Leading Edge
Review

Cell

A Tale of Chromatin and Transcription in 100 Structures

Patrick Cramer[1,*]
[1]Department of Molecular Biology, Max Planck Institute for Biophysical Chemistry, Am Fassberg 11, 37077 Göttingen, Germany
*Correspondence: patrick.cramer@mpibpc.mpg.de
http://dx.doi.org/10.1016/j.cell.2014.10.047

To celebrate a century of X-ray crystallography, I describe how 100 crystal structures influenced chromatin and transcription research.

Introduction

When Max von Laue first illuminated a crystal with X-rays 100 years ago, it was unclear what the obtained diffraction pattern meant. William Henry Bragg and his son Lawrence, however, soon realized that X-ray diffraction provided information about the inner structure of crystals. After decades of elaboration, X-ray crystallography advanced to become the most widely used method for the determination of 3D structures. Its application to biological macromolecules fostered the development of molecular biology in the second half of the 20th century. Crystallography defined biological paradigms such as molecular recognition, enzymatic catalysis, and allosteric regulation.

Until about 30 years ago, researchers could still follow publication of all new crystal structures of biomolecules. But then the pace at which new structures were solved increased rapidly due to the advent of new enabling techniques. Proteins were obtained in recombinant form, and nucleic acids were synthesized in large quantities. Crystals were cryo-cooled to slow down radiation damage. Synchrotron X-ray sources improved, and fast X-ray detection devices emerged. Modern computers and better software for structure determination became available. By now, 100,000 structures have been deposited in the protein database. Many of these revealed the inner workings of molecular machines, allowing researchers to rationalize phenotypes of mutations and to engineer biological processes.

Crystal structures can be like landmarks. They can guide us on our way toward a better understanding of a biological process (Shi, 2014 [this issue of *Cell*]). Landmark structures not only disclose some of life's secrets, but they also open up new frontiers. Here, I describe many of the structures that I consider to be landmark structures in the biology of chromatin, transcription, and epigenetics. I hope the resulting list of about 100 crystallographic structures, along with several structures obtained by other methods, exemplifies how structural information influenced the community and led to new concepts.

How DNA Is Structured

The proposal of the double-helical structure of DNA relied on X-ray diffraction patterns of DNA fibers obtained in the middle of the last century (Watson and Crick, 1953) (Figure 1). The direct observation of a nucleic acid duplex, however, had to await the crystal structure of a transfer RNA (tRNA) from yeast in 1974 (Robertus et al., 1974) (Figure 1). The structure of a B-DNA duplex was solved only after DNA synthesis methods became available (Wing et al., 1980) (Figure 1). Crystal structures of DNA in A-form (Shakked et al., 1981) and Z-form (Wang et al., 1979) highlighted the sequence-dependent conformational flexibility of DNA.

In eukaryotic nuclei, DNA is packaged with histone proteins into chromatin. The fundamental unit of chromatin, the nucleosome core particle, was elucidated structurally in 1984 at a resolution of 7 Å (Richmond et al., 1984). When the resolution reached 2.8 Å, a detailed view of the nucleosome emerged that revealed the DNA conformation and DNA interactions with histones (Luger et al., 1997). The nucleosome core structure confirmed the structure of the free histone octamer (Arents et al., 1991). It further showed that the histone protein tails emerged between and around DNA duplexes to become available for interactions with other nucleosomes or chromatin.

The structure of higher-order chromatin is dynamic, but a complex of four nucleosomes could be crystallized and showed two stacks of nucleosomes and DNA that zigzagged between them (Schalch et al., 2005). Electron microscopy revealed how such tetranucleosome units may be arranged within a 30 nm fiber of chromatin (Song et al., 2014) and provided evidence for helical order in such fibers (Scheffer et al., 2011). Another electron microscopic study provided an alternative fiber model (Robinson et al., 2006). These results explained how extended DNA molecules can be packaged in the cell nucleus but also provided models for how chromatin regulates the accessibility of genes and their transcription. Whereas a compact chromatin structure can cause gene repression, gene activation requires chromatin opening and assembly of the transcription machinery at the promoter.

How DNA Is Recognized

To enable transcription, cells use transcription factors that bind to specific DNA sites. The first crystal structures of transcription factors included the bacterial catabolite activator protein CAP (McKay and Steitz, 1981) (Figure 1) and the bacteriophage lambda proteins cro (Anderson et al., 1981) and repressor (Pabo and Lewis, 1982). These structures contained helix-turn-helix regions that were involved in DNA binding and led to the concept of DNA-binding protein motifs. The studies of the bacteriophage proteins required protein overexpression because these transcription factors could not be isolated from natural sources in quantities required for structural studies.

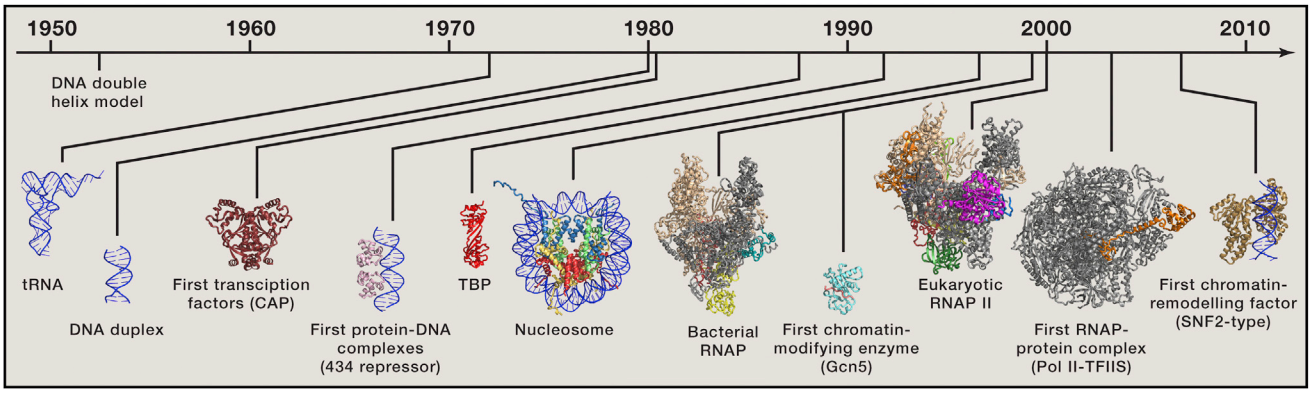

Figure 1. A Selection of Landmark Crystal Structures in the Biology of Chromatin and Transcription

From left to right, the depicted structures are yeast tRNA, a DNA duplex, the bacterial transcription factor CAP, the bacteriophage 434 repressor protein in complex with DNA, the eukaryotic TATA-binding protein TBP, the nucleosome, the bacterial RNA polymerase, the histone acetyltransferase Gcn5, the eukaryotic RNA polymerase 10-subunit core enzyme, the complete 12-subunit RNA polymerase II complex in complex with transcription factor TFIIS, and an archaeal Swi/Snf-type ATPase resembling the catalytic subunit found in many chromatin remodeling complexes. DNA is shown in blue, and proteins are depicted as ribbon models in different colors. For details, please refer to the text.

Structures of DNA-bound transcription factors led to the concept of sequence-specific DNA recognition. DNA complexes of the repressor proteins from bacteriophages 434 (Anderson et al., 1987) (Figure 1) and lambda (Jordan and Pabo, 1988) and of the 434 cro protein (Wolberger et al., 1988) revealed protein helices bound in the DNA major groove. DNA-binding helices were also observed in structures of homeodomains (Kissinger et al., 1990; Otting et al., 1990; Qian et al., 1989). In a ''leucine zipper'' of the GCN4 factor, protein helices in the DNA major groove were extended and used for factor dimerization (Ellenberger et al., 1992; König and Richmond, 1993). The transcription factors recognized target sequences via interactions of amino acid residues with DNA base edges. Such ''direct readout'' can be complemented by ''indirect readout'' of the DNA conformation via protein-DNA backbone contacts (Lesser et al., 1990; Otwinowski et al., 1988).

Later protein-DNA complex structures revealed a variety of DNA-binding structural motifs. The transcription factor NF-κB uses a β barrel fold to contact DNA via protein loops (Becker et al., 1998; Ghosh et al., 1995; Müller et al., 1995). Transcription factors of the zinc finger family recognize DNA with small protein folds that are stabilized by zinc ions (Fairall et al., 1993; Luisi et al., 1991; Marmorstein et al., 1992; Pavletich and Pabo, 1991). Zinc fingers were later used for the design of proteins with new DNA-binding specificities (Choo et al., 1994). This catalyzed the development of protein and genome engineering as new research fields. Zinc fingers were also present in Klf4 (Schuetz et al., 2011), which, together with transcription factors Oct4, Sox2, and c-Myc, enables reprogramming of the genome and generation of induced pluripotent stem cells.

Crystallography also showed how transcription factors bind to adjacent DNA sites for combinatorial gene regulation. The DNA-bound structures of yeast MATα2 interacting with MATa1 (Li et al., 1995) and with MCM1 (Tan and Richmond, 1998) revealed factor-factor interactions that underlie synergistic DNA binding. This concept held for human transcription factors (Piper et al., 1999). Oct4 and Sox2 can also bind to neighboring DNA sites (Reményi et al., 2003). Crystallography also led to a model of an ''enhanceosome'' containing eight transcription factors bound to DNA (Panne et al., 2007). Here, binding of one factor induces a DNA conformation that promotes binding of a neighboring factor.

Transcription factors can also bend DNA dramatically. The bacterial CAP protein bends DNA by 90 degrees to enable specific DNA recognition (Schultz et al., 1991). The eukaryotic TATA box-binding protein (TBP) also introduces a 90 degree bend into DNA (Kim et al., 1993a, 1993b). The integration host factor (Rice et al., 1996) and the mitochondrial transcription factor A (Ngo et al., 2011; Rubio-Cosials et al., 2011) can even bend DNA by 180 degrees, inducing a U-turn. To achieve DNA bending, proteins can use two strategies. They can insert amino acid residues like wedges between DNA base pairs to disrupt base stacking and may also neutralize backbone charges on one side of DNA, which induces bending due to the repulsion of phosphates on the opposite strand.

How DNA Binding Is Regulated

Crystallography further established concepts that explained bacterial gene regulation. In the *trp* operon, the Trp repressor protein inhibits expression of enzymes required for tryptophan biosynthesis when enough of the amino acid is available. The structure of the Trp repressor revealed a homodimer with a DNA-binding helix in each monomer (Schevitz et al., 1985). Binding of the regulator tryptophan alters the relative position of the two helices to enable DNA binding and gene repression (Otwinowski et al., 1988). These studies also showed that water molecules in the protein-DNA interface may contribute to sequence-specific DNA recognition.

Many transcription factors contain not only a DNA-binding domain but also additional domains that can activate transcription or bind other transcription factors or small-molecule regulators. In the bacterial *lac* operon, the Lac repressor protein binds DNA to control the expression of enzymes involved in lactose metabolism. The lac repressor contains a domain that binds

Download English Version:

https://daneshyari.com/en/article/2035367

Download Persian Version:

https://daneshyari.com/article/2035367

Daneshyari.com