Proc. Natl. Acad. Sci. USA
Vol. 86, pp. 7054–7058, September 1989
Evolution

Modern metabolism as a palimpsest of the RNA world

STEVEN A. BENNER*, ANDREW D. ELLINGTON†, AND ANDREAS TAUER

*Laboratory for Organic Chemistry, Eidgenössische Technische Hochschule, CH-8092 Zurich, Switzerland; and †Department of Molecular Biology, Massachusetts General Hospital, Boston, MA 02114

Communicated by F. H. Westheimer, May 15, 1989

ABSTRACT **An approach is developed for constructing models of ancient organisms using data from metabolic pathways, genetic organization, chemical structure, and enzymatic reaction mechanisms found in contemporary organisms. This approach is illustrated by a partial reconstruction of a model for the "breakthrough organism," the last organism to use RNA as the sole genetically encoded biological catalyst. As reconstructed here, this organism had a complex metabolism that included dehydrogenations, transmethylations, carbon-carbon bond-forming reactions, and an energy metabolism based on phosphate esters. Furthermore, the breakthrough organism probably used DNA to store genetic information, biosynthesized porphyrins, and used terpenes as its major lipid component. This model differs significantly from prevailing models based primarily on genetic data.**

Since the discovery of self-splicing RNA (1), molecular biology has become the central focus of speculation concerning early forms of life. Many of these speculations consider genetic structure to the exclusion of most other biochemical data in modeling the "RNA world" (2–5). As discussed elsewhere, this narrow focus leads to interesting but often chemically and biologically implausible models (6–10).

We develop here an alternative approach for generating experimentally testable models of the RNA world, based on metabolic, structural, and mechanistic data from contemporary organisms. Several specific "paradigms," problem solutions covering individual topics in biochemical evolution, are constructed by this approach. These paradigms demonstrate the utility of this broader view and provide elements of a framework for interpreting the "historical" component of modern biochemistry (11).

We begin by assuming that life on earth passed through three episodes (Fig. 1) (12). In the first (the RNA world) (13), RNA was the only genetically encoded component of biological catalysts. The second episode began with the invention of translation-based synthesis of proteins in a "breakthrough organism," the first organism to contain a genetically encoded messenger RNA that directed the synthesis of a protein selectable for its catalytic activity. The third episode comprises the divergent evolution of the "progenote," the most recent common ancestor of all modern forms of life.[a]

This model views modern macromolecular catalysis as a "palimpsest" of an earlier metabolic state, with features that arose recently ("derived traits") superimposed upon features that are remnants of ancient life ("primitive traits"). (A palimpsest is a parchment that has been inscribed two or more times, with the previous texts imperfectly erased and therefore still partially legible.) To describe the biochemistry of these ancient organisms, we must first examine contemporary biochemical traits to distinguish ancient information from information added later. These descriptions are prerequisites for descriptions of the development of metabolism, the origin of translation, and other events that occurred in the RNA world.

If several descendants of an ancient organism can be inspected, a rule of "parsimony" can be used to model the biochemistry of the ancestral organism by extrapolation from the biochemistry of the descendant organisms. The most parsimonious model is one that explains the diversity in the modern descendants by a minimum number of independent evolutionary events. For the progenote, three independent lineages of descendants are known (archaebacteria, eubacteria, and eukaryotes). Thus, a biochemical trait present in all three can be assigned to the progenote. The assignment is strongest when (*i*) the trait is found in several representative organisms from each of the three kingdoms; (*ii*) assignments of homology in various branches of the progenotic pedigree are supported by high information content (preferably sequence data); and (*iii*) aspects of the trait serve no selected function in the modern world.[b] Such assignments are not absolute; if only some criteria are fulfilled, a weaker assignment can be proposed.

Parsimony cannot be similarly used to decide which traits in the progenote are vestiges of the breakthrough organism, as a biochemical description is possible for only a single descendant of the breakthrough organism (the progenote). Thus, chemical criteria are needed. A biochemical trait of the progenote can be assigned to the breakthrough organism most strongly when (*i*) RNA is involved in the trait, (*ii*) the involvement does not reflect the intrinsic chemistry of RNA, and (*iii*) substitution of another structural unit for the RNA unit could, on chemical grounds, provide similar or better biochemical performance.[c]

Using these rules, rRNA can be reliably placed in the progenote and from there in the breakthrough organism. Likewise, RNA cofactors (NAD^+, *S*-adenosylmethionine, CoA, ATP, FAD) all contain fragments of RNA that are present in all lineages descendant from the progenote (20–22); these RNA cofactors can be assigned to the progenote. However, the RNA portions of the cofactors are not intrinsic to the chemical performance of the cofactor.[c] Thus, on

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.

Abbreviation: RNR, ribonucleotide reductase.

[a] As rRNA and tRNA molecules are homologous in all kingdoms, and as rRNA must have been present in the breakthrough organism, the progenote must have been a descendant (or perhaps a contemporary) of the breakthrough organism.

[b] A chemically unique solution to a particular biochemical problem can arise independently more than once. Further, assignments must recognize the possibility of lateral transfer of genetic information between members of divergent branches of an evolutionary tree, a process that occurs with unknown frequency (14–17).

[c] RNA serving a role that could be better performed by proteins is unlikely to arise in a world with proteins; RNA performing roles for which it is intrinsically chemically suited could arise at any time. To evaluate the "intrinsic chemical suitability" of RNA for solving a particular biochemical problem, alternative solutions not involving RNA are compared by using a degree of chemical intuition. For example, pyrophosphate is as good a phosphoryl donor as the RNA cofactor ATP in several kinases (18); *S*,*S*-dimethylthioacetate is as good a methyl donor in enzymes evolved to accept it as the RNA cofactor *S*-adenosylmethionine (19).

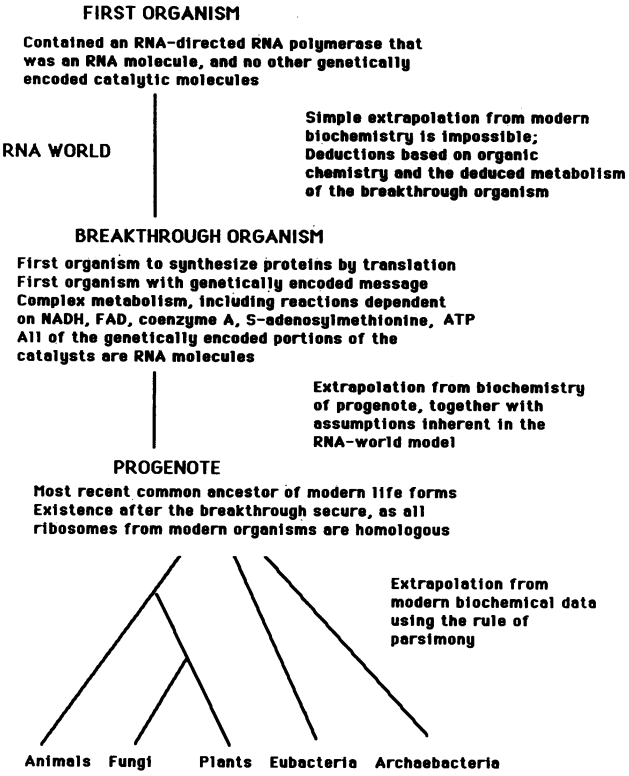

FIG. 1. Stages in the origin of modern life.

chemical grounds, RNA cofactors can be assigned to the breakthrough organism as well (20–22).[d]

In contrast, self-splicing RNA has such an unusual phylogenetic distribution that it cannot be reliably placed in the progenote (23, 24, 57). Furthermore, self-splicing can plausibly be viewed as an adaptive trait for selfish DNA and, therefore, the product of recent evolution. Finally, self-splicing elegantly uses chemistry intrinsic to RNA to solve the problem posed by intron removal. [Non-self-catalyzed splicing, involving small nuclear ribonucleoproteins (snRNPs), is more probably a vestige of the RNA world (25). To the extent that snRNPs can be placed in the progenote (difficult at present, but see ref. 26), they can be assigned to the breakthrough organism. A still stronger argument can be made for RNase P, a ribonucleoprotein in both eukaryotes and prokaryotes (27, 28).] Each fact challenges the belief that self-splicing RNA found in the modern world is a direct, functional descendant of self-splicing RNA in the RNA world.[e]

A model for the breakthrough organism consists of a collection of statements about it, statements that can be examined for internal consistency and that suggest experimentally testable predictions. The considerations described above allow us to begin to construct a model that assigns biochemical traits to the breakthrough organism.[f] At the very least, the presence of many cofactors in the breakthrough organism implies that the breakthrough organism was metabolically complex and contained ribonucleotide enzymes that catalyzed redox reactions, transmethylations, carbon-carbon bond formation, an energy metabolism based on phosphate anhydrides, and carbon-carbon bond forming and breaking reactions. [This is not chemically unreasonable; almost any functionalized macromolecule catalyzes reactions at some rate, although RNA is undoubtedly a poorer catalyst than proteins for most reactions (8).]

We now develop specific arguments that describe in greater detail the metabolism of the last organism to use RNA as the sole genetically encoded biological catalyst ("riboorganism").

The Breakthrough Organism Used DNA

We consider the following facts, together with implications drawn from these facts:

(*i*) Several DNA-dependent enzymes can be reliably assigned to the progenote. In particular, archaebacteria, eubacteria, and eukaryotes contain DNA-dependent RNA polymerases that are quite possibly homologous (30). This implies that the progenote contained a DNA-dependent RNA polymerase and, therefore, had DNA.

(*ii*) For organisms with DNA, endogenous ribonucleotide reductase (RNR), the enzyme that converts ribonucleotides into deoxyribonucleotides, appears to confer strong selective advantage.[g] Trypanosomes and some viruses, organisms that are parasitic in most other respects, carry their own RNR (33). Therefore, the presence of DNA implies the presence of a RNR, while the absence of a RNR implies the absence of DNA. In particular, the conclusion (see above) that the progenote contained DNA implies that the progenote contained a RNR.

(*iii*) However, three (and possibly four) mechanistically different RNRs are known in the modern world (34, 35). This implies that the RNRs in different kingdoms are not homologous, and makes it impossible to assign by parsimony a chemical reaction mechanism to any RNR presumed to have been present in the progenote.[h] Indeed, in the absence of evidence for points *i* and *ii*, it would be reasonable to suggest that the progenote did not contain any RNR and that the mechanistically different RNRs in the modern world reflect the independent origin of several types of RNR after the divergence of the progenote.

Two alternative resolutions of this apparent contradiction can be considered. First, the progenotic RNR may have been a protein homologous to one of the modern RNRs. In this case, the two (or perhaps three) other types of RNR arose by protein-for-protein replacement events after the divergence of the progenote. Alternatively, the RNR in the progenote may

[d] It is conceivable that only a few RNA cofactors emerged in the RNA world, with the rest emerging after the breakthrough and incorporating RNA "handles" to allow protein binding domains used for the early cofactors to be used for the late ones as well. This model is inconsistent with what is generally known about the adaptability of protein binding sites, the structure of biotin, and other modern biochemical details (11). Furthermore, the use of a cofactor implies its biosynthesis, which requires the invention of more binding domains for the biosynthesis of the RNA "handle" than would be saved by reusing binding sites evolved for other cofactors.

[e] The fact that self-splicing RNA can disproportionate small nucleotide oligomers to produce larger ones does not suggest that the RNA molecule is a functional descendant of the first self-replicating RNA molecule. Macromolecular catalysts often catalyze chemical reactions analogous to their selected function at a low level. Furthermore, even if the self-splicing RNA is a direct functional descendant of the original replicase, the replicase activity would have long since disappeared by evolutionary drift after it ceased to serve a selected function.

[f] Interesting analogous problems on unrelated subject matter can be found in historical linguistics (29) and paleontology.

[g] This route to deoxyribonucleotides is a remarkable one, given the chemically plausible alternative route to DNA involving an aldol condensation between glyceraldehyde and acetaldehyde to form 2-deoxyribose followed by the introduction of a base. Both reactions are known in modern enzymology (31, 32).

[h] The sequence similarity displayed in the peptides containing the redox active cysteines in the B-12 and non-B-12 RNRs (35) could indicate homology or sequence convergence, with quite different implications for this model. Here, complete sequences ofseveral B-12-dependent RNRs might be decisive.

have been a ribonucleotide enzyme. In this case, contemporary RNRs all arose by protein-for-RNA replacement events (11).

The replacement of ribonucleotide enzymes by protein enzymes after the invention of translation is expected to occur mostly via a "deletion–replacement" process. A ribonucleotide enzyme is first deleted, focusing selective pressure on the evolution of a protein enzyme. An evaluation of the relative plausibility of these alternative resolutions is based on an evaluation of the relative facility of protein-for-protein replacement events compared to protein-for-RNA events. The frequency of deletion–replacement events depends on two factors: (*i*) the extent to which the deletion is lethal, and (*ii*) the availability of genes for biological macromolecules that, after minor alteration, can assume the deleted function. Deletions are more likely to be lethal if they disrupt a pathway for a metabolite that cannot be obtained in the diet. Macromolecules are more likely to be able to assume the role of the deleted enzyme if they already catalyze a closely related reaction. [Protein-for-protein deletion–replacement events are well known in evolution, where the deletion is tolerated by natural selection (11).]

RNRs are difficult to replace for both reasons. First, a deletion is presumed to be semilethal (point *ii*). Second, even in modern metabolism, few proteins catalyzing reactions similar to that found in RNR are available to replace a deleted RNR. Finally, there appears to be no selective advantage for replacing an enzyme with one mechanism by another. Thus, multiple protein-for-protein replacement events seem implausible for RNR.[i] To the extent that a proteinaceous RNR was present in the progenote, we would expect it to have survived to the modern day in all lineages.

However, in view of the (presumed) advantage of protein over RNA as a catalyst, replacement of RNA by protein is more plausible. Thus, the model for the progenote includes a RNR. If we assume (see above) that new catalytic RNA arose infrequently after the invention of translation, the ribonucleotide enzyme in the progenote should be a descendant of a ribosomal RNR originating in the RNA world. This implies that the breakthrough organism contained a RNR and, therefore, contained DNA.[j]

While this argument is based on parsimony and chemical structure, it is also plausible on general chemical and evolutionary grounds. As with any single metabolic step, ribonucleotide reduction requires less information than translation and should have emerged before translation, especially as the chemical attributes of DNA that make it superior for storing genetic information in the modern world are also likely to have been useful in the RNA world.[k]

[i]Similar arguments explain why a ribonucleotide enzyme such as a RNR survived during the time between the breakthrough and the progenote. Indeed, it is conceivable that some organisms in the modern world have RNRs (or other enzymes involved in DNA biosynthesis) containing catalytic RNA as vestiges of the RNA world.

[j]The absence of one of points *i–iii* prevents this conclusion. For example, the presence of lysine in homologous proteins from all three branches of the tree implies that lysine was used in the progenote. However, the presence of two pathways for biosynthesizing lysine in modern organisms does not imply that lysine was biosynthesized by ribonucleotide enzymes in the progenote, as an analogue to point *ii* does not hold for lysine; many organisms obtain lysine in the diet. Thus, multiple pathways to lysine can be explained by deletion–replacement events, in which lysine was obtained in the diet during the time when the pathway was deleted, and the aminoadipate pathway (because of its chemical similarity to steps in the citric acid cycle) can plausibly be viewed as a result of "pathway capture," whereby enzymes involved in the citric acid cycle underwent mutation to accommodate substrates with slightly different structures, after a biosynthetic pathway reacquired selective advantage.

[k]The information needed before a macromolecule can serve a particular selectable function, estimated from that required for proteins that perform this function in the modern world, is useful for estimating the order of appearance of function in evolution, where functions requiring less information appear first. In organisms with a complex metabolism, once complex function is established for one function, adaptation of that first catalyst to another chemically related function is then facile. We agree with other authors that pieces of the translation machinery must have served another selectable function before serving in the biosynthesis of proteins.

The Breakthrough Organism Biosynthesized Tetrapyrroles

We consider the following facts and their implications:

(*i*) Two pathways exist for the synthesis of 5-aminolevulinate as the first step in the biosynthesis of tetrapyrroles. One (the Shemin pathway) involves a chemically elegant condensation of succinyl-CoA and glycine dependent on a pyridoxal cofactor. The other (the C5 pathway) involves the reduction of an ester of glutamic acid and RNA, followed by rearrangement of glutamate semialdehyde to give aminolevulinic acid (36).

(*ii*) Esters are not intrinsically advantageous intermediates in the biosynthesis of aldehydes from carboxylic acids; the reduction of anhydrides with phosphoric acid is chemically preferable and is the route most used in modern metabolism. Even if esters were desirable intermediates for such reactions, esters with RNA molecules seem to offer no intrinsic advantages over esters with simpler alcohols. Thus, if the C5 pathway can be placed in the progenote, the involvement of RNA in the pathway strongly argues that the C5 pathway (and its products) originated in the RNA world.

(*iii*) The C5 pathway is used to synthesize chlorophyll in photosynthetic eukaryotes and many eubacteria (37) and to synthesize B-12 and factor F430 in archaebacteria (38). Thus, the C5 pathway can be assigned to the progenote.

(*iv*) The Shemin path is found in two kingdoms descendant from the progenote—eukaryotes and eubacteria. Sequence evidence suggests that 5-aminolevulinate synthetases from chicken, yeast, and *Bradyrhizobium japonicum* are homologous (39). In contrast, limited biochemical evidence suggests that some mammalian synthetases in this pathway may not be homologous (40). Further work is necessary to determine the relationship between Shemin pathways in different kingdoms.

The presence in the progenote of an RNA molecule participating in the synthesis of 5-aminolevulinate, in a role not uniquely reflecting the chemistry intrinsic to RNA, supports an assignment of a ribonucleotide enzyme synthesizing 5-aminolevulinate to the breakthrough organism. The role of this intermediate in the metabolism of the breakthrough organism remains uncertain, although a reasonable hypothesis is that it was used in the biosynthesis either of chlorophyll for photosynthesis or of B-12 for methanogenesis.[l] Both

[l]B-12 is an RNA cofactor used in all three kingdoms and is perhaps formed prebiotically (41); an assignment of B-12 to the RNA world is strong. Interestingly, ribosomal proteins apparently are used in the assembly of B-12 in *Escherichia coli* (42). However, apparently the only role for B-12 sufficiently central to explain (in the presence of pyridoxal) the conservation of vestigial RNA in its biosynthesis is in methanogenesis. This implies that the C5 pathway was conserved in a continuous line of methanogens extending from the progenote back to (and perhaps including) the breakthrough organism. Arguing against this picture is the fact that methanofuran, coenzyme M, and factor F430, cofactors in methanogenesis, are not RNA cofactors. Methanopterin and deazaflavin contain a ribose phosphate fragment but no base (43). If RNA cofactors indicate that a pathway arose in the RNA world, these suggest that methanogenesis arose after the breakthrough. Conversely, a continuous line of photosynthetic organisms from the progenote back to the breakthrough organism is suggested by the role for the C5 pathway in chlorophyll biosynthesis. Furthermore, photosynthesis (of some sort) is expected to have emerged in the RNA world, as photosynthesis requires less information than translation, and organisms with a complex metabolism would rapidly exhaust the high-energy molecules created abioti-

roles are sufficiently central to metabolism of the respective organisms that deletion–replacement events are likely to be slow, explaining the conservation of the RNA cofactor for tetrapyrrole biosynthesis in modern methanogens and photo synthetic organisms.

A postulate that oxygenic photosynthesis did or did not arise in the RNA world is critical for assigning geological time to events in Fig. 1 (8). Oxygenic photosynthesis arose ≈2.5 billion years ago. If photosynthesis originated in the RNA world, the breakthrough must have occurred more recently than 2.5 billion years ago. This would imply that the oldest microfossils are fossils of riboorganisms.

The Breakthrough Organism Did Not Synthesize Fatty Acids

We consider the following facts and their implications:

(*i*) Fatty acid synthase complexes from different organisms have different quaternary structures, stereospecificities, substrate specificities, and mechanisms (45). Archaebacteria do not seem to synthesize fatty acids at all (46).[m] Thus, as with ribonucleotide reduction, it is difficult to place fatty acid biosynthesis in the progenote.

(*ii*) However, unlike RNR, homologous enzymes are not yet known in the three kingdoms that use fatty acids as substrates. Therefore, there is no independent evidence that fatty acids were synthesized in the progenote.

(*iii*) The biosynthesis of fatty acids involves biotin and acyl carrier protein (ACP). Both almost certainly did not arise before translation. ACP is a product of translation; it must have emerged after the breakthrough. Chemical and structural features of biotin, reviewed a decade ago by Visser and Kellogg (22), strongly suggest that biotin arose after protein catalysts.

These arguments suggest that fatty acid biosynthesis arose after the breakthrough. However, as lipids of some sort seem to be essential for living cells, an absence of fatty acid biosynthesis in the breakthrough organism creates a compelling need for an alternative source of lipids. This need can be satisfied by terpenes. Conversely, a successful assignment of terpenoid biosynthesis to the breakthrough organism removes a compelling need for fatty acid synthesis there.

The Breakthrough Organism Synthesized Terpenes

We consider the following facts and their implications:

(*i*) Higher terpenes are biosynthesized via similar routes in all three kingdoms, implying that the progenote biosynthesized terpenes, especially higher terpenes (di- and triterpenes).

(*ii*) In the eubacterium *Rhodopseudomonas acidophila*, membranes contain terpenoids covalently joined to RNA fragments, which serve as the polar part of the amphiphilic lipid molecule (50).

(*iii*) However, RNA is not uniquely suited as a polar group in this capacity. Indeed, in most lipids, polar components are not RNA. This suggests that RNA-conjugated lipids were present in the breakthrough organism.

The assignment of terpene biosynthesis to the progenote is stronger than the corresponding assignment of RNA-dependent tetrapyrrole biosynthesis. However, the involvement of vestigial RNA in terpene chemistry is less substantive than in tetrapyrrole biosynthesis; a hopanoid–RNA conjugate is presently known in only a single lineage. Thus, it is more difficult to place terpene biosynthesis in the breakthrough organism. However, terpenoids are themselves constituents of chlorophyll. If chlorophyll can be placed in the breakthrough organism, this strengthens the assignment of terpene biosynthesis to the RNA world. Conversely, if terpene biosynthesis can be assigned to the breakthrough organism, this provides in RNA metabolism an element necessary for the biosynthesis of chlorophyll.

Comparison with Other Models

This is not the first discussion of the RNA world, or the first to appreciate key elements, including the significance of RNA cofactors (2–13, 20–22). However, a model that considers formally the metabolic, structural, and mechanistic features of contemporary organisms differs from earlier models in several ways.

Alternative models nearly always view the period of RNA catalysis as short, the catalytic potential of RNA feeble, and the invention of translation machinery early (2–5, 14, 15, 20–22, 51–56). For example, Darnell and Doolittle (3) recently argued that translation, transcription, reverse transcription, the genetic code, and introns all arose before the synthesis of organic molecules and heterotrophy. Yet, it is difficult to imagine the evolution of an organism capable of assembling the information needed for translation without having, for example, a renewable source of high-energy phosphates or biosynthetic routes to sugars. Futhermore, the information needed for enzymes catalyzing individual metabolic steps is far less than that needed for translation machinery selectable for its ability to produce proteins.

Similarly, some investigators argue that there is "compelling evidence" that life evolved from an RNA world to a ribonucleoprotein (RNA–protein) world to a DNA world (4). However, chemical and metabolic considerations, many discussed above, suggest that DNA emerged before proteins.

Nevertheless, the fundamental disagreement between many alternative models and the one proposed here is methodological. Today, model builders often begin with an arbitrary assumption that a particular detail of the genetic

---

cally. Of course, the conservation of the RNA in the biosynthesis of 5-aminolevulinate can also be explained by assuming that pyridoxal arose late. Interestingly, pyridoxal is not an RNA cofactor, is not obviously a product of prebiotic chemistry, and its biosynthesis lacks characteristics expected for pathways originating in the RNA world (44). Should the biosynthesis and use of pyridoxal in archaebacteria be demonstrated, and (especially) if the enzymes involved are homologous across the three kingdoms, this would permit an argument by parsimony that pyridoxal arose before the progenote, strengthening the case that either methanogenesis or photosynthesis originated in the RNA world. Finally, the RNA involved in the biosynthesis of 5-aminolevulinate is quite similar to glutamate tRNA involved in translation, suggesting that one process was the precursor for the other. If the C5 pathway predates translation, the suggestion is obvious that glutamate tRNA needed for translation was already present and serving a different metabolic role before the breakthrough. The possibility that the C5 RNA cofactor arose from a species that first played a role in translation runs counter to the chemical argument presented in the text.

[m]Isolated reports suggesting that halobacteria contain a straight-chain fatty acid synthetase that incorporates labeled malonyl-CoA into palmitic acid (47) would, if correct, change significantly this argument; either the progenote did synthesize fatty acids, or halobacteria are incorrectly classified with other archaebacteria, as recently suggested by Lake (48). Reproducing the published work, we have found that a small fraction of label from malonyl-CoA fed to extracts of *Halobacter cutirubrum* does indeed appear in a fraction that behaves as a hydrophilic carboxylic acid, as reported. However, careful recrystallization of the phenacyl derivative of this radiolabeled compound with the derivative of palmitic acid as carrier shows (within experimental error) that this species is not labeled palmitic acid. Furthermore, <0.5% of the radioactivity cochromatographed with the derivative of palmitic acid. Details of this work will be reported elsewhere. Independently, others apparently have found no evidence for fatty acid biosynthesis in *Halobacter halobium* (49). However, as these results cannot entirely rule out the biosynthesis of straight-chain fatty acids in archaebacteria, further investigation is warranted. This is another example of how model building is intimately connected with and dependent on experimental evidence.

structure of a modern organism is primitive, disregarding parsimony as a tool for evaluating this assumption and disregarding possible adaptive significance of the trait. This genetic detail then becomes the starting point for an extended extrapolation to structures and in organisms that lived long before the progenote. Such an approach fails to guard against the possibility of recent evolutionary innovation, especially strong if the trait is functional, and especially strong if it is found in highly adaptable organisms (including viruses).

Conclusions

We have shown how structural, metabolic, and mechanistic information drawn from contemporary organisms allows the formulation of an interrelated set of statements about the organism that carried the first genetically encoded mRNA. Because the statements are interrelated, the model can be tested for logical consistency. Because they are dependent on assumptions regarding homology, behavior, and function in modern biological macromolecules, they can be tested experimentally by studies of these macromolecules. Thus, evaluation of these models is possible by methods other than a simple subjective view of their elegance.

In part, the success of such model building depends on a realistic choice of goals. We do not attempt at this stage to describe the origin of life or the origin of translation, events that occurred in the RNA world and whose vestiges largely disappeared as the result of two evolutionary "bottlenecks." Rather, the model building seeks first to reach a more attainable goal: describing the last riboorganism. While a bit less glamorous, the breakthrough organism as reconstructed by this process is not uninteresting from a metabolic point of view.

Nevertheless, reliable models of organisms at both ends of the RNA world, the breakthrough organism and the first riboorganism (developed from studies on prebiotic chemistry), should provide a solid foundation for future efforts to develop a picture of the RNA world itself. Although this process might be viewed as plodding, we believe that it offers the best opportunity for progress in understanding this period of natural history.

We are indebted to many colleagues, including Prof. K. Bloch, Prof. F. Westheimer, Prof. L. Orgel, and several anonymous referees, who read early versions of this manuscript and offered many helpful suggestions for improving it. We are also indebted to Liming Ge for assisting in the work with *H. halobium*.

1. Cech, T. R., Zaug, A. J. & Grabowski, P. J. (1981) *Cell* **27**, 487–496.
2. Weiner, A. M. & Maizels, N. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 7383–7387.
3. Darnell, J. E. & Doolittle, W. F. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 1271–1275.
4. Weiner, A. M. (1988) *Cell* **52**, 155–157.
5. Alberts, B. M. (1986) *Am. Zool.* **26**, 781–796.
6. Benner, S. A. & Ellington, A. D. (1987) *Nature (London)* **329**, 295–296.
7. Benner, S. A. & Ellington, A. D. (1988) *Nature (London)* **332**, 688–689.
8. Benner, S. A., Allemann, R. K., Ellington, A. D., Ge, L., Glasfeld, A., Leanz, G. F., Krauch, T., MacPherson, L. J., Moroney, S., Piccirilli, J. A. & Weinhold E. (1987) *Cold Spring Harbor Symp. Quant. Biol.* **52**, 53–63.
9. Benner, S. A. (1988) in *Redesigning the Molecules of Life*, ed. Benner, S. A. (Springer, Heidelberg), pp. 115–175.
10. Maizels, N. & Weiner, A. M. (1987) *Nature (London)* **330**, 616.
11. Benner, S. A. & Ellington, A. D. (1988) *CRC Crit. Rev. Biochem.* **23**, 369–426.
12. Orgel, L. E. (1968) *J. Mol. Biol.* **3**, 381–393.
13. Gilbert, W. (1986) *Nature (London)* **319**, 818.
14. Carlson, T. A. & Chelm, B. K. (1986) *Nature (London)* **322**, 568–570.
15. Leunissen, J. A. M. & de Jong, W. W. (1986) *J. Mol. Evol.* **23**, 250–258.
16. Syvanen, M. (1985) *J. Theor. Biol.* **112**, 333–343.
17. Lewin, R. (1985) *Science* **227**, 1020.
18. Siu, P. M. L. & Wood, H. G. (1962) *J. Biol. Chem.* **237**, 3044–3051.
19. Cantoni, G. L. (1962) *Methods Enzymol.* **5**, 743–748.
20. White, H. B., III (1976) *J. Mol. Evol.* **7**, 101–104.
21. White, H. B., III (1982) in *The Pyridine Nucleotide Coenzymes*, eds. Everse, J., Anderson, B. & Yu, K.-S. (Academic, New York), pp. 1–17.
22. Visser, C. M. & Kellogg, R. M. (1978) *J. Mol. Evol.* **11**, 171–187.
23. Sogin, M. L., Ingold, A., Karlok, M., Nielsen, H. & Engberg, J. (1986) *EMBO J.* **5**, 3625–3630.
24. Michel, F. & Dujon, B. (1986) *Cell* **46**, 323.
25. Mattaj, I. W. (1984) *Trends Biochem. Sci.* **9**, 434–437.
26. Inouye, M. & Delihas, N. (1988) *Cell* **53**, 5–7.
27. Nichols, M., Soell, D. & Willis, I. (1988) *Proc. Natl. Acad. Sci. USA* **260**, 1379–1383.
28. Guerrier-Takada, C., van Belkum, A., Pleij, C. W. A. & Altman, S. (1988) *Cell* **53**, 267–272.
29. Lehmann, W. P. (1973) in *Historical Linguistics* (Holt, Reinhart, and Winston, New York).
30. Gropp, F., Reiter, W. D., Sentenac, A., Zillig, W., Schnabel, R., Thomm, M. & Stetter, K. O. (1986) *System. Appl. Microbiol.* **7**, 95–101.
31. Hoffee, O. M., Rosen, O. M. & Horecker, B. L. (1965) *J. Biol. Chem.* **240**, 1512–1516.
32. McElwain, M. C. & Pollack, J. D. (1987) *J. Bacteriol.* **169**, 3647–3653.
33. Ator, M. A., Stubbe, J. & Spector, T. (1986) *J. Biol. Chem.* **261**, 3595–3598.
34. Hogenkamp, H. P. C., Follmann, H. & Thauer, R. K. (1987) *FEBS Lett.* **219**, 197–201.
35. Lin, A.-N. I., Ashley, G. W. & Stubbe, J. (1987) *Biochemistry* **26**, 6905–6909.
36. Huang, D.-D., Wang, W.-Y., Gough, S. P. & Kannangara, C. G. (1984) *Science* **225**, 1482–1484.
37. Kannangara, C. G., Gough, S. P., Bruyant, P., Hoober, J. K., Kahn, A. & von Wettstein, D. (1988) *Trends Biochem. Sci.* **13**, 139–143.
38. Friedmann, H. C., Thauer, R. K., Gough, S. P. & Kannangara, C. G. (1987) *Carlsberg Res. Commun.* **52**, 363–371.
39. McClung, C. R., Somerville, J. E., Guerinot, M. L. & Chelm, B. K. (1987) *Gene* **54**, 133–139.
40. Nandi, D. L. (1978) *Arch. Biochem. Biophys.* **188**, 266–271.
41. Eschenmoser, A. (1988) *Angew. Chem. Int. Ed. Engl.* **27**, 5–39.
42. Pezacka, E. & Walerych, W. (1981) *Biochim. Biophys. Acta* **678**, 300–310.
43. Hauska, G. (1988) *Trends Biochem. Sci.* **13**, 2–4.
44. Hill, R. E., Iwanow, A., Sayer, B. G., Wysocka, W. & Spenser, I. D. (1987) *J. Biol. Chem.* **262**, 7463–7471.
45. McCarthy, A. D. & Hardie, D. G. (1984) *Trends Biochem. Sci.* **9**, 60–63.
46. Ross, H. N. M., Collins, M. D., Tindall, B. J. & Grant, W. D. (1981) *J. Gen. Microbiol.* **123**, 75–80.
47. Pugh, E. L., Wassef, M. K. & Kates, M. (1971) *Can. J. Biochem.* **49**, 953–957.
48. Lake, J. A. (1986) *Nature (London)* **319**, 626.
49. Thurl, S. (1986) Dissertation (Ludwig-Maximilians-Universitaet, Munich), p. 3.
50. Neunlist, S. & Rohmer, M. (1985) *Biochem. J.* **228**, 769–771.
51. Visser, C. M. *Origins Life* **14**, 291–300.
52. Woese, C. R. (1967) *The Origins of the Genetic Code* (Harper & Row, New York).
53. Crick, F. H. (1968) *J. Mol. Biol.* **38**, 367–380.
54. Kuhn, H. (1972) *Angew. Chem.* **84**, 838–862.
55. Lewin, R. (1986) *Science* **231**, 545.
56. Orgel, L. E. (1986) *J. Theor. Biol.* **123**, 127–149.
57. Lang, F. B. (1984) *EMBO J.* **3**, 2129–2136.