one approach to treatment is to selectively inhibit reverse transcriptase.) In this case, the virus encodes information on a RNA molecule. In the host cell, viral reverse transcriptase produces a viral DNA molecule using the viral RNA as a template. Such viruses are clinically important and the enzyme, reverse transcriptase, is an important tool in genetic engineering. Nonetheless, the process depicted in Fig. 4.1 is essentially applicable to any cell of commercial importance.

For information storage and exchange to take place, there must be a language. We can conceive of all life as using a four-letter alphabet made up of the nucleotides discussed in Chapter 2 (that is, A, T, G, and C in DNA). All words are three letters long; such words are called *codons*. With four letters and only three-letter words, we have a maximum of 64 words. These words, when expressed, represent a particular amino acid or "stop" protein synthesis.

When these words are put into a sequence, they can make a "sentence" (i.e., a *gene*), which when properly transcribed and translated is a protein. Other combinations of words regulate when the gene is expressed. Carrying the analogy to the extreme, we may look at the complete set of information in an organism's DNA (i.e., the *genome*) as a book. (For the human genome it would be more than 1000 books the size of this one.)

This simple language of 64 words is all that is necessary to summarize your total physical makeup at birth and all your natural capabilities. It is essentially universal, the same for *E. coli* and *Homo sapiens* (humans). This universality has helped us to make great strides in understanding life and is a practical tool in genetic engineering and commercial biotechnology.

Each of the steps in information storage and transfer (Fig. 4.1) requires a macromolecular template. Let us examine how these templates are made and how this genetic-level language is preserved and expressed.

4.3. DNA REPLICATION: PRESERVING AND PROPAGATING THE CELLULAR MESSAGE

The double-helix structure of DNA discussed in Chapter 2 is extremely well suited to its role of preserving genetic information. Information resides simply in the linear arrangement of the four nucleotide letters (A, T, G, and C). *Because G can hydrogen-bind only to C and A only to T, the strands must be complementary* if an undistorted double helix is to result. Replication is *semiconservative* (see Fig. 2.18); each daughter chromosome contains one parental strand and one newly generated strand.

To illustrate the replication process (see Figs. 4.2 and 4.3), let us briefly consider DNA replication in *E. coli*. The enzyme responsible for covalently linking the monomers is DNA polymerase. *Escherichia coli* has three DNA polymerases (named Pol I, Pol II, and Pol III). A DNA polymerase is an enzyme that will link deoxynucleotides together to form a DNA polymer. Pol III enzymatically mediates the addition of nucleotides to an RNA primer. Pol I can hydrolyze an RNA primer and duplicates single-stranded regions of DNA; it is also active in the repair of DNA molecules. The exact role of Pol II is still unclear.

In addition to the enzyme, the enzymatic reaction requires activated monomer and the template. The activated monomers are the nucleoside triphosphates. The formation of