Zipping Up DNA Binding Proteins

By now it is clear that, during cell growth and development, certain proteins are able to bind with specific DNA sites, and thereby activate or inactivate gene expression. On page 1759 of this issue of *Science*, William Landschulz, Peter Johnson, and Steven McKnight from the Carnegie Institution of Washington's Department of Embryology in Baltimore describe a new type of protein structure, which they call the "leucine zipper," that may help to explain how some gene regulatory proteins work.

The Carnegie workers originated the idea of the leucine zipper while studying the structure of a protein that they had identified as a possible gene regulator because it binds specifically to two of the prominent control sites (the CAAT box and the enhancer) of the genes in higher organisms. They gave this protein the designation C/EBP for CAAT-box and enhancer binding protein.

Researchers had previously identified two structural motifs that are used by other gene regulatory proteins to recognize and bind to specific DNA sequences. These motifs are the "helix-turn-helix" structure, which is found primarily in bacterial proteins and consists of two α helices separated by a short nonhelical segment, and the "zinc finger," which is present in some of the gene regulatory proteins of higher organisms and consists of an outlooping of protein sequence held in place by a zinc ion.

McKnight and his colleagues did not find either of these motifs in C/EBP, but a computer search revealed that this protein contains a 35-base-pair sequence that is very similar to a sequence in another regulatory protein, namely the product of the *myc* oncogene.

One feature of the structural similarity was particularly noteworthy. "What we found remarkable," McKnight says, "was that each of these proteins had leucine at every seventh amino acid in the region of similarity." Moreover, three additional proteins with known or suspected gene regulatory activity proved to have the same feature, even though the amino acids sequences between the leucines are otherwise different from those in the leucine repeat regions of C/EBP and the *myc* protein.

The leucine repeat regions in the proteins have the capacity to form a regular α helix, with the leucine side chains projecting out from the helix at regular intervals. The question then is, McKnight says, "What are the leucines doing preserved in all these DNA binding proteins?" Landschulz, Johnson, and McKnight propose that the leucine side chains from one protein molecule interdigitate with those on a second protein, effectively forming a zipper that holds the two molecules together. The researchers further suggest that the dimer is the entity that interacts with DNA and regulates gene expression.

According to their model, the leucine repeats are not directly involved in making the specific DNA contacts, but serve to hold adjacent sections of the proteins in the correct configuration to do this.

At the time the Carnegie workers wrote the *Science* paper these suggestions were almost completely speculative. Since then, however, researchers have at least obtained evidence in support of the idea that the gene regulatory proteins are held together by leucine zippers. For example, Tom Curran of the Roche Institute for Molecular Biology in Nutley, New Jersey, Robert Tjian of the Howard Hughes Medical Institute at the University of California, Berkeley, Peter Vogt of the University of Southern California in Los Angeles, and their colleagues have demonstrated that the *fos* protein associates with the product of *jun* oncogenes, which also has a leucine repeat. Indications are that the complex participates in gene regulation.

In addition, Landschulz and Paul Sigler, a visiting scientist in McKnight's laboratory, have shown that C/EBP itself exists as a dimer. In both cases, the investigators have narrowed the regions through which the protein molecules interact with one another to the sequences containing the leucine repeats.

A role for the leucine zipper in establishing and maintaining DNA recognition sites on the regulatory proteins remains to be established, but one of the virtues of the model is that it allows two different protein molecules to associate, as in the case of the *fos* and *jun* products. The combination of different proteins might produce dimers with new specificities for recognizing gene control sites, McKnight points out, thereby expanding the range of genes that might be regulated by the proteins.

A second advantage of the model is that it can be readily tested by mutating the leucine repeats to see how this influences the interactions of the regulatory proteins and their effects on gene control. ■ JEAN L. MARX

Stress Proteins: Are Links in Disease

A recent study of the immune response to tuberculosis and leprosy in humans has led to some interesting insights into the nature of nonviral infections in general, according to a report by researchers at the Whitehead Institute for Biomedical Research, Cambridge, the University of Pittsburgh, and the Hammersmith Hospital, London. "We have observed an intriguing relationship between stress proteins and the immune response in mycobacterial infections," note Richard Young and his colleagues in their report in the *Proceedings of the National Academy of Sciences*. At first sight, this relationship offers promise in the production of potentially broad-spectrum vaccines.

At least as important, however, is the further implication of a link between stress proteins and nonviral infections: because stress proteins are an abundant and ubiquitous class of proteins, being produced in organisms as phylogenetically distant as humans and pathogenic organisms, there is a strong possibility that they are directly involved in the causation of certain autoimmune diseases. "There are already some data on this," Young told *Science*, citing work on an experimental model of rheumatoid arthritis in rats by Irun Cohen and his colleagues at the Weizmann Institute, Israel, and preliminary results of his own with human patients.

Young and his colleagues initially became involved in these broader issues through a concern for tackling tuberculosis and leprosy. "These diseases afflict 20 to 30 million people and continue to present a global health problem," they write. "To develop more effective tools for the diagnosis and prevention of these diseases, it is important to understand the immune response to infection by mycobacterial pathogens."

Humans infected with these pathogens mount an immune response to a dozen or more proteins produced by the invading organisms, cell-mediated immunity being important in the response. In parallel with work at other laboratories, the Whitehead project was aimed at identifying these proteins, with the view to being then in a position to target vaccines against the pathogens. Six mycobacterial proteins have been pinpointed as being key to infection, in both tuberculosis and leprosy. Genes for these proteins have been cloned and sequenced, results that revealed the connection with stress proteins.

It turns out that of this dozen genes, five have a very close sequence similarity to known stress proteins. "This and other recent reports suggest that infectious agents