CHAPTER ONE

Illustrious Predecessors

SCHRÖDINGER AND VON NEUMANN

In February 1943, at a bleak moment in the history of mankind, the physicist Erwin Schrödinger gave a course of lectures to a mixed audience at Trinity College, Dublin. Ireland was then, as it had been in the days of Saint Columba fourteen hundred years earlier, a refuge for scholars and a nucleus of civilization beyond the reach of invading barbarians. It was one of the few places in Europe where peaceful scientific meditation was still possible. Schrödinger proudly remarks in the published version of the lectures that they were given "to an audience of about four hundred which did not substantially dwindle." The lectures were published by the Cambridge University Press in 1944 in a little book (Schrödinger, 1944) with the title *What is Life*?

Schrödinger's book is less than a hundred pages long. It was widely read and was influential in guiding the thoughts of the young people who created the new science of molecular biology in the following decade. It is clearly and simply written, with only five references to the technical literature and less than ten equations from beginning to end. It is, incidentally, a fine piece of English prose. Although Schrödinger was exiled from his native Austria to Ireland when he was over fifty, he wrote English far more beautifully than most of his English and American contemporaries. He reveals his cosmopolitan background only in the epigraphs that introduce his chapters: three are from Goethe, in German; three are from Descartes and Spinoza, in Latin; and one is from Unamuno, in Spanish. As a sample of his style I quote the opening sentences of his preface:

> A scientist is supposed to have a complete and thorough knowledge, at first hand, of some subjects, and therefore he is usually expected not to write on any topic of which he is not a master. This is regarded as a matter of noblesse oblige. For the present purpose I beg to renounce the noblesse, if any, and to be freed of the ensuing obligation. My excuse is as follows. We have inherited from our forefathers the keen longing for unified, all-embracing knowledge. The very name given to the highest institutions of learning reminds us that from antiquity and throughout many centuries the universal aspect has been the only one to be given full credit. But the spread, both in width and depth, of the multifarious branches of knowledge during the last hundred odd years has confronted us with a queer dilemma. We feel clearly that we are only now beginning to acquire reliable material for welding together the sum-total of what is known into a whole; but, on the other hand, it has become next to impossible for a single mind fully to command more than a small specialized portion of it. I can see no other escape from this dilemma (lest our true aim be lost for ever) than that some of us should venture to embark on a synthesis of facts and theories, albeit with second-hand and incomplete knowledge of some of them, and at the risk of making fools of themselves. So much for my apology.

This apology for a physicist venturing into biology will serve for me as well as for Schrödinger, although in my case the risk of the physicist making a fool of himself may be somewhat greater.

Schrödinger's book was seminal because he knew how to ask the right questions. What is the physical structure of the molecules that are duplicated when chromosomes divide? How is the process of duplication to be understood? How do these molecules retain their individuality from generation to generation? How do they succeed in controlling the metabolism of cells? How do they create the organization that is visible in the structure and function of higher organisms? He did not answer these questions, but by asking them he set biology moving along the path that led to the epoch-making discoveries of the subsequent forty years: to the discovery of the double helix and the triplet code, to the precise analysis and wholesale synthesis of genes, and to the quantitative measurement of the evolutionary divergence of species.

One of the great pioneers of molecular biology who was active in 1943 and is still active today, Max Perutz, dissents sharply from my appraisal of Schrödinger's book (Perutz, 1989). "Sadly," Perutz writes, "a close study of his book and of the related literature has shown me that what was true in his book was not original, and most of what was original was known not to be true even when the book was written." Perutz's statement is well founded. Schrödinger's account of existing knowledge is borrowed from his friend Max Delbrück, and his conjectured answers to the questions that he raised were indeed mostly wrong. Schrödinger was woefully ignorant of chemistry, and in his isolated situation in Ireland he knew little about the new world of bacteriophage genetics that Delbrück had explored after emigrating to the United States in 1937. But Schrödinger never claimed that his ideas were original, and the importance of his book lies in the questions that he raised rather than in the answers that he conjectured. In spite of Perutz's dissent, Schrödinger's book remains a classic because it asked the right questions.

Schrödinger showed wisdom not only in the questions that he asked but also in the questions that he did not ask. He did not ask any questions about the origin of life. He understood that the time was ripe in 1943 for a fundamental understanding of the physical basis of life. He also understood that the time was not then ripe for any fundamental understanding of life's origin. Until the basic chemistry of living processes was clarified, one could not ask meaningful questions about the possibility of spontaneous generation of these processes in a prebiotic environment. He wisely left the question of origins to a later generation.

Now, half a century later, the time is ripe to ask the questions Schrödinger avoided. We can hope to ask the right questions about origins today because our thoughts are guided by the experimental discoveries of Manfred Eigen, Leslie Orgel, and Thomas Cech. The questions of origin are now becoming experimentally accessible just as the questions of structure were becoming experimentally accessible in the 1940s. Schrödinger asked the right questions about structure because his thoughts were based on the experimental discoveries of Timoféeff-Ressovsky, who exposed fruit-flies to X-rays and measured the relationship between the dose of radiation and the rate of appearance of genetic mutations. Delbrück was a friend of Timoféeff-Ressovsky and published a joint paper with him describing and interpreting the experiments (Timoféeff-Ressovsky et al., 1935). Their joint paper provided the experimental basis for Schrödinger's questions. After 1937, when Delbrück came to America, he continued to explore the problems of structure. Delbrück hit on the bacteriophage as the ideal experimental tool, a biological system stripped of inessential complications and reduced to an almost bare genetic apparatus. The bacteriophage was for biology what the hydrogen atom was for physics. In a similar way Eigen became the chief explorer of the problems of the origin of life in the 1970s because he hit on ribonucleic acid (RNA) as the ideal experimental tool for studies of molecular evolution in the test-tube. Eigen's RNA experiments have carried Delbrück's bacteriophage experiments one step further: Eigen stripped the genetic apparatus completely naked, thereby enabling us to study its replication unencumbered by the baggage of structural molecules that even so rudimentary a creature as a bacteriophage carries with it.

Before discussing the experiments of Eigen, Orgel, and Cech in detail, I want to finish my argument with Schrödinger. At the risk, again, of making a fool of myself, I shall venture to say that in his discussion of the nature of life Schrödinger missed an essential point. And I feel that the same point was also missed by Manfred Eigen in his discussion of the origin of life. I hasten to add that in disagreeing with Schrödinger and Eigen I am not disputing the greatness of their contributions to biology. I am saying only that they did not ask all of the important questions.

In Schrödinger's book we find four chapters describing in lucid detail the phenomenon of biological replication and a single chapter describing less lucidly the phenomenon of metabolism. Schrödinger finds a conceptual basis in physics both for exact replication and for metabolism. Replication is explained by the quantum mechanical stability of molecular structures, whereas metabolism is explained by the ability of a living cell to extract negative entropy from its surroundings in accordance with the laws of thermodynamics. Schrödinger was evidently more interested in replication than in metabolism. There are two obvious reasons for his bias. First, he was, after all, one of the inventors of quantum mechanics, and it was natural for him to be primarily concerned with the biological implications of his own brainchild. Second, his thinking was based on Timoféeff-Ressovsky's experiments, and these were biased in the same direction. The experiments measured the effects of X-rays on replication and did not attempt to observe effects on metabolism. Delbrück carried the same bias with him when he came to America. Delbrück's new experimental system, the bacteriophage, is a purely parasitic creature in which the metabolic function has been lost and only the replicative function survives. It was indeed precisely this concentration of attention upon a rudimentary and highly specialized form of life that enabled Delbrück to do experiments exploring the physical basis of biological replication. It was necessary to find a creature without metabolism to isolate experimentally the phenomena of replication. Delbrück penetrated more deeply than his contemporaries into the mechanics of replication because he was not distracted by the problems of metabolism. Schrödinger saw the world of biology through Delbrück's eyes. It is not surprising that Schrödinger's view of what constitutes a living organism resembles a bacteriophage more than it resembles a bacterium or a human being. His single chapter devoted to the metabolic aspect of life appears to be an afterthought put in for the sake of completeness but not affecting the main line of his argument.

The main line of Schrödinger's argument, which led from the facts of biological replication to the quantum mechanical structure of the gene, was brilliantly clear and fruitful. It set the style for the subsequent development of molecular biology. Neither Schrödinger himself nor the biologists who followed his lead appear to have been disturbed by the logical gap between his main argument and his discussion of metabolism. Looking back on his 1943 lectures now with the benefit of half a century of hindsight, we may wonder why he did not ask some fundamental questions that the gap might have suggested to him: Is life one thing or two things? Is there a logical connection between metabolism and replication? Can we imagine metabolic life without replication, or replicative life without metabolism? These questions were not asked because Schrödinger and his successors took it for granted that the replicative aspect of life is primary and the metabolic aspect secondary. As their understanding of replication became more and more triumphantly complete, their lack of understanding of metabolism was pushed into the background. In popular accounts of molecular biology as it is now taught to school children, life and replication have become practically synonymous. In modern discussions of the origin of life it is often taken for granted that the origin of life is the same thing as the origin of replication. Manfred Eigen's view is an extreme example of this tendency. Eigen chose RNA as the working material for his experiments because he wished to study replication but was not interested in metabolism. Eigen's theories about the origin of life are in fact theories about the origin of replication.

It is important here to make a sharp distinction between replication and reproduction. I am suggesting as a hypothesis that the earliest living creatures were able to reproduce but not to replicate. What does this mean? For a cell, to reproduce means simply to divide into two cells with the daughter cells inheriting approximately equal shares of the cellular constituents. For a molecule, to replicate means to construct a precise copy of itself by a specific chemical process. Cells can reproduce, but only molecules can replicate. In modern times, reproduction of cells is always accompanied by replication of molecules, but this need not always have been so in the past.

It is also important to say clearly what we mean when we speak of metabolism. One of my American friends, a professional molecular biologist, told me that it would never occur to him to ask the question whether metabolism might have begun before replication. For him the word metabolism means chemical processes directed by the genetic apparatus of nucleic acids. If the word has this meaning, then by definition metabolism could not have existed without a genetic apparatus to direct it. He said he was astonished when one of his German colleagues remarked that metabolism might have come first. He asked the German how he could entertain such an illogical idea. For the German, there was nothing illogical in the idea of metabolism coming before replication, because the German word for metabolism is *Stoffwechsel*, which translates into English as "stuffchange." It means any chemical process occurring in cells, whether directed by a genetic apparatus or not. My friend tells me that students who learn molecular biology in American universities always use the word metabolism to mean genetically directed processes. That is one reason they take it for granted that replication must come first. I therefore emphasize that in this book I am following the German and not the American usage. I mean by metabolism what the Germans mean by *Stoffwechsel* with no restriction to genetically directed processes.

Only five years after Schrödinger gave his lectures in Dublin, the logical relations between replication and metabolism were clarified by the mathematician John von Neumann (von Neumann, 1948). Von Neumann described an analogy between the functioning of living organisms and the functioning of mechanical automata. His automata were an outgrowth of his thinking about electronic computers. A von Neumann automaton had two essential components; later on, when his ideas were taken over by the computer industry, these were given the names hardware and software. Hardware processes information; software embodies information. These two components have their exact analogues in living cells; hardware is mainly protein and software is mainly nucleic acid. Protein is the essential component for metabolism. Nucleic acid is the essential component for replication. Von Neumann described precisely, in abstract terms, the logical connections between the components. For a complete self-reproducing automaton, both components are essential. Yet there is an important sense in which hardware comes logically prior to software. An automaton composed of hardware without software can exist and maintain its own metabolism. It can live independently for as long as it finds food to eat or numbers to crunch. An automaton composed of software without hardware must be an obligatory parasite. It can function only in a world already containing other automata whose hardware it can borrow. It can replicate itself only if it succeeds in finding a cooperative host automaton, just as a bacteriophage can replicate only if it succeeds in finding a cooperative bacterium.

In all modern forms of life, hardware functions are mainly performed by proteins and software functions by nucleic acids. But there are important exceptions to this rule. Although proteins serve only as hardware, and one kind of nucleic acid, namely deoxyribonucleic (DNA), serves mainly as software, the other kind of nucleic acid, namely RNA, occupies an intermediate position. RNA is both hardware and software. RNA occurs in modern organisms in four different forms with different functions. There is genomic RNA, constituting the entire genetic endowment of many viruses – in particular the AIDS virus. Genomic RNA is unambiguously software. There is ribosomal RNA, an essential structural component of the ribosomes that manufacture proteins. There is transfer RNA, an essential part of the machinery that brings amino acids to ribosomes to be incorporated into proteins. Ribosomal RNA and transfer RNA are unambiguously hardware. Finally, there is messenger RNA, the molecule that conveys the genetic instructions from DNA to the ribosome. It was believed until recently that messenger RNA was unambiguously software, but Thomas Cech discovered in 1982 that messenger RNA also has hardware functions (Cech, 1993). Cech observed messenger RNA molecules that he called ribozymes performing the functions of enzymes. Ribozymes catalyze the splitting and splicing of other RNA molecules. They also catalyze their own splitting and splicing, in which case they are acting as hardware and software simultaneously. RNA is a flexible and versatile molecule with many important hardware functions in addition to its primary software function. Nevertheless it remains true that the overwhelming majority of metabolic functions of modern organisms belong to proteins, and the overwhelming majority of replicative functions belong to nucleic acids.

Let me summarize the drift of my argument up to this point. Our illustrious predecessor Erwin Schrödinger gave his book the title *What is Life*? but neglected to ask whether the two basic functions of life, metabolism and replication, are separable or inseparable. Our illustrious predecessor John von Neumann, using the computer as a metaphor, raised the question that Schrödinger had missed and gave it a provisional answer. Von Neumann observed that metabolism and replication, however intricately they may be linked in the biological world as it now exists, are logically separable. It is logically possible to postulate organisms that are composed of pure hardware and capable of metabolism but incapable of replication. It is also possible to postulate organisms that are composed of pure software and capable of replication but incapable of metabolism. And if the functions of life are separated in this fashion, it is to be expected that the latter type of organism will become an obligatory parasite upon the former. This logical analysis of the functions of life helps to explain and to correct the bias toward replication that is evident in Schrödinger's thinking and in the whole history of molecular biology. Organisms specializing in replication tend to be parasites, and molecular biologists prefer parasites for experimental study because parasites are structurally simpler than their hosts and better suited to quantitative manipulation. In the balance of nature there must be an opposite bias. Hosts must exist before there can be parasites. The survival of hosts is a precondition for the survival of parasites. Somebody must eat and grow to provide a home for those who only reproduce. In the world of microbiology, as in the world of human society and economics, we cannot all be parasites.

When we begin to think about the origins of life we meet again the question that Schrödinger did not ask, What do we mean by life? And we meet again von Neumann's answer, that life is not one thing but two, metabolism and replication, and that the two things are logically separable. There are accordingly two logical possibilities for life's origins. Either life began only once, with the functions of replication and metabolism already present in rudimentary form and linked together from the beginning, or life began twice, with two separate kinds of creatures, one kind capable of metabolism without exact replication and the other kind capable of replication without metabolism. If life began twice, the first beginning must have been with molecules resembling proteins, and the second beginning with molecules resembling nucleic acids. The first protein creatures might have existed independently for a long time, eating and growing and gradually evolving a more and more efficient metabolic apparatus. The nucleic acid creatures must have been obligatory parasites from the start, preying upon the protein creatures and using the products of protein metabolism to achieve their own replication.

The main theme of this book will be a critical examination of the second possibility, the possibility that life began twice. I call this possibility the double-origin hypothesis. It is a hypothesis, not a theory. A theory of the origin of life should describe in some detail a postulated sequence of events. The hypothesis of dual origin is compatible with many theories. It may be useful to examine the consequences of the hypothesis without committing ourselves to any particular theory.

I do not claim that the double-origin hypothesis is true, or that it is supported by any experimental evidence. Indeed my purpose is just the opposite. I would like to stimulate experimental chemists and biologists and paleontologists to find the evidence by which the hypothesis might be tested. If it can be tested and proved wrong, it will have served its purpose. We will then have a firmer foundation of fact on which to build theories of single origin. If the double-origin hypothesis can be tested and not proved wrong, we can proceed with greater confidence to build theories of double origin. The hypothesis is useful only insofar as it may suggest new experiments.

Lacking new experiments, we have no justification for believing strongly in either the single-origin or the double-origin hypothesis. I have to confess my own bias in favor of double-origin. But my bias is based only on general philosophical preconceptions, and I am well aware that the history of science is strewn with the corpses of dead theories that were in their time supported by the prevailing philosophical viewpoints. For what it is worth, I may state my philosophical bias as follows: The most striking fact we have learned about life as it now exists is the ubiquity of dual structure, the division of every organism into hardware and software components, into protein and nucleic acid. I consider dual structure to be prima facie evidence of dual origin. If we admit that the spontaneous emergence of protein structure and nucleic acid structure out of molecular chaos is unlikely, it is easier to imagine two unlikely events occurring separately over a long period than to imagine two unlikely