reports that the fruit fly's ORC2 gene, when put into mutant yeast, can replace some of the functions of yeast ORC2. "Since the complex is so conserved, we think this means that the basic mechanism of [origin site] recognition is conserved," Botchan says.

And the same may be said for the proteins that combine with ORC. One clue comes from a phenomenon called "licensing," discovered in the late 1980s by ICRF's Blow and Ronald Laskey of the University of Cambridge, U.K. Their studies of DNA replication in extracts of *Xenopus* eggs indicated that initiation of DNA replication requires some factor that can only gain access to the chromosomes during mitosis, a time when the nuclear membrane has broken down. Then, once replication has taken place, this licensing factor, as they called it, is somehow lost or inactivated. Once the nuclear membrane has regenerated, it can't reach the chromosomes again until another round of mitosis begins. This behavior, Blow and Laskey proposed, would ensure that DNA could not replicate more than once in any given cycle.

Now some of the molecules needed for licensing are turning out to belong to a family of proteins, known as MCM proteins, that was originally discovered in yeast but is now known to be present in all eukaryotes. This spring, the Blow and Laskey teams, and also that of Haruhiko Takisawa of Osaka University in Japan, showed, among other things, that antibodies to the *Xenopus* MCM3 protein remove licensing-factor activity from *Xenopus* egg extracts, while adding back the MCM proteins restores licensing activity.

The behavior of the MCM proteins in *Xenopus* resembles that of the proteins that combine with yeast ORC to form the prereplication complex. And because MCM proteins are found in yeast, the supposition is that they may in fact be part of that complex. Indeed, there may be another resemblance as well. Blow's group has found an as-yet-unidentified material in *Xenopus* that's needed for replication initiation, in addition to the MCM proteins. "We speculate that it's going to be a frog CDC6," he says, referring to the protein needed for prereplication complex assembly in yeast.

Taken together, the work in yeast and the other eukaryotes is building a picture of a complex initiation machinery centered on ORC. As Stillman describes it, "The concept has emerged that ORC is a landing pad for a whole bunch of other proteins" that come and go at precise times during the cell cycle to regulate DNA replication. And even though researchers still have a way to go to work out the precise identities and functions of all those proteins, they have also come a long way. "Two years ago we had no idea that all these proteins were involved," Stillman says. "Now at least we know what to focus on."

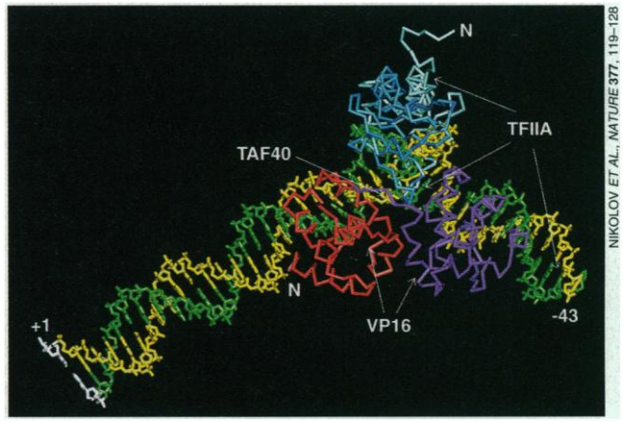
–Jean Marx

ARCHITECTURAL PROTEINS

Protein Sculptors That Help Turn On Genes

Inside a living cell lies perhaps the only place on Earth where a serious accounting problem is solved by something resembling art. In mammals, more than 100,000 genes must blink on or off along the chromosomes at precisely the right times, at just the right intensity, for the body to develop and function normally. But the protein switches that bind to DNA and turn genes on are in such short supply that if each of these proteins, known as transcription factors, had to act by itself, the cell would have far too few of them to give precise control of so many genes. And that's where the cell's artistic talents come into play.

Molecular biologists have known for about a decade that, rather than deploying the gene-regulating proteins one at a time, the cell dispatches groups of them in various combinations to turn on particular genes, thereby greatly expanding the number of distinct switches that can be formed. More recent work now suggests that cells go one step further, using "architectural proteins" to sculpt many of these protein clusters into precise three-dimensional shapes. The architectural proteins, which are themselves part of the transcription factor complexes, usually do this by binding to the DNA and bending it, often sharply, thereby bringing together other members of the complex, which are bound to separate segments of a gene's control region.



Around the bend. DNA bound to TBP *(blue)* and TFIIB *(red, magenta)* is L-shaped. The arrows indicate where other transcription factors in the complex bind.

Without this sculptural talent, the transcription factor clusters would not function normally. "The three-dimensional organization of these complexes is critical for controlling the expression, and level of expression, of various genes," notes Stephen Burley of Rockefeller University. As a result, architectural proteins are turning out to play a key role in physiological processes ranging from inflammation to sex development.

The first indications of the existence of architectural proteins came in the early 1980s when Donald Crothers of Yale University discovered that certain proteins from the bacterium *Escherichia coli* could bend DNA, in one case wrenching 90-degree kinks in the molecule. The significance of that bending wasn't understood at the time, but in the course of his research, Crothers pioneered a method of inferring the existence and location of bends in DNA. The method sifts out bent DNA molecules from straight ones by exploiting the fact that in gel electrophoresis, bent DNA moves more slowly than does straight DNA of the same size.

Crothers's method was soon put to use by Howard Nash and his colleagues at the National Institute of Mental Health in Bethesda, Maryland, who were the first to identify a function for an architectural protein. When certain viruses infect *E. coli*, they insert their DNA into the bacterial genome. At the time, in the mid-1980s, a bacterial protein called integration host factor (IHF) was known to be needed for the insertion, although no one knew exactly what it did. Nash was trying to find out.

After ruling out the possibility that IHF was an enzyme, Nash recalls, he began investigating a structural role for the molecule. He applied Crothers's method and found that IHF bends DNA. He also noted that IHF binds to DNA's minor groove, the small space between its helical twists, an unusual trait for a DNA-binding protein. Nash's subsequent work, together with that of Arthur Landy's group at Brown University, showed that as a result of the bending, two other proteins, bound to separate segments of the bacterial DNA, were able to meet. This interaction, the researchers found, was necessary for viral DNA insertion.

But the DNA bending induced by IHF soon proved to have a much broader role in normal bacterial functioning. In 1990, Sydney Kustu's team at the University of California, Berkeley, noted that the protein

strongly stimulates the transcription of a cluster of genes involved in nitrogen metabolism in *Klebsiella pneumoniae*, a close relative of *E. coli*. What's more, they showed that it bends the DNA at a site located between two other proteins needed for the first step in gene activity, the copying of the genes into RNA: the enzyme RNA polymerase, which does the copying, and the nitrogen fixation regulatory protein, NIFA. As a result of the bending, the two proteins moved together. That meeting, Kustu proposed in a landmark paper in *Cell*, triggered transcription. Kustu's model set a precedent, says gene transcription expert Tom Maniatis of Harvard University: "It was the first example of a DNA-bending mechanism required for gene activation." But it wasn't the last.

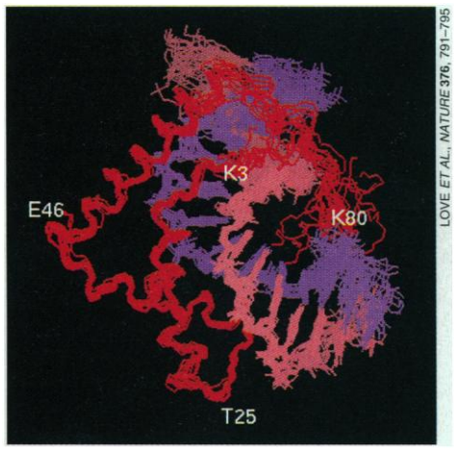
Two years later, a group led by Rudolf Grosschedl of the University of California, San Francisco (UCSF), found evidence that DNA bending plays a role in eukaryotic gene transcription. While studying a protein called LEF-1 (for lymphoid enhancer-binding factor 1), which regulates a gene crucial for the differentiation of key immune cells called T cells, Grosschedl and his colleagues noticed that it shares a peculiar property with IHF: binding to DNA's minor groove. Using Crothers's method, the UCSF team found that LEF-1, like IHF, bends DNA sharply.

Those similarities prompted the researchers to see whether LEF-1 might perform a function similar to that of IHF in assembling protein complexes, Grosschedl says. And they found indirect evidence that it does. They showed that the protein, which they obtained from the mouse, could substitute for IHF in promoting integration of viral DNA into the *E. coli* genome, an action that depends on DNA bending.

Since then, Grosschedl's group has obtained biochemical evidence that DNA bending is also important for LEF-1's normal function as a transcription factor. Those results indicated that LEF-1 plays matchmaker for two critical transcription factors by bending the DNA between them. And they noted that only when the match was made would transcription proceed at a normal pace.

But the most direct evidence for LEF-1's architectural role came just a few months ago. In the 31 August issue of *Nature*, Peter Wright and his co-workers at Scripps Research Institute in La Jolla, California, who worked in collaboration with Grosschedl's group, published an atomic-scale structure of LEF-1 interacting with DNA. The structure, obtained by nuclear magnetic resonance (NMR) methods, reveals that LEF-1 contacts DNA's minor groove, shoves a side chain of one its amino acids between a DNA base pair, and bends the DNA helix a spectacular 120 degrees.

That picture had in fact been forecast a few months earlier when a very similar DNA-bending portrait was found for another protein that, structurally speaking, looks very much like LEF-1. The protein, called SRY (for sex-determining region, Y chromosome), turns on the genes necessary for the formation of the testis and other male sex organs in mammals (*Science*, 29 September, p. 1824). This past June, Marius Clore, Angela Gronenborn, and their colleagues at



Twisted. Binding of the transcription factor LEF-1 *(red)* puts a major kink in DNA.

the National Institutes of Health (NIH) determined the NMR structure of SRY bound to DNA, showing that the protein bows the DNA about 80 degrees. Clore and his colleagues believe the SRY-induced bend snaps together distant proteins into a compact complex that triggers transcription. However, the other components of that complex have not been identified.

While SRY and LEF-1 buttress protein complexes that regulate specific genes, another architectural protein, the transcription factor known as the TATA-binding protein (TBP), acts on many different genes. TBP typically binds to the so-called "TATA box," a short sequence of alternating thymidine (T) and adenosine (A) nucleotides found at the start of most genes. This protein serves as a generic scaffolding that can support different molecular sculptures depending on the gene. "It's like a hook on a wall. You can hang different pictures from it," explains structural biologist Milton Werner, who works in Clore and Gronenborn's lab at NIH.

About 50 proteins fold into each TBP complex, so determining the interactions of all of them is a daunting task. Nevertheless, researchers got their first inkling of how it might form when they determined that TBP bends DNA. That happened in 1993 when two groups—Peter Sigler's at Yale and Burley's at Rockefeller—simultaneously published x-ray crystallography structures of TBP clinging to DNA. The papers revealed "an amazing bend" of 70 degrees in the DNA, says Burley.

In more recent work, Burley's team has gone one step further by showing how at least one part of the transcription machinery fits into the bend. In the September 14 issue of *Nature*, Burley and his colleagues published the crystal structure of DNA bound both to TBP and TFIIB (for transcription factor IIB), a cog in the general transcription machine that includes RNA polymerase. That structure shows an L-shaped fragment of DNA created when TBP clutches the L's outside corner; the TFIIB nestles in the nook of the L. In this case, the bending creates a unique structure that TFIIB recognizes and can hang on to. "It's a beautiful story," comments Werner. "It indicates the structural context of how other proteins assemble on a scaffold."

And the architectural protein story doesn't stop there. At least eight other proteins are accepted members of the group, although their mechanisms of action are less well understood than those of LEF-1, SRY, and TBP. One of these is HMGI(Y), which is required to turn on genes needed for inflammatory responses, such as those encoding cell adhesion molecules that enable white blood cells to stick to, and then slip through, blood vessel walls and into injured tissue.

Recent biochemical work by Maniatis's group suggests that HMGI(Y) binding distorts the DNA, although this bending is much less than that produced by the other architectural proteins. Because HMGI(Y) seems to interact directly with other proteins in the complex, Maniatis thinks this architectural protein functions primarily as a clasp that helps transcription factors bind to DNA and links those factors to each other, rather than using its DNA-bending powers to bring transcription factors together. But until scientists determine the crystal structure of HMGI(Y) interacting with the double helix, the anatomy of this complex will remain unresolved.

Indeed, many questions remain about architectural proteins and their mechanisms of action. One big mystery, for example, is how transcription complexes like the one organized by LEF-1 activate RNA polymerase at the gene they regulate. Amazingly, some of these complexes form thousands of base pairs away from the site where the polymerase binds. In the case of LEF-1, the gap is about 4500 base pairs.

Also unclear is how many architectural proteins exist, although researchers suspect the total may soar into the hundreds. "It remains to be seen how many genes are dependent on architectural proteins," says Grosschedl. "But more and more examples are coming to light." Still, even the examples that have seen the light so far have been enough to illuminate the artistic genius of the cell.

–Ingrid Wickelgren

Ingrid Wickelgren is a free-lance writer based in Brooklyn, New York.