Irregular folding of nucleosomes in the cell
Comment on “Cracking the chromatin code: Precise rule of nucleosome positioning” by Edward N. Trifonov

Hideaki Takata, Kazuhiro Maeshima

Biological Macromolecules Laboratory, Structural Biology Center, National Institute of Genetics, Mishima, Shizuoka, 411-8540 Japan
Department of Genetics, School of Life Science, Graduate University for Advanced Studies (Sokendai), Mishima, Shizuoka, 411-8540 Japan

Received 18 January 2011; accepted 19 January 2011
Available online 20 January 2011
Communicated by M. Frank-Kamenetskii

Keywords: DNA; Nucleosome; Chromatin; Chromosome; Irregular folding; Genome; Repetitive DNA sequences

The human body consists of 60 trillion cells, each of which has a total of 2 meters of DNA in the nucleus with a diameter of approximately 10 µm. DNA carries a negatively charged phosphate backbone that produces electrostatic repulsion between adjacent DNA molecules, making it difficult for DNA to coordinate its own folding [1,2]. For the primary level of folding, a negatively charged DNA molecule is wrapped around the basic core, a histone octamer, which consists of the two sets of histones, H2A, H2B, H3, and H4 proteins, and forms a nucleosome structure [3]. Within the nucleosome fibers, where each nucleosome is connected by linker DNA like “beads on a string” [4], how can the histone octamer “see” DNA sequences? Trifonov have provided a thought-provoking review describing the nucleosome positioning signal present on DNA sequences [5]. Surprisingly, three different approaches revealed the same unique sequence motif (CGRAAATTTYCG)n, or (YRRRRYYYYYR)n in binary form [5]. Since genome functions such as DNA replication, recombination, repair and gene transcription are all dependent upon accessibility to DNA, the nucleosome positioning would represent the “chromatin code”. Therefore, the physico-chemical features of the nucleosome positioning can be analyzed to reveal information regarding the regulation of genome functions.

Several interesting questions surrounding nucleosome positioning have emerged from the review [5]: for example, is the identified sequence motif or the 10–11 base pair periodicity conserved in prokaryotes that do not have nucleosomes? Why are the histone proteins in eukaryotes highly conserved whereas genome sequences and the positioning sequence motif are quite divergent among organisms? Which classes of proteins can recognize the positioning motif, the histone octamer itself, histone chaperones, and/or remodeling proteins? Can the epigenetic mark, histone modifications, affect the nucleosome positioning?
Since the nucleosome structure and precise rule of nucleosome positioning are well described by Trifonov [5], we would like to focus more on the higher order structures formed by nucleosomes. As Trifonov reported, when chromatin is digested with “bulky” nucleases to prepare dinucleosomes, significant size variation of the DNA fragments in the dinucleosomes is observed, suggesting that a diversity of linker DNA length exists in the nucleosome arrays [5]. The size variation of the linker DNA would provide structural flexibility for the nucleosome–nucleosome interaction. This implies that an irregular configuration of nucleosomes is formed, as opposed to the formation of regular 30 nm chromatin fibers. The reason for this is as follows: the formation of the 30 nm chromatin fibers requires the selective binding of nucleosomes, which are spatially close on the DNA strand [6]. The size diversity of the linker DNA would disrupt the selective binding of nucleosomes and formation of the 30 nm chromatin fiber. However, the importance of this linker size variation was not appreciated, as the nucleosome has long been assumed to fold into the regular 30 nm chromatin fibers [6].

Recently, when frozen hydrated human mitotic cells were observed using cryo-electron microscopy, higher order structures, including 30 nm chromatin fibers, were not found [7]. Cryo-electron microscopy remains one of the best ways to visualize chromosomes in their intact, native state [8]. Subsequent image processing revealed an 11 nm, but not 30 nm, periodicity in the obtained images [7]. Thus, we proposed that chromosomes consist mainly of irregularly folded nucleosome fibers lacking 30 nm chromatin fibers, which is similar to a “polymer melt” [6,7,9,10]. Based on this and other available observations we also assumed that a similar state exists in the majority of active interphase nuclei [6,10].

The concept of the polymer melt implies that polymer chains exist in a dynamic state [11]: nucleosome fibers may be constantly moving and rearranging at the local level. We also proposed that these local dynamics would have several advantages in various genome functions such as DNA replication, recombination, repair and gene transcription during interphase, and also chromosome condensation and segregation during mitosis [6,7,10]. For the local movements of each nucleosome, it would be intriguing to investigate how flexible the linker DNA, whose persistence length is approximately 150 nucleotides (50 nm), would be in the cells.

Another interesting point regarding the higher order structure of the genome is the repetitive DNA sequences within the genome itself. Upon completion of the human genome sequencing project we learned that the human genome is full of repetitive DNA sequences [12]. Although we proposed that chromosomes consist mainly of irregularly folded nucleosome fibers, it is reasonable to postulate that the nucleosomes on periodic sequences, which can have regular linker DNA, could form regular fiber structures such as the 30 nm chromatin fibers. How often is the identified sequence motif found in the satellite DNA clustered in the centromere [13], and also in the short interspersed nuclear elements (SINEs) and long interspersed nuclear elements (LINEs) [14], which are scattered throughout the genome? Since the nucleosome positioning would represent the “chromatin code”, investigation of the nucleosome positioning in the repetitive sequences might provide a number of clues for answering the following big questions: Are the repetitive sequences really “junk”? and, Do they have any cellular functions?

References