More Progress on the T Cell Receptor

T cell receptor genes belong to the same superfamily as the genes for antibodies and major histocompatibility antigens

Immunologists have for many years been engaged in a frequently frustrating pursuit of the antigen receptor of T cells, which, together with B cells, are involved in mediating immune responses. Last year, several research groups took two big steps toward solving the problem of the receptor's identity: T cell surface proteins with all the predicted properties of the receptor were isolated; and a gene that appeared to code for a receptor protein was cloned. These studies have now been confirmed and extended. The new results reveal that the T cell receptor belongs to a "superfamily" of molecules that includes antibodies and the major histocompatibility antigens.

T cells play a central role in immune responses, both directly, by killing cells that appear foreign, and also indirectly, by regulating the activities of other immune cells, including the antibody-producing B cells. Of the three major types of surface molecules that are needed for immune reactions—the B cell receptor, the T cell receptor, and the major histocompatibility antigens—only the T cell receptor remained, until now, unknown. The B cells have been known for many years to use antibody molecules as the receptors by which they recognize foreign antigens. The histocompatibility antigens, cell surface molecules that help immune cells to recognize their partners or their targets, began to yield their secrets a few years ago. They were found to be structurally related to antibodies and presumably are derived from the same common ancestor.

But the lack of the T cell receptor left a big gap in the understanding of immune responses. Therefore the identification last year of T cell surface proteins with properties appropriate to the receptor greatly heartened immunologists. As was expected for the T cell receptor, these molecules are found only on the surfaces of T cells, not on B or other cell types. Their structures vary slightly from one T cell clone to another, which is expected for clones that recognize different antigens. And the activities of T cells are specifically blocked by antibodies against the individual proteins.

In due course, structural analysis of these proteins would have led to the cloning of the genes. DNA segments with sequences corresponding to the protein sequences could have been synthesized and used as probes for identifying the genes. In fact, the investigators who cloned the first genes used a different approach—identifying unique T cell messenger RNA's and making DNA copies [complementary DNA's (cDNA's)] of these messengers.

The first public report of the cloning of such a T cell–specific cDNA came from Mark Davis of Stanford University School of Medicine and Stephen Hedrick of the University of California at San Diego, who presented their results, which had been largely obtained while they were both still at the National Institute of Allergy and Infectious Diseases (NIAID), in August of last year at the World Immunology Congress in Kyoto, Japan (*Science*, 23 September 1983, p. 1278). Meanwhile, Tak Mak, Yusuke Yanagi, and their colleagues at the Ontario Cancer Institute in Toronto, were also cloning a cDNA, which was again specific to T cells, this time of human origin. The Davis and Hedrick clone was prepared from mouse cells.

The evidence that these cloned cDNA's encode T cell receptor proteins includes the demonstration that the corresponding genes had undergone rearrangement during the development of the T cells. This situation is analogous to that of antibody genes, which are assembled from three or four separate segments of DNA. Bringing the segments together requires the DNA to rearrange during B cell development.

In addition, the two groups have determined the nucleotide sequences of their cDNA clones and find that the proteins that should be encoded by those sequences bear a strong resemblance to antibody proteins.* The T cell receptor proteins are less closely related to the histocompatibility antigens.

The sequences of the two T cell receptor cDNA clones are very similar to one another. About 80 percent of the nucleotide building blocks in the genes have been conserved. The two clones appear to represent the mouse and human versions of the same gene.

Studies of the T cell receptor itself have revealed that it consists of two different protein chains, which have been designated α and β and range in molecular weight from about 40,000 to 48,000 (*Science*, 29 July 1983, p. 444). The proteins corresponding to the cDNA clones have a molecular weight of approximately 32,000 for the peptides only. This is consistent with those determined for the membrane proteins, which include a carbohydrate portion in addition to their peptide backbones.

> "There was a general feeling that the T cell receptor gene would be a simpler gene. That is not true. It has as many elements as the most complicated immunoglobulin."

In fact, recent evidence shows that the cloned genes code for the β-chain of the receptor. Ellis Reinherz and his colleagues Oreste Acuto and John Smart of the Dana-Farber Cancer Institute of Harvard Medical School have purified enough of the human β-chain to determine the sequence of the first 20 amino acids. The researchers are reasonably certain of the identities of 17 of these. Comparison of this sequence with that predicted by the Mak clone shows, Reinherz says, "that all the residues we are comfortable with are identical, one for one, with the predicted sequence."

The region of identity begins with amino acid 21 of the predicted sequence. The first 20 amino acids presumably form a "leader" sequence, which is needed for transport of the newly synthesized protein to its correct location on the cell membrane but is then removed.

*Y. Yanagi, Y. Yoshikai, K. Leggett, S. P. Clark, I. Aleksander, T. W. Mak, *Nature (London)* **308**, 145 (1984); S. M. Hedrick, D. I. Cohen, E. A. Nielsen, M. M. Davis, *ibid.*, p. 149; S. M. Hedrick, E. A. Nielsen, J. Kavaler, D. I. Cohen, M. M. Davis, *ibid.*, p. 153.

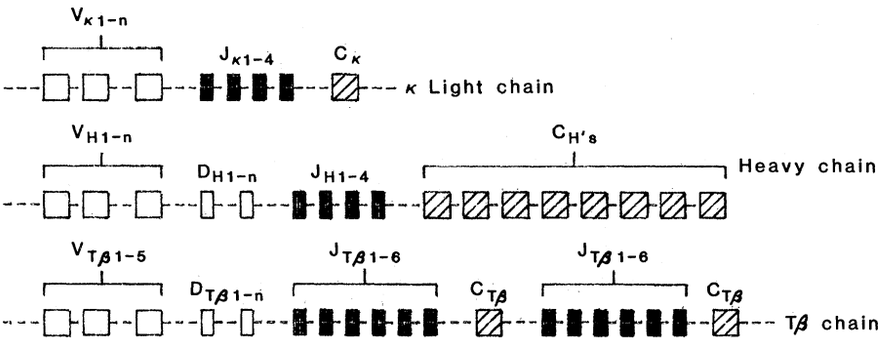

Diagram of unrearranged antibody and T cell receptor genes

The diagram shows, in somewhat simplified form, the genomic patterns of the coding sequences for antibody chains (the upper two lines) and the β-chain of the T cell receptor (the bottom line), before they undergo the DNA rearrangements needed for B and T cell maturation. These rearrangements bring together the V and J coding segments for antibody light chains and the V, D, and J segments for heavy chains and the T cell receptor protein. Attachment to the C region segments occurs at the level of the mRNA's for the proteins. Antibody light chains of the λ class (not shown here) have alternative C region coding segments with their own J regions, a pattern similar to that for the T cell receptor protein.

Because the human and mouse cDNA clones are so similar, it is likely that they both code for β-chains.

The cloned cDNA's can be used as probes for identifying and cloning the genomic DNA that carries the T cell receptor coding segments and for studying the number and arrangement of those segments. This has now been done by Davis and also by Leroy Hood's group at the California Institute of Technology, who used Mak's cDNA clone. They find that the genomic organization of the T cell receptor genes also resembles that of antibody genes.

The lighter of the two protein chains of which antibodies are composed are encoded by three separate DNA segments, which are designated V for variable, J for joining, and C for constant. Together the V and J segments form the part of the light chain that participates in antigen recognition. The antigen-binding region of the heavy chain is encoded in three separate DNA segments, including the D (for diversity) segment in addition to V and J. Heavy chains also have a constant region. The antigen-binding regions of antibody chains vary from one molecule to another whereas the constant region is the same for all chains of a given class.

The DNA rearrangements that occur during B cell maturation involve the joining of the V and J or of V, J, and D segments, a step that contributes to the antibody diversity that is needed to recognize an essentially unlimited number of antigens. Apparently the diversity of T cell receptor proteins is generated by applying mechanisms similar to those that have evolved for the antibody proteins.

Both the Davis and Hood groups find that in the mouse genome the unrearranged segments coding for the J and C regions of the T cell receptor protein are arranged as a cluster of six or so J segments followed by a C segment, and then there is another cluster of six or seven J segments and a second C segment.

In addition, both Hood and Davis have evidence for a D segment in T cell receptor genes. "There was a general feeling that the T cell receptor [gene] would be a simpler gene" Davis says. "That is not true. It has as many elements as the most complicated immunoglobulin."

That does not mean that the heavy chain and T cell receptor genes are the more advanced evolutionarily, however. Hood speculates that they may be the more primitive members of the superfamily and that the light chain evolved by loss of the D segment.

Another point of resemblance between immunoglobulin genes and those of T cell receptors, both groups find, is in the DNA sequences that serve as signals for joining the various segments when the DNA rearranges. "The signals are virtually indistinguishable," notes Hood.

Finally, both Davis and Hood and Mak have mapped the receptor gene to chromosome 6 of the mouse. This chromosome also contains the coding sequence for antibody light chains of the κ class, although this is probably a coincidence. The antibody and T cell receptor proteins are expected to be encoded separately. In the human genome, the T cell receptor gene is not on the chromosome that carries the κ chain genes.

There are some indications that β-chains of the T cell receptor may not vary much from one cell line to another, although not all the data on this point are in yet. For example, Hedrick has found that five lines of cloned T cells all use the same variable region in their β-chains, even though two of them have a slightly different specificity for their target antigens than the other three. The variation in the specificities may be caused by the use of different J and C regions by the two sets of T cells.

Although the investigators located only one V region coding segment with their original probes, Davis and Mak now say that there is evidence for at least five in both the human and mouse genomes. However, according to Davis, so far there is no evidence that somatic mutations, which contribute to the final diversity of antibody molecules, occur in the T cell receptor gene.

If the β-chain does not show a high degree of variability in its structure from cell to cell, then the ability to recognize a large number of antigens may reside in the structure of the α-chain. So far an α-chain gene has not been cloned. Or at least no one is ready to talk about progress in cloning it. But it is a safe bet that much effort is being directed toward this goal by the various laboratories.

Eventually it should be possible to study the function of the T cell receptor proteins by transferring the cloned genes into cells and seeing how this affects their specificity for antigen and other T cell activities. In the meantime, investigators have been using monoclonal and other antibodies to the receptor proteins as probes for function. For example, Hedrick and Davis, with Jonathan Rothbard of Stanford and Ronald Schwartz of NIAID, have shown that antibodies directed against the a synthetic peptide with a sequence corresponding to a J region of the β-chain can block activation of interleukin-2 production by some T cell clones that have been stimulated by appropriate antigens. This is an indication that the antibodies are blocking receptor function.

The close resemblance of the gene for the β-chain of the T cell receptor to antibody genes makes it likely that they evolved from the same ancestral gene. According to Mak, the protein encoded by the human receptor gene resembles the murine and human light chains about equally, a finding which suggests that the antibody and T cell receptor gene families diverged before the species diverged.

In any event, the three gene families—antibody, histocompatibility, and T cell receptor—appear to form a "superfamily," the members of which apparently evolved from a common ancestor to perform different functions that enable the immune system to distinguish foreign antigens from those of the body.

—JEAN L. MARX