Structural studies of chromosomal building blocks

Srinivas Chakravarthy, Rajeswari Edayathumangalam, Uma M. Muthurajan, Robert K. Suto, Cindy L. White, Karolin Luger
Colorado State University (CSU), Department of Molecular Biology, Fort Collins, CO-80523-1870

INTRODUCTION

The fate of a multi-cellular organism depends on the timed and coordinated readout of its genes. At the molecular level, this requires the accurate transcription of a subset of genes from the total complement that is present in each cell. Consequently, certain diseases and developmental disorders are associated with and even caused by aberrant gene expression. The DNA of a single mammalian cell is over two meters in length, but compacts in the cell nucleus to nearly one millionth of this dimension by a hierarchical scheme of folding and compaction into a highly dynamic protein-DNA assembly termed chromatin. Activation of a gene requires its identification within compacted chromatin. Local unpacking and remodeling of chromatin allows access of regulatory proteins and the transcription machinery, resulting in gene activation. Thus, the organization of DNA in chromatin has profound implications for the regulation of gene expression.

High-resolution crystal structures of nucleosome core particles (NCP) from *Xenopus laevis* reveal an octameric histone core around which 147 base pairs of DNA are wrapped in 1.65 superhelical turns [1] (Fig. 1). The histone octamer itself is composed of two copies each of the four histone proteins H2A, H2B, H3, and H4. Massive distortion of the DNA is brought about by the tight interaction between the rigid framework of the histone proteins with the DNA at fourteen independent DNA binding locations [1, 2]. We have previously determined the structure of a nucleosome containing an essential histone variant, from data collected at the ALS [3]. These studies are now being extended to study nucleosomes from other organisms, and are being expanded into new areas.

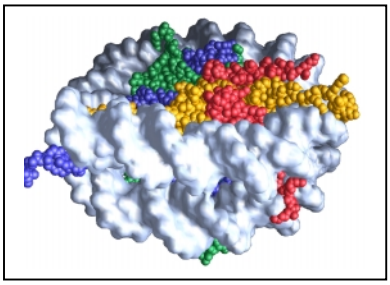

Fig. 1: Structure of the NCP. DNA is shown as a molecular surface in white, protein is shown as a space-filling model in red (H2B), yellow (H2A), blue (H3), and green (H4).

STRUCTURE OF THE YEAST NUCLEOSOME CORE PARTICLE

Fundamental differences between the yeast genome and that of higher organisms suggest that chromatin might be organized in a different manner in yeast. Yeast is a unicellular organism whose entire genome is only ~ 0.5 % the size of that of humans, and its histone proteins are the most divergent among all eukaryotes. The crystal structure of the nucleosome core particle from *Saccharomyces cerevisiae* reveals that the structure and function of this fundamental complex is conserved between single-cell organisms and metazoans [4]. Our results show that yeast nucleosomes are likely to be subtly destabilized as compared to nucleosomes from higher eukaryotes, consistent with the notion that much of the yeast genome remains constitutively open during much of its life cycle. Importantly, minor sequence variations lead to dramatic changes in the way in which nucleosomes pack against each other within the crystal lattice. This has important implications for our understanding of the formation of higher-order chromatin structure and its modulation by post-translational modifications. Finally, the yeast nucleosome core particle provides a structural context by which to interpret genetic data obtained from yeast. Coordinates have been deposited with the Protein Data Bank under accession number 1ID3.

SITE-SPECIFIC RECOGNITION OF NUCLEOSOMAL DNA

The ability of a sequence-specific DNA binding protein to recognize its cognate site in chromatin is restricted by the structure and dynamics of nucleosomal DNA, and by the translational and rotational position of the histone octamer with respect to the binding site. Here we use high-affinity, sequence-specific pyrrole-imidazole polyamides as molecular probes for DNA accessibility in nucleosomes. Sites on nucleosomal DNA facing away from the histone octamer, or even partially facing the histone octamer, are fully accessible and the nucleosomes remain fully folded upon binding. Polyamides only fail to bind where sites are completely blocked by interactions with the histone octamer.

We have determined several high-resolution crystal structures of nucleosome core particles in complex with different hairpin pyrrole-imidazole polyamides (Fig. 2). These structures represent the first nucleosome – ligand co-crystal structures, and provide the first insight into the molecular details of base-specific DNA recognition of nucleosomal DNA. We showed that the binding of ligand does not disrupt any interactions between histones and DNA [5]. All polyamides fit snugly in the minor groove of nucleosomal DNA. Extensive hydrogen bonding between ligand and bases (in addition to non-specific hydrophobic interactions) accounts for the observed high specificity of binding, according to the specificity rules stated by Dervan and colleagues [6]. Local structural changes are imparted on the nucleosomal DNA upon binding. The ligand-induced changes in DNA topology are compensated for by conformational changes in the DNA that are distant from the binding site.

The observed effects of complex formation on the structure of polyamides and nucleosomal DNA have implications for the binding of sequence-specific transcription factors to nucleosomal DNA, and demonstrate a surprising flexibility and plasticity of nucleosomal DNA. Our results demonstrate that much of the DNA in the nucleosome is freely accessible for molecular recognition.

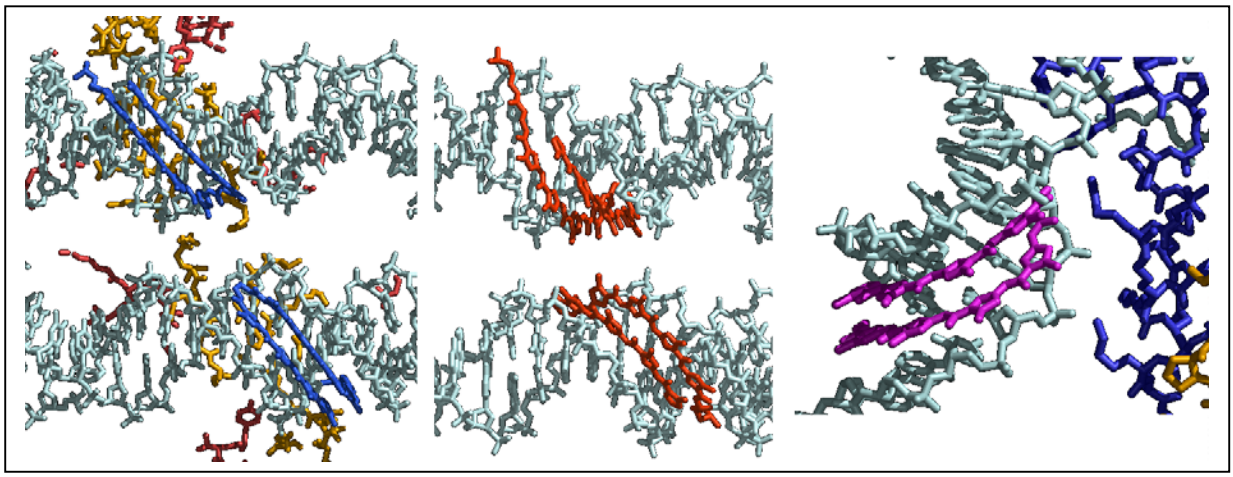

Fig. 2: Close-up view of three NCP-polyamide co-crystal structures. View is from the outside of the nucleosome. Polyamides with different sequence specificities are shown in blue, orange, and magenta, respectively.

This work was funded by grants from the NIH (GM61909), March of Dimes (99-853), Human Frontiers Programs Organization (00059), and by the Searle Scholar award to KL.

Principal Investigator: Karolin Luger, Colorado State University, 970-491-6405, kluger@lamar.colostate.edu.

REFERENCES:

1 Luger, K. *et al*. (1997) *Nature* 389, 251-259

2 Luger, K. and Richmond, T. J. (1998) *Current Opinion in Structural Biology* 8, 33-40

3 Suto, R. K., Clarkson, M. J., Tremethick, D. J. and Luger, K. (2000) *Nat Struct Biol* 7, 1121-1124

4 White, C. L., Suto, R. K. and Luger, K. (2001) *Embo J* 20, 5207-18.

5 Gottesfeld, J. M. *et al*. (2001) *J Mol Biol* 309, 625-39.

6 Dervan, P. B. (2001) *Bioorg Med Chem* 9, 2215-35.