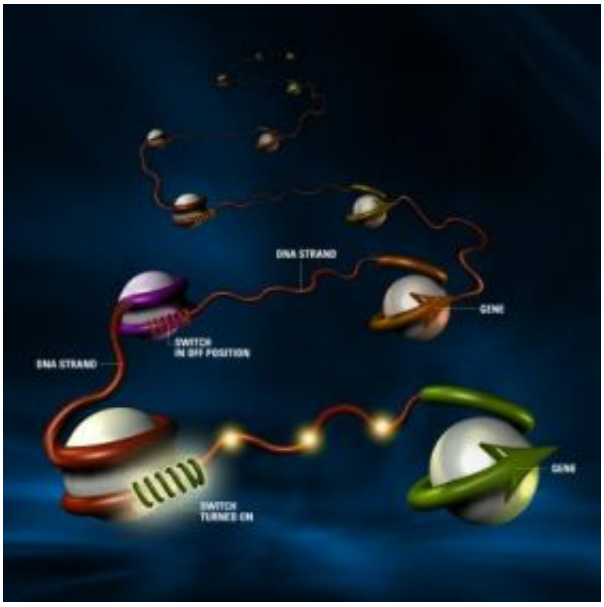
Medical Xpress

Scientists reveal structure of gateways to gene control

March 28 2007



This artist's illustration shows a portion of the genome. A long DNA molecule (red cord) is wrapped up into nucleosome structures with a histone-protein core (white spheres). The gateway to gene transcription spans the control switch (coils) on the left nucleosome to the beginning of the gene (green arrow) on the right nucleosome. Credit: Christina Ullman

Scientists at Penn State University will reveal in the 29 March 2007 issue of the journal *Nature* the first complete high-resolution map of important structures that control how genes are packaged and regulated throughout an entire genome. "For the first time, we are seeing in very high resolution on a genome-wide scale how nucleosomes control the

expression of an organism's genes," said B. Franklin Pugh, professor of biochemistry and molecular biology and the study's lead investigator.

The map pinpoints the locations of certain key gene-controlling nucleosomes -- spool-like structures that wrap short regions of DNA around a protein core. The research suggests how these nucleosomes, positioned at important transcription-promoter sites throughout the cell's DNA, control whether or not a gene's function can be turned on in a particular cell.

The study's many surprising findings together reveal an intimate relationship between the architecture of nucleosome structures and the underlying DNA sequences they regulate. "We now know exactly where these nucleosomes are positioned on the DNA molecule and which DNA building blocks they have wrapped up under their tight control," Pugh said. Among those building blocks, Pugh and his colleagues revealed the architecture of a critical gateway, controlled by the nucleosome, which must be unlocked before a gene can be transcribed.

The study revealed that almost all genes have the same kind of structure where transcription begins, that this beginning contains a critical gateway for transcription, and that the transcription gateway of each gene almost always is located at the same place on a nucleosome. The researchers also discovered some genes whose pattern is somewhat different from this norm, and these unusual sequences also are reported in the Nature paper. "We previously had a low-resolution idea that these structures all could be roughly in the same position, but now this high-resolution map makes it very clear that they really are in exactly the same position. It's a remarkably consistent arrangement," Pugh said.

The study also revealed that the nucleosomes at the transcription-promoter control centers occupy several overlapping positions on the DNA molecule, typically 10 base pairs apart, which exactly matches the

periodic rotation of the DNA double helix. "It is striking how well these positions match with the architecture of the DNA as it wraps around the nucleosome's protein core," Pugh said.

This result powerfully simplifies previous theories about the possible architecture of gene packaging. "There is a certain DNA sequence that shapes the gene's architecture in the same way, producing the same structure in every gene," Pugh said. The overall sequence of DNA building blocks is different in each gene, but the underlying architecture is the same."

To obtain their high-resolution map, the researchers first isolated 322,000 nucleosomes from the 6,000 regions that control gene transcription in the DNA of baker's yeast, S.cerevisiae, an organism widely studied as a model of how human cells work. These promoter nucleosomes are the only ones in the yeast DNA that contain in their core a histone protein called H2A.Z. Led by Pugh and Stephan Schuster, associate professor of biochemistry and molecular biology, the Penn State research team then used antibodies that bind only to this H2A.Z protein as a tool for separating all these promoter nucleosomes from the other parts of the yeast's DNA. Next, the team used a state-of-the-art DNA-sequencing machine to identify, or "read," the sequence of base-pair building blocks along the DNA of each of the H2A.Z nucleosomes. The scientists then pinpointed the original location of the H2A.Z nucleosomes by matching the sequence of each one with the identical sequence on the previously published yeast genome. "Obtaining the exact DNA sequences for all these nucleosomes allows us to precisely map their positions across the entire genome," explains Schuster. The map reveals, for the first time, precisely which DNA sequences are part of the control-center's H2A.Z nucleosome for each gene in the yeast genome. Also for the first time, researchers now have a clear picture of how H2A.Z nucleosomes help to control whether or not a gene can be turned on.

Another discovery is that transcription-control centers tend to be located on the outside edge of the nucleosome and tend to face outward on the DNA helix, allowing the cell's transcription proteins to find them more easily. "This arrangement makes sense, because when signaling proteins arrive at a control center they are well situated to help push the nucleosome out of the way so the reading of the gene can begin," Pugh said.

"Previous research had indicated that DNA sequences located upstream of a gene might be a region that controls whether that gene is read or not, but we did not know the architecture of those sequences -- whether they were exposed and therefore ready for work. Now we know that the gateway to transcription is a part of this control region and that the nucleosome keeps it locked so the gene cannot be turned on until it is needed," Pugh said. When the gene is needed, the cell's molecular machinery loosens the DNA wrapping around the nucleosome, unlocking the transcription gateway to give access to the cell's molecular transcription machinery. "We think that the function of the nucleosome is to control the gateway to transcription," Pugh said.

The research reveals how the pieces of DNA that regulate genes at the transcription-promoter sites are packaged on nucleosomes. The knowledge that these sites are located on the outside edge of the nucleosome spool will help to focus research designed to manipulate gene expression. "Our study has provided a much clearer picture of the architecture of the DNA in the control regions, allowing us to understand much better how genes are regulated, which is important because gene regulation is a critical process for the survival of living things," Pugh explains.

The paper by Pugh's team marks the leading edge of a new wave of anticipated discoveries about gene regulation, made possible by recently developed laboratory equipment for high-volume, or massively parallel,

DNA sequencing. "Traditional DNA sequencing methods processed one DNA strand at a time, but now we can sequence hundreds of thousands of DNA strands at once, rapidly learning incredible amounts of new information," Pugh said.

The knowledge that most genes are packaged basically the same way is powerful information with implications for future research and potential applications. "One implication that I think is important is that we now have a better idea about how packaging the DNA in nucleosomes controls the expression of a gene," Pugh said. "We don't yet know where all the important gene-regulation features are located on the DNA molecule, but now we know we should start looking for some of them on the edges of nucleosomes," Pugh said. "We might even discover some sites that regulate genes that we didn't even know existed."

Source: Penn State