Preface

THE BASIC FEATURES OF THE HEREDITARY MATERIAL, the mechanism of its readout, and much of the encoded biochemical repertoire of cells are conserved in all living organisms on our planet. This is the key to the realization that life, as we know it, descended from a common ancestor. The time from the existence of this ancestor, several billion years ago, contrasts with the time from the emergence of modern humans and language, a mere 50,000–200,000 years ago. For millennia, however, the rich diversity of life obscured appreciation of its essential unity—without scientific investigation it was not obvious that the molecular similarities between an elephant and a tree are as striking as their obvious differences. The facility to appreciate this unity followed from the development of agriculture and complex society a mere 10,000 years ago. We all are privileged to be among the first generation of life that can pose the question of its origin in meaningful molecular terms.

This third edition of *The RNA World* does not deal with the fascinating clash of diverse dogmas associated with different origin-of-life hypotheses formulated before the insights provided by the scientific era. Rather, it focuses on the hypothesis that RNA, acting both as hereditary information and as a catalyst, provided the basis for self-reproducing forms prior to the emergence of our modern biological world. Genetic information now is lodged in a separate storage repository, DNA, and most catalytic functions are mediated by more biochemically versatile proteins.

The second focus of this book is on the modern roles of RNA. Their study is essential for our understanding of biochemistry and our full utilization of biomedical opportunities. In addition, these roles may be "molecular fossils," that provide insights into primordial roles. While mRNA, tRNA, rRNA, and viral RNAs were becoming known starting five decades ago, important new insights into the functions of these RNA species have emerged since the last edition of this book was published in 1999. In fact, the field was on the verge of an explosion of discoveries of major new functions for RNA, hitherto unimagined. It is ribosomal RNA rather than protein that performs peptidyl transfer, linking amino acids encoded by an RNA sequence. Moreover, RNA forms the detailed atomic architecture of the decoding site to effect information transfer from RNA to protein. RNA is primal in this most fundamental mechanism of modern-day decoding, providing key support for the hypothesis that protein-based life emerged from an RNA World without the need for preexisting proteins. Other roles for RNA in our modern world were not anticipated at all. The most widely known of these major new functions is RNA interference—wherein small double-stranded RNA molecules can repress gene expression at the level of mRNA stability, mRNA translation, or mRNA synthesis through changes in chromatin structure. But there are now additional noncoding RNAs being found monthly, many too new to be covered in this volume.

The rapidity with which new noncoding RNAs have been found is due, directly or indirectly, to the various genome projects. Analysis of the transcription of these genomes has revealed much additional RNA with no function yet assigned. Bioinformatics is still relatively primitive in its ability to recognize functional noncoding RNAs within genome sequences. Thus, it is already clear that additional classes of functional RNAs are awaiting discovery, and that RNA World IV, when it appears, will include chapters whose titles we can't even imagine. Some of these new roles for RNA have implications for our vision of an ancient RNA world, and their diversity richly embellishes the pioneering suggestions of Woese, Orgel, and Crick that prebiotic evolution depended on RNA replication.

Before he died last year, Francis Crick had lived to celebrate the 50th anniversary of Watson-Crick base-pairing and the DNA double helix. Fortunately, he lived to see the major advances in our understanding of decoding for which he had so brilliantly laid the groundwork.

We dedicate this volume to Francis Crick, a giant of our times who will go down in history with Darwin.

We thank the many people who made this book possible: the authors, of course, and the staff at Cold Spring Harbor Laboratory Press, especially Inez Sialiano, Patricia Barker, and John Inglis.

R. F. GESTELAND
T.R. CECH
J.F. ATKINS

Foreword to the First Edition

THE TERM "RNA WORLD" ORIGINALLY REFERRED TO a hypothetical time in the evolution of earthly life when there was no elaborate mechanism for protein synthesis such as we have today. As described by Joyce and Orgel (Chapter 1), there were speculations in the 1960s that RNA catalysts existed at that stage in evolution, that RNA was the sole genetic material, and that the standard Watson-Crick pairing was the basis of genetic replication. None of these early authors was smart enough to suggest that relics of these hypothetical catalytic RNAs might still be around today. Indeed, it was speculated that the original ribosomes might have been made solely of RNA but not that their main catalytic activity (the formation of the peptide bond) might still be performed today by RNA alone, as recent evidence seems to suggest.

This hypothesis of an RNA world without protein was largely forgotten but has now become fashionable again because of the remarkable discoveries by Altman and by Cech (described by Cech, Chapter 11) of RNA molecules that do indeed have catalytic activity of one sort or another. These discoveries have removed the chief objection to the RNA world—that RNA by itself cannot act catalytically.

This volume on the RNA world covers a somewhat wider field. It was realized in the 1960s that not all present-day RNA was messenger RNA or viral RNA. The ribosome was known to contain several distinct types of "structural" RNA molecules; tRNAs were familiar, and there were isolated examples of other small RNA molecules. What was not realized was the richness of the present-day RNA world, including snRNPs (as described by Baserga and Steitz, Chapter 14) and now many others. We have come to realize that RNA molecules can occur with many different shapes and can thus have many different functions. Several of the chapters in this book deal with what these structures are, what their functions may be, what makes them stable, and how to predict their structure. Just how many of these RNAs are, by themselves, truly catalytic remains to be seen.

Some of them may use divalent cations to produce both their folding and their catalytic activities, as discussed by Pan et al. (Chapter 12).

Some of the properties of present-day RNA molecules were quite unexpected. One might perhaps have guessed that some limited form of messenger-RNA editing could occur very rarely as a freak, but who would have predicted that such editing would take so many distinct forms and occur as widely as Bass describes in Chapter 15?

That a retrovirus contains a sequence rather like part of a tRNA molecule was a surprise when it was first discovered. In 1987, Maizels and Weiner proposed that such structures tagged RNA genomes for replication in the early RNA world. In Chapter 23, they explore the many possible ramifications of this genome tag hypothesis. In particular, they propose that both tRNA and tRNA-aminoacylation activity were subject to selection before protein synthesis arrived on the scene. Whether one can safely extrapolate from a present-day property of rapidly evolving RNA viruses to as far back in time as the early RNA world remains to be seen.

A good case can be made (Gold et al., Chapter 19) that small lengths of RNA can often form relatively rigid structures more easily than can polypeptides of similar length. New techniques are now being used (Chapter 19 and Chapter 20 [Szostak and Ellington]) to explore this vast "space" of small RNA sequences for interesting properties such as the binding of one specific substrate or another and, in some cases, a particular type of catalytic activity.

The recombinant-DNA revolution of the 1970s was possible because molecular biologists could use the sophisticated products of billions of years of natural evolution—the replicases, restriction enzymes, and so forth—as precise and delicate chemical tools. These molecules were all proteins that acted on nucleic acids in one way or another. Now experimentalists (Chapters 19 and 20) are using a combination of biochemical processes, related to nucleic acid replication, which together embody the mechanism of natural selection. These powerful tools can provide them with new and desirable RNA molecules in the laboratory while they wait. Because these methods do not need whole cells, but only cellular components, they can handle many more individuals at one time and so explore the RNA "space" more quickly. The message is clear: Where possible, let Nature do the work, and if she proves a little slow, take your whip to her.

The exact nature of the early RNA world is now a matter of active debate. What can we learn from the present "relics" and, in particular, can we deduce the composition and nature of each particular type of early RNA from the detailed study of its many descendants today? Can we learn anything from small RNA-related molecules (such as coenzymes) about the general nature of the early RNA world? Early speculations had assumed that the RNA world was rather simple and rather inaccurate. Benner and his colleagues (Chapter 2) now believe it was much more complex and more accurate. As more sequence data accumulate, one would expect these suggestions to become less speculative.

It may be possible to deduce something about some of the activities in the RNA world by a detailed study of the present mechanisms of protein synthesis. For example, Weiss and Cherry (Chapter 3) have proposed an ingenious model for the origins of the smaller ribosomal subunit. They suggest that it was originally an RNA replicase that used oligonucleotides as a substrate, cannibalizing triplets from them for its own replication.

We may, in time, arrive at a rather plausible picture of this early stage in evolution, even if true molecular fossils (that is, actual specimens of the molecules that then existed) are forever unavailable to us because of the ceaseless battering of thermal motion over billions of years.

The details of the transition from the RNA world to the present protein-dominated world are also debatable. For example, was there some sort of protein synthesis before the existence of a primitive ribosome? Did these first ribosomes have some protein, or did the ribosomal proteins arrive a little later?

It may turn out that we will eventually be able to see how this RNA world got started. At present, the gap from the primal "soup" to the first RNA system capable of natural selection looks forbiddingly wide. Was there perhaps a pre-RNA world, with replication based on some even more primitive system?

Such a pre-RNA world might have been of several types. It would presumably have been based on some form of genetic polymer or sheet that was formed more easily than RNA under prebiotic conditions. It might eventually have transcribed its detailed genetic information directly onto RNA to form the RNA world. Alternatively, it might have acted solely as a midwife, setting up the basis of RNA synthesis for some use of its own, with RNA replication then taking over and replacing it. Cairns-Smith has already suggested that this pre-RNA system might have been based on clays, or on some organic polymer (Cairns-Smith, A.G., *Genetic Takeover and the Mineral Origins of Life*. Cambridge University Press [1982]). This time we should be smart enough to ask if there are any relics of this pre-RNA world still with us today.

Finally, a word of heresy: As Moore has pointed out (Chapter 5), the real fossil record suggests that our present form of protein-based life was already in existence 3.6 billion years ago and evolved rather slowly for a billion or so years after that. This leaves an astonishingly short time to get life started. Moreover, the three main lines of descent (see Fig. 3 in Woese and Pace, Chapter 4) seem a very long distance from their hypothetical common ancestor.

The rather far-fetched hypothesis of Directed Panspermia would predict that life was sent here in the form of "bacteria" suitable for growth in anaerobic conditions, *and that several somewhat different forms would probably have been sent at the same time*, in the hope that at least one would survive. All are likely to have evolved originally from a common ancestor (on another planet) that existed some billions of years before the formation of our solar system. Therefore, the final question about the RNA world and the pre-RNA world (if it existed) is: *Where* did it occur? Are we totally confident that our form of life started here, or did it perhaps originate elsewhere in the universe? It might have been easier to start elsewhere because, for example, the atmosphere there was more reducing than the Earth's early atmosphere appears to have been.

The lively and authoritative chapters of this book deal comprehensively both with the hypothetical RNA world and with the complexities of RNA structure and function we find around us today. I recommend it to all molecular biologists and especially to anyone fascinated by the baroque complexity of the nucleic acids and of RNA in particular.

FRANCIS CRICK