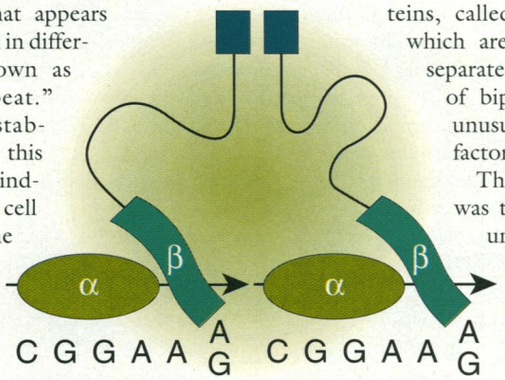
New Role Found for a Common Protein "Motif"

A sequence called the "ankyrin repeat" may be a versatile tether for joining the subunits of regulatory proteins

USUALLY WHEN A NEW GENE IS CLONED, there's but one way the scientists who cloned it can figure out what the protein it encodes does: By finding a sequence similarity to a protein of known function. Every so often, though, the tables are turned when the new protein provides an unexpected insight into the function of sequences already in the databases. That's what Steven McKnight's group now appears to have achieved with the cloning of the genes for two proteins that pair up to make one of the factors the cell uses to regulate gene activity.

Already in the databases was a particular structural motif—a short sequence of amino acids that appears over and over again in different proteins—known as the "ankyrin repeat." The only well-established function for this sequence was in binding proteins to the cell skeleton—until the McKnight group stumbled on it in one of the two transcription factor proteins.



Holding on. *The ankyrin repeat segment* (long rectangle) *binds the* β *chain of GABP to the* α *subunit, so that the complex can bind to DNA.*

Much to their surprise, the researchers found that the ankyrin repeat helps to bind the two proteins together—the first time it's been shown to do that in a transcription factor. Indeed, based on the current work, which is described in articles on pages 762 and 789, McKnight is now proposing that the ankyrin repeat has a still broader role: It may serve, he says, as a general molecular tether, binding together proteins of several types, including nuclear proteins that help to regulate gene expression and embryonic development, as well as cytoplasmic proteins that participate in cell signaling pathways.

"What's really exciting is the diversity of the proteins that the repeats have been found in, and that their function is not restricted to one cellular compartment," says Albert Baldwin of the University of North Carolina in Chapel Hill, who is also studying proteins that contain ankyrin repeats. Ankyrin repeats may, therefore, be part of an extremely versatile system for generating proteins that can regulate numerous cell activities.

McKnight and his colleagues, who work at the Howard Hughes Laboratories at the Carnegie Institution of Washington in Baltimore, found the ankyrin repeat in a hitherto obscure transcription factor that goes by the name GA binding protein (GABP) because it binds to DNA sequences rich in guanine and adenine bases. It was but one of many surprises they got while studying the factor. The first was the discovery that GABP contains two distinct proteins, called alpha and beta, which are encoded by two separate genes. That kind of bipartite structure is unusual for transcription factors, McKnight says.

The second surprise was that the alpha subunit has a sequence similarity to the protein product of the *ets* oncogene. And a third came when the researchers found four copies of the ankyrin repeat, which is 33 amino acids long, in the beta protein. That was surprising because transcription factors work in the nucleus and cell biologists have generally considered the repeat to be a structural characteristic of proteins, such as its namesake ankyrin, that are located in the cell cytoplasm.

And that wasn't the end of the surprises. When the McKnight group removed the ankyrin repeats from the beta protein, they found that it could no longer bind the alpha subunit. The conclusion? The ankyrin repeats are part of the binding site for the alpha protein, which does not contain the repeats. With other protein-binding motifs, McKnight says, "like mates with like." That's what happens, for example, between proteins with the leucine zipper motif, which was also discovered in the McKnight lab, or with the helix-loop-helix motif. The fact that ankyrin repeats don't have to match up with other ankyrin repeats may have been one reason that the sequence's role as a protein-binding motif for transcription factors was missed, even though Kim Nasmyth and Linda Breeden, who were then working at the Molecular Research Council labs in Cambridge, England, had discovered the repeats in two yeast transcription factors in 1987.

A more important reason may have been, however, that the presence of the repeats in transcription factors was simply overshadowed by results with the cytoplasmic proteins. A key finding came in 1990, for example, when Vann Bennett's group at Duke University cloned the ankyrin gene and then went on to show that the repeat is needed to attach ankyrin to tubulin in the cell skeleton. Later, ankyrin repeats turned up in additional cytoplasmic proteins, including the Notch protein of the fruitfly and the *lin12* protein of the roundworm *Caenorhabditis elegans*, both of which are important for the normal embryonic development of those organisms.

And the idea that ankyrin repeats are important for cytoplasmic proteins really got a boost earlier this year, McKnight says, when Baldwin and Stephen Haskill, who's also at North Carolina, cloned the gene for a protein called IκB and found that this protein also contains ankyrin repeats. IκB is so called because it inhibits the action of a transcription factor called NFκB, a key regulator of immune cell activities. The inhibitor apparently works by binding the NFκB and keeping it sequestered in the cytoplasm so that it can't get into the nucleus and activate gene expression, and Baldwin suggested that IκB's ankyrin repeats help it do that. "The thought was that since ankyrin repeats had been shown to interact with tubulin, then it might make sense that IκB is keeping NFκB in the cytoplasm by linking the complex to cytoplasmic structures via these ankyrin repeats," Baldwin explains.

The McKnight findings raise another possibility, however. If the ankyrin repeats make direct protein contacts the way McKnight has proposed, then, Baldwin and McKnight both suggest, IκB may bind NFκB through its ankyrin repeats, and in so doing obscure a signal on the transcription factor that would direct it to the nucleus.

Whether the inhibitor works that way is still unclear. Nor does anyone know yet just how versatile the system of ankyrin repeat proteins is. But the answer might be *very* versatile. Preliminary data from Haskill's lab suggest that many more, as yet unidentified, human proteins contain these repeats. All of which suggest that ankyrin repeats won't be overlooked again. ■ MICHELLE HOFFMAN