HHMI
HOWARD HUGHES MEDICAL INSTITUTE

RESEARCH NEWS

OCTOBER 29, 2009

Studies Begin to Shape New Image of DNA

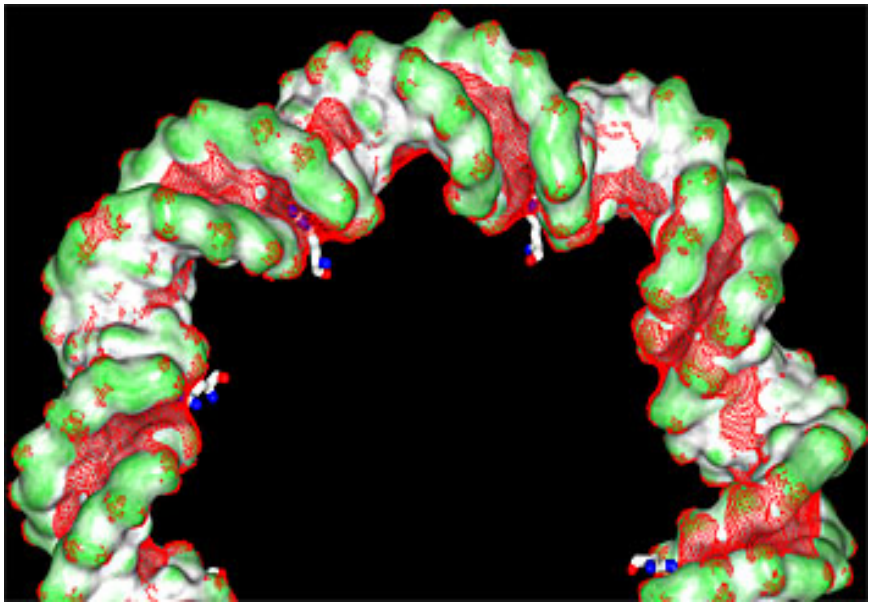

Image Title: The figure illustrates the molecular shape of a region of nucleosomal DNA when wrapped around the histone core, with the narrow minor groove in dark grey. The red mesh shows a surface with negative electrostatic potential. The shape of narrow minor groove regions induces an enhanced negative electrostatic potential, which is read by histone arginines. Such interactions between the protein and DNA contribute to the stabilization of the nucleosome core particle. - Sean West

Most of us carry a mental picture of DNA in its iconic form – the famous double helix unveiled by Francis Crick and James Watson. But researchers are beginning to develop a new picture of DNA that shows the molecule's more dynamic side, which is capable of morphing into a large number of complex shapes. This shape-shifting ability permits proteins to attach and read the right region of DNA so genes can be turned on or off at the proper time.

The findings show that proteins are adept at reading nuances in the shape of the double helix. Those variations in shape transmit information about where

proteins need to bind to make sure the right genes are activated or silenced during development.

“The ideal double helix should not be viewed as a rigid entity but rather seen as a first approximation to a large set of more complex shapes that are recognized by proteins so as to bind to DNA in a sequence-specific fashion,” said Howard Hughes Medical Institute investigator Barry Honig at Columbia University.

"If you now realize that those nucleotides determine which proteins bind the DNA, and they do it in part through their effect on shape, you begin to understand how sensitive and subtle the DNA structure really is, and how this in turn affects how it's being read."
- Barry Honig

Honig and his colleagues have discovered a new mechanism by which proteins recognize specific regions of DNA. Their research is reported in the October 29, 2009, issue of the journal *Nature.*

Even for seasoned biologists, the endless stream of A, T, G and C nucleotides in the genetic code can look like a string of letters in a long book that makes no sense. It can be easy to get lost. But proteins that bind to DNA to turn genes on or off have an innate intelligence -- they know how to read the book. One of their “secrets” is to follow a set of instructions that are hidden inside the DNA sequence.

Scientists have long known that specialized DNA-binding proteins, such as transcription factors that activate and repress genes, look for their docking sites on DNA by scanning the genome for a specific nucleotide sequence that says “bind here.” When proteins recognize that sequence, they bind to DNA and begin to do their jobs. But over the last 20 years, researchers have accumulated evidence that the physical shape of DNA can also influence where and when proteins attach to DNA.

The new studies published in *Nature* by Honig and his colleagues, Remo Rohs, Sean West, Alona Sosinsky, Peng Liu and Richard Mann, extend those results and describe a new recognition strategy that proteins use to identify and bind to DNA. The coiled, complementary strands of DNA form 'major' and 'minor' grooves, to which proteins can bind. The recognition strategy identified by Honig’s team depends on proteins’ ability to read DNA shape – specifically, the width of the minor grooves.

The work originated with an earlier study of how a particular transcription factor, known as a Hox protein, achieves DNA binding specificity. Hox proteins are important for determining the overall body plan in all animals and need to bind to their DNA targets with great specificity, so that they can control the activity of only the appropriate genes. Previous work from a number of labs had revealed common features that all Hox proteins share when they bind to DNA, but scientists did not know how individual Hox proteins distinguish between different binding sites.

In 2007, Honig and collaborators in the labs of geneticist Richard Mann and x-ray crystallographer Aneel Aggarwal compared the structures of two different DNA sequences bound to a Hox protein. One DNA sequence was highly selective for a particular Hox protein, while the other was able to bind multiple Hox proteins. They found that the selective DNA target had a narrower minor groove than is typical for double-helical DNA. In contrast, the less selective DNA sequence had a different, less narrow, minor groove shape.

“The question for us was, why is that important?” Honig says. “What we showed is that the electrostatic potential of the DNA – which is used to attract positive charges – is stronger when a groove is narrow.” Intriguingly, the width and shape of the grooves are affected by single-letter changes in the DNA sequence. In other words, the sequence of DNA determines its precise shape, which then provides a target for the proper protein to bind.

“It is quite surprising actually,” Honig says. “If you now realize that those nucleotides determine which proteins bind the DNA, and they do it in part through their effect on shape, you begin to understand how sensitive and subtle the DNA structure really is, and how this in turn affects how it's being read.”

With proof of this new type of DNA recognition strategy in hand, the team next set out to see whether their concept was generalizable – whether it would apply to proteins outside of the Hox family.

They probed a database of molecular structure information about DNA-protein complexes, looking at how the DNA’s sequence and structure matched up with each of the protein’s amino acid building blocks. They found that narrow minor grooves tended to attract parts of the protein that contained the amino acid arginine, which is positively charged. They saw that there are many arginine binding sites in DNA that have narrower minor grooves and that these have more negative electrostatic potentials that attract positively charged regions of proteins.

“The proteins are actually reading the shape of the DNA through its effect on electrostatic potential,” Honig says. “Sequence determines shape, which determines the affinity for arginines --a mechanism made possible because DNA does not form a perfect double helix.”

Honig's group plans to examine RNA, the intermediate molecule produced when DNA's information is translated into protein. He and his colleagues will be looking for similar relationships between sequence, shape and electrostatic potential. They also hope to use the relationship between DNA sequence and shape to predict which sequences of the DNA are bound by transcription factors and which proteins recognize these regulatory regions.

Honig says he and other researchers are now thinking about DNA in new ways: A separate research group recently compared the shapes of DNA in different species and concluded that these molecular nuances are shaped by evolution. "Taken together with our studies, the results offer a new way to think about DNA," Honig says.