INTRODUCTION

In the early 1970s, R.B. Khesin returned to work with Drosophila after a long hiatus in order to study the mechanisms of eukaryotic gene regulation. Among numerous problems, he chose position effect variegation (PEV), a perfect system for elucidating the mechanism of gene inactivation by changes in chromatin structure. PEV is observed when a gene is placed in heterochromatin by a chromosome rearrangement or a transgene is integrated within a transgenic construct (see [1–3] for a review on PEV and heterochromatin). In either case, the gene is active only in some cells, which results in a mosaic phenotype. Various genetic factors enhance or suppress PEV, suggesting their role in the structure and function of heterochromatin. Most genes found to modify PEV code for chromosomal proteins. The active or inactive state of a gene translocated in heterochromatin is inherited through cell generations. Hence PEV provides an example of epigenetic inheritance, which is based on functional, rather than structural, inheritable changes of the gene activity. Thus, studies of PEV may help to clarify the mechanisms of cell differentiation and development.

I was fortunate to work with Khesin on this problem. He was indeed my mentor in research, and not only in research. The chromatin structure and its formation in Drosophila are still the focus of my work. Here I will briefly review the recent data on the structure and composition of heterochromatin, with a special emphasis on its protein components. Interestingly, many heterochromatin proteins may participate in regulatory events in euchromatin, suggesting that heterochromatin and euchromatin differ less than earlier believed. Another unexpected finding was that these events also involve transposable genetic elements (TGE) [4], which attracted Khesin’ attention in his last years. I will consider the possible genomic role of TGE, which, along with other factors, may contribute to the spatial chromatin arrangement in the interphase nucleus.

STRUCTURAL FEATURES OF EUCHROMATIN AND HETEROCHROMATIN

Active or potentially active genes, regulatory regions, and spacers account for 50–80% of the total genome in various organisms, with only 5–10% genes being transcribed in cells of various types. This part of the genome is known as euchromatin. Euchromatin commonly occupies the central region of the nucleus, is decondensed, and shows diffuse staining in the interphase; is replicated in early S; is enriched in acetylated histones; has less methylated DNA (in vertebrates and in plants); and is more accessible for various agents, including nucleases. The remainder (20–50%) of the genome, which has a lower gene content and a higher content of repetitive sequences, is termed heterochromatin and is opposite to euchroma-
Chromatin is a chain of approximately 146-bp DNA fragments, each wrapped around a histone octamer (core nucleosome) consisting of tetramer (H3 + H4), and two dimers, H2A and H2B (see [5] for a review). The fragments are joined together by short (20–60 bp) linkers, which are associated with histone H1 or with HMG proteins. Nucleosome position is commonly sequence-independent. Histones interact with each other and with DNA (mostly through their N-terminal tails), which leads to a higher-order folding of the basic chain. The resulting structure may be stabilized or destabilized by other proteins and/or by posttranslational histone modification. The most common in vivo modifications are acetylation, methylation, phosphorylation, ubiquitination, and polyadenosylphosphorylation (see [6] for a review). Modification usually involves the N-terminal region and changes the net charge and hydrophobicity of a histone. This modulates its interaction with DNA, other histones, and nonhistone proteins. Histone modification is catalyzed by enzymes that introduce or remove certain groups at one or more positions. The enzymes, along with other proteins (e.g., gene-specific activators and repressors), are present in large multimeric chromatin-remodeling complexes.

Chromatin-remodeling complexes of another type facilitate the positioning of nucleosomes behind the replication fork and/or destabilize nucleosomes (see [7, 8] for a review). Various complexes may coexist in cells of one type and even share one or more subunits. The most important feature is that complexes often include a subunit possessing the ATPase activity. The activity may be induced by nucleosomes, DNA, or both in different complexes. A complex may activate as well as suppress many genes. Notwithstanding the great number of relevant works, it is still poorly understood how do the complexes act and, in particular, how are they directed to their target. Presumably, target recognition depends on other regulatory proteins and/or on the nucleotide sequence of corresponding genes.

It is clear that chromatin-remodeling complexes of either type are essential, since nucleosomes commonly, though not always, impede the access of regulatory proteins and RNA polymerase to their target sites in enhancers and promoters. In addition, nucleosomes suppress or even block RNA elongation by RNA polymerase.

In heterochromatin, nucleosomes are structurally similar but are more regularly and widely spaced [9, 10]. Their site-specific and total nuclease sensitivity is substantially lower. These properties suggest that nucleosome positions and internucleosomal interactions are somewhat changed in heterochromatin compared with euchromatin. The changes may be due to interactions with other proteins or to another pattern of histone modification. Thus core histones are enriched in histone H4 acetylated at Lys-12, and histone H3 is often methylated at Lys-9 in heterochromatin, while H4 Lys-5 and Lys-8 are acetylated in euchromatin [6]. It is unclear whether special chromatin-remodeling complexes arrange nucleosomes in the heterochromatin-specific manner. Interestingly, a heterochromatin-like spacing was observed for nucleosomes assembled in vitro by two Drosophila complexes, CHRAC and ACF (see [8, 11] for a review).

Euchromatin and heterochromatin differ in arrangement in the interphase nucleus (see [12, 13] for a review). Euchromatic genome regions are mostly in the interior and have a certain structural organization, since chromosome positions are to an extent fixed. Genes located on different chromosomes and regulated by one group of protein activators may interact with each other, since many activators are nonuniformly distributed in the nucleus and seen as spots upon immune staining (see [14] for a review). In contrast, heterochromatin mostly forms one or more large blocks at the periphery of the nucleus. The functional role and the factors involved in this spatial arrangement of interphase chromatin attract the increasing attention of researchers.

HETEROCHROMATIN PROTEINS

Heterochromatin is concentrated in pericentric chromosome regions, which possibly provides for the normal structure and/or function of the centromere (e.g., see [3]). Pericentric heterochromatin contains satellite DNA and moderate repeats. Some satellite DNA fractions are chromosome-specific, others occur on all chromosomes, though in varying amounts. Formation of detectable heterochromatin blocks was earlier attributed to a physical interaction between repetitive sequences. More likely, the formation is due to abundant heterochromatin proteins, most of which are known to produce multimeric complexes.

Proteins that specifically interact with satellite DNA were found in mammals and in Drosophila [15, 16]. However, the first found specific heterochromatin protein, Drosophila HP1 [17], only nonspecifically, if at all, binds to DNA [18]. However, its presence in heterochromatin is functional, since PEV is suppressed by HP1 mutations and enhanced by an increase in HP1 (see [19, 20] for a review). An HP1 homolog was found in yeast; mouse, human, and Drosophila cells each have three HP1 homologs [21]. On evidence of immune staining, one of these is mostly in pericentric heterochromatin, another is equally associated with both chromatin fractions, and the third is mostly in euchromatin. The three proteins can physically interact with each other. Protein–protein interactions involve at least two HP1 domains, a