JOURNAL OF VIROLOGY, Sept. 1993, p. 5426–5434
0022-538X/93/095426-09$02.00/0

Vol. 67, No. 9

Endonucleolytic Cleavages and DNA-Joining Activities of the Integration Protein of Human Foamy Virus

ARMIN PAHL AND ROLF M. FLÜGEL*

Abteilung Retrovirale Genexpression, Angewandte Tumorvirologie, Deutsches Krebsforschungszentrum, 69120 Heidelberg, Germany

Received 12 February 1993/Accepted 15 June 1993

The bacterial expression plasmids, pET3b and pET16b, that contain the integrase domain of the human foamy virus (HFV) reverse transcriptase were constructed and expressed in *Escherichia coli*. The histidine-tagged HFV IN protein was purified to near homogeneity by single-step Ni^{2+} chelate affinity chromatography. HFV-specific proteins of 39 and 120 kDa from virus-infected cells reacted with antisera raised against the recombinant IN protein. Purified recombinant HFV IN protein was active as an endonuclease specifically cleaving two nucleotides from a 20-bp oligodeoxynucleotide substrate that mimics the authentic 5′ ends of HFV DNA. Substrates with mutations relatively close to the cleavage site were less efficiently cleaved or not cleaved at all compared with the HFV U5 DNA end. The purified recombinant protein was active as integrase with double-stranded oligodeoxynucleotide substrates. The reverse reaction of DNA strand transfer, the disintegration activity, was shown by efficient cleavage of an intermediate Y-shaped oligodeoxynucleotide. In the presence of Mn^{2+} as the preferred divalent cation, oligodeoxynucleotides were specifically and efficiently cleaved. In contrast, endonucleolytic cleavages in the presence of Mg^{2+} ions led to a broad range of reaction products with the His-tagged HFV IN protein. After further purification of the HFV IN by cation-exchange chromatography, the unspecific degradation of oligonucleotide substrate in the presence of Mg^{2+} was not detectable.

Foamy or spumaviruses constitute the third genus of the family *Retroviridae*. They are complex both in structure and in function, since novel *bel* genes besides *gag*, *pol*, and *env* were identified (for recent reviews, see references 14, 15, and 33). The prototype human foamy virus (HFV) was obtained from a patient with a nasopharyngeal carcinoma (1). The complete DNA sequence of this isolate has been determined, and an infectious HFV DNA clone was constructed and analyzed (16, 31). The *pol* gene of all known foamy viruses encodes several distinct domains (20, 25). The amino-terminal part contains the protease region (15, 35) that is followed by the reverse transcriptase domain, a relatively extensive hinge or tether region, the ribonuclease H, and the integrase (IN) domains. Retroviral IN proteins play an essential role in integration, a key step in the life cycle of retroviruses (2, 4, 9, 17, 18, 23). Foamy viruses encode an integrase protein that has a high degree of homology to IN sequences of different retroviruses and retroid elements (12, 16, 24–26, 32, 36). The highest degree of homology was found to be that to the IN sequence of Moloney murine leukemia virus (MLV) (16, 21) including the HHCC metal finger domain where both the HFV and the MLV sequences diverge substantially from other retroviral IN sequences. This domain was reported to bind Zn^{2+} in vitro (5). In contrast, the conserved, centrally located acidic amino acid residues are almost all conserved in foamy virus IN proteins (11, 12, 24–26, 34, 36). These residues have been shown to be critical for endonuclease and integrase activity in the case of human immunodeficiency virus type 1 (HIV-1) and Rous sarcoma virus IN proteins (11, 12, 25). During integration, a two-step reaction is required. In the processing reaction as the first step, the IN protein introduces nicks and removes two nucleotides from the 3′ end of each viral DNA strand, exposing the conserved CA dinucleotides (4, 23). In the second step, the DNA-joining reaction, the two processed 3′ hydroxyl termini of viral DNA are joined to recessed 5′ phosphate ends of host DNA (10, 22). Since the integration efficiency of foamy viruses is unknown and the majority of the viral DNA after HFV infection of fibroblast cells is an unintegrated form (14, 15, 33), it is of interest to examine whether or not the wild-type HFV IN is biologically active, capable of processing and DNA joining, and to compare its substrate specificity with that of the well-studied integrases of Rous sarcoma virus, MLV, and HIV (4, 10–13, 17, 22–25, 28, 38, 40, 41).

Here we report on the molecular cloning, bacterial expression, purification, and functional characterization of the integrase protein of HFV which in vitro has both endonucleolytic and DNA-joining activities.

MATERIALS AND METHODS

Cells and virus. The propagation of HFV in human embryonic lung (HEL) cells was done as recently described (16).

Bacteria and plasmids. Plasmids were transformed in *Escherichia coli* JM109 for cloning and strain BL21 (DE3) for expression (39). Transformed bacteria were grown in Luria broth medium in the presence of 50 mg of carbenicillin per ml. To construct the recombinant plasmid pET3bIN, pET3b (39) was digested with *Nde*I and plasmid C55, which carries the HFV integrase sequence (16), was digested with *Eco*RV. The 1,222-bp *Eco*RV fragment was ligated first to *Nde*I linkers and thereafter into the *Nde*I-cloning site of the pET3b vector. The orientation of the insertion was determined by restriction analysis and dideoxy sequencing (Fig. 1). Recombinant protein HFVpET3bIN contained two additional amino acid residues, methionine and asparagine, at the amino terminus that were introduced by the cloning procedure.

Purification of HFVpET3bIN protein. To purify denatured protein for immunization of rabbits, BL21 (DE3) pET3bIN

* Corresponding author.

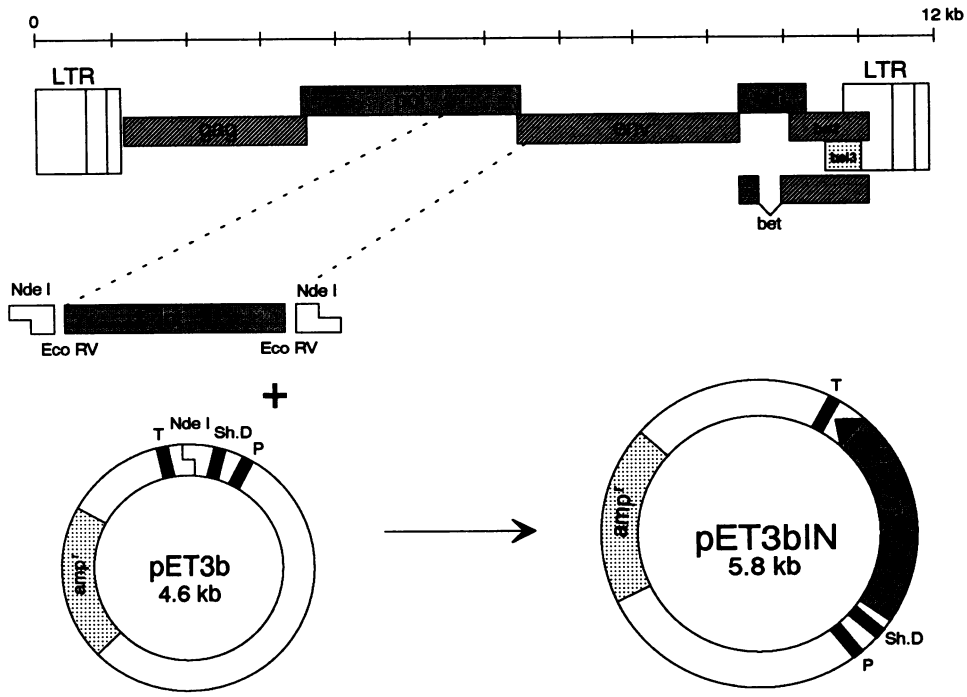

FIG. 1. Molecular cloning of HFV *int*. The upper part shows the HFV proviral map. The IN domain of *pol* was obtained by cloning the 1,222-bp *Eco*RV HFV DNA fragment (plus the *Nde*I linker) into the *Nde*I site of pET3b. Plasmid pET3bIN was used for expression in *E. coli* (for details, see Materials and Methods). P, T7 promoter; Sh.D, Shine-Dalgarno consensus sequence; T, T7 terminator (39). For recombinant plasmid pET16bIN, pET16b was used instead of pET3b.

cells were cultivated and induced as described below under "Labeling of the recombinant HFV IN protein." The bacterial culture was centrifuged, and the pellet was lysed in 50 mM Tris (pH 8.0)–200 mM NaCl–2 mM EDTA–1 mM phenylmethylsulfonyl fluoride (PMSF)–10 mM benzamidine–0.2 mg of lysozyme per ml–1% Triton X-100 and then sonicated on ice until the viscosity decreased. The suspension was layered on top of a 40% sucrose cushion and centrifuged at 12,000 × *g* for 30 min. The pellet was dissolved in 8 M urea–0.5 M NaCl–0.5 M Tris (pH 8.0)–1 mM EDTA–1 mM dithiothreitol–1 mM PMSF–10 mM benzamidine and centrifuged at 12,000 × *g* for 10 min. The supernatant that contained more than 50% recombinant HFV IN protein was electrophoresed on a 12% polyacrylamide gel in the presence of 0.1% sodium dodecyl sulfate (SDS) (27). The band containing the IN protein was excised and eluted from the gel with a biotrap as described by the supplier (Schleicher & Schuell). The gel and biotrap procedure was repeated to yield a highly purified recombinant HFV IN protein which was used to immunize rabbits.

Labeling of the recombinant HFV IN protein. BL21 (DE3) pET3bIN bacteria from an overnight culture were diluted 1:50 in Luria broth medium (supplemented with 100 mg of ampicillin per ml) and cultivated at 37°C. At an optical density value at 600 nm of 1.0, isopropyl-β-D-thiogalactoside (IPTG) was added to a final concentration of 1 mM. Rifampin (1 μg/ml) and L-[^{35}S]methionine (1,000 Ci/mmol, 10 mCi/ml) were added (1 μl/ml of culture) 90 min later. The culture was grown for another 2.5 h. Bacterial lysates were analyzed by SDS-polyacrylamide gel electrophoresis (PAGE) (12% polyacrylamide gel) and visualized by exposure to an X-ray film for 2 days.

Construction, growth, and purification of HFVpET16bIN protein. To obtain enzymatically active IN protein, plasmid pET16b (19) (Novagen) was digested with *Nde*I and *Bam*HI. The DNA fragment containing the HFV IN was obtained from pHSRV clone C55 (16) by polymerase chain reaction, using the Pfu DNA polymerase (Stratagene) that has been reported to be of higher fidelity than the *Taq* DNA polymerase (31). The sequences of the antisense and the sense primers were 5′-CAGTATAATTGGATCCTTCTG-3′ and 5′-CAGGGTCATCATATGAAAGGA-3′, respectively (HFV genome positions 5418 to 5438 and 6592 to 6612). Polymerase chain reaction was done at 92°C for 1 min, 52°C for 45 s, and 72°C for 90 s; the cycle was repeated 35 times. The resulting polymerase chain reaction fragments were digested with *Nde*I and *Bam*HI and ligated into the corresponding cloning sites of pET16b. The correct orientation of the insert was confirmed by dideoxy sequencing of both ends of the vector-HFV insertion borders. Plasmid pET16bIN was used to express HFV IN in *E. coli* BL21 (DE3). To purify enzymatically active HFV IN protein in greater amounts, 1 liter of BL21 (DE3) pET16bIN cells was grown at 37°C in Luria broth medium (supplemented with 50 mg of carbenicillin per ml) to an optical density value at 600 nm of 0.7 to 0.8. To induce IN protein expression, IPTG was added to a final concentration of 1 mM; bacteria were grown for another 4 h and harvested by low-speed centrifugation. The pellet was resuspended in 40 ml of 20 mM Tris (pH 7.6)–1 M NaCl–10% glycerol–1 mM PMSF–5 mM imidazole (IMAC5). Cells were lysed with a Dounce homogenizer, sonicated for 3 min, and centrifuged at 39,000 × *g* and 4°C for 20 min to remove cell debris. The supernatant was loaded on a column containing Ni^{2+}-charged His Bind TMResin (Novagen). The flow rate was 10 ml/h. The resin was washed with IMAC200 (the same as IMAC5 except with 200 mM imidazole), and the

TABLE 1. Efficiency of endonucleolytic activity of the HFV IN protein[a]

Substrate	Origin	% Uncleaved	% Cleaved	Activity (%)
5'ATTGTCATGGAATTTTGTAT 3'TAACAGTACCTTAAAACATA* ↑	HFV U5 wt LTR	38	62	100
5'ATTGTGGTGGAATGCCACTA 3'TAACACCACCTTACGGTGAT* ↑	HFV U3 wt LTR	67	33	53
5'ATTG***TG***GTGGAATTTTGTAT 3'TAAC***AC***CACCTTAAAACATA*	HFV U5 56 LTR	76	24	40
5'ATTGTCATGGAATTT***ACA***AT 3'TAACAGTACCTTAAA***TGT***TA*	HFV U5 1678 LTR	35	65	104
5'ATTGTGGTGGAATG***CTGT***TA 3'TAACACCACCTTAC***CACA***AT*	HFV U3 1678 LTR	73	27	43
5'AT***AC***TCATGGAATTTTGTAT 3'TA***TG***AGTACCTTAAAACATA*	HFV U5 34 LTR	93	7	11
5' TGTGGTGGAATGCCACTAGA 3' ACACCACCTTACGGTGATCT*	HFV U3 Δ12 LTR	100	0	0
5'ACTG***CTAGAGA***TTT***CCACA*** 3'T***GACGATCT***CTA***AAAAGGTGT*** 5'***CGGTGACGTTCCGGGGTGCG*** 3'***GCCACTGCAAGGCCCCACGC***	HIV-1 U5 wt LTR	100	ND	ND
	HFV U5 random LTR	100	0	0

[a] The activity is expressed as the percentage of cleaved oligonucleotides relative to the amounts cleaved in the case of HFV U5 wild-type LTR substrate, which was arbitrarily set to 100%. The yields of oligonucleotides were determined by optical densitometry. Vertical arrows indicate the cleavage sites of wild-type HFV LTR termini. Mutated nucleotides different from the HFV wild-type LTR sequences are printed in boldface italics. The asterisk marks the ^{32}P label. Assays were done in the presence of Mn^{2+} ions under standard assay conditions; for details, see Materials and Methods. ND, not detectable; wt, wild type.

specifically bound protein was eluted with IMAC500. The eluate was dialyzed against 25% glycerol–50 mM Tris (pH 7.6)–20 mM β-mercaptoethanol–0.1 mM EDTA–1 mM PMSF and stored at −70°C or, alternatively, further purified by cation-exchange chromatography on a Mono S column (Pharmacia). To elute the IN protein, a linear salt gradient from 0.2 to 1.0 M NaCl in HEPES (*N*-2-hydroxyethylpiperazine-*N*'-2-ethanesulfonic acid) buffer at pH 7.6 (10 mM β-mercaptoethanol) was employed. The HFV IN protein was recovered from a peak that was eluted at 600 mM NaCl.

Immunodetection, Western immunoblotting, and PAGE. SDS-PAGE (27) was used in monitoring HFV IN protein expression in BL21 (DE3) cells, in purification by Ni^{2+} chelate chromatography, and in detecting IN protein in infected HEL and COS-7 cells. Rabbit polyclonal antibodies raised against purified recombinant HFV IN protein were used for Western blotting. The immunoreactive proteins were detected by alkaline phosphatase-conjugated protein A and subsequently stained with diaminobenzidine. Determination of protein concentration was done according to Bradford, with bovine serum albumin as the standard (3). Marker proteins were prestained standards from GIBCO-BRL.

Synthesis and purification of oligodeoxynucleotides. Oligodeoxynucleotides were synthesized on a Millipore Cyclon Plus apparatus, gel purified, and eluted from the gel slices as described previously (37). Oligodeoxynucleotide concentrations were spectrophotometrically determined at an optical density of 260 nm.

Endonucleolytic cleavage assay (23). A gel-purified synthetic 20-bp oligodeoxynucleotide corresponding to the 3' end of the U5 region of the plus strand of the HFV long terminal repeat (LTR) (see V2, Fig. 5A) was labeled at the 5' terminus with T4 polynucleotide kinase and [γ-^{32}P]ATP (6,000 Ci/mmol). The labeled strand was annealed with a fourfold excess of a gel-purified synthetic 20-mer oligodeoxynucleotide corresponding to the 3' end of the minus-strand U5 region of the HFV LTR (see V1, Fig. 5A). Subsequently, 1.3 pmol of V1-V2 was used in a reaction mixture containing 20 mM MOPS [3-(*N*-morpholino)propane sulfonic acid], 85 mM KCl, 10 mM dithiothreitol, 25% glycerol, 100 μg of bovine serum albumin per ml, 6 mM $MgCl_2$ or 6 mM $MnCl_2$, and 370 ng of HFV IN in a total volume of 10 μl (pH 7.2). The mixture was incubated for 90 min at 37°C and stopped by addition of 10 μl of dye-containing formamide and heated to 95°C for 5 min. Samples were loaded (4 μl each) on a 12% denaturing polyacrylamide gel and electrophoresed in 89 mM Tris–89 mM boric acid–2 mM EDTA. Reaction products were visualized by autoradiography for 1 or 2 h. The same standard assay conditions were used for analysis of different substrates except that instead of HFV U5 LTR, the substrates listed in Table 1 were employed. The intensities of the bands on the autoradiogram were quantitated by an optical densitometric scanner (Gilford Response).

Integration assays (10, 23). To assay for DNA-joining activity, an 18-bp oligonucleotide corresponding to the 3' end of the HFV U5 plus strand but lacking two nucleotides at the 3' terminus (see V3, Fig. 5B) was labeled at the 5' end and annealed to a fourfold excess of V1. The reaction conditions used were identical to those of the endonucleolytic cleavage assay except that 13.6 pmol of annealed V1-V3 was employed as a substrate. The time of exposure was 15 h.

Disintegration assays. The disintegration assay was done under the same conditions as the endonuclease assay except that 13.6 pmol of a Y-shaped oligonucleotide (see Fig. 5C) (according to Chow et al. [8]) which was shortened by 2 bp at the 5' end of T1 was used as substrate. To generate the Y-shaped oligomer, 13.6 pmol of the T1 target oligomer was labeled with ^{32}P at its 5' end. Subsequently, 54.4 pmol of V1, V4, and T2 (see Fig. 5) was added, heated to 80°C, and slowly cooled to 30°C. V4 was a 33-mer and is derived from V3 (see Fig. 5B) plus a target sequence (8). The Y-shaped oligomer was precipitated and used for disintegration assays.

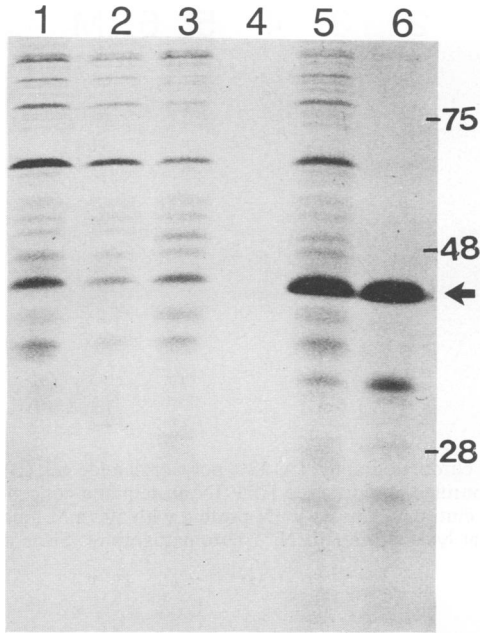

FIG. 2. Autoradiogram of [^{35}S]methionine-labeled recombinant HFV IN protein expressed in *E. coli* BL21 (DE3) cells. Expression was induced with 1 mM IPTG. Lanes: 1, uninduced, no plasmid; 2, induced, no plasmid; 3, uninduced, with pET3b; 4, induced, with pET3b; 5, uninduced, with pET3bIN; 6, induced, with pET3bIN. The arrow marks the 41-kDa HFV IN protein. Molecular sizes (in kilodaltons) of marker proteins are shown in the right margin.

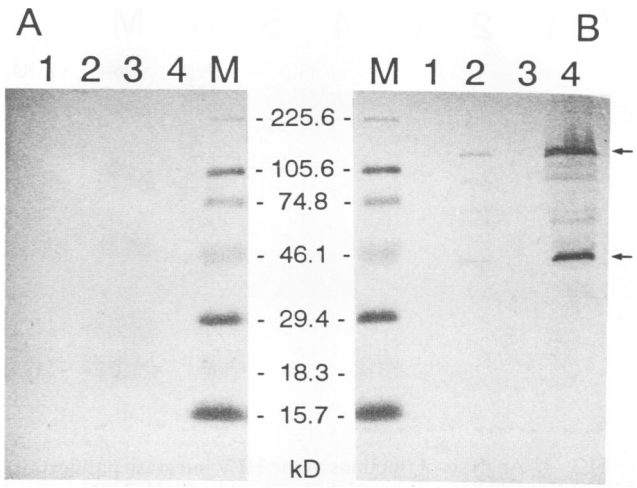

FIG. 3. Immunodetection of wild-type HFV IN protein in infected HEL and transfected COS cells with rabbit anti-HFV IN antibodies. Shown are Western blots incubated with preimmune serum (A) and with HFV IN antiserum (B). Extracts are from mock-infected HEL cells (lane 1), HFV-infected HEL cells (lane 2), mock-transfected COS cells (lane 3), and HFV-transfected COS cells (lane 4). The upper arrow marks the *pol* precursor protein; the lower arrow marks the HFV IN protein. The calculated molecular mass of the HFV IN protein is 38 kDa. Sizes of protein markers (lanes M) are given in the center.

RESULTS

Cloning and expression of the HFV IN protein. To obtain sufficient amounts of HFV IN protein by bacterial expression, the HFV *int* gene was cloned into the pET3b vector (Fig. 1). Protein sequence comparisons between known retroviral integrases showed that most, if not all, encode a metal finger domain that starts approximately 10 amino acid residues downstream of the amino terminus of HFV IN (12, 16, 22, 24). To ascertain that the recombinant HFV IN protein included all amino acid residues necessary for the different enzymatic activities of the integrase, the NH_2 terminus was set 37 residues upstream of the HHCC metal finger motif. Retroviral integrases are synthesized by proteolytic cleavage of the *gag-pol* precursor and lack their own initiation codons. For bacterial expression of HFV IN, an ATG codon was introduced upstream of the IN region. It was expected that termination occurred at the authentic stop codon of the wild-type HFV *pol* reading frame. The resulting recombinant plasmid shown in Fig. 1 was designated pET3bIN.

To confirm the expression of cloned HFV IN protein by plasmid pET3bIN, pET3bIN-transformed *E. coli* cells were labeled with [^{35}S]methionine and induced with IPTG in the presence of rifampin so that transcription occurs only from the T7 promoter by T7 RNA polymerase. Figure 2 shows that a single protein is synthesized in large amounts with a molecular mass of 41 kDa in close agreement with the calculated value of 41.6 kDa.

Immunodetection of the HFV IN protein in HFV-infected cells. The IN protein was purified to homogeneity by repeated SDS-PAGE and subsequent electroelution. Rabbit antiserum against HFV IN was prepared and used to detect the wild-type HFV IN protein in two different cell types, in HFV-infected HEL cells and in COS-7 cells transfected with the infectious pHSRV13 DNA clone (30). Figure 3 shows the immunoblot of HFV-infected HEL cells and COS cells transfected with proviral DNA. The immunodetection of the IN protein was weaker in HFV-infected HEL cells than in transfected COS cells. The apparent molecular masses of the detected immunoreactive proteins were 39, 60, 62, 81, and approximately 120 kDa. The 39- and the 120-kDa proteins were stained more intensely than the other HFV gene products. It is likely that the mature HFV IN protein corresponds to the 39-kDa band and that the *pol* precursor corresponds to the 120-kDa band as reported recently (35). There was no detectable immunoreaction of the preimmune serum with any proteins of either mock-infected HEL cells, HFV-infected HEL cells, or mock-transfected or transfected COS cells. The hyperimmune serum did not react with any bands in mock-infected HEL cells or mock-transfected COS cells (Fig. 3).

Purification of enzymatically active HFV IN protein. To show that recombinant HFV IN protein is enzymatically active, purification by affinity chromatography was used. To this end, HFV IN from pET3bIN was recloned into the pET16b vector that codes for a stretch of 10 amino-terminal histidine residues that selectively bind to a Ni^{2+} chelate resin (19). The recombinant HFV IN protein was purified by high-salt extraction from bacterial lysates followed by Ni^{2+} affinity chromatography as described in Materials and Methods. The single-step purification resulted in milligram quantities of electrophoretically homogeneous and soluble HFV IN protein under nondenaturing conditions (Fig. 4A). Western blotting of the same gel with the antiserum against HFV IN protein revealed that the purified protein was indeed the HFV IN protein (Fig. 4B).

In vitro endonucleolytic cleavage reactions. Purified HFVpET16bIN protein was assayed for endonuclease, inte-

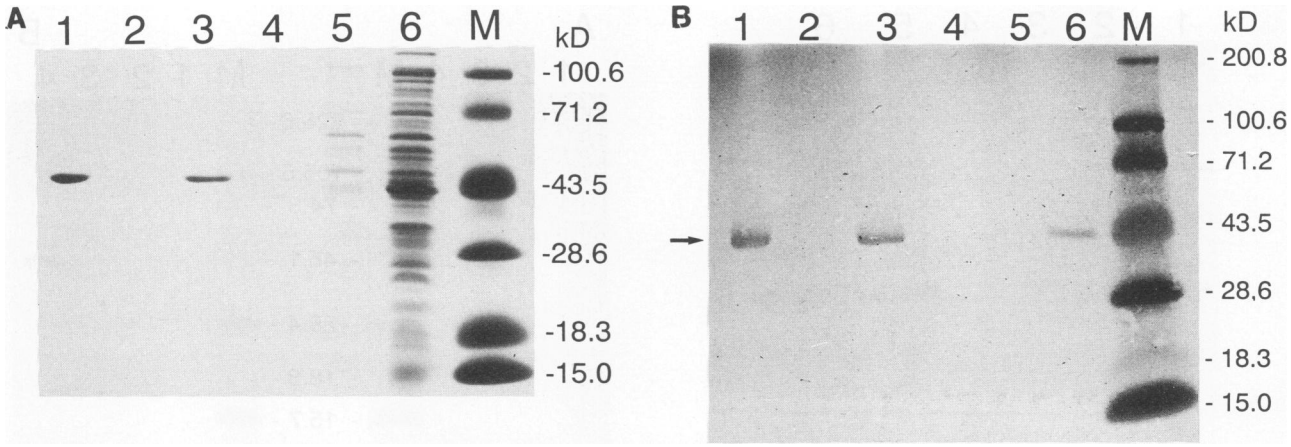

FIG. 4. Analysis of fractions of the HFV integrase purification by Ni^{2+} affinity chromatography. (A) 12% polyacrylamide gel; (B) Western blot of the gel shown in panel A, incubated with HFV IN antibodies. Lanes: 1, purified and dialyzed HFV IN protein at a concentration of 70 μg/ml; 2, elution of bound proteins from the column with 100 mM EDTA; 3, elution of the HFV IN protein with 500 mM imidazole; 4, washing with 200 mM imidazole; 5, flowthrough of soluble proteins; 6, bacterial lysates prior to Ni^{2+} chromatography. Sizes of marker proteins (lanes M) are given in the right margins.

grase, and disintegration activity. HFV endonuclease is expected to specifically remove two nucleotides (T and A) from the 3′ ends of the DNA plus strand of the U5 region and the minus strand of the U3 region as predicted from the HFV sequence. To assay for endonucleolytic cleavages, a ^{32}P-end-labeled plus strand of the 20-bp duplex oligonucleotide representing the 3′ end of the U5 LTR sequence was used as a substrate (Fig. 5A). The products of the endonuclease reaction were electrophoretically separated. When the HFV integrase was incubated in the presence of Mn^{2+}, two nucleotides were in fact specifically removed from the 3′ end of the plus strand. The reaction resulted in an 18-mer as the predominant product (Fig. 6A). When the DNA cleavage reaction was carried out with HFV U3 LTR (Table 1) instead of HFV U5 LTR as a substrate under the same conditions, it was found that the 18-mer product was also formed specifically but with lower efficiency (Fig. 6B; Table 1). In the presence of Mg^{2+} ions and with HFV U5 LTR as a substrate, a range of degradation products of various sizes was found, indicating that a contaminating nuclease might be responsible for the unspecific degradation observed (Fig. 7A), in contrast to the products obtained in the presence of Mn^{2+}. A similar result was obtained with HFV U3 LTR (not shown). To ascertain that the unspecific pattern of degradation is not due to a contaminating nuclease that happens to copurify during Ni^{2+} affinity chromatography, the HFV IN protein was further purified by cation-exchange chromatography. The peak fraction of HFV IN protein was recovered and again assayed for endonuclease activity in the presence of either Mn^{2+} or Mg^{2+}. The results of the endonuclealytic cleavage reaction with HFV IN purified by Ni^{2+} affinity and subsequent cation-exchange chromatography shown in Fig. 7B clearly indicate that the pattern of unspecific degradation that had been observed with Mg^{2+} (Fig. 7A) did not occur. Instead, the HFV LTR substrate was not cleaved at all (Fig. 7B). This indicates that the unspecific degradation by the His-tagged HFV enzyme in the presence of Mg^{2+} was very likely due to traces of a contaminating nuclease which was effectively removed by the second purification step of cation-exchange chromatography. The preference of the HFV IN for Mn^{2+} ions is reminiscent of the known cation specificity of foamy reverse transcriptase (29).

Site and substrate specificity of the HFV endonuclease activity. To determine optimal reaction conditions for endonuclease activity, DNA cleavage assays with increasing ratios of HFV IN protein to HFV U5 LTR substrate were carried out at a reaction time of 90 min. Figure 8 shows the result of an analysis with 1.3 pmol of HFV U5 LTR and increasing concentrations of the IN protein while other parameters including pH, concentrations of Mn^{2+}, buffer, and salt were kept invariant. It was found that HFV IN protein concentrations of 370 ng (ratio of 1 pmol of substrate

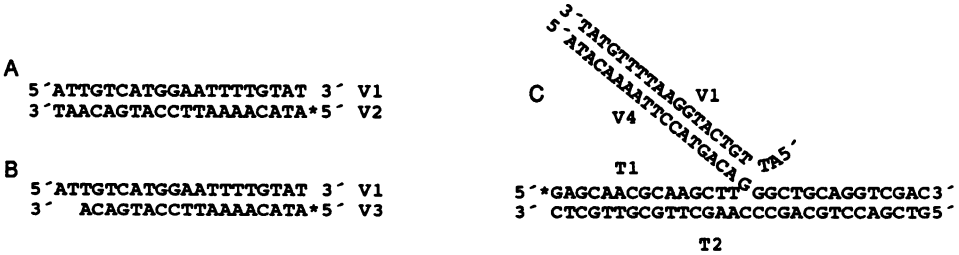

FIG. 5. Oligodeoxynucleotides used as substrates for the endonuclease (A), integration (B), and disintegration (C) assays. The asterisks mark the radioactive ^{32}P label of the oligonucleotides. The substrates in panels A and B are given by the 3′ end of the HFV U5 LTR region. The target sequences were taken from reference 8 and shortened by 2 bp. V4 is a 33-mer that consists of V3 plus a target oligomer sequence. For details, see text.

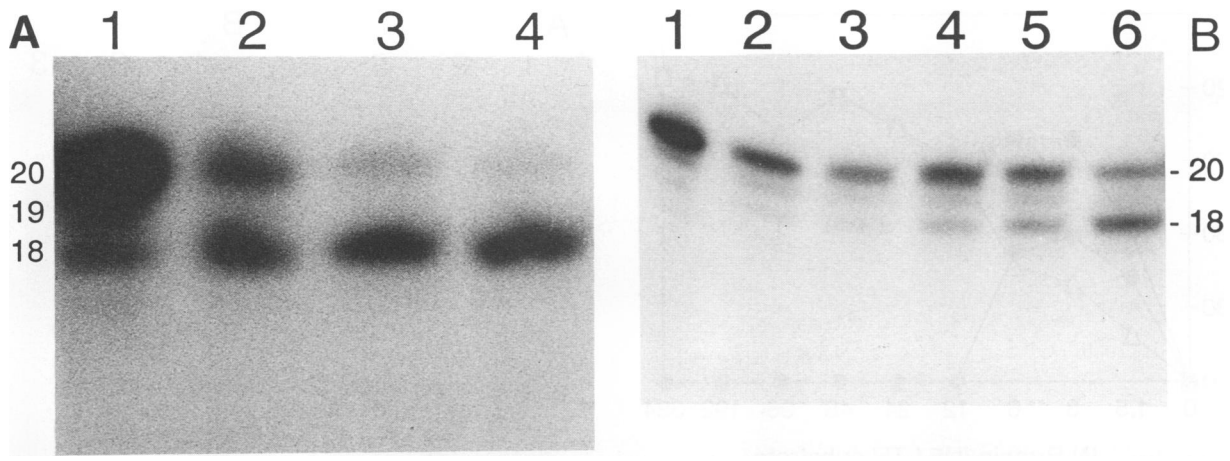

FIG. 6. Kinetics of the endonuclease assay. Assays were done as described in Materials and Methods in the presence of Mn^{2+}. In panel A, HFV U5 LTR was used as substrate. Aliquots were taken after 5 min (lane 1), 20 min (lane 2), 60 min (lane 3), and 90 min (lane 4). In panel B, HFV U3 LTR was used, and aliquots were taken after 0 min (lane 1), 5 min (lane 2), 10 min (lane 3), 20 min (lane 4), 40 min (lane 5), and 90 min (lane 6); reactions were stopped by formamide followed by heating to 95°C and separation on a 12% polyacrylamide gel. Numbers in the margins indicate lengths of products in nucleotides.

to 8 pmol of HFV IN) resulted in a site-specific cleavage of two nucleotides (Fig. 6). At lower IN concentrations, the apparent site specificity which is due to a contaminating nuclease activity was lost (see preceding paragraph); instead, reaction products of various lengths were observed, and at higher IN concentrations site-specific cleavage was observed (Fig. 8). During a kinetic analysis, the rate of DNA cleavage activity began to plateau after 90 min at 37°C (Fig. 6). Consequently, we chose to use 370 ng of HFV IN protein, 1.3 pmol of substrate, and 90 min as standards to assay for DNA cleavage activity of HFV IN. To gain insight into the specificity of the HFV enzyme, different oligodeoxynucleotide substrates were assayed (Table 1). With the wild-type oligonucleotide (HFV U5 LTR), the yield of the reaction products with the recessed ends after a 90-min reaction time was taken as 100% of HFV endonuclease activity.

The results described so far indicate that HFV IN is capable of cleaving both the U5 and the U3 ends, which have 2 of 11 nucleotide mismatches. It was expected that the HFV IN protein might tolerate some variability in substrate sequence. To determine the relative importance of specific nucleotides in the wild-type HFV sequence, several base-pair substitutions were introduced into the oligomer substrate. Mutations of HFV U5 close to the cleavage site (HFV U5 34 LTR in Table 1) decreased the nuclease activity approximately 10-fold. This result is consistent with data from several groups who observed that similar substrate mutations in other retroviral LTRs led to greatly diminished DNA cleavage activity (6, 23). Substitutions of nucleotides

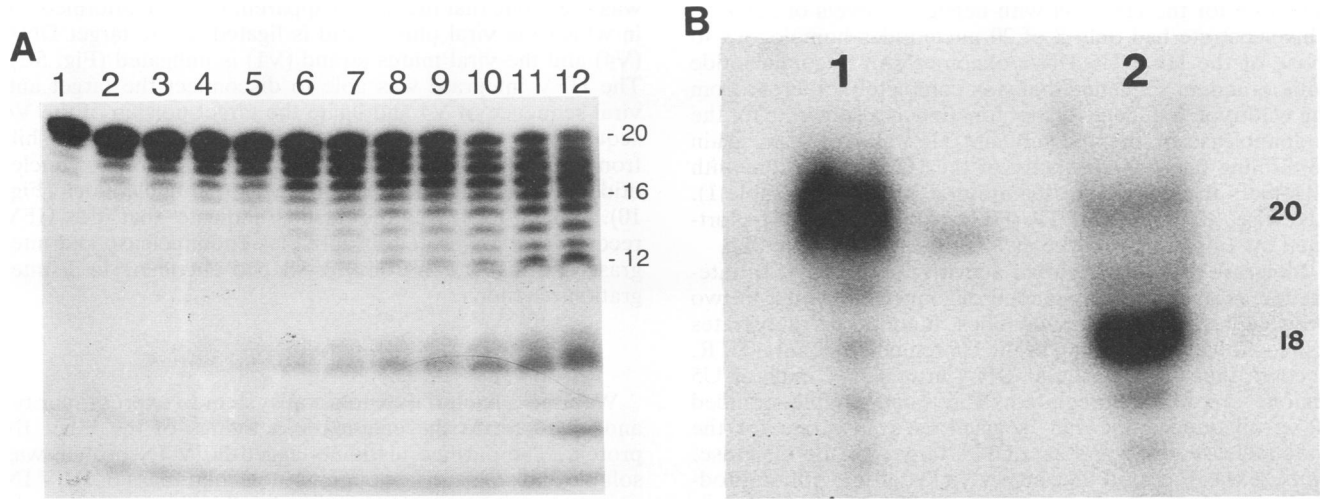

FIG. 7. (A) Kinetics of the endonuclease assay in the presence of Mg^{2+}. Aliquots of the reaction products were taken after 0.5, 5, 10, 15, 20, 30, 40, 50, 60, 80, 100, and 120 min (lanes 1 to 12) and processed as described in the legend to Fig. 6. (B) Endonucleolytic cleavages were carried out with the HFV IN highly purified by both Ni^{2+} affinity and cation-exchange chromatography (see Materials and Methods) in the presence of Mg^{2+} (lane 1) and Mn^{2+} (lane 2). Minor oligonucleotide spots of sizes smaller than 20 bp (lane 1) and 18 bp (lane 2) are due to self-decay. Numbers in the right margins mark lengths in nucleotides.

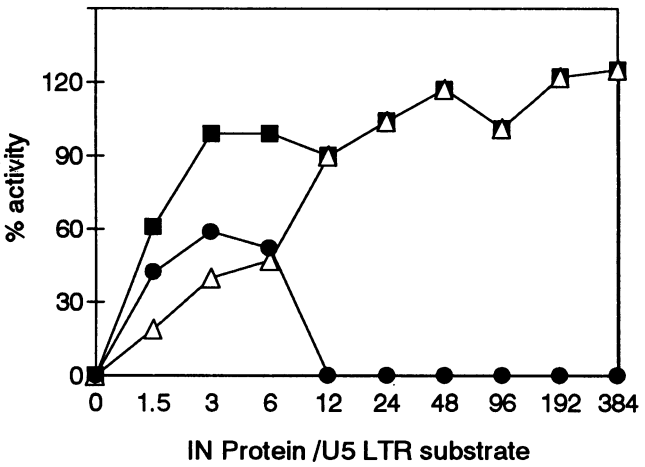

FIG. 8. DNA cleavage activity of increasing concentrations of HFV IN protein plotted as the molar ratio of HFV IN to HVF U5 LTR substrate against the concentration of reaction products relative to that of HFV U5 LTR set to 100% (Table 1). Concentrations were quantitated with a Gilson Response scanner. Symbols present total amounts of products cleaved (■), the 18-mer oligonucleotide (▲), and unspecific reaction products (●).

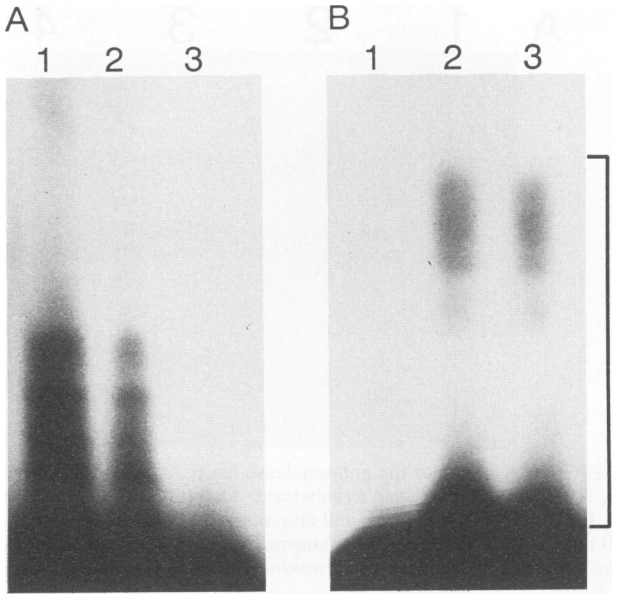

FIG. 9. Autoradiogram of the reaction products of the HFV integrase. The bracket marks the region of integration products. (A) Integrase reaction in the presence of Mn^{2+} (lanes 1 and 2); double the amount was loaded in gel pocket 1 compared with lane 2. Lane 3 shows the control reaction containing buffer and labeled substrate, omitting HFV IN protein. (B) Integrase reaction under the same conditions as in panel A with Mg^{2+} replacing Mn^{2+} ions (lanes 2 and 3); double the amount of reaction products was loaded into pocket 2 compared with pocket 3. Lane 1 shows the integrase reaction without HFV IN.

at positions 4 and 5 (that match the corresponding nucleotides of HFV U3 LTR) upstream of the cleavage site (HFV U5 67 LTR in Table 1) yielded a twofold reduction in activity. To determine the influence of nucleotides further upstream of the cleavage site, HFV U5 1678 LTR (Table 1), which has three substitutions at positions 16, 17, and 18 from the cleavage site, was used as a substrate. An activity similar (104%) to that with wild-type U5 ends was detected. However, when a correspondingly mutated HFV U3 substrate (HFV U3 1678 LTR) was used, only 43% activity was recovered, compared with 53% for the wild-type U3 sequence (Table 1). Thus, the value of 43% actually corresponds to 82% of the wild-type HFV U3 sequence.

It is of interest that an oligonucleotide corresponding to the wild-type sequence of HIV-1 U5 LTR did not serve as a substrate for the HFV IN with negligible levels of activity. This substrate had only 8 of 20 nucleotides homologous to those of the HFV U5 DNA oligomer. An oligonucleotide with a random sequence that was completely different from the wild-type sequence did not function as a substrate for the endonuclease of the recombinant HFV IN protein, again confirming the site specificity of the HFV IN protein with oligomers that mimic the foamy viral DNA ends (Table 1). Likewise, an HFV U3 LTR (HFV U3 Δ12, Table 1) shortened by 2 bp did not serve as a substrate.

Integrase and disintegration activities of HFV IN. In integration assays, double-stranded oligonucleotides act in two ways. First, the oligonucleotides function as substrates representing the 3′ termini of the U5 region of the viral LTR. Second, they mimic cellular DNA in which 3′ ends of U5 regions are to be integrated. The same double-stranded 20-bp oligonucleotide can be used as a substrate for the endonuclease assay and as a DNA target for the integrase. Since it was reported that higher yields of integration products were obtained when the product with 3′ recessed ends was used as a substrate in the integration assay (11), we used the corresponding staggered HFV U5 substrate (Fig. 5B) for the integration reaction. The resulting products of the integrase reactions were longer than the starting substrate and encompass a range of oligodeoxynucleotides of different lengths that, integrated into themselves, are indicated by brackets in Fig. 9. It is noteworthy that in contrast to the endonucleolytic cleavage reaction, the integration reaction did not show a preference for either of the divalent cations used.

In the disintegration (or reversal of integration), a Y-shaped oligodeoxynucleotide (8) of four different strands was used, one that mimics an apparent reaction intermediate in which the viral plus strand is ligated to the target DNA (V4) and the viral minus strand (V1) is unligated (Fig. 5C). The HFV integrase was able to disconnect the target and viral sequence of V4 and ligate the target portion of the V4 sequence to the T1 oligomer. This resulted in a band shift from 15 nucleotides of the labeled starting oligodeoxynucleotide T1 to a single labeled product of 30 nucleotides (Fig. 10). Taken together, the results indicate that the HFV recombinant IN protein is active as endonuclease and integrase and that it can utilize Y-shaped oligomers in disintegration reactions.

DISCUSSION

We used a bacterial expression system to express, purify, and characterize the enzymatic activities of the HFV IN protein. The purified histidine-tagged HFV IN protein was soluble and stable. The apparent molecular mass of HFV IN protein in wild-type HFV-infected cells was approximately 39 kDa, consistent with that predicted from the IN protein sequence and with that reported recently (35). Thus, the HFV IN is substantially larger than the HIV-1 IN protein, which was reported to have a molecular mass of 32 kDa (11,

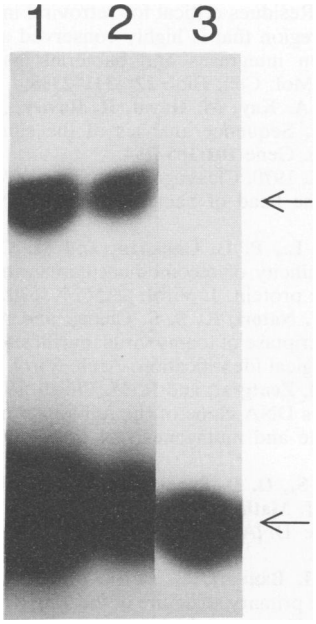

FIG. 10. Autoradiogram of disintegration reaction products of the HFV IN protein in the presence of Mn^{2+}. The lower arrow marks the substrate (15-mer), and the upper arrow marks the reaction product (30-mer). Lanes 1 and 2, disintegration assay; lane 3, control assay omitting HFV IN protein.

40, 41). The apparent molecular masses of both recombinant HFV IN proteins (41 and 44 kDa) reported here are distinctly higher than that of wild-type HFV. At least two factors account for the different values. The proteolytic cleavage site that generates the amino terminus of native HFV IN is still unknown, and the site chosen for our constructs is probably upstream of the authentic processing site. Alternatively and/or additionally, there might be processing at the carboxy terminus of HFV IN protein as reported for other retroviral IN proteins (9). Secondly, the tag of 10 histidine residues and the additional amino acids introduced at the amino terminus by the cloning procedure used contribute to the slightly higher value of 44 kDa for the HFVpET16bIN protein.

It is interesting that the spumavirus enzyme is enzymatically active in the endonuclease, integration, and disintegration reaction when expressed in *E. coli* as recombinant protein even with additional flanking residues. The DNA cleavage activity reached maximal values when the ratio of substrate HFV U5 LTR to HFV IN protein was approximately 1:8, indicating that the foamy viral enzyme can probably function as an oligomeric protein. At substrate/HFV IN ratios below 1:8, the contaminating nuclease is obviously active, indicating that at higher HFV IN concentrations, the LTR substrate seems to be protected by the IN protein against unspecific degradation.

By using the His-tagged purified IN protein, the site-specific activity of the enzyme for the model substrates was analyzed. It was observed that two nucleotides were specifically removed from the HFV U5 substrate duplex oligodeoxynucleotide. This result is not unexpected, since it was also reported previously for the endonuclease of Rous sarcoma virus, MLV, and HIV-1 (7, 10, 13, 23). It is remarkable that the HFV DNA cleavage activity is highly specific for Mn^{2+}. The preference for Mn^{2+} required for specific cleavage by the HFV IN protein is characteristic and pronounced. In the presence of Mn^{2+}, the HFV endonuclease seems to stop after two nucleotides are released from the substrate. Of note, when HFV IN protein that had been purified by both Ni^{2+} and cation-exchange chromatography was used, no endonuclease activity of the HFV IN was detectable in the presence of Mg^{2+}. The 5′ end of the HFV U3 LTR functioned as a substrate with about 50% of the efficiency of the HFV U5 LTR under the same conditions. The reason for this is unclear, but it is interesting to compare it with the data reported for HIV-1 IN cleavage of its cognate HIV-1 LTR ends. It was recently reported that the U3 ends of HIV-1 DNA are better substrates than the U5 termini (6, 11, 13, 38). The difference in the substrate functions of the two HFV ends probably reflects the difference in sequence. The sequence of the HFV U3 end encodes part of the *bel3* protein and therefore may be constrained between two competing necessities for both integration and coding requirements. Thus, HFV U3 coding for *bel3* in that respect is comparable to HIV-1 U3 that encodes *nef* (6). This might well be a more general feature of the U3 ends of complex retroviruses that have additional requirements in the U3 regions in contrast to those of less complex retroviruses. When the HFV DNA U5 substrate was mutated, we found that nucleotides at positions close to the cleavage site are critical for the activity of the HFV endonuclease. Mutations at positions 1 and 2 in the HFV U5 LTR still served as substrates (11% activity), whereas mutations at positions 4 and 5 upstream of the cleavage site had about 40% of the activity of wild-type DNA ends. Mutations even further upstream of the cleavage site retained full functional capacity to serve as substrate. In addition, it was found that the HIV-1 U5 LTR ends barely served as substrates and that random sequences did not function as substrates. These results are consistent with data reported for other retroviral IN proteins (4, 6, 11, 23, 40, 41).

ACKNOWLEDGMENTS

We thank Jürgen Kleinschmidt for advice during cation-exchange chromatography, Hans-Georg Kräusslich for critically reading the manuscript, and H. zur Hausen for support.

This work was supported by grants from the Bundesministerium für Forschung und Technologie (01KI 8901/6) and the Commission of the European Community (TS2-CT187-0186-D).

REFERENCES

1. **Achong, B. G., P. W. A. Mansell, M. A. Epstein, and P. Clifford.** 1971. An unusual virus in cultures from a human nasopharyngeal carcinoma. J. Natl. Cancer Inst. **46**:299–307.
2. **Bowerman, B., P. O. Brown, J. M. Bishop, and H. E. Varmus.** 1989. A nucleoprotein complex mediates the integration of retroviral DNA. Genes Dev. **3**:469–478.
3. **Bradford, M. M.** 1976. A rapid and sensitive method for quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. **72**:248–254.
4. **Brown, P. O.** 1991. Integration of retroviral DNA. Curr. Top. Microbiol. Immunol. **157**:19–48.
5. **Burke, C. J., G. Sanyal, M. W. Bruner, J. A. Ryan, R. L. LaFemina, H. L. Robbins, A. S. Zeft, R. Middaugh, and M. Cordingley.** 1992. Structural implications of spectroscopic characterization of a putative zinc finger peptide from HIV-1 integrase. J. Biol. Chem. **267**:9636–9644.
6. **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.
7. **Bushman, F. D., T. Fujiwara, and R. Craigie.** 1990. Retroviral DNA integration directed by HIV integration protein in vitro. Science **249**:1555–1558.

8. **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.
9. **Coffin, J. M.** 1990. Replication of retroviridae, p. 1437–1500. *In* B. N. Fields (ed.), Virology. Raven Press, New York.
10. **Craigie, R., T. Fujiwara, and F. Bushman.** 1990. The IN protein of Moloney murine leukemia virus processes the viral DNA ends and accomplishes their integration in vitro. Cell **62:**629–637.
11. **Drelich, M., R. Wilhelm, and J. Mous.** 1992. Identification of amino acid residues critical for endonucleases and integration activities of HIV-1 IN protein in vitro. Virology **188:**459–468.
12. **Engelman, A., and R. Craigie.** 1992. Identification of conserved amino acid residues critical for human immunodeficiency virus type 1 integrase function in vitro. J. Virol. **66:**6361–6369.
13. **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.
14. **Flügel, R. M.** 1991. Spumaviruses: a group of complex retroviruses. J. Acquired Immune Defic. Syndr. **4:**739–750.
15. **Flügel, R. M.** 1992. The molecular biology of human spumavirus, p. 193–214. *In* B. C. Cullen (ed.), Human retroviruses—frontiers in molecular biology. Oxford University Press, Oxford.
16. **Flügel, R. M., A. Rethwilm, B. Maurer, and G. Darai.** 1987. Nucleotide sequence analysis of the *env* gene and its flanking regions of the human spumaretrovirus reveals two novel genes. EMBO J. **6:**2077–2084.
17. **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.
18. **Grandgenett, D. P., and S. R. Mumm.** 1990. Unraveling retrovirus integration. Cell **60:**3–4.
19. **Hoffmann, A., and R. G. Roeder.** 1991. Purification of his-tagged proteins in non-denaturing conditions suggests a convenient method for protein interaction studies. Nucleic Acids Res. **19:**6337–6338.
20. **Johnson, M. S., M. A. McClure, D.-F. Feng, J. Gray, and R. F. Doolittle.** 1986. Compure analysis of retroviral *pol* genes: assignment of enzymatic functions to specific sequences and homologies with nonviral enzymes. Proc. Natl. Acad. Sci. USA **83:**7648–7651.
21. **Jonsson, C. B., G. A. Donzella, and M. J. Roth.** 1993. Characterization of the forward and reverse integration reactions of the Moloney murine leukemia virus integrase protein purified from *E. coli*. J. Biol. Chem. **268:**1462–1469.
22. **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.
23. **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.
24. **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.
25. **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.
26. **Kupiec, J.-J., A. Kay, M. Hayat, R. Ravier, J. Peries, and F. Galibert.** 1991. Sequence analysis of the simian foamy virus type 1 genome. Gene **101:**185–194.
27. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227:**680–685.
28. **LaFemina, R. L., P. L. Callahan, and M. Cordingley.** 1991. Substrate specificity of recombinant human immunodeficiency virus integrase protein. J. Virol. **65:**5624–5630.
29. **Liu, W. T., T. Natori, K. S. S. Chang, and A. M. Wu.** 1977. Reverse transcriptase of foamy virus: purification of the enzyme and immunological identification. Arch. Virol. **55:**1987–2000.
30. **Löchelt, M., H. Zentgraf, and R. M. Flügel.** 1991. Construction of an infectious DNA clone of the full-length human spumaretrovirus genome and mutagenesis of the *bel 1* gene. Virology **184:**43–54.
31. **Lundberg, K. S., D. D. Shoemaker, M. W. W. Adams, J. M. Short, and E. J. Mathur.** 1991. High-fidelity amplification using a thermostable DNA polymerase from *Pyrococcus furiosus*. Gene **108:**1–6.
32. **Maurer, B., H. Bannert, G. Darai, and R. M. Flügel.** 1988. Analysis of the primary structure of the long terminal repeat and the *gag* and *pol* genes of the human spumaretrovirus. J. Virol. **62:**1590–1597.
33. **Mergia, A., and P. A. Luciw.** 1991. Replication and regulation of primate foamy viruses. Virology **184:**475–482.
34. **Mergia, A., K. E. S. Shaw, J. E. Lackner, and P. A. Luciw.** 1990. Relationship of the *env* gene and the endonuclease domain of the *pol* genes of simian foamy type 1 and human foamy virus. J. Virol. **64:**406–410.
35. **Netzer, K.-O., A. Schliephake, B. Maurer, R. Watanabe, A. Aguzzi, and A. Rethwilm.** 1993. Identification of *pol*-related gene products of human foamy virus. Virology **192:**336–338.
36. **Renne, R., E. Friedl, M. Schweizer, U. Fleps, R. Turek, and D. Neumann-Haefelin.** 1992. Genomic organization and expression of simian foamy virus type 3 (SFV-3). Virology **186:**597–608.
37. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed., p. 11.23–11.40. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
38. **Sherman, P. A., M. L. Dickson, and J. A. Fyfe.** 1992. Human immunodeficiency virus type 1 integration protein: DNA sequence requirements for cleaving and joining reactions. J. Virol. **66:**3593–3601.
39. **Studier, F. W., A. H. Rosenberg, J. J. Dunn, and J. W. Dubendorff.** 1990. Use of RNA polymerase to direct expression of cloned genes. Methods Enzymol. **185:**60–89.
40. **Vincent, K. A., V. Ellison, S. A. Cow, 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.
41. **Vink, C., D. C. V. 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.