Proc. Natl. Acad. Sci. USA
Vol. 85, pp. 1134–1135, February 1988
Evolution

An all-purine precursor of nucleic acids

(evolution/metabolism/bases/coenzymes)

GÜNTER WÄCHTERSHÄUSER

8000 Munich 2, Tal 29, Federal Republic of Germany

Communicated by Karl R. Popper, October 26, 1987

ABSTRACT **The theory is proposed that the pyrimidines in extant nucleic acids are postenzymatic substitutes for their isoelectronic and isogeometric position 3-bonded purine analogs xanthine and isoguanine, which were sibling products in a preenzymatic *de novo* purine pathway.**

Orgel's suggestion (1) that, in the evolution of life, the emergence of protein enzymes was preceded by the emergence of nucleic acids seems promising. Experimental attempts to demonstrate the possibility of an enzyme-free self-organization of RNA from prebiotic components have run into difficulties, however. Thus, Orgel concludes that RNA must have been preceded by a simpler ancestral nucleic acid (2). Such a precursor is proposed here.

The experimental difficulties appear to be due to the requirement of a pyrimidine base, since in the absence of a polymerase, pyrimidine nucleotides do not oligomerize onto a complementary template. But purine nucleotides do so oligomerize (3).

To circumvent these difficulties, the all-purine precursor base pairs shown in Fig. 1 are postulated. Each consists of one of the standard N9-bonded purines and one of the postulated N3-bonded purines xanthine and isoguanine. These base pairs are isoelectronic and isogeometric with the standard mixed purine–pyrimidine base pairs (Watson–Crick base pairs). It is my hypothesis that the Watson–Crick base pairs were preceded by one or both of the proposed all-purine base pairs. (Incidentally, hypoxanthine may substitute for guanine and 2-aminoadenine for adenine.)

My hypothesis is speculative, but not wholly because N3-bonded xanthine of unknown function has been found in prokaryote and eukaryote cells. Indeed, 3-ribosylxanthine 5′-phosphate has been demonstrated to be biosynthesized from xanthine and phosphoribosyl pyrophosphate with a uridine pyrophosphorylase (4), and 3-ribosylxanthine as well as 3-deoxyribosylxanthine are biosynthesized from xanthine and ribose 1-phosphate or deoxyribose 1-phosphate with a uridine phosphorylase (5). It is of interest that these phosphorylases ribosylate indiscriminately both uracil and position 3 of xanthine. Isoguanine and isoguanosine (N9-bonded) also have been isolated from natural sources (6, 7), but ribosylation in position 3 has not been demonstrated.

The lack of success of the attempts to demonstrate enzyme-free oligomerization of pyrimidine nucleotides has been well explained by the low stacking energies of the pyrimidine bases (3). With the base pairs proposed here, this problem disappears since only purines with high stacking energies are involved. Further, in analogy to the higher propensity for homocrystallization as compared with cocrystallization, it can be expected that an all-purine nucleic acid will exhibit a higher propensity for self-organization than that of mixed purine–pyrimidine structure.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

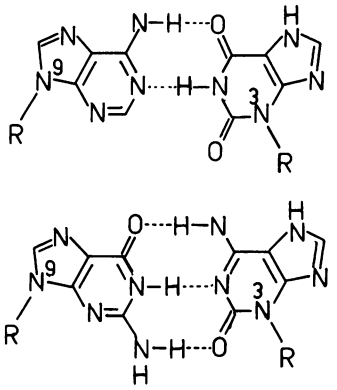

FIG. 1. Proposed base pairs. The upper base pair consists of a standard N9-bonded adenine and a proposed N3-bonded xanthine. The lower base pair consists of a standard N9-bonded guanine and a proposed N3-bonded isoguanine. R represents a backbone moiety derived from ribose, deoxyribose, or a ribose precursor.

The proposed base pairs cannot be accommodated within a standard A-RNA double-helix structure because of the sterical hindrance between N9 of the purine moiety and O4′ and O5′ of the ribose moiety. For this reason, a precursor nucleic acid with the all-purine base pairs must have had a more accommodating backbone structure. In this respect, it is of interest that DNA has a more versatile backbone structure than RNA. A still more flexible and accommodating backbone structure could be obtained with open-chain prochiral precursor nucleotides as has been suggested recently (8, 9).

In the mixed purine–pyrimidine structures of RNA and DNA, the stacking energy is highly variable and sequence-dependent. In the postulated all-purine structures, all sequences are expected to have a high stacking energy, and the constraints of a rigid all-purine stacking structure may prove to be the main factor determining the conformation of the nucleic acid backbone and also the base sequence, which perhaps may be regarded as an early place-holder of what later became the code.

It is a special merit of the postulated all-purine base pairs that an additional metabolic pathway to the N3-bonded purines does not have to be invoked. The extant *de novo* purine pathway is dependent on highly specific enzymes. An ancestral pathway with less-specific enzymes or even without enzymes (10) would necessarily produce both N3-bonded and N9-bonded purines in two parallel branch pathways as shown in Fig. 2. This branching of the pathway is due to the tautomerism of the formylglycinamidine ribonucleotide (FGAM). Both sibling pathways are strictly parallel in terms of reaction types and required reaction conditions.

The postulated base pairs suggest a revision in the account of the origin of life. After the emergence of a polymerase enzyme, the substitution of the N3-bonded purines by the isogeometric and isoelectronic pyrimidines should have become possible in spite of a looser stacking structure. Since

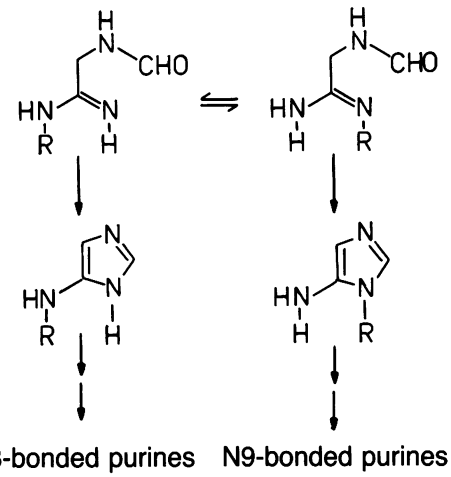

FIG. 2. Branching point in a proposed aboriginal *de novo* purine pathway: the tautomerism of the formylglycinamidine ribonucleotide giving rise to two sibling pathways, to N3-bonded and N9-bonded purines. R represents a moiety of ribose or a ribose precursor.

pyrimidines would be produced by a *de novo* pathway less energy-demanding than the *de novo* purine pathway, the substitution of the purines by the pyrimidines would have meant a considerable saving of energy with the extra benefit of greater structural versatility.

The hypothesis proposed here can be readily tested. Further, it satisfies a methodological criterion: that of an increased explanatory power compared with precursor hypotheses (11). Specifically, it offers explanations of the following three observed facts of the central metabolism: (*i*) the restriction of the nucleic acids to the four bases adenine, guanine, cytidine and uracil (or thymidine); (*ii*) the predominance of purine-related coenzymes; and (*iii*) the well-known striking difference between the *de novo* purine and pyrimidine pathways.

It is suggested that these facts can be explained with the aid of the proposed hypothesis, if we assume that the extant *de novo* purine pathway is a vestige of an aboriginal preenzymatic metabolism and that it emerged before the emergence of the genetic code and of genetically coded enzymes. Accordingly, the coenzymes, which emerged as catalysts in the evolution of this preenzymatic metabolism (e.g., ATP, GTP, FMN, FAD, NADH, CoA, coenzyme F_{420}, pterine coenzymes, and thiamine pyrophosphate), were derivatives of purines and purine precursors rather than of pyrimidines. The aboriginal *de novo* purine pathway resulted necessarily in N9-bonded as well as N3-bonded purines. All of these purines entered into the first nucleic acid structures. After the establishment of a genetic code and of coded enzymes, the *de novo* pyrimidine pathway emerged as a typical postenzymatic pathway. The two pyrimidines cytidine and uracil (or thymidine) came to be used in the nucleic acids simply because of the fact that only these could function as isoelectronic and isogeometric substitutes for the N3-bonded purines xanthine and isoguanine. It was only after the establishment of the *de novo* pyrimidine pathway that uracil and cytidine could enter as carriers into the more recent extensions of the sugar and phospholipid pathways. This historical explanation relies partly on selection and partly on chemical restriction.

I wish to express my gratitude to Prof. Sir Karl Popper for his help in the development and formulation of this hypothesis.

1. Orgel, L. E. (1968) *J. Mol. Evol.* **38,** 381–393.
2. Orgel, L. E. (1986) *Origins Life* **17,** 27–34.
3. Inoue, T. & Orgel, L. E. (1983) *Science* **219,** 859–862.
4. Hatfield, D. & Wyngaarden, J. B. (1964) *J. Biol. Chem.* **239,** 2580–2592.
5. Zimmerman, M. (1966) *J. Biol. Chem.* **241,** 4914–4916.
6. Cherbuliez, E. & Bernhard, K. (1932) *Helv. Chim. Acta* **15,** 464–471; 978–980.
7. Purrmann, R. (1940) *Annalen* **544,** 182–186.
8. Spach, G. (1984) *Origins Life* **14,** 433–437.
9. Schwartz, A. W. & Orgel, L. E. (1985) *Science* **228,** 585–587.
10. Franks, M., Green, C. P., Shaw, G. & Litchfield, G. J. (1966) *J. Chem. Soc C,* 2270–2274.
11. Popper, K. R. (1963) *Conjectures and Refutations* (Routledge & Kegan Paul, London), 4th Ed.