JOURNAL OF VIROLOGY, Mar. 1994, p. 1468–1474
0022-538X/94/$04.00+0

Vol. 68, No. 3

# Activities of the Feline Immunodeficiency Virus Integrase Protein Produced in *Escherichia coli*

CORNELIS VINK, KARIN H. VAN DER LINDEN, AND RONALD H. A. PLASTERK*

*Division of Molecular Biology, The Netherlands Cancer Institute, 1066 CX Amsterdam, The Netherlands*

Received 4 October 1993/Accepted 30 November 1993

**Retroviral DNA integration requires the activity of at least one viral protein, the integrase (IN) protein. We cloned and expressed the integrase gene of feline immunodeficiency virus (FIV) in *Escherichia coli* as a fusion to the *malE* gene and purified the IN fusion protein by affinity chromatography. The protein is active in site-specific cleavage of the viral DNA ends, DNA strand transfer, and disintegration. FIV IN has a relaxed viral DNA substrate requirement: it cleaves and integrates FIV DNA termini, human immunodeficiency virus DNA ends, and Moloney murine leukemia virus DNA ends with high efficiencies. In the cleavage reaction, IN exposes a specific phosphodiester bond near the viral DNA end to nucleophilic attack. In vitro, either $H_2O$, glycerol, or the 3′ OH group of the viral DNA terminus can serve as nucleophile in this reaction. We found that FIV IN preferentially uses the 3′ OH ends of the viral DNA as nucleophile, whereas HIV IN protein preferentially uses $H_2O$ and glycerol as nucleophiles.**

Integration of a double-stranded DNA copy of the retroviral RNA into the DNA of the infected cell is essential for efficient replication of retroviruses. The following steps are involved in retroviral DNA integration (for recent reviews, see references 1, 5, 14, 15, 26, and 46). After infection of a susceptible cell, the RNA of the virus is copied into a double-stranded, blunt-ended DNA molecule, a process catalyzed by the viral enzyme reverse transcriptase. A few nucleotides (generally two) are subsequently removed from the 3′ ends of the viral DNA. After transfer to the nucleus, the newly generated viral 3′ OH ends are coupled to phosphate groups in both strands of the host DNA. These phosphate groups are 4 to 6 bp apart from each other (43). The non-base-pairing nucleotides at the 5′ ends of the viral DNA (which flank the single-stranded gaps in the host DNA) are removed in a repair step that is presumably carried out by cellular enzymes. The single-stranded gaps are probably also filled in in this step. Finally, the integrated retrovirus (or provirus) has lost (in general) 2 bp from each end and is flanked by short, direct duplications of the target DNA. The length of these duplications is specific for the integrating retrovirus (43). The proviral DNA is transcribed by host enzymes, and the resulting RNA is translated into protein or packaged into virions as new viral genomes. After budding from the cellular membrane, these virions can infect other permissive cells.

The only protein known to be required for integration is the viral integrase (IN) protein (6, 18, 19, 36). Purified recombinant IN both specifically cleaves the viral DNA ends and couples these ends to target DNA (2, 6, 18). Studies on human immunodeficiency virus type 1 (HIV-1) and HIV-2 IN have indicated that specific cleavage of the viral DNA termini (or donor cut) involves hydrolysis of the phosphodiester bond directly 3′ of a conserved 5′-CA-3′ sequence, resulting in release of a dinucleotide with a 5′ phosphate and a 3′ OH group (5′ p-GT-OH 3′ [10, 49]). The integration or DNA strand transfer reaction also proceeds in one step; a phosphodiester bond in the target DNA is broken while a new one is made between viral and target DNA (10). IN also possesses disintegration activity, which can be regarded as the reversal, in chemical terms, of the strand transfer reaction (4).

Analyses of mutants of HIV-1, HIV-2, and Rous sarcoma virus IN have identified three distinct regions in the protein. (i) The N terminus may play a role in recognition of the viral DNA ends (40, 44). (ii) The central region of the protein is conserved among all IN proteins and is also present in the transposases of several bacterial transposons (11, 20). This region is sufficient to carry out disintegration, and therefore it probably contains the complete (and single) active site of IN (3, 45). (iii) The C terminus of IN contains a DNA-binding domain; this domain binds both the viral DNA ends and nonspecific DNA with approximately equal affinity (32, 45, 51), as does the full-length protein (20, 23, 27, 32, 39, 51).

Both kinetic experiments (17) and in vitro complementation assays (9, 42) have indicated that the IN proteins are active as oligomers. Although IN forms predominantly dimers in solution (16, 17, 36, 24, 44), it has not yet been established whether this is indeed the active oligomeric state of the protein.

The IN proteins of several retroviruses have been studied extensively. There are three main reasons why we set out to characterize the IN protein of feline immunodeficiency virus (FIV). (i) Solving the three-dimensional structure of (part of) the IN protein by X-ray analysis first requires that crystals of the protein be grown. To increase the chance of crystal formation, it is recommended that several IN proteins from various sources be studied. (ii) The FIV system is the smallest natural animal model for infections with lentiviruses and might therefore serve as an important model to test the effects of inhibitors of IN on infections with these retroviruses (8, 28). Inhibitors of IN, however, will first have to be identified in in vitro assays with purified IN protein. (iii) The IN proteins specifically recognize their corresponding substrates (HIV-1 IN cleaves and integrates the HIV-1 DNA ends more efficiently than Moloney murine leukemia virus [MoMLV] ends), but sequence-specific DNA binding has not been observed. Also, it is not known which domains of the protein determine the DNA substrate specificity. Comparison of related integrases with different substrate specificities, and possibly the study of chimeric FIV–HIV-1 IN proteins, could lead to a better understanding of the substrate recognition by IN.

* Corresponding author. Mailing address: Division of Molecular Biology, The Netherlands Cancer Institute, Plesmanlaan 121, 1066 CX Amsterdam. Phone: 31 205122081. Fax: 31 206172625. Electronic mail address: RPLAS@NKI.NL.

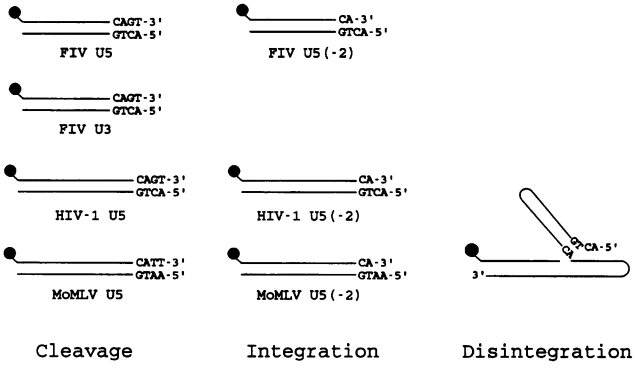



FIG. 1. Oligonucleotide substrates used in cleavage, integration, and disintegration assays. The viral end sequence is shown to the right of each substrate. Substrate FIV U5 represents the U5 terminus of FIV; it is a 20-mer of the sequence 5′-CGGGCCGAGAACTTCG CAGT-3′, annealed to its complement. Substrate FIV U3 is a 20-mer of the sequence 5′-GTTCCAGTACTCATCCCAGT-3′ and its complement. The sequence of substrate HIV-1 U5 (a 28-mer) is 5′-TTAGTCAGTGTGGAAAATCTCTAGCAGT-3′ and its complement. The MoMLV U5 substrate (28-mer) is 5′-GACTA CCCGTCAGCGGGGGTCTTTCATT-3′ and its complement. The integration substrates are identical in sequence to the respective cleavage substrates except for the lack of two nucleotides at the 3′ ends of the top strands. The disintegration substrate consists of two annealed oligonucleotides, a 12-mer (5′-GCAGCAGCTAGG-3′) and a 44-mer (5′-**ACTGCTAGTTCTAGCA**GCAGCCGTTCGGCTGCCC TAGCTGCTGC-3′; the viral DNA part of this oligonucleotide is indicated in boldface. The position of the $^{32}P$ label is represented by a closed circle. Cleavage reactions were also tested with substrates which were labeled at the 3′ end of the top strand, by incorporation of $^{32}P$ between the 3′-terminal G and T residues of substrates FIV U5 and HIV-1 U5. The FIV sequence is from reference 38, the HIV-1 sequence is from reference 50, and the MoMLV sequence is from reference 37.

Here we report the expression, purification, and characterization of the FIV IN protein. FIV IN is highly active in site-specific cleavage. However, the protein is less active than HIV-1 IN in DNA strand transfer. FIV IN has a more relaxed substrate sequence requirement than HIV-1 IN and is active in the presence of $Mn^{2+}$ but not $Mg^{2+}$.

## MATERIALS AND METHODS

**Construction of an FIV IN expression vector.** The IN gene of FIV was amplified by PCR, using the oligonucleotides AB3857 (5′-CACGG**GGATCC**TCCTCTTGGGTTGACAGA ATTG-3′; nucleotides 4397 to 4399 [38] are underlined and encode the N-terminal amino acid of IN [8]) and AB3858 (5′-GTGCCC**CTGCAG**TCACTCATCCCCTTCAGGAAGAG-3′; the underlined nucleotides are complementary to the stop codon of IN, positions 5240 to 5242 of the FIV sequence [38]). These primers introduce unique *Bam*HI (AB3857) and *Pst*I (AB3858) restriction sites (in boldface) in the amplified fragment. As the template, we used total DNA of FIV producer cell line FL-4 (52) (kindly provided by A. Pfauth, The Netherlands Cancer Institute, Amsterdam). DNA was isolated as described by Miller et al. (25). The PCR product was purified, digested with *Bam*HI and *Pst*I, and ligated into the *Bam*HI- and *Pst*I-digested vector pMAL-c (New England Biolabs). In the resulting plasmid (pRP817), the FIV IN is fused at its 5′ end to the 3′ end of the *Escherichia coli malE* gene.

**Expression and purification of the FIV IN fusion protein.** Construct pRP817 was introduced into *E. coli* JM101 (31). The

```
          1                                                  50
HIV IN    FLDGIDKAQD EHEKYHSNWR AMASDFNLPP VVAKEIVASC DKCQLKGEAM
FIV IN SSWVDRIEEAEI NHEKFHSDPQ YLRTEFNLPK MVAEEIRRKC PVCRIRGEQV
          1                                                  52

          51                                                 100
HIV IN    HGQVDCSPGI WQLDCTHLEG KVILVAVHVA SGYIEAEVIP AETGQETAYF
FIV IN    GGQLKIGPGI WQMDCTHFDG KIILVGIHVE SGYIWAQIIS QETADCTVKA
          53                                                 102

          101                                                150
HIV IN    LLKLAGRWPV KTIHTDNGSN FTGATVRAAC WWAGIKQEFG IPYNPQSQGV
FIV IN    VLQLLSAHNV TELQTDNGPN FKNQKMEGVL NYMGVKHKFG IPGNPQSQAL
          103                                                152

          151                                                200
HIV IN    VESMNKELKK IIGQVRDQAE HLKTAVQMAV FIHNFKRKGG IGGYSAGERI
FIV IN    VENVNHTLKV WIRKFLPETT SLDNALSLAV HSLNFKRRGR IGGMAPYELL
          153                                                202

          201                                                250
HIV IN    VDIIATDIQT KELQKQITKI QNFRVYYRDS RNSLWKGPAK LLWKGEGAVV
FIV IN    AQQESLRIQ. DYFSAIPQKL QAQWIYYKDQ KDKKWKGPMR VEYWGQGSVL
          203                                                251

          251                                          288
HIV IN    IQD.NSDIKV VPRRKAKIIR DYGKQMAGDD CVASRQDED
FIV IN    LKDEEKGYFL IPRRHIRRVP EPCALPEGDE
          252                                  281
```

FIG. 2. Alignment of the amino acid sequences of the 281-residue FIV IN and 288-residue HIV-1 IN. Amino acids that are identical between the proteins are in boldface; residues Gln-116 and Gly-145 of FIV IN are underlined (see Discussion).

resulting strain was grown overnight at 37°C in LB medium (31) containing 50 μg of ampicillin per ml. The culture was diluted 1:100 in 150 ml of TB medium (31) with ampicillin and grown at 37°C to an optical density at 600 nm of 0.8. Protein production was induced by addition of isopropyl-β-D-thiogalactopyranoside (IPTG) to a final concentration of 0.3 mM. After 2.5 h, the bacteria were harvested by centrifugation and resuspended in 5 ml of 10 mM sodium phosphate (pH 7.2)–1 M NaCl–1 mM EDTA–1 mM β-mercaptoethanol–1 mg of lysozyme per ml. The suspension was sonicated and centrifuged for 20 min at 10,000 rpm in a Beckman HB-4 rotor. To the supernatant, 5 ml of 10 mM sodium phosphate (pH 7.2)–1 mM EDTA–1 mM β-mercaptoethanol was added, and the material was loaded onto a 1.5-ml amylose column (New England Biolabs). The column was washed with 5 ml of buffer

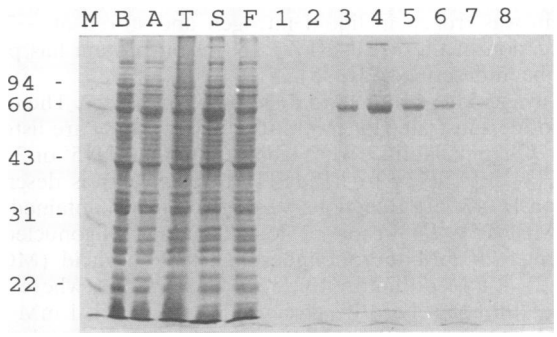



FIG. 3. Purification of MBP-FIV IN. Samples include crude extract from bacteria harboring plasmid pRP817, which expresses MBP-FIV IN, before (lane B) and after (lane A) induction with IPTG, the total sonication suspension (lane T), and the supernatant after centrifugation of this suspension (lane S). The sonication supernatant was loaded onto an amylose resin column. The flowthrough from the column is shown in lane F, and the elution fractions are shown in lanes 1 to 8. One microliter of the elution fractions was loaded. Fractions 3 to 7 were pooled, dialyzed, and tested for activity as described in Materials and Methods. Lane M, molecular size markers. Molecular sizes are indicated at the left in kilodaltons. Electrophoresis was carried out in a SDS–10% polyacrylamide gel.

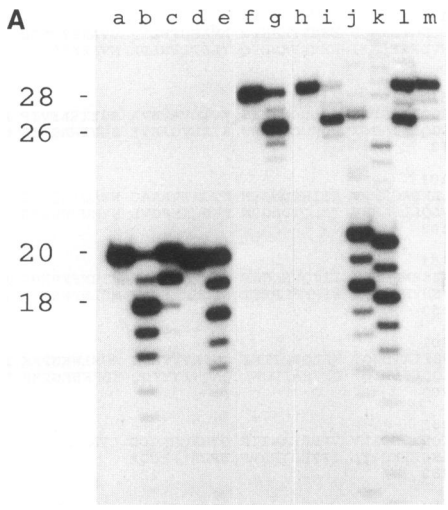
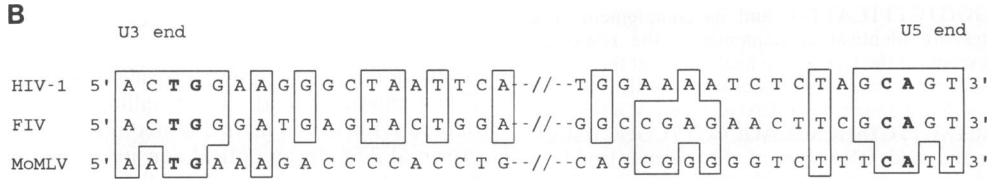



FIG. 4. Specific cleavage of 5′-end-labeled substrates by FIV IN. (A) Substrate FIV U5 (Fig. 1) was incubated without protein (lane a), with FIV IN (lane b), with FIV IN in the presence of $Mg^{2+}$ instead of $Mn^{2+}$ (lane c), or with HIV-1 IN (lane j). Substrate FIV U3 was incubated without protein (lane d), with FIV IN (lane e), or with HIV-1 IN (lane k). Substrate HIV-1 U5 was incubated without protein (lane f), with FIV IN (lane g), or with HIV-1 IN (lane l). Substrate MoMLV U5 was incubated without protein (lane h), with FIV IN (lane i), or with HIV-1 IN (lane m). The lengths of the substrates (20 and 28 nucleotides) and of the specific cleavage products (18 and 26 nucleotides) are indicated at the left. Products with a lower mobility than the substrates in lanes j and k are the results of DNA strand transfer. The samples were electrophoresed on a denaturing 15% polyacrylamide gel. (B) Alignment of the viral DNA ends of FIV (38), HIV-1 (50), and MoMLV (37). Only one of the two strands is shown. Nucleotides that are identical are boxed. The conserved 5′-CA-3′ sequences (or complementary 5′-TG-3′ sequences) are indicated in boldface.




A (10 mM sodium phosphate [pH 7.2], 0.5 M NaCl, 1 mM EDTA, 1 mM β-mercaptoethanol) supplemented with 0.25% Tween 20 and subsequently with 8 ml of buffer A. The maltose-binding protein (MBP)–IN fusion protein was eluted with 5 ml of buffer A containing maltose at 10 mM. Eight fractions of 0.5 ml were collected and analyzed by sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis. Peak fractions were pooled and concentrated three- to fourfold by dialysis to 10 mM Tris-HCl (pH 7.5)–0.5 M NaCl–0.5 mM EDTA–1 mM β-mercaptoethanol–40% glycerol. The protein was frozen in liquid nitrogen and stored at −70°C. Purification of the MBP–HIV-1 IN fusion protein has previously been described (22, 45).

**Cleavage, integration, and disintegration assays.** The oligonucleotide substrates that were used in this study are listed in Fig. 1. Oligonucleotides were labeled at either the 5′ or 3′ end (with [γ-$^{32}$P]ATP or [α-$^{32}$P]dTTP, respectively) as described previously (47, 49). Reactions were in 10 μl and contained 0.68 μM MBP-FIV IN fusion protein, 0.02 μM oligonucleotide substrate, 20 mM morpholinepropanesulfonic acid (MOPS; pH 7.2), 3 mM dithiothreitol, 3 mM $MnCl_2$ (or, where indicated, 5 mM $MgCl_2$), 10% glycerol, 50 mM NaCl, 1 mM Tris, 50 μM EDTA, and 0.1 mM β-mercaptoethanol. Incubations were for 1 h at 30°C. Integration reaction mixtures were preincubated for 10 min at room temperature, after which 5 pmol of unlabeled nonspecific target DNA was added (a double-stranded oligonucleotide of the sequence 5′-GGGG TCCTATAGAACTTTCCCACACTGT-3′ and its complement). Reactions were stopped by addition of 10 μl of formamide loading dye (95% formamide, 20 mM EDTA, 0.05% bromophenol blue, 0.05% xylene cyanol) and analyzed by denaturing polyacrylamide gel electrophoresis followed by autoradiography.

**Southwestern (DNA-protein) blot analysis.** Southwestern blots were generated essentially as previously described (51). As the specific probe, plasmid pRP131, was used. This plasmid contains the terminal U3 and U5 sequences of HIV-1 (see reference 48 for a full description of this plasmid). As the nonspecific probe, plasmid PiAN7 (34) was used; this plasmid is the backbone of plasmid pRP131. The plasmids were labeled by random priming in the presence of [α-$^{32}$P]dATP (31).

## RESULTS

**Cloning, expression, and purification of FIV IN.** The FIV IN gene was amplified by PCR, using DNA from FIV producer cell line FL-4 (52) as the template. The primers were chosen so as to incorporate unique endonuclease restriction sites in the amplified product, thereby facilitating cloning of the product into expression vector pMAL-c. The resulting plasmid was introduced into *E. coli* and encodes a MBP-FIV IN fusion protein. As the N-terminal amino acid of FIV IN, residue Ser-840 (8) was chosen. This residue has been identified as the N-terminal residue of native FIV IN, after proteolytic processing of a Gag-Pol polyprotein precursor (8). The sequence of the 281-amino-acid FIV IN protein is aligned with the 288-amino-acid HIV-1 IN sequence in Fig. 2. The proteins are approximately 37% identical. FIV IN contains two additional (Ser) residues at the N terminus but is nine amino acids shorter than the HIV-1 protein at the C terminus. The nine (additional) residues at the C terminus of HIV-1 IN can be deleted without effect on the activities of the protein in vitro (45).

The MBP-FIV IN fusion was expressed, and purified by lysis of the bacteria in high salt, followed by affinity chromatography over an amylose resin column (Fig. 3; see Materials and Methods). The purified protein, which has a molecular weight of approximately 73,000, was ~80% pure, as judged from Coomassie blue-stained gels.

**Cleavage of 5′-end-labeled substrates.** Site-specific cleavage activity of the purified FIV IN fusion protein was tested on several radiolabeled oligonucleotide substrates (Fig. 1). The substrates are identical in sequence to the viral DNA ends of FIV, HIV-1, and MoMLV and are labeled with $^{32}P$ at the 5′ ends of the strands that are expected to be cleaved by IN. The termini of the FIV DNA were predicted on the basis of similarity with the sequences of other retroviruses. As shown in Fig. 4A, FIV IN cleaves both the U5 and U3 FIV DNA termini, generating a prominent product which is two nucleotides shorter than the substrate. FIV IN cleaves the U3 substrate less efficiently and less specifically than the U5 substrate (compare lanes b and e in Fig. 4A); a similar observation has previously been made for the HIV-1 IN protein (2, 35, 39). The FIV IN fusion protein shows efficient cleavage of two nucleotides from the viral substrate only in the presence of $Mn^{2+}$, not in the presence of $Mg^{2+}$ (lane c). In the presence of $Mg^{2+}$, only nonspecific cleavage activity was observed. Purified recombinant IN proteins of HIV and MoMLV show similar divalent cation dependence (6, 36).

Although the FIV, HIV-1, and MoMLV substrates differ considerably in sequence (Fig. 4B), these substrates are cleaved by FIV IN with high efficiencies (Fig. 4A, lanes b, g, and i), showing that FIV IN has a relaxed DNA substrate requirement in donor cutting. Note that cleavage of the HIV and MoMLV substrates results in a lower amount of background products (e.g., the product at the −1 position) than cleavage of the FIV substrates. This is, however, probably due to the lengths of the various substrates: it was found for HIV-1 IN that cleavage of 28-bp substrates gave rise to a lower background than cleavage of 15-bp substrates of identical terminal sequence (47). Similarly, the reactions with 20-bp FIV substrates and HIV-1 IN (Fig. 4A, lanes j and k) have a higher background than the reaction with the 28-bp HIV-1 substrate and HIV-1 IN (lane l).

Interestingly, the HIV-1 IN protein cleaves both FIV substrates as efficiently as the HIV-1 U5 substrate but cleaves the MoMLV substrate rather poorly (Fig. 4A, lane m), as has previously been described (39, 47).

**Cleavage of 3′-end-labeled substrates.** To identify the products that are released from the viral DNA substrate during cleavage, we tested substrates that were radiolabeled within the dinucleotide portion at the 3′ side of the IN cleavage site. Both an FIV and a HIV-1 U5 substrate were tested. Figure 5 shows that cleavage of these substrates by FIV IN in the presence of $Mn^{2+}$ gives rise to one major product and two minor products. These different products have previously been identified in studies of HIV-1 IN as (i) a linear dinucleotide, which is released during hydrolysis of the phosphate group 3′ of the conserved 5′-CA-3′ sequence, (ii) a product in which glycerol is covalently coupled to the 3′-terminal dinucleotide via a 5′ phosphodiester bond, and (iii) a circular dinucleotide, generated by nucleophilic attack of the 3′ OH group of the viral DNA terminus at the target phosphate group (10, 49). FIV IN cleaves both the FIV and the HIV substrate in similar fashion (Fig. 5, lanes c and g). However, there is difference in the relative amounts of the various cleavage products produced by FIV and HIV-1 IN. HIV-1 IN predominantly generates the linear dinucleotide and glycerol-derived product (lanes i and k), whereas FIV IN mainly generates the cyclic dinucleotide product. Obviously, FIV and HIV-1 IN have different preferences for the various nucleophiles in site-specific cleavage ($H_2O$, glycerol, or the viral 3′ OH end).

**Integration and disintegration activity of FIV IN.** Integration activity of FIV IN was tested with precleaved substrates that represent the viral DNA U5 ends of FIV, HIV-1, and MoMLV (Fig. 6A). FIV IN integrates both the FIV and HIV-1 substrates (Fig. 6A, lanes b and e) and also, albeit with reduced efficiency, the MoMLV substrate (lane h) into a target oligonucleotide. The protein is active only in the presence of $Mn^{2+}$ in the reaction mixture. The pattern of integration products that is formed by the FIV protein clearly differs from the pattern generated by HIV-1 IN (compare lanes b and i and lanes e and j). This result shows that these proteins have different integration site preferences. The integration activity of FIV IN does not seem to parallel its high activity in site-specific cleavage; in the cleavage reactions (Fig. 4), FIV IN converts approximately 85% of the substrate into the specific recessed product, whereas in integration, less than 10% of the substrate is integrated into target DNA. In contrast, the activity of HIV-1 IN in integration (Fig. 6A, lane j) reflects its activity in donor cutting rather well: in both reactions, ~50% of the substrate oligonucleotide is converted into product. As was the case for donor cutting, the integration activities of both FIV and HIV-1 IN are higher with the longer (26-bp) HIV-1 substrate than with the smaller (18-bp) FIV substrate.

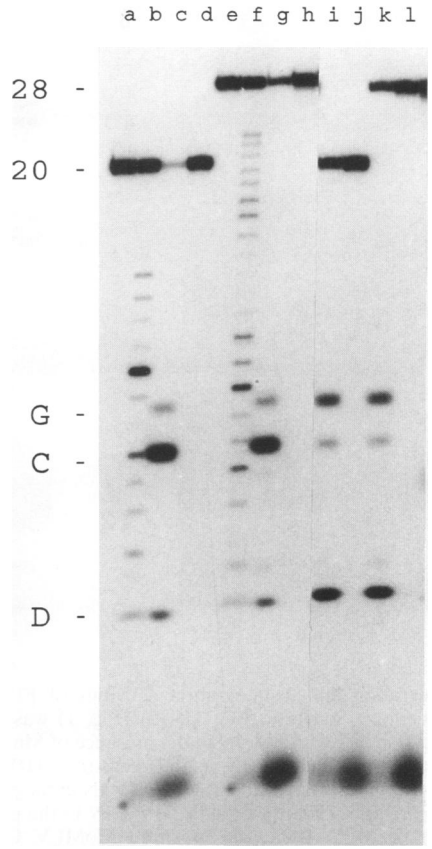



FIG. 5. Specific cleavage of 3′-end-labeled substrates by FIV IN. Substrate FIV U5 (Fig. 1) was incubated without protein (lane 1), with DNase I (as size marker; lane b), with FIV IN in the presence of $Mn^{2+}$ (lane c) or $Mg^{2+}$ (lane d), or with HIV-1 IN in the presence of $Mn^{2+}$ (lane i) or $Mg^{2+}$ (lane j). Substrate HIV-1 U5 was incubated without protein (lane e), with DNase I (lane f), with FIV IN in the presence of $Mn^{2+}$ (lane g) or $Mg^{2+}$ (lane h), or with HIV-1 IN in the presence of $Mn^{2+}$ (lane k) or $Mg^{2+}$ (lane l). The lengths of the substrates are shown on the left (in nucleotides). The specific dinucleotide products, the linear form (D), the cyclic form (C), and the glycerol-derived product (G), are also indicated. Reaction products were separated on a denaturing 24% polyacrylamide gel.

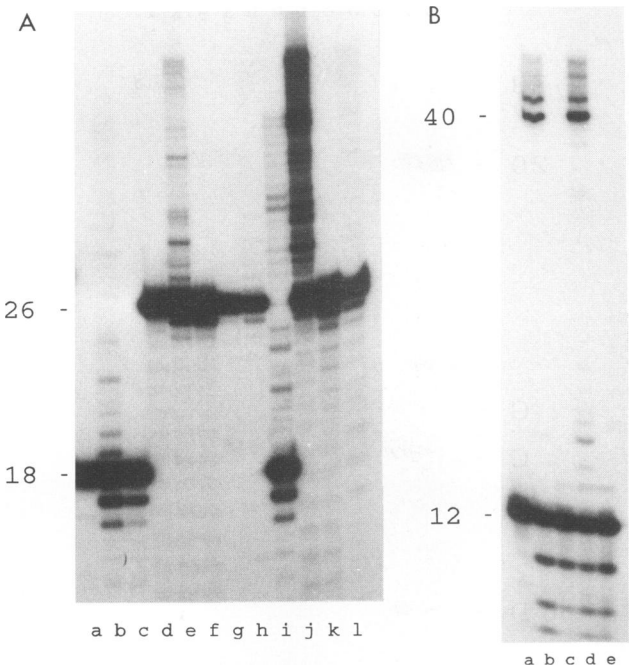



FIG. 6. Integration and disintegration activities of FIV IN. (A) Integration reactions. Substrate FIV U5(-2) (Fig. 1) was incubated without protein (lane a), with FIV IN in the presence of $Mn^{2+}$ (lane b) or $Mg^{2+}$ (lane c), or with HIV-1 IN (lane i). Substrate HIV-1 U5(-2) was incubated without protein (lane d), with FIV IN in the presence of $Mn^{2+}$ (lane e) or $Mg^{2+}$ (lane f), or with HIV-1 IN in the presence of $Mn^{2+}$ (lane j) or $Mg^{2+}$ (lane k). Substrate MoMLV U5(-2) was incubated without protein (lane g), with FIV IN (lane h), or with HIV-1 IN (lane l). The lengths of the various substrates (18 and 26 nucleotides) are indicated on the left. Integration products are visible as products with a lower mobility than the substrates. Samples were electrophoresed on a denaturing 12% polyacrylamide gel. (B) Disintegration reactions. The disintegration substrate was incubated without protein (lane a), with FIV IN in the presence of $Mn^{2+}$ (lane b) or $Mg^{2+}$ (lane c), or with HIV-1 IN in the presence of $Mn^{2+}$ (lane d) or $Mg^{2+}$ (lane e). The lengths of the substrate and the product are indicated at the left in nucleotides. The ladder of bands that is observed near the product band is probably caused by secondary structures within the product (4).

Disintegration can be seen as the reversal, in chemical terms, of the integration reaction (4). Disintegration activity is tested by using a branched DNA substrate which resembles the product of integration of a viral DNA end into a target DNA (Fig. 1). In analogy with the integration reaction (10), disintegration is probably a one-step polynucleotidyl transfer reaction, in which a new phosphodiester bond is generated (in the target part of the substrate) while another one is broken (between viral and target DNA), resulting in release of the viral DNA part of the substrate. As shown in Fig. 6B, FIV IN is active in disintegration in the presence of $Mn^{2+}$ (lane b) but not in the presence of $Mg^{2+}$ (lane c). Although a low level of disintegration activity of HIV-1 IN could be detected in the presence of $Mg^{2+}$ (after prolonged autoradiography of the gel in Fig. 6B), HIV-1 IN is also mainly active in the presence of $Mn^{2+}$ (compare lanes d and e).

**DNA-binding properties of FIV IN.** The ability of FIV IN to discriminate between viral and nonviral DNA was investigated by Southwestern blot analysis, using as the probe either a plasmid containing viral end sequences or a plasmid that lacks any viral sequences (Fig. 7). The FIV IN protein binds both the viral probe (Fig. 7C, lane d) and the nonviral probe (Fig. 7B, lane d) approximately equally well. Similar results have been obtained for the IN proteins from HIV-1 and HIV-2 (Fig. 7B and C, lanes c) (23, 32, 39, 51), MoMLV (29), and avian myeloblastosis virus (27).

## DISCUSSION

The IN protein of FIV was expressed as an MBP fusion protein in *E. coli* and purified in one step. The purified protein is active in DNA binding, cleavage of the viral DNA ends, integration of these ends into target DNA, and disintegration. FIV IN shows a relaxed sequence requirement for site-specific cleavage and integration of viral DNA ends; the protein is active on FIV, HIV-1, and MoMLV DNA termini. In contrast, the HIV-1 IN protein is active on HIV-1 and FIV ends but is barely active on MoMLV substrates.

A clear difference in activity between FIV and HIV-1 IN is the choice of nucleophiles in the site-specific cleavage reaction in vitro. HIV IN can use either $H_2O$, glycerol, or the 3′ OH end of the oligonucleotide substrate as nucleophile. Under the reaction conditions that we describe here, HIV IN prefers to use $H_2O$ and glycerol as nucleophiles. FIV IN, however, preferentially uses the 3′ OH end of the viral DNA substrate as nucleophile, resulting in release of a cyclic dinucleotide product. A similar preference for the 3′ OH viral DNA end as nucleophile has previously been found for a certain class of HIV-2 IN active-site mutants, so-called alcoholysis mutants (41). Wild-type HIV-2 IN cleaves the viral oligonucleotide substrates similarly to HIV-1 IN (Fig. 5) (41, 49). The introduction of certain point mutations in the active site of the protein, however, alters the relative amounts in which the various dinucleotide products are formed. Mutation of either amino acid His-114, Asn-117, Tyr-143, Gln-146, Ser-147, or Gln-148 of HIV-2 IN results in the alcoholysis cleavage phenotype (41). Interestingly, the FIV IN cleavage phenotype can be explained by comparing the amino acid sequences of HIV and FIV IN. Changing amino acid Tyr-143 to Leu in HIV-2 IN causes the altered cleavage pattern of the protein. This Tyr residue is conserved among IN proteins from a large number of retroviruses, including HIV-1, MoMLV, simian immunodeficiency virus (21), bovine immunodeficiency virus (13), and feline leukemia virus (7). However, the residue analogous to Tyr-143 in HIV IN is Gly-145 in the FIV IN protein (see the alignment in Fig. 2). The presence of a Gly residue at position 145 in FIV IN could therefore explain the difference in nucleophile choice of the HIV IN proteins and the FIV protein. In addition, a Gln residue is present at position 116 in FIV IN, which corresponds to position 114 of the HIV IN proteins; as described above, alteration of this less well conserved (His) residue also results in a different use of nucleophiles by the mutant protein (41).

Another property that is shared between FIV IN and HIV-2 IN alcoholysis mutants, and which is probably directly related to the choice of nucleophiles, is the divalent cation dependence. Purified, soluble HIV-1 and HIV-2 IN are inactive in the presence of $Mg^{2+}$ (instead of $Mn^{2+}$; [2, 23, 24, 36, 39, 44, 47]). After immobilization to thiopropyl-Sepharose, however, these proteins are active in the presence of both $Mn^{2+}$ and $Mg^{2+}$ (41, 49). Interestingly, in the presence of $Mg^{2+}$, IN almost exclusively selects $H_2O$ and a relatively low amount of glycerol as nucleophile but not the viral 3′ OH end (41, 49). This could explain why the (immobilized) HIV-2 alcoholysis mutants (41) and (immobilized) FIV IN are active only with $Mn^{2+}$ (46a): the use of the 3′ OH group of the viral DNA end

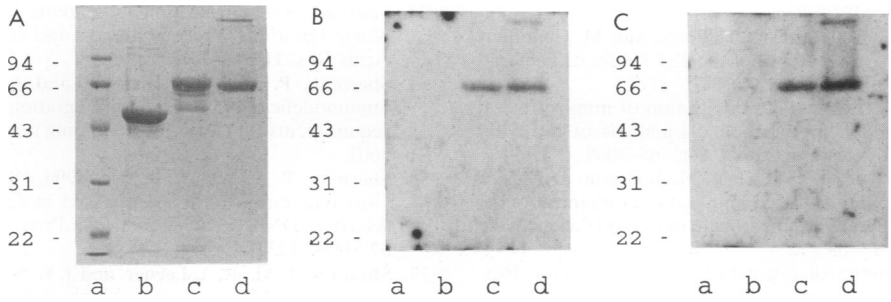



FIG. 7. Southwestern blot analysis of FIV IN. (A) Coomassie blue-stained gel with purified MBP–β-galactosidase α (negative control; lane b), MBP-HIV-1 IN (lane c), and MBP-FIV IN (lane d). Gels similar to the one shown in panel A were transferred to nitrocellulose, and DNA-binding activity of the proteins on the filter was tested by incubation with a radiolabeled probe that does not contain viral DNA (B) or a probe that contains viral DNA end sequences (C). DNA binding was detected by autoradiography. Lane a, Molecular weight markers. The molecular sizes of the marker proteins are indicated at the left in kilodaltons.

as nucleophile in site-specific cleavage is an artifact resulting from the nonphysiologically high concentration of $Mn^{2+}$ and is not found in the presence of $Mg^{2+}$. Van Gent et al. (41) have suggested that the amino acid residues which, when altered, give rise to the alcoholysis phenotype are probably required for efficient cleavage by IN in vivo, since they are essential for the donor cut reaction under physiological ion conditions (low $Mn^{2+}$ concentration). The alcoholysis mutations, however, can be present in naturally occurring IN proteins of FIV, suggesting that IN proteins that carry these alcoholysis mutations are functional in vivo. As in FIV IN, a Gly residue is present in Rous sarcoma virus IN (33) at the position that is analogous to Tyr-143 of HIV, and a Trp residue is at that position in caprine arthritis-encephalitis virus (30). Elucidation of the three-dimensional structure of IN, complexed with (viral) DNA in the presence of various divalent cations, might clarify how the various amino acid side chains in the catalytic pocket of IN influence the choice of nucleophile in site-specific cleavage.

An apparent discrepancy in the activities of the FIV IN protein is the high activity of the protein in site-specific cleavage and the relatively low activity in DNA strand transfer or integration. In site-specific cleavage, FIV IN converts approximately 85% of the oligonucleotide substrate into product, whereas in integration, less than 10% of the substrate is converted. It could be possible that the FIV IN protein that we tested is hindered in integration because of the N-terminal MBP part of the MBP-FIV IN fusion protein. However, MBP–HIV-1 IN fusion proteins are indistinguishable in activity from purified, nonfusion HIV-1 IN (42, 45). Furthermore, we expressed FIV IN from an infectious molecular clone of FIV as a fusion with six His residues N terminally of the IN protein. This protein showed the same characteristics as the MBP-FIV IN fusion protein that we describe here (46a). A similar difference between activity levels in the cleavage and integration reaction was found for one of the HIV-2 IN alcoholysis mutants, Asn-117→Ile (40).

There are obvious differences between the in vitro conditions under which activity of IN is tested and the conditions under which IN is active in vivo. For instance, specific activity of FIV IN is detectable only in the presence of nonphysiological concentrations of $Mn^{2+}$ (1 to 3 mM). In the integration reaction in vivo, IN will have to promote the concerted, pairwise integration of two viral DNA termini into target DNA in order to establish the proviral state. However, we were not able to detect these pairwise integration events with purified recombinant FIV IN, using both a PCR-based assay and a genetic integration assay (6, 12, 18, 46a). Similar observations were made for the HIV-1 and HIV-2 IN proteins (for a review, see reference 46). It is possible that IN molecules and viral DNAs do not assemble properly in vitro or that additional proteins are required for either formation of proper nucleoprotein complexes or stabilization of these complexes. Detailed molecular analysis of the complexes that are generated in vivo could clarify this.

## ACKNOWLEDGMENTS

We thank Sean Colloms, Fusinita van den Ent, and Ramon Puras Lutzke for critical reading of the manuscript. We thank A. Pfauth for the gift of cell line FL-4.

This work was supported by the Netherlands Organization for Scientific Research, by the Netherlands AIDS Foundation, and by a grant from Glaxo Group Research Limited to C.V.

## REFERENCES

1. **Brown, P. O.** 1990. Integration of retroviral DNA. Curr. Top. Microbiol. Immunol. **157:**19–48.
2. **Bushman, F. D., and R. Craigie.** 1991. Activities of human immunodeficiency virus (HIV) integration protein in vitro: specific cleavage and integration of HIV DNA. Proc. Natl. Acad. Sci. USA **88:**1339–1343.
3. **Bushman, F. D., A. Engelman, I. Palmer, P. Wingfield, and R. Craigie.** 1993. Domains of the integrase protein of human immunodeficiency virus type 1 responsible for polynucleotidyl transfer and zinc binding. Proc. Natl. Acad. Sci. USA **90:**3428–3432.
4. **Chow, S. A., K. A. Vincent, V. Ellison, and P. O. Brown.** 1992. Reversal of integration and DNA splicing mediated by integrase of human immunodeficiency virus. Science **255:**723–726.
5. **Craigie, R.** 1992. Hotspots and warm spots: integration specificity of retroelements. Trends Genet. **8:**187–190.
6. **Craigie, R., T. Fujiwara, and F. D. Bushman.** 1990. The IN protein of Moloney murine leukemia virus processes the viral DNA ends and accomplishes their integration in vitro. Cell **62:**829–837.
7. **Donahue, P. R., E. A. Hoover, G. A. Beltz, N. Riedel, V. M. Hirsch, J. Overbaugh, and J. I. Mullins.** 1988. Strong sequence conservation among horizontally transmissible, minimally pathogenic feline leukemia viruses. J. Virol. **62:**722–731.
8. **Elder, J. H., M. Schnolzer, C. S. Hasselkus-Light, M. Henson, D. A. Lerner, T. R. Phillips, P. C. Wagaman, and S. B. H. Kent.** 1993. Identification of proteolytic processing sites within the Gag and Pol polyproteins of feline immunodeficiency virus. J. Virol. **67:**1869–1876.
9. **Engelman, A., F. D. Bushman, and R. Craigie.** 1993. Identification of discrete functional domains of HIV-1 integrase and their organization within an active multimeric complex. EMBO J. **12:**3269–3275.
10. **Engelman, A., K. Mizuuchi, and R. Craigie.** 1991. HIV-1 DNA integration: mechanism of viral DNA cleavage and DNA strand

transfer. Cell **67**:1211–1221.
11. **Fayet, O., P. Ramond, P. Polard, M. F. Prère, and M. Chandler.** 1990. Functional similarities between the IS3 family of bacterial insertion elements? Mol. Microbiol. **4**:1771–1777.
12. **Fujiwara, T., and R. Craigie.** 1989. Integration of mini-retroviral DNA: a cell-free reaction for biochemical analysis of retroviral integration. Proc. Natl. Acad. Sci. USA **86**:3065–3069.
13. **Garvey, K. J., M. S. Oberste, J. E. Elser, M. J. Braun, and M. A. Gonda.** 1990. Nucleotide sequence and genome organization of biologically active proviruses of the bovine immunodeficiency-like virus. Virology **175**:391–409.
14. **Goff, S. P.** 1992. Genetics of retroviral integration. Annu. Rev. Genet. **26**:527–544.
15. **Grandgenett, D. P., and S. R. Mumm.** 1990. Unravelling retrovirus integration. Cell **60**:3–4.
16. **Grandgenett, D. P., A. C. Vora, and R. D. Schiff.** 1978. A 32,000-dalton nucleic acid-binding protein from avian retrovirus cores possesses DNA endonuclease activity. Virology **89**:119–132.
17. **Jones, K. S., J. Coleman, G. W. Merkel, T. M. Laue, and A. M. Skalka.** 1992. Retroviral integrase functions as a multimer and can turnover catalytically. J. Biol. Chem. **287**:16037–16040.
18. **Katz, R. A., G. Merkel, J. Kulkosky, J. Leis, and A. M. Skalka.** 1990. The avian retroviral IN protein is both necessary and sufficient for integrative recombination in vitro. Cell **63**:87–95.
19. **Katzman, M., R. A. Katz, A. M. Skalka, and J. Leis.** 1989. The avian retroviral integration protein cleaves the terminal sequences of linear viral DNA at the in vivo sites of integration. J. Virol. **63**:5319–5327.
20. **Khan, E., J. P. G. Mack, R. A. Katz, J. Kulkosky, and A. M. Skalka.** 1991. Retroviral integrase domains: DNA binding and the recognition of LTR sequences. Nucleic Acids Res. **19**:851–860.
21. **Kornfeld, H., N. Riedel, G. A. Viglianti, V. Hirsch, and J. I. Mullins.** 1987. Cloning of HTLV-4 and its relation to simian and human immunodeficiency viruses. Nature (London) **326**:610–613.
22. **Kulkosky, J., K. S. Jones, R. A. Katz, J. P. G. Mack, and A. M. Skalka.** 1992. Residues critical for retroviral integrative recombination in a region that is highly conserved among retroviral/retrotransposon integrases and bacterial insertion sequence transposases. Mol. Cell. Biol. **12**:2331–2338.
23. **LaFemina, R. L., P. L. Callahan, and M. G. Cordingley.** 1991. Substrate specificity of recombinant human immunodeficiency virus integrase protein. J. Virol. **65**:5624–5630.
24. **Leavitt, A. D., R. B. Rose, and H. E. Varmus.** 1992. Both substrate and target oligonucleotide sequences affect in vitro integration mediated by human immunodeficiency virus type 1 integrase protein produced in *Saccharomyces cerevisiae*. J. Virol. **66**:2359–2368.
25. **Miller, S. A., D. D. Dykes, and H. F. Polesky.** 1988. A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res. **16**:1215.
26. **Mizuuchi, K.** 1992. Polynucleotidyl transfer reactions in transpositional DNA recombination. J. Biol. Chem. **267**:21273–21276.
27. **Mumm, S. R., and D. P. Grandgenett.** 1991. Defining nucleic acid-binding properties of avian retrovirus integrase by deletion analysis. J. Virol. **65**:1160–1167.
28. **Pedersen, N. C., E. W. Ho, M. L. Brown, and J. K. Yamamoto.** 1987. Isolation of a T-lymphotropic virus from domestic cats with an immunodeficiency-like syndrome. Science **235**:790–793.
29. **Roth, M. J., N. Tanese, and S. P. Goff.** 1988. Gene product of Moloney murine leukemia virus required for proviral integration is a DNA-binding protein. J. Mol. Biol. **203**:131–139.
30. **Saltarelli, M., G. Querat, D. A. Konings, R. Vigne, and J. E. Clements.** 1990. Nucleotide sequence and transcriptional analysis of molecular clones of caprine arthritis encephalitis virus which generate infectious virus. Virology **179**:347–364.
31. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
32. **Schauer, M., and A. Billich.** 1992. The N-terminal region of HIV-1 integrase is required for integration activity. Biochem. Biophys. Res. Commun. **186**:874–880.
33. **Schwartz, D. E., R. Tizard, and W. Gilbert.** 1983. Nucleotide sequence of Rous sarcoma virus. Cell **32**:853–869.
34. **Seed, B.** 1983. Purification of genomic sequences from bacteriophage libraries by recombination and selection in vivo. Nucleic Acids Res. **11**:2427–2447.
35. **Sherman, P. A., M. L. Dickson, and J. A. Fyfe.** 1992. Human immunodeficiency virus type 1 integration protein: DNA sequence requirements for cleavage and joining reactions. J. Virol. **66**:3593–3601.
36. **Sherman, P. A., and J. A. Fyfe.** 1990. Human immunodeficiency virus integration protein expressed in *Escherichia coli* possesses selective DNA cleaving activity. Proc. Natl. Acad. Sci. USA **87**:5119–5123.
37. **Shinnick, T. M., R. A. Lerner, and J. G. Sutcliffe.** 1981. Nucleotide sequence of Moloney murine leukaemia virus. Nature (London) **293**:543–548.
38. **Talbott, R. L., E. E. Sparger, K. M. Lovelace, W. M. Fitch, N. C. Pedersen, P. A. Luciw, and J. H. Elder.** 1989. Nucleotide sequence and genomic organization of feline immunodeficiency virus. Proc. Natl. Acad. Sci. USA **86**:5743–5747.
39. **Van Gent, D. C., Y. Elgersma, M. W. J. Bolk, C. Vink, and R. H. A. Plasterk.** 1991. DNA binding properties of the integrase proteins of human immunodeficiency viruses types 1 and 2. Nucleic Acids Res. **19**:3821–3827.
40. **Van Gent, D. C., A. A. M. Oude Groeneger, and R. H. A. Plasterk.** 1992. Mutational analysis of the integrase protein of human immunodeficiency virus type 2. Proc. Natl. Acad. Sci. USA **89**:9598–9602.
41. **Van Gent, D. C., A. A. M. Oude Groeneger, and R. H. A. Plasterk.** 1993. Identification of amino acids in HIV-2 integrase involved in site-specific hydrolysis and alcoholysis of viral DNA termini. Nucleic Acids Res. **21**:3373–3377.
42. **Van Gent, D. C., C. Vink, A. A. M. Oude Groeneger, and R. H. A. Plasterk.** 1993. Complementation between HIV integrase proteins mutated in different domains. EMBO J. **12**:3261–3267.
43. **Varmus, H., and P. Brown.** 1989. Retroviruses, p. 52–108. *In* D. E. Berg and M. M. Howe (ed.), Mobile DNA. American Society for Microbiology, Washington, D.C.
44. **Vincent, K. A., V. Ellison, S. A. Chow, and P. O. Brown.** 1993. Characterization of human immunodeficiency virus type 1 integrase expressed in *Escherichia coli* and analysis of variants with amino-terminal mutations. J. Virol. **67**:425–437.
45. **Vink, C., A. A. M. Oude Groeneger, and R. H. A. Plasterk.** 1993. Identification of the catalytic and DNA-binding region of the human immunodeficiency virus type 1 integrase protein. Nucleic Acids Res. **21**:1419–1425.
46. **Vink, C., and R. H. A. Plasterk.** 1993. The human immunodeficiency virus integrase protein. Trends Genet. **9**:433–437.
46a. **Vink, C., K. H. van der Linden, and R. H. A. Plasterk.** Unpublished data.
47. **Vink, C., D. C. van Gent, Y. Elgersma, and R. H. A. Plasterk.** 1991. Human immunodeficiency virus integrase protein requires a subterminal position of its viral DNA recognition sequence for efficient cleavage. J. Virol. **65**:4636–4644.
48. **Vink, C., D. C. van Gent, and R. H. A. Plasterk.** 1990. Integration of human immunodeficiency virus types 1 and 2 DNA in vitro by cytoplasmic extracts of Moloney murine leukemia virus-infected mouse NIH 3T3 cells. J. Virol. **64**:5219–5222.
49. **Vink, C., E. Yeheskiely, G. A. van der Marel, J. H. van Boom, and R. H. A. Plasterk.** 1991. Site-specific hydrolysis and alcoholysis of human immunodeficiency virus DNA termini mediated by the viral integrase protein. Nucleic Acids Res. **19**:6691–6698.
50. **Wain-Hobson, S., P. Sonigo, O. Danos, S. Cole, and M. Alizon.** 1985. Nucleotide sequence of the AIDS virus, LAV. Cell **40**:9–17.
51. **Woerner, A. M., M. Klutch, J. G. Levin, and C. J. Marcus-Sekura.** 1992. Localization of DNA binding activity of HIV-1 integrase to the C-terminal half of the protein. AIDS Res. Hum. Retroviruses **8**:2433–2437.
52. **Yamamoto, J. K., C. D. Ackley, H. Zochlinski, H. Louie, E. Pembroke, M. Torten, H. Hansen, R. Munn, and T. Okuda.** 1991. Development of IL-2 independent feline lymphoid cell lines chronically infected with feline immunodeficiency virus: importance for diagnostic reagents and vaccines. Intervirology **32**:361–375.