Journal of General Virology (2001), **82**, 139–148. Printed in Great Britain

Characterization of chimeric enzymes between caprine arthritis–encephalitis virus, maedi–visna virus and human immunodeficiency virus type 1 integrases expressed in *Escherichia coli*

Nicola Berger, Astrid E. Heller, Klaus D. Störmann and Eberhard Pfaff

Federal Research Centre for Virus Diseases of Animals, Institute for Immunology, Paul-Ehrlich-Strasse 28, D-72076 Tübingen, Germany

In order to investigate the functions of the three putative lentiviral integrase (IN) protein domains on viral DNA specificity and target site selection, enzymatically active chimeric enzymes were constructed using the three wild-type IN proteins of caprine arthritis–encephalitis virus (CAEV), maedi–visna virus (MVV) and human immunodeficiency virus type 1 (HIV-1). The chimeric enzymes were expressed in *Escherichia coli*, purified by affinity chromatography and analysed *in vitro* for IN-specific endonuclease and integration activities on various DNA substrates. Of the 21 purified chimeric IN proteins constructed, 20 showed distinct site-specific cleavage activity with at least one substrate and six were able to catalyse an efficient integration reaction. Analysis of the chimeric IN proteins revealed that the central domain together with the C terminus determines the activity and substrate specificity of the enzyme. The N terminus appears to have no considerable influence. Furthermore, an efficient integration activity of CAEV wild-type IN was successfully demonstrated after detailed characterization of the reaction conditions that support optimal enzyme activities of CAEV IN. Also, under the same *in vitro* assay conditions, MVV and HIV-1 IN proteins exhibited endonuclease and integration activities, an indispensable prerequisite of domain-swapping experiments. Thus, the following report presents a detailed characterization of the activities of CAEV IN *in vitro* as well as the analysis of functional chimeric lentiviral IN proteins.

Introduction

Integration of the viral DNA into the host cell DNA is an essential step in retrovirus replication (LaFemina *et al.*, 1992; Sakei *et al.*, 1993). As shown for human immunodeficiency virus type 1 (HIV-1), this integration reaction is a prerequisite for a productive infection (Ansari-Lari *et al.*, 1995; Engelman *et al.*, 1995; Englund *et al.*, 1995; Wiskerchen & Muesing, 1995; Taddeo *et al.*, 1996). *In vitro*, the only enzyme required for catalysing all steps of integration is the viral integrase (IN) protein (Katzman *et al.*, 1989; Craigie *et al.*, 1990; Katz *et al.*, 1990). Following reverse transcription to convert the genomic retroviral RNA into a double-stranded blunt-ended viral DNA, the IN protein cleaves both 3′-terminal viral DNA ends after a highly conserved CA dinucleotide and removes the terminal two nucleotides. This site-specific cleavage reaction, also referred to as endonucleolytic cleavage, endonuclease reaction or 3′ processing (Katzman *et al.*, 1989; Craigie *et al.*, 1990; Bushman & Craigie, 1991), generates recessed 3′-OH groups at both ends of the linear viral DNA. This is assumed to occur *in vivo* in the preintegration complex before entry into the nucleus, where the subsequent insertion reaction takes place. In a one-step transesterification reaction, IN catalyses the nucleophilic attack of the 3′-OH groups of the processed viral DNA to break phosphodiester bonds in the host DNA and simultaneously joins the viral and host DNA together. This concerted reaction is referred to as strand transfer, integration or DNA joining (Bushman *et al.*, 1990; Craigie *et al.*, 1990; Bushman & Craigie, 1991; Engelman *et al.*, 1991). Both the specific cleavage and strand transfer reactions can be investigated *in vitro* using double-stranded oligonucleotides representing the U_3 and U_5 ends of the viral long terminal repeats (LTRs) (Katzman *et al.*, 1989).

Recombinant IN proteins of various retroviruses have been studied *in vitro* (Katzman *et al.*, 1989; Bushman *et al.*, 1990; Craigie *et al.*, 1990; Drelich *et al.*, 1992; Van Gent *et al.*, 1992; Pahl & Flügel, 1993; Störmann *et al.*, 1995; Shibagaki *et al.*, 1997). Comparison of the amino acid sequences, com-

Author for correspondence: Eberhard Pfaff.
Fax +49 7071 967 303. e-mail eberhard.pfaff@tue.bfav.de

plementation experiments (Engelman *et al.*, 1993; Van Gent *et al.*, 1993) and mutation analysis of HIV-1 IN, the most intensively analysed IN, revealed three different domains of the enzyme. The N-terminal region possesses a HHCC zinc finger motif that is highly conserved within retroviral IN proteins and retrotransposons (Johnson *et al.*, 1986; Khan *et al.*, 1991; Burke *et al.*, 1992; Bushman *et al.*, 1993). It may therefore function in the recognition and interaction of the viral DNA ends, as is known for other DNA-recognizing and DNA-binding proteins (e.g. transcription factors) (Khan *et al.*, 1991; Van Gent *et al.*, 1992; Vincent *et al.*, 1993; Vink *et al.*, 1993). The central region of the enzyme is thought to be the catalytic domain, which was identified between amino acid residues 50 and 194 of HIV-1 IN (Vink & Plasterk, 1993). Mutations of any of three conserved amino acids (Asp64, Asp116 and Glu152) forming the essential DD(35)E-motif completely abolish all enzyme activities (Engelman & Craigie, 1992; Van Gent *et al.*, 1992). The C-terminal domain is the least conserved region of the enzyme. It is thought to be involved in DNA binding, but the specific DNA recognition site has not been clearly identified to date. Similar affinities for binding specific as well as unspecific DNA are reported (Engelman *et al.*, 1994; Puras-Lutzke *et al.*, 1994). In order to adapt the identified enzyme domains with the assumed one-step reaction mechanism of IN (Engelman *et al.*, 1991), models were developed whereby IN is active as either an oligomer or a multimer (Asante-Appiah & Skalka, 1999; Esposito & Craigie, 1999). Complementation experiments (Engelman *et al.*, 1993; Van Gent *et al.*, 1993) and X-ray structure analysis of the central domain (Dyda *et al.*, 1994; Wlodawer, 1999) strengthen this one-step assumption. The central domain and the C terminus of the enzyme seem to be essential for the formation of oligomers (Engelman *et al.*, 1993; Kalpana & Goff, 1993; Van Gent *et al.*, 1993; Barsov *et al.*, 1996; Jenkins *et al.*, 1996).

Domain-swapping experiments between wild-type enzymes are useful to define and carefully characterize functional enzyme domains (Yagil *et al.*, 1995; Shibagaki *et al.*, 1997; Katzman & Sudol, 1995, 1998; Dildine *et al.*, 1998; Tasara *et al.*, 1999). The prerequisites for analysis of such chimeric enzymes are that the wild-type enzymes must show enzyme-specific activities under identical reaction conditions and that these results are well distinguishable in order to recognize the influence of the domains derived from the different wild-type enzymes.

We have reported the successful expression, purification and analysis of wild-type IN proteins of maedi–visna virus (MVV) German strain 461, caprine arthritis–encephalitis virus (CAEV) and HIV-1 (Störmann *et al.*, 1995). These three IN proteins exhibited the full repertoire of *in vitro* activity characteristic of retroviral IN proteins and showed basic similarities in their endonuclease and integration activities, as well as distinct differences in respect to substrate specificity and substrate turnover under identical reaction conditions. These findings meet the aforementioned requirements of domain-swapping experiments and therefore encouraged us to construct chimeras between these three wild-type IN proteins. Chimeric enzymes were expressed in *Escherichia coli* and purified by Ni^{2+}-affinity chromatography. Their enzymatic activities were analysed *in vitro* on different viral DNA sequences in order to obtain more information about the functions of the different domains of lentiviral IN proteins, especially with regard to the location and specificity of the virus and host DNA-binding sites.

Methods

■ **PCR mutagenesis and cloning of chimeric IN proteins.** The IN sequences of CAEV strain 75-G63, MVV German strain 461 and HIV-1 strain NL4-3 isolates were cloned into the pQE60 expression vector as previously described (Störmann *et al.*, 1995). The enzyme domain borders were determined for HIV-1 IN at amino acid positions 50 and 194 (Vink & Plasterk, 1993) and were identified by alignment analysis for MVV and CAEV IN proteins at positions 50 and 196. For the construction of chimeric IN proteins, we introduced unique *Bam*HI (at aa 50) and *Avr*II (at aa 194 or 196) restriction sites into wild-type IN sequences using overlapping PCR mutagenesis (Horton *et al.*, 1993). PCR was performed, utilizing a Bio-Med 60 thermal cycler, with 2·5 U *Taq* DNA polymerase (Stratagene) as previously described (Störmann *et al.*, 1995). Amplification conditions used for PCR were 30 cycles of melting (2 min at 94 °C), annealing (1 min at 60 °C) and extension (2 min at 72 °C). Oligonucleotide primers used for amplification were synthesized on a Millipore Expedite DNA synthesizer.

To introduce the *Bam*HI restriction site, PCR was performed with the primer pairs PQE1P/IN1M and IN1P/PQE1M and wild-type IN sequences cloned into the pQE60 vector. In an overlapping PCR extension reaction, both PCR products together with primers PQE1P and PQE1M were used to generate the entire mutated IN1 sequence (INmut1). In a second overlapping PCR mutagenesis, the *Avr*II restriction site was introduced using the purified INmut1 PCR product as the DNA template and the primer pairs PQE1P/IN2M and IN2P/PQE1M. Amplification with the primers PQE1P and PQE1M and both PCR products as templates resulted in mutated wild-type IN sequences containing the unique *Bam*HI and *Avr*II restriction sites between the enzyme domains. After purification, digestion with *Nco*I and *Hin*dIII and cloning into the pQE60 vector, the three domains of the various wild-type IN sequences were reciprocally exchanged using the restriction enzymes *Nco*I/*Bam*HI, *Bam*HI/*Avr*II or *Avr*II/*Hin*dIII in order to generate 24 different chimeric IN sequences. Restriction enzymes and T4 DNA ligase were obtained from New England Biolabs and were used according to the manufacturer's specifications followed by standard cloning procedures (Sambrook *et al.*, 1989). The amplified and cloned sequences were verified by sequencing with a T7 sequencing kit (Pharmacia) according to the dideoxy sequencing method of Sanger *et al.* (1977).

Primers used for PCR amplification were PQE1P, 5′ AATTCATTAAAGAGGAGAAATTAA**CCATGG** 3′, and PQE1M, 5′ AGCTAATT**AAGCTT**AGTGATGGTGATGGTG 3′, which include the *Nco*I and *Hin*dIII restrictions sites indicated in bold, respectively. Wild-type IN-specific primers were CIN1P, 5′ AGAAAGGACACCTGCCG**GGATCC**GAGGAGGAAACAAAAGA 3′, and CIN2P, 5′ ATAAAAAGAAAGGGTGG**CCTAGG**GACAAGCCCTATGGATA 3′, for CAEV IN; MIN1P, 5′ AAATAAAGCGCCTAGT**GGGATCC**GGGGAAGTAATAAAAGG 3′, and MIN2P, 5′ ATAAAAAGAAAGGGTGG**CCTAGG**GACAAGCCCTATGGACA 3′, for MVV IN; and HIN1P, 5′ TCAGCTAAAAGGGGAAG**GGATCC**ATGGACAAGTAGACTGT 3′, and HIN2P, 5′ TTTAAAAGAAAAGGGGG**CCTAGG**GGGGTA-

CAGTGCAGGGG 3′, for HIV-1 IN. Restriction sites (*Bam*HI and *Avr*II) that were introduced are indicated in bold and nucleotides that were mutated in order to create those unique restriction sites are underlined. The sequences of the IN1M and IN2M reverse primers were complementary to the respectively listed IN1P and IN2P primers.

■ **Expression and purification of chimeric IN proteins.** Expression and purification of IN proteins were performed as described by Drelich *et al*. (1992) and Störmann *et al*. (1995).

■ **Assay for endonucleolytic cleavage and integration activities.** IN activities were tested *in vitro* using radiolabelled double-stranded 20-mer oligonucleotides as DNA substrates. The sequences of these oligonucleotides correspond to the outer U_3 and U_5 LTR regions of CAEV 75-G63, MVV 461 and HIV-1 NL4-3 viral DNA as described by Störmann *et al*. (1995). The 5′ termini of the plus strands were labelled using T4 polynucleotide kinase (Biolabs) and [γ-^{32}P]ATP (5000 Ci/mmol, ICN), purified using the PCR purification kit (Qiagen) and annealed with their complementary strands in 10 mM Tris–HCl, pH 7·6 and 150 mM NaCl. In a standard IN activity assay, 1 pmol of DNA substrate was incubated with 10 pmol (0·4 μg) of purified IN in 10 μl reaction buffer (25 mM Tris–HCl, pH 8·0 and 1 mM DTT) supplemented with 2 mM $MnCl_2$. All components, including IN and DNA substrate, were added as tenfold concentrated stock solutions. Therefore, the final concentration of NaCl in the reaction was 65 mM. To prevent the formation of oxidation products, $MnCl_2$ was prepared and stored as a separate 20 mM $MnCl_2$ stock solution (pH 4–5). After an incubation time of 90 min at 37 °C, the reaction was stopped by adding 10 μl of dye-containing formamide (95 % formamide, 20 mM EDTA, 0·05 % bromophenol blue and 0·05 % xylene cyanol). Reaction products of 2 μl aliquots were heated for 5 min to 95 °C and analysed on 15 % denaturing polyacrylamide gels. Wet gels were autoradiographed at −70 °C. Products of the cleavage activity were detected after 3–10 h whereas integration products appeared after an extended radiographic exposure of 3–5 days. Furthermore, the efficiency of the cleavage reaction was quantified by scanning the gel with a Bio Imaging Analyser System (Fujifilm). For determination of the substrate turnover, the signals were analysed by the TINA 2.0 software (Raytest).

Results

Construction, expression and purification of MVV 461, CAEV and HIV-1 chimeric IN proteins

The three different lentiviral IN cDNA sequences had previously been cloned into the *E. coli* pQE60 expression vector allowing IN expression as C-terminally tagged hexahistidine fusion proteins (Störmann *et al*., 1995). By alignment analysis of the amino acid sequences of HIV-1 IN functional domains (Vink & Plasterk, 1993), the postulated functional domains for the IN proteins of CAEV and MVV could also be identified. To construct the chimeric IN enzymes, we introduced unique restriction sites by PCR mutagenesis between the postulated functional IN domains at amino acids 50 and 194 for HIV-1 IN and at amino acids 50 and 196 for MVV and CAEV IN. The N terminus, the core region and the histidine-tagged C terminus of wild-type IN proteins were reciprocally exchanged and combined to produce 24 different chimeric IN proteins. These IN proteins are later referred to as CCC, MMM and HHH where the three letters represent the three domains of the enzymes. For example, MCH is the chimeric IN with the N-terminal domain of MVV IN, the central domain of CAEV IN and the C terminus of HIV-1 IN.

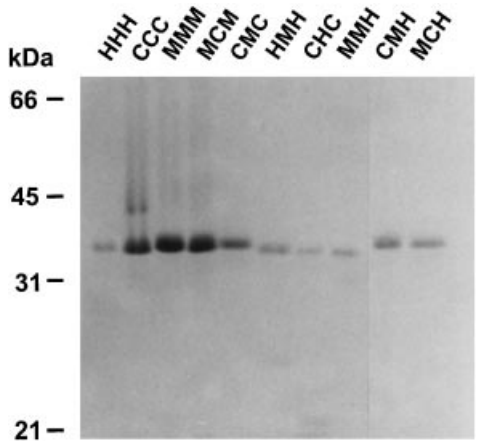


Fig. 1. Gel analysis of purified chimeric IN proteins. Each chimeric IN (4 μl), indicated above the lanes, was separated by SDS–PAGE after purification. The gel was stained with Coomassie brilliant blue. Differences in migration depend on the particular combinations of the wild-type domains and their variable lengths.

After induced expression and purification by affinity chromatography under non-denaturing conditions, each IN protein was analysed on a 12·5 % SDS–PAGE gel (Laemmli, 1970). Elution profiles, protein yield concentrations and purity varied depending on the chimeric construct. Fig. 1 shows the analysis of some of the purified chimeric IN proteins. A total of 21 chimeric IN proteins were obtained with a purity greater than 90 %. The remaining three, HHM, HMC and HCM, could be expressed, but not purified by Ni^{2+}-affinity chromatography.

Characterization and optimization of the specific cleavage and integration reaction of wild-type CAEV IN

Both IN-specific reactions, specific cleavage and strand transfer, can be investigated *in vitro* using radiolabelled double-stranded oligonucleotides that mimic the viral DNA termini. The substrates were cleaved at the conserved CA dinucleotide by the specific IN endonuclease reaction and shortened by the removal of two nucleotides at the 3′ end (see Fig. 6*a*). In the subsequent strand transfer reaction, as outlined in Fig. 5(*a*), the IN protein joins the processed 3′-OH ends to another oligonucleotide molecule that serves as target DNA. The reaction products can be visualized by denaturing gel electrophoresis followed by radiographic exposure.

In previous studies, we had shown the enzyme activities of CAEV, MVV and HIV-1 IN *in vitro* using oligonucleotides that mimicked the virus LTR U_3 and U_5 DNA ends of CAEV 75-G63, MVV 461 and HIV-1 NL4-3 (Störmann *et al*., 1995). Using identical reaction conditions, the three wild-type IN proteins had been found to be active in site-specific cleavage with their authentic virus substrates but had demonstrated distinct differences on heterologous substrates. Concerning the

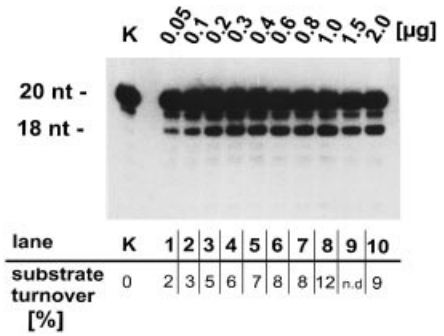


Fig. 2. Endonuclease reaction of CAEV IN against enzyme concentration. To investigate the influence of enzyme concentration on efficiency of cleavage, different concentrations of CAEV IN were incubated under standard conditions with 1 pmol of the CAEV U_3 substrate. Incubation of the CAEV U_3 substrate without IN (K) is the negative control. The autoradiography shows the gel analysis of the reaction products. Numbers on the left indicate the length and position of the substrate (20 nt) and the expected cleavage products (18 nt). The determined substrate turnovers are listed as rounded values in the grid below. n.d, Not done.

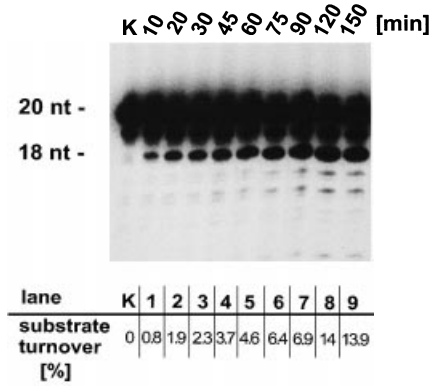


Fig. 3. Time-course of the endonuclease reaction of CAEV IN. Cleavage activity of CAEV IN on the CAEV U_3 substrate under standard reaction conditions was stopped by adding formamide at different time-points. The reaction products were analysed as described. The negative control (K) is the reaction without IN incubated for 150 min. The grid below shows the substrate turnover depending on the reaction time. Positions of the substrate and the reaction products on the autoradiography are indicated on the left.

strand transfer reaction, only MVV and HIV-1 IN had been able to integrate the cleavage products into an acceptor DNA, whereas CAEV IN had not exhibited any detectable integration activity in these studies (Störmann *et al.*, 1995). However, analysis of the DNA joining activity of the chimeric IN proteins also required the activity of the CAEV wild-type enzyme. Therefore, we investigated the reaction conditions that support the optimal activities of CAEV IN. Since CAEV IN has not been analysed in detail, we characterized the endonuclease and strand transfer reaction of CAEV IN on its virus U_3 substrate and optimized assay parameters such as enzyme concentration, reaction time and influence of divalent cations.

Analysis of CAEV IN concentration (Fig. 2) in the *in vitro* assay revealed a linear slope of cleavage activity of up to

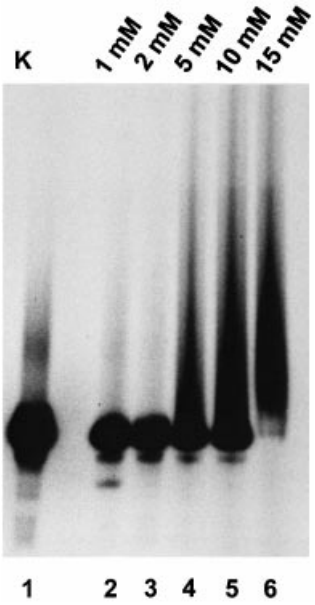


Fig. 4. Influence of Mn^{2+} ions on substrate oligonucleotides. The CAEV U_3 substrate was incubated for 90 min at 37 °C in standard reaction buffer without IN but with different concentrations of $MnCl_2$. Incubation of the CAEV U_3 substrate without $MnCl_2$ (K) acts as the negative control. The reactions were analysed on a denaturing polyacrylamide gel as described for IN assay reaction products.

0·6 µg enzyme per pmol of oligonucleotide substrate. At higher enzyme concentrations, the cleavage reaction seemed to reach saturation (Fig. 2, lanes 6–10). Integration products could be detected from a concentration of 0·3 µg/pmol and higher (data not shown). In order to obtain both enzyme-specific activities and to avoid saturation effects, we subsequently used an enzyme concentration of 0·4 µg/pmol substrate for CAEV IN. The time-course experiment shown in Fig. 3 demonstrated a time-dependent increase of substrate cleavage, which reached its maximum from 120 to 150 min. Longer radiographic exposure of the gel displayed strand transfer products after a reaction time of 45 min (data not shown). The incubation time for standard reactions was then set to 90 min. We also addressed the concentration of divalent cations on the reaction conditions of CAEV IN. Using Mn^{2+} instead of Mg^{2+} ions in the assay results in an increased endonuclease activity of CAEV, MVV and HIV-1 IN (Störmann *et al.*, 1995). When analysing different concentrations of Mn^{2+} ions, we observed oligonucleotide substrate complexes, which appeared as a marked smear when separating the DNA substrates on a denaturing gel (Fig. 4). Mn^{2+} ions are stable only in acidic solutions. In a neutral and basic environment $Mn(OH)_2$, which is very sensitive to oxidation, is formed. The oxidation process results in brown-coloured, insoluble Mn(III)- and Mn(IV)-oxides (Hollemann & Wiberg, 1984). An atomspectroscopic analysis carried out with a tenfold stock solution (50 mM $MnCl_2$, 250 mM Tris–HCl, pH 8·0 and 10 mM DTT) indicated a reduction of Mn^{2+} ions in the reaction mixture to 30% after one freeze–thaw cycle and incubation of the onefold solution

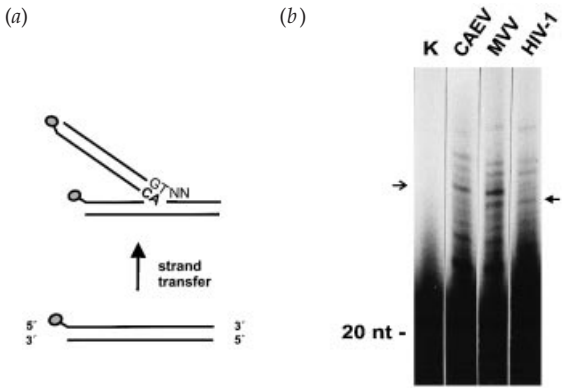


Fig. 5. (*a*) Schematic diagram of the integration reaction. Radiolabelled 5′ ends of the double-stranded oligonucleotides are shown as circles. The conserved CA dinucleotide is in bold. The site-specific cleavage product generated by the endonuclease activity (see Fig. 6*a*) is integrated into another substrate molecule. (*b*) Strand transfer activities of the wild-type IN protein from MVV, CAEV and HIV-1 on the CAEV U_3 substrate. Reactions were performed under standard reaction conditions and analysed as described. Integration products appear as a ladder of bands larger than the initial 20-mer substrate. Differences of integration patterns indicate different target site selections of IN on the same substrate. The arrow on the right points out a common integration product of MVV and HIV-1 IN that is missed in the reaction of CAEV IN. The arrow on the left indicates a preferred target site of CAEV and MVV IN that is missed in the reaction of HIV-1 IN.

for 90 min at 37 °C. DTT used as a reducing agent was unable to completely prevent the oxidation process. In subsequent reactions, we therefore used 2 mM $MnCl_2$, where CAEV IN exhibits high endonuclease and integration activities and a Mn precipitate could not be detected (Fig. 4, lane 3).

Applying these optimized reaction parameters, 0·4 mg (~ 10 pmol) enzyme was incubated with 1 pmol DNA substrate in the reaction buffer supplemented with 2 mM $MnCl_2$ for 90 min at 37 °C. Under these conditions, CAEV IN exhibits a very efficient cleavage activity on both authentic U_3 and U_5 DNA substrates and, furthermore, is able to catalyse an efficient strand transfer reaction (Fig. 5*b*).

Activities of MVV and HIV-1 IN under the optimized assay conditions

In order to check whether the CAEV IN assay parameters were also valid for MVV and HIV-1 IN, we tested their activities on all six DNA substrates. In brief, analysing HIV-1 IN revealed no changes in substrate specificity and enzyme activities as previously described (Bushman & Craigie, 1991; Drelich *et al.*, 1992; LaFemina *et al.*, 1991; Sherman *et al.*, 1992; Störmann *et al.*, 1995). Enzyme activity was higher on the HIV-1 U_5 substrate than on the U_3 substrate. However, under the applied reaction conditions, the cleavage activities of HIV-1 IN on all six substrates were weaker when compared to those of CAEV and MVV IN. An integration activity could be observed with the HIV-1 U_5 and the CAEV U_3 substrate. MVV and CAEV IN cleaved all six substrates with high efficiencies, but the highest activity occurred on their authentic U_3 and U_5 substrates and, surprisingly, on the HIV-1 U_5 substrate. Both MVV and CAEV wild-type IN proteins were able to catalyse strand transfer reactions with all U_3 substrates with a distinct preference for the CAEV U_3 substrate. Fig. 5 shows the integration activities of CAEV and MVV IN as well as of HIV-1 IN with the CAEV U_3 substrate. The different patterns of integration products indicate divergent target site selection of wild-type IN proteins on the same substrate.

Specific endonuclease activities of chimeric IN on different viral DNA substrates

In order to determine the function of the IN domains with respect to viral DNA specificity, each chimeric IN was tested for its processing activity with all six U_3 and U_5 substrates. Fig. 6(*b*) shows the results of the specific cleavage reaction of chimeric CAEV/MVV IN (CM-IN) on the CAEV U_3 and MVV U_3 substrates. Fig. 6(*b*) also demonstrates distinct differences in the cleavage activities of the chimeric IN proteins. MMC and CMM, for example, exhibit only a very weak endonuclease activity on the CAEV U_3 substrate, whereas MCC and CCM cleave the CAEV U_3 substrate with nearly the same efficiency as CCC. Also, MCM and CMC show distinct cleavage reactions on the CAEV U_3substrate. Comparing the cleavage activities of the chimeric IN proteins on the CAEV U_3 and MVV U_3 substrate, differences in substrate specificities of the enzymes are clearly discernible. MCM, MCC and CCM, for example, cleave the CAEV U_3 substrate more efficiently than the MVV U_3 substrate, whereas CMM, MMC and CMC prefer the MVV U_3 substrate. Since wild-type CAEV and MVV IN showed their highest cleavage efficiencies on their homologous U_3 and U_5substrates, respectively, the substrate specificities of the chimeric CM-IN proteins were analysed by the comparison of their cleavage activities on the CAEV U_3/U_5 and MVV U_3/U_5 substrates. The results of the substrate specificities of the various CM-IN proteins are summarized in Fig. 6(*c*). Comparison of cleavage efficiencies on the U_3 substrate revealed that the central domain of the chimeric CM-IN proteins determines which substrate is preferentially used. Regarding the processing activities on the two U_5 substrates, the chimeric IN proteins containing the central domain from MVV IN (MMC, CMM, CMC) also demonstrated higher activities on the MVV U_5 than on the CAEV U_5 substrate. CCM and MCC cleaved both U_5 substrates with nearly the same efficiency. Only MCM, which was expected to prefer the CAEV U_5 substrate, cleaved MVV U_5 more efficiently.

The cleavage activities of CAEV/HIV-1 (CH-IN) and MVV/HIV-1 (MH-IN) chimeric IN proteins were analysed with regard to the CAEV U_3 and MVV U_3 substrates, respectively. The activities with the two HIV-1 DNA substrates were not evaluated, since HIV-1 U_3 is an unsuitable substrate for all three wild-type IN proteins (Störmann *et al.*, 1995) and the HIV-1 U_5 substrate is cleaved even more efficiently by MVV and CAEV IN than their own authentic U_5

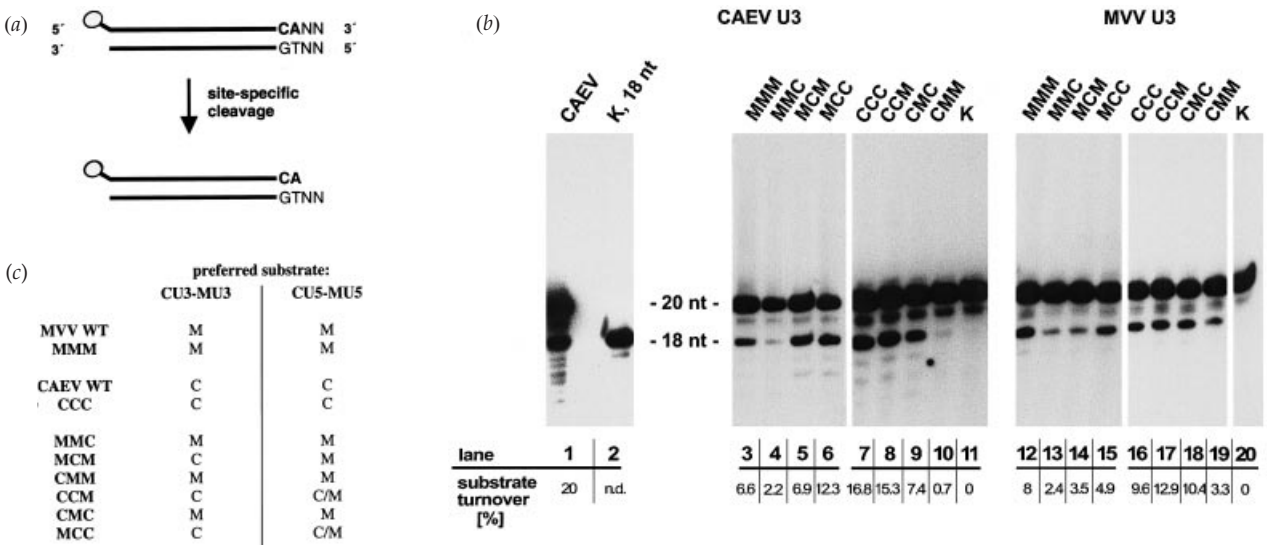


Fig. 6. (*a*) Schematic diagram of the IN cleavage reaction. Radiolabelled 5′ ends of the double-stranded oligonucleotides are shown as circles. The conserved CA dinucleotide is in bold. In a site-specific cleavage reaction, IN removes the two nucleotides downstream of CA. (*b*) Endonuclease activities of the chimeric CM-IN proteins on the CAEV U_3 (lanes 3–10) and MVV U_3 (lane 12–19) substrates. Reactions were performed under standard conditions and analysed on a denaturing gel. Lane 1, cleavage reaction of CAEV IN on the CAEV U_3 substrate; lane 2, double-stranded ^{32}P-labelled 18-mer oligonucleotide incubated in a standard assay without IN; K, CAEV U_3 (lane 11) or MVV U_3 (lane 20) substrates incubated without IN. The length and position of substrates (20 nt) and expected products (18 nt) are indicated. The determined substrate turnovers are listed in the grids below. n.d, Not done. (*c*) Substrate specificities of wild-type (WT) and chimeric CM-IN proteins. The grid shows the cleavage preferences of the listed chimeric IN proteins for CAEV- or MVV-derived substrates. The activities on the CAEV U_3 substrate were compared to those on the MVV U_3 substrate and the activities on the CAEV U_5 substrate to those on the MVV U_5 substrate. M, preference for MVV derived substrates; C, preference for the CAEV derived substrates.

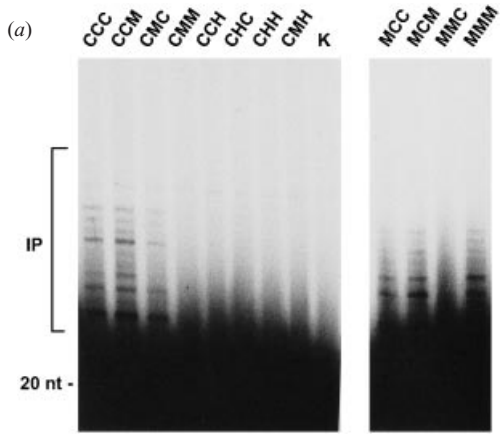
Table 1. Endonuclease activities of wild-type (WT) IN and the chimeric CH- and MH-IN proteins on CAEV and MVV substrates

Site-specific cleavage was assayed under standard conditions. The efficiency of the reaction on the MVV U_3 and CAEV U_3 substrates was quantified as described. Cleavage is scored as efficient (++), distinct (+), weak but detectable (+/−) or no endonuclease activity (−).

Chimeric MH-IN	MVV U_3	Chimeric CH-IN	CAEV U_3
MVV WT	++	CAEV WT	++
MMM	++	CCC	++
HIV-1 WT	+/−	HIV-1 WT	+/−
HHH	+/−	HHH	+/−
HMM	++	HCC	+
MHM	+/−	CHC	−
MMH	−	CCH	++
MHH	+/−	CHH	+/−
HMH	+/−	HCH	+/−

substrates. The results of the analysis of the endonuclease activities of CH- and MH-IN with CAEV and MVV DNA substrates are summarized in Table 1. These results again support the assumption that the central domain of IN has a strong influence on substrate specificity. Whereas the substrate preference of HCC and CCH closely matched that of wild-type CAEV IN and HMM reacted like MVV IN, chimeric IN proteins with the central domain from HIV-1 IN showed no (CHC) or only weak (MHM, MHH, CHH) processing activity on CAEV U_3 or MVV U_3 substrates. In contrast, MMH, which was expected to cleave the MVV U_3 substrate, showed no cleavage activity. The processing activities of HCH and HMH on the MVV U_3 substrate were weak but detectable. HHC was the only chimeric enzyme that seemed to be completely inactive, since no activity could be detected with all six substrates. A distinct influence of the N-terminal domain on the activity of the chimeric IN proteins could not be recognized. However, the C terminus seems to modulate the reaction efficiency. An increase in enzyme activity was found when the C terminus was derived from the same wild-type IN as the central region: HMM and HCC exhibited a very efficient cleavage activity in contrast to MMH or HCH.

Analysis of the chimeric IN proteins consisting of three varied domains, each derived from a different wild-type IN, confirmed the importance of the central IN domain for substrate specificity. Therefore, MCH processed the CAEV U_3 substrate with the highest efficiency of all the tested DNA substrates and CMH showed the best activity on the MVV U_3



(b)

	CAEV U3
MVV WT	+++
MMM	++
MMC	-
MCM	++
MCC	++
MMH	-
MHM	-
MHH	-
MCH	++
MHC	-
CAEV WT	+++
CCC	+++
CCM	+++
CMC	+++
CMM	-
CCH	-
CHC	-
CHH	-
CMH	-
CHM	-
HIV-1 WT	++
HHH	+
HMM	-
HMH	-
HCC	+
HCH	-
HHC	-

Fig. 7. Integration activities of the chimeric IN proteins on the CAEV U_3 substrate. Reactions were performed under standard conditions and analysed as described. (*a*) Gel analysis of the strand transfer reaction of some of the chimeric IN proteins (indicated above the lanes). Integration products appear as a ladder of bands larger than the initial 20-mer substrate (IP). K, incubation of the CAEV U_3 substrate without IN (negative control). (*b*) The integration activity of each tested chimeric IN on the CAEV U_3 substrate is scored as very efficient (+++), good (++), weak (+) and none (−). WT, wild-type IN.

substrate. In contrast, MHC and CHM, with central domains from HIV-1 IN, exhibited very unspecific and weak endonuclease activities.

Strand transfer activities of chimeric IN proteins

The ability of IN to link the processed 3′-OH viral DNA ends covalently with the host DNA can be shown in the same *in vitro* reaction as the specific cleavage activity. The integration products appear as a ladder of bands larger than the initial 20-mer substrate. As described for the endonuclease activity, the substrate specificities of the chimeric IN proteins were tested with all six DNA substrates in order to analyse the influence of the enzyme domains on the integration activity. Only a few chimeric IN proteins were able to catalyse the strand transfer reaction with at least one substrate, which was mainly the CAEV U_3 oligonucleotide. The integration reaction of some chimeric IN proteins are shown in Fig. 7 (*a*). The results of all enzymes tested with the CAEV U_3 substrate are summarized in Fig. 7 (*b*). HCC and MCH were the only chimeric IN proteins with domains from HIV-1 IN showing a strand transfer reaction on the CAEV U_3 substrate. The chimeric MCM, MCC, CCM and CMC IN proteins exhibited integration activities closely matching the activities shown by the respective wild-type IN, whereas the reaction of CMM and MMC showed no strand transfer reactivity. In order to exclude the possibility that a missed or weak endonuclease reaction is the reason for lacking an integration activity, we analysed some chimeric IN proteins exhibiting only weak or no cleavage activities with preprocessed 18-mer oligonucleotides. No integration products were detectable in any of those analysed.

Differences in the integration product patterns as found for the wild-type IN proteins (Fig. 5 *b*) could be detected only in the case of CMC IN. The pattern created by CMC in strand transfer reactions with the CAEV U_3 substrate differs markedly from that created by CCC and CCM (Fig. 7 *a*). This suggests that CMC has a varying target site selection for integration. Unfortunately, similar differences on the integration products of other active chimeric IN proteins could not be observed.

Discussion

The integration of the viral cDNA into the host cell genome is essential for retrovirus replication. All reactions required for this process are catalysed by the viral IN proteins and can be investigated *in vitro* with oligonucleotides that mimic the viral DNA ends (Katzman *et al.*, 1989). Three domains within the lentiviral IN protein could be identified and characterized by complementation and mutation analysis (Engelman *et al.*, 1993; Van Gent *et al.*, 1993; Vink & Plasterk, 1993). Though various models were developed, the details of the molecular integration mechanism are still unknown. In order to evaluate the function of the supposed domains with regard to recognition and specificity of viral and host DNA, we constructed chimeric IN proteins where the three domains of CAEV, MVV and HIV-1 IN were reciprocally exchanged. Chimeric IN proteins were expressed in *E. coli* as C-terminal hexahistidine-tagged proteins, purified by Ni^{2+}-affinity chromatography and analysed for their enzyme activities and substrate specificity on different oligonucleotide substrates.

We were able to express and purify 21 of 24 constructed chimeric IN proteins with a purity higher than 90%. However, the expression levels and purification profiles of the chimeric IN proteins differed greatly. The obtained concentration and purity varied depending on the chimeric construct (Fig. 1). Chimeric IN proteins with domains from CAEV and MVV IN showed an expression and purification profile similar to that of

the wild-type IN enzymes. The expression and yield of the purified chimeric HIV-1 IN proteins were much lower. Differences in purification profiles may be due to modifications of the enzyme conformation caused by the interaction of domains derived from the different wild-type IN proteins. The amino acid sequences of CAEV and MVV IN show an identity of over 70% whereas HIV-1 IN has an identity of only 30% as compared to CAEV and MVV IN (Störmann *et al.*, 1995). Changes in the secondary or tertiary structure of the enzyme could influence solubility and therefore the purification profile. In order to prepare the central domain from HIV-1 IN for X-ray structure analysis, Dyda *et al.* (1994) were able to increase the solubility of that domain by an exchange of only one amino acid. Lower expression levels of chimeric IN proteins with domains of HIV-1 IN may refer to the previously observed lower expression of HIV-1 IN as compared to that of CAEV and MVV IN. This could probably be caused by a less efficient usage of HIV-1 IN codons in *E. coli* (Holler *et al.*, 1993). Furthermore, the yield of chimeric IN proteins with central domains from HIV-1 IN is reduced by the occurrence of an additional N-terminal truncated expression product, which is translated from an internal Shine–Dalgarno sequence but is not purified under the applied conditions (Störmann *et al.*, 1995).

In contrast to previous reports (Störmann *et al.*, 1995), the wild-type IN of CAEV 75-G63 exhibited a distinct strand transfer activity under the applied reaction conditions. We analysed the endonuclease and integration reaction of CAEV IN depending on the enzyme concentration and incubation time and then optimized the assay parameters. Analysis of the IN-specific activities of MVV and HIV-1 IN proteins also revealed distinct endonuclease and integration activities of both wild-type IN proteins on several DNA substrates under identical reaction conditions.

To judge the influence of the three enzyme domains, we investigated the endonuclease and integration activities of chimeric IN proteins on different oligonucleotide substrates. The results of the site-specific cleavage reactions of the chimeric IN proteins revealed that the central domain determines the activity and substrate specificity. The N terminus does not contribute to the reaction specificity. The cleavage reactions of the chimeric CH- and MH-IN proteins indicated that the C-terminal domain of the enzyme may also have an influence on substrate specificity and cleavage efficiency. This influence may become more distinct in the case of CH- and MH-IN proteins than in that of CM-IN proteins because sequence differences between HIV-1 and CAEV/MVV IN are greater than those between CAEV and MVV IN. The chimeric IN between MVV 461 and HIV-1 IN proteins showed the same activities on the MVV U_3 substrate as reported for chimeric enzymes between the MVV Iceland strain 1514 and HIV-1 IN proteins (Katzman & Sudol, 1995).

The dominance of the central domain on enzyme activity observed in this study corresponds to previous studies that determined the core domain as the enzymatically active centre of IN (Engelman & Craigie, 1992; Van Gent *et al.*, 1992). The insignificance of the N terminus to substrate specificity supports the assumption that this domain does not contribute to the specific binding of viral DNA (Khan *et al.*, 1991; Vink *et al.*, 1993). Reports where the core domain and the C terminus are found to be responsible for the recognition of viral DNA ends (Esposito & Craigie, 1998) as well as the analysis of chimeric IN proteins with an extended central domain (Katzman & Sudol, 1998) support our observed influence of the C-terminal domain on substrate specificity and enzyme activity. In recent reports, the catalytic domain of HIV-1 IN is defined between amino acids 50 and 212 (Asante-Appiah & Skalka, 1999; Esposito & Craigie, 1999).

In contrast to distinct endonuclease activities of nearly all constructed chimeric IN proteins, the strand transfer activity appears to be more sensitive. While all but one of the 21 chimeric IN proteins showed a cleavage activity, only six were able to catalyse the strand transfer reaction. Four of those six were chimeric CM-IN proteins. The divergence of the conformation of chimeric enzymes from that of wild-type IN protein, as discussed in context with the various purification profiles, may also have an influence on the integration activity because, predominantly, chimeric CH- and MH-IN proteins showed no integration activity. The integration reaction is a much more complex mechanism than the cleavage reaction. It requires the simultaneous co-ordination of recognition and binding of both viral and substrate DNA and the catalysis of covalent joining. *In vivo*, this combined reaction occurs within the preintegration complex. Therefore, and as described in various reaction models (Engelman *et al.*, 1993; Van Gent *et al.*, 1993; Vincent *et al.*, 1993; Barsov *et al.*, 1996; Katzman & Sudol, 1998), IN may function as a multimer. It is conceivable that modifications of the conformation of chimeric IN can influence the formation of multimers and thereby the integration activity. The identification of the central domain and the C terminus as essential regions for the formation of oligomers (Engelman *et al.*, 1993; Kalpana & Goff, 1993; Van Gent *et al.*, 1993; Barsov *et al.*, 1996; Jenkins *et al.*, 1996) support our finding that both domains are responsible for activity and substrate specificity. Complementation experiments with chimeric IN would help to clarify this assumption.

We thank the Institut für Anorganische Chemie of the University of Tübingen for performing the atomspectroscopic analysis. We also thank Susanne Kempter and Sandra Bordel for their excellent technical assistance and Dr J. Cox for critical reading of the manuscript. This work was supported by a grant from the DFG (PF 257/2-2).

References

Ansari-Lari, M. A., Donehower, L. A. & Gibbs, R. A. (1995). Analysis of human immunodeficiency virus type 1 integrase mutants. *Virology* **211**, 332–335.

Asante-Appiah, E. & Skalka, A. M. (1999). HIV-1 integrase: structural organization, conformational changes and catalysis. *Advances in Virus Research* **52**, 351–369.

Barsov, E. V., Huber, W. E., Marcotrigiano, J., Clark, P. K., Clark, A. D., Arnold, E. & Hughes, S. H. (1996). Inhibition of human immunodeficiency virus type 1 integrase by the Fab fragment of a specific monoclonal antibody suggests that different multimerization states are required for different enzymatic functions. *Journal of Virology* **70**, 4484–4494.

Burke, C. J., Sanyal, G., Bruner, M. W., Ryan, J. A., Lafemina, R. L., Robbins, H. L., Zeft, A. S., Middaugh, C. R. & Cordingley, M. G. (1992). Structural implications of spectroscopic characterization of a putative zinc finger peptide from HIV-1 integrase. *Journal of Biological Chemistry* **267**, 9639–9644.

Bushman, F. D. & Craigie, R. (1991). Activities of human immunodeficiency virus (HIV) integration protein *in vitro*: specific cleavage and integration of HIV DNA. *Proceedings of the National Academy of Sciences, USA* **88**, 1339–1343.

Bushman, F. D., Fujiwara, T. & Craigie, R. (1990). Retroviral DNA integration directed by HIV integration protein *in vitro*. *Science* **249**, 1555–1558.

Bushman, F. D., Engelman, A., Palmer, I., Wingfield, P. & Craigie, R. (1993). Domains of the integrase protein of human immunodeficiency virus type 1 responsible for polynucleotidyl transfer and zinc binding. *Proceedings of the National Academy of Sciences, USA* **90**, 3428–3432.

Craigie, R., Fujiwara, T. & Bushman, F. (1990). The IN protein of Moloney murine leukemia virus processes the viral DNA ends and accomplishes their integration *in vitro*. *Cell* **62**, 829–837.

Dildine, S. L., Resress, J., Jolly, D. & Sandmeyer, S. B. (1998). A chimeric Ty3/Moloney murine leukemia virus integrase protein is active *in vivo*. *Journal of Virology* **72**, 4297–4307.

Drelich, M., Wilhelm, R. & Mous, J. (1992). Identification of amino acid residues critical for endonuclease and integration activities of HIV-1 IN protein *in vitro*. *Virology* **188**, 459–468.

Dyda, F., Hickman, A. B., Jenkins, T. M., Engelman, A., Craigie, R. & Davies, D. R. (1994). Crystal structure of the catalytic domain of HIV-1 integrase: similarity to other polynucleotidyl transferases. *Science* **266**, 1981–1985.

Engelman, A. & Craigie, R. (1992). Identification of conserved amino acid residues critical for human immunodeficiency virus type 1 integrase function *in vitro*. *Journal of Virology* **66**, 6361–6369.

Engelman, A., Mizuuchi, K. & Craigie, R. (1991). HIV-1 integration: mechanism of viral DNA cleavage and DNA strand transfer. *Cell* **67**, 1211–1221.

Engelman, A., Bushman, F. D. & Craigie, R. (1993). Identification of discrete functional domains of HIV-1 integrase and their organisation within an active multimeric complex. *EMBO Journal* **12**, 3269–3275.

Engelman, A., Hickman, A. B. & Craigie, R. (1994). The core and carboxyl-terminal domains of the integrase protein of human immunodeficiency virus type 1 each contribute to nonspecific DNA binding. *Journal of Virology* **68**, 5911–5917.

Engelman, A., Englund, G., Orenstein, J. M., Martin, M. A. & Craigie, R. (1995). Multiple effects of mutations in human immunodeficiency virus type 1 integrase on viral replication. *Journal of Virology* **69**, 2729–2736.

Englund, G., Theodore, T. S., Freed, E. O., Engelman, A. & Martin, M. A. (1995). Integration is required for productive infection of monocyte-derived macrophages by human immunodeficiency virus type 1. *Journal of Virology* **69**, 3216–3219.

Esposito, D. & Craigie, R. (1998). Sequence specificity of viral end DNA binding by HIV-1 integrase reveals critical regions for protein–DNA interaction. *EMBO Journal* **17**, 5832–5843.

Esposito, D. & Craigie, R. (1999). HIV integrase structure and function. *Advances in Virus Research* **52**, 319–333.

Holleman, A. F. & Wiberg, E. (1984). In *Lehrbuch der Anorganischen Chemie, 91–100: Verbessorte und Stark Erweiterte Auflage Von Nils Wiberg*, pp. 1110–1117. Berlin & New York: Walter De Gruyter.

Holler, T. P., Foltin, S. K., Ye, Q.-Z. & Hupe, D. J. (1993). HIV-1 integrase expressed in *Escherichia coli* from a synthetic gene. *Gene* **136**, 323–328.

Horton, R. M., Ho, S. N., Pullen, J. K., Hunt, H. D., Cai, Z. & Pease, L. R. (1993). Gene splicing by overlap extension. *Methods in Enzymology* **217**, 270–279.

Jenkins, T. M., Engelman, A., Ghirlando, R. & Craigie, R. (1996). A soluble active mutant of HIV-1 integrase: multimerization involves the core and C-terminal domains. *Journal of Biological Chemistry* **271**, 7712–7718.

Johnson, M. S., McClure, M. A., Feng, D.-F., Gray, J. & Doolittle, R. F. (1986). Computer analysis of retroviral *pol* genes: assignment of enzymatic functions to specific sequences and homologies with nonviral enzymes. *Proceedings of the National Academy of Sciences, USA* **83**, 7648–7652.

Kalpana, G. V. & Goff, S. P. (1993). Genetic analysis of homomeric interactions of human immunodeficiency virus type 1 integrase using the yeast two-hybrid system. *Proceedings of the National Academy of Sciences, USA* **90**, 10593–10597.

Katz, R. A., Merkel, G., Kulkosky, J., Leis, J. & Skalka, A. M. (1990). The avian retroviral IN protein is both necessary and sufficient for integrative recombination *in vitro*. *Cell* **63**, 87–95.

Katzman, M. & Sudol, M. (1995). Mapping domains of retroviral integrase responsible for viral DNA specificity and target site selection by analysis of chimeras between human immunodeficiency virus type 1 and visna virus integrases. *Journal of Virology* **69**, 5687–5696.

Katzman, M. & Sudol, M. (1998). Mapping viral DNA specificity to the central region of integrase by using functional human immunodeficiency virus type 1/visna virus chimeric proteins. *Journal of Virology* **72**, 1744–1753.

Katzman, M., Katz, R. A., Skalka, A. M. & Leis, J. (1989). The avian retroviral integration protein cleaves the terminal sequences of linear viral DNA at the *in vivo* sites of integration. *Journal of Virology* **63**, 5319–5327.

Khan, E., Mack, J. P. G., Katz, R. A., Kulkosky, J. & Skalka, A. M. (1991). Retroviral integrase domains: DNA binding and the recognition of LTR sequences. *Nucleic Acids Research* **19**, 851–860.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685.

LaFemina, R. L., Callahan, P. L. & Cordingley, M. G. (1991). Substrate specificity of recombinant human immunodeficiency virus integrase protein. *Journal of Virology* **65**, 5624–5630.

LaFemina, R. L., Schneider, C. L., Robbins, H. L., Callahan, P. L., LeGrow, K., Roth, E., Schleif, W. A. & Emini, E. A. (1992). Requirement of active human immunodeficiency virus type 1 integrase enzyme for productive infection of human T-lymphoid cells. *Journal of Virology* **66**, 7414–7419.

Pahl, A. & Flügel, R. M. (1993). Endonucleolytic cleavages and DNA-joining activities of the integration protein of human foamy virus. *Journal of Virology* **67**, 5426–5434.

Puras-Lutzke, R. A., Vink, C. & Plasterk, R. H. A. (1994). Characterization of the minimal DNA-binding domain of the HIV integrase protein. *Nucleic Acids Research* **22**, 4125–4131.

Sakei, H., Kawamura, M., Sakuragi, J.-I., Sakuragi, S., Shibata, E., Ishimoto, A., Ono, N., Ueda, S. & Adachi, A. (1993). Integration is essential for efficient gene expression of human immunodeficiency virus type 1. *Journal of Virology* **67**, 1169–1174.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*, 2nd edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

Sanger, F., Nicklein, S. & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences, USA* **74**, 5463–5467.

Sherman, P. A., Dickson, M. L. & Fyfe, J. A. (1992). Human immunodeficiency virus type 1 integration protein: DNA sequence requirements for cleaving and joining reactions. *Journal of Virology* **66**, 3593–3601.

Shibagaki, Y., Holmes, M. L., Appa, R. S. & Chow, S. A. (1997). Characterization of feline immunodeficiency virus integrase and analysis of functional domains. *Virology* **230**, 1–10.

Störmann, K. D., Schlecht, M. C. & Pfaff, E. (1995). Comparative studies of bacterially expressed integrase proteins of caprine arthritis–encephalitis virus, maedi–visna virus and human immunodeficiency virus type 1. *Journal of General Virology* **76**, 1651–1663.

Taddeo, B., Carlini, F., Verani, P. & Engelman, A. (1996). Reversion of a human immunodeficiency virus type 1 integrase mutant at a second site restores enzyme function and virus infectivity. *Journal of Virology* **70**, 8277–8284.

Tasara, T., Amacker, M. & Hubscher, U. (1999). Intramolecular chimeras of the p51 subunit between HIV-1 and FIV reverse transcriptases suggest a stabilizing function for the p66 subunit in the heterodimeric enzyme. *Biochemistry* **38**, 1633–1642.

Van Gent, D. C., Oude Groeneger, A. A. M. & Plasterk, R. H. A. (1992). Mutational analysis of the integrase protein of human immunodeficiency virus type 2. *Proceedings of the National Academy of Sciences, USA* **89**, 9598–9602.

Van Gent, D. C., Vink, C., Oude Groeneger, A. A. M. & Plasterk, R. H. A. (1993). Complementation between HIV integrase proteins mutated in different domains. *EMBO Journal* **12**, 3261–3267.

Vincent, K. A., Ellison, V., Chow, S. A. & Brown, P. O. (1993). Characterization of human immunodeficiency virus type 1 integrase expressed in *Escherichia coli* and analysis of variants with amino-terminal mutations. *Journal of Virology* **67**, 425–437.

Vink, C. & Plasterk, R. H. A. (1993). The human immunodeficiency virus integrase protein. *Trends in Genetics* **9**, 433–437.

Vink, C., Oude Groeneger, A. A. M. & Plasterk, R. H. A. (1993). Identification of the catalytic and DNA-binding region of the human immunodeficiency virus type 1 integrase protein. *Nucleic Acids Research* **21**, 1419–1425.

Wiskerchen, M. & Muesing, M. A. (1995). Human immunodeficiency virus type 1 integrase: effects of mutations on viral ability to integrate, direct viral gene expression from unintegrated viral DNA templates, and sustain viral propagation in primary cells. *Journal of Virology* **69**, 376–386.

Wlodawer, A. (1999). Crystal structures of catalytic core domains of retroviral integrases and role of divalent cations in enzymatic activity. *Advances in Virus Research* **52**, 335–350.

Yagil, E., Dorgai, L. & Weisber, R. A. (1995). Identifying determinants of recombination specificity: construction and characterization of chimeric bacteriophage integrases. *Journal of Molecular Biology* **252**, 163–177.

Received 29 February 2000; Accepted 19 September 2000