Proc. Natl. Acad. Sci. USA
Vol. 84, pp. 2693–2697, May 1987
Biochemistry

Relocation of a protease-like gene segment between two retroviruses

(gene transfer/retrovirus genes/homology/phylogenetic trees)

M. A. McCLURE, M. S. JOHNSON, AND R. F. DOOLITTLE

Department of Chemistry, University of California, San Diego, La Jolla, CA 92093

Contributed by R. F. Doolittle, January 2, 1987

ABSTRACT **An anomalous sequence in certain lentiviruses was found to be related to a region in a completely different part of the simian retrovirus type I (SRV-I) and its close relative, the hamster intracisternal A particle (IAP-H18). The segment is not present in the human immunodeficiency virus (HIV), which is also a lentivirus, nor is it found in any one of a dozen other retroviruses whose sequences have been reported. These observations imply that a horizontal transfer of newly acquired genetic information has taken place between an SRV-I-type virus and one of the lentivirus type, and that this event occurred more recently than did the divergence of members of this latter group and HIV. Comparison of the viral nucleic acid sequences that encode these segments revealed the presence of imperfect direct nucleotide repeats resembling the retroviral endonuclease cleavage sites at the 5′ and 3′ ends of these regions.**

During the course of a large-scale computer comparison of retroviral protein sequences, an unexpected sequence similarity was encountered between two rather distantly related groups of viruses. A segment located within the polymerase region of two lentiviruses—visna (1) and equine infectious anemia virus (EIAV) (2)—was not present at this position in any of a dozen other retroviruses and retrotransposons examined. The putative protein sequence was searched against a large protein sequence data base, and, surprisingly, the only sequence retrieved corresponded to another anomalous region adjacent to the protease of the simian retrovirus type I SRV-I (3). It was also present in the closely related hamster intracisternal particle IAP-H18 (4). The SRV-I segment had been noticed by Power *et al.* (3), who suggested that it was the result of a tandem duplication of the protease gene. Inspection and alignment of these unassigned segments, which amount to about 130 amino acids each, leaves no doubt that the two sets are homologous. The generally accepted relationship of the viruses involved is such, however, that the relative locations of the segments are most readily explained by a horizontal gene transfer. Apparently a nucleic acid segment from one virus was excised and integrated into a different location in another quite distantly related one. The intrigue of these anomalous occurrences led us to inquire about the origin of the segment itself, as well as to ponder a mechanism of translocation and its structure–function consequences.

We began by aligning and comparing the two sets of anomalous segments themselves to find exactly how similar they are. That sequences were available from two different viruses in each case allowed us to measure how fast the segments were evolving relative to other retrovirus proteins. We also compared and aligned the proteases from the same viruses and then aligned these with the anomalous segments. As it happens, the question of whether the original extra segment was in fact the result of a tandem duplication is central to the whole issue.

We also aligned a number of other sequences including the various polymerase proteins—reverse transcriptase, ribonuclease H, and endonuclease sequences—from an assortment of retroviruses. These were used both for the construction of phylogenetic trees and for the delineation of the exact boundaries of the anomalous segments. Finally, we examined the nucleic acid sequences, as opposed to the amino acid sequences, of these regions in an effort to find clues to the mechanism of relocation. Indeed, the junctions of the segments involved contain direct sequence repeats that are similar to those found at the terminals of the retroviruses themselves.

In this article we provide a quantitative analysis of the relationship among the various protease-like segments and the authentic viral proteases and construct phylogenetic trees based on other more highly conserved viral gene products of representative retroviruses. The analysis points to horizontal transfer of genetic information. We also describe a simple model that explains how the relocation of the gene segment may have taken place.

METHODS

The computer used in this study was a DEC 11/730 VAX computer running the UNIX (Berkeley 4.3) operating system. The search program used a moving window of 40 residues and a table of weighted values taken from the Mutation Matrix of Dayhoff (5). Binary amino acid sequence alignments were determined by previously described procedures (6). Alignments and distance values for the phylogenetic trees were determined as described in Feng and Doolittle (7); nucleic acid comparisons were performed with the SEQH program of Goad and Kanehisa (8).

The protein sequence data bases utilized were the 1986 version of NEWAT (9) and release 10.0 of the National Biomedical Research Foundation Atlas (10). Nucleic acid sequences were taken from the 1986 version of GenBank* (Los Alamos National Laboratory). Additional retrovirus sequences utilized in this study are: the human T-cell leukemia viruses, HTLV-I (11) and HTLV-II (12), bovine leukemia virus (BLV) (13), Rous sarcoma virus (RSV) (14), human immunodeficiency virus (HIV) (15), and Moloney murine leukemia virus (Mo-MLV) (16).

RESULTS

Designation of Gene Arrangement. A diagrammatic depiction of the gene arrangement in these retroviruses is presented in

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: EIAV, equine infectious anemia virus; HTLV-I and -II, human T-cell leukemia viruses I and II; BLV, bovine leukemia virus; RSV, Rous sarcoma virus; HIV, human immunodeficiency virus; Mo-MLV, Moloney murine leukemia virus; SRV-I, simian retrovirus type I.

*National Institutes of Health (1986) Genetic Sequence Databank: GenBank (Research Systems Div., Bolt, Beranek, and Newman, Cambridge, MA), Tape Release 46.0.

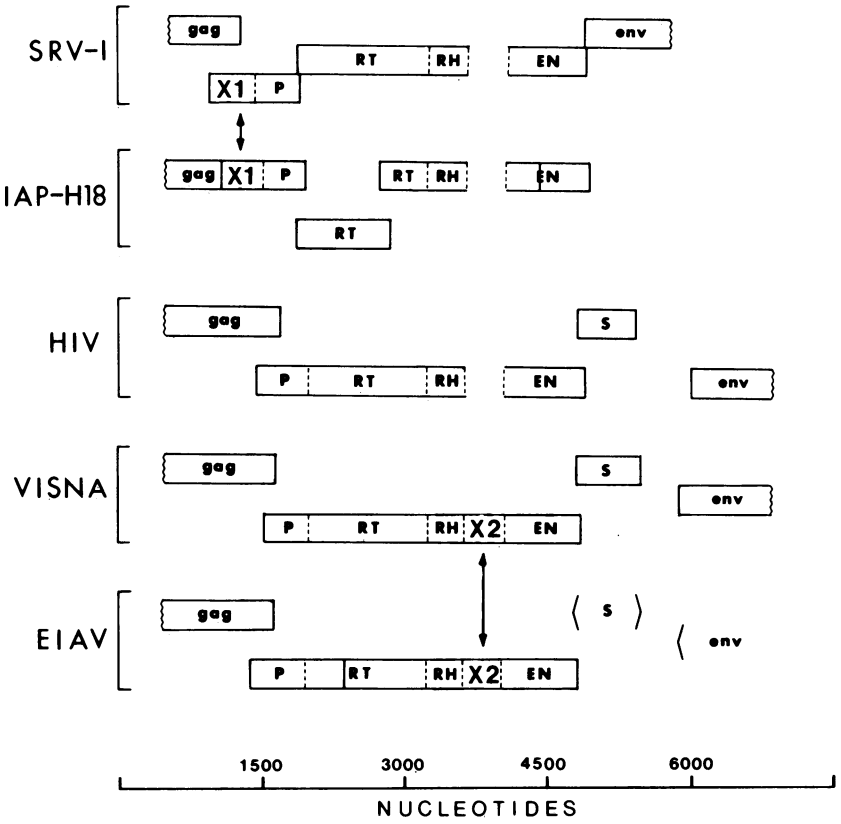

FIG. 1. Diagrammatic representation of the gene order derived from the three reading frames of SRV-I, IAP-H18, HIV, visna virus (VISNA), and EIAV. The protease-like sequences (X1) in SRV-I and IAP-H18 are adjacent to the protease gene (P), while the equivalent sequences (X2) are situated between the ribonuclease H (RH) and endonuclease (EN) coding regions (17) of the polymerase in both visna virus and EIAV. The relocated segment is not present in HIV. The boundaries of the *gag* (gag), reverse transcriptase (RT), short open reading frame (S), and *env* (env) coding regions are shown for reference purposes. Although the entire EIAV sequence is not yet available, probable positions of an *env* and short open reading frame are indicated between the angle brackets. The coding regions are based on the frame of the *gag* genes, all of which have been arbitrarily assigned to the first reading frame. Solid vertical lines indicate translation, initiation, and termination signals; dashed lines indicate gene-product functional divisions. SRV-I, IAP-H18, and HIV are interrupted between the RNase H and endonuclease to emphasize the absence of the protease-like segment within these *pol* regions.

Fig. 1. The diagram is based on numerous published accounts providing evidence for the reading frames encoding various gene products, including the *gag* region, protease, polymerase, and envelope protein genes. We have divided the polymerase gene into three segments, including the experimentally identified endonuclease region, on the one hand, and the reverse transcriptase and ribonuclease H segments, on the other, the boundaries of which we recently proposed on the basis of sequence similarity to the *Escherichia coli* ribonuclease H (17). In the following discussion, for purposes of keeping track, the protease-like sequences from the simian virus and hamster particle are denoted "X1" and the equivalent but translocated segments in the two lentiviruses are denoted "X2" (Fig. 1). The authentic protease sequence segments previously identified in retroviruses (18–20) are denoted "P."

Sequence Alignments. Three different sets of amino acid alignment data were examined as an aid to understanding the proposed relocation event. First, the translocated segments in the lentiviruses (X2) were examined and found to share common ancestry with the simian virus and hamster particle protease-like segments (X1) (Fig. 2). The sequence resemblances are of the order of 30% (Table 1). When the data were subjected to a conventional jumbling test (6), the authentic alignment scores were about 10 standard deviations above the mean score of the randomized comparisons (Table 1).

Second, the viral proteases were compared among themselves to determine how similar they are to one another and to ascertain whether or not the protease-like regions are changing faster than the proteases themselves, as might be expected if there were a relaxed selection after loss of function. As it happens, the segments appear to be changing at about the same rate as the authentic proteases, in both groups the "nearest-relative-similarities" being of the order of 35–40% (Table 1).

Finally, the degree of similarity between the authentic proteases (P) and the duplicated regions (X1 or X2) was considered. This was necessary both to establish that the new segments did in fact evolve by tandem duplication of a protease gene and to determine when that event might have occurred. In this regard, the extensive sequence divergence that has taken place among the protease and protease-like segments makes it impossible to obtain statistically significant scores by pairwise alignment. On the other hand, a multiple alignment of all the segments revealed that conserved remnants of their common heritage are still much in

```
SRV-I X1        SLWGGQLCSSQQKQPISKLTRAT PGSAGLDLSSTSHTVLTPEMGPQAL
IAP-H18 X1           RNMGTNFRKAIKRKRFPRNLRNGLACRS  DFLLMPQMNVQPV
                       **  *   **        ** **        *   *
VISNA X2         SEIFLAKEGR  GILQKR     AEDAGYDLICP QEISIPAGQVKRI
EIAV X2          EEIMLAYQGT  QIKEKR     DEDAGFDLCVP YDIMIPVSDTKII

SRV-I X1        STGIYGPLPPNTFGLILGRSSITIKGLQVYPGVIDNDYTGEIKIMAK  A
IAP-H18 X1      PVHSPGPLPPATIGLILGRGSLTLQGLIIYPGIVDPYHKEEIQVLCSSPR
                **     **** ** * * **   *** *  ****  ******* *
VISNA X2        AIDLKINLKKDQWAMIGTKSSFANKGVFVQGGIIDSGYQGTIQVVIY  N
EIAV X2         PTDVKIQVPPNSFGWVTGKSSMAKQGLLINGGIIDEGYTGEIQVICTNIG

SRV-I X1        VNN IVTVPQGNRIAQLILLPLIET    DNKVQQPYRGQGSFGSSDIYW
IAP-H18 X1      G   VFSIKQGDRIAQLVLPPSLGD    GETYTLQKRAMGSSGSDSAYL
                 **  * ****   *********            **   ***   **
VISNA X2        SNNKEVVIPQGRKFAQLILMPLIHEELEPWGETRKTERGEQGFGSTGMYW
EIAV X2         KSNIKLI  EGQKFAQLIILQHHSNSRQPWDENKISQRGDKGFGSTGVFW
```

FIG. 2. Alignment of the duplicated region (X1) from SRV-I (residues 22–164) and IAP-H18 (residues 1–120) with relocated segments (X2) from visna virus (VISNA; residues 687–825) and EIAV (residues 331–468). Residue positions start at the open reading frame coding for each region, except for IAP-H18, where the first 14 amino-terminal residues are derived from the end of the next open reading frame. Asterisks denote positions in the alignment where at least one residue is in common between the SRV-I/IAP-H18 group and the VISNA/EIAV group.

Table 1. Resemblance (% identity) and significance (in SD) between various pairs of protease-like sequences (X1, X2) and retroviral proteases (P)

	Protease-like sequence					Retroviral protease			
	SRV-I (X1)	IAP-H18 (X1)	VISNA (X2)	EIAV (X2)		SRV-I (P)	IAP-H18 (P)	VISNA (P)	EIAV (P)
SRV-I (X1)	—	34.6%	28.0%	28.3%	SRV-I (P)	—	38.9%	26.0%	20.5%
IAP-H18 (X1)	14.2 SD	—	24.1%	22.6%	IAP-H18 (P)	17.2 SD	—	19.2%	23.0%
VISNA (X2)	8.8 SD	5.7 SD	—	42.8%	VISNA (P)	4.5 SD	2.0 SD	—	34.2%
EIAV (X2)	11.7 SD	10.4 SD	22.1 SD	—	EIAV (P)	3.8 SD	6.3 SD	9.9 SD	—

The standard deviations are a measure of how much greater the alignment scores of genuine alignments are in comparison with scores obtained from alignment of random sequences of the same lengths and compositions. Scores greater than 3 SD are generally considered as significant.

evidence (Fig. 3). It should be noted that for all pairwise alignment combinations, the protease-like segments are more closely related among themselves than they are to any of the authentic viral proteases.

Phylogenetic Relationships. The horizontal nature of the transfer of the protease-like segment from a simian virus ancestor to a visna virus predecessor is graphically depicted on a sequence-based phylogenetic tree (Fig. 4*B*). Two trees were constructed, one based on alignment of the reverse transcriptase sequences and another based on the endonuclease sequences. These two gene products are among the most conserved in retroviruses in general. The tree generated from the reverse transcriptase data is similar to the one published by Gonda *et al.* (21) in that the lentivirus branch order is: visna, HIV, and EIAV (Fig. 4*A*). In contrast, the endonuclease-based tree produces a different branch order for the lentiviruses (Fig. 4*B*). The differences in the two trees reflect the limits of accuracy inherent in current tree-building schemes when only a single gene product is used. Indeed, ancillary data are often needed to resolve such discrepancies. In this instance, the occurrence of the protease-like segment in visna and EIAV, but not in HIV, indicates that the endonuclease-based tree more accurately reflects the relationship for these three lentiviruses.

Nucleic Acid Sequence Analysis. The relocation of a gene segment in one group of retroviruses to a distinctly different region in another group (Fig. 1) led us to compare the nucleic acid sequences directly—as opposed to the inferred amino acid sequences—in search of mechanistic clues. To this end it was necessary to define precisely the amino and carboxyl termini of the translated segments. Inspection of the ribonuclease H/endonuclease junctions from an assortment of retroviruses, on the one hand, and a comprehensive multiple alignment of their proteases, on the other, pinpointed the termini of the unassigned segments. Moreover, the amino terminus of the endonuclease of HIV (22) has been determined experimentally and is consistent with the expected location of the carboxyl termini of the relocated segments.

Comparison of the DNA sequences defined by the predicted amino and carboxyl termini and 50 nucleotides of additional flanking sequences to insure completeness revealed an intriguing set of similarities in the terminal regions (Fig. 5). The boundaries of the duplicated segments (X1) and relocated segments (X2) were identified on the basis of the reported terminal sequences of adjacent genes. In this regard, the upstream sequence of visna and EIAV encodes the ribonuclease H (RH in Fig. 1) and the downstream region encodes the endonuclease (EN in Fig. 1). In SRV-I, the

FIG. 3. Multiple alignment of the duplicated segment (X1) and translocated segment (X2) and the previously identified proteases. Identical residues between the segments (X1 and/or X2) and proteases (P) are circled. The lined residue clusters indicate the remnants common among the retroviral proteases (19, 20). The sequences of the protease-like segments (X1 and X2) included are the same as those shown in Fig. 2; the protease sequences are numbered as follows: SRV-I, residues 164–300; IAP-H18, residues 121–255; visna virus (VISNA), residues 43–164; and EIAV, residues 85–207.

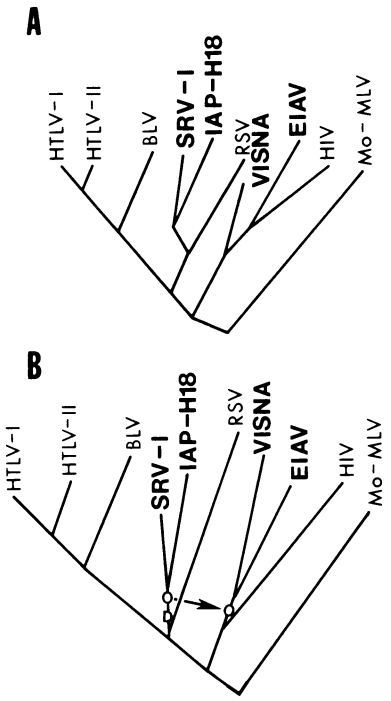

FIG. 4. Phylogenetic trees depicting the evolutionary relationship of SRV-I, IAP-H18, visna virus (VISNA), and EIAV (bold print) and representative retroviruses (HTLV-I, HTLV-II, BLV, RSV, HIV, and Mo-MLV). Trees were constructed from distance values after alignment of the reverse transcriptase sequences (*A*) and endonuclease sequences (*B*). On the endonuclease-based tree, "D" marks the occasion of the internal duplication leading to the protease-like segment, and the arrow denotes its horizontal transfer from an ancestor of SRV-I and IAP-H18 to a predecessor of visna virus and EIAV.

upstream region is the beginning of the open reading frame that codes for both the putative duplication and the authentic protease; in IAP-H18 this region is the 3′ end of the preceding open frame. In both of the latter situations, the downstream regions code for a protease (P) (Fig. 5).

An imperfect direct repeat sequence of 11–14 nucleotides is located at the ends of both the duplicated and relocated gene segments (Fig. 5). It is defined by the trinucleotides CAG at the 5′ end and TGG at the 3′ end. The consensus sequence of these direct repeats is CAG(N_n)TAT(N_n)TGG (n = 0–5 nucleotides; N = an unspecified nucleotide); in all four cases examined, it falls at or within one residue of the predicted amino acid alignment data.

DISCUSSION

We have found that a gene segment located between the ribonuclease H and the endonuclease genes of visna and its close relative EIAV (Fig. 1) is related to a different region in SRV-I type viruses, itself thought to be the result of a tandem duplication of its protease gene (3) (Table 1 and Fig. 2). The notion that the segment originated as a tandem duplication is supported by its location and equivalent size, as well as by a multiple alignment of all the sequences involved (Fig. 4) that reveals remnants of the most conserved residue clusters found in retroviral proteases (19, 20).

The functional significance of these protease-like segments is unclear. In either case the coding sequences are "in frame" with the adjacent genes and are not set off by terminator codons, indicating that the corresponding amino acids are not only expressed but are made as distinctly different parts of the retroviral polyprotein. Whether or not subsequent processing releases them as independent proteins remains to be determined. The fact that the protease-like segments are changing at about the same rate as the authentic proteases implies a genuine function independent of where the segment is located in the polyprotein.

The overall phylogenetic relationship of these retroviruses indicates that the presence of the protease-like region in the lentiviruses is the result of horizontal transfer of the corresponding gene segment (Fig. 4*B*). Although a search of the protein sequence data base did not reveal any additional relationships, the duplicated segment is also present in the recently published sequence of human endogenous retrovirus (23), another close relative of SRV-I. The segment will no doubt be found in other close relatives of both SRV-I and visna-type viruses. Significantly, the relocated segment is not found in HIV (Fig. 1), indicating that the relocation event must have occurred after the divergence of HIV from the other lentiviruses (Fig. 4*B*). The alternative possibility in which HIV lost the segment seems unlikely.

How the protease-like segment was horizontally transferred from its site of duplication to a nonhomologous region in another virus group is open to speculation. Classical homologous recombination cannot account for the translocation because the sequences surrounding the extra segments are not at all similar (Fig. 1). Although nonhomologous recombination could conceivably have taken place, the presence of 5′ and 3′ direct repeats at the termini of the translocated segments suggests an alternative explanation: to wit, a mechanism similar to one proposed for the conservative transposition observed in some bacterial systems (24).

The important point here is that the "direct repeats" are similar to the endonuclease cleavage sites (25) of the circular long-terminal-repeat junctions of retroviruses in that the invariant dinucleotides CA and TG are preserved on either side of the protease-like regions (Fig. 5). The variation in the nucleotide number and composition between these dinucleotides is well within that generally found in the long-terminal-repeat junctions of retroviruses and retrotransposons. Given that an endonuclease was responsible for "freeing" the gene for the protease-like segment, how were the ends of the direct repeats regenerated at the new location? The mechanism

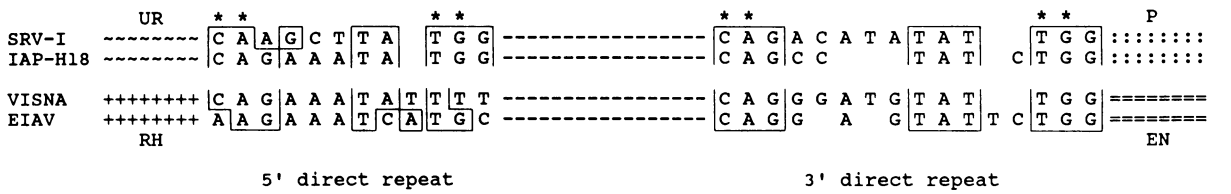

FIG. 5. Schematic representation of the nucleotide sequences of the protease-like segments (X1 and X2). The boxed regions indicate the nucleotides conserved at each end of the gene segment. Asterisks denote the positions of the dinucleotides CA and TG, common to the circular long-terminal-repeat junctions of all retroviruses (25). In the case of visna virus (VISNA) and EIAV, the nucleotides upstream are denoted RH for ribonuclease H (indicated by "+"), while those downstream are denoted EN for endonuclease (indicated by "="). The SRV-I and IAP-H18 upstream regions are denoted UR for upstream region (indicated by "~"), and the downstream regions are denoted P for active protease (indicated by ":"). Nucleotide positions are numbered beginning with the 5′ end of the genomes and indicate the first and last position of the protease-like gene segment: SRV-I, bases 2345–2784; IAP-H18, bases 2471–2777; VISNA, bases 3753–4165; and EIAV, bases 3573–3986.

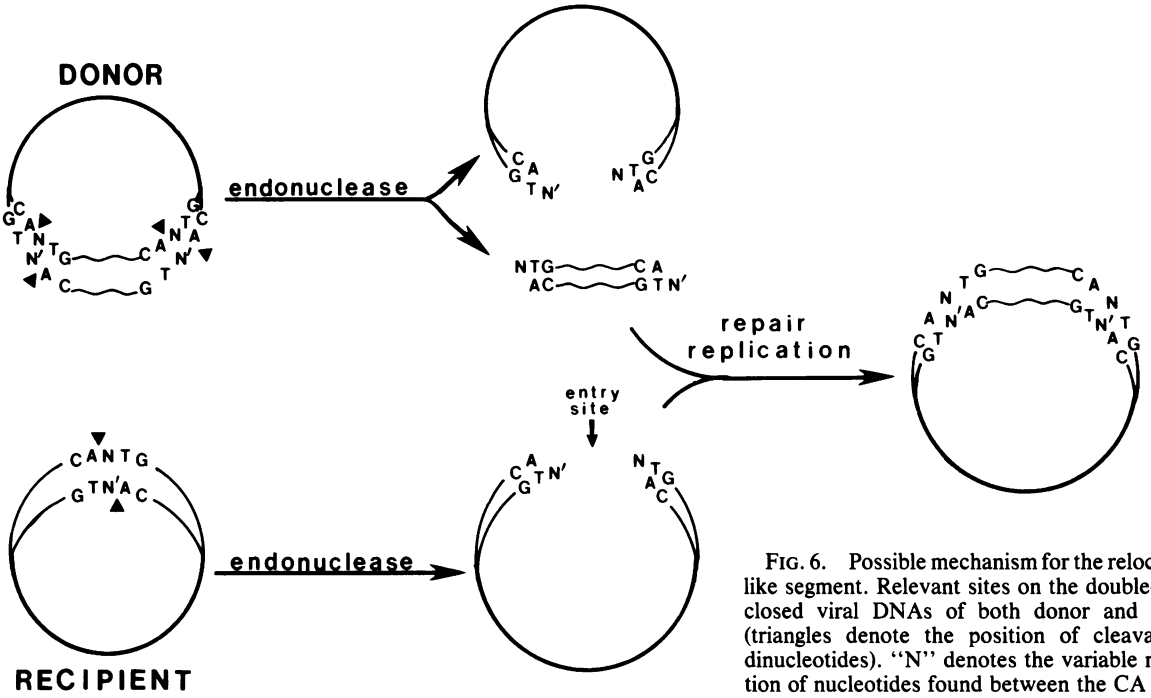

FIG. 6. Possible mechanism for the relocation of the protease-like segment. Relevant sites on the double-stranded, covalently closed viral DNAs of both donor and recipient are shown (triangles denote the position of cleavage at the invariant dinucleotides). "N" denotes the variable number and composition of nucleotides found between the CA and TG.

illustrated in Fig. 6 readily accounts for the presence of these conserved ends. As a result of a tandem duplication, a donor molecule has two sequences that are recognizable by an endonuclease. A recipient has only a single cleavage-recognition sequence that can provide an entry site for the released gene segment. Interestingly, examination of several ribonuclease H/endonuclease junction regions in retroviruses that do not have the relocated segment revealed that HIV has the sequence CAGTGCTG near the 3′ end of its ribonuclease H gene. As such, it is a potential recipient for this type of gene transfer.

Thus, a tandem duplication event generated a gene segment bounded by endonuclease cleavage sites, thereby allowing the release of the intervening segment. The chance occurrence of an endonuclease cleavage site near the junction of two genes on a different viral genome resulted in the capture of the nonhomologous gene segment. Obviously, coinfection of a host organism by the two different viral genomes ancestral to this event must have occurred for the transfer to have taken place. The importance and frequency of retroviral gene segment relocation, whether or not mediated by the "illegitimate transposition" mechanism suggested here, remains to be determined. It is clear, however, that there has been a relocation of gene segments between different viruses, and the newly arranged genomes have survived natural selection.

We thank Da Fei Feng for helpful discussion and programming assistance during the course of this study. This work was supported by a grant from the American Cancer Society and National Institutes of Health Grant GM-34434.

1. Sonigo, P., Alizon, M., Staskus, K., Klatzmann, D., Cole, S., Danos, O., Retzel, E., Tiollais, P., Haase, A. & Wain-Hobson, S. (1985) *Cell* **42,** 369–382.
2. Stephens, R. M., Casey, J. W. & Rice, N. R. (1986) *Science* **231,** 589–594.
3. Power, M. D., Marx, P. A., Bryant, M. L., Gardner, M. B., Barr, P. J. & Luciw, P. A. (1986) *Science* **231,** 1567–1572.
4. Ono, M., Toh, H., Miyata, T. & Awaya, T. (1985) *J. Virol.* **55,** 387–394.
5. Dayhoff, M. O., Schwartz, R. M. & Orcutt, B. C. (1978) *Atlas of Protein Sequence and Structure*, ed. Dayhoff, M. O. (Natl. Biomed. Res. Found., Washington, DC), Vol. 5, Suppl. 3, pp. 345–358.
6. Feng, D.-F., Johnson, M. S. & Doolittle, R. F. (1985) *J. Mol. Evol.* **21,** 112–125.
7. Feng, D.-F. & Doolittle, R. F. (1986) *J. Mol. Evol.*, in press.
8. Goad, W. B. & Kanehisa, M. I. (1982) *Nucleic Acids Res.* **10,** 247–263.
9. Doolittle, R. F. (1981) *Science* **214,** 149–159.
10. George, D. G., Barker, W. C. & Hunt, L. T. (1986) *Nucleic Acids Res.* **14,** 11–15.
11. Seiki, M., Hattori, S., Hirayama, Y. & Yoshida, M. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 3618–3622.
12. Shimotohno, K., Takahashi, Y., Shimizu, N., Gojobori, T., Golde, D. W., Chen, I. S. Y., Miwa, M. & Sugimura, T. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 3101–3105.
13. Rice, N. R., Stephens, R. M., Burny, A. & Gilden, R. V. (1985) *Virology* **142,** 357–377.
14. Schwartz, D. E., Tizard, R. & Gilbert, W. (1983) *Cell* **32,** 853–869.
15. Ratner, L., Haseltine, W., Patarca, R., Livak, K. J., Starcich, B., Josephs, S. F., Doran, E. R., Rafalski, J. A., Whitehorn, E. A., Baumeister, K., Ivanoff, L., Pettewat, S. R., Jr., Pearson, M. L., Lautenberger, J. A., Papas, T. S., Ghrayeb, J., Chang, N. T., Gallo, R. C. & Wong-Staal, F. (1985) *Nature (London)* **313,** 277–284.
16. Shinnick, T. M., Lerner, R. A. & Sutcliffe, J. G. (1981) *Nature (London)* **293,** 543–548.
17. Johnson, M. S., McClure, M. A., Feng, D.-F., Gray, J. & Doolittle, R. F. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 7648–7652.
18. Van Beveren, C., Coffin, J. & Hughes, S. (1985) in *RNA Tumor Viruses, 2/Supplements and Appendixes*, eds. Weiss, R., Natalie, T., Varmus, H. & Coffin, J. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), 2nd Ed., pp. 589–594.
19. Sagata, N., Yasunaga, T. & Ikawa, Y. (1984) *FEBS Lett.* **178,** 79–82.
20. Toh, H., Ono, M., Saigo, K. & Miyata, T. (1985) *Nature (London)* **315,** 691.
21. Gonda, M. A., Braun, M. J., Clements, J. E., Pyper, J. M., Wong-Staal, F., Gallo, R. C. & Gilden, R. V. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 4007–4011.
22. Lightfoote, M. M., Coligan, J. E., Folks, T. M., Fauci, A. S., Martin, M. A. & Venkatesan, S. (1986) *J. Virol.* **60,** 771–775.
23. Ono, M., Yasunaga, T., Miyata, T. & Ushikubo, H. (1986) *J. Virol.* **60,** 589–598.
24. Grindley, N. D. F. & Reed, R. R. (1985) *Ann. Rev. Biochem.* **54,** 880–887.
25. Duyk, G., Longiaru, M., Cobrinik, D., Kowal, R., deHaseth, P., Skalka, A. M. & Leis, J. (1985) *J. Virol.* **56,** 589–599.