Perspective

Some biochemical thoughts on the RNA world

Konrad Bloch

Ten years ago, Walter Gilbert contemplated "an RNA world containing only RNA molecules that serve to catalyze the synthesis of themselves" [1]: in other words, self-replicating systems that function in the absence of protein and DNA.

A recent comprehensive monograph describes the current understanding of the above concept and surveys recent progress concerning the biochemistry and molecular biology of RNA [2]. Here I highlight some facts about the role of the four RNA bases in contemporary intermediary metabolism, and then offer some thoughts on what implications these facts have for the evolution of our present DNA/protein-based world from the postulated RNA world, a perspective rarely mentioned in the discussion of the RNA world's central themes such as replication.

We must start by looking at research at the beginning of the second half of this century, when biochemists began to focus on the monomeric precursors of the larger cell constituents: fats, carbohydrates, proteins and nucleic acids.

Before the advent of isotopic tracers, these syntheses were thought to be simply a reversal of the corresponding lytic processes. How naive, in retrospect! Eventually, in the 1950's, these views became untenable, not because they were rejected as unreasonable but because of chance discoveries pointing to obligatory monomer activation by nucleotides and their derivatives prior to polymerization. The critical event in several of these chance discoveries was reagent contamination.

Accidental discoveries

Although it was recognized early on as the universal source of chemical energy, ATP, commercial or home-made, proved to be of dubious purity. No one initially suspected that preparations of ATP might contain other nucleoside triphosphates. This unwarranted confidence came to an end in the 1950's with dramatic consequences.

At the time Kornberg and Pricer [3] and Kennedy's laboratory [4] independently investigated the biosynthesis of phospholipids, but with divergent results. Both groups examined the energy-dependent conversion of precursors to phosphatidylcholine.

In the Kornberg laboratory, amorphous ATP (Pabst) was used as a reagent. Somewhat later, Kennedy and Weiss repeated and confirmed the results obtained in Kornberg's laboratory, but in addition they tested samples of the newly available crystalline ATP. These ATP specimens gave entirely negative results, pointing to a contaminant in the amorphous ATP as the source of the activity; this was later shown to be CTP. The cytosine-containing nucleotide proved to be the energy source that nature had selected not only for the synthesis of phosphatidylcholine, but also for phosphatidylethanolamine and other membrane phospholipids.

The discovery of GTP as an energy source closely resembled the accidental encounter of CTP. In studies in the 1970's on a glycogen-sensitive adenylic cyclase from rat liver membranes [5], Martin Rodbell found that ATP was required for activation, but in unusally high concentrations. Moreover, the responses were variable depending on the source of ATP. Aware of the experiences of Kornberg and Kennedy, Rodbell traced the variable activities of ATP samples to contamination with GTP.

Equally inadvertent, the discovery of the importance of uridine triphosphate came about when Luis Leloir found a thermostable cofactor necessary for the enzymatic breakdown of lactose [6]. By chance, his coworker Caputto found a published ultraviolet spectrum of uridine that proved to be identical with that of the cofactor needed for lactose breakdown, now known as UDP-galactose.

A unique example

Before proceeding with a discussion of the functions of nucleotide derivatives, I call attention to one apparently unique example of RNA itself as a carrier or activator in the contemporary synthesis of a small molecule.

Until recently, the key intermediate in porphyrin biosynthesis, δ-aminolevulinate, was thought to be synthesized solely by a pathway originating with succinate and glycine [7]. Surprisingly, a totally different pathway yielding δ-aminolevulinate has emerged more recently. It was discovered in plants and bacteria [8], and starts from glutamyl-tRNA. In some organisms, for example the alga *Euglena gracilis*, the two pathways are active side by side.

Undoubtedly the tRNA-dependent mechanism is more ancient in origin. In any event, this mechanism, which involves a tRNA-linked metabolic intermediate as the ribocofactor in tetrapyrrole biosynthesis, has remained exceptional.

A second, seemingly relevant, example of RNA involvement in intermediary metabolism excited the biochemical community 35 years ago, but turned out to be notorious rather than notable. W.M. Bates, in the laboratory of F. Lipmann, claimed to have isolated an RNA derivative of the dipeptide γ-glutamylcysteine. This was after Paul Zamecnik's group had demonstrated that amino-acyl-tRNA is the intermediate in protein synthesis. Alas, Bates' 'discovery' was a spectacular fraud!

Clues from coenzymes

How do present-day cells use nucleotides and their derivatives? In

Table 1

Some roles of the four RNA bases in intermediary metabolism

Base	Intermediate	Function
Adenine	Aminoacyladenylate	Activated intermediate in forming aminoacyl-tRNA
	ADP-glucose	Glucosyl donor for cellulose synthesis in plants
	Cyclic AMP	Second messenger, mediates responses to hormones such as glucagon
	Adenosylhopane	Component of bacterial membranes
Uracil	UDP-glucose and UDP-galactose	Activated donors of sugars in the synthesis of lipopolysaccharide, sucrose, lactose, glycogen and glycoproteins
	UDP-N-acetyl-glucosamine	Precursor in bacterial peptidoglycan synthesis
Cytosine	CDP-diacylglycerol	Activated intermediate in *de novo* synthesis of phosphoglycerides
	CDP-glucose	Glucosyl donor for cellulose synthesis in plants
Guanine	GDP-mannose	Unknown function in yeast
	GDP-glucose	Glucosyl donor for cellulose synthesis in plants
	Cyclic GMP	Second messenger
	GTP	Cofactor for G-proteins and GTPases

addition to the uses described above and in Table 1, all endergonic reactions require a source of chemical energy, and for biosynthetic events this is ordinarily provided by cleavage of a phosphodiester bond in ATP. This choice can be rationalized given the information that, under simulated prebiotic conditions, adenine is formed from HCN in high yield [9].

Similarly, numerous enzyme-catalyzed processes require small molecules as cofactors, and among the universal cofactors or coenzymes those containing adenine prevail: NAD, NADP, FAD, FMN, coenzyme A and the adenosylcobalamine moiety of vitamin B_{12}. Harold White [10,11] pointed out this dominant role of adenine in coenzyme chemistry, and suggested that other coenzymes (thiamine pyrophosphate, tetrahydrofolate, pyridoxalphosphate) might have been derived from adenine later in evolutionary time.

Thymine, the DNA base

In contrast to adenine and other RNA bases, the DNA base thymine and thymine nucleotides can be excluded from the discussion of nucleotide-linked metabolic reactions. Thymine-containing coenzymes are not known to exist, and there has been no evidence for the involvement of such coenzymes in contemporary intermediary metabolism.

The present pathway of thymine nucleotide synthesis does, however, have implications for early evolution. Thymine and the pathway for its synthesis must have arisen at some, presumably late, stage in the proposed transition from an RNA to a DNA world.

Thymine derives from the RNA base uracil by methylation of α-UMP, in a reaction requiring tetrahydrofolate (THF) as the methyl donor [12]. In turn, THF is synthesized from two precursors, 6-methylpterin, derived from guanine, and *p*-aminobenzoic acid, one of the several aromatic amino acids produced by the shikimic acid pathway.

These metabolic events, leading to the synthesis of thymine nucleotides, must therefore be added to the catalytic repertoire of RNA enzymes present in any RNA world; presumably these ribozymes would have arisen towards the end of the RNA world. As Lubert Stryer puts it [13], "The reactions catalyzed by ribonucleotide reductase and thymidilate synthetase are recapitulations of the transition from an RNA world to one in which DNA became the store of genetic information".

Attesting to their ancient origins, both enzymes are strictly anaerobic, and probably date to the era, about 2×10^9 years ago, before free oxygen accumulated in the atmosphere. The evolution of DNA would also have required the synthesis of deoxyribose, for which ribose still appears to be the sole source [14].

Although the above comments provide some clues as to how the RNA world may have evolved into the DNA world, we are still ignorant of much of nature's logic. For example, why does one of the four RNA bases and not another provide the activating group for precursors in biosynthetic and regulatory processes? Equally puzzling is the observation that, in some instances, nature has seen fit to employ different means for achieving the same end, such as ADP-glucose in some instances and UDP-glucose in others for polysaccharide synthesis.

Perhaps this unpredictability, the mystery of nature's logic, explains why chemists, until recently, have refrained from exploring biochemical processes. It was François Jacob who said that "tinkering", rather than strategic planning at the desk, used to distinguish the biologist's from the chemist's approach to gain new knowledge.

Acknowledgements

I thank Walter Gilbert, Leslie Orgel and Frank Westheimer for valuable suggestions. The writing of this article was aided by the Chemical Research Fund of Harvard University.

References

1. Gilbert, W. (1986). The RNA world. *Nature* **319**, 618.
2. Gesteland, R.F. & Atkins, J.F. (1993). *The RNA World.* Cold Spring Harbor Press, Cold Spring Harbor, NY.
3. Kornberg, A. & Pricer, W.E. (1953). Enzymatic esterification of α-glycerophosphate by long chain fatty acids. *J. Biol. Chem.* **204**, 345–357.
4. Kennedy, E.P. (1962). The metabolism and function of complex lipids. *The Harvey Lectures Series* **57**, 143–171.

5. Rodbell, M., Krans, H.M., Pohl, S.L. & Birnbaumer L. (1971). The glucagon-sensitive adenyl cyclase system in plasma membranes of rat liver. IV. Effects of guanylnucleotides on binding of ^{125}I-glucagon. *J. Biol. Chem.* **246**, 1872–1876.
6. Leloir, F. (1970). Two decades of research on the biosynthesis of saccharides. In *Les Prix Nobel.* pp. 178–188, Impremerial Royal, Stockholm.
7. Shemin, D. (1955). The biosynthesis of porphyrins. *The Harvey Lectures* **50**, 258–284.
8. Weinstein, J.D. & Beale, S.I. (1983). Separate physiological roles and subcellular compartments for two tetrapyrrole biosynthetic pathways in *Euglena gracilis*. *J. Biol. Chem.* **258**, 6799–6807.
9. Oró, J. & Kimball, A.P. (1961). Synthesis of purines under possible primitive earth conditions. I. Adenine from hydrogen cyanide. *Arch. Biochem. Biophys.* **94**, 217–227.
10. White, H.B. III (1976). Coenzymes as fossils of an earlier metabolic state. *J. Mol. Evol.* **7**, 101–104.
11. White, H.B. III (1982). Evolution of coenzymes and the origin of pyridine nucleotides. In *The Pyridine Nucleotide Coenzymes.* (Everse, J., Anderson, B., You, K., eds), pp. 1–17, Academic Press, New York.
12. Friedkin, M. (1973). Thymidylate synthetase. *Adv. Enzymol.* **38**, 235–292.
13. Stryer, L. (1988). *Biochemistry.* (3rd edn), W.H. Freeman and Co., San Francisco, CA.
14. Reichard, P. (1995). To be there when the picture is painted. *Annu. Rev. Biochem.* **64**, 1–28.

Address: Department of Chemistry, Harvard University, 12 Oxford Street, Cambridge, MA 02138, USA.