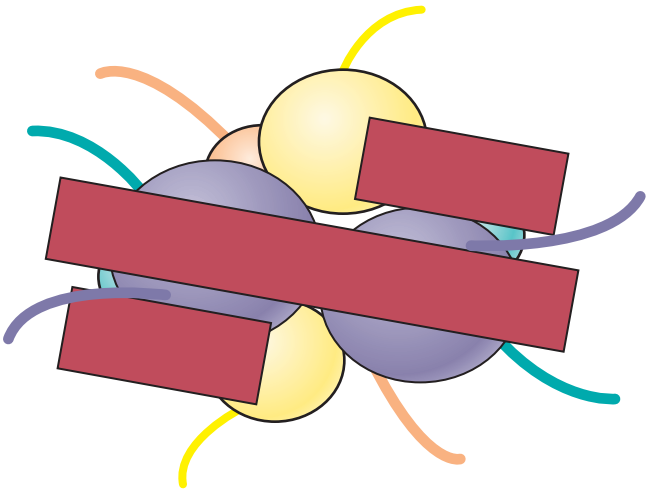

FIGURE 1-4 Schematic representation of a nucleosome. Rectangular blocks represent the DNA strand wrapped around the core that consists of eight histone proteins. Each histone has a protruding tail that can be modified to repress or activate transcription.

Epigenetic Control of Gene Expression

Complex regulatory networks revolve around transcription factors, nucleosomes, chromatin structure, and epigenetic markings. *Epigenetics* refers to heritable changes in gene expression without changes in the DNA sequence. Examples include DNA methylation, gene silencing, chromatin remodeling, and X-chromosome inactivation. This form of inheritance involves alterations in gene function without changes in DNA sequence. The chemical marking of DNA methylation is cell specific and developmentally regulated. Methylation of the 5′-CpG dinucleotide by specific methyl transferases, which occurs in 70% of the mammalian genome, is another mechanism of gene regulation. Steric hindrance from the bulky methyl group of 5′-methylcytosine precludes occupancy by transcription factors that stimulate or attenuate gene expression. Most genes are found in CpG islands, reflecting sites of gene activity across the genome.

In an analogous manner, modifications of histone by phosphorylation, methylation, ubiquitination, and acetylation are transmitted and reestablished in an inheritable manner. It is conceivable that other epigenetic mechanisms do not involve genomic modifications of DNA. For example, modification of the gene encoding the estrogen receptor-α has been implicated in gene silencing at 5-methylcytosine (^{5m}C) sites of multiple downstream targets in breast cancer cells. Powerful approaches are being developed to examine feedback and feed-forward loops in the transmission of epigenetic markings.

The concept that dynamic modifications (i.e., DNA methylation and acetylation) of histones or epigenesis contribute in part to tumorigenic potential for progression has already been translated into therapies. Histone acetyltransferases (HATs) and histone deacetyltransferases (HDACs) play antagonistic roles in the addition and removal of acetylation in the genome. Genome-wide analysis of HATs and HDACs is beginning to provide important insights into complex modes of gene regulation. Several inhibitors of histone deacetylases with a range of biochemical and biologic activities are being developed and tested as anticancer agents in clinical trial. Results of phase I clinical trials have suggested that these drugs are well tolerated. Inhibition of deacetylase remodels chromatin assembly and reactivates transcription of the genome. Because the mechanisms of actions of HDACs extend to apoptosis, cell cycle control, and cellular differentiation, current clinical trials are seeking to determine the efficacy of these novel reagents in the drug compendium for human cancers.

GENETIC SEQUENCE VARIATION, POPULATION DIVERSITY, AND GENETIC POLYMORPHISMS

A stable, heritable change in DNA is defined as a *mutation*. This strict contemporary definition does not depend on the functional relevance of the sequence alteration and implicates a change in the primary DNA sequence. Considered in a historical context, mutations were first defined on the basis of identifiable changes in the heritable phenotype of an organism. As biochemical phenotyping became more precise in the mid-20th century, investigators demonstrated that many proteins exist in more than one form in a population, and these forms were viewed as a consequence of variations in the gene coding for that protein (i.e., allelic variation). With advances in DNA-sequencing methods, the concept of mutation evolved from one that could be appreciated only by identifying differences in phenotype to one that could precisely be defined at the level of changes in the structure of DNA. Although most mutations are stably transmitted from parents to offspring, some are genetically lethal and cannot be passed on. The discovery of regions of the genome that contain sequences that repeat in tandem a highly variable number of times (i.e., tandem repeats) suggests that some mutations are less stable than others. These tandem repeats are further described later.

The molecular nature of mutations is varied (Table 1-1). A mutation can involve the deletion, insertion, or substitution of a single base, all of which are referred to as *point mutations*. Substitutions can be further classified as *silent* when the amino acid encoded by the mutated triplet does not change, as *missense* when the amino acid encoded by the mutated triplet changes, and as *nonsense* when the mutation leads to premature termination of translation (i.e., stop codon). Occasionally, point mutations can alter the processing of precursor mRNA by producing alternate splice sites or eliminating a splice site. When a single- or double-base deletion or insertion occurs in an exon, a frameshift mutation results, usually leading to premature termination of translation at a now in-frame stop codon. The other end of the spectrum of mutations includes large deletions of an entire gene or a set of contiguous genes; deletion, duplication, and translocation of a segment of one chromosome to another; or duplication or deletion of an entire chromosome. These chromosomal mutations play a large role in the development of many cancers.

Each individual possesses two alleles, one from each parent, for any given gene locus. Identical alleles define homozygosity and nonidentical alleles define heterozygosity for any gene locus. The heritability of these alleles follows typical mendelian rules. With a clearer understanding of the molecular basis of mutations and of allelic variation, their distribution in populations can be analyzed precisely by following specific DNA sequences.