (potentially) prebiotic conditions has been demonstrated by numerous observations. Apart from formaldehyde, hydrogen cyanide, cyanogen, and cyanoacetylene would appear to be among the most important, all three molecular reservoirs of chemical energy, regularly produced whenever excess energy batters material containing carbon, nitrogen, and hydrogen, observed in interstellar space, and, *inter alia* also in the atmosphere of the Saturn satellite Titan (see Figure 4). The energy storage capacity of these and related molecules reflects itself in the triple bonds of their chemical formulae. These, to the chemist, stand for a deficit in bonding energy, and hence for the molecule's propensity to undergo addition reactions with ease. It is remarkable how some of these reactions appear to aim specifically at biomolecular building blocks. Figure 5 illustrates this for the oligomerization of five molecules of HCN to the thermodynamically much more stable adenine, and for the exothermic conversion of cyanoacetylene together with HCN and ammonia into the dinitrile of aspartic acid, further convertible by exothermic hydrolysis into aspartic acid itself. Hydrolysis of cyanogen can lead to cyanate and thence (as discovered by Woehler[16]) to urea; in the ammonolysis of cyanogen, cyanamide is formed. That compound, a tautomer of carbodiimide, is an example of an activating reagent for reactions requiring the elimination of water. In this connection it is worth recognizing that HCN and cyanogen are not only what we believe were prominent actors in chemical processes of prebiological times, they were also, as a matter of fact, objects of the first steps in the history of organic synthesis.[17]

Not only adenine but the entire family of nucleic acid bases (with the exception of thymine) are to be regarded as *elementary* structures from a chemical viewpoint. Figure 6 depicts a short excerpt from the chemistry of formation of these compounds.[10,11,18,19] There are not just one but several elementary routes to the four bases: these differ in details, *e.g.* the specific constitution of the reacting building blocks or the timing of their

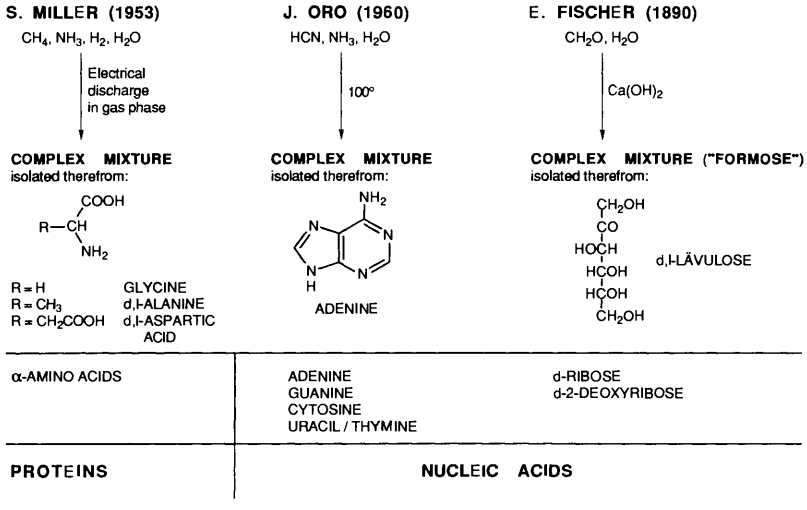


Figure 3

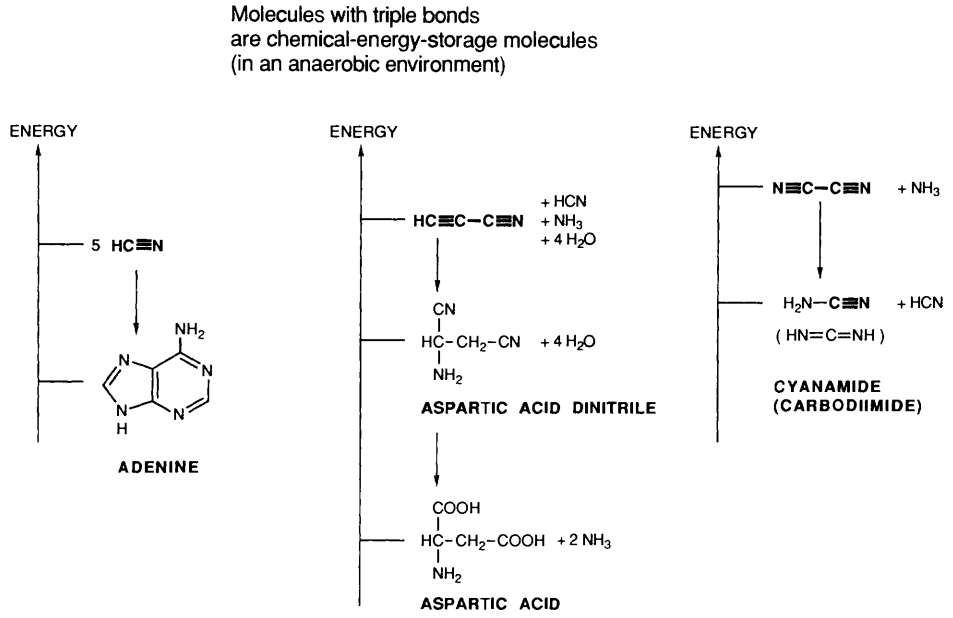


Figure 5

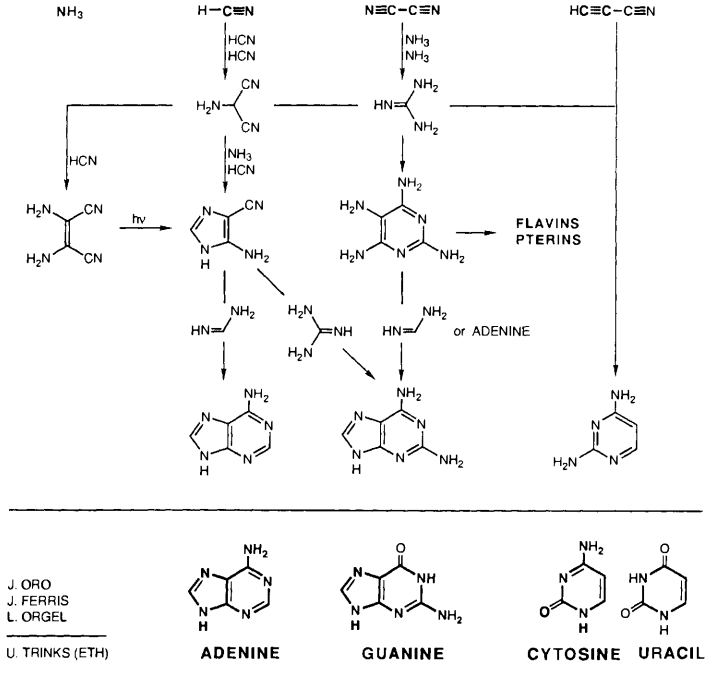


Figure 6

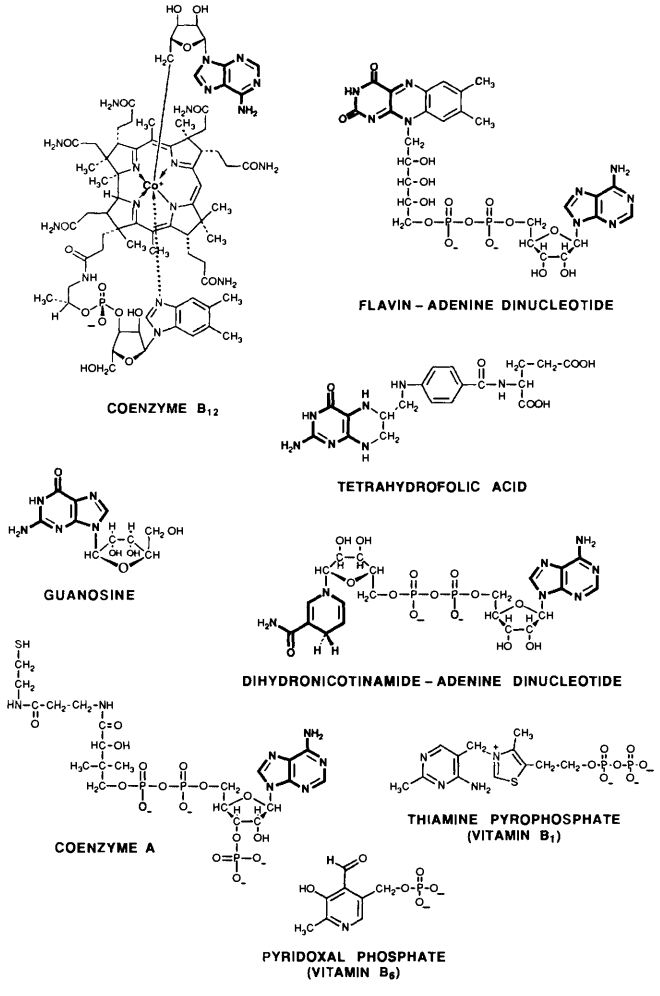


Figure 7

participation, especially also the timing of the participation of water in partial hydrolyses, yet in principle they follow the same basic channels of constitutional assembly. Schemes of this kind cannot be understood to represent specific proposals on exactly how and under what conditions biomolecular units were prebiologically formed. What they do is to point at an intrinsic structural propinquity of biomolecules and elementary starting materials and to demonstrate the existence of direct synthesis channels connecting them.

Figure 7 reminds us of the structures of a selection of cofactors (or coenzymes), molecules small but of exquisite structural complexity essential to life. By way of sophisticated operational communion with specific proteins, coenzymes are capable of fulfilling basic metabolic functions by reactions which are chemically highly demanding, e.g. reactions which change the oxidation state of metabolic intermediates, or in which carbon skeletons are constructed or modified. The choice of structures depicted in Figure 7 is dictated by the presence of constitutional patterns which, from a chemical point of view, imply *elementary* structural features in the sense referred to above. Among others, such structural features are: the recurring adenine nucleus,

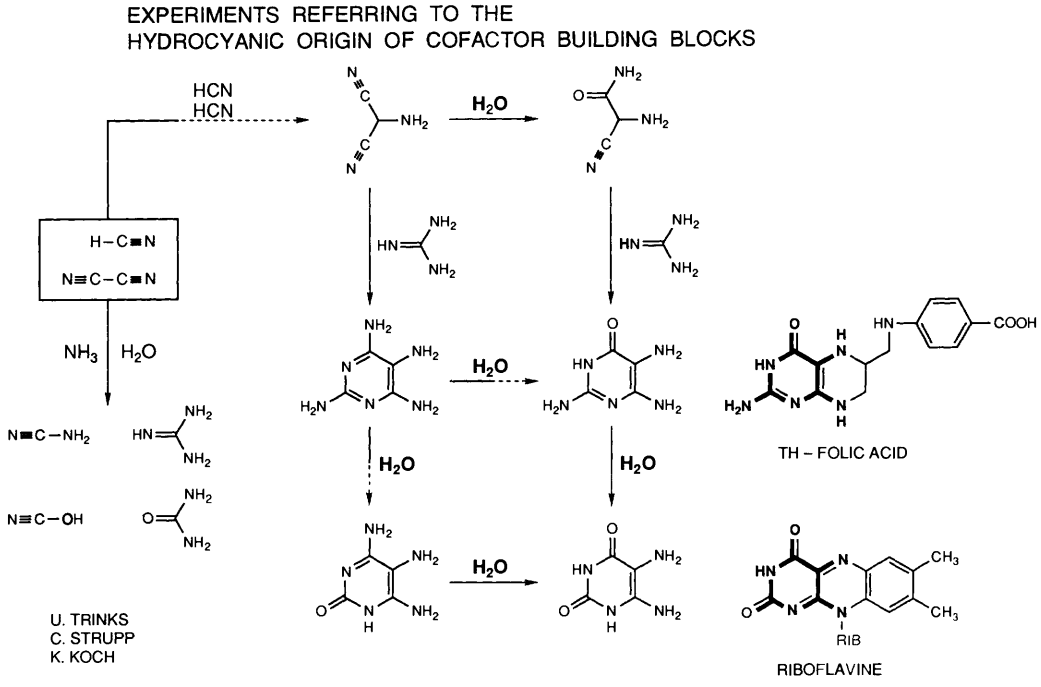


Figure 8

heterocycles that are structurally related to guanine, and the sugar unit of ribose, most often present in phosphorylated form, reminiscent of the role of this sugar as a building block of ribonucleic acid. The concept of coenzyme structures as 'molecular fossils' of an early phase of life has been alluded to in the literature,[20,21] and synthetic analyses of some of them have been carried out for the sake of demonstrating experimentally the *elementary* nature of their structures or of some of their structural features.[21] Here we look at the specific example of riboflavin, vitamin B_2, the functionally crucial structural part of the coenzyme flavin-adenine dinucleotide (FAD).

Figure 8 summarizes structural interconnections between hydrogen cyanide and cyanogen and a family of heterocyclic bases, tetra-aminopyrimidine and partially hydrolysed derivatives of it; these interconnections – except for isolation of the (highly reactive) HCN trimer from oligomerizing HCN* – have been verified experimentally.[18,19,22,23] The molecular patterns of two of the heterocycles are re-encountered in the structures of the co-factors tetrahydrofolic acid and riboflavin (Figure 8, right hand side). The latter, a vitamin and central co-enzyme of metabolic redox reactions, derives in its contemporary biosynthesis from guanosine on the one hand, and from ribose derivatives on the other. While these connections would appear to be chemically obvious with respect to the heterocyclic part and the side chain of the riboflavin structure, the origin of its carbocyclic part has for long been an enigma. Figure 9 summarizes the present state of knowledge on that subject.[24,25]

The chemically dramatic part in the biosynthesis of riboflavin is the constitutional disproportionation, catalysed by riboflavin synthetase, of two molecules of the bicyclic 6,7-dimethyl-8-D-ribityl lumazine to give one molecule of the tricyclic riboflavin and one of the monocyclic 5-amino-6-D-ribitylamino-uracil (compound A in Figure 9). The latter compound on its part, together with a C_4-molecule (see Figure) derived from ribulose-1,5-diphosphate, serves as starting point for the enzymatic formation of the self-same bicyclic lumazine derivative.[25,26] Each one of these biosynthetic transformations comprises a sequence of reactions, which by their complexity would appear to be prime examples demonstrating what one could refer to as the indispensability of enzymes. *But the really extraordinary thing about these two complex biosynthetic steps is the fact, that they can take place* in vitro *without enzymes.* This has been discovered by Wood[27,28] for the second step and shown recently in our laboratory to be the case also for the first[23] (see also[29]). Heating the monocyclic ribityl-uracil derivative in aqueous solution at pH 7 for 30 minutes at 120 °C in the absence of air produces (together with other products), the bicyclic lumazine derivative,[23] and if a neutral aqueous solution of the latter in turn is heated for 6 hours to 100 °C riboflavin is produced.[27] That of course is what natural product chemists would call a *biomimetic* (chemo)synthesis. However, in view of the foregoing it might also be appropriate to say that the two enzymatic reaction steps constitute a characteristic example of a *chemomimetic* biosynthesis. Enzymes do not create new reactions – they optimize existing ones;[30] and were we in need of a further example to illustrate this principle – here is a good one.

We could say that the molecular structures of the two reaction partners 5-amino-6-D-ribitylamino-uracil and D-ribulose-1,5-diphosphate intrinsically contain all the information necessary for a constitutional self-assembly of the riboflavin structure, since they are capable of producing it under (in principle) one single set of elementary reaction conditions without further instruction. Whether then we should refer to riboflavin as having an *elementary* structure will depend on whether it will be possible to find out if and under what conditions the (reductive!) transformation of for instance 5,6-diamino-uracil with ribose into 5-amino-6-ribitylamino-uracil can be effected by non-enzymatic means.

The role of pentose sugars in the riboflavin context brings us to the ribose problem.

There have been revived discussions recently on the prebiological origin of ribose, occasioned by the discovery of ribozymes.[31,32] Understandably, this advent has given impetus to the view that ribonucleic acids were there at the beginning, being capable both of carrying genetic information and of catalytic activity. Against that, views have been expressed (and rightly so), by researchers familiar with the problems of prebiotic chemistry, which emphatically question a prebiological availability of oligoribonucleotides.[3b,33,34] In particular, they have pointed to the complexity of the product mixture resulting from the formose reaction and its low content of (racemic) ribose. Figure 10 illustrates this by a gas chromatogram published by Decker *et al.*[15] of a formose sugar mixture (after derivatization);

* The partial hydrolysis of HCN trimer in water (pH 8) to the much more stable monoamide is very effectively catalysed by (*e.g.*) formaldehyde.[22]

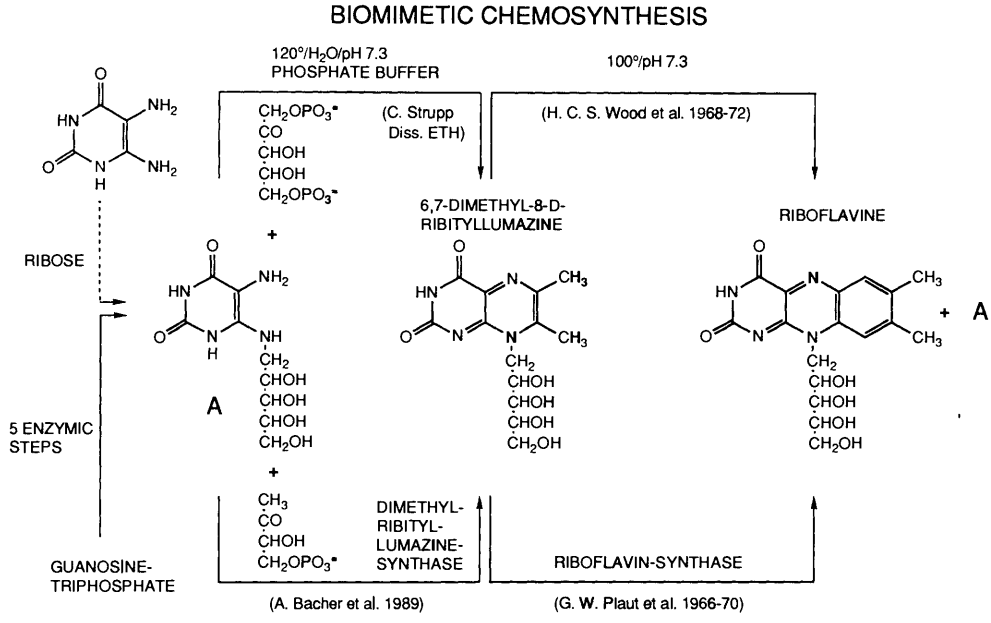


Figure 9

and Figure 11 supplies additional emphasis to the impression created by its gallery of structural formulae of the sugar constituents of that mixture, as identified by gas chromatographic and mass spectral data. Here one finds trioses, tetroses, pentoses, hexoses, heptoses, of both aldose and ketose type, straight chain as well as branched sugars, and incidentally also (racemic) ribose.

The array of structural formulae in Figure 12 shows the stations which the train of constitutional self-assembly en route from racemic ribose on the line all the way to an oligonucleotide residue would have to halt at, were it to travel on the rails defined by conventional chemistry (phosphorylation, nucleotidation, phosphate activation, polycondensation). Also, this figure, by its mirror image arrangement, points at the phenomenon of molecular chirality (handedness) and its consequences for molecular synthesis and the associated problem of stereochemical selection. The sugar units in a natural oligoribonucleotide all have the same handedness (*i.e.* specifically belong to the left one of the two mirror halves in Figure 12). In the growth of such a nucleotide chain the units successively to be added on would have to be specifically selected from two enantiomeric forms which would offer themselves with equal chance.

Figure 13 is designed to bring home to a non-chemist 'what might happen to a reacting chiral molecule on encountering a (chiral) partner with the wrong handedness'. It would probably feel the way we would, were we to offer someone our right hand, to be shaken by our partner's left. In a world of molecules and in the language of the chemist a handshake would be termed a *stereoselective act*; normally, a right hand shakes a right hand, the left to left handshake (less commonly encountered) would be its '*enantiomorph*', but for left handers of equal physiological significance. Against that, the same act right to left would be '*diastereomorphic*', that is to say not in the nature of a mirror image, but of different structure and thus for both right and left handers with physiologically different sensation (for molecules: different physical and chemical properties).

The enormous expansion of molecular multiplicity (and hence of information storage potential) in organic matter due to chirality is of such basic importance in both chemistry and biology that we may try to illustrate one of its consequences for the structure and formation of biopolymers by an additional allegorical experiment with a physiological flavour (see Figure 14). Suppose we are given the task of creating little men, beginning with a basic part of the body (which is itself suppo-

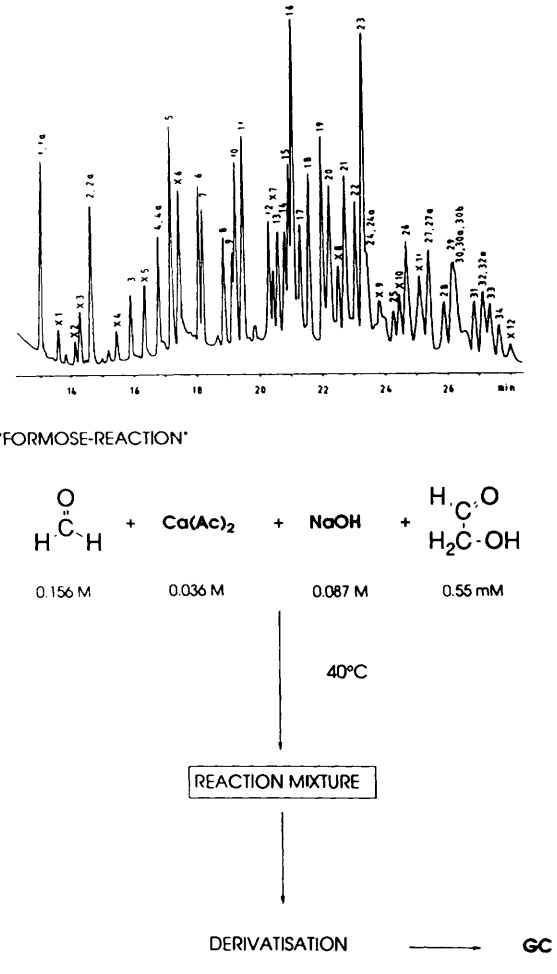


Figure 10 Gas chromatogram of n-butoxime trifluoracetyl derivatives of carbohydrates arising in the formose reaction.
(Reproduced by permission from P. Decker, H. Schweer, and R. Pohlmann, *J. Chromatogr.*, 1982, **244**, 281.)

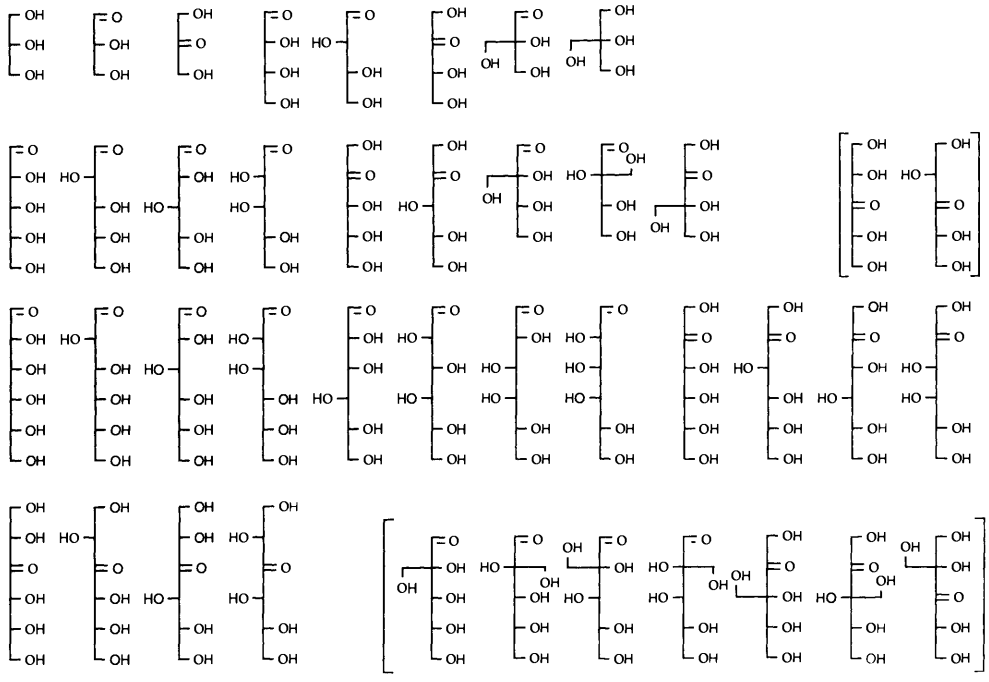


Figure 11

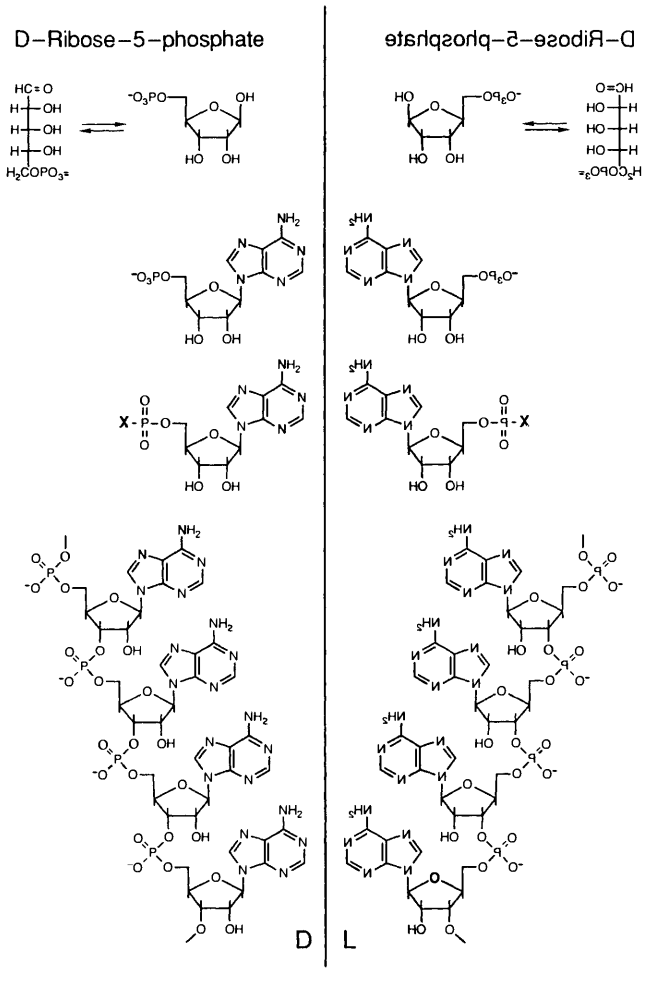


Figure 12

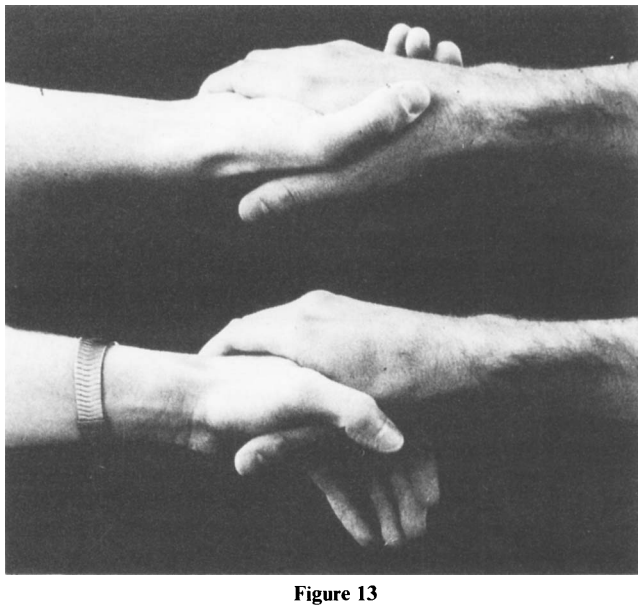


Figure 13

sedly chiral), and then adding limbs: hands, feet, and other organs. These addends, chiral themselves, would all be at our disposal as a mixture of mirror image (enantiomorphic and not superimposable) forms in a 1:1 ratio. As soon as we add a pair of hands there are four possibilities and only one of them is 'correct', *i.e.* a right hand at the right side of the trunk and a left hand at the left. Little men with two right hands or two left hands, or the left on the right and *vice versa*, would be considered 'freaks' or 'malformed'. If we now consider adding both hands and feet there are 16 alternatives, and again only one is 'correct'. In fact, when the number of organ pairs is n, then the total number of possibilities, only one of which is physiologically 'sound', amounts to 2^{2n}, which goes to show how drastically the number of possibilities rises with increasing number of added units.

The allegory of Figure 14* is meant to illustrate for the non-

* Figure 14 dates back to a lecture on 'B_{12}-Synthesis' presented before a largely non-chemical audience at the session of the Zürcher Naturforschende Gesellschaft on January 21, 1974.

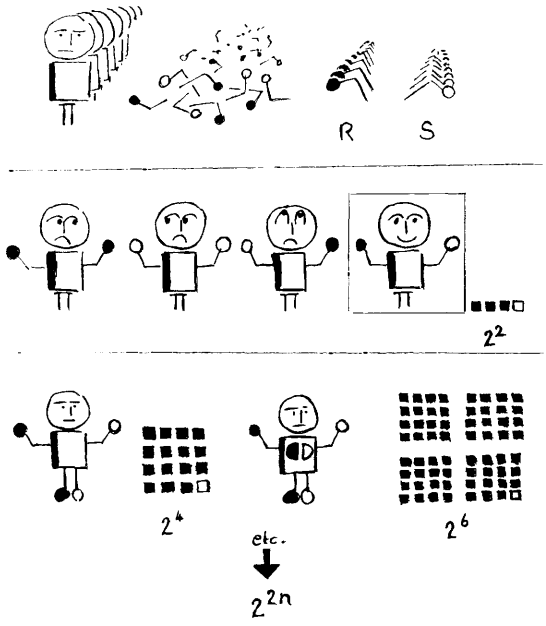


Figure 14

chemist the problems of *enantio-* and *diastereoselectivity* in chemical synthesis. Today, such problems are topics of research in the laboratories of organic chemistry the world over. The standard of achievement is set by Nature itself, *i.e.* by the phenomenon of enzymes performing with enantio- and diastereoselectivity at a level of near perfection. In biological synthesis, enzymes provide the external instruction required for structural selectivity and complexity. By sheer definition, no reliance on participation of (evolved) enzymes as external instructors can be made when it comes to the task of conceiving structural models of prebiological synthesis. Much speaks, however, in favour of the idea[35] that naturally occurring minerals may have had to play a role as external instructors. An extreme view goes so far as to propose an inorganic pre-evolution of mineral catalysts.[36]

The question of dependence *versus* non-dependence on external instruction by (structured) catalysts is a central issue in the context of prebiological synthesis. Instruction for diastereoselectivity need not necessarily come from outside. It can be of intrinsically intramolecular origin, part of the chemical properties of reactant and product molecules and, therefore, of the structure-forming process. There are many such examples in organic chemistry, and we shall discuss one based on results obtained in our laboratory, on the ribose problem.

When living Nature deals with sugars it does so almost exclusively using phosphorylated forms. This well-known fact together with our interest in the chemistry of α-amino nitriles has led us to consider a modification of the formose reaction that would lead not to free, but directly to phosphorylated sugars. This has involved a detailed experimental study of the aldol reaction of glycolaldehyde phosphate in the absence and in the presence of formaldehyde. As suggested in Figure 15 such an aldomerization should proceed in a more clear-cut manner than the normal formose reaction, because one of the important reasons for the complexity of the latter process, namely the aldose–ketose equilibration under the conditions of the reaction, is suppressed. Figure 16 shows our experimental results.[37] In the absence of formaldehyde only the eight possible diastereomeric (racemic) hexo-aldose-2,4,6-triphosphates are formed; whereas in the presence of half an equivalent of formaldehyde the major fraction of the product mixture consists of the four diastereomeric (racemic) pento-aldose-2,4-diphosphates. These aldolizations are diastereoselective to a surprisingly high degree: without formaldehyde the main product is allose 2,4,6-triphosphate whereas with formaldehyde it is ribose 2,4-diphosphate (Figure 17). In both sugars the centres of chirality along the

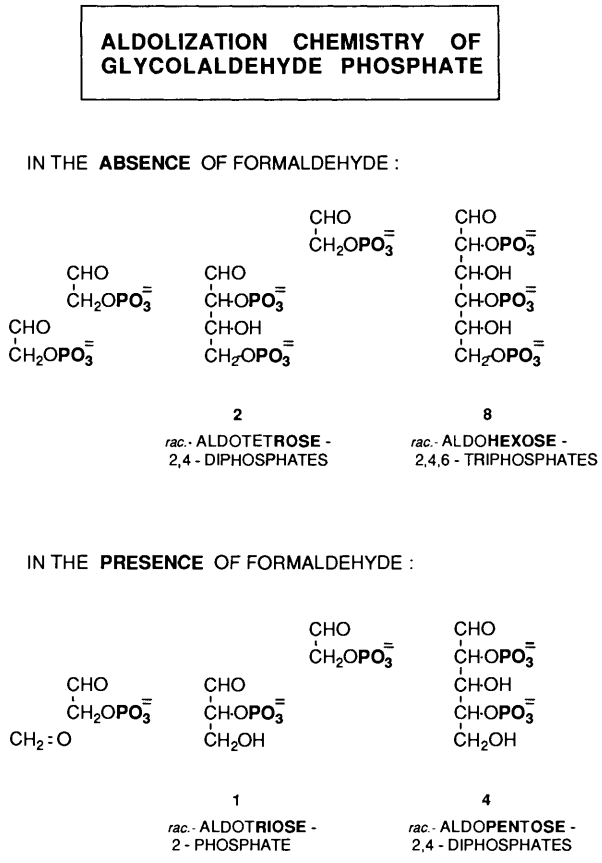


Figure 15

carbon chain show the same spatial arrangement (configuration) of the substituents; it is the configuration chosen by Nature in the process of selecting ribose among the four diastereomeric pentoses as the building block of her nucleic acids. Here then, we have an example where – so to say – the instruction required for a preferential stereochemical course of a chemical synthesis comes from within the reacting system itself. We believe that we know the nature of this instruction; it is essentially what the chemist calls the minimization of steric interference between substituents in the transition state of a kinetically controlled reaction. This intrinsic type of stereocontrol in forming the ribose structure is of course less proficient than the type of external control seen in reactions catalysed by enzymes. It is, however, self-contained and thus of particular significance in context with the problems of a prebiological ribose synthesis.

It would appear that 2,4(6)-sugar phosphates could well be better candidates for prebiological sugar formation than the free sugars themselves. Not only can they be formed much more selectively, but they are also much more stable under basic conditions.[37] Moreover, sugar phosphates would offer distinct advantages over free sugars when it came to the task of separating and concentrating sugar components out of complex mixtures of organic materials; anionic sugar phosphates are attractive candidates for selective adsorption on surfaces and interfaces of minerals.

Can glycolaldehyde phosphate be considered a potentially prebiological natural product? We encountered this compound in our experimental studies on the chemistry of α-aminonitriles,[38] specifically on aziridine-2-carbonitrile,[39] a potentially prebiological precursor of α-aminonitriles and, hence, of α-amino-acids. By reaction with phosphoric acid it forms the phosphate of serine-nitrile and, this in turn, gives glycolaldehyde phosphate by a retro-Strecker synthesis.[40] Also of interest in this context is our recent finding[41] that oxirane-2-carbonitrile – another potential α-amino-acid precursor – reacts smoothly and regiospecifically with inorganic phosphates to form the cyano-

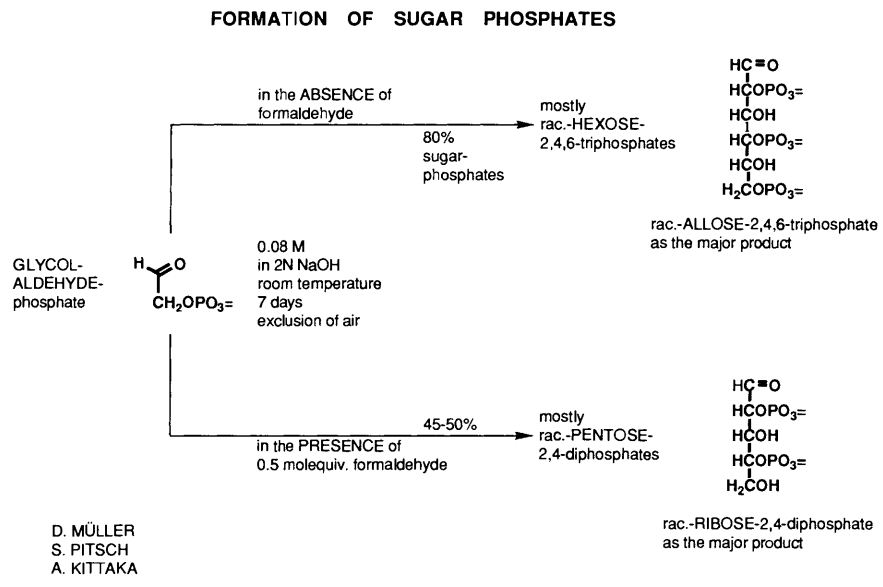


Figure 16

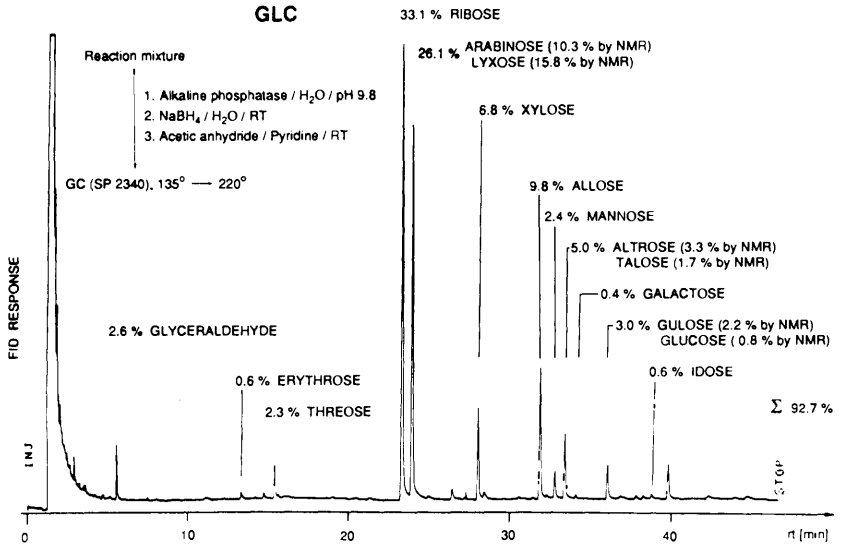


Figure 17 Peracetylated polyalcohols derived from aldomerization of glycolaldehyde phosphate in the presence of formaldehyde.

hydrin of glycolaldehyde phosphate, and that it does so in basic as well as in neutral aqueous solution (Figure 18). The possibility of a photochemical formation of oxirane-2-carbonitrile from carbon monoxide and acetonitrile would be particularly intriguing in this connection; however, as far as we know, nothing is known about it.

It was the facile and diastereoselective formation of hexose-2,4,6-triphosphates from glycolaldehyde phosphate which at the time had induced us to ask ourselves the question why Nature had chosen pentoses and not hexoses as sugar building blocks of nucleic acids. Judging from chemical experience such as described above hexoses are hardly less elementary molecular structures than pentoses; and potentially prebiological formation of the former should have had a comparable if not even better chance. There must be factors based on structure, reactivity, and finally on biological function, which are responsible for what we implicitly assume to be the biological superiority of pentose nucleic acids as against the non-existent (or extinct?) hexose nucleic acids. Insofar as biological properties reflect physical and chemical properties of molecules such factors should be experimentally ascertainable, namely, by the chemical synthesis of hexose nucleic acids, and by the study of their properties and systematic comparison of these with the corresponding properties of the natural nucleic acids. That is the goal we have set ourselves in a project which we have been pursuing in our laboratory in Zurich since 1986. What follows is an account of some of our results.[41—46]

The type of hexose nucleic acid of interest to us here has a molecular backbone that is no different from that of a natural nucleic acid save for the substitution of a hexose-pyranose for a pentose-furanose ring. The structure envisaged is exemplified by that of the D-allopyranosyl oligonucleotide shown in Figure 19. So far, however, most of our work has been done with hexapyranosyl oligonucleotides derived from 2′,3′-dideoxy-allose. They were the first to be synthesized and studied for reasons of simplicity of their synthesis. Their structure differs from that of oligonucleotides of the DNA type merely by the presence of an additional CH_2 group in the sugar ring, and hence we refer to them as homo-DNA oligonucleotides (Figures 19 and 20).

Our experiments on the synthesis of homo-DNA oligonucleotides were preceded by a qualitative conformational analysis of

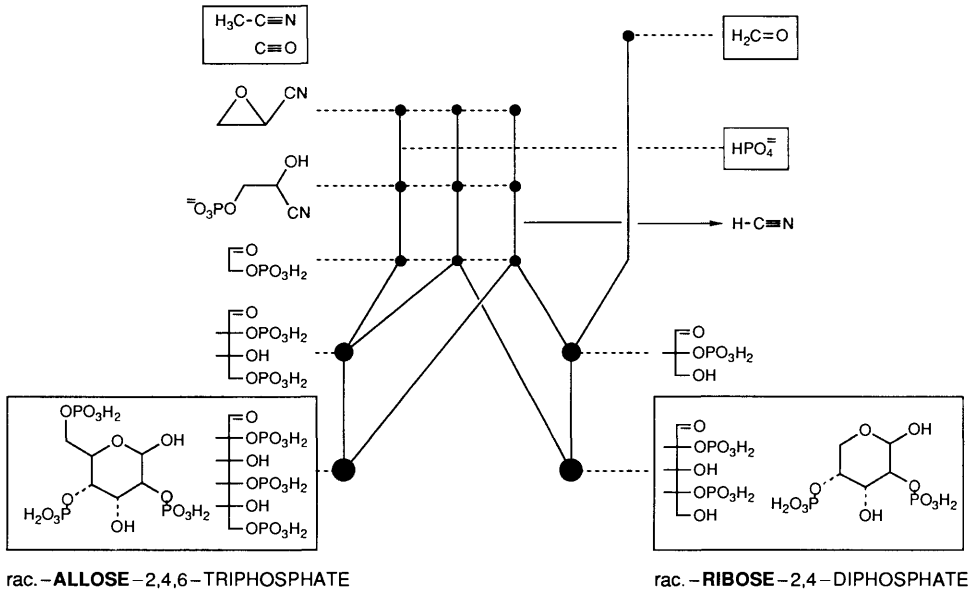


Figure 18

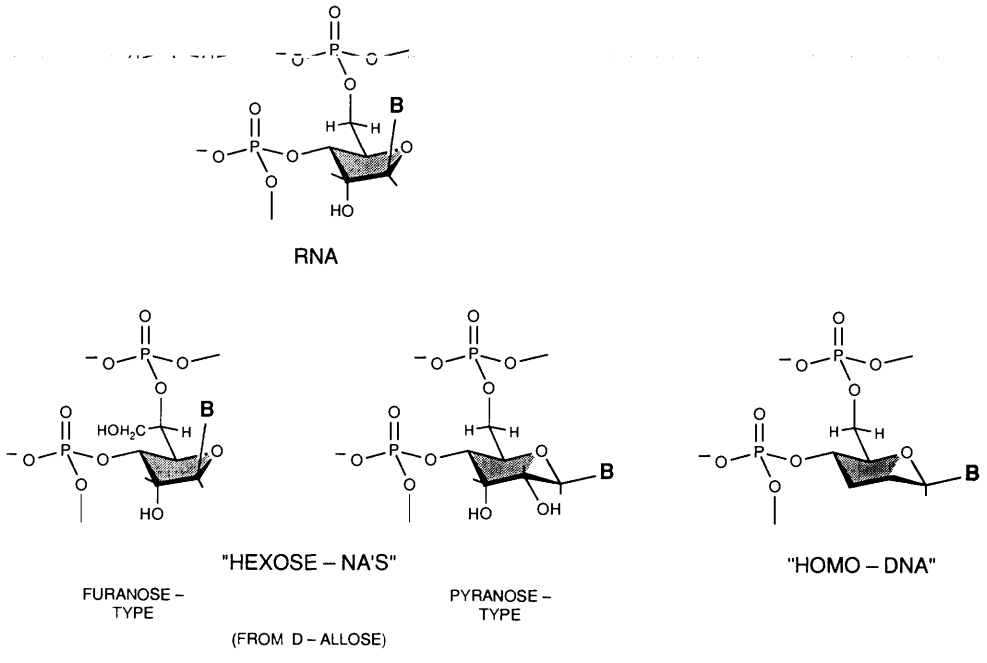


Figure 19

their structure, the results of which were as surprising as they were stimulating. They have indicated that the backbone of both homo-DNA single strands and homo-DNA duplexes in their idealized conformation should be *linear* and not helical[42,43] (Figure 21). Another outcome of this conformational study was to realise that the helical form of the backbone of DNA duplexes is first and foremost an outcome of the five-memberedness of the sugar ring. The critical structural parameter is the endocyclic torsion angle δ which, in a six-membered pyranose ring, automatically assumes a value around 60° (the value required for linearity of the backbone *cf.* Figure 21), but not in the five-membered furanose ring where this torsion angle is always greater than 60°, as a consequence of the angle strain which induces flattening of the ring. Figure 22 shows models of a DNA-(left) and a homo-DNA-(right) single strand, both constructed using identical (idealized) parameters, except for the endocyclic backbone torsional angle which is set at 60° in the homo-DNA model, and at 140° (average value in B-DNA) in the DNA model. The rotation that corresponds to the δ angle shift from 60° to 140° determines the twist of the helix (righthanded in the case of a D-sugar).

Homo-DNA strands with complementary sequences of the bases guanine, cytosine, adenine, and thymine (or uracil) were found to form antiparallel duplexes, whose melting points are uniformly higher (by up to 40 °C) than those of analogous oligonucleotide duplexes of the natural series (Figures 23 and 24). *Homo-DNA strands pair more strongly than DNA-strands.* The thermodynamic parameters[47] of a series of such duplexes show this greater stability of homo-DNA duplexes to be due not to higher hydrogen bridge or base stacking energies (ΔH values are only 2/3 of those of the DNA analogues) but to lower (negative) entropy changes: *pairing of homo-DNA strands is less unfavourable entropically than that of DNA strands* (see example in Figure 24). Various factors are presumably responsible for this; among them, the homo-DNA strand with its rigid pyranose chairs may be more highly pre-organized conformationally

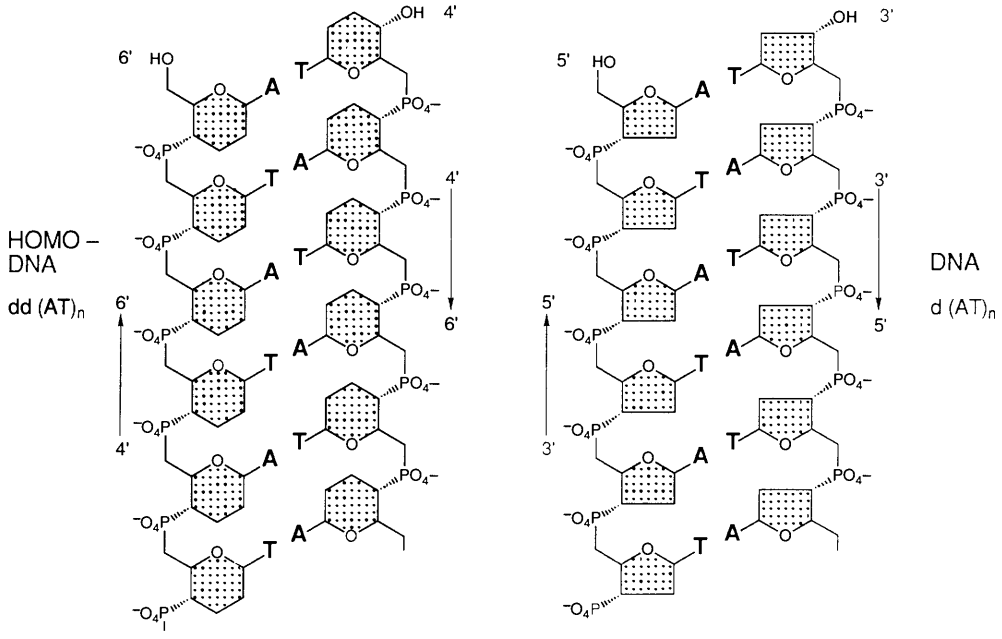
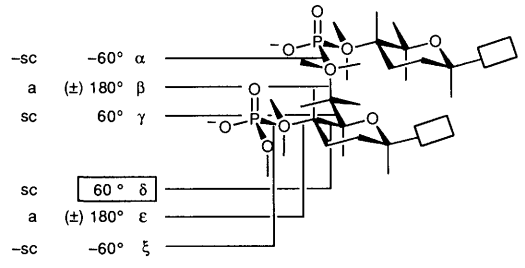


Figure 20

HOMO – DNA : QUALITATIVE CONFORMATIONAL ANALYSIS



AMONG A TOTAL OF 486 (= $2 \cdot 3^5$) POSSIBLE, IDEALIZED CONFORMATIONS OF A HOMO – DNA MONOMER BACKBONE UNIT THERE ARE **THREE** STERICALLY LEAST UNFAVORABLE CONFORMATIONS *).

ONLY **ONE** OF THEM IS REPETITIVE AND THEREFORE PRONE TO BASE PAIRING IN AN OLIGOMER DUPLEX.

THIS **UNIQUE** CONFORMATION GIVES RISE TO A HOMO – DNA BACKBONE THAT IS **LINEAR** (IN ITS IDEALISED FORM).

AT THE SAME TIME, IT BELONGS TO THE **TYPE** OF CONFORMATION THAT OCCURS IN DUPLEXES OF NATURAL A – DNA.

THE HELICAL SHAPE OF THE LATTER APPEARS AS A DIRECT CONSEQUENCE OF THE **FIVE-MEMBEREDNESS** OF THE SUGAR FURANOSE RING ($\delta > 60°$).

*) ACCORDING TO (ONLY) THREE QUALITATIVE CONFORMATIONAL CRITERIA:
— SINGLE BONDS STAGGERED THROUGHOUT
— 1,5–REPULSIONS MINIMIZED
— PHOSPHODIESTER CONFORMATIONS ACCORDING TO ANOMERIC EFFECT

Figure 21

towards duplex formation than the conformationally less rigid DNA strand with its flexible furanose rings.

All experimental data collected so far on the structure of homo-DNA complexes containing adenine, guanine, thymine (as well as uracil), and cytosine as bases are in accord with the structure assignment of antiparallel duplexes comprising base pairs of the normal Watson–Crick type.[45,48] A comprehensive multidimensional NMR study on the structure of $dd(AAAAATTTTT)_2$ has been carried out by G. Otting in K. Wüthrich's laboratory at ETH[46] (Figure 25). Most interestingly, the study reveals a conformation of the duplex backbone which *in type* is surprisingly similar to the one which had been singled out by the qualitative conformational analysis of the homo-DNA single strand (Figure 21). We are still uncertain how far the shape of the duplex deviates from linearity; an *X*-ray crystallographic analysis of crystals of the self-complementary octamer $dd(CGAATTCG)_2$ is under way in the laboratory of T. Richmond (ETH).

A characteristic and important property of homo-DNA oligonucleotides was found to be the phenomenon of purine–purine pairing. For example, the hexamers of homo-deoxyadenosine as well as the hexamer of homo-deoxyguanosine show cleanly sigmoidal melting curves (UV, CD) with remarkably high melting temperatures, in sharp contrast to the corresponding hexamers in the DNA-series. According to all the experimental evidence presently at hand on a series of homo-DNA oligonucleotides containing only adenine and guanine as bases, such oligomers form antiparallel duplexes held together by purine–purine pairing of the Hoogsteen type shown in Figure 26.[45,49,50]

The adenine–adenine as well as guanine–guanine self-pairings are stronger than the adenine–thymine pairing. This is illustrated, for example, by the fact that the $dd(A_8)_2$ Hoogsteen duplex is not transformed into a Watson–Crick duplex in the presence of the single strand of ddT_8. Importantly, base pairing in the homo-DNA series follows pairing selectivity rules which are different from those operating in natural DNA (Figure 27). This implies that the Watson–Crick rules are not only the consequence of nucleic acid base-innate hydrogen-bridging, but also an outcome of the sugar unit furanose structure.

We have extended our studies on homo-DNA purine–purine pairing to purines which do not occur in natural DNA but which nonetheless are *elementary* structures from a chemical point of view; these bases are 2,6-diaminopurine (D), isoguanine (I), and xanthine (X) (Figure 28).[51] Antiparallel complementary homo-DNA sequences of guanine and isoguanine, as well as of 2,6-diaminopurine and xanthine show sigmoidal melting curves (UV, CD) with melting temperatures so high that we must assume that we are dealing with Watson–Crick duplexes of the type shown in Figure 29. On the other hand, 2,6-diaminopurine and isoguanine can also pair with themselves. All we know presently suggests that these cases of self-pairing are analogous

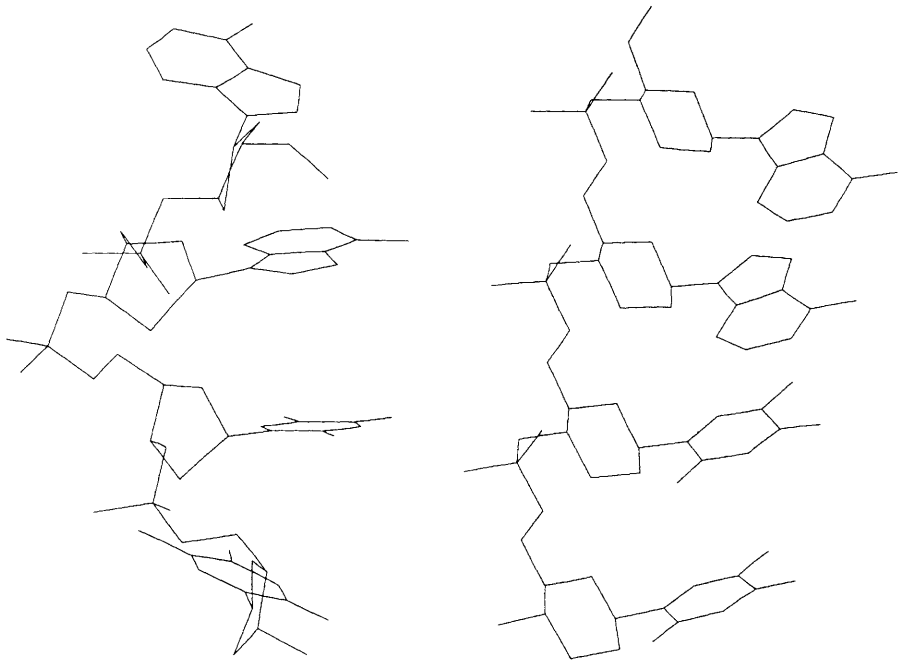
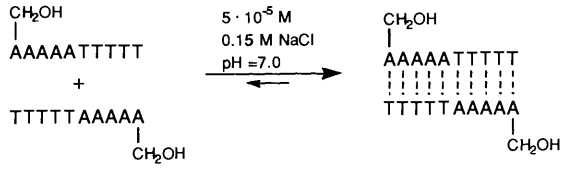


Figure 22

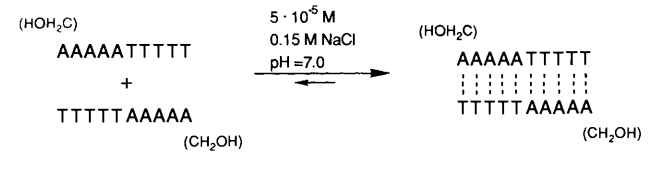


Homo-DNA strands with complementary purine–pyrimidine base sequences form ANTIPARALLEL duplexes.

Homo-DNA duplexes are thermodynamically MORE STABLE than corresponding DNA duplexes.
This higher duplex stability is not due to a higher enthalpy of formation, but to a less negative (less unfavorable) ENTROPY change.

Figure 23

(and isomorphous) to the pairing of adenine to adenine and, therefore, of the Hoogsteen type as shown in Figure 30. The table in Figure 31 gives the most recent overview* on all the purine–purine pairing combinations tested so far in the homo-DNA series in terms of melting (= de-pairing) temperatures (UV) of hexamers of indicated constitutions under standardized conditions.[48—51] Against a background of weak purine–purine interactions there emerge two distinct groups of strong purine–purine pairings distinguished by melting temperatures clustering either around 60° (the Watson–Crick group) or around 40° (the Hoogsteen group). Much of our experimental effort at present goes into consolidating our knowledge on these pairing phenomena and into attempts at establishing constitutional assignments. One aspect of these phenomena is of particular interest: the possibility that homo-DNA Watson–Crick duplexes comprising purine base sequences capable of *dichotomic* purine–purine pairing of the observed type (see Figure 32) may carry the potential for duplex replication *via* Hoogsteen-mediated information transfer** (Figure 33). Sequence information of a Wat-



	T_m(°C)	$\Delta G^{25°}$	ΔH	$T\Delta S^{25°}$
dA_5T_5	33	-7.1	-61.3	-54.2
ddA_5T_5	55	-10.5	-44.0	-33.5

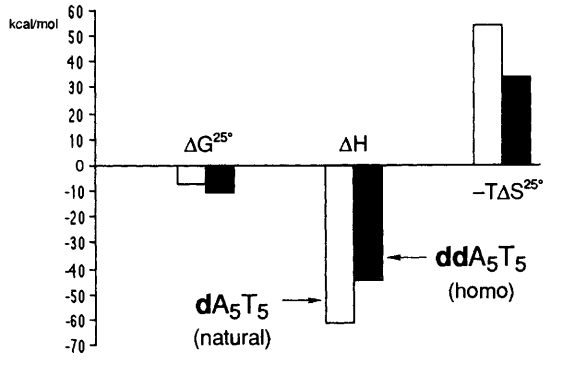


Figure 24

son–Crick duplex consisting of complementary sequences of guanine and isoguanine bases could in principle be retrieved directly from the duplex by Hoogsteen-reading without prior separation of the two strands for information exposure. This hypothetical scheme of what would look like a 'primitive' version of oligonucleotide replication raises questions which can and will be tackled experimentally; towards this end, Orgel's pioneering experiments on template controlled oligoribonucleotide synthesis[2b,53,56,34] can serve as a guide.

* The table is an up-dated version (September 1991) of the one published in the German version of the lecture.

** For an example of 'Hoogsteen-mediated information transfer' see *e.g.* Dervan's case[52] of oligonucleotide chain ligation within a oligonucleotide triplex. For recent chemical studies on non-enzymic structure replication see the work of Orgel,[53] v. Kiedrowski,[54] and Rebek.[55]

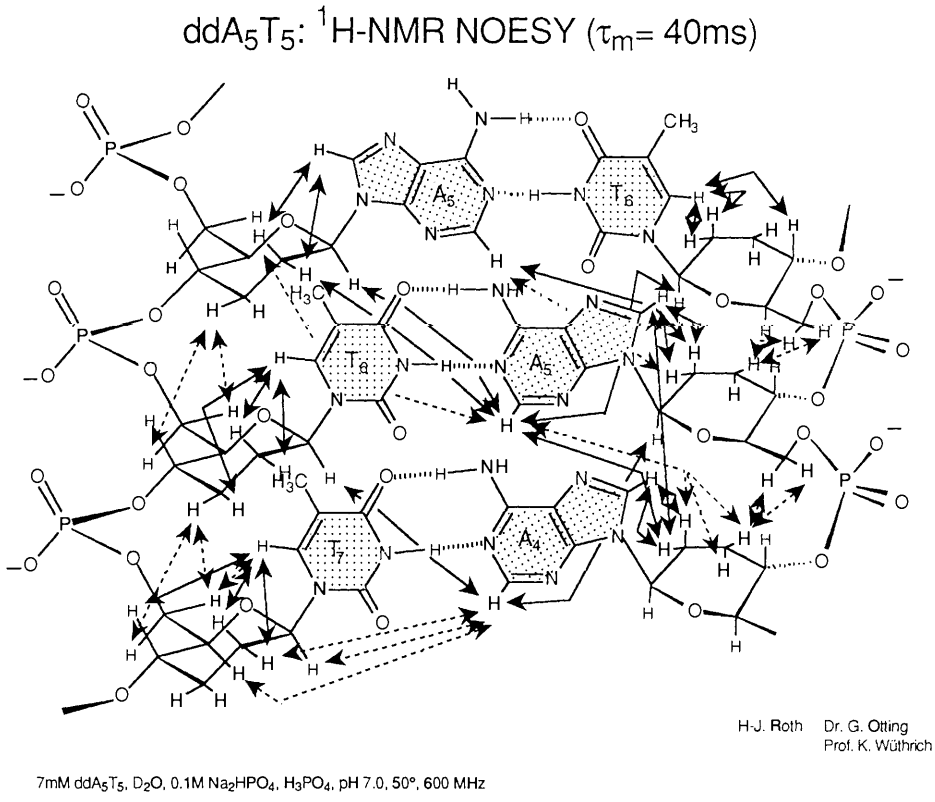


Figure 25

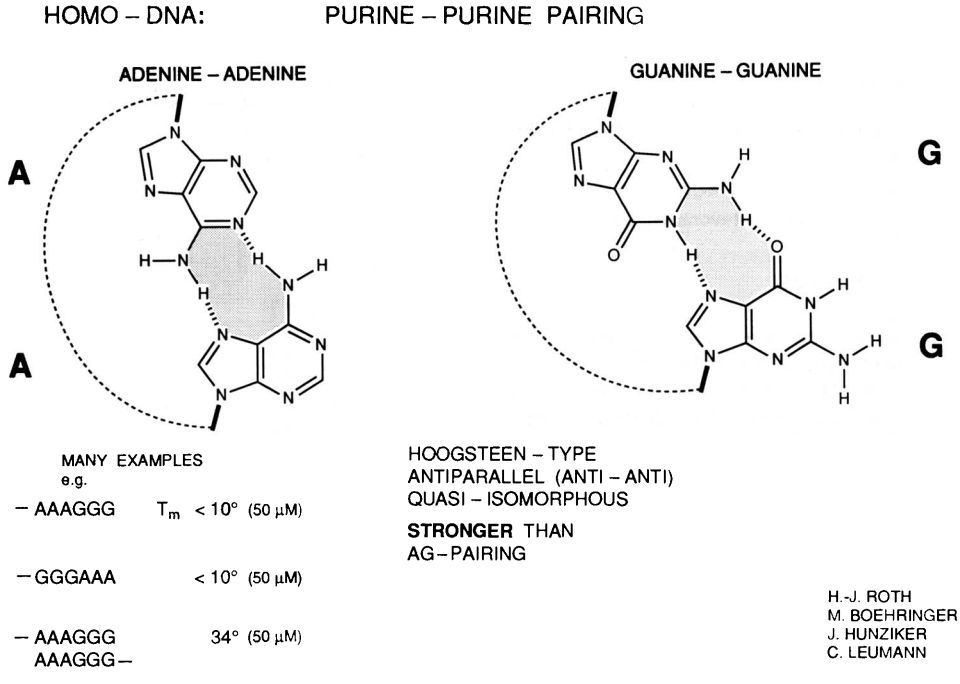


Figure 26

Our work on the model system homo-DNA can provide and already has provided important information on the consequences of substituting a furanose by a pyranose ring in DNA; it cannot, however, take the place of experimental work on nucleic acid variants derived from *potentially natural hexoses* such as for example allose or altrose as sugar building blocks. Homo-DNA's 2′,3′-dideoxy-hexapyranose cannot possibly be regarded as a potentially prebiological natural product. It does not have an *elementary* structure according to the criteria which we believe are valid in this context – quite the opposite of the seemingly 'more complex' structures of genuine hexoses like allose or altrose (Figure 34).

Work on allopyranosyl-oligonucleotides is progressing in our laboratory.[57] We have learned how to synthesize them and we already know some of their properties. In some ways they are similar to, and in others quite different from those of homo-DNA oligonucleotides. They exhibit purine–purine pairing like the latter, but the stability of duplexes in this series is uniformly

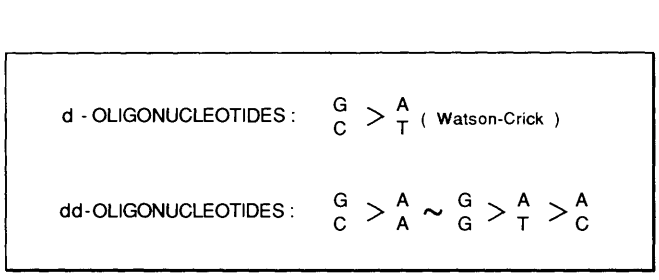


Figure 27

much lower. We believe we understand why this should be so and, in this context, we have proceeded towards the synthesis of oligonucleotides derived from 2-deoxy-allose[58] as well as altrose.[59]

Homo-DNA is an artificial and self-contained pairing system, self-contained because we know that there is no communication by pairing between homo-DNA and DNA oligonucleotide strands containing complementary base sequences.[45,48] There are probably many other functioning and possibly also self-contained pairing systems of the oligonucleotide type waiting to be discovered and to have their properties studied. A comprehensive knowledge of the structure and properties of such systems will deepen our understanding of the structural basis for the functioning of the natural nucleic acids and will illuminate their special place in evolution. Such studies may be considered to belong to the experimental part of what we might call *chemical hermeneutics* of the constitution of the natural nucleic acids.

The work which has been described in this lecture is based on contributions of outstanding collaborators to whom A. E. expresses his deep appreciation and gratitude. They are (in chronological order): on sugar phosphates, Dr. E. Wagner, D. Mueller, Dr. M. Goebel, Dr. E. Pombo-Villar, Dr. A. Kittaka, S. Pitsch, Dr. K. Atsumi; on purines, pyrimidines, folic acid, and riboflavin, U. Trinks, Dr. P. Dunn, C. Strupp, K. Koch; on homo-DNA, Dr. C. Leumann, M. Boehringer, H. J. Roth, J. Hunziker, Dr. M. Goebel, Dr. R. Chandran Krishnan, F. Giger, Dr. W. Fraser, U. Diederichsen, Katrin Groebke, and Ling Peng; on allopyranosyl-NA, R. Fischer and A. Helg.

Contributions to molecular modelling of homo-DNA oligonucleotides have been made by Prof. K. Müller (Basel), Prof. M. Dobler (ETH), and Dr. M. Billeter (ETH), to NMR spectroscopy by Dr. G. Otting (ETH), P. D. Dr. B. Juan (ETH), and Dr. H. Widmer (Basel). A. E. also acknowledges the participation of Prof. K. Wüthrich and Prof. T. Richmond in the project and, finally, he would like to thank Dr. J. Schreiber, Dr. D. Felix, Dr. E. Zass, and Dr. K. Zimmermann for their helpfulness.

The work has been financially supported by the Swiss National Science Foundation, by Ciba-Geigy AG, Basel, and by Firmenich & Cie in Geneva.

References

1 (*a*) M. Eigen, *Naturwissenschaften*, 1971, **58**, 465; (*b*) H. Kuhn, *Angew. Chem.*, 1972, **84**, 838; *Int. Ed. Engl.*, 1972, **11**, 798; (*c*) I. Prigogine and I. Stengers, 'Order out of Chaos', New Science Library, Shambhala/Boulder and London, 1984.
2 (*a*) S. L. Miller and L. E. Orgel, 'The Origins of Life on the Earth', Prentice-Hall, Englewood Cliffs, N.Y. (USA) 1974; (*b*) L. E. Orgel and R. Lohrmann, *Acc. Chem. Res.*, 1974, **7**, 368.
3 (*a*) H. Kuhn and J. Waser, *Angew. Chem.*, 1981, **93**, 495; *Int. Ed. Engl.*, 1981, **20**, 500; (*b*) G. F. Joyce, A. W. Schwartz, S. L. Miller, and L. E. Orgel, *Proc. Natl. Acad. Sci. USA.*, 1987, **84**, 4398.
4 J. D. Watson and F. H. C. Crick, *Nature*, 1953, **171**, 737.
5 M. Jackson, 'Life Story', BBC Film, 1987.
6 J. D. Watson, 'The Double Helix', Atheneum, New York, 1968.
7 E. Chargaff, *Experientia*, 1950, **6**, 201.
8 L. Pauling and R. B. Corey, *Arch. Biochem. Biophys.*, 1956, **65**, 164.
9 S. L. Miller, *Science*, 1953, **117**, 528.
10 J. Oró, *Biochem. Biophys. Res. Commun.*, 1960, **2**, 407.
11 J. P. Ferris and L. E. Orgel, *J. Am. Chem. Soc.*, 1966, **88**, 1074.
12 A. Butlerow, *Liebigs Ann. Chem.*, 1861, **120**, 295.
13 (*a*) E. Fischer, *Ber. Dtsch. Chem. Ges.*, 1888, **21**, 988; (*b*) E. Fischer, *ibid.*, 1890, **23**, 370.
14 L. Hough and J. K. N. Jones, *J. Chem. Soc.*, 1951, 3191.

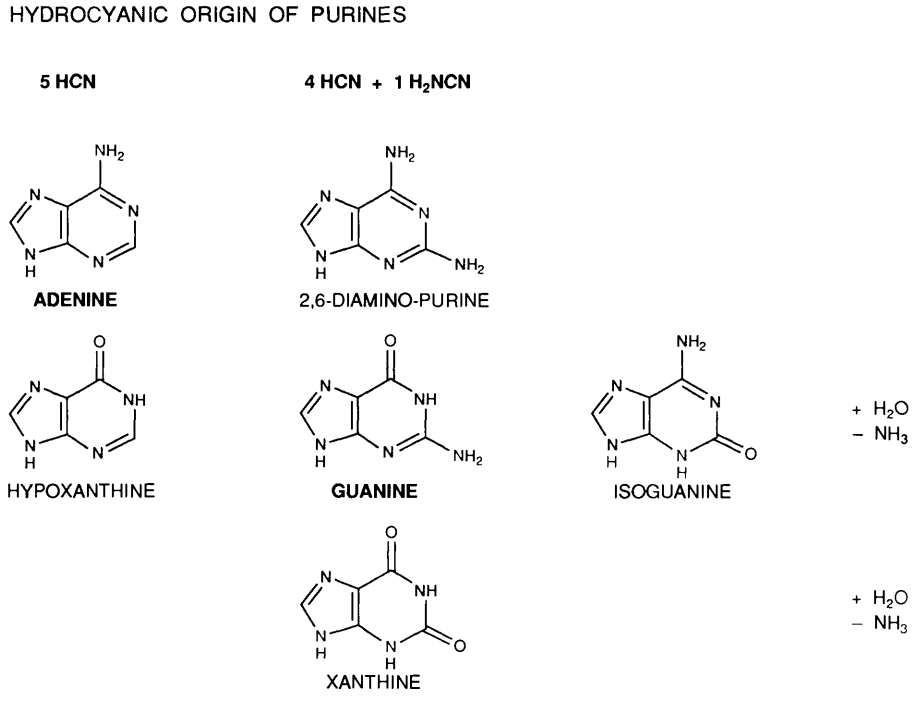


Figure 28

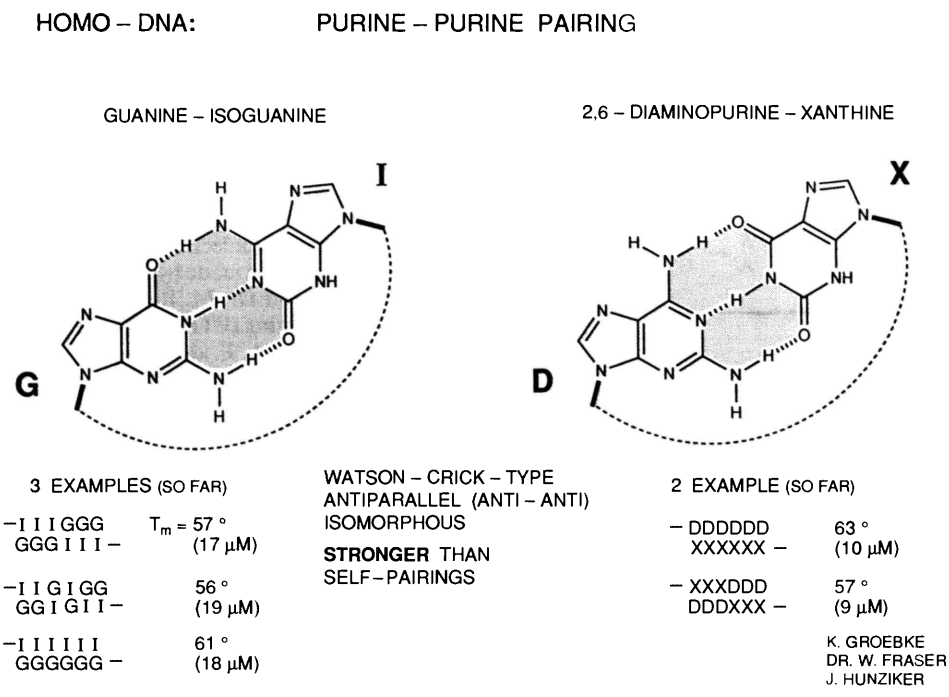


Figure 29

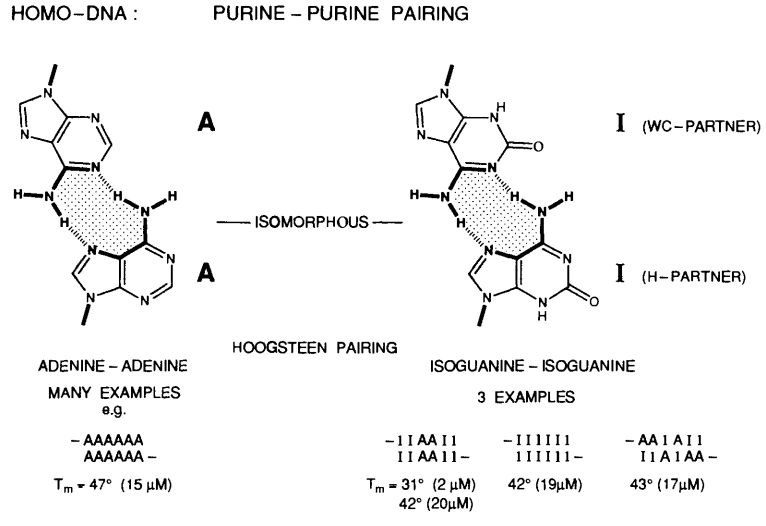


Figure 30

15 P. Decker, H. Schweer, and R. Pohlmann, *J. Chromatogr.*, 1982, **244**, 281.
16 F. Wöhler, *Poggendorfs Ann. Phys. Chem.*, 1828, **12**, 253.
17 H. Bauer, *Naturwissenschaften*, 1980, **67**, 1.
18 J. P. Ferris and W. J. Hagan, Jr., *Tetrahedron*, 1984, **40**, 1093.
19 U. Trinks, 'Zur Chemie der Aminopyrimidine', Thesis, ETH Zürich, No. 8368 (1987).
20 R. E. Eakin, *Proc. Natl. Acad. Sci. USA*, 1963, **49**, 360; L. E. Orgel and J. E. Sulston, in 'Prebiotic and Biochemical Evolution', ed. A. P. Kimball and J. Oró, North-Holland Publ. Co., Amsterdam–London, 1971, p. 89; H. B. White, *J. Mol. Evol.*, 1976, **7**, 101; G. A. M. King, *Biosystems* 1980, **13**, 23; C. M. Visser, *Bioorg. Chem.*, 1980, **9**, 411; *Origins of Life*, 1982, **12**, 165.
21 A. Eschenmoser, *Angew. Chem.*, 1988, **100**, 5; *Int. Ed. Engl.*, 1988, **27**, 5.
22 K. Koch, Thesis ETH Zürich, in preparation.
23 C. Strupp, Thesis ETH Zürich, in preparation.
24 D. W. Young, *Nat. Prod. Rep.*, 1986, 395.
25 R. Volk and A. Bacher, *J. Am. Chem. Soc.*, 1988, **110**, 3651.
26 G. W. E. Plaut, C. M. Smith, and W. L. Altworth, *Ann. Rev. Biochem.*, 1974, **43**, 899.
27 T. Rowan and H. C. S. Wood, *Proc. Chem. Soc.*, 1963, 21.
28 T. Paterson and H. C. S. Wood, *J. Chem. Soc., Perkin Trans. 1*, 1972, 1051.
29 D. H. G. Crout and J. A. Hadfield, *J. Chem. Soc., Chem. Commun.*, 1987, 742.
30 J. R. Knowles, *Nature*, 1991, **350**, 121.
31 S. Altman, *Angew. Chem.*, 1990, **102**, 735; *Int. Ed. Engl.*, 1990, **29**, 749 (Nobel lecture).
32 T. R. Cech, *Angew. Chem.*, 1990, **102**, 745; *Int. Ed. Engl.*, 1990, **29**, 759 (Nobel lecture).
33 R. Shapiro, *Origins of Life*, 1988, **18**, 71.
34 G. F. Joyce, *Cold Spring Harbor Symp. Quant. Biol.*, 1987, **52**, 41.
35 J. D. Bernal, 'The Origin of Life', Weidenfeld and Nicholson, London, 1967.
36 A. G. Cairns-Smith, 'Genetic Takeover and the Mineral Origin of Life', Cambridge University Press, 1982.
37 D. Müller, S. Pitsch, A. Kittaka, E. Wagner, C. E. Wintner, A. Eschenmoser, *Helv. Chim. Acta*, 1990, **73**, 1410.
38 G. Ksander, G. Bold, R. Lattmann, C. Lehmann, T. Früh, Yi-Bin Xiang, K. Inomata, H.-P. Buser, J. Schreiber, E. Zass, and A. Eschenmoser, *Helv. Chim. Acta*, 1987, **70**, 1115.
39 S. Drenkard, J. Ferris, and A. Eschenmoser, *Helv. Chim. Acta*, 1990, **73**, 1373.

HOMO – DNA : PURINE – PURINE PAIRING

MELTING TEMPERATURES (°C) OF HEXAMER – DUPLEXES:

15-2 μM
150mM
10mM Tris pH

dd-BBBBBB
BBBBBB-dd

dd-BBBBBB
BBBBBB-dd

dd-BB**B**BB**B**
BB**B**BB-dd

dd-BB**B**BB**B**
BBBBB-dd

		A	H	D	X	I	G
ADENINE	A	47					
HYPOXANTHINE	H	<10	—				
2,6-DIAMINOPURINE	D	41		36			
XANTHINE	X	14		**63**	—		
ISOGUANIN	I	43		39	16	42	
GUANINE	G	<15	<15	<15	18	**61**	38

~40 HOOGSTEEN (BIDENTATE)

~60 WATSON – CRICK (TRIDENTATE)

— NO PAIRING OBSERVED

KATRIN GROEBKE, MARKUS BOEHRINGER, HANS-JÖRG ROTH, JÜRG HUNZIKER, ULF DIEDERICHSEN, DR. WILLIAM FRASER, DR. CHRISTIAN LEUMANN

Figure 31

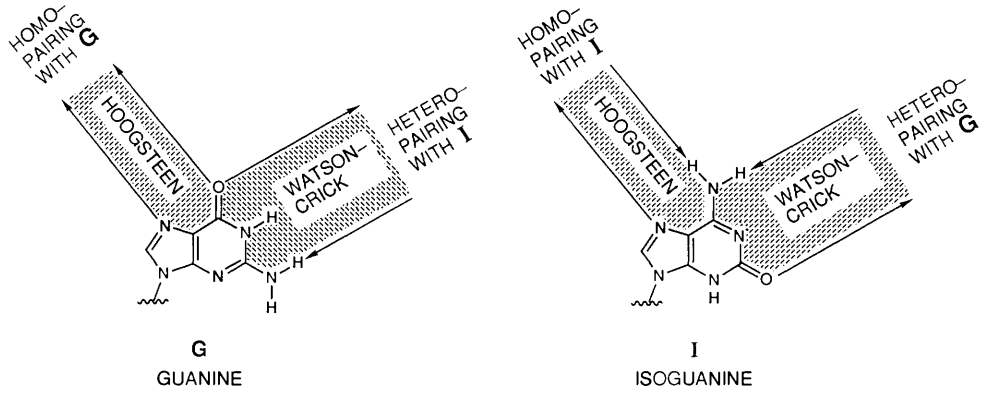


Figure 32 Dichotomic purine–purine pairing in hexose nucleic acids.

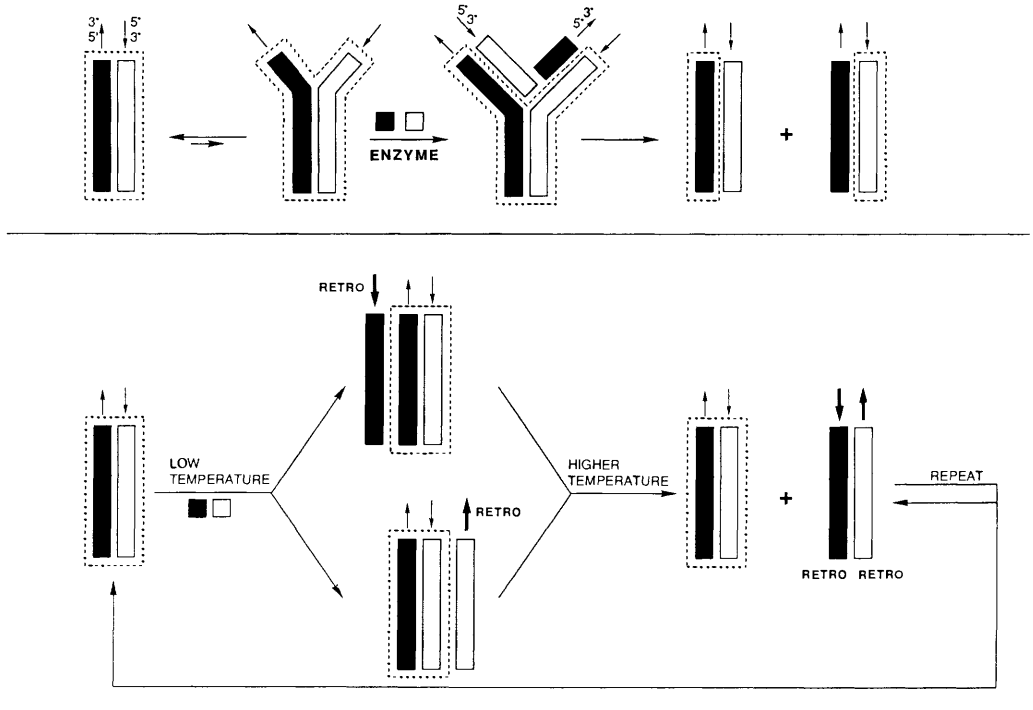


Figure 33

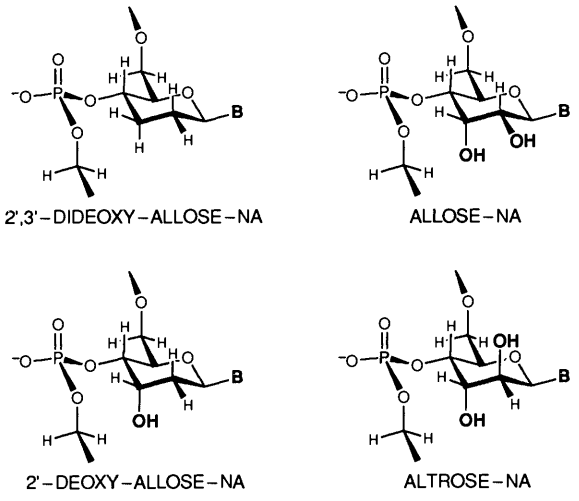


Figure 34

40 E. Wagner, Yi-Bin Xiang, K. Baumann, J. Gück, and A. Eschenmoser, *Helv. Chim. Acta*, 1990, **73**, 1391.
41 S. Pitsch (ETH), unpublished work, 1990.
42 A. Eschenmoser, *Nachr. Chem. Tech. Lab.*, 1991, **39**, 795.
43 A. Eschenmoser and M. Dobler, 'Warum Pentose- und nicht Hexose-Nukleinsäuren? Teil I. Einleitung und Problemstellung, Konformationsanalyse für Oligo-2',3'-dideoxy-glucopyranosyl-nukleotide ('Homo-DNS') sowie Betrachtungen zur Konformation von A- und B-DNS' (to appear in *Helv. Chim. Acta*, 1992).
44 M. Böhringer, H. J. Roth, J. Hunziker, M. Göbel, R. C. Krishnan, A. Giger, B. Schweizer, J. Schreiber, C. Leumann, A. Eschenmoser, 'Warum Pentose- und nicht Hexose-Nukleinsäuren? Teil II. Oligo-2',3'-dideoxyglucopyranosyl-nukleotide ("Homo-DNS"): Herstellung' (to appear in *Helv. Chim. Acta*, 1992).
45 J. Hunziker, H. J. Roth, M. Böhringer, A. Giger, M. Göbel, R. C. Krishnan, B. Jaun, C. Leumann, and A. Eschenmoser, 'Warum Pentose- und nicht Hexose-Nukleinsäuren? Teil III. Oligo-2',3'-dideoxy-glucopyranosylnukleotide ("Homo-DNS"): Paarungseigenschaften' (to appear in *Helv. Chim. Acta*, 1992).
46 H. J. Roth, C. Leumann, A. Eschenmoser, G. Otting, M. Billeter, K. Wüthrich, 'Warum Pentose- und nicht Hexose-Nukleinsäuren? Teil IV. Oligonukleotide aus 2',3'-Dideoxy-D-glucopyranose-Bausteinen ('Homo-DNS'): ^{1}H- ^{13}C- ^{31}P-NMR-spektroskopische Untersuchung von dd(AAAAATTTTT)$_2$' (to appear in *Helv. Chim. Acta*, 1992).
47 L. A. Marky and K. J. Breslauer, *Biopolymers*, 1987, **26**, 1601.
48 H. J. Roth, 'Homo-DNS: Herstellung, Paarungseigenschaften und Struktur von Adenin/Thymin-haltigen Oligonukleotiden'. Thesis, ETH Zürich No. 9591, (1992).
49 M. Böhringer, 'Synthese und strukturelle Charakterisierung von Adenin-Oligonukleotiden mit 2,3-Dideoxy-D- und L-Glycopyranosen als Zuckereinheiten', Thesis, ETH Zürich, No. 9377 (1991).
50 J. Hunziker, 'Homo-DNS: Herstellung und Paarungseigenschaften von Guanin-haltigen Oligonukleotiden'. Thesis ETH Zürich, in preparation (1992).
51 K. Groebke, W. Fraser, U. Diederichsen, Ling Peng, J. Hunziker, and C. Leumann (ETH), unpublished work, 1991.
52 K. J. Luebke and P. B. Dervan, *J. Am. Chem. Soc.*, 1989, **111**, 8733.
53 T. Inoue and L. E. Orgel, *Science*, 1983, **219**, 859; T. Inoue, G. F. Joyce, K. Grzeskowiak, L. E. Orgel, J. M. Brown, and C. B. Reese, *J. Mol. Biol.*, 1984, **178**, 669; W. S. Zielinski and L. E. Orgel, *Nature*, 1987, **327**, 346.
54 G. v. Kiedrowski, *Angew. Chem.*, 1986, **98**, 932; *Int. Ed. Engl.*, 1986, **25**, 932; G. v. Kiedrowski, B. Wlotzka, and J. Helbing, *Angew. Chem.*, 1989, **101**, 1259; *Int. Ed. Engl.*, 1989, **28**, 1235.
55 T. Tjivikua, P. Ballester, and J. Rebek Jr., *J. Am. Chem. Soc.*, 1990, **112**, 1249.
56 L. E. Orgel, *Cold Spring Harbor Symp. Quant. Biol.*, 1987, **52**, 9.
57 R. Fischer, A. Helg, A. Giger, and C. Leumann (ETH), unpublished work, 1991.
58 Work by Dr. R. Hammer (ETH).
59 Work by Dr. G. Rehberg, K. Groebke, and Dr. P. P. Adam.