

color to change from red to white. In the latter case, the character is *eye color*, and the phenotype is *white* eye color. However, genes and phenotypes are not always in one-to-one correspondence. For example, in *E. coli* there are seven genes involved in the biosynthesis of the amino acid tryptophan from a precursor molecule. Mutations in any of these seven genes might lead to a Trp^- phenotype (which requires the addition of tryptophan for growth in the minimal medium). Similarly, a phenotype such as height or stature in *Homo sapiens* is controlled by a number of different genes, this time in a quantitative rather than all-or-none manner.

A given gene (corresponding to a particular locus) may have alternative forms called **alleles**, as described in Section 1.3.2. These alleles differ in DNA sequence, and these differences lead to differences in amino acid sequence. For example, individuals affected by sickle cell anemia have a beta-globin gene in which the glutamine normally present as the sixth amino acid residue is replaced by valine. This altered beta-globin gene is one allele of beta-globin, and the normal gene (wild-type gene) is another allele. There are other beta-globin alleles in the human population.

Genes are transcribed from DNA to form RNA molecules (including mRNA, a very important class of RNA). The DNA strand that is complementary to the mRNA is called the **template strand**, while the DNA strand whose sequence corresponds to the mRNA sequence is called the **coding strand**. DNA features (elements) found 5' relative to the coding sequence are considered to be "upstream," and elements that occur 3' relative to the coding sequence are referred to as "downstream." Diagrams of prokaryotic and eukaryotic genes are presented in Fig. 1.7.

Prokaryotic genes have a number of component elements. Going in the 5' to 3' direction (i.e., in the direction of transcription), we encounter sites for binding proteins that control expression of the gene, a **promoter** where transcription initiates, the uninterrupted coding sequence (that eventually is translated into an amino acid sequence), and translational terminators. Sometimes the coding sequences for two or more polypeptide chains will be transcribed in succession from the same promoter to produce a single polycistronic mRNA. A set of contiguous genes transcribed from the same promoter into a single mRNA molecule constitutes an **operon**. Operons are common in prokaryotes but uncommon in complex eukaryotes.

Eukaryotic genes are much more complicated than prokaryotic genes. They contain **exons**, which are segments of gene sequences that are represented in the processed mRNA. All segments of the coding sequence that will eventually be translated into protein appear in exons. **Introns** are noncoding DNA segments that separate the exons, and the RNA corresponding to introns is removed from the initial transcript by an excision process called **splicing**. Eukaryotic genes have more extensive control regions that include binding sites for transcription factors and enhancer sequences (that together regulate the level of transcription) and the "core" promoter where the transcription complex assembles. Eukaryotic transcription terminates nonspecifically down-