More About the HTLV's and How They Act

Identification of the trans*-activating genes of all three HTLV's may lead to a better understanding of cellular growth control*

The human T-lymphotropic viruses (HTLV's) have been closely linked to two serious human diseases—adult T-cell leukemia and acquired immune deficiency syndrome (AIDS). That alone would be sufficient to make them the subjects of an intense research effort. But, like other viruses that infect mammalian cells, they are also interesting because of what they may reveal about the innermost secrets of those cells.

All three HTLV's infect T lymphocytes, which are needed for many immune responses. Infection by HTLV-I or -II, the viruses that have been associated with leukemias and lymphomas, leads to uncontrolled cell proliferation. In contrast, HTLV-III, the AIDS virus which is also called lymphadenopathy/AIDS virus (LAV) and AIDS-associated retrovirus (ARV), kills the cells it infects.

Finding out how the viruses produce these opposite effects could therefore give an improved understanding of cellular growth control. As William Haseltine of Harvard's Dana-Farber Cancer Institute has said, "These viruses may have put their fingers on the pulse of the replicative cycle of the T4 cell." The possibility that the information might also provide the means of interrupting the often fatal progression of the diseases caused by the viruses does not diminish interest in the research.

A recent profusion of reports, most of them appearing in *Science* over the past 3 or 4 weeks, has yielded new clues to the ways in which the HTLV's act. Earlier research had indicated that all three of the viruses produce *trans*-activating proteins that stimulate the expression of viral genes. The proteins might also alter cellular gene expression, thus disrupting growth control in infected cells.

It will now be easier to test this hypothesis. The viral genes coding for the *trans*-activating proteins have been definitively identified in all three viruses, thus confirming that HTLV-III, like HTLV-I and -II, does have such a gene. "It reinforces the notion that these retroviruses are unique in having this function," says Flossie Wong-Staal of the National Cancer Institute (NCI).

HTLV-III provided a surprise, however. Its *trans*-activating gene is in an unexpected location and is distinctly different in structure from those of HTLV-I and -II. The difference may help explain the diametrically opposed effects of HTLV-I and -II on the one hand and of HTLV-III on the other. It may also be new grist for the mills of those who already think that the AIDS virus is too unlike HTLV-I and -II to carry the same family name (*Science*, 22 March, p. 1449).

Nevertheless, despite the differences in the *trans*-activating genes, the new work shows that transcription of all three into the corresponding messenger RNA's (mRNA's), which is the first step in the synthesis of the proteins, follows the same unusual pattern. Finally, investigators are defining the control sequences on the target genes that respond to *trans*-activation. These are also unusual and may be representatives of a new class of genetic regulatory elements.

> "At a minimum the viruses have all evolved a *trans*-activating mechanism for speeding up their own replication."

When the original evidence for *trans*-activation by HTLV-I and -II was obtained about a year ago, the prediction was that the protein causing this activity would be encoded in a region of the viral genomes that had been designated "pX" because its function was unknown. The right-hand half of the pX region, which is sometimes designated the *lor* (for long open reading frame) region, codes for a protein with a molecular weight of about 40,000 and is very similar in the two viruses, a finding which may mean that it has an essential function.

Two groups, using different approaches, have now confirmed that the *x-lor* region codes for the *trans*-activating proteins of HTLV-I and -II. Haseltine, Joseph Sodroski, and their Dana-Farber colleagues have shown that the *x-lor* region of either virus is all that is needed to stimulate the expression of genes under the control of the viral regulatory sequences (*1*).

In addition, in a report in this issue of *Science* (p. 54) Irvin Chen, Dennis Slamon, and their colleagues at the University of California School of Medicine in Los Angeles describe HTLV-II mutants in which the *x-lor* region can no longer make its protein product. Transcription of the viral genes is greatly reduced in the mutants, which fail to replicate. "This is direct evidence that the *x-lor* gene is necessary for replication because it is necessary for transcriptional activation," Chen says. Both the Haseltine (*2*) and the Chen (*3*) groups have found that the *x-lor* proteins are largely located in the cell nucleus, the appropriate site for proteins that alter gene transcription.

Chen and his colleagues could not detect the *trans*-activating protein in the HTLV-I particle. Apparently the virus does not carry it into newly infected cells but must make it after arrival there. The results with the *x-lor* mutant suggest that a low level of transcription occurs even in the absence of the *trans*-activating protein. If the same is true for the normal, wild-type virus, then it may initially produce a small amount of the protein, thus stimulating more transcription and producing a positive feedback that allows efficient virus replication.

In any event, now that the function of the *x-lor* region has been established, Haseltine proposes that it be given the designation *tat*—for *trans*-activating transcriptional gene—with the subscripts I or II to indicate the specific virus.

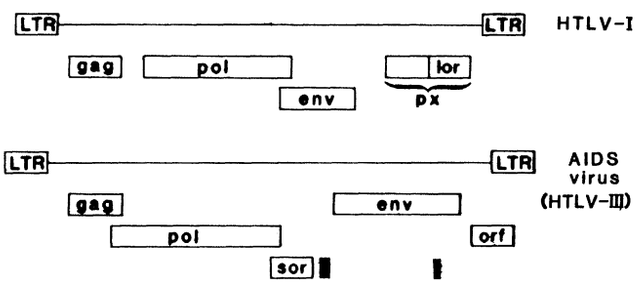
If the HTLV-III genome followed the same organizational pattern as HTLV-I and -II, then its *trans*-activating protein ought to have been encoded in the open reading frame that is located to the right of the viral envelope (*env*) gene and extends into the right-hand long terminal repeat (LTR). That is not what has been found, however.

In this issue, Wong-Staal and her NCI colleagues (p. 69) and Sodroski, Haseltine, and their colleagues (p. 74) report that they have identified the HTLV-III *tat* gene. It consists of three exons, the first of which does not code for protein structure. The second and third exons encode a total of 86 amino acids, which makes the protein about one-fourth the size of the HTLV-I and -II *trans*-activating proteins. The second exon, which codes for 72 amino acids, is located approximately in the middle of the viral

genome between a short open reading frame, designated *sor*, and the start of the *env* gene. That region had not previously been thought to have protein-coding capabilities. The remaining 14 amino acids are encoded by a small exon located within the *env* gene. Only the second exon is absolutely required for *trans*-activation.

The sequence of the HTLV-III protein does not resemble those of any other protein, including the HTLV-I and -II *tat* proteins. It is rich in basic amino acids, a finding that is consistent with its proposed role in altering gene expression.

Investigators who favor the inclusion of the AIDS virus in the HTLV family and those who oppose it can both find support in these results. The gene encoding the AIDS virus *trans*-activating protein is clearly different from those of HTLV-I and -II. Nevertheless, the identification of such a gene in the AIDS virus buttresses the arguments of those who think that it is truly an HTLV. "At a minimum the viruses have all evolved a *trans*-activating mechanism for speeding up their own replication," Haseltine explains. "It is a new property these viruses have and they have new genes to go along with it."



Organization of the HTLV-I and -III genomes. The black bars indicate the regions coding for the trans-*activating protein of the AIDS virus.*

Among the RNA-containing viruses, only the three HTLV's and the related bovine leukemia virus have definitively been shown to have *trans*-activating capabilities. Although there has been a report of *trans*-activation by Rous sarcoma virus, other investigators, including Haseltine, have had trouble confirming the result and this situation awaits further clarification. The HTLV's and BLV may in fact more closely resemble certain DNA-containing viruses in their mode of action than other retroviruses. For example, the DNA viruses SV40 and adenovirus, both of which can transform appropriate cell types, make *trans*-activating proteins.

Synthesis of the *trans*-activating gene messengers provides another point of similarity between HTLV-III and the other two HTLV's. The Haseltine and Chen groups (*4*) and also that of Mitsuaki Yoshida at the Tokyo Cancer Institute (*5*) have found that the *tat* genes of HTLV-I and -II consist of three exons separated by two introns, an organization similar to that of the HTLV-III gene, despite its other differences. Formation of the corresponding mRNA's thus requires two splicing steps to remove the introns and join the exons. This is a very unusual mode of synthesis for a retroviral mRNA. Two exons and a single splicing event is more typical.

In addition, the first four nucleotides coding for the *trans*-activating proteins of HTLV-I and -II are located not in *x-lor* region, which forms the third exon, but come from the start of the *env* gene, which is some distance away in the second exon. "There is a completely novel pattern for generating the *x-lor* mRNA," Chen points out. The unusual splicing pattern may provide a means by which the HTLV's can control the synthesis of their *trans*-activating proteins.

The supposition is that the *trans*-activating proteins of HTLV-I and -II are the transforming proteins of the viruses. In addition to stimulating transcription of viral genes, they may perhaps increase the transcription of cellular genes that stimulate cell division, thus producing the uncontrolled growth characteristic of transformed cells. Although the potential cellular targets of the HTLV-I and -II *trans*-activating proteins are still unknown, the gene for the interleukin-2 receptor remains a good candidate. Transformed cells have greatly increased numbers of this receptor, which mediates the growth stimulating effects of the lymphokine interleukin-2.

However, the *trans*-activating proteins by themselves may not be sufficient to generate a cancerous tumor, Chen points out. Although cultured cells that have been transformed by HTLV-I or -II make the corresponding mRNA's, tumor cells that have been isolated from leukemic patients do not. The *trans*-activating proteins may be needed early in oncogenesis to give the cells the ability to divide continuously. This would be consistent with the role proposed for the *trans*-activating proteins of such DNA viruses as SV40 and adenovirus. Other, additional steps may then be required to complete the cancerous transformation.

In contrast to the role proposed for the HTLV-I and -II *trans*-activating proteins, the *trans*-activating protein of HTLV-III may turn off a gene needed for cell division or stimulate one that suppresses cell division. "Now that we have the proteins isolated we can test this hypothesis," Haseltine says.

A complete understanding of *trans*-activation requires identification of the regulatory sequences on the target genes that respond to the *tat* proteins. Haseltine and his colleagues have identified these sequences for the HTLV-I genome and have found that they, like so many other HTLV features, are unusual (*6*).

Most genes have at least two regulatory elements, a promoter that helps to determine accurately the start site for transcription and an enhancer that governs the rate of transcription from the promoter. In retroviruses these regulatory regions are in the LTR's at the ends of the genome.

The Haseltine group has found that the regulatory sequence for HTLV-I *trans*-activation, which they call the *tar* (for *trans*-acting responsive) sequence, lies between nucleotide −159 and nucleotide +315 (counting from the transcription start site). It includes part of the enhancer, which extends from nucleotide −350 to nucleotide −55, plus the promoter, which is located about 35 nucleotides before the start site. But neither the enhancer nor the promoter is by itself sufficient for *trans*-activation.

Also unusual, Haseltine stresses, is the efficacy with which this regulatory region seems to work. *Trans*-activation by the HTLV's gives greater increases in transcription than other systems, including the SV40 system, usually do. "We think that it may be a new kind of regulatory sequence that will be found in other genes that need to be turned on in a hurry," Haseltine says. He suggests that it may be similar to another regulatory sequence that was recently found by Nam-Hai Chua of Rockefeller University and his colleagues in a totally unrelated gene coding for a plant photosynthetic protein. All in all, the HTLV's are proving to be estimable guides to the molecular biology of the mammalian cell.

—JEAN L. MARX

References

1. J. Sodroski, C. Rosen, W. C. Goh, W. Haseltine, *Science* **228**, 1430 (1985).
2. W. C. Goh *et al.*, *ibid.* **227**, 1227 (1985).
3. D. J. Slamon *et al.*, *ibid.* **228**, 1427 (1985).
4. W. Wachsman *et al.*, *ibid.*, p. 1534 (1985).
5. M. Seiki, A. Hikikoshi, T. Taniguchi, M. Yoshida, *ibid.*, p. 1532 (1985).
6. C. A. Rosen, J. G. Sodroski, W. A. Haseltine, *Proc. Natl. Acad. Sci. U.S.A.*, in press.